Evaluating climate and air pollution control policies in emerging economies

UNIVERSITY OF EAST ANGLIA

SCHOOL OF ECONOMICS

HAIBO ZHANG

100207669

A thesis submitted in fulfillment of the requirements for the degree of

Doctor of Philosophy

February 2024

Abstract

Environmental regulations are important tools to tackle the issues of climate change and air pollution that are increasingly severe in emerging economies. Using appropriate policy instruments plays an important role in realising the targets. In this thesis, we assess the effectiveness of environmental regulations in China, which is one of the largest emerging economies in the world.

In Chapter 2, we assess the effectiveness of early climate policy in China by causally evaluating the impact of the Low-carbon City Pilot (LCCP) on city-level per-capita CO_2 emissions and CO_2 intensity of GDP over the period 2003-2017. The idiosyncrasies of the policy design pose significant challenges for causal identification, which we overcome within a synthetic control framework. Contrary to previous contributions, our results suggest that the LCCP had no significant impact on either carbon emissions or intensity. The main takeaway of our empirical investigation is that even in emerging economies, effective environmental policy requires transparent, quantifiable targets, and credible enforcement.

Chapter 3 revisits the impact of he LCCP on environmental efficiency using a city-level panel dataset from 2003 to 2016. The unique design of the policy calls into question the credibility of the existing empirical analysis based on standard methods. We suggest an alternative identification framework based on synthetic control method. Contrary to the existing literature, our results suggest that the LCCP had no statistically significant increase on environmental efficiency. Nevertheless, for the first time we find a learning effect that instead increased the non-treated cities' efficiency in the short-run. We conduct a series of robustness checks to validate our results.

Chapter 4 investigates the carbon leakage induced by the air pollution control policy that focuses on $PM_{2.5}$ mitigation in China – *Action Plan for Prevention and Control of Air Pollution*. We employ a one-to-one nearest neighbour matching technique to overcome the significant challenge posed by the policy design. Our findings demonstrate unambiguous evidence that the Action Plan resulted in significant leakage of 151 thousand tonnes of CO_2 emissions each year. This translates to an annual increase of CO_2 emissions by around 4.4% in the surrounding regions. We validate our empirical findings through a battery of tests. We also explore the heterogeneity of our analysis and investigate the potential economic benefits and the possible channels.

Acknowledgements

This four-year PhD career is a truly life-changing experience, and it would not have been possible for me without the support that I have received.

First and foremost, I would like to express my deepest gratitude to my supervisors, Prof. Corrado Di Maria and Dr. Bahar Ghezelayagh, for their dedicated help during this time. It was my fortunate of having their guidance during my PhD career. I would not have been able to submit one of the chapters to the *Journal of Environmental Economics and Management* for potential publication without your help. Corrado, thank you for your effort in polishing our work, and I still remember the time that we spent together in Switzerland and Cyprus. I need to apologise, though, for not arranging a better treat for you and your family in Shanghai. The one that we went was awful. Bahar, thank you for your effort in helping me on the coding matters. I need to apologise for giving you too much chocolate when you were on diet. Anyway, I hope you could enjoy a bit whenever you can, because they are great.

Second, I am deeply grateful to Dr. Mike Brock. I had been working as an associate tutor for his module *Introductory Microeconomics* for one year and a half, and it was a pleasure. Once I could not teach the Thursday's seminar because my left eye had extremely strong foreign body sensation due to conjunctivitis. He was very careful of my condition and just took the class. The support, suggestion and company that he offered are considerably helpful. Particularly, Mike, thank you for your treat at Dyrrah.[1] Now I know I could get drunk on three ciders.

Third, I would like to thank my badminton team members. Besides Mike, the names of the others are Dr. Adam Kennerley, Ann Vicens, Shanfei Zhang and Dr. Stefania Sitzia, etc. Note that names are sorted alphabetically, so there is no bias at all. There is so much fun when having them around, not just playing badminton. Many thanks for the company and I wish so much we can gather around someday.

Fourth, I would like to thank my families. It was very inconvenient to fly back to China because everything was difficult during the Covid-19 pandemic. Luckily, it is almost there.

[1] Dyrrah is a coffee shop located in the Earlham Shopping Centre, NR2 3PD. They do amazingly great Italian pizza, and the one that I like the most is called 'Cotto e Funghi'.

A special thank to my girlfriend, Fei Liu, who graduated from our school two years ago. We started our relationship on 16 November 2019, and started our long-distance relationship on 15 December 2021. It is never easy to say goodnight over a screen every day. Fei, thank you for your patience and company over these years. Will you marry me?

"Yes, I do."

Contents

List of Figures

List of Tables

List of Acronyms

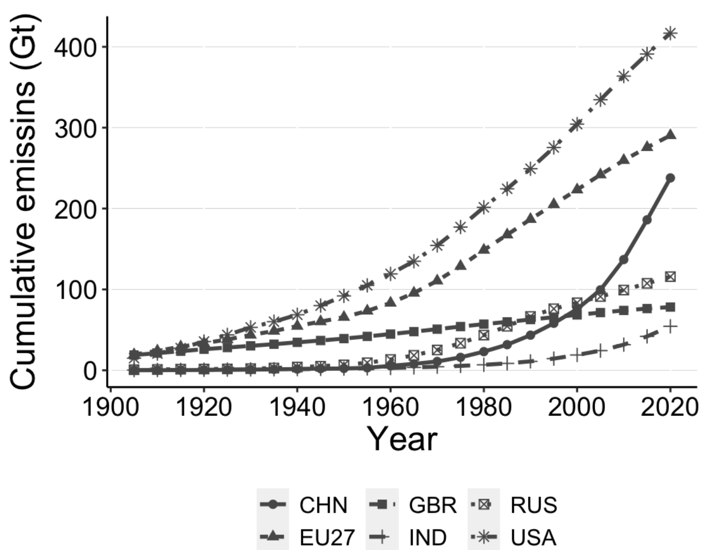
ADL	Auto-regressive distributed lag
AR5	The Fifth Assessment Report
ASCM	Augmented synthetic control method
ATT	Average treatment effect of the treated
CDM	Clean Development Mechanism
CEADs	Carbon Emission Accounts and Datasets
CNY	Chinese Yuan
DALYs	Deaths and disability-adjusted life-years
DEA	Data envelopment analysis
DiD	Difference-in-differences
DRC	Regional Development and Reform Commission
EPA	Environmental Protection Agency
ETS	Emission trading scheme
EU	European Union
FYP	Five-Year Plan
GDP	Gross domestic product
IPCC	Intergovernmental Panel on Climate Change
LCCP	Low-Carbon City Pilot
MEE	Ministry of Ecology and Environment of the People’s Republic of China
NASA	National Aeronautics and Space Administration
NBS	National Bureau of Statistics
NCSC	National Center for Climate Change Strategy and International Cooperation
NDRC	National Development and Reform Commission
NPC	National People’s Congress

OSHA	Occupational Safety and Health Administration
PM	Particulate matter
R&D	Research and development
RGGI	Regional Greenhouse Gas Initiative
SCM	Synthetic control method
SCPRC	State Council of the People's Republic of China
SFA	Stochastic frontier analysis
STUVA	Stable treatment unit value assumption
USA	United States of America

Chapter 1

Introduction

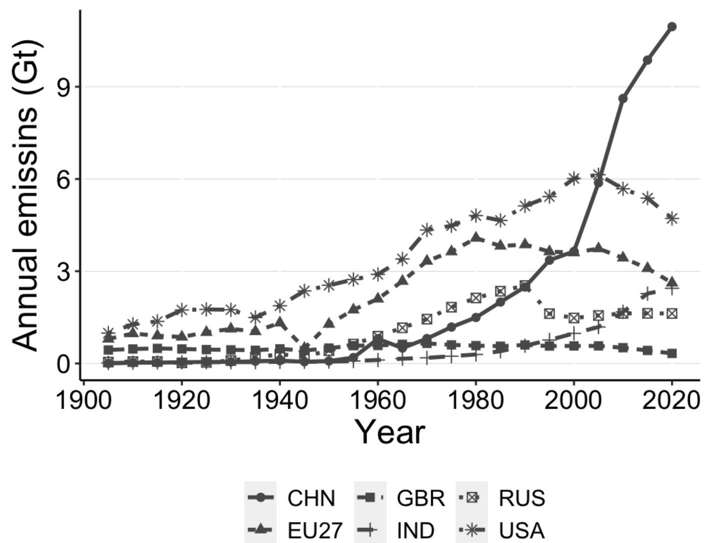
Weather and climate extremes have resulted in substantial economic losses (see Estrada et al., 2015, for example). The connection between these occurrences and anthropogenic climate change has become more pronounced since the release of the Fifth Assessment Report (AR5), and it is virtually certain the anthropogenic CO_2 emissions are the main driver (Masson-Delmotte et al., 2021). However, the debate between developed and emerging economies on who should shoulder the responsibility of reducing emissions has never stopped.



Notes: This figure displays the cumulative CO_2 emissions from several major economies worldwide, including Brazil, 27 members in the European Union, India, United States of America, China, Great Britain, and Russia.

FIGURE 1.1: Cumulative CO_2 emissions

Figure 1.1 displays the cumulative CO_2 emissions from several major economies worldwide. As of 2021, the US and EU members were the top two contributor to cumulative CO_2 emissions, with values of approximately 422 gigatonnes and 293 gigatonnes, respectively. More accurate estimates were reported by Wei et al. (2012), where developed and emerging countries had contributed about 60–80% and 20–40% to climate change by 2005, respectively. It clearly suggests that the developed economies ought to contribute more in mitigating climate change, due to their historical CO_2 emissions. Indeed, most of the existing literature discussing the effectiveness of environmental regulations focuses on the developed economies.



Notes: This figure displays the annual CO_2 emissions from several major economies worldwide, including Brazil, 27 members in the European Union, India, United States of America, China, Great Britain, and Russia.

FIGURE 1.2: Annual CO_2 emissions

Nevertheless, the growth of CO_2 emissions in the foreseeable future is expected to originate in emerging economies (Wei et al., 2012). Figure 1.2 displays the annual CO_2 emissions from several major economies worldwide. When the developed economies had started to show downward-sloping trends, the trends for emerging economies, especially China, were still surging. In 2021, China's annual CO_2 emissions reached 11 gigatonnes, more than the sum of the statistics from the US and EU members. Therefore, deploying effective climate policy in emerging economies is equivalently important, and distilling the relevant experience would be very much helpful for a coordinated development between the Global North and the Global South.

Unlike market-based or command-and-control instruments, voluntary approaches are less popular adopted in correcting for the market failure induced by the externalities of anthropogenic activities. The existing literature evaluating the performance of voluntary approaches offers a rather limited number of contributions. Prasad and Mishra (2017) evaluated the impact of voluntary environmental management standards ISO 14001 on Indian iron and steel sector. Using sample of 76 firms from 2006 to 2012, they find a significantly positive relationship between voluntary compliance and reduction in emissions intensity.

Deploying voluntary approaches to address environmental problems, from the theoretical point of view, is typically viewed as second-best instruments, since they do not meet the economist's ideal for a optimised policy instrument (Segerson, 2013). In fact, the effectiveness of using voluntary approaches largely depends on how they are designed and whether they can play an important role in environmental policy mix (Segerson and Miceli, 1998). Early empirical studies suggest that a strong regulatory threat and a reliable monitoring are important, since voluntary approaches are unlikely to be effective or efficient, if incentives for participation are weak or participation entails a limited pollution abatement commitment (Alberini and Segerson, 2002). More recent studies underscore the importance of how the voluntary regulation coordinates with other policies, by lending support to the proposition that self-regulation could be effective in yielding improvements on environmental performance within the current regulatory system, rather than as a substitute for that system (Lyon and Maxwell, 2019).

In Chapter 2 and 3, we evaluate a voluntary climate policy – the *Low-Carbon City Pilot* (LCCP), in one of the largest emerging economies in the world, China. We characterise the LCCP as 'voluntary' because the selection into treatment is largely associated with other unobserved characteristics. Indeed, regional authorities are often incentivised to perform key tasks that the regime deems essential (Landry et al., 2018). Such mechanism might have allowed the local officials to self-select themselves into the LCCP. While this policy has been heavily documented in the existing literature suggesting positive results, the setups in these studies are broadly similar. Noticeably, almost all studies use the canonical difference-in-differences (DiD) approach. However, the LCCP was adopted in different cities at different points of time. As econometrically demonstrated by Goodman-Bacon (2021), using the canonical DiD design in this context would lead to biased estimates. This methodological shortcoming due to the policy design of the LCCP pose challenges to causally identify the effect, which we will discuss with details in Chapter 2 and 3.

We investigate different questions in Chapter 2 and 3. Specifically, we evaluate the effectiveness of the LCCP in Chapter 2, with a focus on CO_2 emissions per capita and carbon intensity. We develop a complete identification framework that allows us to discuss the potential confoundedness that we are aware of, and explore the heterogeneity of our results. In Chapter 3, we follow the developed identification framework to investigate its impact on environmental efficiency. The existing literature broadly gauge the efficiency using data envelopment analysis (DEA) with slacks-based measure. While informative, the estimates may be over-estimated, since the efficiency is not separated from random shocks due to the deterministic nature of the DEA. We approach the question using stochastic frontier analysis that provides more economic interpretation of the estimates.

Apart from climate change, concentration of air pollutants is another market failure caused by the externalities of anthropogenic activities. Just like what the developed economies faced in the last century, air pollution is also a critical challenge for emerging economies.[1] China in 2013 began with haze and ended with haze. In January, the $PM_{2.5}$ concentration covered almost one quarter of China's land area. Over 600 million people were exposed to the polluted atmosphere with a record-high hourly maximum of 791 $\mu g/m^3$ in Beijing, more than 50 times higher than the Air Quality Guidelines 2021 published by WHO (2021) – 15 $\mu g/m^3$. Another occurrence was at the end of 2013, where the instantaneous concentration in Nanjing reached 943 $\mu g/m^3$.

In chapter 4, we study an air pollution control policy in China – *Action Plan for Prevention and Control of Air Pollution* (Action Plan). This policy was specifically issued to ease the growing concentration of particulate matter with a diameter of 2.5 micrometers or less ($PM_{2.5}$) in China. The effectiveness of the Action Plan has been well documented in the literature. For example, Liu et al. (2020b) examine the influence of the Action Plan on air quality across 16 districts in Beijing. Employing a first difference approach, they observe a substantial yearly reduction of approximately 10% in concentrations of SO2, PM10, $PM_{2.5}$, and CO. Yu et al. (2022) explore a comparable inquiry, compiling a city-level dataset spanning from 2008 to 2018. Employing a DiD design and propensity score matching, they report significant reductions of 18.4% in SO2 emissions and 24.7% in PM2.5 concentration due to the Action Plan. Additionally, Wu (2023), using a city-level panel dataset, evaluate the Action

[1] The pollution levels during the infamous London smog event in the winter of 1952, for instance, were 5-19 times higher than current regulatory standards and guidelines (Bell and Davis, 2001). The severe and prolonged consequences of this event resulted in approximately 12,000 excess deaths.

Plan's impact employing a triple-difference estimator and arrive at similar findings.

We assess the socioeconomic impact of Action Plan with a focus on an undesirable output in economic activities – CO_2. Specifically, we investigate whether the Action Plan led to increases of CO_2 emissions in the neighbouring regions. While closely related to our motivation in Chapter 2 and 3, our empirical analysis in Chapter 4 is also well aligned with the literature that focuses on the socioeconomic dimensions induced by environmental regulations. Walker (2013), for instance, focusing on the labor market in the United States of America (USA), investigate whether the 1990 Clean Air Act Amendments led to worker reallocation. Using an array of micro-level panel datasets with a triple-difference estimator, he find that workers in newly regulated plants experienced substantial. Another example is from a performance evaluation system that was constructed for mitigating SO_2 concentration. Using city-level statistics and a DiD approach, Chen et al. (2018) find that the system significantly reduced SO_2 emissions, but at the cost of the GDP growth rate. They show that local bureaucrats are willing to trade of economic performance to achieve emissions reduction goal.

Some studies have focused on the socioeconomic impact of the Action Plan, therefore they are more closely related to our work. Using door-to-door survey data from 302 households in six villages, Barrington-Leigh et al. (2019) assess the impact of the Action Plan on household energy use and expenditure, well-being and indoor environmental quality, by comparing the the treated (coal ban in place with subsidised heating system) and the untreated (no ban nor subsidy) households. They find positive impacts of the Action Plan on all outcomes. However, these benefits are sensitive to household wealth, where fewer benefits were found in low-income districts. Mei et al. (2021) investigate the impact of the Action Plan on real estate industry. They estimate a triple-difference estimator, using housing transaction data from 2011 to 2015 and administrative data on all power plants in Beijing. They find that the Action Plan led to a marginally significant price premium of 11% for properties close to coal-fired power plants. Using a multi-regional input-output model and an atmospheric chemical transport model, Fang et al. (2019) evaluate the impact of the Action Plan on primary $PM_{2.5}$ and secondary precursor emissions. They find that the Action Plan in fact lead to a leakage of the air pollutants, especially in neighboring provinces. Our work in Chapter 4 shares similarities with Fang et al. (2019), but we will approach the question in a different way and complement their study. We investigate whether the Action Plan led to increases on a 'co-pollutant' – CO_2 emissions, by examining the impact of the geographical proximity.

Our empirical investigation is also well aligned with the literature on carbon leakage, where most studies focus on international protocols and carbon markets. Carbon leakage from the Clean Development Mechanism (CDM) was suggested by Rosendahl and Strand (2011), where they find that the unilateral climate policy affect market equilibrium in energy and product markets, increasing emissions elsewhere. Aichele and Felbermayr (2015) find significant carbon leakage induced by the Kyoto Protocol, where the binding commitments have increased the embodied carbon imports from non-committed countries by around 8%, and the emission intensity of their imports has risen by about 3%. No statistically significant carbon leakage was found from the EU ETS (Naegele and Zaklan, 2019; Dechezleprêtre et al., 2022). Using data from German multinational firms, Koch and Mama (2019) find that the ETS-regulated firms have substantially increased the number of their affiliates outside the EU, which is suggestive of leakage of carbon in the future. Using transmission data from the national electricity grid, Fell and Maniloff (2018) find significant leakage of electricity generation from the Regional Greenhouse Gas Initiative (RGGI) where, the reduction of capacity utilisation in the RGGI region is compensated by cleaner generation in RGGI-surrounding regions. Results from the Japanese sub-national ETS are more interesting – ETS entities also reduced their emissions from the unregulated facilities in ETS-free regions (Sadayuki and Arimura, 2021). Cui et al. (2023) and He and Chen (2023) investigate the leakage of carbon induced by the China's ETS pilots, both concluding that the pilots significantly increased carbon emissions of the non-ETS firms that belong to the same ownership network as ETS ones. Although Zhu et al. (2022) suggest that this leakage effect is not related to administrative boundaries, their conclusions are less convincing, since they exclude some ETS-cities and use aggregated data. We will add to the literature and discuss our contribution with details in Chapter 4.

1.1 Outline

This thesis consists of three independent papers, which we start by assessing the effectiveness of the LCCP, with a focus on CO_2 emissions per capita and GDP CO_2 intensity in Chapter 2. These outcome variables provide direct comparability across different administrative divisions as well as being immediately related to the long-run relationship between CO_2 emissions and economic growth. The LCCP was adopted in different cities at different points of time, and the treated units were likely self-selected into the treatment due to the policy design. Taken together, we need identification strategy that incorporates variation in treatment timing and selection bias. In what follows, we use to the *partially pooled* SCM as our main identification strategy. We set out a robust design and correct for the mis-perception reported in the existing literature. We clarify the potential threats to our analysis, and use an alternative dataset to examine the robustness of our results. This dataset allows us to unprecedentedly decompose the treatment effect into sectoral level. We also document the relevant discussion and the choice of policy instruments.

In Chapter 3, we continue the analysis, with a focus on assessing the impact of the LCCP on environmental efficiency. We specify an enhanced hyperbolic distance function, then estimate the efficiency by using SFA. To account for the learning effect that might have promoted the diffusion of low-carbon mitigation from the pilots to the non-pilot cities, we adopt the timing-based approach suggested by Miller (2023). We apply the *partially pooled* SCM to control for the selection into treatment, and explore the potential heterogeneity of our results. We document relevant discussion at the end of the analysis.

In Chapter 4, we investigate the socioeconomic impact induced by the Action Plan that focuses on $PM_{2.5}$ mitigation. Our identification framework builds upon the policy design where the mandates were set at different levels for different regions. We use a nearest neighbour matching technique based on Mahalanobis distance to control for the systematic differences between the treated and control units. We clarify potential confounding factors in our analysis, and explore the heterogeneity of the leakage effect. We also investigate the economic benefits brought by the leakage of carbon, and explore the possible channels that could be attributed to.

Chapter 2

Climate Policy in Emerging Economies: Evidence from China's Low-Carbon City Pilot

2.1 Introduction

With the urgent need to effectively tackle climate change now beyond doubt (Masson-Delmotte et al., 2021), a fierce debate has broken out between developed and emerging economies on who should shoulder the responsibility – and the costs – of reducing carbon emissions. Undeniably, western economies ought to bear the responsibility for their historical emissions; at the same time almost all the growth in global energy demand – and therefore emissions – over the coming decades is expected to originate in emerging markets (Wei et al., 2012). Sound policies are therefore needed in both the Global North and the Global South to ensure that the expansion in human activity is finally decoupled from greenhouse gas emissions. Most of the research that aims to evaluate the effectiveness and the consequences of climate policy has to date focused on developed economies, while much less is known about how climate policies perform in emerging economies. Gaining a sufficient understanding of whether climate policies are working in the fastest-growing emerging economies is nevertheless vital to the debate on how to share the mitigation burden among countries.

In this chapter, we contribute to this debate by analysing the impact of the *Low-Carbon City Pilot*

(LCCP) – the first national climate policy introduced in China. Launched in 2010 by the National Development and Reform Commission (NDRC), the pilot was introduced to 'develop and demonstrate' the pathways that would help to accelerate the transition to a low-carbon economy (NCSC, 2020, In Chinese). The LCCP is particularly relevant, from our point of view, because it has been identified as an effective template for other countries to emulate (e.g. Hong et al., 2021).[1]

Given the specific focus of the LCCP on facilitating the shift to a low-carbon economy, in what follows we ask whether it indeed had a significant mitigation effect on both per-capita carbon emissions and carbon intensity of GDP.[2] These questions have not been satisfactorily answered in the literature so far. A rich literature has so far focused on efficiency and productivity effects, concluding that the LCCP had modest but statistically significant positive impacts, yet only a few studies have directly considered carbon emissions and, to the best of our knowledge, no study has directly addressed per-capita emissions.[3] Yu et al. (2019), Huo et al. (2022), and Tu et al. (2022) investigate directly the impact of the LCCP on carbon emissions; Feng et al. (2021), Zhou and Zhou (2021) and Hong et al. (2021) focus instead on emissions intensity, and are therefore closer in spirit to our investigation. None of these papers presents a credible framework for causal inference, however. Virtually all of the papers mentioned here adopt (some version of) the difference-in-differences (DiD) approach and, therefore, fail to address the idiosyncratic design of the LCCP, where the selection of the cities into the treatment group was far from random and the treatment staggered over time (e.g. Goodman-Bacon, 2021). The only exception to this is represented by Yu et al. (2019), who focus on Guangdong Province as a case study over the period 2010-2015. They construct a synthetic counterfactual for Guangdong and conclude that the LCCP reduced carbon emissions by approximately 10%. Their study is unsatisfactory, however, since they do not account for the simultaneous introduction of China's emissions trading scheme (ETS) pilots. Because of these methodological

[1]The LCCP was introduced in response to China's commitment, at the 2009 Copenhagen Conference of the Parties (COP15), to reduce by 2020 the CO_2 intensity of its GDP by 40-45% relative to its 2005 levels.

[2]Our outcome variables have the advantage of providing direct comparability across different administrative divisions as well as being immediately related to the long-run relationship between CO_2 emissions and economic growth.

[3]Most of the existing literature has used methods linked to productivity analysis such as Data Envelopment Analysis (DEA) to provide estimates of changes in efficiency and productivity that they would then link to the LCCP. Cheng et al. (2019), Yu et al. (2021), and Wen et al. (2022), for example, all point to positive, albeit limited, impacts of the LCCP on technical efficiency. Others, who used measures of productivity as their outcome of interest – such as Yao and Shen (2021) and (Zhou and Zhou, 2021)– conclude that the impact of the LCCP was less clear cut, and could have even been negative.

shortcomings, these papers fail to convincingly gauge whether the LCCP has been effective in kick-starting China's low-carbon transition. Our main contribution is, therefore, to provide a robust design for the causal identification of the impact of the LCCP on both per-capita carbon emissions and the carbon intensity of GDP and to present credible results to inform the debate on climate policy effectiveness in emerging economies.

To analyse the impact of the LCCP, we construct a unique dataset that merges socioeconomic and energy-related data. Overall, our dataset comprises detailed information on socioeconomic and environmental indicators, as well as CO_2 emissions for 245 Chinese prefecture-level cities over the period 2003-2017.[4] This dataset allows us to causally assess the impact of the LCCP on emissions per capita and carbon intensity for the administrative units treated in the first two waves (in 2010 and 2012, respectively).

A serious challenge to naïve identification in the context of the LCCP is that, as discussed in more detail below, the selection into treatment is not random. We overcome this problem by adopting an approach based on recent developments in the field of synthetic control method (SCM) that uses the pool of cities outside the LCCP to create credible counterfactuals that match the (pre-treatment) outcome variables of the treated ones (Ben-Michael et al., 2022). We then estimate the treatment effect by comparing the actual post-treatment outcomes of the treated cities to the relevant synthetic controls. We perform multiple tests to validate our identification strategy and conduct several robustness checks to shore up confidence in our empirical findings.

Our work complements the existing literature along three dimensions. First, as discussed above, we identify and overcome a range of potential challenges to causal inference that arise from the idiosyncratic design and the timing of the LCCP, thereby presenting empirical evidence which corrects the record in the literature on the actual effectiveness of the LCCP. Second, given that the implementation of the LCCP is largely voluntary, we contribute one state-of-the-art piece to the scant empirical literature that evaluates voluntary environmental policy instruments (e.g. Borck and Coglianese, 2009; André and Valenciano-Salazar, 2022). Third, we take a step forward in the

[4] In this context 'cities' is our short-hand for administrative divisions that comprise an urban centre and the surrounding county-level divisions. In China, there are three levels of administrative divisions: province-level, prefecture-level and county-level. Province-level divisions are the highest administrative level. In total, there are 34 province-level divisions, including 23 provinces, 5 autonomous regions, 4 municipalities and 2 special administrative regions. Prefecture-level cities are subordinate to the province-level division and comprise 293 prefecture-level cities, 30 autonomous prefectures, 7 prefectures and 3 leagues.

literature by assessing the impact of policy on carbon emissions in China by applying the methodology developed by Shan et al. (2017), which is based on the Intergovernmental Panel on Climate Change (IPCC) Guidelines, to construct an alternative emission inventory (IPCC, 2006). This alternative dataset not only allows us to examine the sensitivity of our results to changes in the data source but also enables us to look closer at the sectoral impacts of the LCCP and to discuss the potential for fuel-switching.

Overall, we find no evidence that the LCCP had significant impact in terms of reducing per-capita emissions or carbon intensity of GDP. This conclusion is drawn through examinations on the identification and a series of robustness checks. While these results contrast sharply with the results found elsewhere in the literature, they are not surprising when put in the context of a regulation that is fundamentally voluntary, provides no binding targets and lacks enforcement. We conclude that this early policy experiment did not deliver on its stated goals, at least not in terms of promoting a rapid de-coupling of economic growth from carbon emissions.

The rest of the chapter develops as follows, in Section 2.2, we describe the policy background and discuss the specific design characteristics that complicate causal identification in this case. Building on this, we discuss the identification strategy and the data in Section 2.3. Section 2.4 is devoted to the discussion of the main empirical results, their validity and some robustness checks. Section 2.5 discusses the potential economic mechanism and the sectoral impacts. Finally, section 2.6 summarises and concludes.

2.2 Policy background

Starting from a relatively low level of technological development, China's fast economic growth has come at the cost of severe environmental consequences over the last five decades (Smil, 1993). The sheer scale of China's economy has also meant that its rapidly increasing CO_2 emissions have greatly contributed to a rise in atmospheric concentrations of greenhouse gases with significant global impacts (Grimm et al., 2013).

In 2007, recognising the severity of this problem, China issued its National Climate Change Program (NDRC, 2007). This was followed in 2008 by the white paper on the country's actions and strategy on climate change (SCPRC, 2008). In 2009, following on the commitments agreed to within the framework of the 2009 United Nations Climate Change Conference, the State Council for the first time announced a target of reducing the carbon intensity of its GDP by 40–45% by 2020 compared to the 2005 level (SCPRC, 2009). This emissions mitigation target was then incorporated into the national 12th Five-Year Plan (FYP) (2011-2015) for the very first time, at the same time setting a binding target of 17% reduction in CO_2 emissions per unit of GDP between 2011 to 2015 (NPC, 2011).[5] Within the framework of the 12th FYP, each province was assigned a mitigation target, according to its socioeconomic characteristics and growth trajectories. When the 13th FYP (2016-2020) was published in 2016, the reduction target for the carbon intensity of GDP was set at 18% between 2016 to 2020 and further decomposed into different targets for each city (NPC, 2016).

Against this backdrop, the NDRC launched the LCCP, designed to accelerate the transition to a low-carbon economy and demonstrate pathways to achieve ambitious carbon reduction goals for the benefit of other cities. On 19 July 2010, the NDRC issued a ‘Notice on the Piloting Work of Low-carbon Provinces and Cities’ and then the first wave of the pilot started (NCSC, 2020, in Chinese). This first phase included two municipalities, five provinces, and six prefecture-level cities. The second wave began two years later and covered two municipalities, one province, and 26 prefecture-level cities. Finally, the third wave was introduced in 2017 and focused on prefecture-level cities and smaller administrative divisions. In total, eight additional county-level divisions (seven counties and one district) and 35 prefecture-level cities were included in the pilot scheme in the final stage.

It should be noted that, according to the NDRC, these pilot cities and provinces were selected based on their geographic, social and economic diversity, rather than being identified at random (NCSC, 2013, in Chinese). Moreover, in choosing the pilot locations account was taken of any ongoing work in low-carbon development and of any expression of interest by the regions to be part of the pilot.[6] Naturally, this process was also prone to political bargaining and manipulation. Therefore, assignment to treatment cannot be thought of as random by any stretch of the imagination, which poses

[5] The FYPs are a series of regulations in China, focusing on devising social and economic development guidelines for the entire country. The first Five-Year Plan (1953-1957) was implemented in 1953, the latest and current one is the 14th Five-Year Plan (2021-2025), introduced in 2021.

[6] Baoding and Shanghai, for example, had both been working with the World Wildlife Fund (WWF) on the ‘Low-Carbon City Initiative’ pilot to reduce CO_2 emissions since 2008, two years prior to the LCCP implementation. They were included in the first and second wave, respectively.

a serious challenge to our empirical investigation.

Rather than being assigned binding targets or given specific mandates, by the central government, each pilot division had significant flexibility in defining its own mitigation targets as long as they were consistent with the overarching FYP mandates. In particular, they were free to decide on the allocation of abatement across sectors. As mandated by the NDRC, the pilot cities were required to compile an explicit low-carbon development plan, which would articulate the measures needed to promote an effective local low-carbon economy, accelerate the establishment of a low-carbon industrial system, build a management system for greenhouse gas emission statistics, and encourage low-carbon lifestyles and green consumption patterns. To date, however, publicly available information on the overarching implementation process and any specific guidance offered to the local authorities remain scarce. Therefore, we collected additional information by scouring the official websites of the regional municipal people's governments, wherever available.[7]

In the majority of cases, we found that targets were set in terms of carbon intensity, the share of non-fossil energies, retiring outdated power plants, and forest coverage rate. Specific efforts were made to compile greenhouse gas inventories, decarbonise farming, public transportation systems and construction, introduce green nudges, and promote wetland conservation. For some of these measures, targets were set in some cities. For instance, Shijiazhuang was treated in the second wave and its online agenda clearly states that the share of 'new energy automobiles' in the personal transportation system should exceed 90% by 2015.[8]

To conclude, unlike traditional policy instruments, the LCCP is by and large a voluntary program, without mandated enforcement. In this sense, we would not expect it to have much impact, based on the evidence available in the literature (Borck and Coglianese, 2009). The mitigation pathways were devised by the regional authorities based on their regional economies and their local preferences. While mitigation pathways differ across treated units, almost all cities had targets on CO_2 emissions or GDP CO_2 intensity, and a few of which were more stringent than FYP mandates (see

[7]We managed to find online agendas for 20 of the 40 regulated administrative units, including two municipalities, three provinces, and 15 prefecture-level cities. As an example, see the online agenda (in Chinese) in Ningbo: `https://www.ningbo.gov.cn/art/2013/4/28/art_1229541831_59033042.html`. For the cities that did not publish agendas online or whose agendas are untraceable, we contacted the Regional Development and Reform Commission (DRC) for additional information. Based on their response, these cities either did not have a specific agenda or their agendas have been incorporated as a part of the 12th FYP.

[8]The term 'new energy automobiles' is often used by the Chinese government to refer to plug-in hybrid electric vehicles, battery electric vehicles, fuel cell electric vehicles, as well as liquefied natural gas vehicles. The exact definition may vary depending on the regional governments, however.

Khanna et al., 2014, for a detailed analysis of the first wave). In what follows, we, therefore, focus on assessing whether the LCCP was effective in bringing about additional mitigation, compared to elsewhere in the country, with a focus on emissions per capita and the CO_2 intensity of GDP.

2.3 Identification strategy and data

As discussed in the introduction, much of the existing literature on the LCCP employs a DiD approach to estimate the average treatment effect. Our empirical investigation, therefore, starts by replicating these efforts within a DiD framework.

Mindful of recent contributions that warn against using standard fixed-effect methods in the presence of heterogeneous treatment effects (e.g. Goodman-Bacon, 2021; Baker et al., 2022), and keeping in mind the staggered adoption of the LCCP, however, we adopt the dynamic DiD framework for intertemporal treatment effects proposed by De Chaisemartin and d'Haultfoeuille (2022).[9] To the best of our knowledge we are the first to use this methodology in this context.

While we believe that the use of dynamic DiD estimators *à la* De Chaisemartin and d'Haultfoeuille (2022) could control for the issue of heterogeneous treatment effects, it is clear that a naïve identification of the impact of the LCCP based on DiD methods would still be flawed, due to the non-random nature of the process whereby cities were included in the pilot. In fact, the pilot cities selected themselves, at least in part, into the pilot group and were otherwise chosen based on characteristics – such as their current level of industrialisation and their energy intensity – that are clearly correlated to the outcomes we seek to evaluate. To overcome these issues, we design our identification strategy around the pooled SCM recently introduced in the literature.

Generally speaking, SCMs estimate the treatment effect by constructing synthetic counterfactuals and comparing them to the actual outcomes for the treated units. The synthetic control is constructed by assigning weights to selected units drawn from the pool of control units (donors) so that the synthetic controls closely match the outcome of the treated units in the pre-treatment phase (Abadie and Gardeazabal, 2003; Abadie et al., 2010). While the SCM was originally designed to study a single treated unit, a number of recent contributions suggest possible extensions of

[9]Given the staggered adoption and the substantial differences in the treated units, heterogeneous treatment effects are indeed likely.

the SCM to multiple treated units (Dube and Zipperer, 2015; Galiani and Quistorff, 2017; Donohue et al., 2019). Estimating weights that minimise the average pre-treatment imbalance across different treatment units, however, could produce an almost perfect fit for the average, while leading to poor unit-specific fits.[10] On the other hand, focusing on a separate synthetic control for each treatment unit and estimating the average treatment effect on the treated could yield good fits for the unit-specific predictors while producing a poor balance for the average.[11] Recently, Ben-Michael et al. (2022) have instead proposed the so-called *partially-pooled SCM*, which seeks to mitigate such biases within a staggered treatment framework. Their method decomposes the error of the average treatment effect on the treated (ATT) estimate into errors stemming from the pooled fit and the unit-specific fits and then proceeds to minimise a weighted combination of the two. See Appendix A for the technical details. In an extension of their basic model, Ben-Michael et al. (2022) further recommend incorporating auxiliary covariates to insure a good pre-treatment fit not only for the main outcome variable of interest but also for other key characteristics of the units of analysis. In what follows, we adopt this augmented partially-pooled approach for staggered treatment as it fits well with the need to ensure a good fit across a range of treated units that are heterogeneous by design, and that are treated at different points in time.

2.3.1 Data

Our outcome variables of interest are the CO_2 emissions per capita (in ton/person) and the CO_2 intensity of GDP (in ton/10,000 CNY). Emissions per capita are calculated by dividing the regional CO_2 emissions by resident population, and the CO_2 intensity of GDP is calculated as CO_2 emissions per 10,000 CNY of regional GDP.[12]

The most challenging part of the data collection is to find reliable information on city-level CO_2 emissions. While in general preferable, estimates of emissions based on the IPCC guidelines are

[10]Kreif et al. (2016), for example, follow a similar approach and construct an aggregate treated unit and match the average pre-treatment aggregate outcome using weighted controls.

[11]For instance, Dube and Zipperer (2015) propose a modified SCM by converting the estimates to elasticities by ranking them based on the treatment intensity and aggregating the elasticities across different treatments. Similarly, Galiani and Quistorff (2017) and Donohue et al. (2019) focus on finding separate synthetic control for each of the treated units and then estimate the average treatment effect on the treated (ATT) by averaging the unit-specific SCM estimates.

[12]The use of the resident population *en lieu* of the registered population is generally recommended as it better reflects actual economic activities in China. This is also in line with the practice for calculating GDP per capita adopted by China's National Bureau of Statistics since 2004 (NBS, 2004, in Chinese).

only available for a limited set of cities due to the lack of complete data on city-level energy use (see Shan et al., 2017, for a discussion). In what follows, we, therefore, use the widely used data of Chen et al. (2020), that provide county-level carbon emissions data based on nighttime light data from satellite imagery. The emission inventories include 2,735 counties and districts in around 350 administrative divisions from 1997 to 2017. We obtain the data from the Carbon Emission Accounts Datasets and aggregated the CO_2 emissions at the city level (CEADs, 2020).

NCSC (2020, in Chinese) provides us with the information we need to construct our treatment indicators, which distinguish between the cities treated in each successive wave of the LCCP. Due to data availability, however, our data spans the period 2003-2017 and, therefore, omits the third wave of treatment. We exclude from our sample all the cities that will be subject to treatment in the third wave of the LCCP so that the control group more correctly reflects the 'never-treated' status of the non-LCCP cities.[13]

To construct the synthetic controls for the treated cities, we first use the values of the outcome variables – per-capita CO_2 emissions and CO_2 intensity of GDP – in the pre-treatment period. Abadie (2021) warns of the dangers of matching only on pre-treatment outcomes, which may lead to overfitting to noise and introduce potential sources of bias. We, therefore, introduce additional covariates in our predictor set that we use to try and balance against systematic differences between the treated cities and weighted donor units. We include per-capita GDP (in 10,000 CNY), the industrialisation rate, i.e. the GDP share of the secondary sector, social fixed asset investments (in 10 billion CNY), and industrial SO_2 discharges (in 10,000 ton) as additional predictors. As a robustness check, in what follows we also include employment (million people) and investment in science and technology (billion CNY) to the predictor set to gauge the sensitivity of our results to changes in the predictor set. All these data come from the *China City Statistical Yearbook* (NBS, 2017), and the monetary values are normalised to 2010 CNY. All our data was also cross-checked with the relevant data from prefectural and provincial statistical yearbooks – which may be accessed via the cities' or provinces' municipal bureau of statistics – to ensure accuracy and consistency.

In our baseline results, we classified as treated all the cities that were included either directly or

[13] Using the third-wave cities as donor units for the previous waves implies that we assume that they did not prepare in any way ahead of the regulation, i.e. that there is no anticipation effect. This is questionable, however, given the possibility of political bargaining and the strong connections between regional authorities and the central government.

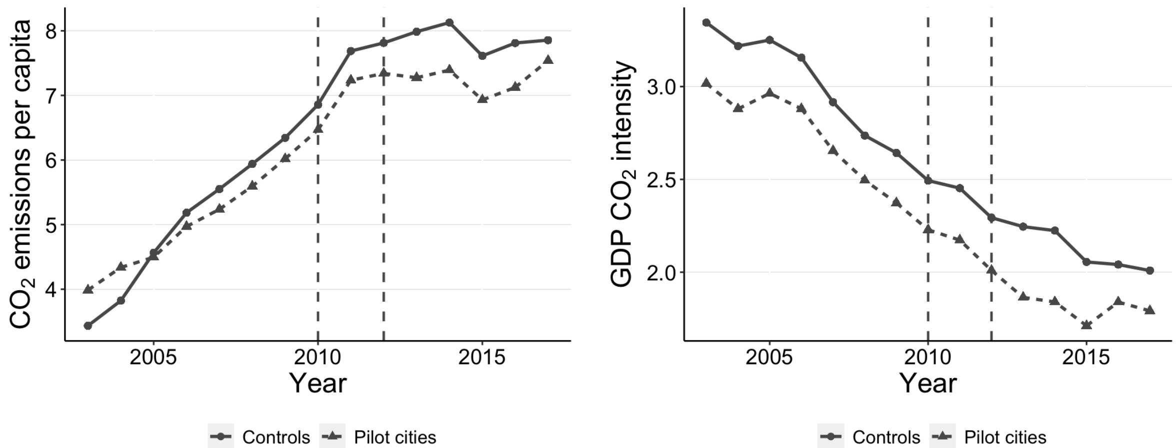
TABLE 2.1: Descriptive statistics, 2003-2017.

	Mean	Std. dev.	Min.	Max.	Obs.
Panel A: Treated cities					
Outcome variable:					
CO_2 emissions per capita (ton/person)	6.13	4.20	0.35	32.86	1,230
GDP CO_2 intensity (ton/10,000 CNY)	2.31	1.39	0.19	10.12	1,230
Socioeconomic measurement:					
GDP per capita (10,000 CNY)	3.28	2.37	0.11	15.41	1,230
Employment (million people)	0.50	0.56	0.06	4.64	1,230
Industrialisation rate (%)	47.67	9.92	18.14	84.39	1,230
Social fixed asset investment (10 billion CNY)	8.46	9.60	0.26	65.30	1,230
Expenditure on science and technology (billion CNY)	0.48	1.76	0.00	34.42	1,228
Industrial SO_2 discharge (10,000 ton)	5.09	4.69	0.01	33.90	1,225
Panel C: Donor cities					
Outcome variable:					
CO_2 emissions per capita (ton/person)	6.44	5.63	0.46	55.22	2,443
GDP CO_2 intensity (ton/10,000 CNY)	2.60	1.58	0.52	15.07	2,443
Socioeconomic measurement:					
GDP per capita (10,000 CNY)	2.90	2.35	0.16	20.24	2,443
Employment (million people)	0.35	0.29	0.04	3.80	2,443
Industrialisation rate (%)	48.56	11.84	2.66	90.97	2,443
Social fixed asset investment (10 billion CNY)	7.45	7.85	0.20	63.59	2,443
Expenditure on science and technology (billion CNY)	0.20	0.37	0.00	4.82	2,443
Industrial SO_2 discharge (10,000 ton)	5.38	4.86	0.01	33.19	2,428

Notes: The table shows means, standard deviations, minimum values, maximum values and the number of observations. Panel A displays the statistics of the cities under either the first or second wave; panel B displays the statistics of the cities that are never included in any treatment pool under the LCCP.

indirectly, i.e. via municipality-, province- or prefecture-level treatment in either the first or second wave. As a control group, we use all cities that were not treated in either wave. An important caveat in creating reliable synthetic controls is that both pre-treatment outcomes and additional predictors of the treated unit should fall in the convex hull of the donor units (as indicated by the minimum and maximum). As explained in Section 2.2, however, the assignment to treatment is not random, so that the treated cities are on average cleaner and more advanced, making it impossible to create close matches on some of the measurements. We, therefore, identify and exclude as outliers the cities of Beijing, Tianjin, Shanghai, Suzhou, Guangzhou, Shenzhen and Chongqing, for which no plausible donors exist. After the adjustment, we are left with 245 cities, 82 of which were included in the LCCP in either the first or the second wave.

Table 2.1 provides the descriptive statistics of the variables used in the analysis, divided by treatment status, over the period 2003-2017. Although as mentioned the treated cities exhibit better economic and environmental performances, their minimum and maximum values fall approximately in the support of the donor cities for most measurements. We are therefore confident in fitting reliable synthetic counterfactuals that closely match the treated cities' historical outcomes and additional predictors.



Notes: The figure shows the trends of the outcome variables, CO_2 emissions per capita and GDP CO_2 intensity, respectively.

FIGURE 2.1: Outcome variables

Figure 2.1 displays the trends of the outcomes over the study period. In the pre-treatment periods, the trend for GDP CO_2 intensity is approximately parallel. For emissions per capita, however, the pilot cities outpaced the control cities in the early periods, then develop approximately parallel afterwards. In the post-treatment periods, both outcomes diverge, although the differences seem to diminish in the last few periods.

2.4 Empirical results

We begin this section by presenting the results we obtain within the DiD framework discussed in Section 2.3. This approach allows us to clarify the placement of our contribution within an existing literature that has mostly relied on naïve DiD estimations, before moving on to discussing the results that emerge from our preferred synthetic-control-based methodology.

Table 2.2 and Figure 2.2 present the results of a staggered DiD estimation, following the methodology introduced by De Chaisemartin and d'Haultfoeuille (2022). The goal of this procedure is to capture the aggregate effect of the LCCP on the outcome variables of interest over the first two waves.[14] The results in the Table suggest that the LCCP had a statistically significant impact on per-capita CO_2 emissions, with a reduction of 0.38 ton per capita – about 7% less than the average emissions in the pre-treatment phase – whereas there is no significant effect on the carbon intensity of GDP, compared to the control group.

TABLE 2.2: Estimates of staggered difference-in-differences

	ATT estimate	Std. err.	*p*-value
CO_2 emissions per capita	-0.363***	0.139	0.009
GDP CO_2 intensity	-0.061	0.041	0.140

Note: (i) The table displays the estimates of staggered difference-in-differences estimations of CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

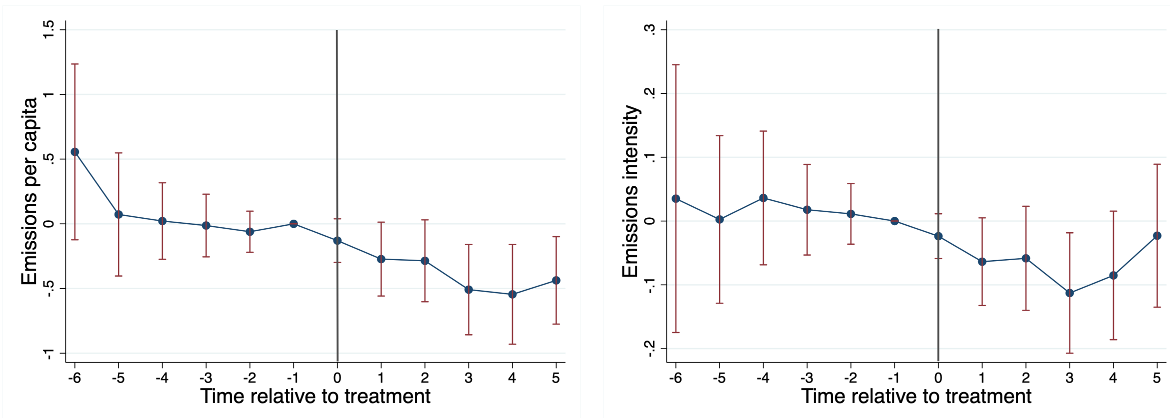
Figure 2.2 plots the evolution over time of the impacts of the policy across the first two waves of the LCCP and shows downward-sloping trends over time, at least initially. This pattern is particularly pronounced for per-capita emissions. For the emissions measure, the results become significantly negative at $t = 3$ and remain so until the end of the horizon. The results are less clear-cut for CO_2 intensity. The DiD estimates are negative in the short run, albeit only significantly so at $t = 3$, and rebound strongly towards zero at the end of the time window. Alarmingly, however, for both outcomes, the estimates in the treatment period suggest that they are the continuation of trends started well before $t = 0$, thus violating the 'parallel-trends' assumption needed for DiD identification. These results confirm that naïve regressions of this type are not the ideal approach to identify causality in this context.

While problematic, our emissions results are broadly consistent with the existing literature that considers the role of the LCCP in mitigating carbon emissions using a DiD approach. For example, Huo et al. (2022) and Tu et al. (2022) find that the LCCP reduces CO_2 emissions by 2-3%.[15] Our null

[14] In the figure, the treatment effects are normalised relative to the beginning of the corresponding treatment period, i.e. t=0 represents 2010 for wave I, and 2012 for Wave II.

[15] Huo et al. (2022) do not control for the staggered nature of the treatment, nor do they account for the non-random nature of the selection into treatment, both of which bias their results, and call their identification strategy into question.

results on the impact of the LCCP on the carbon intensity of GDP, however, contrast both with the findings of Feng et al. (2021) and Zhou and Zhou (2021), who argue that the LCCP has *increased* the carbon content of GDP, and those of Hong et al. (2021), who instead find a significant reduction of energy consumption relative to GDP.[16] These differences might be due to the heterogeneity of treatment effects, which have been shown to give rise to biased estimates in the presence of staggered treatments (Baker et al., 2022, e.g.).



Notes: The figure shows the results of intertemporal difference-in-differences estimations on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP (De Chaisemartin and d'Haultfoeuille, 2022). The effects are normalised relative to the beginning of the corresponding treatment, i.e. 2010 for Wave I and 2012 for Wave II.

FIGURE 2.2: Plots of the staggered difference-in-differences effects

Overall, our assessment of this first set of results is that even if they represent an improvement on the current state-of-the-art, in that they at least address the potential biases in the estimated treatment effects due to the staggered nature of the treatment, they still fall short of providing a convincing identification framework for the causal effects of the LCCP. Indeed, it is clear that – as argued in Section 2.2 – the selection into the LCCP is not random. As a consequence, the identification strategy that underlies the DiD efforts discussed above is unsatisfactory. In view of this

While Tu et al. (2022) account for the staggered treatment, they also fail to control for the selection into treatment aspect. Neither study, moreover, discusses the potential misattribution of the effect that arises from the partial overlap of the LCCP with the ETS pilots, so their identification strategy is questionable.

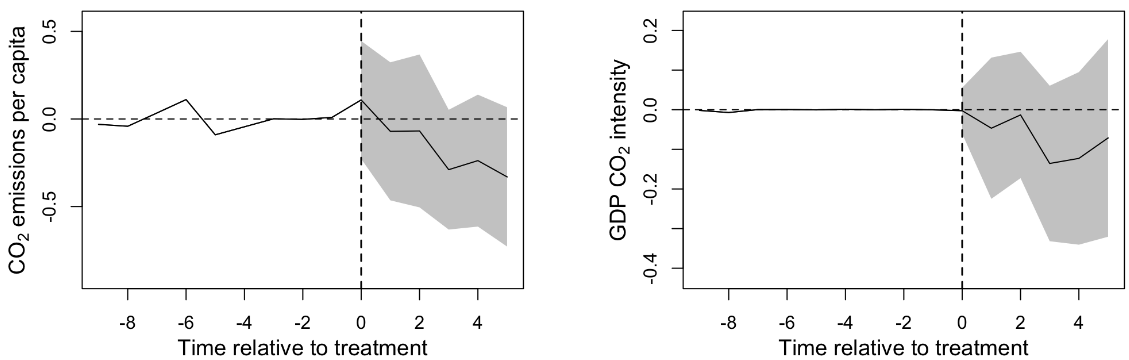
[16]We note here that, taken together, these results would imply that China moved to a much more carbon-intensive energy mix *as a consequence of the LCCP*, which is hard to believe. These studies, however, suffer from a number of limitations that might explain their somewhat erratic conclusions. In particular, neither Feng et al. (2021) nor Hong et al. (2021) control for the staggered nature of the treatment, while Zhou and Zhou (2021) focuses on Wave II only. Neither of the two latter studies controls for the non-random treatment selection, and all fail to account for policy overlaps. Overall, their identification strategies are not very convincing, which might explain their contrasting results.

discussion, we now move on to the main part of our analysis, where we apply the partially-pooled SCM introduced by Ben-Michael et al. (2022) to the LCCP.

TABLE 2.3: Estimates of the staggered synthetic control – baseline

	ATT estimate	Std. err.	p-value
CO_2 emissions per capita	-0.148	0.174	0.395
GDP CO_2 intensity	-0.065	0.077	0.399

Notes: (i) The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.



Notes: The figure shows the results of the staggered synthetic control method on per-capita CO_2 emissions and GDP CO_2 intensity. The effects are normalised relative to the beginning of treatment, i.e. 2010 for Wave I and 2012 for Wave II.

FIGURE 2.3: Plots of the staggered synthetic control method – baseline

The results of the partially-pooled, staggered synthetic control procedure run for the first two waves of the LCCP are presented in Table 2.3. The table provides the point estimates, the standard errors and the p-values for the treatment effects.[17] Relative to the non-pilot cities, these results show that the LCCP had no statistically significant effect at conventional levels on the treated ones. Figure 2.3 plots the estimates of the effects over time. Overall, the pre-treatment fits are satisfactory

[17] The conventional practice to claim statistical inference of synthetic control method is to run a number of falsification tests. Specifically, one can estimate treatment effects $\hat{\tau}_j$ for each of the $j = 2, \ldots, N$ donor units following the main specification, using the remaining $N - 2$ donor units. Here we follow Ben-Michael et al. (2022) and provide statistical inference using the leave-one-unit-out jackknife approach. See the online appendix of Ben-Michael et al. (2022) for more details.

and, based on the confidence intervals plotted, the treatment effects on both measures remain insignificant throughout the treatment period.

These findings are clearly at odds with the ones currently available in the literature, so in the remainder of this section, we delve deeper into the data to shore up our confidence that these results are indeed correct and robust.

2.4.1 Challenges to identification

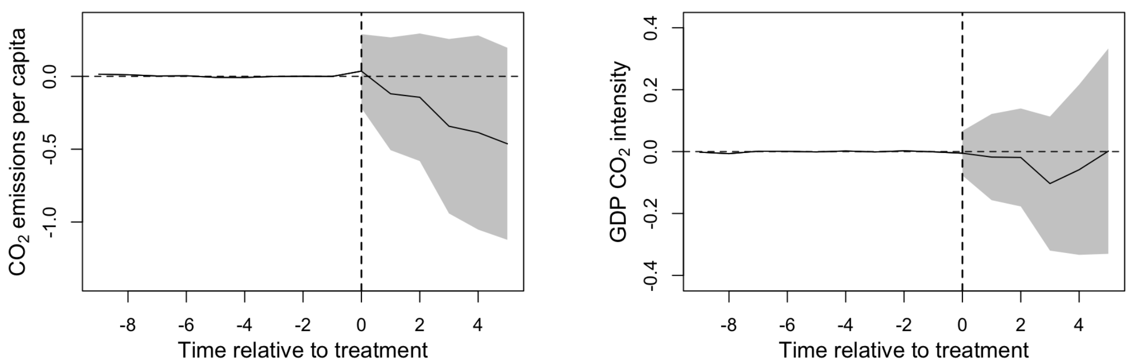
One of the most critical challenges to identification in the context of the LCCP derives from the fact that several policy initiatives aimed at decoupling carbon emissions from economic growth were undertaken in China around the same time as the LCCP. For example, shortly after the introduction of the LCCP, the Chinese authorities started discussing the introduction of emissions trading as a climate change mitigation tool. Beginning in 2011, with trading commencing in 2013, seven emissions trading scheme (ETS) pilots were launched. The pilots included one prefecture-level city (Shenzhen), two provinces (Hubei and Guangdong) and four municipalities (Beijing, Shanghai, Tianjin, Chongqing).[18] Due to the fact all ETS-regulated cities are also treated by the LCCP, making it impossible to attribute any treatment effect to the LCCP alone causally.

To control for this confoundedness, we exclude all ETS-regulated cities from our sample, leaving us with a total sample of 214 cities, 51 of which were treated under the LCCP. Using this restricted sample, we run our SCM model once again to confirm the validity of our design.

Figure 2.4 and Table 2.4 report the results of this exercise. The effect of excluding the cities treated by the ETS pilots is relatively small. Compared to the baseline discussed in Table 2.3, the changes in the estimated coefficients are small and they remain insignificant, with the p-value for the carbon intensity increasing to 0.699. For completeness, we repeat the same exercise using the staggered difference-in-differences approach of De Chaisemartin and d'Haultfoeuille (2022) and report it in the lower half of Table 2.4. In this case, the results are quite striking as the coefficient of the per-capita emissions becomes much smaller and strongly insignificant compared to the ones presented in Table 2.2.

[18] The cap covered around 40% of the total CO_2 emissions in each division, including a range of entities and industries (Swartz, 2016). The empirical literature has suggested that the ETS pilots reduced CO_2 emissions by around 15.5% (Hu et al., 2020).

Taken together, these findings suggest that the inclusion of the ETS pilot cities in the LCCP treatment group might lead to significant biases in the results of DiD estimates. The fact that we find no evidence that our SCM baseline results are significantly impacted by them suggests that the pooled SCM methodology may be more robust to this type of overlap than other approaches. It is worth noting that, to the best of our knowledge, none of the significant results reported in the literature control for the policy overlap discussed here. This strongly suggests that taking them at face value might lead to misleading conclusions.



Notes: The figure shows the results of the staggered synthetic control method on per-capita CO_2 emissions and GDP CO_2 intensity on a restricted sample that excludes all the cities taking part in the ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE 2.4: Staggered synthetic control estimation – controlling for policy overlap

TABLE 2.4: Controlling for policy overlap – excluding ETS cities

	ATT estimate	Std. err.	*p*-value
Panel A: Staggered synthetic control			
CO_2 emissions per capita	-0.236	0.230	0.304
GDP CO_2 intensity	-0.034	0.087	0.699
Panel B: Staggered difference-in-differences			
CO_2 emissions per capita	-0.070	0.147	0.632
GDP CO_2 intensity	-0.081	0.055	0.138

Notes: (i) The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP on a restricted sample that excludes all the cities taking part in the ETS pilots. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

As discussed in Section 2.2, alongside the LCCP the Chinese government was making concurrent efforts to decarbonise the economy, via the increasingly stringent targets mandated by the FYPs. Our identification framework implicitly assumes that treated and donor units are assigned similar reduction targets under the FYPs, thus not biasing our estimates of the impact of the LCCP. To test whether this assumption holds, we collect information on the reduction targets mandated for each of the cities in our sample under both the twelfth and thirteenth FYPs. We then perform equivalence tests for the average reduction targets to ensure that the FYPs' mandates do not introduce biases to our estimates above.

Table 2.5 reports the results of these tests. As the t-statistics and the p-values suggest, we cannot reject the null hypothesis that the reduction targets are equal between the two groups. This implies that our results above are not likely to be driven by differences in the reduction targets in the FYPs..[19]

TABLE 2.5: Testing differences in reduction targets under the 12th and 13th Five-Year Plans

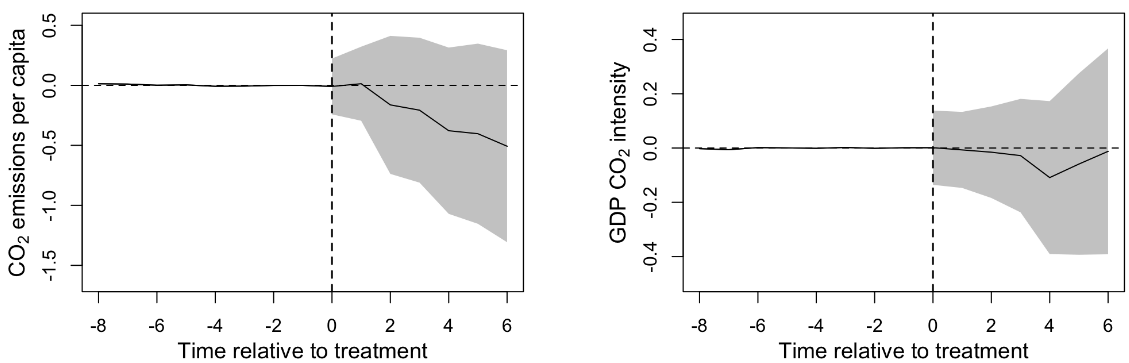
	Donor units Mean	Treated units Mean	Diff.	t-statistic	p-value
12th Five-Year Plan (2011-15)	17.04	17.08	-0.04	-0.18	0.86
13th Five-Year Plan (2016-20)	18.54	18.88	-0.35	-0.96	0.34

Notes: (i) The table reports the results of the t-test for the equality of means between the treated and donor units for the carbon emissions reduction targets set by 12th and 13th Five-Year Plans. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Another possible limitation of our identification strategy is that the LCCP was introduced following an earlier announcement and the selection process of suitable pilot candidates was also rather

[19] As indicated by one anonymous reviewer, another policy – the Two Control Zones (TCZ) policy – might be another possible source of confoundedness here. While the TCZ has been shown to have been successful at reducing polluting emissions (e.g. Cai et al., 2016), however, it had only a minimal overlap with the LCCP. On the one hand, it stopped running in 2010, the first year in which the LCCP was introduced; on the other hand, its goal was to mitigate acid rains by reducing SO_2 emissions from coal combustion, rather than focusing on carbon emissions. The main consequences of SO_2 regulation were the closure of older coal-fired boilers and a switch to lower-sulfur coal (including washed coal). According to Zhang et al. (2016), these behavioural responses to the regulation have significantly contributed to decoupling economic growth from SO_2 discharge in China. The effect of the TCZ on CO_2 emissions in the control zones is more uncertain, however, because switching from high-sulfur coal to cleaner coal does not necessarily reduce CO_2 emissions (e.g. Zhang et al., 2016). In fact, low-sulfur coal has a higher net caloric value than dirtier coal and therefore produces more CO_2 emissions per unit of weight during combustion (Shan et al., 2018b). As discussed by Glomsrød and Taoyuan (2005), moreover, switching from dirtier to cleaner coal has complex system-wide implications, which might even lead to an increase in CO_2 emissions. For the sake of completeness, however, we have re-run our SCM analysis excluding the TCZ cities, to control for any policy overlap; we also repeated our analysis using only the set of cities treated under the TCZ, to isolate the potential *additional* effects of the LCCP. In both cases, we fail to identify any impact from the LCCP. If anything, the results are even more insignificant than our baseline ones. See Table D.1 in Appendix for the summary of the results, and Figures D.1 and D.2 for the synthetic control fits

slow. From this point of view, our choice to start the treatment period from the official inception dates of wave I and II – in 2010 and 2012, respectively – might be considered naïve. It is indeed plausible that at least in some of the treated cities, both officials and economic agents might have been aware of their future treatment status through their own lobbying for selection into the pilot or other political connections. If this were indeed the case and at least some of the pilot cities had taken early actions to prepare for the pilot, this could introduce biases in the selection of donors. Selecting donors with lower emissions would then potentially lead to an attenuation of the estimated effect, and to insignificant results. To control for this potential bias, we conduct our analysis again, this time moving the notional start of the treatment to one year prior to the official start of the pilot.[20]



Notes: The figure shows the results of examining the anticipation effect on per-capita CO_2 emissions and GDP CO_2 intensity using staggered synthetic control method. The effects are normalised relative to the beginning of treatment, i.e. 2009 for wave I and 2011 for Wave II.

FIGURE 2.5: Controlling for potential anticipation effect – alternative policy start

Figure 2.5 and Panel A of Table 2.6 present the results of the above discussion. While the development trajectories are not subject to major changes, we find that the estimates attenuate for both outcomes compared to Table 2.3. In the presence of an anticipation effect, we would instead expect larger estimates and smaller p-values, because by backdating the treatment start date, the anticipation effect would be incorporated into the treatment effect. Overall, we find no evidence to support the existence of a significant anticipation effect. For completeness, we also exclude the

[20]While a two-year anticipation effect seems excessive in this context, for completeness we also performed this analysis moving the treatment date up by two years. The results do not change qualitatively, as the treatment effect remains insignificant for both outcomes.

cities treated under the ETS pilots. The results are reported in Panel B of Table 2.6 and in Appendix (see Figure D.3). Again, we find no indication that an anticipation effect might have taken place.

TABLE 2.6: Controlling for potential anticipation effect – alternative policy start

	ATT estimate	Std. err.	*p*-value
Panel A: Baseline sample			
CO_2 emissions per capita	0.030	0.329	0.927
GDP CO_2 intensity	-0.061	0.091	0.502
Panel B: Excluding the ETS-regulated cities			
CO_2 emissions per capita	-0.236	0.268	0.377
GDP CO_2 intensity	-0.033	0.110	0.766

Notes: (i) The table displays the estimates of examining the anticipation effect on per-capita CO_2 emissions and GDP CO_2 intensity using staggered synthetic control method. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Lastly, we focus on the potential for treatment spillovers to the control group, which would compromise our identification. On the one hand, it is possible that the pilot was successful in identifying, developing and demonstrating low-carbon pathways that may have been adopted by other cities. This would potentially lead to reductions in both outcomes among treated and control units. On the other hand, the introduction of the LCCP might have increased the cost of carbon emissions in the treatment regions and pushed economic activities towards areas with less stringent environmental regulations, thus leading to carbon leakage. In this case, emissions would increase in the destination cities alongside economic activity.

To test for the presence of these treatment spillovers, and assuming that any spillover is more likely to occur in cities 'close' to the pilot ones, we first excluded from the donor pool cities that are in close geographical proximity to the pilots from our sample.[21] Using this restricted sample, we rerun our synthetic control estimations for both outcomes. Next, we restrict the donor pool to include the neighbouring cities only and repeat the analysis. The overall idea here is that, in the presence of treatment spillovers, this latter set of results ought to be less significant than the former.

[21] Specifically, we drop all control units that share a border with a treated city.

TABLE 2.7: Controlling for treatment spillovers

	ATT estimate	Std. err.	*p*-value
Panel A: Excluding neighboring cities			
CO_2 emissions per capita	-0.304	0.372	0.415
GDP CO_2 intensity	-0.030	0.142	0.830
GDP per capita	0.104	0.146	0.476
Panel B: Using neighboring cities as donor units			
CO_2 emissions per capita	-0.168	0.332	0.612
GDP CO_2 intensity	-0.001	0.094	0.993
GDP per capita	0.062	0.142	0.660

Notes: (i) The table displays the estimates of examining the treatment spillovers on CO_2 emissions per capita, GDP CO_2 intensity, and GDP per capita for the first two waves of the LCCP using staggered synthetic control method. Panel A shows the results excluding neighbouring cities from our sample; Panel B shows the results using only neighbouring cities as donor units. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Finally, to control for possible leakage effects, we also run an additional test using per-capita GDP as an outcome that allows us to identify economic leakage.[22]

Table 2.7 presents the results of our investigation into treatment spillovers.[23] All the estimates remain insignificant, although – consistent with the idea that spillovers are more likely in neighbouring cities – point estimates in Panel B seem to be rather attenuated. Overall, we discard the idea that treatment spillovers or carbon leakage drive our insignificant results in the baseline.

2.4.2 Robustness checks

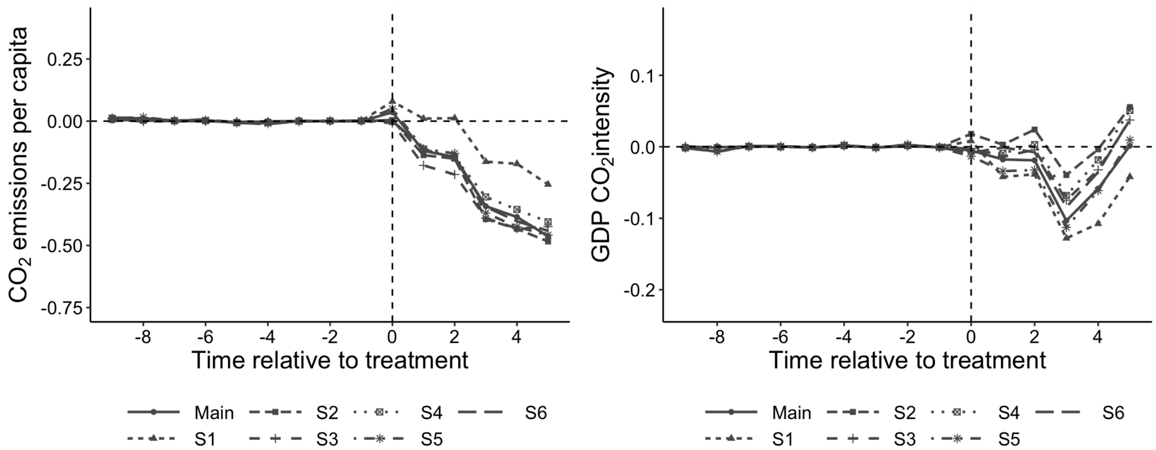
Having acknowledged the possible challenges to our identification strategy and having found that they do not invalidate our approach, we now start our discussion of the robustness of our results to several possible changes in the data. For the remainder of this section, we work with a restricted dataset from which we have removed the ETS-regulated cities, for cleaner identification.

Our first step is to make sure that the main results are not driven by the set of predictors used to construct the synthetic controls in our main specification. In what follows, we repeat our estimates

[22] To achieve treatment-control balance in the GDP per capita analysis, we use employment, industrialisation rate, social fixed asset investment and expenditure in science and technology as additional predictors in the construction of the synthetic control.

[23] See Figures D.4 - D.6 in Appendix for the synthetic control fits.

with different sets of predictors, starting from matching on outcomes only. We then expand the predictor set one variable at a time, until we have used all the variables at our disposal. The complete set of predictors includes the two original outcome variables, GDP per capita, industrialisation rate, social fixed asset investments, industrial SO_2 discharges, employment, and expenditure on science and technology. If the results do not change substantially, we can conclude that the selection of the predictors does not drive our SCM results.



Notes: The figure shows the results of examining the sensitivity to different predictor sets on CO_2 emissions per capita and GDP CO_2 intensity using a staggered synthetic control method. Effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE 2.6: Robustness checks – changing predictor sets

Table 2.8 and Figure 2.6 present the results of our sensitivity analysis to the different predictor sets. We find that including or excluding predictors only marginally changes the point estimates, and no estimate comes close to being significant. Overall, this exercise shows that our results are extremely robust across predictor sets.

We next look into possible differential effects across the first two waves of the LCCP that might be hidden by the staggered treatment analysis of Table 2.3. We replicate our previous analysis separately for each wave. Panel A and B in Table 2.9 show the estimates for the treatment effects on CO_2 emissions per capita and the CO_2 intensity of GDP for the different waves.[24] The estimates are broadly consistent with our baseline results above in that they confirm that the LCCP had no statistically significant effect in the first wave for both outcomes, and for the carbon intensity of GDP in the second wave. The coefficient for emissions per capita in the second wave, however, is much

[24]See Figures D.7 and D.8 in Appendix for the synthetic control fits.

TABLE 2.8: Robustness checks – changing predictor sets

	Baseline	S1	S2	S3	S4	S5	S6
CO_2 emissions per capita	-0.236	-0.081	-0.266	-0.274	-0.221	-0.241	-0.241
	(0.223)	(0.209)	(0.204)	(0.216)	(0.227)	(0.226)	(0.280)
GDP CO_2 intensity	-0.034	-0.058	0.010	-0.015	-0.008	-0.041	-0.016
	(0.087)	(0.135)	(0.095)	(0.108)	(0.087)	(0.096)	(0.099)

Notes: (i) The table displays the estimates of examining the sensitivity to different predictor sets on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP using staggered synthetic control method. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.
Baseline: Predictor set as in the main results.
S1: Only outcome variables in the pre-treatment periods as predictors.
S2: Outcome variables and GDP per capita as predictors.
S3: Outcome variables, GDP per capita and industrialisation rate as predictors.
S4: Outcome variables, GDP per capita, industrialisation rate and social fixed asset investment as predictors.
S5: Outcome variables, GDP per capita, industrialisation rate, social fixed asset investment, industrial SO_2 discharge and employment predictors.
S6: Outcome variables, GDP per capita, industrialisation rate, social fixed asset investment, industrial SO_2 discharge, employment and expenditure on science and technology as predictors.

larger than the one in the baseline and marginally significant, with a *p*-value of 0.085. While these results *per se* do not change our overall assessment of the policy, it might suggest that any benefits of the LCCP are rather muted in the short to medium term but might take longer to materialise. The difficulty with this type of reasoning, of course, is that the counterfactual might become rather less convincing over longer periods of time, akin to a violation of the Stable Unit Treatment Value Assumption (SUTVA).

As a further test for the robustness of our results, we now distinguish between cities that are assigned to treatment directly (which we refer to as city-level treatment) versus cities that are assigned treatment status as part of a province-level treatment assignment. The rationale for this further test is the two types of treatments might differ with respect to the enforcement pressure.[25] Table 2.10 reports on the outcome of this test, showing that the treatment effect is insignificant, irrespective of the level of their assignment into treatment.[26]

Our next robustness check is conducted to ensure that our insignificant results do not arise because

[25] We thank one anonymous reviewer for suggesting this additional test.
[26] See Figures D.9 and D.10 in Appendix for the synthetic control fits.

TABLE 2.9: Checking for robustness — results for individual waves

	ATT estimate	Std. err.	p-value
Panel A: LCCP first wave			
CO_2 emissions per capita	-0.274	0.340	0.420
GDP CO_2 intensity	0.050	0.146	0.732
Panel B: LCCP second wave			
CO_2 emissions per capita	-0.511*	0.297	0.085
GDP CO_2 intensity	0.010	0.114	0.930

Notes: (i) The table displays the estimates of the synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for individual waves of the LCCP. Panel A shows the results of the first wave; panel B shows the results of the second wave. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

TABLE 2.10: Checking for robustness – heterogeneous treatment levels

	ATT estimate	Std. err.	p-value
Panel A: City-level treatment			
CO_2 emissions per capita	-0.410	0.316	0.195
GDP CO_2 intensity	-0.012	0.084	0.882
Panel B: Province-level treatment			
CO_2 emissions per capita	-0.248	0.496	0.617
GDP CO_2 intensity	-0.036	0.124	0.769

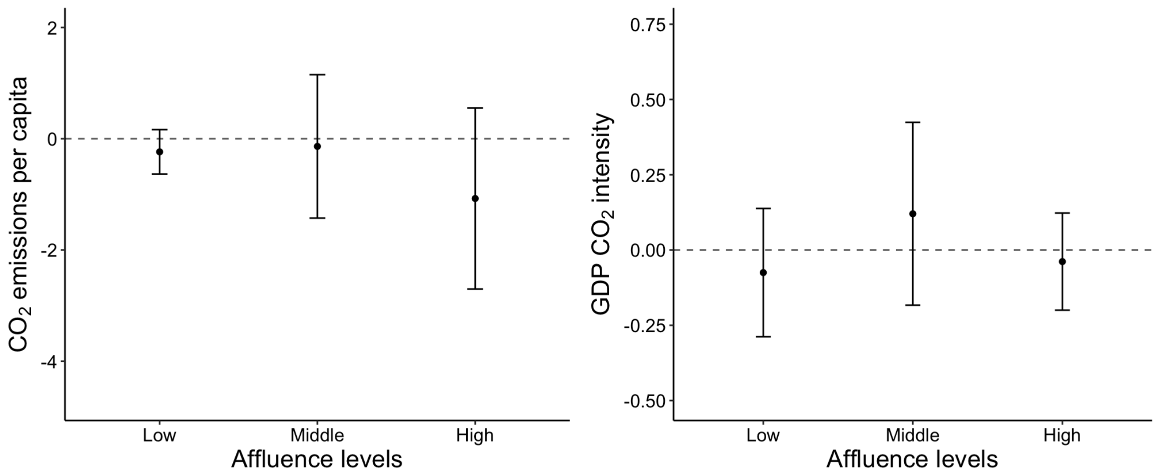
Notes: (i) The table displays the estimates in different administrative levels on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP using staggered synthetic control method. Panel A shows the results of the city-level treatment; panel B shows the results of the province-level treatment. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

of an averaging of heterogeneous treatment effects across units. In particular, we are concerned that, given the significant differences that exist across more developed regions and less developed ones in China, our aggregate results might not be very informative as to the actual impact of the LCCP. We, therefore, group the treated cities by affluence level and by geographical position before running our SCM tests again separately for each group – Table D.3 in Appendix D provides the details.

Starting with the results by income level, we use the cities' GDP per capita to proxy for the units' level of economic development, grouping them by mean GDP per capita between 2003 and 2017.

Specifically, we define three groups: cities with mean GDP per capita smaller than 35K CNY are defined as low-income cities; those between 35K and 65K CNY are defined as middle-income cities; those in excess of 65K CNY are defined as high-income cities.

Figure 2.7 – the estimates can be found in Panel A of Table D.2 in Appendix D – reports the results of this exercise.[27] Once again, the treatment effects remain clearly insignificant across all groups for both outcome variables although the precision of the estimates varies greatly.



Notes: The figure shows the results of differential effects in cities with different affluence levels on CO_2 emissions per capita and GDP CO_2 intensity using the staggered synthetic control method.

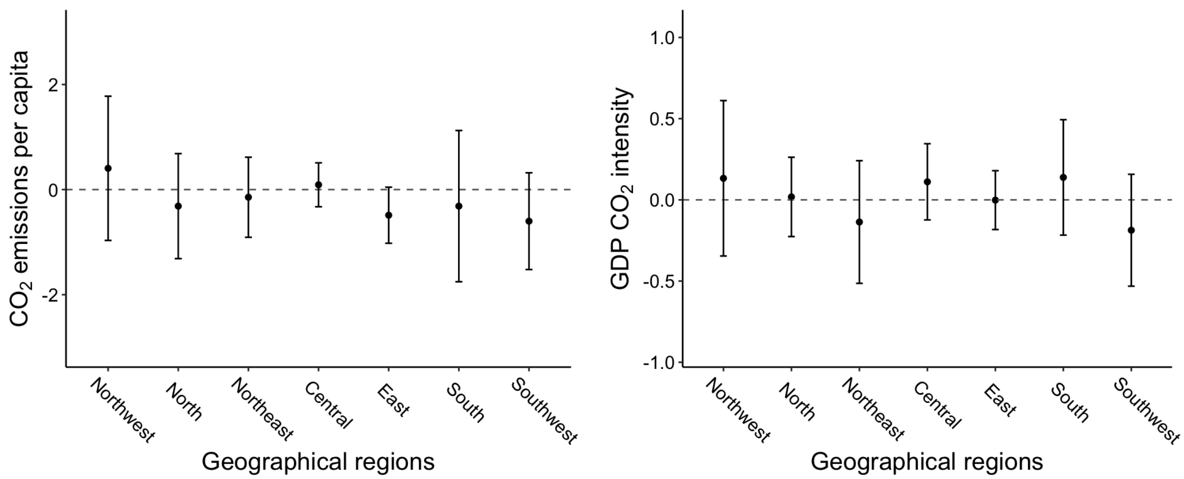
FIGURE 2.7: Checking for robustness -– cities with different affluence levels

We now turn to possible heterogeneous impacts across different regions in China. Chinese regions differ substantially from each other by their different degree of reliance on coal, and the quality of their infrastructures, for example. These differences make it likely that carbon emissions mitigation would happen at different rates. We classify the treated cities into regions according to the framework for Chinese human geography proposed by Fang et al. (2017). Based on the cities' location, we are able to estimate treatment effects across seven regions.

Figure 2.8 – and Panel B of Table D.2 in Appendix D – reports the results of this analysis.[28] Also in this case, the results suggest that the LCCP had no significant effect on carbon intensity.

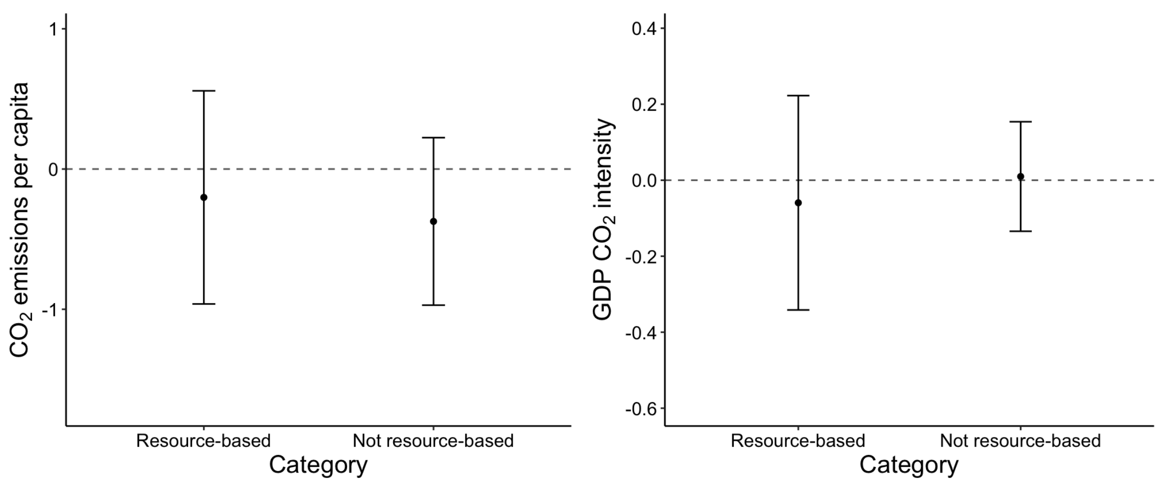
Lastly, we explore whether resource-based cities behave differently from non-resource-based cities. We define cities as resource-based if their dominant industries are based on the exploitation and

[27] See Figures D.11 - D.13 in Appendix for the synthetic control fits.
[28] See Figures D.14 - D.20 in Appendix for the synthetic control fits.



Notes: The figure shows the results of differential effects in cities in different regions on CO_2 emissions per capita and GDP CO_2 intensity using the staggered synthetic control method.

FIGURE 2.8: Checking for robustness –– cities in different geographical regions



Notes: The figure shows the results of the differential effects on resource-based and non-resource-based cities' CO_2 emissions per capita and GDP CO_2 intensity using the staggered synthetic control method.

FIGURE 2.9: Checking for robustness –– resource-based and non-resource-based cities

processing of local natural resources, based on the classification contained in the National Sustainable Development Plan for Resource-based Cities (2013–2020) issued by the State Council.[29] We perform SCMs separately for each group.

[29] See the Development Plan at `http://www.gov.cn/zwgk/2013-12/03/content_2540070.htm` (in Chinese). 262 administrative units were classified as resource-based cities, including 126 prefecture-level divisions, 120 county-level divisions, and 16 districts.

Figure 2.9 and Panel C of Table D.2 present the results.[30] The estimated treatment effects are insignificant for both categories and outcomes, again suggesting no evidence of any significant effect.

Having come so far, we are confident that our identification strategy is correct and that the methodology we deploy is appropriate for the case study at hand. We are, however, also conscious that, while the data we used so far has been extensively used in the literature, they are far from perfect. Indeed, the county-level CO_2 emission inventories our data are constructed from might be problematic, as they are down-scaled to the county level starting from provincial carbon emissions estimates based on nighttime light data. One of the problems, of course, is that nighttime light data are only able to offer a direct proxy for the electricity used for illumination and any other extrapolation (to the level of economic activity or the overall energy demand and carbon emissions) is at best the result of a noisy procedure (Shan et al., 2018a,b, 2020). Fortunately, an alternative is available in the form of consumption-based CO_2 emissions estimates using the IPCC guidelines with updated emission factors from survey studies in China.[31] The energy consumption data necessary to compile the new emission inventories are collected from the respective city-level statistical yearbook (e.g. Beijing Municipal Bureau of Statistics, 2021; Shanghai Municipal Bureau of Statistics, 2021), which also allows us to decompose the aggregate emissions into emissions from 17 different fossil fuels, 47 socioeconomic sectors, and cement production. In this section, we use these alternative emission inventories to examine the sensitivity of our results to changes in emissions data.

Using this alternative data presents us with a trade-off, however. On the one hand, the data have been shown to be more accurate and reliable; on the other hand, by relying on city-level energy consumption estimates for its construction, it only allows the construction of a narrower and shorter panel dataset. The new dataset covers the period 2005-2016 and a total of 122 cities (45 treated, 77 donor units). We report the descriptive statistics in Table 2.11, alongside the corresponding descriptive statistics from our original dataset. Overall, the two sets of emission data appear noticeably different, especially in terms of the minimum-maximum spread. This is likely because the original emissions data obtained by downscaling the nighttime light data may average out the extreme values.

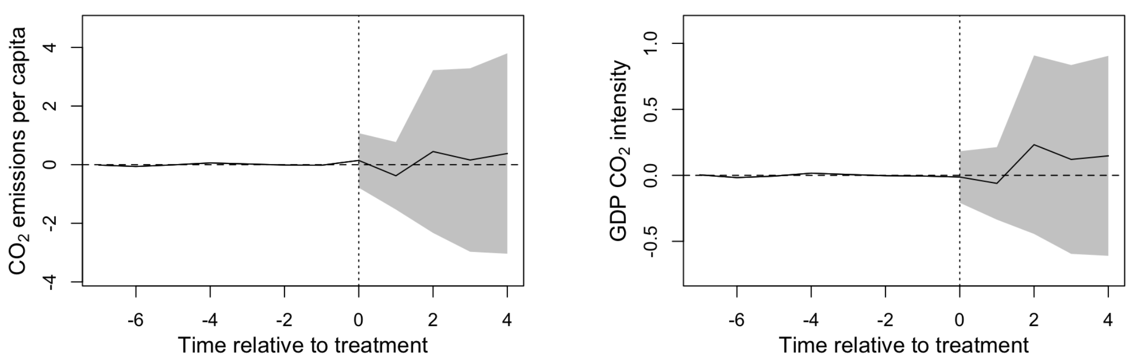
We examine the sensitivity of our results to using different datasets by applying the SCM using the

[30] See Figures D.21 and D.22 in Appendix for the synthetic control fits.
[31] See Appendix B for a discussion of how this inventory is constructed.

TABLE 2.11: Comparison of emissions data between different sources

	Mean	Std. dev.	Min.	Max.	Obs.
Panel A: Cities received treatment					
IPCC Guidelines:					
CO_2 emissions per capita (ton/person)	6.94	4.72	0.51	39.74	540
GDP CO_2 intensity (ton/10K CNY)	1.97	1.44	0.35	10.27	540
Nighttime light:					
CO_2 emissions per capita (ton/person)	6.13	2.71	1.73	14.05	540
GDP CO_2 intensity (ton/10K CNY)	1.75	0.91	0.31	5.87	540
Panel B: Donor units					
IPCC Guidelines:					
CO_2 emissions per capita (ton/person)	12.90	19.75	0.53	177.34	924
GDP CO_2 intensity (ton/10K CNY)	3.17	3.29	0.37	28.63	924
Nighttime light:					
CO_2 emissions per capita (ton/person)	8.24	7.01	1.14	53.54	924
GDP CO_2 intensity (ton/10K CNY)	2.24	1.28	0.57	10.99	924

Notes: The table compares the means, standard deviations, minimum and maximum values as well as the number of observations using data collected using the IPCC Guidelines and data based on nighttime light data. Panel A displays the values of the pilot cities. Panel B displays the values of the never-treated cities.



Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity, using outcome variables calculated based on the IPCC Guidelines using city-level statistics on energy use.

FIGURE 2.10: Checking for robustness – alternative emissions data

IPCC data as the basis to construct alternative outcome variables. For comparability, we use the same covariates and definition of the treatment group and exclude cities regulated by China's ETS

TABLE 2.12: Checking for robustness – alternative emissions data

	ATT estimate	Std. err.	*p*-value
Panel A: Staggered estimation			
CO_2 emissions per capita	0.148	1.080	0.891
GDP CO_2 intensity	0.085	0.239	0.721
Panel B: LCCP first wave			
CO_2 emissions per capita	-0.992	0.959	0.301
GDP CO_2 intensity	-0.103	0.237	0.664
Panel C: LCCP second wave			
CO_2 emissions per capita	1.005	1.939	0.604
GDP CO_2 intensity	0.262	0.383	0.494

Notes: (i) The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP. The outcome variables are calculated based on the IPCC Guidelines using city-level statistics on energy use. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

pilots or whose predictors do not fall in the convex hull.[32] Figure 2.10 and Table 2.12 present the results.[33] Despite the change in data, the results are consistent with those in Figure 2.4 in that the treatment effects are clearly insignificant. The only difference of relevance is that using this alternative data, the marginal significance of the treatment effect on per-capita emissions in the second wave vanishes. To conclude this section, we believe that our robustness checks support the idea that our main results are correct in that the LCCP had negligible effects on the treated cities.

2.5 Discussion and sectoral analysis

Until now, we have focused our attention on gauging the effect of the LCCP on two key variables of interest in climate policy debates, namely CO_2 emissions per capita and the carbon intensity of GDP. Consistently, our efforts in this chapter show that the introduction of this pilot scheme had no significant differential effect on the treated cities. Indeed, there seems to be no doubt that the effect of the policy has been negligible.

[32] We exclude 25 ETS-regulated cities and two outliers (Suzhou and Qingdao) from the sample. After the exclusion, we have 18 treated units and 77 donor units left.

[33] See Figures D.23 and D.24 in Appendix for the synthetic control fits for individual waves.

In this section, we focus on a few potential channels that we would expect to underpin the 'demonstration' effect expected of the policy. In particular, we investigate signs of an increase in investment activity in treated cities versus their untreated counterparts. We look for investments in both physical and knowledge capital as we would expect that the LCCP would provide incentives to innovation in treated areas, and/or that older machinery and infrastructure would need replacing to support a low-carbon transition.

We use data on 'Expenditure on Science and Technology' and 'Social fixed asset investment' as proxies for the type of investment activities discussed above. Table 2.13 reports the results of our staggered synthetic control estimations.[34] We find no evidence that LCCP cities are investing more than other cities, at least at this level of aggregation.

TABLE 2.13: Expenditure on science and technology and social fixed capital

	ATT estimate	Std. err.	*p*-value
Expenditure on science and technology	0.005	0.032	0.880
Social fixed asset investment	0.551	0.502	0.272

Notes: (i) The table displays the estimates of the staggered synthetic control method on each of the outcomes for the first two waves of the LCCP on a restricted sample that excludes all the cities taking part in the ETS pilots. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

One possible explanation for this pattern could be that low-carbon investments might have simply crowded out other types of investments, leaving the total unchanged. Unfortunately, we have no disaggregated investment data to test for this.

Even if aggregate investment remained constant, we should be able to infer the presence of low-carbon structural changes via changes in sectoral emissions. If any sector became relatively greener in LCCP cities than in control ones, we should observe changes in CO_2 emissions patterns across sectors. Similarly, any greening of economic activity should be flagged up by fuel switching, e.g. moving from coal to gas in manufacturing or a reduction in oil consumption in the transportation sector. The data constructed following the IPCC methodology discussed in the previous section

[34]We exclude four treated cities in evaluating the impact of the LCCP on investment in science and technology. These cities exhibit large numbers on this outcome, which we are unable to find appropriate counterfactuals. See Figure D.25 in Appendix for the synthetic control fits

provides an unprecedented wealth of information in this context. We next use this data to present a sectoral analysis of the impact of the LCCP.

TABLE 2.14: Sectoral analysis by fuel type

Sectors	CO_2 emissions			
	Total	Coal products	Gas	Oil products
Agriculture	-0.060 (0.044)	-0.021 (0.039)	0.000 (0.001)	0.011 (0.030)
Mining	-0.673 (0.656)	-0.790 (0.622)	0.026 (0.019)	-0.006 (0.009)
Light manufacturing	-0.161 (0.182)	-0.078 (0.145)	-0.002 (0.036)	-0.008 (0.012)
Heavy manufacturing	-0.318 (1.364)	-0.314 (1.034)	0.121 (0.350)	-0.156 (0.157)
High-tech manufacturing	0.005 (0.035)	0.011 (0.036)	-0.001 (0.006)	-0.006 (0.005)
Energy supply sector	-3.106 (2.033)	-3.239* (1.966)	0.077 (0.093)	0.003 (0.004)
Construction	-0.004 (0.039)	-0.012 (0.013)	-0.001 (0.001)	0.021 (0.019)
Transportation	-0.062 (0.190)	-0.022 (0.023)	-0.008 (0.017)	-0.161 (0.268)
Service sector	-0.046 (0.204)	-0.127 (0.183)	0.035 (0.028)	0.045 (0.053)
Household usage	0.049 (0.187)	-0.028 (0.129)	-0.004 (0.040)	-0.060 (0.047)

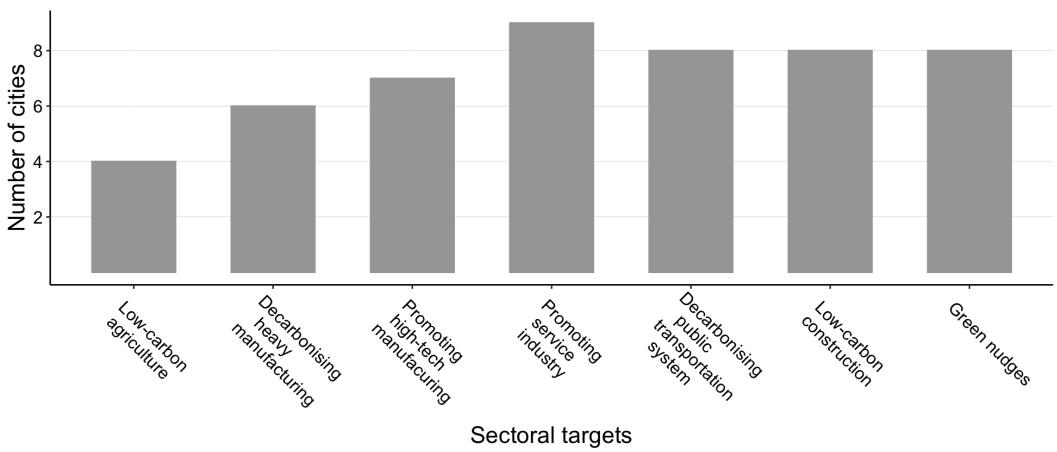
Notes: (i) The table shows the treatment effects on CO_2 emissions for the first two waves of the LCCP using the staggered synthetic control method. Results are divided by fuel type and economic sector. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Given the lack of sector-specific GDP data, we focus here on sectoral-level CO_2 emissions as our outcome variable of interest. In what follows, for ease of exposition, we group the 47 socioeconomic sectors and 17 fossil fuels available to us into broader categories – see Table D.4 and Table D.5 in Appendix D for the details. As before, we exclude the ETS-regulated cities and the cities whose predictors do not fall in the convex hull of the donor set.[35]

[35] Apart from the ETS-regulated cities, we exclude Shenyang from the sample for the evaluation of CO_2 emissions from oil products in agriculture; Shijiazhuang for the evaluation of coal products in light manufacturing and the evaluation of oil products in high-tech manufacturing; Ningbo and Wenzhou for the evaluation of oil products in the power supply

Table 2.14 summarises the synthetic control estimates for the impact of the LCCP on CO_2 emissions for each broader sector, by fuel. Even at this level of disaggregation, we fail to find any evidence of a low-carbon transition brought about by the LCCP. This is surprising, given the degree of flexibility afforded to each city to focus its efforts on specific sectors, or on specific energy uses.

In fact, reading through the details contained in the online agendas published by the LCCP pilot cities, we found a surprising degree of consistency in the type of targets they set (See also Khanna et al., 2014). Figure 2.11 provides the distribution of the sectoral targets across the 11 cities for which we are able to locate an online agenda. Most of these cities published targets aimed at promoting the service sector, decarbonising the public transportation system, boosting low-carbon construction and introducing green nudges.[36]



Notes: The figure shows the distribution of cities according to their LCCP sectoral mitigation targets, based on the information contained in their online agendas.

FIGURE 2.11: Distribution of cities by LCCP sectoral-level target

Our next step is to focus on the cities whose agendas we used to construct Figure 2.11. Using this (admittedly small) set of cities, we once again drill down to the sector/fuel level. Table 2.15 provides the results of this more focused analysis, including a sectoral-level analysis by a level-of-treatment split, similar to our discussion in Table 2.10.

sector; Shijiazhuang and Xi'an for the evaluation of gas in the service industry; Shenyang, Dalian and Qingdao for the evaluation of oil products in the service industry; Hangzhou and Xi'an for the evaluation of gas in household usage. These cities are all outliers based on the selected predictors.

[36]As discussed in the footnote in section 2.2, for the cities that did not publish online agendas or their online agendas are not traceable, we contacted the regional DRC for additional information. For these cities, however, we were unable to distinguish the details of their sectoral targets, either because they do not have specific ones, or because their agendas have been subsumed into the 12th FYP.

TABLE 2.15: Sectoral analyses by fuel type

Sectors	CO_2 emissions			
	Total	Coal products	Gas	Oil products
Agriculture	-0.024 (0.070)	0.022 (0.052)	0.000 (0.001)	0.031 (0.052)
Heavy Manufacturing	3.135 (2.705)	2.197 (1.928)	0.606 (0.590)	-0.220 (0.360)
High-tech Manufacturing	0.046 (0.062)	0.046 (0.057)	0.007 (0.011)	0.001 (0.007)
Construction	0.023 (0.069)	-0.005 (0.019)	-0.002 (0.002)	0.027 (0.026)
Transportation	-0.383* (0.226)	-0.020 (0.039)	0.005 (0.031)	-0.321 (0.286)
Service sector	0.100 (0.274)	0.012 (0.257)	0.060** (0.029)	0.064 (0.073)
Household usage	0.308 (0.249)	0.071 (0.167)	0.023 (0.060)	-0.026 (0.084)

Notes: The table shows the treatment effects on CO_2 emissions for the first two waves of the Low Carbon-City Pilot using the staggered synthetic control method. Results are divided by fuel type and economic sector. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Overall, while acknowledging that the data we use represent just a subset of the overall population of treated cities, using consumption-based data reveals the surprising result that the LCCP had no statistically significant effect on sector-level emissions across China, even in sectors that were set clear targets. The only exception is a statistically significant increase in emissions related to natural gas use in the Service sector, which might signal some degree of fuel switching to a cleaner fuel. Interestingly, also the treatment-level analysis also returns null results, suggesting that the intensity of enforcement did not play a role either.

Coming towards the end of our analysis, we must conclude that the LCCP had no significant impact on the carbon emissions of the treated cities, either at an aggregate level or at a sectoral level. We want to conclude our analysis by checking whether we can find evidence that the LCCP might have led to an increase in the costs of production in treated cities. Not having a direct way to assess these costs, we look at the level of employment across treated and non-treated cities, as well as their GDP. Our last results, in Table 2.16 show that, neither in the full sample nor among the cities that have published online agendas, we can find any evidence that the LCCP made any difference to the treated cities.[37]

TABLE 2.16: Staggered synthetic control estimation – other outcomes

	ATT estimate	Std. err.	p-value
Panel A: Full sample			
GDP	0.641	3.558	0.857
GDP per capita	0.134	0.129	0.298
Employment	-0.039	0.056	0.488
Panel B: Cities with published agendas			
GDP	-0.936	4.738	0.843
GDP per capita	0.156	0.146	0.285
Employment	-0.047	0.053	0.375

Notes: (i) The table displays the estimates of the staggered synthetic control method on other outcomes for the first two waves of the LCCP on a restricted sample that excludes all the cities taking part in the ETS pilots. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

2.6 Concluding remarks

In this chapter, we set out to evaluate the effectiveness of early climate policy efforts in the largest emerging economy in the world. We focus on the LCCP because it was the first policy implemented to mitigate climate change and because it was hailed as the first significant step taken by the Chinese authorities in the transition to a more sustainable development path. From this point of view, a careful assessment of the policy's impacts is essential to make (further) progress towards mitigating climate change. Our focus is, moreover, motivated by the fact that, despite its idiosyncratic design, the LCCP has been recognised as a success story in much of the literature (e.g. Feng et al., 2021; Hong et al., 2021; Huo et al., 2022; Tu et al., 2022). Our results, unfortunately, hardly support this optimistic view. We find no evidence that the LCCP led to a significant reduction in carbon emissions per capita, nor did it have a significant impact on the carbon intensity of GDP.

[37] As previously mentioned, we exclude treated cities exhibiting large values in the outcomes of interest for which we are unable to find appropriate counterfactuals. In Panel A, two cities are excluded when evaluating the impact of the LCCP on GDP, and one city is excluded when assessing the impact on employment. See Figures D.26 - D.28 in Appendix for the synthetic control fits.

Our results contrast with the existing literature on the LCCP that emphasises a small but generally positive impact of the pilot.[38] We are confident, however, that our work benefits from a more careful identification strategy and better accounts for the impacts of overlapping policies. In particular, our use of the synthetic control method increases our confidence that our counterfactuals are not biased by the non-random nature of the selection of the cities into the pilot. Furthermore, we are particularly careful in controlling for the impacts of China's ETS pilots, which partially overlap with the LCCP, and for the measures contained in the 12th and 13th Five-Year Plans, both of which have received little attention in the literature and might have been important confounding factors in the results published to date.

In our analysis, besides considering the main outcomes of interest, we forensically discuss both the identification strategy and the robustness of the baseline results. We also include an analysis of the main channels through which the demonstration role of the policy would likely play out. We find no evidence that the LCCP led to an increase in investment in either physical capital or science and technology, both of which would be expected to play a key role in any low-carbon transitions. We also test for evidence that the LCCP might have put pressure on pilot cities through an increase in production costs. Indeed, neither the level of economic activity nor the level of employment shows any deviation from the relevant counterfactual. Using the rich sectoral level energy consumption data collected from the respective city-level statistical yearbook (e.g. Beijing Municipal Bureau of Statistics, 2021; Shanghai Municipal Bureau of Statistics, 2021), we are able to construct emissions data for different economic sectors and by fossil fuel type. Using this data we are the first to be able to discuss the sectoral impact of the LCCP. Our analysis shows that, even at such a disaggregated level, we cannot identify any impact of the LCCP.

Overall, our work leads us to conclude that the measures introduced by the treated cities as part of the LCCP failed to generate a differential response by the economic agents operating in their jurisdictions. That is not to say, of course, that China's climate policy efforts had no mitigating effects, based on our results, however, we can clearly conclude that – in the context of a country that was starting to ready itself for a lower-carbon future – the LCCP failed to mobilise sufficient resources, political attention and creativity to galvanise a low-carbon transition. Given the general lack of evidence of any significant change over time, across regions and economic sectors, we must

[38] It is fair to point out that our results also contradict the findings of the contributions that find significant and negative impacts linked to the LCCP (e.g Zhou and Zhou, 2021; Feng et al., 2021).

conclude that the LCCP's design was simply not conducive to generating sufficient incentives to cause a significant response across the economy.

While our results run counter the existing literature, our conclusions in fact are well aligned with the theoretical priors on the likely impacts of the LCCP. The LCCP was designed and introduced fundamentally as a voluntary scheme, which the administrators of Chinese cities might sign up to. The policy provided only vague ambitions to 'demonstrate pathways' to a transition to the low-carbon economy. The scheme also lacked explicit mandates in terms of the instruments to use and had no specific quantitative target. The policy also lacked any actual enforcement mechanism. On all these grounds, we would indeed not expect the policy to have made much of a difference to the choices of the agents in the economy. From this point of view, the main lesson to be drawn from our analysis is that, even among emerging economies, the design of effective environmental policy requires the careful setting of transparent and quantifiable targets, the introduction of economic instruments that affect economic incentives, and credible enforcement mechanisms.

Chapter 3

Climate policy and environmental efficiency

3.1 Introduction

As the climate emergency takes centre stage, calls for the introduction of effective climate policies have been growing louder across the world (Shukla et al., 2022). With most of the growth in global energy demand – and therefore emissions – over the coming decades expected to originate in emerging markets (Wei et al., 2012), an understanding of whether climate policies are leading to a more efficient use of carbon resources in the fastest-growing emerging economies is particularly important to the debate on how to best approach the issue of mitigation.

Built on the background that we introduced in Chapter 2, in this chapter we continue our analysis of the LCCP on low-carbon economy transition, instead focusing on its impact on environmental efficiency measured with a stochastic frontier analysis (SFA) framework. By allowing for explicit trade-offs among inputs and outputs, it provides a more holistic view of the overall environmental efficiency performance of the units and a more nuanced understanding of the transition to a low-carbon economy.

The existing literature on the LCCP provides results as Yu et al. (2019), Hong et al. (2021) and Huo et al. (2022) to the improvements of the LCCP on carbon emissions, whereas Zhang et al. (2022a) provides more convincing causal evidence that the LCCP had no impact. Ma et al. (2021); Yang (2023) support that the LCCP has led to significant increases on low carbon innovation. More closely related to our focus here, Cheng et al. (2019) investigate the impact of the LCCP on efficiency growth, using a prefecture-level panel dataset over the period 2007-2016. They use data envelopment analysis (DEA) with a slack-based approach to measure environmental efficiency, with CO_2 emissions as undesirable output. Based on a standard difference-in-difference (DiD) framework, they find that the LCCP significantly promoted efficiency growth. Methodologies adopted by Liu et al. (2020a), Fu et al. (2021), Yu et al. (2021), Shi and Xu (2022), Wen et al. (2022), Zhang et al. (2022b), Wang et al. (2023) and Yang et al. (2023) are broadly consistent with Cheng et al. (2019), all suggesting significant increases on environmental efficiencies. Chen et al. (2021) instead investigate the impact of the LCCP on productivity over the period 2005-2015. They estimate productivity based on the uniform semi-parametric estimation method proposed by Olley and Pakes (1992). Using a standard DiD framework with propensity score matching, they find that the LCCP has significantly increased firms' productivity.

While these contributions seem to suggest favourable outcomes delivered by the LCCP, the identification frameworks proposed in the literature so far seem worth challenging on at least three grounds. First, by its very nature as a voluntary measure, the assignment to treatment within the LCCP cannot be assumed to be random by any stretch of the imagination. Second, recent advances in econometric theory advise against using the standard two-way fixed effects method in the presence of heterogeneous treatment effects (e.g. Goodman-Bacon, 2021; Baker et al., 2022). To the best of our knowledge, most – if not all – studies fail to recognise the staggered nature of the LCCP. Such modelling issue might bias the estimated treatment effect in the existing literature.[1] Third, as claimed by the Chinese government, one of the aims of the LCCP is to demonstrate the low-carbon transition to other non-pilot cities (NCSC, 2020, in Chinese). Constructing a meaningful counterfactual in this context is therefore questionable, due to the potential diffusion of low-carbon process from the pilots to non-pilot cities.

In this chapter, we revisit the impact of the LCCP by carefully accounting for the potential biases mentioned above. We first assemble a dataset including socioeconomic measurements and CO_2 emissions for 260 cities over the period 2003-2016. We then specify an enhanced hyperbolic distance function proposed by Cuesta et al. (2009), and estimate the environmental efficiency using

[1] Given the staggered adoption and the substantial differences in the treated units across waves, heterogeneous treatment effects are indeed likely.

SFA. To estimate the treatment effect while controlling for the learning effect, we adopt the timing-based approach suggested by Miller (2023), using cities treated earlier or later as controls for one another. We use dynamic DiD to start our empirical investigation, then move on to employing the *partially pooled* synthetic control method (SCM) proposed by Ben-Michael et al. (2022) to control for the selection into treatment. Finally, we conduct several robustness checks to show that our empirical investigation is reliable.

We contribute to the literature in several ways. First, we build on the recent advances in the literature on distance function and efficiency to provide a more accurate measurement of environmental efficiency than the ones currently reported in the literature. Specifically, we adopt the enhanced hyperbolic distance function approach of Cuesta et al. (2009) and Mamardashvili et al. (2016), allowing the units to contract inputs and undesirable output, at the same time expanding the desirable output. In addition, contrary to the literature that uses DEA to estimate efficiency, we use SFA, which allows us to distinguish statistical noise from inefficiency. Second, we introduce a more careful identification framework that accounts for the learning effect of the treatment while controlling for the selection bias. The timing-based approach also allows us to gauge the learning effect. To the best of our knowledge, we are the first to investigate such unique policy design.

Contrary to the existing literature, our analysis leads us to conclude that the LCCP did not significantly increase treated cities' environmental efficiency. Nevertheless, we find evidence of a positive learning effect in the second wave of the LCCP, although it is not persistent and only significant in the short run. We conclude that the LCCP might not be stringent enough to promote a consistent transition to a low-carbon economy.

This chapter is organised as follows. We explain the methodology and data in Section 3.2, and then present our results in Section 3.3. Section 3.4 concludes.

3.2 Methodologies

3.2.1 Environmental efficiency

Following recent advances in the field of productivity analysis, we exploit the enhanced hyperbolic distance function introduced by Färe et al. (1985) and Färe et al. (1989), and adapted by Cuesta et al. (2009) and Mamardashvili et al. (2016) to measure environmental efficiency. We measure the outcome to the production frontier following a hyperbolic path by contracting the use of input and undesirable output, at the same time increasing the desirable output.

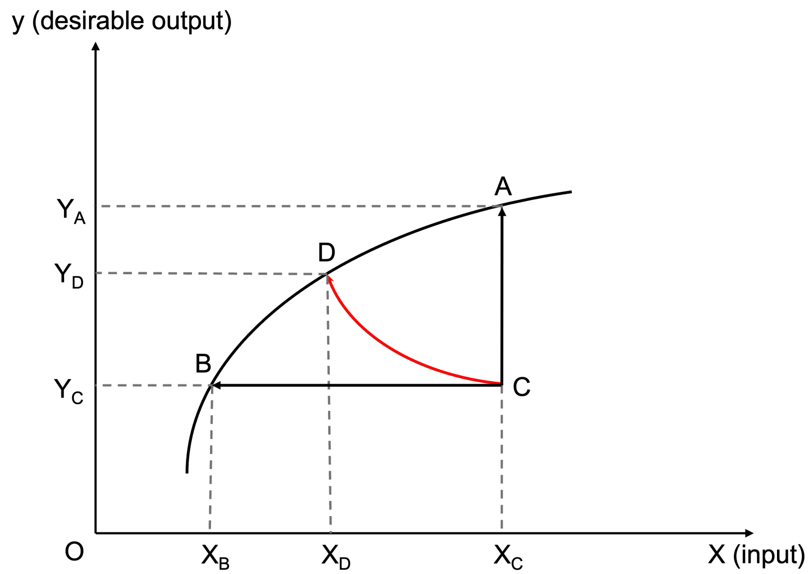


FIGURE 3.1: Graphical demonstration of the allocative adjustment

Figure 3.1 illustrates the methodology. Unit C sits inside of the production possibility set and is therefore technically inefficient. From the output-oriented point of view, the unit can increase its desirable output from Y_C to Y_A while holding the use of input X_C unchanged. Similarly, under the input-oriented approach the unit might decrease its input use from X_C to X_B, while holding the level of output Y_C unchanged. The economic implications are that the output-orientation is consistent with the revenue-maximisation behaviour, whereas the input-orientation links to the cost-minimisation behaviour (Cuesta and Zofío, 2005). The projection of interest in this chapter performs a hyperbolic path onto the frontier D, by contracting the input and undesirable output from X_C to X_D, and at the same time increases its desirable output from Y_C to Y_D.

In what follows, we specify the enhanced hyperbolic distance function and estimate the efficiency following Cuesta et al. (2009) as:

$$D(x, y, b) = \min_{\theta}\{\theta : (\theta x, \frac{y}{\theta}, \theta b) \in T\}, \tag{3.1}$$

where T is the production possibility set, and θ is the efficiency of interest that equiproportionately expands desirable output y and contract input x and undesirable output b.

Following Mamardashvili et al. (2016), we define the enhanced hyperbolic efficiency as:

$$HE_{it} = D(x, y, b). \tag{3.2}$$

Using the almost homogeneity property yields:

$$D(\frac{x}{\theta}, \theta y, \frac{b}{\theta}) = \theta D(x, y, b) \quad \forall \quad \theta > 0. \tag{3.3}$$

Letting $\theta = \frac{1}{y_M}$, where y_M refers to the M^{th} output, we have

$$D(x \cdot y_M, \frac{y}{y_M}, b \cdot y_M) = \frac{1}{y_M} \cdot D(x, y, b). \tag{3.4}$$

After taking logarithms on both sides, we obtain

$$lnD(x, y, b) = lnD(x \cdot y_M, \frac{y}{y_M}, b \cdot y_M) + lny_M. \tag{3.5}$$

Substitute equation (5) back into equation (2) yields

$$-lny_M = lnD(x \cdot y_M, \frac{y}{y_M}, b \cdot y_M) - lnHE_{it} + v_{it}, \tag{3.6}$$

which is an estimable form of the enhanced hyperbolic distance function, where v_{it} is statistical noise.

The enhanced hyperbolic distance function can be estimated in either a parametric or a non-parametric framework. DEA proposed by Charnes et al. (1978) is commonly used when it comes to non-parametric framework. In contrast, parametric estimation relies on SFA. While DEA does not require a functional form, its deterministic nature makes separating efficiency change from random shock impossible. In addition, DEA estimates are serial-correlated, which complicates inference in the potential second stage (Simar and Wilson, 2007).

We specify a translog hyperbolic distance function with three inputs x_{kit} including labor, capital and energy with $k \in \{1,2,3\}$, GDP as the desirable output y_{it} and CO_2 emissions as the undesirable output b_{it}:

$$\begin{aligned} -ln y_{it} &= \alpha_0 + \sum_{k=1}^{3} \alpha_k ln(x_{kit} y_{it}) + \frac{1}{2} \sum_{k=1}^{3} \sum_{l=1}^{3} \alpha_{kl} ln(x_{kit} y_{it}) ln(x_{lit} y_{it}) \\ &+ \beta_b ln(b_{it} y_{it}) + \frac{1}{2} \beta_{bb} (ln(b_{it} y_{it}))^2 + \sum_{k=1}^{3} \beta_{kb} ln(x_{kit} y_{it}) ln(b_{it} y_{it}) + \epsilon_{it}, \end{aligned} \tag{3.7}$$

where i identifies the different cities and t for the different time periods.

To control for intertemporal technical change, we augment the hyperbolic distance function with a time trend t.[2] We assume that the technical change is associated with the production factors:

$$\begin{aligned} -ln y_{it} &= \alpha_0 + \sum_{k=1}^{3} \alpha_k ln(x_{kit} y_{it}) + \frac{1}{2} \sum_{k=1}^{3} \sum_{l=1}^{3} \alpha_{kl} ln(x_{kit} y_{it}) ln(x_{lit} y_{it}) \\ &+ \beta_b ln(b_{it} y_{it}) + \frac{1}{2} \beta_{bb} (ln(b_{it} y_{it}))^2 + \sum_{k=1}^{3} \beta_{kb} ln(x_{kit} y_{it}) ln(b_{it} y_{it}) \\ &+ \gamma_t t + \frac{1}{2} \gamma_{tt} t^2 + \sum_{k=1}^{3} \gamma_{kt} t ln(x_{kit} y_{it}) + \eta_{tb} t ln(b_{it} y_{it}) + \epsilon_{it}. \end{aligned} \tag{3.8}$$

The composed error term ϵ_{it} comprises a non-negative inefficiency term u_{it} and a random noise term v_{it}:

$$\epsilon_{it} = u_{it} + v_{it}. \tag{3.9}$$

We estimate the enhanced hyperbolic distance function SFA, and measure the efficiency HE_{it} using the estimator proposed by Battese and Coelli (1988) in a maximum likelihood estimation framework:

$$HE_{it} = E(e^{-u_{it}} | \epsilon_{it}). \tag{3.10}$$

Following Orea et al. (2015), we centre the production inputs at the sample mean to facilitate convergence. We assume a half-normal distribution for u_{it} and a normal distribution for v_{it}. Following

[2] Specifically, change in environmental efficiency is related to efficiency change and technical change, where the later component may shift the production frontier which, makes the efficiency of interest less comparable across periods. Our exercise here is equivalent to adding a year fixed effect in estimating the environmental efficiency, which allows the efficiency estimation to be more accurate and comparable.

Mamardashvili et al. (2016), we allow u_{it} and v_{it} to be heteroskedastic:

$$\sigma^2_{u,it} = e^{z_i^{'}\rho}, \tag{3.11}$$

$$\sigma^2_{v,it} = e^{w_i^{'}\tau}, \tag{3.12}$$

where z_i and w_i are the variables that respectively affect the variance of the environmental efficiency and random noise, and ρ and τ are the vectors of parameters to be estimated. To explain the variance of the environmental efficiency across cities, we use expenditure on science and technology, as it directly links to research and development activities (Xiong et al., 2020). We introduce a regional indicator that indexes different geographical positions for different cities to explain the variance of the random noise, following the regionalisation framework of China proposed by Fang et al. (2017).

3.2.2 Data

We collect a range of socioeconomic measurements for 285 cities from 2003 to 2016 to measure the efficiency and balance the systematic differences between the treated cities and weighted counterfactuals. In particular, we extract GDP (billion CNY), GDP per capita (10 thousand CNY), employment measured as people employed in urban units (million people), industrialisation rate measured by the share of GDP from the secondary sector (%), social fixed asset investment (10 billion CNY), and electricity usage (billion kWh) from the *China City Statistical Yearbook* (NBS, 2017). Monetary values are normalised as constant 2010 CNY or USD.

While the above statistics provide us with rich city-level information, certain variables require additional calculations as they are not readily available. In particular, we need data on capital stock and CO_2 emissions to measure environmental efficiency. We estimate capital stock using the perpetual inventory method, assuming a 4% depreciation rate, as suggested by Zhang (2008) and Jidong et al. (2014):

$$K_t = I_t + (1 - \delta)K_{t-1}, \tag{3.13}$$

where I_t is investment and δ refers to the constant depreciation rate of the capital stock. The initial capital stock is estimated following Harberger (1988):

$$K_0 = \frac{I_0}{\delta + g}, \tag{3.14}$$

where I_0 refers to the initial investment, and g refers to the average GDP growth rate over the study period.

One of the challenge that we face is to find reliable emissions data. While the estimates of CO_2 emissions based on the IPCC Guidelines are generally preferable, they are only available for a very limited set of cities, due to the lack of complete city-level statistics. We resort to the CO_2 emission inventories provided by Chen et al. (2020) as an indicator of undesirable output. The data was estimated based on nighttime light data from satellite imagery, including 2,735 counties and districts in around 350 administrative divisions from 1997 to 2017.[3] We obtain the data from the Carbon Emission Accounts Datasets and aggregated the CO_2 emissions at the city level (CEADs, 2020).

We then meticulously clean the data. First, we cross-check the data with prefectural and provincial statistical yearbooks, which are accessed via the municipal bureau of statistics in different administrative divisions, to ensure the best accuracy. Second, for data quality assurance, we exclude 18 cities (five of which were treated by the first or the second wave of the LCCP) with substantial misreported statistics from our sample. We also correct occasional missing and misreported values by employing linear interpolation for electricity usage across 72 observations out of 3,836.

We estimate environmental efficiency using production inputs as employment, capital stock, and electricity usage. We respectively use GDP and CO_2 emissions as desirable output and undesirable output. Given our specification in function 3.8, a negative coefficient suggests that the variable positively contributes to the efficiency of interest. Table 3.1 presents the frontier estimates derived from the enhanced hyperbolic distance function, where the coefficients broadly have expected signs and tend to statistically significant.

To construct counterfactuals that closely match the treated cities, we use GDP per capita, industrialisation rate, social fixed asset investment, and CO_2 emissions as additional predictors. A caveat of

[3]We decide not to use the final year of this emissions data, because the rules of compiling the *China City Statistical Yearbook* changed in the year 2017. Some city-level statistics are therefor inconsistent and are not comparable across years.

TABLE 3.1: Frontier estimates

Variable	Coefficient	Std. err.	p-value
α_1	-0.154***	0.004	0.000
α_2	-0.281***	0.005	0.000
α_3	-0.056***	0.002	0.000
α_{11}	0.098***	0.012	0.000
α_{12}	-0.101***	0.011	0.000
α_{13}	0.016***	0.005	0.001
α_{22}	0.136***	0.015	0.000
α_{23}	-0.021***	0.005	0.000
α_{33}	-0.025***	0.004	0.000
β_b	-0.037***	0.004	0.000
β_{bb}	-0.039***	0.011	0.001
β_{1b}	0.018*	0.010	0.063
β_{2b}	-0.014	0.011	0.210
β_{3b}	0.022***	0.004	0.000
γ_t	-0.001**	0.001	0.048
γ_{tt}	0.007***	0.000	0.000
γ_{1t}	0.000	0.001	0.886
γ_{2t}	-0.010***	0.001	0.000
γ_{3t}	0.004***	0.000	0.000
η_{tb}	0.006***	0.001	0.000
Constant	-0.023*	0.012	0.055
σ_u^2			
R&D	0.139***	0.000	
Constant	-5.844***	0.565	0.055
σ_v^2			
Regionalisation	0.155***	0.016	0.000
Constant	-5.310***	0.117	0.000

Note: (i) The table reports enhanced hyperbolic distance frontier estimates. Year fixed effects are included in all estimations. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively. (iii) R&D is an abbreviation for the expenditure on science and technology.

creating reliable synthetic controls is that the pre-treatment outcome and predictors of the treated unit should approximately fall into the convex hull of the donor units. In what follows, we exclude eight treated cities from our sample. These cities significantly outperform the non-pilot ones along many dimensions, which in no circumstances could we find counterfactuals. Moreover, as detailed in Section 2.4.1 of Chapter 2, the policy overlap due to the China's emission trading scheme (ETS) pilots introduces significant confoundedness. Therefore, we exclude 32 ETS-regulated cities (of

TABLE 3.2: Descriptive statistics

	Mean	Std. dev.	Min	Max	N
Panel A. Descriptive statistics for the Wave I cities					
Outcome variable:					
Environmental efficiency	0.97	0.01	0.93	0.98	378
Socioeconomic measurements:					
GDP (billion CNY)	111.82	131.65	5.06	965.16	378
GDP per capita (10 thousand CNY)	2.91	2.06	0.11	10.60	378
Employment (million people)	0.46	0.49	0.06	2.93	378
Industrialisation rate (%)	48.85	10.95	20.27	81.09	378
Electricity usage (billion kWh)	6.61	8.68	0.08	58.40	378
Social fixed asset investment (10 billion CNY)	7.68	9.07	0.26	51.41	378
CO_2 emissions (million ton)	22.51	16.65	1.76	84.43	378
Panel B. Descriptive statistics for the Wave II cities					
Outcome variable:					
Environmental efficiency	0.96	0.01	0.94	0.98	280
Socioeconomic measurements:					
GDP (billion CNY)	156.57	160.87	9.22	854.05	280
GDP per capita (10 thousand CNY)	3.34	2.03	0.42	10.29	280
Employment (million people)	0.48	0.37	0.06	1.75	280
Industrialisation rate (%)	46.66	10.68	18.57	81.09	280
Electricity usage (billion kWh)	6.77	6.83	0.23	41.66	280
Social fixed asset investment (10 billion CNY)	10.13	10.57	0.44	63.60	280
CO_2 emissions (million ton)	29.60	22.40	2.70	99.84	280
Panel C. Descriptive statistics for the non-pilot cities					
Outcome variable:					
Environmental efficiency	0.96	0.01	0.92	0.98	2,226
Socioeconomic measurements:					
GDP (billion CNY)	105.69	102.78	3.88	785.70	2,226
GDP per capita (10 thousand CNY)	2.84	2.32	0.23	20.24	2,226
Employment (million people)	0.35	0.28	0.04	2.21	2,226
Industrialisation rate (%)	49.19	11.85	2.66	90.97	2,226
Electricity usage (billion kWh)	4.72	5.87	0.08	56.51	2,226
Social fixed asset investment (10 billion CNY)	7.01	7.32	0.20	59.70	2,226
CO_2 emissions (million ton)	21.69	16.70	1.63	108.48	2,226

Note: Table shows means, standard deviations, minimum values, maximum values and number of observations of outcome variable and socioeconomic measurements from 2003 to 2016. Monetary values are normalised as constant 2010 CNY or USD.

which all were treated by the first or the second wave of the LCCP) for a clearer identification. Lastly, we exclude 29 cities that were treated by the LCCP third wave implemented in 2017, to circumvent any potential confoundedness introduced by the anticipation effect. Overall, we have 2,870 observations, including 46 treated and 159 non-pilot cities over the period 2003-2016.

Panel A of Table 3.2 displays the statistics of the cities treated by the first wave (Wave I cities, hereafter); *Panel B* displays the statistics of the cities treated by the second wave (Wave II cities, hereafter); *Panel C* displays the statistics of the non-pilot cities.[4] These three groups are comparable and could serve as counterfactuals for each other, providing us an opportunity for the timing-based approach that we will elaborate below.

3.2.3 Identification strategy

One of the primary motivations of launching the LCCP is to chart a viable path for transitioning to a low-carbon economy. Experiences gleaned from high-income cities' transitions may not be directly applicable to low-income counterparts due to their idiosyncratic characteristics. Therefore, the critical task of selecting the most suitable candidates becomes pivotal in propagating low-carbon mitigation. As elaborated in the technical report published by the National Center for Climate Change Strategy and International Cooperation (NCSC), the diversity in social and economic status is one of the key factors influencing the selection process (NCSC, 2013, in Chinese).

Another factor to consider is regional representativeness, which is linked to unobserved regional differences not captured by social or economic indicators. For example, South and Northwest China have distinct climate characteristics that influence the types of crops grown. These subtle distinctions within industries can give rise to intricate cross-industry implications when implementing low-carbon mitigation measures. In fact, the pilot cities were primarily concentrated in South and East China, but are also dispersed throughout the country.

The fact that the selection into the treatment was affected by the aforementioned factors suggests that the learning effect was factored in by the policy makers and that they introduced carbon mitigation process to diffuse to the non-pilot cities. Therefore, it is challenging to construct a meaningful counterfactual in conventional wisdom, since the non-pilot cities are, in fact, partially treated.

Figure 3.2 displays the development trajectories for three different groups: the short dashed line with triangles represents the Wave I cities; the long dashed line with squares represents the Wave II cities; the solid line with circles represents the non-pilot cities. Overall, the pre-trends are approximately parallel across the three groups. When the trend for Wave I cities started to accelerate

[4] Note that Yan'an was treated twice, since it was treated at province-level in the first wave and city-level in the second wave. We include this city in both Panel A and B, thus the sum of 'N' is larger than the number of observations.

after the introduction of the first wave, the trends for Wave II cities and the non-pilot remained unchanged. The almost flat post-trend also suggests that the treatment effect might be stable over the entire post-treatment period. Shortly after the introduction of the second wave, both Wave II cities and the non-pilot cities experienced a noticeable short-term increase. This short-term effect did not seem to persist, however, and completely disappeared in the final period.[5]

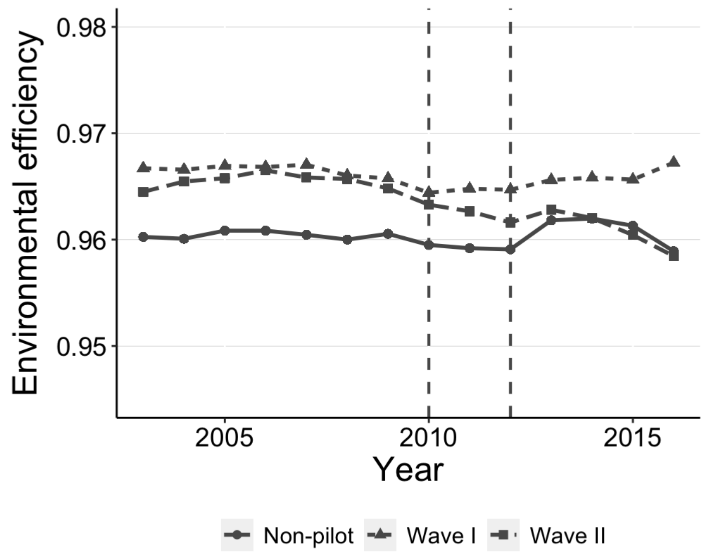


FIGURE 3.2: Mean environmental efficiency

The figure presented above seem to provide evidence that the trends of the earlier and later treated groups offer reasonably suitable counterparts for each other. First, the non-pilot cities serve as a reasonable counterfactual for the Wave I cities from 2003 to 2012, for which the non-pilot cities' trends are virtually and continuously flat after the first wave, at least visibly showing no indication of the learning effect being diffused. Second, the Wave I cities serve as a reasonable counterfactual for the Wave II cities and for evaluating the learning effect from 2010 to 2016, for which all three groups share a same treatment history within this period.

In what follows, we employ the timing-based approach recommended by Miller (2023). We select two distinct time windows where cities treated earlier or later serve as controls for one another. First, we evaluate the first wave of the LCCP using the time window that spans from 2003 to 2012,

[5]The fact that we do not observe such short-term increase after the introduction of the first wave, likely because of the treatment levels. In Wave I, most cities were assigned treatment at province-level, whereas in Wave II, most cities were assigned at city-level. Province-level treatment has larger jurisdictional area, which may increase the transaction cost, therefore limiting the diffusion of low-carbon mitigation.

where the non-pilot cities serve as controls. Second, we evaluate the second wave of the LCCP and the learning effect using the time window that spans from 2010 to 2016, where the Wave I cities serve as controls.

We employ the dynamic DiD as our first method to assess the effects of the LCCP. Following the timing-based approach, we specify three dynamic DiD models. The first model evaluates the first wave of the LCCP, which we formally express as

$$Y_{it} = \alpha + \sum_{\substack{-7 \leq k \leq 2 \\ k \neq -1}} \beta_k^{Wave\ I} \times \text{Cities}_i^{Wave\ I} \times \text{Treatment}_{t+k}^{Wave\ I} + \delta X_{it} + \lambda_i + \theta_t + \epsilon_{it}. \tag{3.15}$$

i includes Wave I cities and those non-pilot; t spans from 2003 to 2012. $\text{Cities}_i^{Wave\ I}$ and $\text{Treatment}_{t+k}^{Wave\ I}$ are dummy variables that respectively index the Wave I cities (one if the city was treated by the first wave and zero otherwise) and the introduction of the first wave (one if the first wave has been implemented and zero otherwise). X_{it} is a vector of city-level control variables. λ_i and θ_t are the city-level and time fixed effects. ϵ_{it} is the random error. The coefficient of interest is $\beta_k^{Wave\ I}$, which we specify with leads and lags, and estimate relative to the year before the implementation, $k = -1$

The second model evaluates the second wave of the LCCP. The regression is expressed as

$$Y_{i't'} = \alpha + \sum_{\substack{-2 \leq k' \leq 4 \\ k' \neq -1}} \beta_{k'}^{Wave\ II} \times \text{Cities}_{i'}^{Wave\ II} \times \text{Treatment}_{t'+k'}^{Wave\ II} + \delta X_{i't'} + \lambda_{i'} + \theta_{t'} + \epsilon_{i't'}. \tag{3.16}$$

i' differs from i as it includes both Wave I and Wave II cities, where we use Wave I cities as controls. $\beta_{k'}^{Wave\ II}$ is the coefficient of interest, which we specify with different numbers of leads and lags.

The third model evaluates the learning effect, which we consider the non-pilot cities as treated cities, and use the Wave I cities as controls. We formally express the model as

$$Y_{i''t'} = \alpha + \sum_{\substack{-2 \leq k' \leq 4 \\ k' \neq -1}} \beta_{k'}^{Non-pilot} \times \text{Cities}_{i''}^{Non-pilot} \times \text{Treatment}_{t'+k'}^{Wave\ II} + \delta X_{i''t'} + \lambda_{i''} + \theta_{t'} + \epsilon_{i''t'}. \tag{3.17}$$

i'' therefore includes the Wave I cities and the non-pilot cities. We estimate the coefficient of interest $\beta_{k'}^{Non-pilot}$ and specify the same leads and lags as model 3.16.

As introduced earlier, the selection into treatment is intentional and lacks randomness. Even worse, the voluntary nature of the LCCP suggests that the treated cities are likely self-selected into the treatment. Thus, the dynamic DiD is less satisfactory in causally identifying the effect of the LCCP.

In what follows, for each analysis we use the *partially pooled* SCM proposed by Ben-Michael et al. (2022) to mitigate these biases. We believe that their proposal outperforms other SCM proposals because it seeks an intermediate balance between the *pooled* SCMs and the *separate* SCMs, where the imbalances in both proposals determine the error of a weighting estimator for the average effect Ben-Michael et al. (2022). See Section 2.3 of Chapter 2 and Appendix A for the technical details.

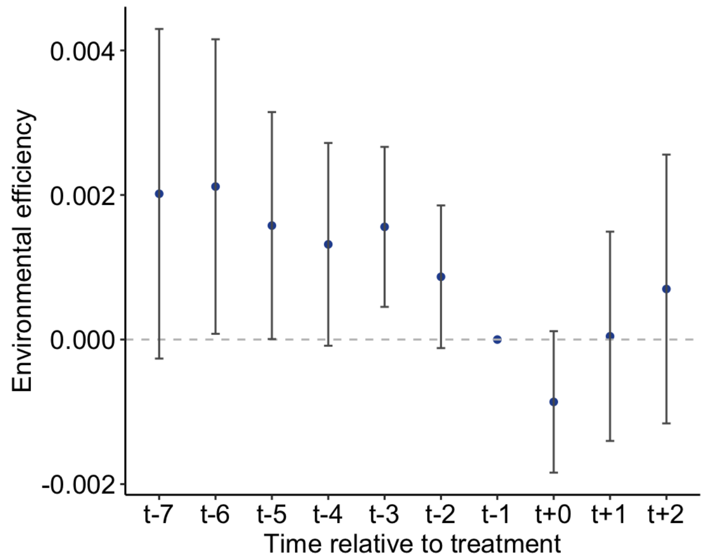
3.3 Main results

In this section, we present the results from the dynamic DiD and SCM for the three different analyses, which we start by investigating the effect of the LCCP first wave. Figure 3.3 shows the results of the LCCP first wave on environmental efficiency using the dynamic DiD approach. While the post-treatment estimates are overall insignificant, the point estimates seem to decrease over time. The persistence of this trend before $t = 0$ contradicts the parallel-trend assumption. This violation of the identification assumption diminishes the informativeness of our post-treatment estimates. This result suggests that the selection into treatment is not random, therefore naïve identification based on standard methods are not the ideal approach to identify causality.

We move on to the *partially pooled* SCM, and report the results in Figure 3.4. The pre-treatment trends are virtually flat, showing satisfactory balance between the treated cities and the counterfactual. Although the point estimates seem to increase over time, the average treatment effect is insignificant, as indicated by the 95% confidence interval.[6]

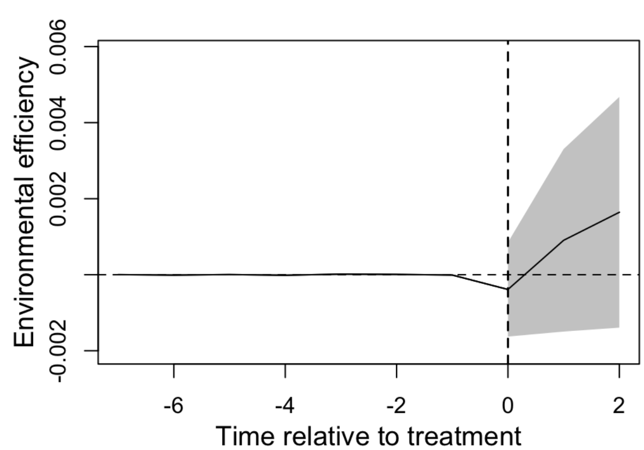
We then move on to the LCCP second wave, which we first display the results of the dynamic DiD in Figure 3.5. The post-treatment estimates are statistically insignificant, and they exhibit a negligible upward-sloping trend. Using the Wave I cities as the control group, we have only one pre-treatment estimate that is statistically positive. Similar to the previous analysis, we find that the parallel trend assumption is violated, which motivates us to the SCM to control for the selection into treatment.

[6]As discussed in the footnote in the Section 2.4 of Chapter 2, we follow Ben-Michael et al. (2022) to provide statistical inference using the leave-one-unit-out jackknife approach. See the online appendix of Ben-Michael et al. (2022) for more details.



Notes: The figure shows the effect of the LCCP first wave on environmental efficiency using the dynamic difference-in-differences approach.

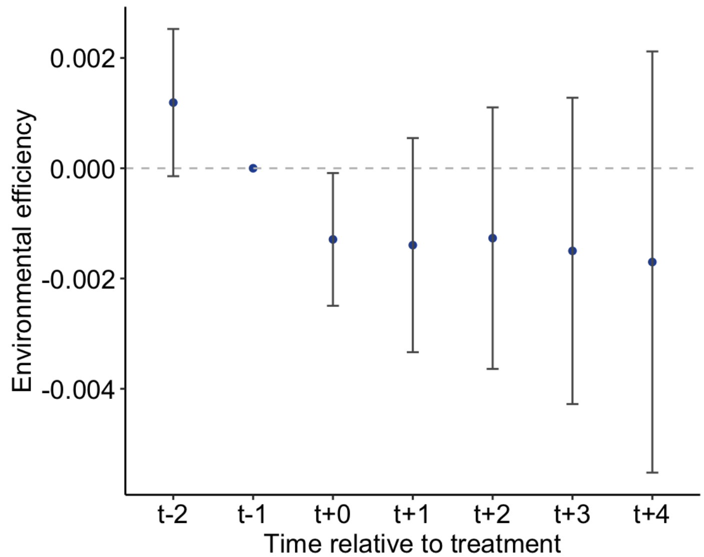
FIGURE 3.3: Effect of the LCCP first wave using a difference-in-differences approach



Notes: The figure shows the effect of the LCCP first wave on environmental efficiency using the *partially pooled* synthetic control method.

FIGURE 3.4: Effect of the LCCP first wave using the synthetic control method

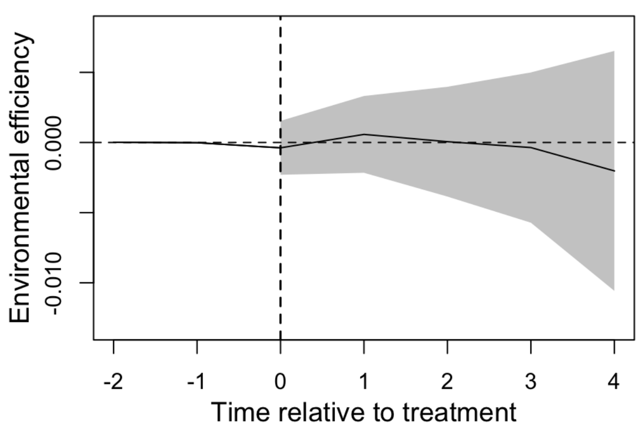
Figure 3.6 shows the results of the SCM application. The post-treatment estimates hover around the x-axis, suggesting that the LCCP second wave had no significant effect on environmental efficiency



Notes: The figure shows the effect of the LCCP second wave on environmental efficiency using the dynamic difference-in-differences approach.

FIGURE 3.5: Effect of the LCCP second wave using a difference-in-differences approach

for the treated cities.



Notes: The figure shows the effect of the LCCP second wave on environmental efficiency using the *partially pooled* synthetic control method.

FIGURE 3.6: Effect of the LCCP second wave using the synthetic control method

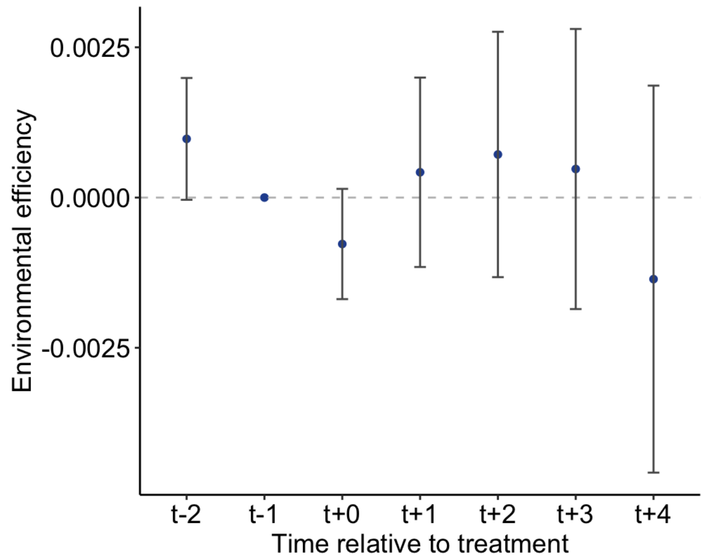
The above results clearly contradict the existing literature that points to a significantly positive impact of the LCCP (e.g. Cheng et al., 2019; Liu et al., 2020a; Fu et al., 2021; Yu et al., 2021; Shi and Xu, 2022; Zhang et al., 2022b; Wang et al., 2023; Yang et al., 2023). Despite the fact that we use SFA instead of DEA, there are a number of limitations that might explain the difference. First, none of the previous studies discuss the potential misattribution due to the introduction of the China's ETS pilots. Second, almost all studies do not control for the variation in treatment timing. Therefore, their estimates are likely to be inconsistent and unsatisfactory. One of the exceptions is Yu et al. (2021) who adopt the staggered DiD framework developed by Callaway and Sant'Anna (2021). However, they do not control for the selection bias, and their verification of the parallel trend assumption is incorrectly specified (Sun and Abraham, 2021; Roth et al., 2023).[7] Another exception is Fu et al. (2021) who only focus on the second wave. Although their estimated average treatment effect is significantly positive at 5% level, the estimator is only statistically significant at $t = 4$, showing very limited effect in improving efficiency.

Using the Wave I cities as the control group, we next evaluate the learning effect. Figure 3.7 presents the results from dynamic DiD. Albeit insignificant, the post-treatment estimates suggest a short-run learning effect, where the control units' environmental efficiency increased after the introduction of the LCCP second wave. Same as before, however, the pre-trend clearly runs against the parallel-trend assumption, which motivates us to the application of the SCM.

The results of the SCM is displayed in Figure 3.8. After controlling for the selection into treatment, we find that the learning effect is significantly positive at 5% at $t+1$. However, similar to the results from the DiD, this learning effect is not persistent and soon decreases to insignificant at $t + 2$ and negative at $t+4$. Overall, the average treatment effect of the treated (ATT) seems to be insignificant.

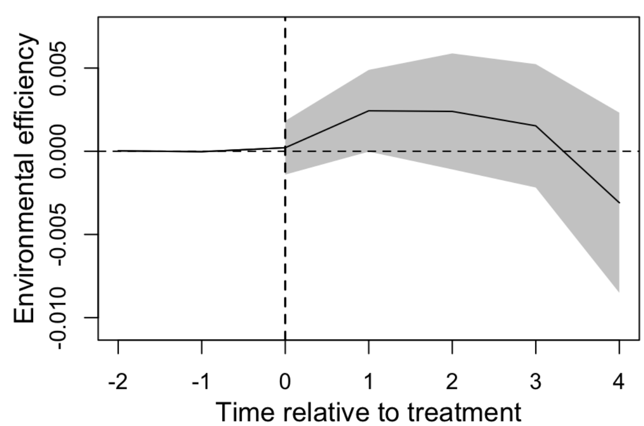
Table 3.3 summarises the results from the SCM estimations. Consistent with the plots before, the estimates suggest that the LCCP had no statistically significant impact on environmental efficiency.

[7] Specifically, Yu et al. (2021) use a dynamic DiD estimator to verify the parallel trend assumption. Given the staggered nature of the LCCP and the heterogeneous dynamic treatment effects across cohorts, however, their coefficients are difficult to interpret Sun and Abraham (2021). In this case, the estimator also fails to yield consistent estimates Roth et al. (2023).



Notes: The figure shows the learning effect on environmental efficiency using a dynamic difference-in-differences approach.

FIGURE 3.7: Investigating the learning effect using a difference-in-differences approach



Notes: The figure shows the learning effect on environmental efficiency using the *partially pooled* synthetic control method.

FIGURE 3.8: Investigating the learning effect using the synthetic control method

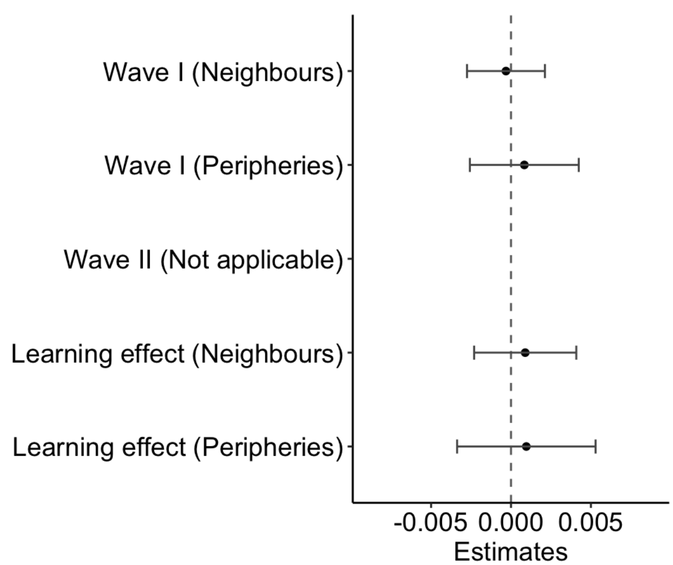
TABLE 3.3: Synthetic control method estimations

	ATT estimate	Std. error	*p*-value
LCCP first wave	0.0007	0.0010	0.485
LCCP second wave	-0.0004	0.0021	0.842
Learning effect	0.0007	0.0016	0.670

Note: (i) The table displays the synthetic control method estimation on environmental efficiency for the three analyses. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

3.3.1 Robustness checks

We start by dividing the non-pilot cities into two groups. Specifically, we define the non-pilot cities that share a border with those treated as 'Neighbours'. For the others we define as 'Peripheries'. Our main concern is that the learning effect might be more significant for the non-pilot cities that are geographically closer to those treated, since the diffusion of low-carbon mitigation might depend on how far they are away from those treated. We apply such maneuver to the evaluations of the first wave and the learning effect.

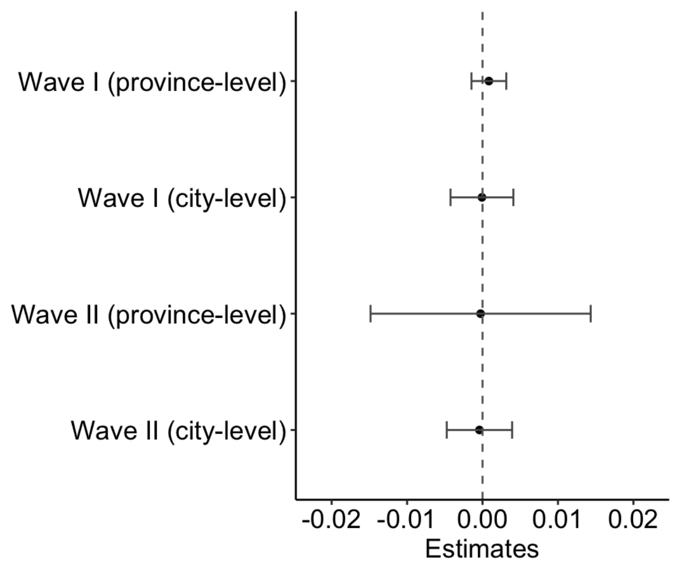


Notes: The figure shows the heterogeneity analysis by dividing the non-pilot cities into two groups, depending on whether they share a border with those treated ones.

FIGURE 3.9: Heterogeneity analysis – neighbours and peripheries

Figure 3.9 displays the results of the above exercise.[8] The estimates are insignificant for all analyses, showing that our main results are not sensitive to such concern.

Next, we differentiate the treated cities based on the treatment levels that they received. Specifically, we divide the treated cities into two groups, depending on whether they were assigned to treatment directly (which we refer to as city-level treatment) or assigned treatment status as part of a province-level treatment assignment. We then re-run our synthetic control estimation for each group, each wave.



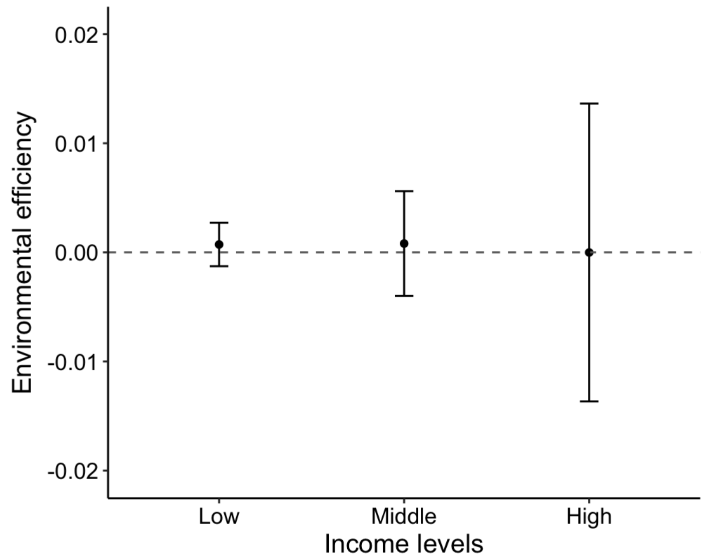
Notes: The figure shows the heterogeneity analysis by dividing the treated cities into two groups, depending on whether they were assigned to treatment directly or assigned treatment status as part of a province-level treatment assignment.

FIGURE 3.10: Heterogeneity analysis – treatment levels

Figure 3.10 displays the results.[9] Same as before, the estimates are insignificant for all analysis, suggesting that the effects of treatment do not differ from cities that were treated at different levels.

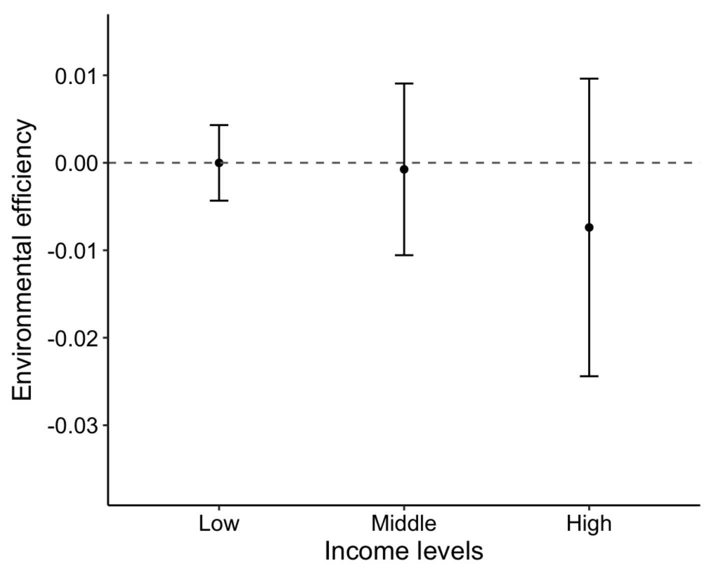
Lastly, for each of the analysis we group the cities in the treatment group based on their income levels, which we use the cities' mean GDP per capita between 2003 and 2016 to proxy. Specifically, we define three groups: cities with mean GDP per capita smaller than 35K CNY are defined as low-income cities; those between 35K and 65K CNY are defined as middle-income cities; those in excess

[8]See Figures E.1 and E.2 in Appendix E for the synthetic control fits.
[9]See Figures E.3 and E.4 in Appendix E for the synthetic control fits.



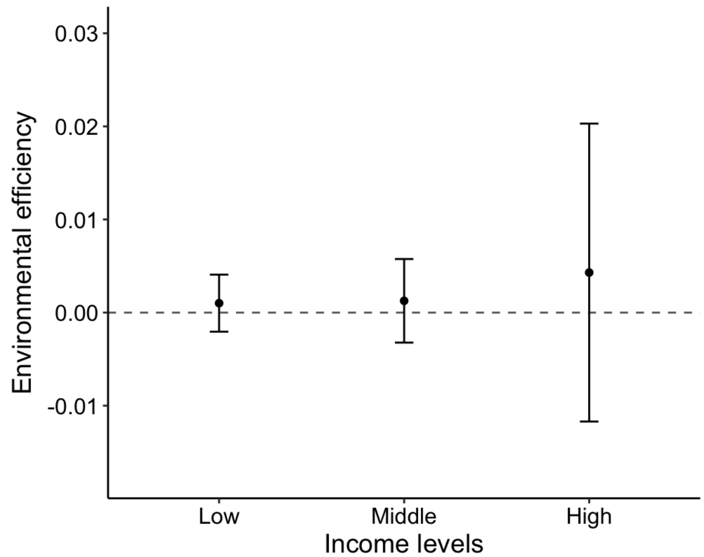
Notes: The figure shows the heterogeneity analysis for the cities treated by the first wave of the LCCP, by grouping the treated cities based on their mean GDP per capita between 2003 and 2016.

FIGURE 3.11: Heterogeneity analysis – effect of the LCCP first wave by income levels



Notes: The figure shows the heterogeneity analysis for the cities treated by the second wave of the LCCP, by grouping the treated cities based on their mean GDP per capita between 2003 and 2016.

FIGURE 3.12: Heterogeneity analysis – effect of the LCCP second wave by income levels



Notes: The figure shows the heterogeneity analysis of the learning effect, by grouping the cities in the treatment group based on their mean GDP per capita between 2003 and 2016.

FIGURE 3.13: Heterogeneity analysis – the learning effect by income levels

of 65K CNY are defined as high-income cities.

Figures 3.11-3.13 display the results.[10] The effects show clear trends across different income levels. The estimates in Figures 3.11 and 3.12 become more negative when the income levels increase. These trends seem to suggest that the LCCP is more positive for less affluence regions, possibly due to the fact that they are less efficient in optimising the use of production factors, which in turn higher up their marginal efficiency growth (Griffith et al., 2004). In Figure 3.13 the estimates are more positive for high-income cities and less when the income level decreases. This trend seems to suggest that the diffusion of low-carbon mitigation is more easily adopted for affluence regions. From statistical point of view, however, these interpretations are with very limited value, since all estimates are statistically insignificant.

3.4 Concluding remarks

In this chapter, we set out a credible framework to revisit the impact of the LCCP on environmental efficiency. Contrasting to the existing literature that uses data envelopment analysis to estimate

[10] See Figures E.5 - E.13 in Appendix E for the synthetic control fits.

efficiency, we estimate efficiency via a enhanced hyperbolic distance function using stochastic frontier analysis, which provides more economic interpretation of the results. The most significant challenge that we face is that, besides the selection into treatment and variation in treatment timing, the LCCP has clear learning effect that promote the diffusion of low-carbon to the non-pilot cities. We adopt a timing-based approach recommended by Miller (2023) with staggered synthetic control methodology proposed by Ben-Michael et al. (2022) to overcome these problems.

We find that the LCCP was not statistically significant in improving environmental efficiency, which unfortunately run against the existing literature that points to a significantly positive impact of the LCCP (e.g. Cheng et al., 2019; Liu et al., 2020a; Fu et al., 2021; Yu et al., 2021; Shi and Xu, 2022; Wen et al., 2022; Zhang et al., 2022b; Wang et al., 2023; Yang et al., 2023). Benefiting from our identification strategies and the framework that we set up in Chapter 2, we are confident that our results are more reliable. We also conducted a series of tests to ensure the robustness of our results.

For the first time, we find that the LCCP second wave promoted the non-pilot cities' environmental efficiency, although this effect is only significant in the short run and soon decreases to insignificance after $t + 3$. Such learning effect is not found in the first wave of the LCCP. We attribute the difference to the treatment levels. In Wave I, most cities were assigned treatment at province level. In contrast, in Wave II most cities were assigned treatment at city level. Province-level treatment has larger jurisdictional area, which may increase the transaction cost, therefore limiting the diffusion of low-carbon mitigation.

From a policy design perspective, our results make sense. The LCCP was introduced fundamentally as a voluntary scheme, where cities self-selected themselves into the treatment with vague ambitious. The policy itself also lacks explicit mandates and quantifiable targets. Therefore, the measures of the LCCP may not be stringent enough to kick-start the low-carbon transition for the treated cities, and ending up with null effect in this context does not surprise us much. Nevertheless, our results suggest that the relevant experience has been successfully adopted by the non-pilot cities in the short run. However, we are unable to explore the possible channels that developed the transition pathways due to the lack of relevant information and disaggregated data, but could be a fruitful avenue for future research.

Chapter 4

Unintended carbon leakage induced by pollution control policy in China

4.1 Introduction

Environmental regulation is an important tool to mitigate air pollutants concentration, and the adoption of air pollution control policies have been widely seen across economies Kuklinska et al. (2015); Balakrishnan et al. (2019); Yu et al. (2021). However, theses policies often have unintended consequences on many socioeconomic dimensions Feng et al. (2024). While relatively small compared to the estimated benefits, the 1990 Clean Air Act Amendments has been suggested to have significant and substantial re-allocative cost for workers in newly regulated plants Walker (2013). Similarly, evidence from the performance evaluation system that was constructed for cutting SO_2 emissions in China show that it slowed down economic growth rate (Chen et al., 2018). Clearly, understanding how air quality policies impact economies from a cost-benefit perspective is crucial for devising well-crafted regulations in the future.

In this chapter, we add to the relevant literature by investigating the socioeconomic impact of an air quality policy in one of the largest emerging economies in the world with a focus on an undesirable output in economic activities. Specifically, we ask whether the *Action Plan for Prevention and Control of Air Pollution* (Action Plan, hereafter) in mega-cities (Beijing and Tianjin) led to increases in CO_2 emissions in the surrounding regulated province (Hebei Province). This policy was launched by the State Council of China in 2013 to ease the growing air pollutants concentration, by setting explicit mandates on the densities of particulate matter with a diameter of 2.5 micrometers or less ($PM_{2.5}$) in different levels for different regions.

In what follows, we construct a unique dataset that documents detailed socioeconomic indicators and CO_2 emissions for 171 counties (37 districts and 134 counties) in Hebei Province over the period 2007-2017.[1] Our treatment group consists of counties in the neighbouring cities that share a border with Beijing and Tianjin; our control group includes counties in the peripheries, i.e. cities that do not share a border with Beijing and Tianjin. We employ a nearest neighbour matching technique that allows us to pair the treated units with their closely matched controls, which mitigates the bias introduced by the region-specific idiosyncratic characteristics. With this matched set, we can assess the leakage of carbon using a difference-in-differences (DiD) approach, by comparing the CO_2 emissions between the counties in the neighbouring cities and those in the peripheries.

To the best of our knowledge, we contribute to the literature along at least three dimensions. First, we add to the literature that focuses on the socioeconomic impact of environmental regulations. Our empirical findings are well aligned with, for instance, Walker (2013) and (Chen et al., 2018), highlighting the importance of comprehensive assessment in decision making process. Second, we contribute to the literature that focuses on the policy evaluation of the Action Plan, where many studies have suggested improvement on air quality (e.g. Liu et al., 2020b; Zhao et al., 2020; Wang et al., 2021; Yu et al., 2021). Some studies, such as those by Barrington-Leigh et al. (2019) and Mei et al. (2021), investigate the socioeconomic impacts of the Action Plan, and thus have a closer thematic alignment with our study. Specifically, Fang et al. (2019) find that the induced improvement in the regulated regions (Beijing, Tianjin, and Hebei Province) is at the cost of leakage of air pollutants in other neighbouring provinces. We complement their study by providing empirical evidence on leakage of carbon inside the regulation zone, with a focus between the mega-cities (Beijing and Tianjin) and the surrounding province (Hebei Province). Third, we complement Duvivier and Xiong (2013) who focuses on transboundary pollution in Hebei Province in China. They find that polluting firms are more likely to set up in the border counties than in the interior ones. This preliminary

[1]Here, 'counties' is the short-hand for county-level divisions that consist of 1301 counties, 977 districts, 117 autonomous counties, 49 banners, 3 autonomous banners, 1 special district and 1 forestry area. In China, there are three levels of administrative divisions: province-level divisions, prefecture-level divisions and county-level divisions, of which province-level divisions are the ones with the highest administrative status. County-level divisions subordinate to prefecture-level divisions. For brevity, we will use the term 'cities' as a shorthand for prefecture-level divisions henceforth.

study is not satisfactory, however, since they only focus on the overall probability. We complement their study by quantitatively evaluating the effect, examining the heterogeneity, and explore the economic benefits and leakage channels, with a focus on a 'co-pollutant' – CO_2 emissions.

We find a significant carbon leakage of 151 thousand tonnes of CO_2 emissions each year. This translates to an annual increase of CO_2 emissions of around 4.4% in the surrounding counties. We examine the potential challenges to our identification. We then discuss the robustness of our results, by using another set of covariates, different number of match size, and alternative matching estimator. We also perform heterogeneity analysis, by grouping the neighbouring counties based on their affluence levels and distances to Beijing and Tianjin. Our further results suggest additional economic benefits delivered by the Action Plan. These changes are likely attributed by secondary sector, where its share of GDP and annual gross product respectively increased by 5.2% and 906 million CNY, without crowding out other economic sectors.

The chapter develops as follows. Section 4.2 introduces the institutional background. Section 4.3 reviews the relevant literature. Section 4.4 elaborates the identification strategies and data. Section 4.5 and 4.6 presents empirical results and the relevant discussion. Section 4.7 concludes.

4.2 Institutional background

In January 2013, China encountered one of the worst air pollution events in history. The continuous haze weather covered almost one quarter of China's land area and over 600 million people were exposed to the polluted atmosphere. The record-high $PM_{2.5}$ concentration reached an hourly maximum of 791 $\mu g/m^3$ in Beijing, along with other extensively elevated atmospheric pollutants and satellite-derived aerosol optical depths (Andersson et al., 2015). The severe air pollution has produced massive damage to people's physical health, which increased total mortality induced by stroke and cardio-respiratory diseases (e.g. Chen et al., 2012; Rohde and Muller, 2015; Ebenstein et al., 2017; Gu et al., 2019).

Given this background, the State Council issued the *Action Plan for Prevention and Control of Air Pollution* in September 2013 (MEE, 2013). As a command-and-control regulation, it mandated that the density of inhalable particulate matters in all municipalities and prefecture-level cities would be decreased by more than 10% by 2017 relative to the levels in 2012. In particular, the densities of $PM_{2.5}$ in Beijing-Tianjin-Hebei region, Yangtze River Delta and Pearl River Delta would be respectively decreased by 25%, 20% and 15%. The annual average concentration of fine particulate matter in Beijing was additionally required to reduce to around 60 μg/m^3.[2]

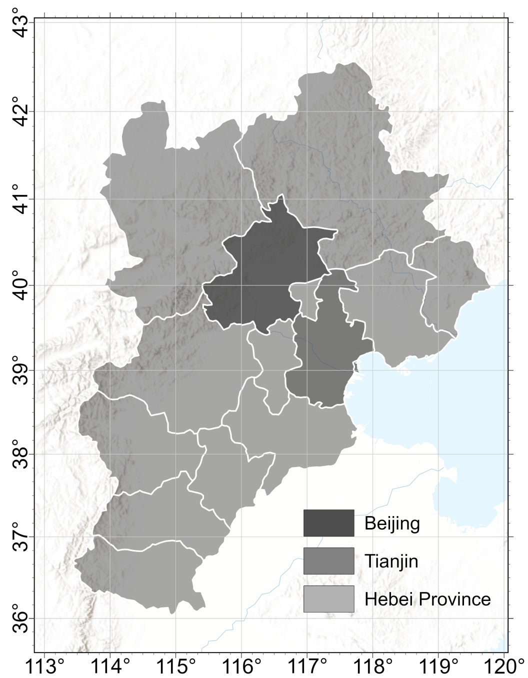


FIGURE 4.1: Geographical location of Beijing, Tianjin and Hebei Province

While the reduction targets set for Tianjin and Hebei are the same, the enforcements are likely sensitive to the differences in economic development and environmental quality. The public interest theory predicts that regulation protects and benefits the public at large (Stigler, 1971; Posner, 1974), by maximising social welfare to prevent market failures (Hantke-Domas, 2003). When air pollution control policy is supplied for inefficient market practices, the demand would be the willingness to pay for air pollution mitigation (Posner, 1974). The market equilibrium is realised when the supply and the demand intersects, which translates to different enforcement levels for different entities. Empirical evidence suggests that Tianjin has the highest ratio of willingness to pay to

[2]In 2013, the annual average concentration of fine particulate matter in Beijing was 89.5 μg/m^3. This additional requirement translates to a reduction of the density by around 33%.

income for air pollution mitigation in China (Sun et al., 2016). Therefore, it is theoretically conceivable to expect more stringent enforcement in Tianjin relative to Hebei, as this would maximise the social welfare of the environmental regulation market.

Figure 4.1 displays the geographical location of our study area, where Beijing and Tianjin are clearly surrounded by Hebei Province. Such geographical characteristic provides us an unique institutional context to investigate the relevant question. The differences in enforcement levels across regions bolster our confidence in expecting carbon leakage. In what follows, we assess the carbon leakage, with a focus on county-level CO_2 emissions.

4.3 Literature review

Our study is well aligned with the strand of literature that focuses on the policy evaluation of the Action Plan. A large body of existing literature has suggested improvement on air quality resulted by the Action Plan. Liu et al. (2020b), for instance, investigate the impact of the Action Plan on air quality using data from 16 districts in Beijing. Using a first difference approach, they find that the Action Plan significantly reduced concentrations of SO_2, PM_{10}, $PM_{2.5}$ and CO by around 10% each year. Yu et al. (2022) look into the similar question by assembling a city-level dataset from 2008 to 2018. Using a DiD design and a propensity score matching technique, they find that the Action Plan significantly reduced SO_2 emissions and $PM_{2.5}$ concentration by 18.4% and 24.7%, respectively. Also using a city-level panel dataset, Wu (2023) assess the impact of the Action Plan using a triple-difference estimator, and find similar results.

A number of papers have investigated the socioeconomic impact of the Action Plan, therefore they are more closely related to our work. Using survey data from 302 households in three districts in Beijing, Barrington-Leigh et al. (2019) suggest increased benefits on indoor temperature, indoor air pollution and life satisfaction. These benefits are however contingent on household wealth, where there were fewer benefits for households in low-income district. Mei et al. (2021) investigate the impact on real estate industry, using housing transaction data from 2011 to 2015 and administrative data on all power plants in Beijing. They estimate a triple-difference estimator, and find that the Action Plan led to a marginally significant price premium of 11% for properties close to coal-fired power plants. Particularly, Fang et al. (2019) use a multi-regional input-output model and an atmospheric chemical transport model to evaluate the impact on primary $PM_{2.5}$ and secondary precursor emissions. They find that the Action Plan reduced primary $PM_{2.5}$ and secondary precursor emissions in the regulation zone, but at the cost of leakage of air pollutants in other provinces, especially for the ones that are neighbouring to the regulated regions.

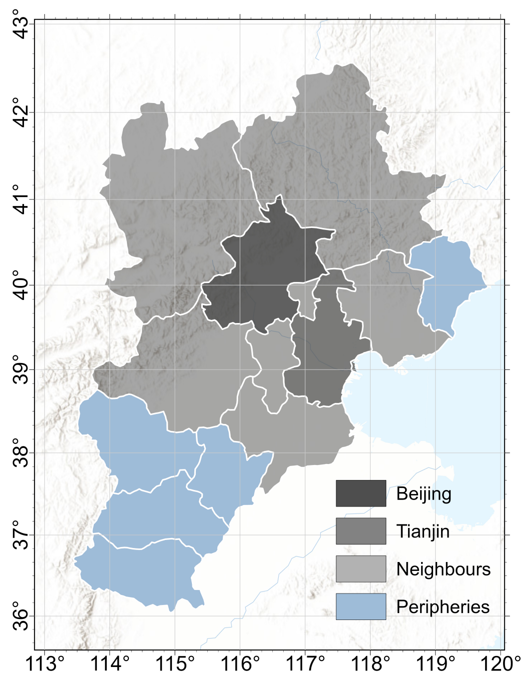
Our study is also closely related to the strand of literature on carbon leakage, where the relevant literature yet mainly focuses on international protocols and carbon markets. Little attention has been paid to alternative policy instruments. Studies focusing on Clean Development Mechanism (CDM) (Rosendahl and Strand, 2011) and Kyoto Protocol (Aichele and Felbermayr, 2015) suggest significantly positive carbon leakage. Findings on carbon markets are however contradictory. Despite the broader empirical support for the EU ETS from Naegele and Zaklan (2019) and Dechezleprêtre et al. (2022), Koch and Mama (2019) find that the regulated firms on average have increased their number of affiliates outside the EU, which seems to suggest future carbon leakage. While the findings from the Regional Greenhouse Gas Initiative (RGGI) in the US suggest carbon leakage in RGGI-surrounding regions (Fell and Maniloff, 2018), no displacement is found for the Japanese regional ETSs (Sadayuki and Arimura, 2021). As for the China's ETS pilots, leakage is found within firm ownership networks (He and Chen, 2023; Cui et al., 2023), but not found across administrative boundaries (Zhu et al., 2022).

4.4 Identification strategy and data

4.4.1 Identification strategy

We investigate the carbon leakage from Beijing and Tianjin induced by the Action Plan by examining the impact of the geographical proximity. Our identification strategy builds upon the investigation on transboundary pollution in China in Duvivier and Xiong (2013), who suggest that polluting firms are more likely to set up near the border. Carbon leakage likely follows the similar pattern, where the neighbouring counties have higher probability of becoming the focal points (Paroussos et al., 2015).

Figure 4.2 demonstrate our identification strategy. We define the cities that share a border with Beijing or Tianjin as neighbours, and then define their subordinated counties as neighbouring counties; cities that do not share a border with Beijing or Tianjin are defined as peripheries. We assess the leakage using a DiD approach, by comparing the county-level CO_2 emissions between the neighbouring counties and those in the peripheries.



Notes: The figure illustrates our identification strategy. Specifically, cities that share a border with Beijing or Tianjin are defined as neighbours; cities that do not share borders with Beijing or Tianjin are defined as peripheries. We assess the effect of carbon leakage by comparing the county-level CO_2 emissions across the neighbours and peripheries.

FIGURE 4.2: Identification framework

For county i in year t, we estimate the effect of carbon leakage via the following regression:

$$Y_{it} = \alpha + \sum_{\substack{-6 \le k \le 4 \\ k \ne -1}} \beta_k \times \text{Neighbour}_i \times \text{Treatment}_{t+k} + \delta X_{it} + \lambda_i + \theta_t + \epsilon_{it}. \tag{4.1}$$

Y_{it} is the outcome variable, i.e. county-level CO_2 emissions, of county i in year t. Neighbour_i and Treatment_t are the dummy variables respectively indexing neighbouring counties (one if the county is a neighbour and zero otherwise) and the Action Plan (one if the Action Plan has been implemented and zero otherwise). β_k is the coefficient of interest that estimates the dynamic effect of the leakage of carbon. X_{it} is the county-specific control variables, including economic and infrastructure indicators. λ_i and θ_t are the county-level and time fixed effects. ϵ_{it} is the random error. We specify leads and lags in our event study model to respectively examine the parallel trend assumption and observe how the leakage effect evolve over time. Note that the effects are estimated relative to the year before the implementation, $k = -1$. Given the detailed control variables and fixed effects, we identify the yearly effect by comparing the outcome across very similar counties.

The key identifying assumption, of course, is that the CO_2 emissions would develop in parallel trends between the neighbouring counties and those in the peripheries in the absence of treatment. Nevertheless, the neighbouring counties are geographically closer to Beijing and Tianjin, leading to a geographical advantage that may allow the neighbouring counties to have easier access to production inputs, e.g. labor, capital and energy. While the distances between the neighbouring counties and Beijing and Tianjin are time-invariant variables, they may have a time-varying impact on the outcome, which might have allowed the neighbouring counties to outpace the controls in the peripheries in the growth of CO_2 emissions, leading to a violation of the parallel trend assumption and add bias to our DiD estimate.

In what follows, we employ a one-to-one nearest neighbour matching technique.[3] For each neighbouring county, we match it with the county that situates in peripheries and has the shortest Mahalanobis distance. Due to the large number of neighbouring counties, we allow the matching with replacement, to ensure that each neighbouring county has the closest counterfactual. To further ensure the systematic balance of the covariates, we standardise the mean difference between the neighbouring counties and their counterfactuals using the standard deviations and sample mean across all neighbouring units. We then plot the standardised difference over the pre-treatment periods.

Even though the matching allows us to pair the neighbouring counties with their counterfactuals in the peripheries along a number of dimensions, it does not provide causal identification without a well-formulated research design. We need to ensure that the neighbours did not select themselves

[3]We conduct this using the matching methods proposed by Imai et al. (2021) for panel data. See the technical details in Appendix C.

as ‘neighbours’. In what follows, we reviewed the historical administrative adjustments of the units in Hebei Province. We find that the adjustments mostly happened in the last century. In our study period (2007-2017), two counties and three districts were abolished, and their administrative areas were merged into other nearby units without beyond city-level jurisdictions.[4] Some counties were renamed without any changes on their land areas. No administrative units were merged across cities. Overall, we are confident that our identification is not driven by this potential confoundedness (we examine the sensitivity of our results to this in Section 4.5.3).

4.4.2 Data

Our outcome variable is the county-level CO_2 emissions in Hebei Province. While in general preferable, estimates of emissions based on the IPCC guidelines are only available for a very limited set of counties (IPCC, 2006), owing to the insufficient and often incomparable county-level energy use information (Chen et al., 2020). In what follows, we use the county-level CO_2 emissions data estimated by Chen et al. (2020), who downscale the provincial energy-carbon emissions based on the nighttime light data from satellite imagery. The emission inventories include 2,735 counties and districts in around 350 administrative divisions from 1997 to 2017. We obtain the emission inventories from the Carbon Emission Accounts Datasets (CEADs, 2020).

To assess the question of interest, ideally, we need counterfactuals that closely mimic the treated units’ outcome and socioeconomic performance. In what follows, we collect both economic and infrastructure attributes to balance the systematic differences between the treatment and control groups, including GDP per capita (thousand CNY), share of GDP taken by secondary sector (%), share of GDP taken by social fixed asset investment (%), share of GDP taken by fiscal expenditure (%), highway per land area (km), telephones (fixed and mobile) per capita (unit) and beds in health care institutions per thousands (bed). These data were collected from the *Hebei Statistical Yearbook* (NBS, 2018). We have crosschecked the data with the relevant data from prefecture-level statistical

[4] Four cities were involved in such adjustments. (i) Shijiazhuang: Qiaodong District was abolished in 2014, and its administrative area was split and merged into Chang’an District and Qiaoxi District. (ii) Baoding: In 2015, Nanshi District and Beishi District were abolished and merged together as Lianchi District (new). (iii) Handan: Handan County was abolished in 2016, and its administrative area was split and merged into Hanshan District and Congtai District. (iv) Zhangjiakou: Xuanhua County was abolished in 2016. Its administrative area was merged into Xuanhua District.

yearbooks to ensure the best possible accuracy, and have normalised all monetary measurements to 2010 CNY.

By merging the above data with the county-level CO_2 emissions inventory, we assemble a unique county-level dataset that documents detailed statistics of the counties in Hebei Province spanning from 2007 to 2017. In total, there are 171 counties in our sample, of which 92 are the neighbouring counties. To control for the potential omitted variables bias and to best predict the neighbouring counties' CO_2 emissions, we use all economic and infrastructure indicators as control variables in a DiD approach and the nearest neighbour matching estimation.

TABLE 4.1: Descriptive statistics, 2007-2017

	Neighbours Mean	Peripheries Mean	Uncond. diff.	*t*-test
Outcome variable				
CO_2 emissions (million tonnes)	3.86 (2.73)	3.45 (2.39)	0.41*** (0.12)	3.43
Economic indicators				
GDP per capita (thousand CNY)	30.35 (20.54)	25.92 (13.79)	4.43*** (0.81)	5.46
Share of GDP taken by secondary sector (%)	46.94 (14.42)	47.68 (12.90)	-0.74 (0.65)	-1.14
Share of GDP taken by social fixed asset investment (%)	83.75 (41.01)	88.00 (36.18)	-4.26** (1.81)	-2.35
Share of GDP taken by fiscal expenditure (%)	14.91 (9.28)	12.73 (6.74)	2.18*** (0.38)	5.79
Infrastructure indicators				
Highway per land area (km)	1.08 (0.48)	1.43 (0.47)	-0.35*** (0.02)	-14.40
Telephones (fixed and mobile) per capita (unit)	0.15 (0.08)	0.12 (0.09)	0.02*** (0.00)	5.32
Beds in health care institutions per thousands (unit)	3.39 (2.32)	3.22 (2.08)	0.17 (0.11)	1.52

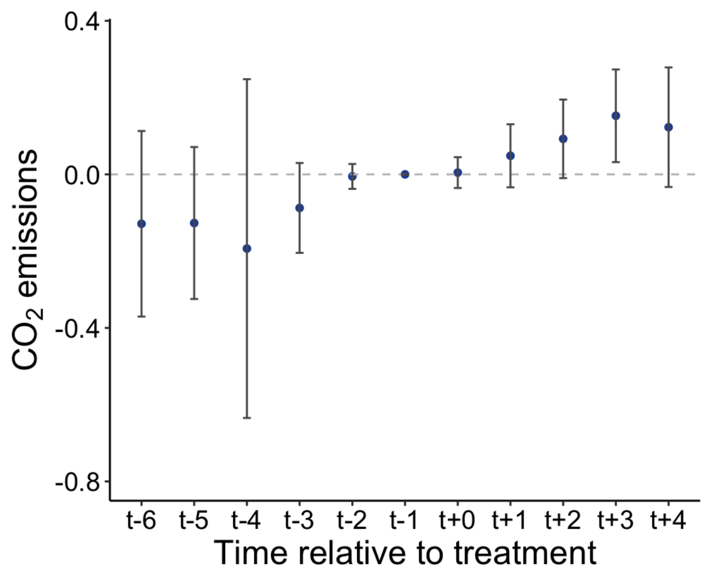
Notes: The table presents the means and standard deviations, along with the unconditional differences and results of the Student's t-test, for specific outcome and socioeconomic indicators in neighbouring counties compared to those in the peripheries.

Table 4.1 displays the statistics of the outcome variable and socioeconomic indicators. As suggested by the means and unconditional differences, neighbouring counties significantly differs from those in the peripheries along most variables except share of GDP taken by secondary (%) and beds in health care institutions per thousands (unit). The share of GDP taken by secondary sector is near

50% for both groups, indicating that their regional economies are dominated by secondary sector. Note that the share of GDP taken by social fixed asset investment is over 200% for some counties, because of the city-level inter-county transfer payment (see Table F.1 for the descriptive details).

4.5 Results

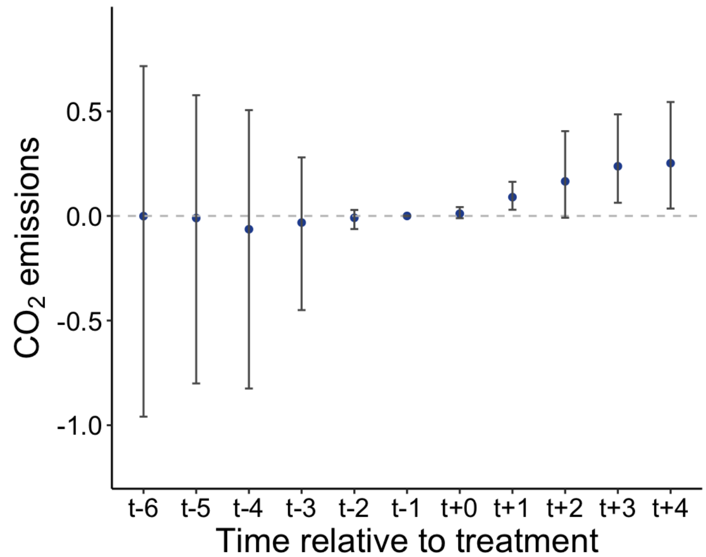
4.5.1 Main results



Notes: The figure reports the results of a difference-in-differences approach. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% confidence interval.

FIGURE 4.3: Carbon leakage – difference-in-differences approach

We start this section by presenting the baseline results estimated using the dynamic DiD approach. Results are displayed in Figure 4.3. The post-treatment estimates suggest an upward-sloping trend for CO_2 emissions. Specifically, the point estimate is statistically significant at $t + 3$, as suggested by the 95% confidence interval. However, the average treatment effect seems to be insignificant, as suggested by averaging the point estimates across the study period. More importantly, there is clear pre-trend before the implementation of the Action Plan, which violates the parallel trend assumption, although the estimates are insignificant.



Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.4: Carbon leakage – with Mahalanobis matching

Clearly, the naïve regression is not the ideal approach to identify causality in this context. To overcome the bias introduced by the region-specific idiosyncratic characteristics, we move on to the nearest neighbour matching technique to pair the best possible controls to the neighbouring counties. Figure 4.4 presents the results of applying the matching method. The outcome in the pre-treatment periods is virtually flat, suggesting satisfactory balance between the neighbouring counties and the counterfactuals. The treatment effects become significantly positive at $t+1$ and remain so until the end of the horizon. Relative to the counties in the peripheries, these results show that the Action Plan caused statistically significant carbon leakage in the neighbouring counties.

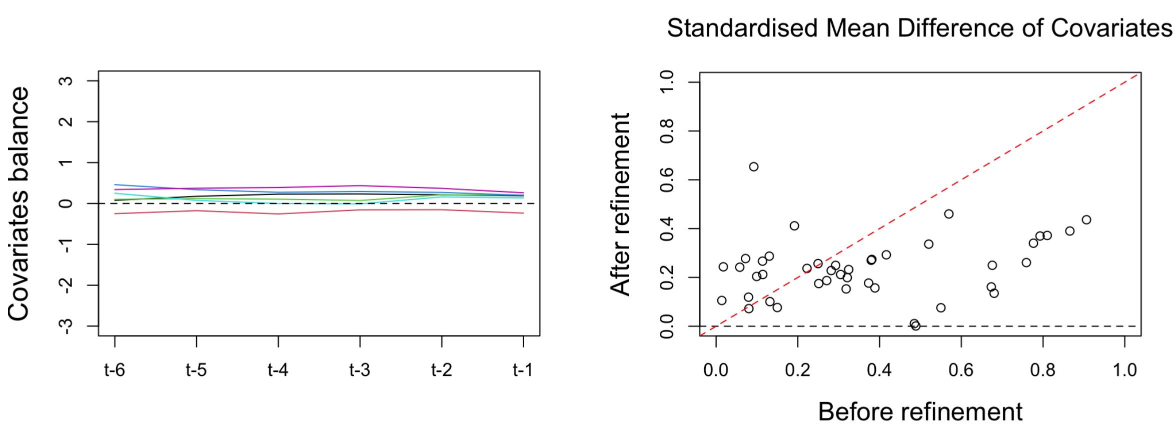
We report the average treatment effect in Table 4.2. As suggested by the estimates and bootstrapped confidence intervals, the treatment effect is 0.151, being significantly positive at 5%, implying an annual leakage of 151 thousand tonnes of CO_2 emissions, which translates to an increase of CO_2 emissions around 4.4% in the neighbouring counties.

We further examine the covariates balance between the neighbouring counties and their counterfactuals. Figure 4.5 displays the matching quality for the main results. The left panel shows the

TABLE 4.2: Estimates of carbon leakage – with Mahalanobis matching

	Average treatment effects
Estimate	0.151**
90% CI	(0.040, 0.283)
95% CI	(0.022, 0.323)
99% CI	(-0.005, 0.400)

Note: (i) The table displays the estimate of the nearest neighbour matching on county-level CO_2 emissions. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.



Notes: (i) The figure displays the matching quality for the main results. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE 4.5: Covariates balance for the main results.

covariates balance that is measured in the unit of standard deviation. We find that the covariates balance is generally within one standard deviation after the matching, which is satisfactory. The right panel shows how the matching method improves the covariates balance between the neighbouring counties and the counterfactuals, where control units are assigned equal weights before the matching, and are then matched one-to-one to a treated unit that has the shortest Mahalanobis distance. We compare the covariates balance before and after the refinement using the absolute value of standardised mean difference of covariates, where a dot below the 45 degree line implies improvement. Overall, we find that the matching improves balance for most control variables in the pre-treatment periods.

4.5.2 Challenges to identification

While we believe that the data and methods used are appropriate for this case study, certain issues could potentially compromise our identification. To address these concerns, we have conducted a more thorough analysis of the data to bolster our confidence in the empirical findings.

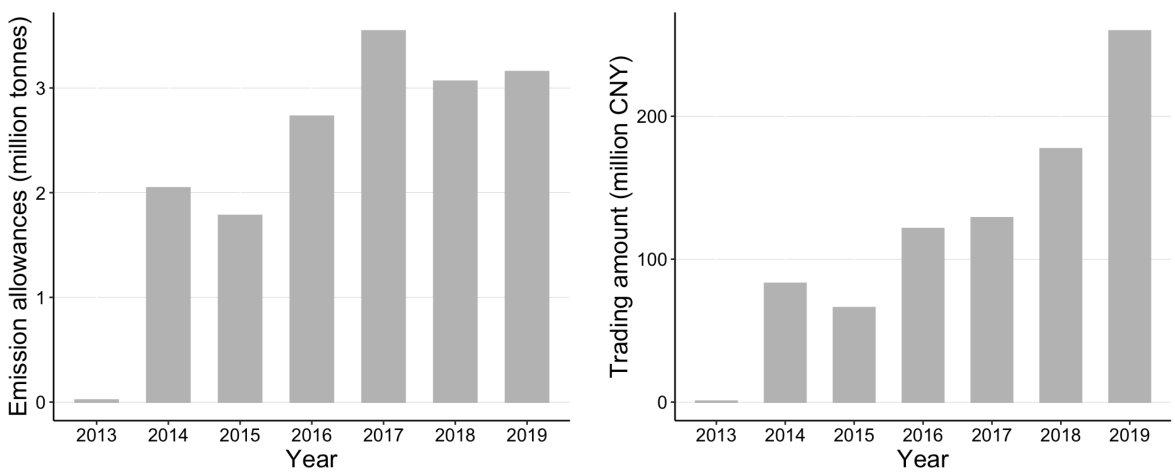
First and foremost, at the time of issuing the Action Plan, Chinese authorities introduced emission trading scheme (ETS) to mitigate climate change. Starting in 2011, with trading commencing in 2013, seven ETS pilots were launched, including one prefecture-level city (Shenzhen), two provinces (Hubei and Guangdong) and four municipalities (Beijing, Shanghai, Tianjin, Chongqing).[5] The ETS pilots in Beijing and Tianjin pose critical challenge to our identification since, without appropriately controlling for such policy overlap, it is impossible to causally attribute any leakage to the Action Plan.[6]

To control for this confoundedness, we exploit two unique features of the China's ETS pilots. First, mindful that the trading scheme was announced in 2011, and empirical results have suggested that ETS-regulated firms had started to significantly reduce CO_2 emissions (Cui et al., 2021). Such anticipatory reduction of CO_2 emissions due to the announcement of the ETS pilots, as suggested by Cui et al. (2023), is likely at the cost of leakage of carbon to the entities that are within a same firm ownership networks with those ETS-regulated ones but locate outside the pilots regions. Thus, we can investigate the carbon leakage induced by the announcement of the ETS pilots, by moving the start of the treatment back to 2011. If we find null effect, we can conclude that the announcement of the ETS pilots did not increase neighbouring counties' CO_2 emissions, therefore do not add confoundedness to our identification.

[5] The cap covered around 50% of the total CO_2 emissions in each treated division, including a range of entities and industries (Cui et al., 2021). The empirical literature has suggested that the ETS pilots reduced CO_2 emissions by around 15.5% (Hu et al., 2020).

[6] Zhu et al. (2022) suggest that China's ETS pilots did not lead to transboundary carbon leakage. However, due to the fact that Beijing and Tianjin are not included in their survey area, little is known in terms of the carbon leakage induced by the carbon markets in these two cities. Also, additional discretion needs to be taken, as the ETS pilots have been suggested to cause carbon leakage onto the entities that are within a same firm ownership networks with those ETS-regulated ones but locate outside the pilots regions (Cui et al., 2023; He and Chen, 2023).

Second, we follow the empirical findings reported by He and Chen (2023) who also suggest leakage of carbon due to the ETS pilots, but their estimate is only significant during the trading phase. Although our focus is not to determine which study is more credible, we need to ensure that our identification is not sensitive to any potential confoundedness. In fact, the ETS pilots in Beijing and Tianjin came into force at the end of 2013, bringing in very limited enforcement onto the ETS-regulated firms in the current year.[7] Thus, by moving the start of the treatment to 2014, we can investigate the mixed carbon leakage of the Action Plan and the ETS pilots, which we expect an increase on the coefficient relative to the one in 2013, if there is any leakage related to the ETS pilots. This is because by moving the the start of the treatment forwards, any leakage due to the ETS pilots in Beijing and Tianjin would be incorporated into the next year's estimate. If we find null effect, we can conclude that the implementation of the ETS pilots did not increase neighbouring counties' CO_2 emissions, therefore being conclusive combined with our exercise discussed before.



Notes: The figure displays the trading emission allowances (left, in million tonnes) and trading amount (right, in million CNY) documented in Beijing Carbon Emission Exchange and Tianjin Carbon Emission Exchange. Monetary values are normalised to 2013 CNY.

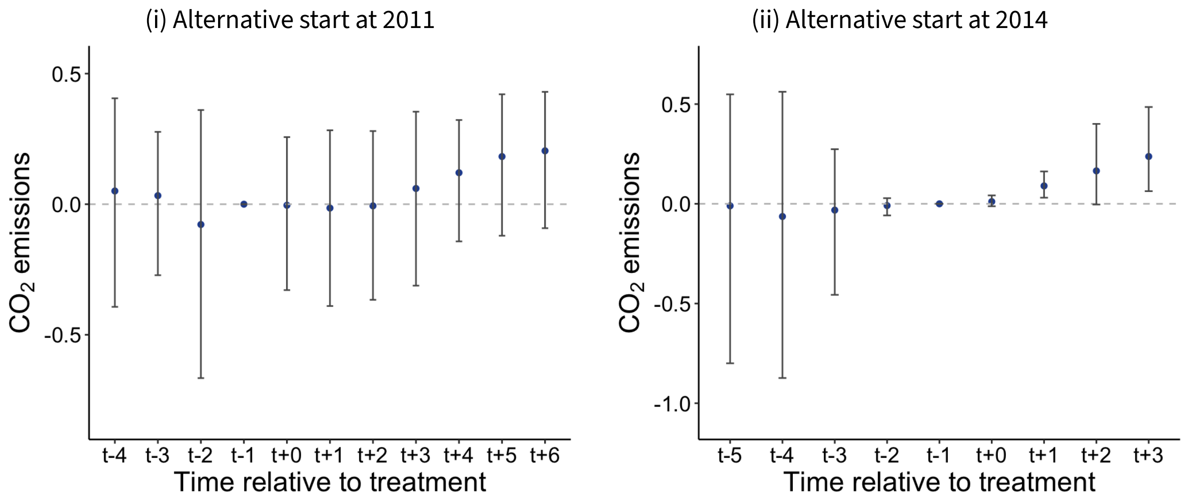
FIGURE 4.6: The ETS pilots in Beijing and Tianjin

Figure 4.6 provides the statistical evidence for our inference.[8] The left panel displays the traded emission allowances (million tonnes). The allowance was 20 thousand tonnes in 2013 as opposed to 2 million tonnes in 2014, showing that the carbon markets were very inactive in the previous

[7]The ETS pilots in Beijing and Tianjin came into force on 28th and 26th November 2013, respectively.

[8]We obtained the trading data documented in Beijing Carbon Emission Exchange and Tianjin Carbon Emission Exchange from the China Stock Market & Accounting Research (CSMAR) database. Monetary values are normalised to 2013 CNY. We also have data on trading date, trading type and prices, but decide not to display for brevity.

year. Similarly, as indicated by the statistics in the right panel , the trading amount in 2013 was 624 thousand CNY as opposed to 83 million CNY in 2014, again suggesting that the carbon markets were in a very preliminary stage.



Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions. Treatment effect is normalised relative to the beginning of the treatment. (ii) The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.7: Carbon leakage – examination of the ETS pilots in Beijing and Tianjin

We implement both exercises to ensure that our results are not driven by this potential confoundedness. Figure 4.7 displays the examination of the overlapping policies.[9] In the left panel, we move the start of the treatment to 2011 to investigate the carbon leakage induced by the announcement of the ETS pilots in Beijing and Tianjin. The pre-treatment balance is satisfactory, suggesting the assumption of parallel trends that is necessary for identification. The point estimates in the post-treatment periods are insignificant, suggesting null effect of carbon leakage due to the announcement. In the right panel, we move the treatment forwards to 2014 to investigate the mixed carbon leakage induced by the announcement and Action Plan. Again, the balance is satisfactory in the pre-treatment periods. The post-treatment estimates, albeit significantly positive, seem to have smaller magnitudes relative to our main results in Table 4.2.

We report the estimates of the examination in Table 4.3 to have a closer look of the the above exercise. Consistent with the insights in Figure 4.3, there is no statistically significant effect when

[9]Note that there are only nine time periods in the right panel of Figure 4.7. This is due to the fact that one may need to concern about the validity of parallel trends over longer time horizons. See the detailed discussion in Roth et al. (2023).

TABLE 4.3: Estimates of carbon leakage – examination of the ETS pilots in Beijing and Tianjin

	Alternative start at 2011	Alternative start at 2014
Estimate	0.077	0.126**
90% CI	(-0.213, 0.287)	(0.035, 0.239)
95% CI	(-0.277, 0.321)	(0.020, 0.263)
99% CI	(-0.347, 0.384)	(-0.013, 0.314)

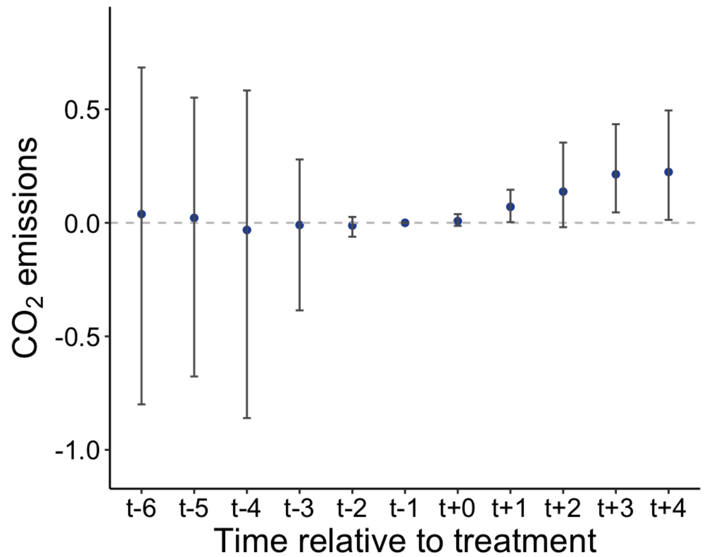
Note: (i) The table displays the estimate of the nearest neighbour matching on county-level CO_2 emissions. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

moving the treatment back to 2011, suggesting that the announcement of the ETS pilots in Beijing and Tianjin did not lead to increases in CO_2 emissions in the neighbouring counties. The estimate of moving the treatment to 2014 is 0.126, albeit still significantly positive at 5% level, smaller than our main results reported in Table 4.2 that is 0.151. Overall, we find no evidence of carbon leakage induced by the ETS pilots, showing that our main results are not confounded by the overlapping policies.

We further plot the covariates balances in Figures G.1 and G.2 in Appendix G. As suggested by the left panels in both figures, the standardised mean differences are within one standard deviation for all covariates. The right panels suggest that there are improvements for most control variables in the pre-treatment periods. Although the balance noticeably deteriorates for one control in one of the pre-treatment period, the standardised mean difference is still within one standard deviation, showing satisfactory balances for both exercises.

Next, we move on to discussing the potential confoundedness brought by the administrative adjustments discussed in Section 4.4.1. Including the involved counties and districts in our analysis, however, may lead to a noisy estimate, due to the inconsistent statistics. In what follows, we repeat our analysis, this time excluding the involved units, to ensure that our results are not driven by such potential confoundedness.

Figure 4.8 reports the results of the above exercise. The pre-treatment balance is satisfactory. Consistent with our main results before, the post-treatment estimates suggest significantly positive carbon leakage. Table 4.4 displays the estimate for the administrative adjustments. Relative to our



Notes: The figure reports the results of the administrative adjustments, using the nearest neighbour matching. The treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.8: Carbon leakage – examination of the administrative adjustments

main results, the treatment effect negligibly attenuates from 0.151 to 0.148, still being significantly positive at 5% level. Figure G.3 in Appendix G further displays the matching quality. Same as before, the standardised mean differences are satisfactory for all covariates, and there is noticeable improvement of the balance after the matching. Overall, we are confident that our results are not driven by the historical administrative adjustments.

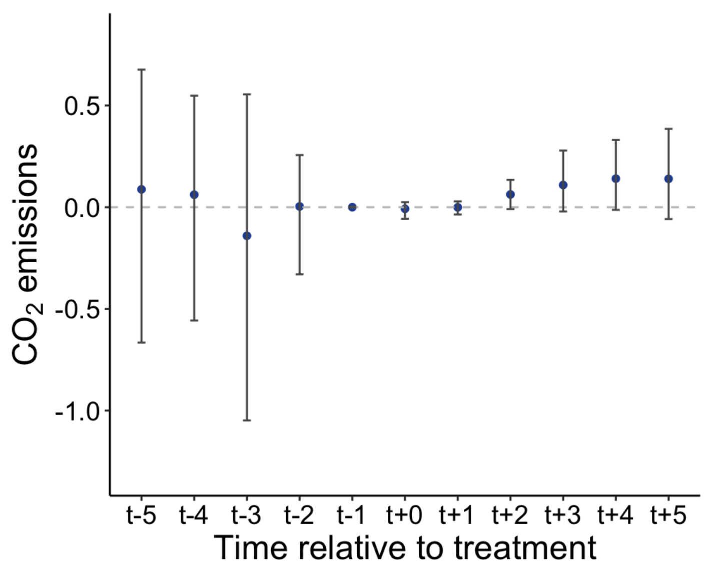
Even though the Action Plan was not introduced following an earlier announcement, the regional authorities might have been aware of potential air pollution control policy through their political connections, which may attenuate the estimated treatment effect. To provide a cleaner investigation and control for this potential concern, we repeat our analysis, this time moving the the treatment to one year prior to its official start.

Figure 4.9 reports the results, where the pre-treatment balance is satisfactory. The point estimates in the post-treatment periods seem to have significantly attenuated, relative to our main results in Figure 4.4. Table 4.5 reports the estimates. After the adjustment, the treatment effect decreases

TABLE 4.4: Estimates of carbon leakage – examination of the administrative adjustments

	Administrative adjustments
Estimate	0.148**
90% CI	(0.038, 0.289)
95% CI	(0.020, 0.331)
99% CI	(-0.006, 0.423)

Note: (i) The table displays the estimates for the administrative adjustments, using the nearest neighbour matching. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.



Notes: The figure reports the results of the announcement effect, using the nearest neighbour matching. The treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.9: Carbon leakage – examination of the announcement effect

from 0.151 to 0.074, and becomes insignificant. In principle, we would expect larger and more significant estimate in the presence of announcement effect, because by backdating the treatment start date, the announcement effect would be incorporated into the estimate. The matching quality is displayed in Figure G.4 in Appendix G, where the standardised mean differences are generally within one standard deviation for all control variables. The improvement of covariates balance is also substantial after the matching, showing satisfactory matching quality. Overall, we find no evidence of the announcement effect.

TABLE 4.5: Estimates of carbon leakage – examination of the announcement effect

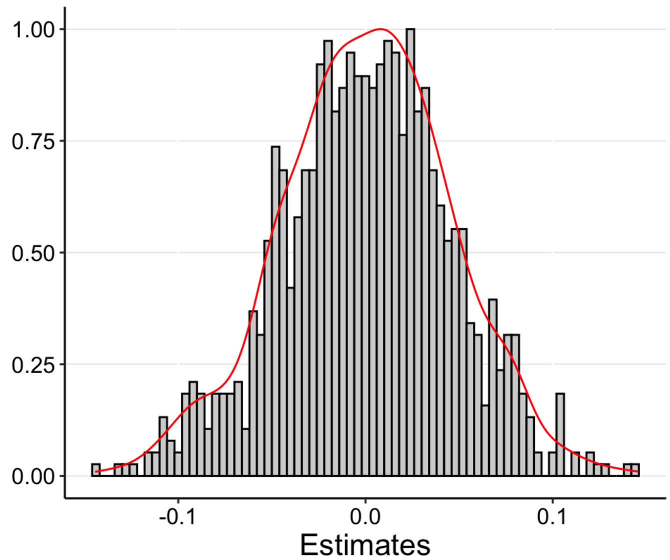
	Announcement effect
Estimate	0.074
90% CI	(-0.012, 0.169)
95% CI	(-0.025, 0.193)
99% CI	(-0.060, 0.238)

Note: (i) The table displays the estimates for the announcement effect, using the nearest neighbour matching. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

4.5.3 Robustness check

Having validated our empirical findings, we now move on to testing the robustness of our results. We start this section by conducting a falsification test, to ensure that our significant results do not come from nowhere. Specifically, we randomly draw 92 counties from the dataset without replacement to construct an 'alternative' set of neighbouring counties. We use the 'alternative' neighbouring counties as the treatment group, and then apply the nearest neighbour matching. The average placebo effect should not be statistically different from zero. We iterate this exercise 1000 times.

Figure 4.10 shows the placebo effects, where the average treatment effects are plotted on the x-axis. The effects centre around zero and approximate to normal distribution with sample mean and standard deviation respectively as 0.000 and 0.045. This result is consistent with our expectation, suggesting that our estimates are not driven by other confounders.



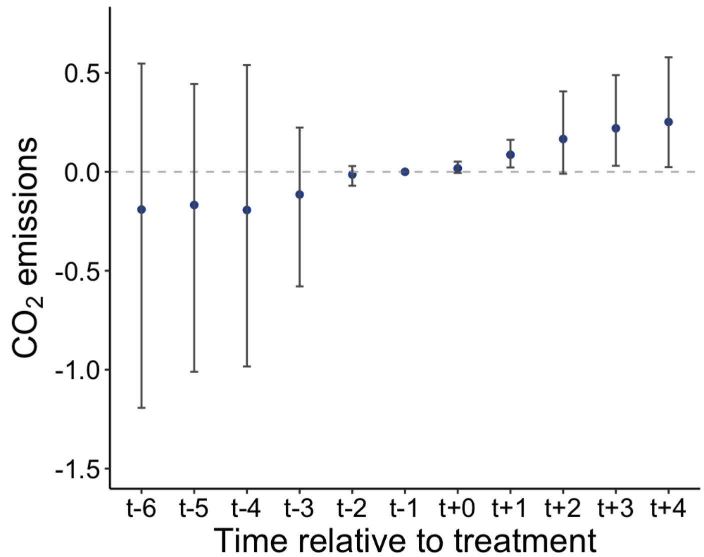
Notes: The figure reports the distribution of the placebo effects. Specifically, we randomly draw 92 counties from the dataset, and apply the nearest matching estimator to the 'alternative' treatment group using both economic and infrastructure indicators as control variables.

FIGURE 4.10: Placebo effects

Although we provided theoretical context and empirical evidence as our rationale to expect more stringent enforcement in Tianjin relative to Hebei in Section 4.2, the way of inference may be considered as 'hand-waving', as there has been no direct evidence that explicitly show that the enforcements are in different levels. For a cleaner analysis, we re-define neighbouring cities as the ones who only share a border with Beijing. We then exclude the neighbouring cities of Tianjin from our sample, to limit any potential leakage that may dampen our analysis.

Figure 4.11 reports this exercise, where the pre-treatment balance is satisfactory. The covariates balance is within one standard deviation for all covariates, as suggested by Figure G.5 in Appendix F. The post-treatment estimates are significantly positive and do not seem to attenuate relative to our main results in Figure 4.4.

For a closer look, we display the estimate in Table 4.6. The treatment effect negligibly declines from 0.151 to 0.148, still being statistically significant at 5%, suggesting that our identification is not sensitive to such concern.



Notes: The figure reports the results of re-defining the neighbouring cities, using the nearest neighbour matching. The treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.11: Carbon leakage – re-define the neighbouring cities

TABLE 4.6: Estimates of carbon leakage – re-define the neighbouring cities

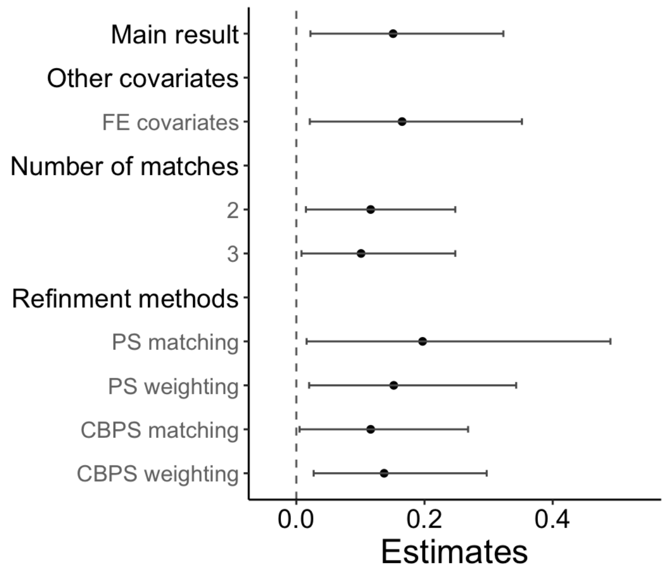
	Re-define the neighbouring cities
Estimate	0.148**
90% CI	(0.038, 0.299)
95% CI	(0.019, 0.327)
99% CI	(-0.011, 0.386)

Note: (i) The table displays the estimates for the administrative adjustments, using the nearest neighbour matching. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Next, we examine the sensitivity of our results to different refinement specifications, to mitigate the potential judgement calls involved in the identification strategy.

We start by changing the control variables in the nearest neighbour matching. Specifically, we instead use GDP per capita (thousand CNY), share of GDP taken by secondary sector (%), share of

GDP taken by fiscal expenditure (%), highway per land area (km), and telephones (fixed and mobile) per capita (unit) as the control variables. These variables are significantly correlated with the county-level CO_2 emissions, identified by the dynamic DiD approach.



Notes: (i) The figure reports the results of changing the refinement specifications. For convenient comparison, we display the main result at the top, and then proceed to the changes below. The error bars are constructed by the 95% quantiles of the bootstrapped estimates. (ii) We use abbreviations for brevity. In 'FE covariates', we use as control variables the significant variables identified by the dynamic DiD approach. In 'PS matching', 'PS weighting', 'CBPS matching', and 'CBPS weighting', we estimate the effect of carbon leakage instead using propensity score matching, propensity score weighting, covariate balancing propensity score matching, and covariate balancing propensity score weighting, respectively.

FIGURE 4.12: Different specifications

Second, we examine the sensitivity of our results to different number of match size. Most worries about the reduced power since one-to-one matching would inevitably discard a large number of observations. Additionally, the matching is allowed with replacement due to the large number of neighbouring units, which may raise a concern that a same county in the peripheries is repeatedly matched to different neighbouring ones. To ensure the robustness of our results, we conduct the test by increasing the number from one up to three. We expect slightly attenuated treatment effects, because the matching would converge to a regression where equal weights are assigned to all control units as we increase the match size.

Third, to ensure that our results are consistently robust across different refinement methods, we respectively estimate the treatment effect using propensity score matching, propensity score weighting, covariate balancing propensity score matching, and covariate balancing propensity score weighting (Imai and Ratkovic, 2014). We expect similar estimates and significance levels across different methods.

Figure 4.12 reports the results of changing the refinement specifications.[10] Overall, the treatment effects are significantly positive across all specifications, suggesting that our results are extremely robust.[11]

4.5.4 Heterogeneity analysis

Owing to the significant differences across neighbouring counties, our aggregate results may not be very informative in reflecting the actual carbon leakage. Therefore, as the last robustness check, we explore the heterogeneity of our results, by grouping the neighbouring counties based on their affluence levels and distances to Beijing and Tianjin.

We start by grouping the neighbouring counties based on their affluence levels, measured by the mean GDP per capita across our study period. We define as high-income counties whose mean GDP per capita are over 50k, middle-income counties whose GDP per capita are from 20k to 50k, and low-income counties whose GDP per capita are smaller than 20k.

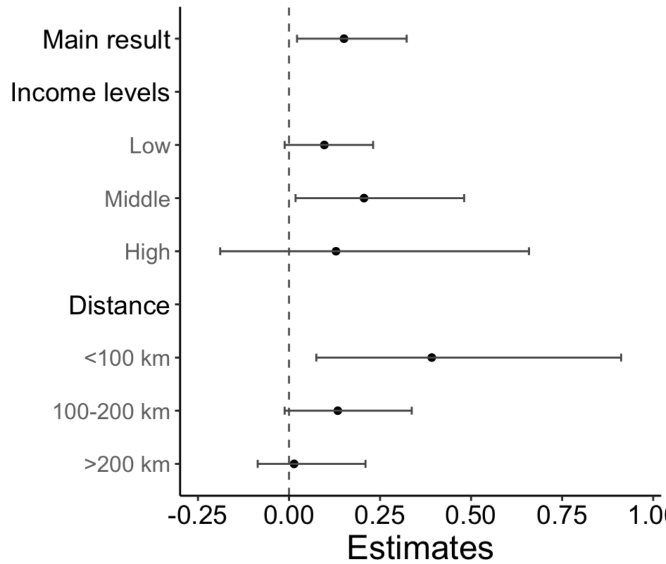
Second, we group the neighbouring counties based on their distances to Beijing and Tianjin, using the mean of the Euclidean distances measured by dropping two pins on Baidu Maps: one on the Beijing or Tianjin Municipal People's Government, and another one on the county-specific Municipal People's Government of interest.[12] Specifically, we group the neighbouring counties into four

[10] We use abbreviations for brevity. In 'FE covariates', we use as control variables the significant variables identified by the dynamic DiD approach. In 'PS matching', 'PS weighting', 'CBPS matching', and 'CBPS weighting', we estimate the effect of carbon leakage instead using propensity score matching, propensity score weighting, covariate balancing propensity score matching, and covariate balancing propensity score weighting, respectively.

[11] See Figures F.1 - F.4 in Appendix F for visualizations of the results. Figures G.6 - G.12 in Appendix G illustrate the quality of various refinements. Notably, the covariates balances show a deterioration in some exercises, especially with propensity score matching (Figure G.9) and covariate balancing propensity score matching (Figure G.11). Despite these observations, the approximately parallel pre-trends across all specifications suggest that our main results are robust, indicating that these alternative methodologies, while not optimal for this case study, still provide informative insights.

[12] We drop the pin on the former address of the Beijing Municipal People's Government, since the Beijing Municipal People's Government moved to the city subcenter in Tongzhou District in January, 2019. The former address is 2 Zhengyi Road, Dongcheng District, Beijing.

categories: (i) close counties whose distances are within 100 km; (ii) nearby counties whose distances range from 100 to 200 km; (iii) distant counties whose distances exceed 200 km. We expect the geographically closer counties to have larger and more significant carbon leakage, relative to the neighbouring counties that locate farther away (Paroussos et al., 2015).



Notes: The figure reports the results of heterogeneity analyses. For convenient comparison, we display the main result at the top. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE 4.13: Heterogeneity analysis

Figure 4.13 reports the results of the heterogeneity analysis discussed above.[13] As suggested by the estimates, we find that the carbon leakage is only relevant to the middle-income counties, with the effect is significantly positive at 5% level. The estimated coefficient is 0.206, implying that the Action Plan led to significant carbon leakage of 206 thousand tonnes of CO_2 emissions each year, equivalent to an annual increase by 5.6%. From the distance point of view, the leakage is only significantly positive for the close counties at 5% level with the coefficient of 0.392 . These estimates suggest significant carbon leakage of 392 thousand tonnes of CO_2 emissions per year, equivalent to annual increases by 8.3%.

[13]See Figures F.5 - F.10 in Appendix F for the visualisations of the results, where the parallel trend assumption holds for all checks. Figures G.13 - G.18 in Appendix G show the qualities of the refinements, where there are substantial improvements after the matching.

The city-level idiosyncratic characteristics may allow the neighbouring cities to substantially differ from each other, leading to carbon leakage at different rates. To uncover the city-specific carbon leakage and identify the main driver, we further decompose the above aggregate effects into disaggregate level. Results are reported in Figure F.11 in Appendix F. As suggested by the estimates, the leakage is mostly relevant to Baoding and Langfang, where the treatment effects are both statistically significant at 5% level with estimated treatment effects of 0.158 and 0.423. This results suggest that the Action Plan caused significant carbon leakage of 158 thousand tonnes and 423 thousand tonnes in these two cities, equivalent to annual increases of CO_2 emissions by 5.9% and 8.6%, respectively.[14] The estimated coefficient is also significantly positive for the high-income counties in Chengde, but is not significant at the average level.

4.6 Further results

Until now, we have been focusing on assessing the carbon leakage of the Action Plan, using the neighbouring counties' CO_2 emissions as the outcome variable. Based on the above tests, we are convinced that our identification is correct for the case study, and our empirical findings make sense in the question of interest.

In this section, we investigate economic benefits induced by the leakage and possible leakage channels that contribute to the leakage. We start by using alternative outcome variables to uncover these economic benefits. Specifically, we investigate the impact on physical capital by employing social fixed asset investment (in billion CNY) as the outcome measure, based on the assumption that neighbouring counties may be incentivised to invest in complementary infrastructure as a response to carbon leakage. Additionally, we include GDP (billion CNY) and GDP per capita (thousand CNY) as alternative outcome variables to examine the economic benefits associated with the carbon leakage.[15]

[14]Notably, the coefficient is significantly negative for the high-income counties in Langfang. However, there is only one neighbouring county and the plot suggests violation of parallel trend assumption.

[15]For this analysis, we use the same control variables as previously, with the exception of the model for GDP per capita. In this case, our control variables include the shares of GDP accounted for by the primary, secondary, and tertiary sectors (%), the share of GDP allocated to fiscal expenditure (%), highway density per land area (km), and the number of telephones (fixed and mobile) per capita (unit).

TABLE 4.7: Further results – economic benefits

	GDP	GDP per capita	Physical capital
Estimate	0.819**	2.225**	0.655
90% CI	(0.177, 1.495)	(0.503, 4.302)	(-1.147, 2.252)
95% CI	(0.049, 1.643)	(0.248, 4.733)	(-1.481, 2.489)
99% CI	(-0.266, 1.918)	(-0.306, 5.522)	(-2.183, 2.914)

Note: (i) The table displays the estimate of the nearest neighbour matching on alternative outcome variables. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Table 4.7 displays the results.[16] As suggested by the estimates, we find that the neighbouring counties' GDP and GDP per capita increased by around 819 million CNY and 2.2 thousand CNY relative to the counties in the peripheries, with significance level at 5% level for both estimations. These estimates translate to annual increases on the outcomes by around 7.8% and 8.8%, respectively. However, we find no statistically significant effect on physical capital.

TABLE 4.8: Further results – leakage channels

	Primary sector	Secondary sector	Tertiary sector
Panel A: share of GDP taken by sector (%)			
Estimate	-1.389	5.159*	-3.913
90% CI	(-4.010, 0.680)	(0.170, 11.675)	(-8.904, 0.021)
95% CI	(-4.563, 1.054)	(-0.424, 12.962)	(-10.086, 0.559)
99% CI	(-5.689, 1.729)	(-1.264, 17.308)	(-11.998, 1.560)
Panel B: gross product by sector (billion CNY)			
Estimate	0.100	0.906**	0.015
90% CI	(-0.131, 0.326)	(0.134, 1.900)	(-0.580, 0.533)
95% CI	(-0.176, 0.372)	(0.015, 2.151)	(-0.707, 0.612)
99% CI	(-0.274, 0.473)	(-0.181, 2.616)	(-0.895, 0.778)

Note: (i) The table displays the estimate of the nearest neighbour matching on alternative outcome variables. The confidence intervals are constructed by the 90%, 95% and 99% quantiles of the bootstrapped estimates. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

[16] See Figures F.12 - F.14 in Appendix F for the visualisations, and Figures G.19 - G.21 in Appendix G for the quality of the refinement.

Next, we focus on the possible leakage channels that contribute to the carbon leakage. While preferable, sector-specific energy consumption or CO_2 emissions are not available, because of the incomplete statistics at the county level. Therefore, we instead use shares of GDP respectively taken by the primary, secondary and tertiary sectors (%) as the outcome variables. We further use the gross product of the primary, secondary and tertiary sectors (billion CNY) as the alternative outcomes, to further identify whether the increase (decrease) crowds out (in) other sectors.[17]

Table 4.8 displays the results.[18] Albeit statistically significant at 10% level, we find that the Action Plan led to a structural change in the neighbouring counties, where the share of secondary sector rose by 5.2%. We report a significant increase on the gross product of secondary sector around 906 billion CNY. This significant increase did not crowd out the primary or tertiary sector, however, as indicated by the correspondingly insignificant estimates.

4.7 Concluding remark

In this chapter, we set out to assess the impact of an air pollution control policies in China – *Action Plan for Prevention and Control of Air Pollution* – on an undesirable output in economic activities, CO_2 emissions. Exploiting its policy design that the mitigation mandates were set at different levels in different regions, we are able to investigate the question of interest by comparing the CO_2 emissions from the counties in the neighbouring cities with those from the counties in the peripheries. The overlapping policy of the China's ETS pilots poses significant challenge to our identification, which we overcome by delving into the policy details of the pilots. We are also aware of the potential biases introduced by the historical administrative adjustments and the announcement effect. Furthermore, we uncover substantial heterogeneity along various dimensions, and explore the economic benefits and leakage channels.

We find that the Action Plan led to a significant carbon leakage of 151 thousand tonnes of CO_2 emissions each year. This translates to an annual increase of CO_2 emissions of 4.4% in the neighbouring

[17] We use GDP per capita (thousand CNY), shares of GDP respectively taken by primary, secondary and tertiary sectors (%), highway per land area (km), and telephones (fixed and mobile) per capita (unit) as control variables, wherever possible.

[18] See Figures F.15 - F.20 in Appendix F for the visualisations, and Figure G.22 - G.27 in Appendix G for the quality of the refinement.

counties. Our empirical findings are extremely robust and survived all relevant checks. The leakage of carbon is mostly relevant to Baoding and Langfang, where the Action Plan increased 158 thousand tonnes and 423 thousand tonnes of CO_2 emissions, equivalent to annual increases of CO_2 emissions by 5.9% and 8.6%, respectively. The leakage of carbon also brought additional economic benefits in the neighbouring counties, where their GDP and GDP per capita respectively rose by 819 million CNY and 2.1 thousand CNY, mostly contributed by secondary sector, where its share of GDP and gross product respectively increased by 5.2% and 906 million CNY, without crowding out other economic sectors.

Our empirical findings suggest an inflow of economic activities in the neighboring units induced by more stringent enforcement in the mega-cities. While this maneuver has brought economic benefits to the neighboring units, there may be uncovered costs associated with it that are beyond the scope of this analysis. Exploring these potential costs could be a fruitful avenue for future research. The specific channels of the inflow could involve relocating production across administrative boundaries or demand outsourcing. Unfortunately, we are unable to clarify which mechanism is more pronounced due to the lack of disaggregated data. Regardless, this does not alter the main idea of this chapter — an optimal environmental regulation should account for leakage channels. This emphasizes the need for a comprehensive assessment in a general equilibrium setting.

Chapter 5

Concluding remarks

In this thesis, we assess climate and air pollution control policies in emerging economies, with a focus on the *Low-Carbon City Pilot* (LCCP) and the *Action Plan for Prevention and Control of Air Pollution* (Action Plan) in China.

In Chapter 2, we investigate the effectiveness of the LCCP on CO_2 emissions per capita and CO_2 intensity of GDP. Contrary to the existing literature, we find that the LCCP had no statistically significant effect on either of the outcomes. Our results are robust to a series of tests. We also use an alternative set of data as outcomes, following an improved version of IPCC Guidelines reported by Shan et al. (2017). This set of data allows us to unprecedentedly explore the sectoral impact of the LCCP. We also delve into the policy design, and differentiate the units with explicit political agendas. Our results surprisingly suggest that the LCCP had no sector-level emissions, even for those cities with explicit political agendas.

In Chapter 3, we follow our developed identification framework and assess the impact of the LCCP on environmental efficiency. While we find that the LCCP had no statistically significant effect on efficiency, we find that the non-treated cities' efficiencies were closely associated with the implementation of the LCCP. Our results suggest that the second wave of the LCCP had statistically significant effect in increasing the non-treated cities' efficiencies, although only in the short run.

In Chapter 4, we move on to assessing the socioeconomic impact of the Action Plan, by investigating whether the policy led to increases of CO_2 emissions in the neighbouring counties. Our results suggest significant leakage of 151 thousand tonnes of CO_2 emissions each year, equivalent to an annual increase of CO_2 around 4.4%. We attributed the increases to the secondary sector, where its gross product and share of GDP respectively increased by 906 million CNY and 5.2%, without crowding out other economic sectors. The inflow of economic activities respectively rose neighbouring counties' GDP and GDP per capita by 819 million CNY and 2.1 thousand CNY.

The conclusions of Chapter 2 and 3 unfortunately run against to the existing literature. The differences mostly derive from the policy design of the LCCP, where cities are self-selected into the treatment, and are treated at different points of time. We believe that our empirical findings benefit from a more robust design of identification, which is tailored for the case study of the LCCP. Indeed, effects are in general insignificant once we control for the possible confounding factors.

One may worry that the learning effect that we reported in Chapter 3 may undermine our effort in Chapter 2. In fact, the LCCP mostly worked as a demonstration tool to promote transition to a low-carbon economy, rather than explicitly focusing on reducing CO_2 emissions. Moreover, we find no indication of treatment spill-overs for CO_2 emissions per capita and carbon intensity due to the LCCP. Therefore, we conclude that the LCCP has been focusing on optimising the use of production factors. Our use of environmental efficiency precisely reflects such transition, where we find that the second wave of the LCCP indeed promoted the diffusion of low-carbon mitigation. The effect is, unfortunately, not persistent and only statistically significant in the short run.

Our results lead us to conclude that emerging economies do not behave significantly different from the developed ones. The empirical findings are well aligned with the theory of environmental regulation. That is, quantifiable targets and the introduction of clear instruments affect the incentives of the economic agents. Credible enforcement still represent the key elements for effective environmental policy.

In Chapter 4, we find inflow of economic activities from the mega-cities to the neighbouring counties. Our concern is consistent with those reported in the literature, that such outsourcing of energy demand or relocation of heavy emitters may offset the benefits delivered by the intervention. Indeed, a well-crafted policy design should be devised in a general equilibrium setting that accounts for such potential leakage channels. While generally preferable, monitoring data of air pollutants are only available from 2014 onwards, thus, we are unable to investigate the leakage of air pollutants. The lack of appropriate data constraints our research at this scope, but could be a fruitful research revenue in the future.

This thesis could be extended in a number of dimensions. First, we could add more energy consumption data. The fact that the CO_2 emissions compiled following IPCC Guidelines have a shorter panel and less cities is because the relevant statistics are stored offline in the municipal bureau of statistics. Recruiting more researchers into our group might be helpful in adding more observations. Second, we could derive the data of air pollutants concentration by downscaling the global estimates of surface $PM_{2.5}$ reported by satellite images from National Aeronautics and Space Administration (NASA). Such transformation requires additional administrative data and programming, but is promising for studying the leakage of air pollutants.

Bibliography

Abadie, A. (2021). Using synthetic controls: Feasibility, data requirements, and methodological aspects. *Journal of Economic Literature*, 59(2):391–425.

Abadie, A., Diamond, A., and Hainmueller, J. (2010). Synthetic control methods for comparative case studies: Estimating the effect of California's tobacco control program. *Journal of the American statistical Association*, 105(490):493–505.

Abadie, A., Diamond, A., and Hainmueller, J. (2015). Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510.

Abadie, A. and Gardeazabal, J. (2003). The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1):113–132.

Aichele, R. and Felbermayr, G. (2015). Kyoto and carbon leakage: An empirical analysis of the carbon content of bilateral trade. *Review of Economics and Statistics*, 97(1):104–115.

Alberini, A. and Segerson, K. (2002). Assessing voluntary programs to improve environmental quality. *Environmental and Resource Economics*, 22(1):157–184.

Andersson, A., Deng, J., Du, K., Zheng, M., Yan, C., Sköld, M., and Gustafsson, Ö. (2015). Regionally-varying combustion sources of the January 2013 severe haze events over eastern China. *Environmental Science & Technology*, 49(4):2038–2043.

André, F. J. and Valenciano-Salazar, J. A. (2022). Voluntary carbon neutral programs. Adoption and firms' strategies. *Journal of Cleaner Production*, page 135191.

Baker, A. C., Larcker, D. F., and Wang, C. C. (2022). How much should we trust staggered difference-in-differences estimates? *Journal of Financial Economics*, 144(2):370–395.

Balakrishnan, K., Dey, S., Gupta, T., Dhaliwal, R., Brauer, M., Cohen, A. J., Stanaway, J. D., Beig, G., Joshi, T. K., Aggarwal, A. N., et al. (2019). The impact of air pollution on deaths, disease burden, and life expectancy across the states of India: the Global Burden of Disease Study 2017. *The Lancet Planetary Health*, 3(1):e26–e39.

Barrington-Leigh, C., Baumgartner, J., Carter, E., Robinson, B. E., Tao, S., and Zhang, Y. (2019). An evaluation of air quality, home heating and well-being under beijing's programme to eliminate household coal use. *Nature Energy*, 4(5):416–423.

Battese, G. E. and Coelli, T. J. (1988). Prediction of firm-level technical efficiencies with a generalized frontier production function and panel data. *Journal of Econometrics*, 38(3):387–399.

Beijing Municipal Bureau of Statistics (2021). Beijing statistical yearbook. `https://nj.tjj.beijing.gov.cn/nj/main/2021-tjnj/zk/indexch.htm` Last accessed: July 07, 2023.

Bell, M. L. and Davis, D. L. (2001). Reassessment of the lethal london fog of 1952: novel indicators of acute and chronic consequences of acute exposure to air pollution. *Environmental health perspectives*, 109(suppl 3):389–394.

Ben-Michael, E., Feller, A., and Rothstein, J. (2022). Synthetic controls with staggered adoption. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 84(2):351–381.

Borck, J. C. and Coglianese, C. (2009). Voluntary environmental programs: assessing their effectiveness. *Annual Review of Environment and Resources*, 34:305–324.

Cai, X., Lu, Y., Wu, M., and Yu, L. (2016). Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. *Journal of Development Economics*, 123:73–85.

Callaway, B. and Sant'Anna, P. H. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2):200–230.

CEADs (2020). County-level CO_2 Emissions in China during 1997–2017. `https://ceads.net/` Last accessed: March 15, 2023.

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. *European journal of operational research*, 2(6):429–444.

Chen, H., Guo, W., Feng, X., Wei, W., Liu, H., Feng, Y., and Gong, W. (2021). The impact of low-carbon city pilot policy on the total factor productivity of listed enterprises in China. *Resources, Conservation and Recycling*, 169:105457.

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.

Chen, R., Kan, H., Chen, B., Huang, W., Bai, Z., Song, G., and Pan, G. (2012). Association of particulate air pollution with daily mortality: the China Air Pollution and Health Effects Study. *American Journal of Epidemiology*, 175(11):1173–1181.

Chen, Y. J., Li, P., and Lu, Y. (2018). Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China. *Journal of Development Economics*, 133:84–101.

Cheng, J., Yi, J., Dai, S., and Xiong, Y. (2019). Can low-carbon city construction facilitate green growth? Evidence from China's pilot low-carbon city initiative. *Journal of cleaner production*, 231:1158–1170.

Cuesta, R. A., Lovell, C. K., and Zofío, J. L. (2009). Environmental efficiency measurement with translog distance functions: A parametric approach. *Ecological Economics*, 68(8-9):2232–2242.

Cuesta, R. A. and Zofío, J. L. (2005). Hyperbolic efficiency and parametric distance functions: with application to Spanish savings banks. *Journal of Productivity Analysis*, 24(1):31–48.

Cui, J., Wang, C., Wang, Z., Zhang, J., and Zheng, Y. (2023). Carbon leakage within firm ownership networks. *Available at SSRN 4514971*.

Cui, J., Wang, C., Zhang, J., and Zheng, Y. (2021). The effectiveness of China's regional carbon market pilots in reducing firm emissions. *Proceedings of the National Academy of Sciences*, 118(52):e2109912118.

De Chaisemartin, C. and d'Haultfoeuille, X. (2022). Difference-in-differences estimators of intertemporal treatment effects.

Dechezleprêtre, A., Gennaioli, C., Martin, R., Muûls, M., and Stoerk, T. (2022). Searching for carbon leaks in multinational companies. *Journal of Environmental Economics and Management*, 112:102601.

Donohue, J. J., Aneja, A., and Weber, K. D. (2019). Right-to-carry laws and violent crime: A comprehensive assessment using panel data and a state-level synthetic control analysis. *Journal of Empirical Legal Studies*, 16(2):198–247.

Dube, A. and Zipperer, B. (2015). Pooling multiple case studies using synthetic controls: An application to minimum wage policies. *Available at SSRN 2589786*.

Duvivier, C. and Xiong, H. (2013). Transboundary pollution in China: A study of polluting firms' location choices in Hebei province. *Environment and Development Economics*, 18(4):459–483.

Ebenstein, A., Fan, M., Greenstone, M., He, G., and Zhou, M. (2017). New evidence on the impact of sustained exposure to air pollution on life expectancy from China's Huai River Policy. *Proceedings of the National Academy of Sciences*, 114(39):10384–10389.

Estrada, F., Botzen, W. W., and Tol, R. S. (2015). Economic losses from US hurricanes consistent with an influence from climate change. *Nature Geoscience*, 8(11):880–884.

Fang, C., Liu, H., Luo, K., and Yu, X. (2017). Process and proposal for comprehensive regionalization of Chinese human geography. *Journal of Geographical Sciences*, 27(10):1155–1168.

Fang, D., Chen, B., Hubacek, K., Ni, R., Chen, L., Feng, K., and Lin, J. (2019). Clean air for some: Unintended spillover effects of regional air pollution policies. *Science advances*, 5(4):eaav4707.

Färe, R., Grosskopf, S., and Lovell, C. K. (1985). *The measurement of efficiency of production*, volume 6. Springer Science & Business Media.

Färe, R., Grosskopf, S., Lovell, C. K., and Pasurka, C. (1989). Multilateral productivity comparisons when some outputs are undesirable: a nonparametric approach. *The Review of Economics and Statistics*, pages 90–98.

Fell, H. and Maniloff, P. (2018). Leakage in regional environmental policy: The case of the regional greenhouse gas initiative. *Journal of Environmental Economics and Management*, 87:1–23.

Feng, T., Sun, Y., Shi, Y., Ma, J., Feng, C., and Chen, Z. (2024). Air pollution control policies and impacts: A review. *Renewable and Sustainable Energy Reviews*, 191:114071.

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.

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.

Galiani, S. and Quistorff, B. (2017). The synth_runner package: Utilities to automate synthetic control estimation using synth. *The Stata Journal*, 17(4):834–849.

Glomsrød, S. and Taoyuan, W. (2005). Coal cleaning: a viable strategy for reduced carbon emissions and improved environment in China? *Energy Policy*, 33(4):525–542.

Goodman-Bacon, A. (2021). Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, 225(2):254–277.

Griffith, R., Redding, S., and Reenen, J. V. (2004). Mapping the two faces of R&D: Productivity growth in a panel of OECD industries. *Review of economics and statistics*, 86(4):883–895.

Grimm, N. B., Chapin III, F. S., Bierwagen, B., Gonzalez, P., Groffman, P. M., Luo, Y., Melton, F., Nadelhoffer, K., Pairis, A., Raymond, P. A., et al. (2013). The impacts of climate change on ecosystem structure and function. *Frontiers in Ecology and the Environment*, 11(9):474–482.

Gu, H., Cao, Y., Elahi, E., and Jha, S. K. (2019). Human health damages related to air pollution in China. *Environmental Science and Pollution Research*, 26(13):13115–13125.

Hantke-Domas, M. (2003). The public interest theory of regulation: non-existence or misinterpretation? *European Journal of Law and Economics*, 15:165–194.

Harberger, A. C. (1988). Perspectives on capital and technology in less-developed countries. *Estudios de Economía (Chile)*.

He, L. and Chen, K. (2023). Does China's regional emission trading scheme lead to carbon leakage? Evidence from conglomerates. *Energy Policy*, 175:113481.

Hong, M., Chen, S., and Zhang, K. (2021). Impact of the “low-carbon city pilot” policy on energy intensity based on the empirical evidence of Chinese cities. *Frontiers in Environmental Science*, 9:717–737.

Hu, Y., Ren, S., Wang, Y., and Chen, X. (2020). Can carbon emission trading scheme achieve energy conservation and emission reduction? Evidence from the industrial sector in China. *Energy Economics*, 85:104590.

Huo, W., Qi, J., Yang, T., Liu, J., Liu, M., and Zhou, Z. (2022). Effects of China’s pilot low-carbon city policy on carbon emission reduction: A quasi-natural experiment based on satellite data. *Technological Forecasting and Social Change*, 175:121422.

Imai, K., Kim, I. S., and Wang, E. H. (2021). Matching Methods for Causal Inference with Time-Series Cross-Sectional Data. *American Journal of Political Science*.

Imai, K. and Ratkovic, M. (2014). Covariate balancing propensity score. *Journal of the Royal Statistical Society: Series B: Statistical Methodology*, pages 243–263.

IPCC (2006). IPCC Guidelines for National Greenhouse Gas Inventories. Technical report, IPCC.

Jidong, W., Ning, L., and Peijun, S. (2014). Benchmark wealth capital stock estimations across China’s 344 prefectures: 1978 to 2012. *China Economic Review*, 31:288–302.

Khanna, N., Fridley, D., and Hong, L. (2014). China’s pilot low-carbon city initiative: A comparative assessment of national goals and local plans. *Sustainable Cities and Society*, 12:110–121.

Koch, N. and Mama, H. B. (2019). Does the EU Emissions Trading System induce investment leakage? Evidence from German multinational firms. *Energy Economics*, 81:479–492.

Kreif, N., Grieve, R., Hangartner, D., Turner, A. J., Nikolova, S., and Sutton, M. (2016). Examination of the synthetic control method for evaluating health policies with multiple treated units. *Health economics*, 25(12):1514–1528.

Kuklinska, K., Wolska, L., and Namiesnik, J. (2015). Air quality policy in the US and the EU–a review. *Atmospheric Pollution Research*, 6(1):129–137.

Landry, P. F., Lü, X., and Duan, H. (2018). Does performance matter? Evaluating political selection along the Chinese administrative ladder. *Comparative Political Studies*, 51(8):1074–1105.

Liu, C., Zhou, Z., Liu, Q., Xie, R., and Zeng, X. (2020a). Can a low-carbon development path achieve win-win development: Evidence from China's low-carbon pilot policy. *Mitigation and Adaptation Strategies for Global Change*, 25(7):1199–1219.

Liu, Z., Chen, X., Cai, J., Baležentis, T., and Li, Y. (2020b). The impact of "coal to gas" policy on air quality: Evidence from Beijing, China. *Energies*, 13(15):3876.

Liu, Z., Guan, D., Wei, W., Davis, S. J., Ciais, P., Bai, J., Peng, S., Zhang, Q., Hubacek, K., Marland, G., et al. (2015). Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature*, 524(7565):335–338.

Lyon, T. P. and Maxwell, J. W. (2019). "Voluntary" approaches to environmental regulation. In *Economic institutions and environmental policy*, pages 75–120. Routledge.

Ma, J., Hu, Q., Shen, W., and Wei, X. (2021). Does the low-carbon city pilot policy promote green technology innovation? Based on green patent data of Chinese a-share listed companies. *International Journal of Environmental Research and Public Health*, 18(7):3695.

Mamardashvili, P., Emvalomatis, G., and Jan, P. (2016). Environmental performance and shadow value of polluting on Swiss dairy farms. *Journal of Agricultural and Resource Economics*, pages 225–246.

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S. L., Péan, C., Berger, S., Caud, N., Chen, Y., Goldfarb, L., Gomis, M., et al. (2021). Climate change 2021: the physical science basis. *Contribution of working group I to the sixth assessment report of the intergovernmental panel on climate change*, page 2.

MEE (2013). The State Council Issues Action Plan on Prevention and Control of Air Pollution Introducing Ten Measures to Improve Air Quality. `https://english.mee.gov.cn/News_service/infocus/201309/t20130924_260707.shtml` Last accessed: March 22, 2023.

Mei, Y., Gao, L., Zhang, W., and Yang, F.-A. (2021). Do homeowners benefit when coal-fired power plants switch to natural gas? Evidence from Beijing, China. *Journal of Environmental Economics and Management*, 110:102566.

Miller, D. L. (2023). An Introductory Guide to Event Study Models. *Journal of Economic Perspectives*, 37(2):203–230.

Naegele, H. and Zaklan, A. (2019). Does the EU ETS cause carbon leakage in European manufacturing? *Journal of Environmental Economics and Management*, 93:125–147.

NBS (2004). 2004.01: Financial news. http://www.stats.gov.cn/tjzs/tjsj/tjcb/zgjjjqyb/200406/t20040609_36901.html Last accessed: March 8, 2023.

NBS (2017). *China City Statistical Yearbook*. China statistics press.

NBS (2018). *Hebei Statistical Yearbook*. China Statistics Press.

NCSC (2013). Review of the Low-Carbon City Pilot. Technical report.

NCSC (2020). Notice on the Piloting Work of Low-Carbon Provinces and Cities. http://www.ncsc.org.cn/SY/dtsdysf/202003/t20200319_769716.shtml Last accessed: June 10, 2023.

NDRC (2007). China's National Climate Change Programme. Technical report.

NPC (2011). The Twelfth Five-Year Plan on National Economy and Social Development. Technical report.

NPC (2016). The Thirteenth Five-Year Plan on National Economy and Social Development. Technical report.

Olley, S. and Pakes, A. (1992). The dynamics of productivity in the telecommunications equipment industry.

Orea, L., Llorca, M., and Filippini, M. (2015). A new approach to measuring the rebound effect associated to energy efficiency improvements: An application to the US residential energy demand. *Energy economics*, 49:599–609.

Paroussos, L., Fragkos, P., Capros, P., and Fragkiadakis, K. (2015). Assessment of carbon leakage through the industry channel: the EU perspective. *Technological Forecasting and Social Change*, 90:204–219.

Posner, R. (1974). Theories of Economic Regulation. *Bell Journal of Economics*, 5(2):335–358.

Prasad, M. and Mishra, T. (2017). Low-carbon growth for Indian iron and steel sector: exploring the role of voluntary environmental compliance. *Energy Policy*, 100:41–50.

Rohde, R. A. and Muller, R. A. (2015). Air pollution in China: mapping of concentrations and sources. *PloS one*, 10(8):e0135749.

Rosendahl, K. E. and Strand, J. (2011). Carbon leakage from the Clean Development Mechanism. *The Energy Journal*, 32(4).

Roth, J., Sant'Anna, P. H., Bilinski, A., and Poe, J. (2023). What's trending in difference-in-differences? A synthesis of the recent econometrics literature. *Journal of Econometrics*.

Sadayuki, T. and Arimura, T. H. (2021). Do regional emission trading schemes lead to carbon leakage within firms? Evidence from Japan. *Energy Economics*, 104:105664.

SCPRC (2008). China's policies and actions for addressing climate change. Technical report.

SCPRC (2009). China to cut 40 to 45% GDP unit carbon by 2020. Technical report.

Segerson, K. (2013). Voluntary approaches to environmental protection and resource management. *Annual Review of Resource Economics*, 5(1):161–180.

Segerson, K. and Miceli, T. J. (1998). Voluntary environmental agreements: good or bad news for environmental protection? *Journal of environmental economics and management*, 36(2):109–130.

Shan, Y., Guan, D., Hubacek, K., Zheng, B., Davis, S. J., Jia, L., Liu, J., Liu, Z., Fromer, N., Mi, Z., et al. (2018a). City-level climate change mitigation in China. *Science advances*, 4(6):eaaq0390.

Shan, Y., Guan, D., Liu, J., Mi, Z., Liu, Z., Liu, J., Schroeder, H., Cai, B., Chen, Y., Shao, S., et al. (2017). Methodology and applications of city level CO_2 emission accounts in China. *Journal of Cleaner Production*, 161:1215–1225.

Shan, Y., Guan, D., Zheng, H., Ou, J., Li, Y., Meng, J., Mi, Z., Liu, Z., and Zhang, Q. (2018b). China CO_2 emission accounts 1997–2015. *Scientific data*, 5(1):1–14.

Shan, Y., Huang, Q., Guan, D., and Hubacek, K. (2020). China CO_2 emission accounts 2016–2017. *Scientific data*, 7(1):1–9.

Shanghai Municipal Bureau of Statistics (2021). Shanghai statistical yearbook. `https://tjj.sh.gov.cn/tjnj/20220309/0e01088a76754b448de6d608c42dad0f.html` Last accessed: July 07, 2023.

Shi, X. and Xu, Y. (2022). Evaluation of China's pilot low-carbon city program: a perspective of industrial carbon emission efficiency. *Atmospheric Pollution Research*, 13(6):101446.

Shukla, P., Skea, J., Slade, R., amd R. van Diemen, A. A., McCollum, D., Pathak, M., Some, S., Vyas, P., Fradera, R., Belkacemi, M., Hasija, A., Lisboa, G., Luz, S., and Malley, J., editors (2022). *Climate Change 2022: Mitigation of Climate Change. Contribution of Working Group III to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, UK and New York, NY, USA.

Simar, L. and Wilson, P. W. (2007). Estimation and inference in two-stage, semi-parametric models of production processes. *Journal of Econometrics*, 136(1):31–64.

Smil, V. (1993). *China's Environmental Crisis: An Enquiry into the Limits of National Development*. Routledge.

Stigler, G. J. (1971). The theory of economic regulation. *The Bell journal of economics and management science*, pages 3–21.

Sun, C., Yuan, X., and Yao, X. (2016). Social acceptance towards the air pollution in China: evidence from public's willingness to pay for smog mitigation. *Energy Policy*, 92:313–324.

Sun, L. and Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2):175–199.

Swartz, J. (2016). China's national emissions trading system. *ICTSD Series on Climate Change Architecture*, pages 20–23.

Tu, Z., Cao, Y., and Kong, J. (2022). The Impacts of Low-Carbon City Pilot Projects on Carbon Emissions in China. *Atmosphere*, 13(8):1269.

Walker, W. R. (2013). The transitional costs of sectoral reallocation: Evidence from the clean air act and the workforce. *The Quarterly journal of economics*, 128(4):1787–1835.

Wang, J., Li, Z., Ye, H., Mei, Y., Fu, J., and Li, Q. (2021). Do China's coal-to-gas policies improve regional environmental quality? A case of Beijing. *Environmental Science and Pollution Research*, 28:57667–57685.

Wang, L., Shao, J., and Ma, Y. (2023). Does China's low-carbon city pilot policy improve energy efficiency? *Energy*, 283:129048.

Wei, T., Yang, S., Moore, J. C., Shi, P., Cui, X., Duan, Q., Xu, B., Dai, Y., Yuan, W., Wei, X., et al. (2012). Developed and developing world responsibilities for historical climate change and CO_2 mitigation. *Proceedings of the National Academy of Sciences*, 109(32):12911–12915.

Wen, S., Jia, Z., and Chen, X. (2022). Can low-carbon city pilot policies significantly improve carbon emission efficiency? Empirical evidence from China. *Journal of Cleaner Production*, 346:131131.

WHO (2021). *WHO global air quality guidelines: particulate matter ($PM_{2.5}$ and PM_{10}), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. World Health Organization.

Wu, W. (2023). Is air pollution joint prevention and control effective in China—evidence from "Air Pollution Prevention and Control Action Plan". *Environmental Science and Pollution Research*, pages 1–15.

Xiong, W., Han, Y., Crabbe, M. J. C., and Yue, X.-G. (2020). Fiscal expenditures on science and technology and environmental pollution: Evidence from China. *International Journal of Environmental Research and Public Health*, 17(23):8761.

Yang, G. (2023). Can the low-carbon city pilot policy promote firms' low-carbon innovation: Evidence from China. *Plos one*, 18(1):e0277879.

Yang, X., Yang, X., Zhu, J., Jiang, P., Lin, H., Cai, Z., and Huang, H. (2023). Achieving co-benefits by implementing the low-carbon city pilot policy in China: Effectiveness and efficiency. *Environmental Technology & Innovation*, 30:103137.

Yao, Y. and Shen, X. (2021). Environmental protection and economic efficiency of low-carbon pilot cities in China. *Environment, Development and Sustainability*, 23(12):18143–18166.

Yu, C., Kang, J., Teng, J., Long, H., and Fu, Y. (2021). Does coal-to-gas policy reduce air pollution? Evidence from a quasi-natural experiment in China. *Science of The Total Environment*, 773:144645.

Yu, X., Shen, M., Wang, D., and Imwa, B. T. (2019). Does the low-carbon pilot initiative reduce carbon emissions? Evidence from the application of the synthetic control method in Guangdong Province. *Sustainability*, 11(14):3979.

Yu, Y., Dai, C., Wei, Y., Ren, H., and Zhou, J. (2022). Air pollution prevention and control action plan substantially reduced PM2.5 concentration in China. *Energy Economics*, 113:106206.

Zhang, H., Di Maria, C., Ghezelayagh, B., and Shan, Y. (2022a). Evaluating climate policies in emerging economies: evidence from the china's low-carbon city pilot.

Zhang, H., Feng, C., and Zhou, X. (2022b). Going carbon-neutral in China: Does the low-carbon city pilot policy improve carbon emission efficiency? *Sustainable Production and Consumption*, 33:312–329.

Zhang, J. (2008). Estimation of China's provincial capital stock (1952–2004) with applications. *Journal of Chinese Economic and Business Studies*, 6(2):177–196.

Zhang, Z., Xue, B., Pang, J., and Chen, X. (2016). The decoupling of resource consumption and environmental impact from economic growth in China: spatial pattern and temporal trend. *Sustainability*, 8(3):222.

Zhao, S., Hu, B., Gao, W., Li, L., Huang, W., Wang, L., Yang, Y., Liu, J., Li, J., Ji, D., et al. (2020). Effect of the "coal to gas" project on atmospheric NOX during the heating period at a suburban site between Beijing and Tianjin. *Atmospheric Research*, 241:104977.

Zhou, S. and Zhou, C. (2021). Evaluation of China's low-carbon city pilot policy: Evidence from 210 prefecture-level cities. *Plos one*, 16(10):e0258405.

Zhu, J., Ge, Z., Wang, J., Li, X., and Wang, C. (2022). Evaluating regional carbon emissions trading in China: effects, pathways, co-benefits, spillovers, and prospects. *Climate Policy*, 22(7):918–934.

Appendix A

Partially pooled synthetic control method

We suppose a panel dataset with an outcome Y_{it} for $1, ..., N$ units in $1, ..., T$ time periods. Each unit has a potential treatment outcome $Y_{it}(s)$ in each treatment time s, for $s = 1, ..., T$. $j = 1, ..., J$ indexes the units that are treated in time T_i with an order $T_1 \leq T_2 \leq ... \leq T_N$, and non-zero never treated units are indexed by $N_0 = N - J$ with $T_i = \infty$. k indexes event time relative to treatment time T_j by $k = t - T_j$, and ATT is estimated k periods after the treatment start. For each treated unit, we consider the outcome up to $L_j \leq T_j - 1$ periods before treatment, with $L \equiv max_{j \leq J} L_j$ referring to the maximum number of lagged outcomes. The ATT in original SCM is therefore expressed as:

$$ATT_k = \frac{1}{J} \sum_{j=1}^{J} \Big(Y_{j,T_j+k}(T_j) - Y_{j,T_j+k}(\infty) \Big). \tag{A.1}$$

When the ATT is estimated in the presence of multiple treated units through unit-specific fits, the average pre-treatment root mean square error q^{seq} across the J treated units is given by:

$$q^{seq} = \sqrt{\frac{1}{J} \sum_{j=1}^{J} \left[\frac{1}{L_j} \sum_{\ell=1}^{L_j} \left(Y_{j,T_j-\ell} - \sum_{i=1}^{N} \hat{\gamma}_{ij} Y_{i,T_j-\ell} \right)^2 \right]}. \tag{A.2}$$

Alternatively, ATT for multiple treated units can also be estimated by performing the average pre-treatment fit. The imbalance q^{pool} for the average of the treated units is expressed as:

$$q^{pool} = \sqrt{\frac{1}{L}\sum_{\ell=1}^{L}\left[\frac{1}{J}\sum_{T_j>\ell}\left(Y_{j,T_j-\ell} - \sum_{i=1}^{N}\hat{\gamma}_{ij}Y_{i,T_j-\ell}\right)\right]^2}. \tag{A.3}$$

The partially pooled SCM looks for weights that minimise a convex combination of imbalance in separate SCM q^{sep} and pooled imbalance q^{pool}:

$$\min_{\gamma_1,\ldots,\gamma_J} \nu(q^{pool})^2 + (1-\nu)(q^{sep})^2 + \lambda\sum_{j=1}^{J}\sum_{i=1}^{N} f(\gamma_{ij}). \tag{A.4}$$

Both of the pooled SCM and separate SCM are nested in this optimisation problem with a hyper-parameter $\nu \in [0,1]$ that equals 0 and 1 for each of them respectively (Ben-Michael et al., 2022). Specifically, $\lambda\sum_{j=1}^{J}\sum_{i=1}^{N} f(\gamma_{ij})$ is a term that penalises the weights toward uniformity over a hyper-parameter λ (Abadie et al., 2015). In the presence of perfect pre-treatment fit, the choice of penalty can be important since the optimisation problem may have multiple solutions.

Appendix B

Emission inventories using the IPCC Guidelines

Recent contributions in the literature have used the method developed by the IPCC to calculate CO_2 emissions, i.e. they multiply energy consumption by standard emissions factors (IPCC, 2006). However, recent survey data from 602 samples from 100 different mining areas that cover the majority of China's coal production suggests that the default emission factors proposed by the IPCC are on average 40% higher than than the actual values for China (Liu et al., 2015; Shan et al., 2018b). In addition, most studies do not take the CO_2 emissions from industrial processes into account. In the year 2016, the aggregate CO_2 emissions in China was 9,217.15 Mt, 7.6% of which are emissions due to chemical reactions linked to industrial processes rather than due to fossil fuels combustion (Shan et al., 2020). To correctly assess the amount of carbon emissions across cities, it is, therefore, necessary to both use the revised emission factors and to include process emissions.

In this paper, we, therefore, follow Shan et al. (2017) and calculate CO_2 emissions for each of our observations using the updated emission factors to compile the CO_2 emission inventories. The data for compiling the CO_2 emission inventory for each city is collected from the respective city-level statistical yearbook, which allows us to decompose the aggregate emissions into emissions from 17 different fossil fuels, 47 socioeconomic sectors and cement production.

Formally, the CO_2 emissions from fossil fuel combustion are calculated as:

$$CE_{\text{Energy},pt} = \sum_i \sum_j CE_{ptij} = \sum_i \sum_j AD_{ptij} \times NCV_{pti} \times CC_{pti} \times O_{ptij}, \tag{B.1}$$

where p denotes cities; t denotes the year; i indexes the 17 different fossil fuel types in the data and j indexes the 47 different economic sectors. AD_{ptij} represents the activity data, i.e. the physical quantity of fuel i consumed by sector j; NCV_i represents the net caloric value, i.e. is the heat value for each physical unit of the fossil fuel; CC_i represents CO_2 emissions per unit of the net caloric value of the fossil fuel; O_{ij} represents the oxygenation rate, which is the oxidation rate in the process of fossil fuel combustion.

Similarly, the CO_2 emissions from industrial processes can be expressed as:

$$CE_{\text{Process},pt} = \sum_m CE_{ptm} = \sum_m AD_{ptm} \times EF_m, \tag{B.2}$$

where m indexes the 7 different industrial processes for which we have information. AD_{ptm} denotes the production (in physical quantity) from industrial process m and EF_m denotes the corresponding emission factors. Table B.1 summarises the net caloric values and the emission factors for calculating CO_2 emissions from both fossil fuel combustion and industrial processes. For the combustion emissions, we used the oxygenation rates provided by Shan et al. (2018b).

TABLE B.1: Emission factors for CO_2 emissions calculations.

No.	Fossil fuel types	NCV_i	CC_i	Industrial process	EF_t
1	Raw coal	0.21	96.51	Cement production	0.4985
2	Cleaned coal	0.26	96.51		
3	Other washed coal	0.15	96.51		
4	Briquette	0.18	96.51		
5	Coke	0.28	115.07		
6	Coke oven gas	1.61	78.80		
7	Other gas	0.83	78.80		
8	Other coking products	0.28	100.64		
9	Natural gas	3.89	56.17		
10	Crude oil	0.43	73.63		
11	Gasoline	0.44	69.30		
12	Kerosene	0.44	71.87		
13	Diesel oil	0.43	74.07		
14	Fuel oil	0.43	77.37		
15	Other petroleum products	0.51	74.07		
16	Liquefied petroleum gas (LPG)	0.47	63.07		
17	Refinery gas	0.43	73.33		

Note: "Briquettes" includes briquettes and gangue. "Other gas" includes blast furnace gas, converter gas and other unclassified gas. "Other petroleum products" includes naphtha, lubricants, paraffin, white spirit, bitumen asphalt, petroleum coke and other unclassified petroleum products.

Appendix C

Mahalanobis distance

We calculate the average Mahalanobis distance between the treated units and each control following the expression below:

$$MD_{jt}(i') = \frac{1}{L}\sum_{l=1}^{L}\sqrt{(X_{j,t-l} - X_{i',t-l})^{\intercal}\Sigma_{j,t-1}^{-1}(X_{j,t-l} - X_{i',t-l})}, \tag{C.1}$$

where i' is the matched control unit for each treated unit $j = 1, ..., J$. $\sum_{j,t-1}^{-1}$ is the sample covariance matrix of X_{jt} that is the vector of control variables that one wishes to control for. With the non-negative integer $l = 1, ..., L$ as lags, we compute the distance using the included control variables, then average it across the study period.

After creating a matched set M_{it}, we compute the DiD estimate for each treated unit and then average it across all treated observations. For brevity, we omit the non-negative weight assigned to each treated observation, since it equals to one whatsoever due to our specification of one-to-one nearest neighbour matching. Following Imai et al. (2021), we specify the DiD estimator as

$$\hat{\delta}(F, L) = \frac{1}{\sum_{j=1}^{J}\sum_{t=L+1}^{T-F} D_{jt}}\sum_{j=1}^{J}\sum_{t=L+1}^{T-F} D_{jt}\Big\{(Y_{j,t+F} - Y_{j,t-1}) - (Y_{i',t+F} - Y_{i',t-1})\Big\}, \tag{C.2}$$

where $D_{jt} = 1$ only if observation (j, t) switches to treated unit from the control condition at time $t - 1$ to time t and has at least one matched control unit.

For each treated observation, the covariate balance for variable x at the pre-treatment period $t-l$ is defined as

$$B_{jt}(x,l) = \frac{X_{j,t-l,x} - X_{i',t-l,x}}{\sqrt{\frac{N_1}{N_1-1}(X_{i',t-l,x} - \overline{X}_{t-l,x})^2}}, \tag{C.3}$$

where N_1 is the total number of treated observations, and $\overline{X}_{t-l,x} = \sum_{i=1}^{N} D_{i,t-l,x}/N$. We then aggregate this balance across all treated observations for each control variable and pre-treatment period:

$$\overline{B}(x,l) = \frac{1}{N_1}\sum_{i=1}^{N}\sum_{t=L+1}^{T-F} D_{it}B_{jt}(x,l). \tag{C.4}$$

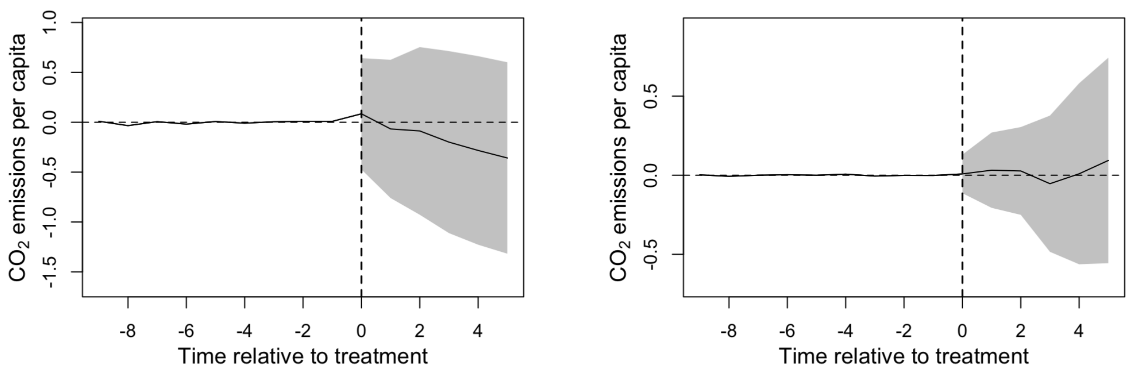
Appendix D

Additional figures and tables for Chapter 2

TABLE D.1: Controlling for policy overlap – Two Control Zones

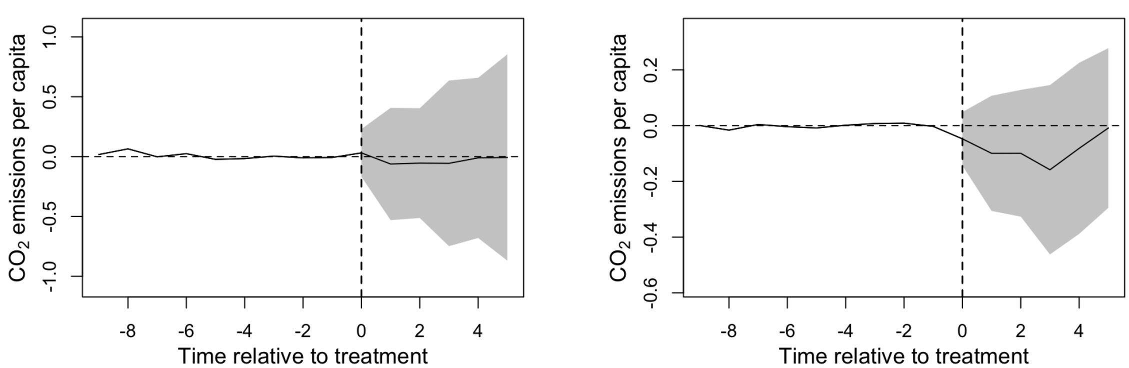
	ATT estimate	Std. err.	*p*-value
Panel A: Excluding Two Control Zones cities			
CO_2 emissions per capita	-0.152	0.384	0.692
GDP CO_2 intensity	0.019	0.173	0.913
Panel B: Restrict to Two Control Zones cities			
CO_2 emissions per capita	-0.027	0.267	0.919
GDP CO_2 intensity	-0.083	0.117	0.478

Notes: (i) The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.



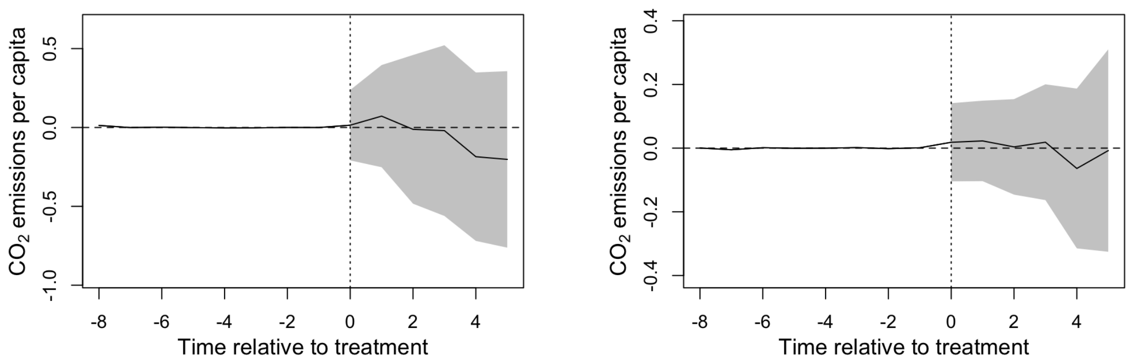
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on a sample that excludes all Two Control Zones cities. The effects are normalised relative to the beginning of treatment, i.e. 2009 for wave I and 2011 for Wave II.

FIGURE D.1: Controlling for policy overlap – excluding Two Control Zones cities



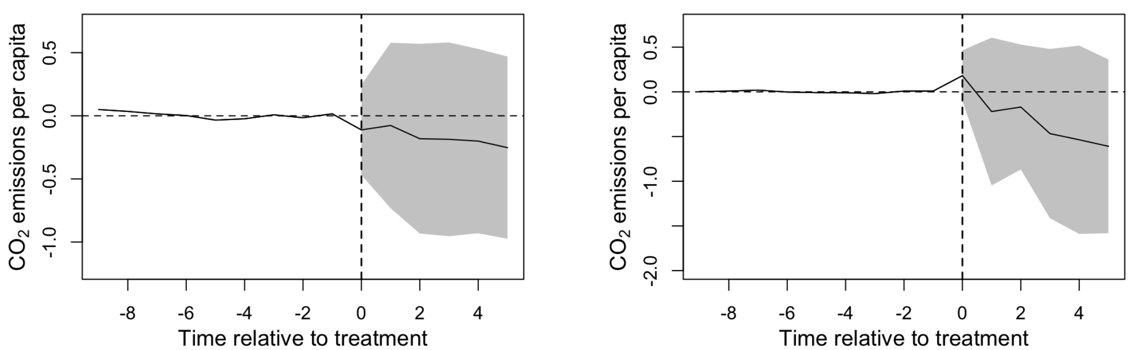
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on a sample that is restricted to Two Control Zones cities. The effects are normalised relative to the beginning of treatment, i.e. 2009 for wave I and 2011 for Wave II.

FIGURE D.2: Controlling for policy overlap – Two Control Zones cities



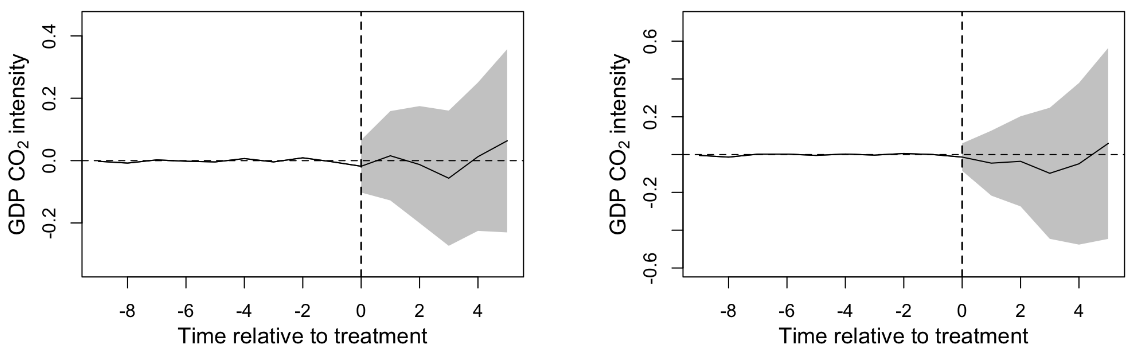
Notes: The figure shows the results of examining the anticipation effect on CO_2 emissions per capita and GDP CO_2 intensity using staggered synthetic control method on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2009 for wave I and 2011 for Wave II.

FIGURE D.3: Controlling for potential anticipation effect – alternative policy start



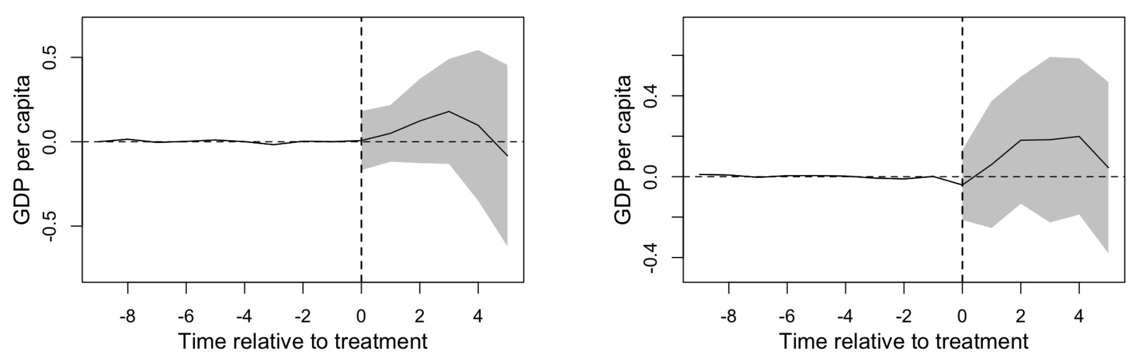
Notes: The figure shows the results of examining the treatment spillovers on CO_2 emissions per capita using staggered synthetic control method. The left panel excludes neighbouring cities, and the right panel uses neighbouring cities as donor units. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.4: Controlling for treatment spillovers – CO_2 emissions per capita



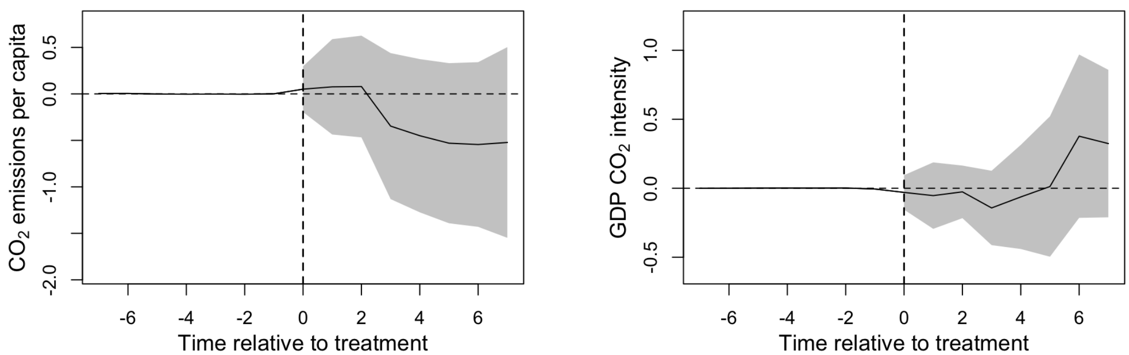
Notes: The figure shows the results of examining the treatment spillovers on GDP CO_2 intensity using staggered synthetic control method. The left panel excludes neighbouring cities, and the right panel uses neighbouring cities as donor units. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.5: Controlling for treatment spillovers – GDP CO_2 intensity



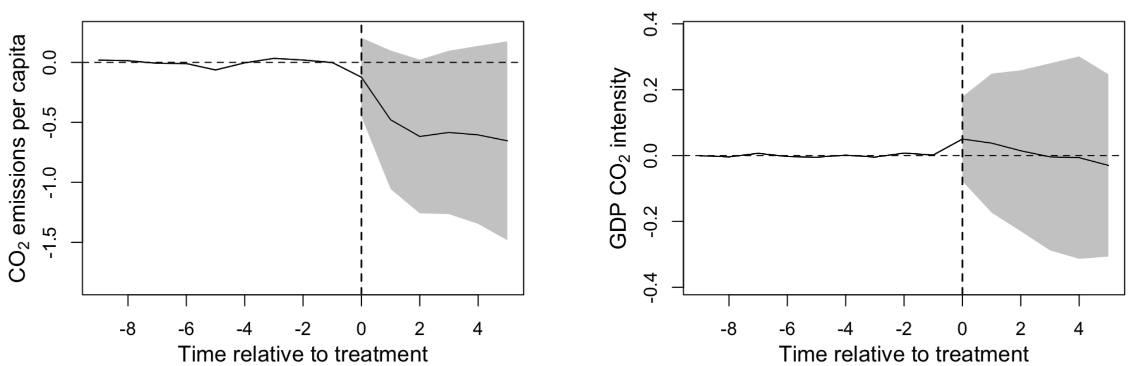
Notes: The figure shows the results of examining the treatment spillovers on GDP per capita using staggered synthetic control method. The left panel excludes neighbouring cities, and the right panel uses neighbouring cities as donor units. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.6: Controlling for treatment spillovers – GDP per capita



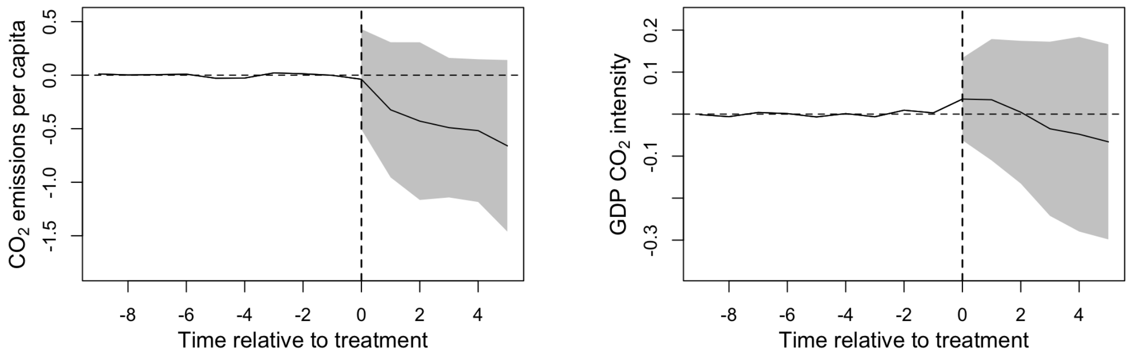
Notes: The figure shows the results of the first wave on CO_2 emissions per capita and GDP CO_2 intensity using synthetic control method on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment.

FIGURE D.7: Checking for robustness – results for the first wave



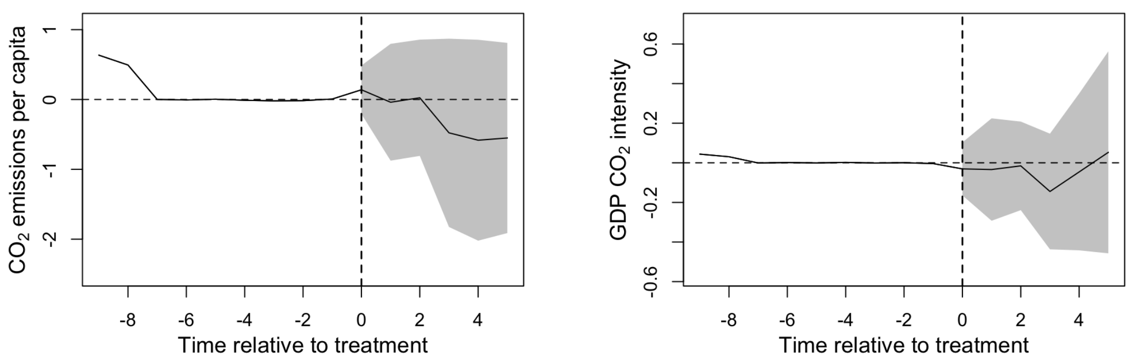
Notes: The figure shows the results of the second wave on CO_2 emissions per capita and GDP CO_2 intensity using synthetic control method on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment.

FIGURE D.8: Checking for robustness – results for the second wave



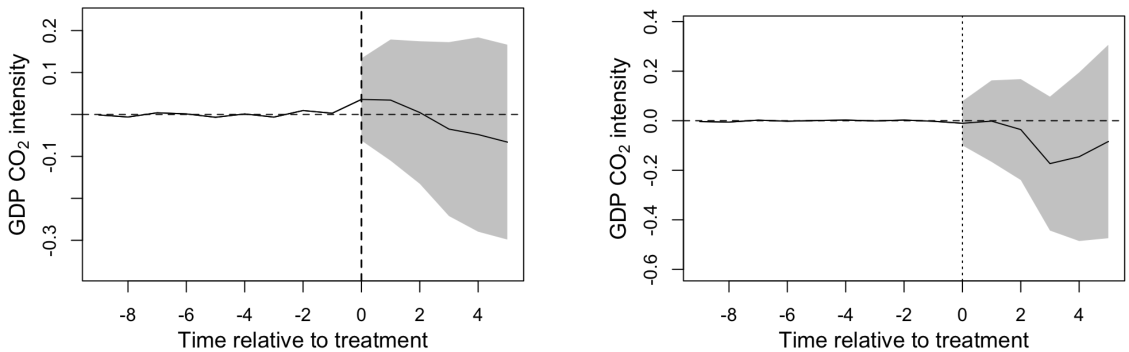
Notes: The figure shows the results of the city-level treatment on CO_2 emissions per capita and GDP CO_2 intensity using staggered synthetic control method on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.9: Checking for robustness – city-level treatment



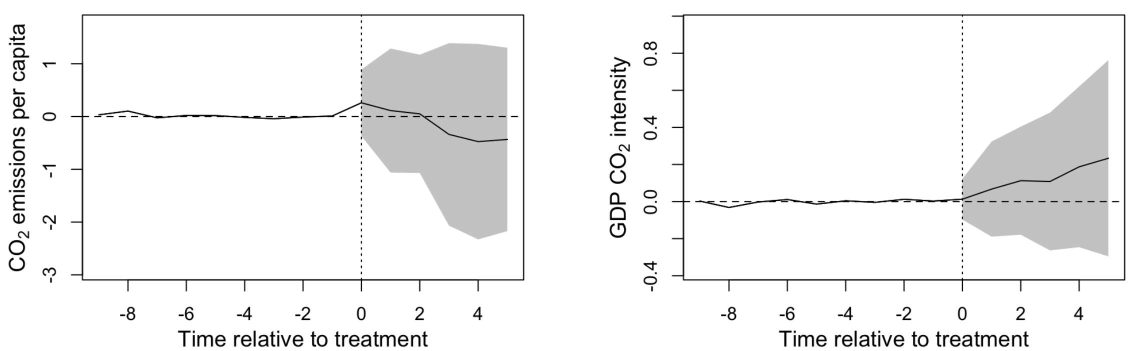
Notes: The figure shows the results of the province-level treatment on CO_2 emissions per capita and GDP CO_2 intensity using staggered synthetic control method on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.10: Checking for robustness – province-level treatment



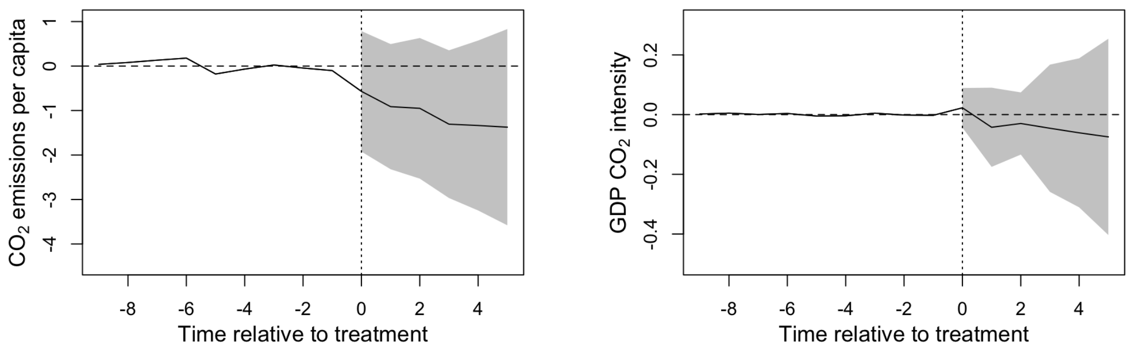
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on low-income cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.11: Checking for robustness – low-income cities



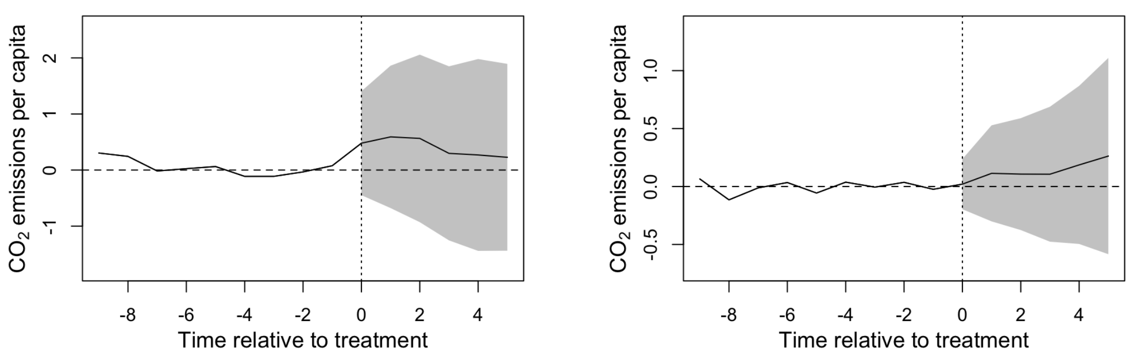
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on middle-income cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.12: Checking for robustness – middle-income cities



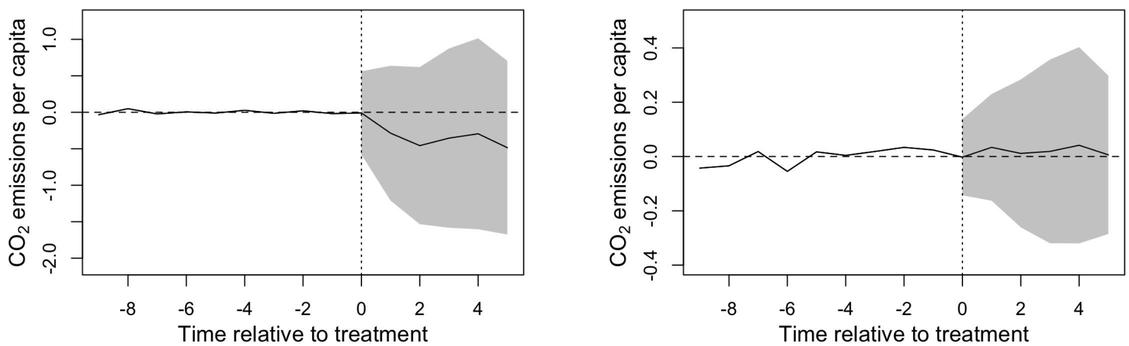
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on high-income cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.13: Checking for robustness – high-income cities



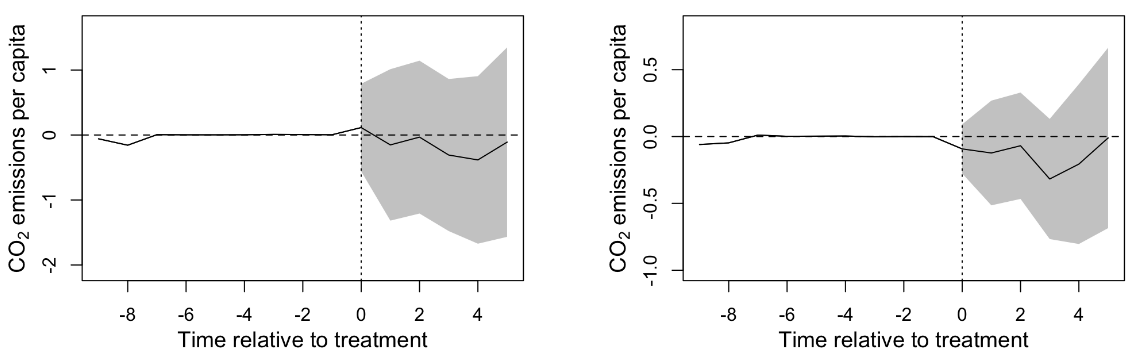
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on northwestern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.14: Checking for robustness – northwestern cities



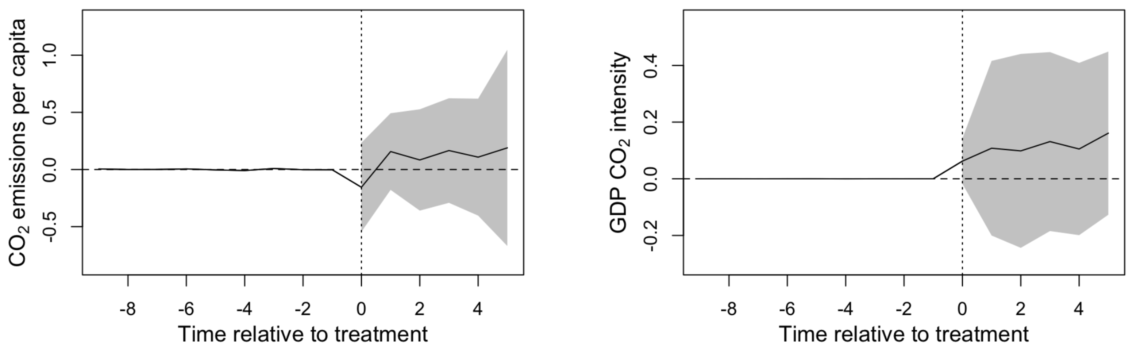
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on northern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.15: Checking for robustness – northern cities



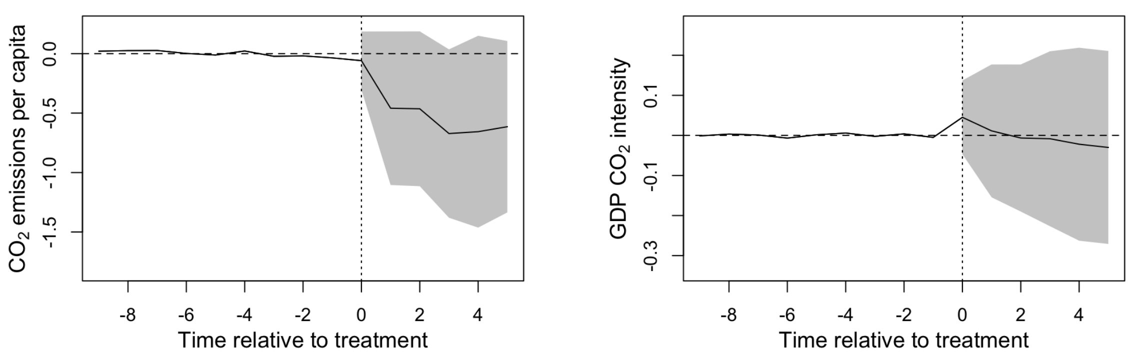
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on northeastern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.16: Checking for robustness – northeastern cities



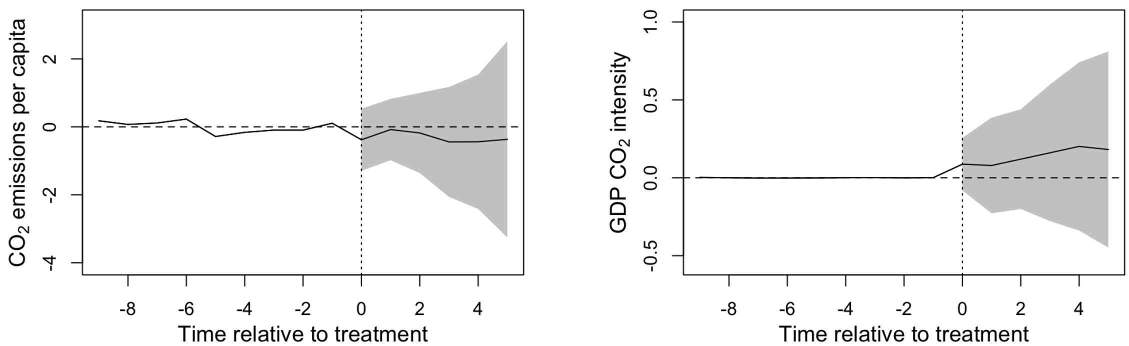
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on central cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.17: Checking for robustness – central cities



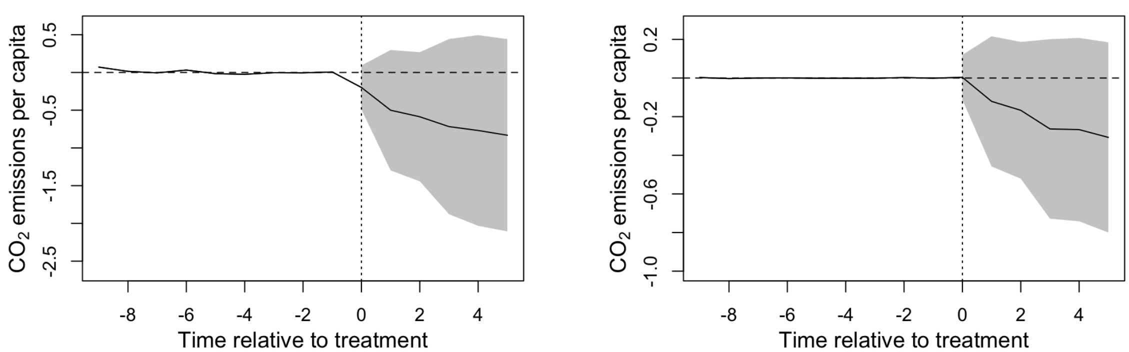
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on eastern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.18: Checking for robustness – eastern cities



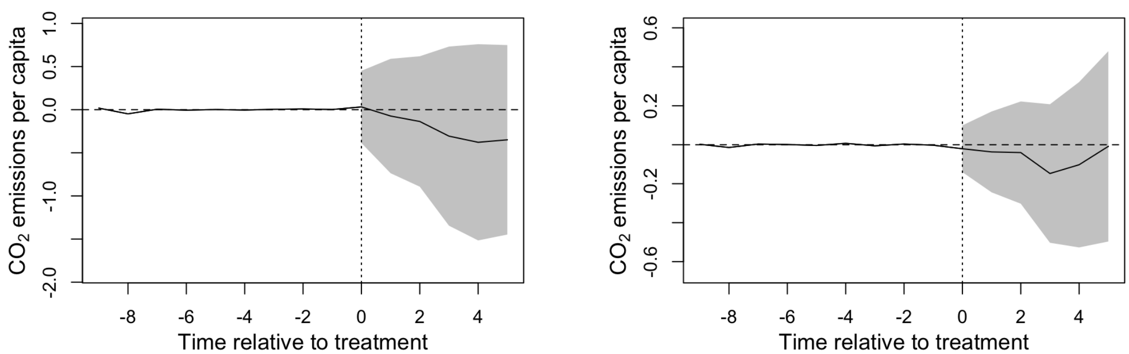
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on southern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.19: Checking for robustness – southern cities



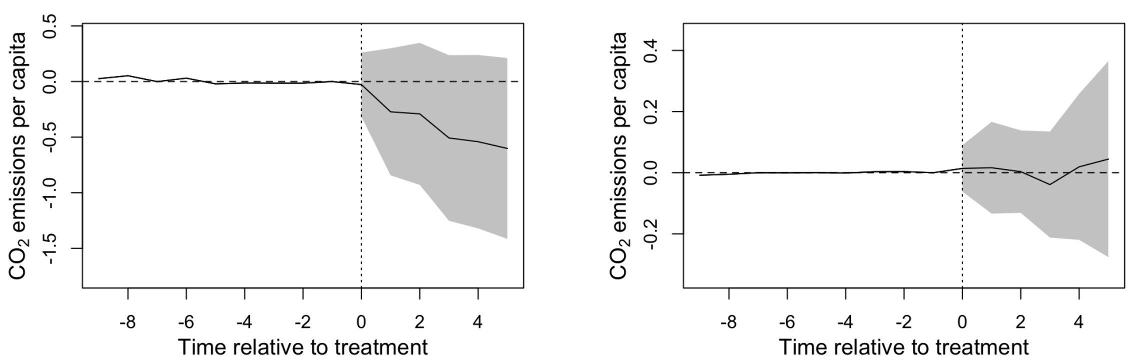
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on southwestern cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.20: Checking for robustness – southwestern cities



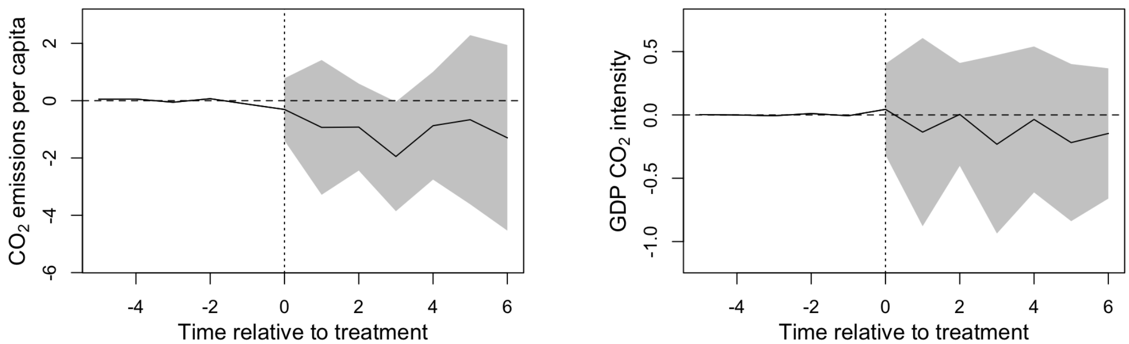
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on resource-based cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.21: Checking for robustness – resource-based cities



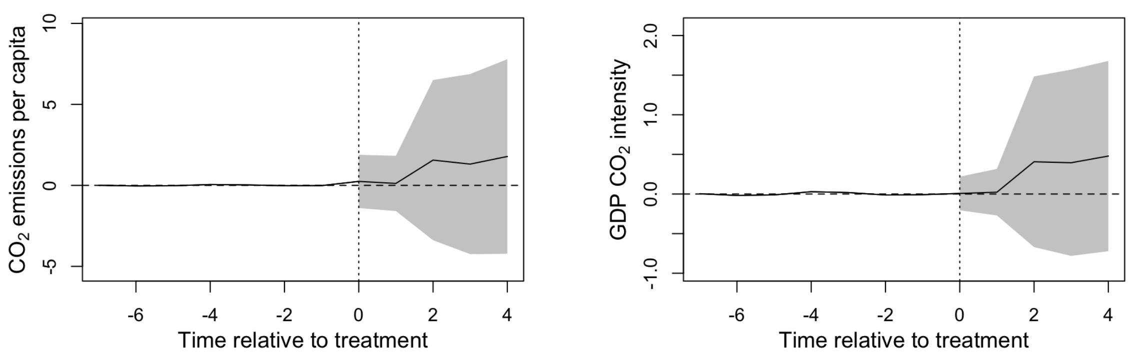
Notes: The figure shows the results of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity on non-resource-based cities which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.22: Checking for robustness – non-resource-based cities



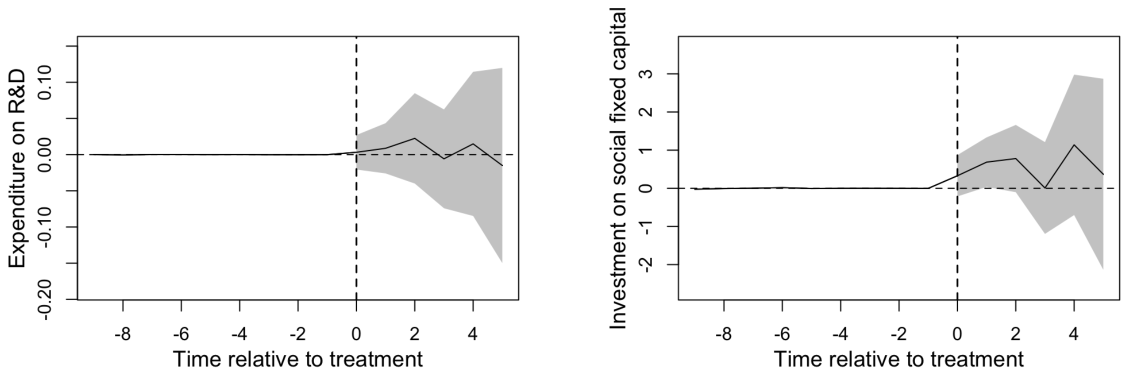
Notes: The figure shows the results of the first wave on CO_2 emissions per capita and GDP CO_2 intensity using synthetic control method on an alternative dataset which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment.

FIGURE D.23: Alternative emissions data – results for the first wave



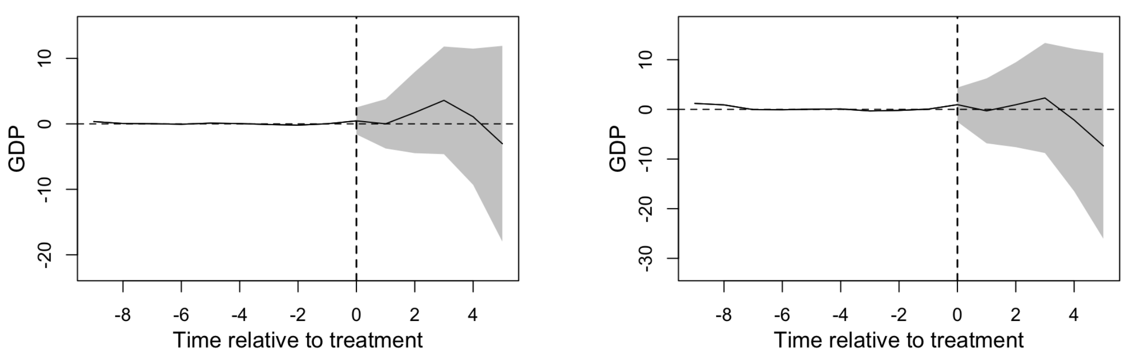
Notes: The figure shows the results of the second wave on CO_2 emissions per capita and GDP CO_2 intensity using staggered synthetic control method on an alternative dataset which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.24: Alternative emissions data – results for the second wave



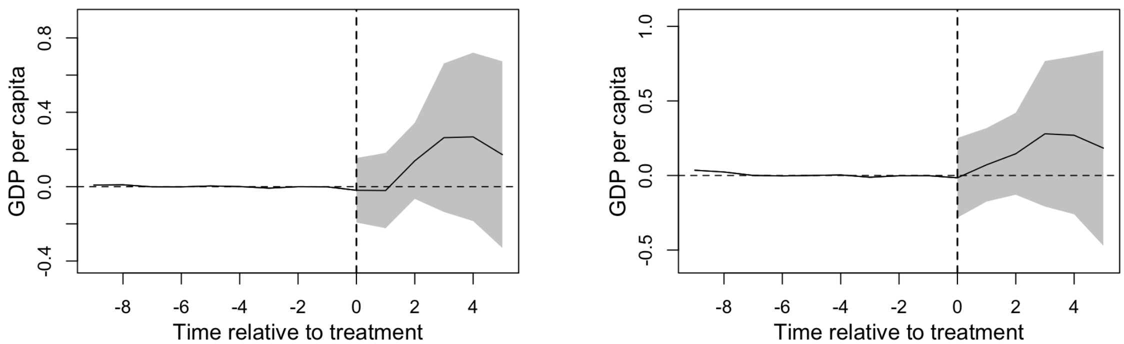
Notes: The figure shows the results of the staggered synthetic control method on expenditure on science and technology and social fixed asset investment on a sample which excludes all cities treated under China's ETS pilots. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.25: Alternative outcome – expenditure on science and technology and social fixed capital



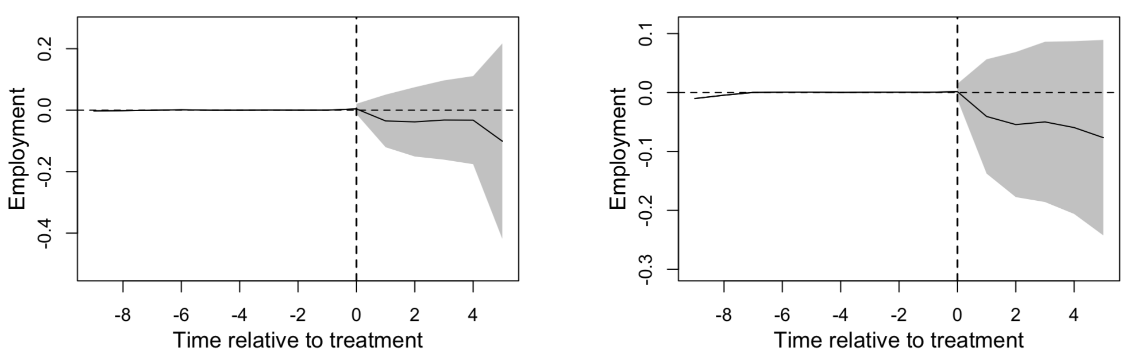
Notes: The figure shows the results of the staggered synthetic control method on gross domestic product on a sample which excludes all cities treated under China's ETS pilots. The left panel displays the results for the entire sample, while the right panel is restricted to cities that have published agendas. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.26: Alternative outcome – gross domestic product



Notes: The figure shows the results of the staggered synthetic control method on gross domestic product per cpaita on a sample which excludes all cities treated under China's ETS pilots. The left panel displays the results for the entire sample, while the right panel is restricted to cities that have published agendas. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.27: Alternative outcome – gross domestic product per capita



Notes: The figure shows the results of the staggered synthetic control method on employment on a sample which excludes all cities treated under China's ETS pilots. The left panel displays the results for the entire sample, while the right panel is restricted to cities that have published agendas. The effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

FIGURE D.28: Alternative outcome – employment

TABLE D.2: Staggered synthetic control estimation - excluding ETS pilot cities

	CO_2 emissions per capita	GDP CO_2 intensity
Panel A. Different affluence levels		
Low-income cities	-0.236	-0.075
	(0.204)	(0.109)
Middle-income cities	-0.138	0.120
	(0.658)	(0.155)
High-income cities	-1.075	-0.039
	(0.830)	(0.082)
Panel B. Different geographical regions		
Northwest China	0.405	0.133
	(0.700)	(0.244)
North China	-0.314	0.018
	(0.510)	(0.125)
Northeast China	-0.145	-0.136
	(0.389)	(0.193)
Central China	0.091	0.111
	(0.213)	(0.120)
East China	-0.488*	-0,002
	(0.272)	(0.092)
South China	-0.315	0.138
	(0.734)	(0.181)
Southwest China	-0.601	-0.187
	(0.470)	(0.176)
Panel C. City category		
Resource-based cities	-0.202	-0.059
	(0.388)	(0.144)
Non-resource-based cities	-0.374	0.010
	(0.305)	(0.074)

Notes: (i) The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP on a sample which excludes all cities treated under China's ETS pilots. (ii) *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

TABLE D.3: Distribution of LCCP cities by affluence levels and geography

	China's Low-Carbon City Pilot	
	First wave	Second wave
Panel A. Affluence levels		
High-income cities	Hangzhou, Xiamen	Zhenjiang, Ningbo
Middle-income cities	Anshan, Fushun, Benxi, Yingkou, Panjin, Nanchang, Kunming, Xi'an, Yan'an, Yulin (Shaanxi)	Jilin, Wenzhou, Qingdao, Kunming, Yan'an, Jinchang, Ürümqi
Low-income cities	Baoding, Dandong, Jinzhou, Fuxin, Liaoyang, Tieling, Huludao, Guiyang, Qujing, Baoshan, Zhaotong, Lijiang, Pu'er, Lincang, Tongchuan, Baoji, Xianyang, Weinan, Hanzhong	Shijiazhuang, Qinhuangdao, Jincheng, Hulunbuir, Huai'an, Chizhou, Nanping, Jingdezhen, Ganzhou, Guilin, Haikou, Guangyuan, Zunyi
Panel B. Geographical distribution		
North China	Baoding	Shijiazhuang, Qinhuangdao, Jincheng, Qingdao
Northeast China	Anshan, Fushun, Benxi, Dandong, Jinzhou, Yingkou, Fuxin, Liaoyang, Panjin, Tieling, Huludao	Hulunbuir, Jilin
East China	Hangzhou	Huai'an, Zhenjiang, Ningbo, Wenzhou, Chizhou
South China	Xiamen	Nanping, Guilin, Haikou
Central China	Nanchang	Jingdezhen, Ganzhou
Southwest China	Guiyang, Kunming, Qujing, Baoshan, Zhaotong, Lijiang, Pu'er, Lincang	Guangyuan, Zunyi, Kunming
Northwest China	Xi'an, Tongchuan, Baoji, Xianyang, Weinan, Yan'an, Hanzhong, Yulin (Shaanxi)	Yan'an, Jinchang, Ürümqi

Notes: The table displays the list of LCCP cities by affluence levels and geographic locations. Note that Yulin may refer to multiple prefecture-level cities, therefore, we use Yulin (Shaanxi) to avoid confusion.

TABLE D.4: Economic sectors

No.	Economic sectors	Category
1	Farming, Forestry, Animal Husbandry, Fishery and Water Conservancy	Agriculture
2	Coal Mining and Dressing	Mining
3	Petroleum and Natural Gas Extraction	
4	Ferrous Metals Mining and Dressing	
5	Nonferrous Metals Mining and Dressing	
6	Non-metal Minerals Mining and Dressing	
7	Other Minerals Mining and Dressing	
8	Logging and Transport of Wood and Bamboo	Light Manufacturing
9	Food Processing	
10	Food Production	
11	Beverage Production	
12	Tobacco Processing	
13	Textile Industry	
14	Garments and Other Fibre Products	
15	Leather, Furs, Down and Related Products	
16	Timber Processing, Bamboo, Cane, Palm Fibre & Straw Products	
17	Furniture Manufacturing	
18	Papermaking and Paper Products	
19	Printing and Record Medium Reproduction	
20	Cultural, Educational and Sports Articles	
21	Medical and Pharmaceutical Products	
22	Petroleum Processing and Coking	Heavy Manufacturing
23	Raw Chemical Materials and Chemical Products	
24	Chemical Fibre	
25	Rubber Products	
26	Plastic Products	
27	Non-metal Mineral Products	
28	Smelting and Pressing of Ferrous Metals	
29	Smelting and Pressing of Nonferrous Metals	
30	Metal Products	
31	Ordinary Machinery	
32	Equipment for Special Purposes	
33	Transportation Equipment Manufacturing	
34	Electric Equipment and Machinery	High-tech Manufacturing
35	Electronic and Telecommunications Equipment	
36	Instruments, Meters, Cultural and Office Machinery	
37	Other Manufacturing Industry	
38	Scrap and waste	
39	Production and Supply of Electric Power, Stream and Hot Water	Power Supply Sector
40	Production and Supply of Gas	
41	Production and Supply of Tap Water	
42	Construction	Construction
43	Transportation, Storage, Post and Telecommunication Services	Service industry
44	Wholesale, Retail Trade and Catering Services	
45	Other Service Sectors	
46	Urban Resident Energy Usage	Household usage
47	Rural Resident Energy Usage	

Notes: The table shows the economic sectors and categorisation. In general, we categorise 47 economic sectors into nine broader categories, partly following the suggestion in Shan et al. (2018b).

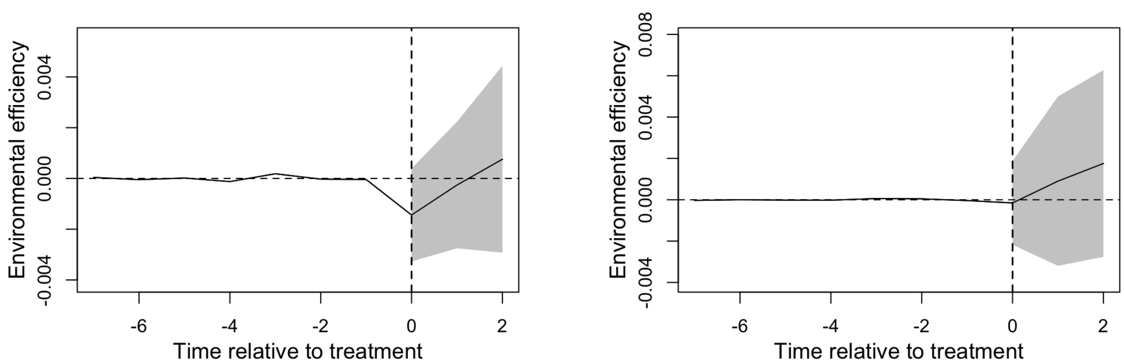
TABLE D.5: Fossil fuel types

No.	Fossil fuel types	Category
1	Raw Coal	Coal Products
2	Cleaned Coal	
3	Other Washed Coal	
4	Briquettes	
5	Coke	
6	Other Coking Products	
7	Coke Oven Gas	Gas
8	Other Gas	
9	Liquefied Petroleum Gas	
10	Refinery Gas	
11	Natural Gas	
12	Crude Oil	Oil products
13	Gasoline	
14	Kerosene	
15	Diesel Oil	
16	Fuel Oil	
17	Other Petroleum Products	Petroleum Products

Note: The table shows the fossil fuel types and categorisation. In general, we categorise 17 types of fossil fuel into 4 broader categories. “Briquettes” includes briquettes and gangue. “Other gas” includes blast furnace gas, converter gas and other unclassified gas. “Other petroleum products” includes naphtha, lubricants, paraffin, white spirit, bitumen asphalt, petroleum coke and other unclassified petroleum products.

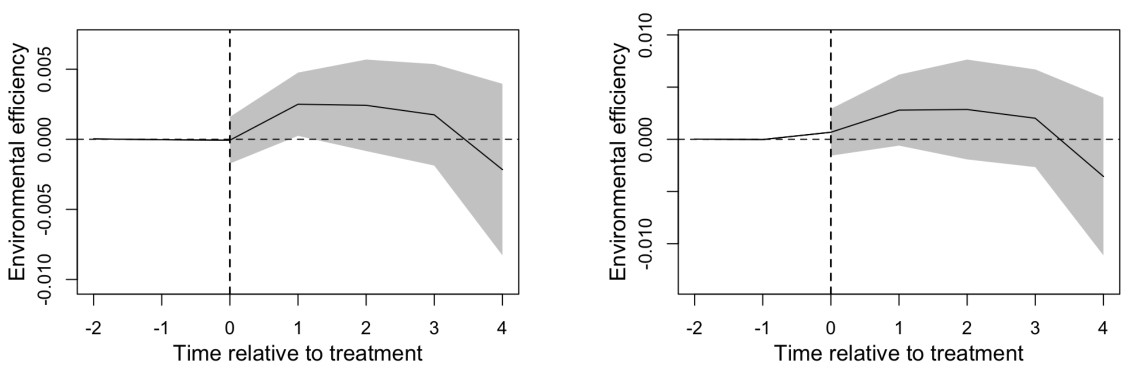
Appendix E

Additional figures and tables for Chapter 3



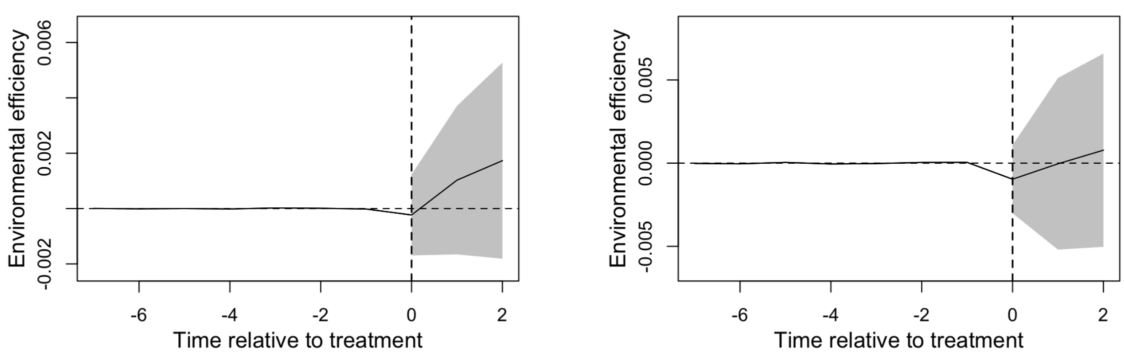
Notes: The figure shows the heterogeneity analysis for the LCCP first wave, by dividing the non-pilot cities into two groups, depending on whether they share a border with those treated ones.

FIGURE E.1: Heterogeneity analysis – neighbours and peripheries for the first wave



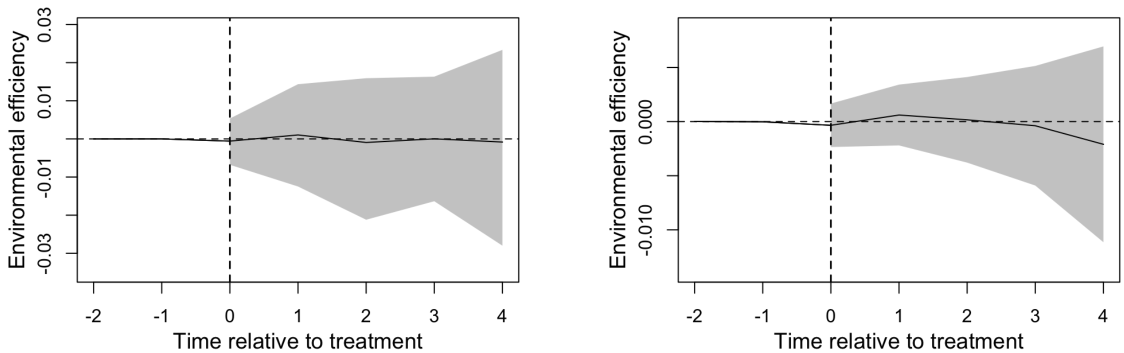
Notes: The figure shows the heterogeneity analysis for the learning effect, by dividing the non-pilot cities into two groups, depending on whether they share a border with those treated ones.

FIGURE E.2: Heterogeneity analysis – neighbours and peripheries for the learning effect



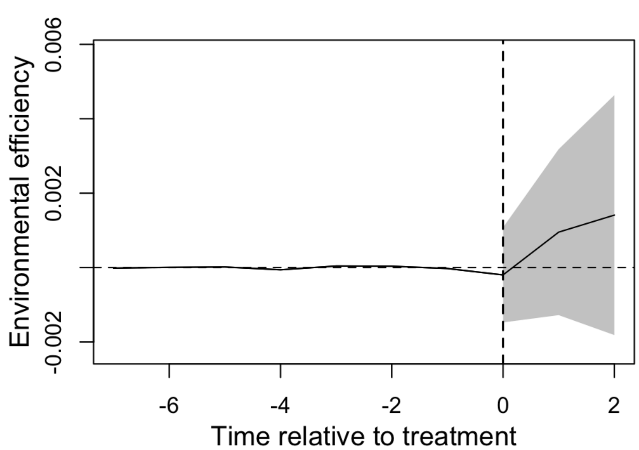
Notes: The figure shows the heterogeneity analysis for the LCCP first wave, by dividing the treated cities into two groups, depending on whether they were assigned to treatment directly or assigned treatment status as part of a province-level treatment assignment.

FIGURE E.3: Heterogeneity analysis – treatment levels for the first wave



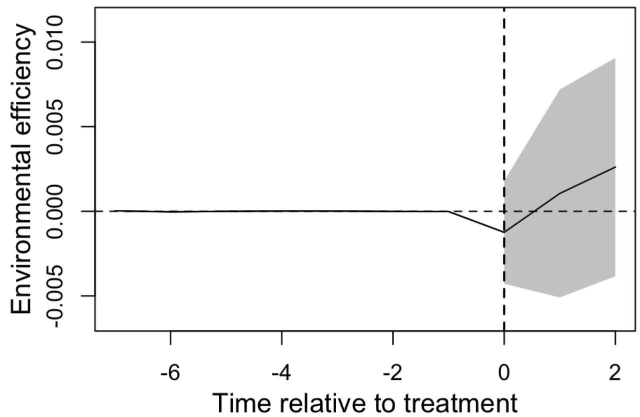
Notes: The figure shows the heterogeneity analysis for the LCCP second wave, by dividing the treated cities into two groups, depending on whether they were assigned to treatment directly or assigned treatment status as part of a province-level treatment assignment.

FIGURE E.4: Heterogeneity analysis – treatment levels for the second wave



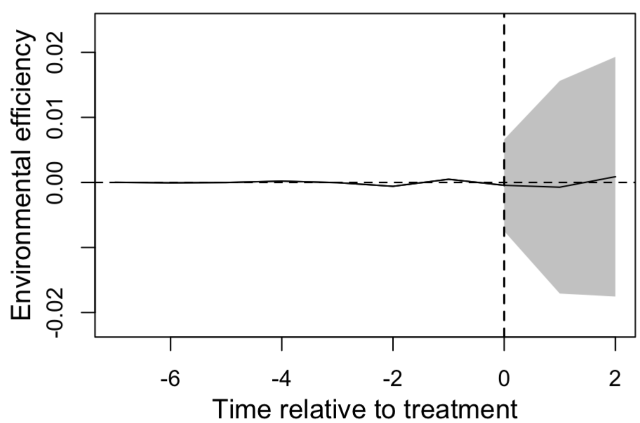
Notes: The figure shows the effect of the LCCP first wave on low-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.5: Heterogeneity analysis – LCCP first wave on low-income cities



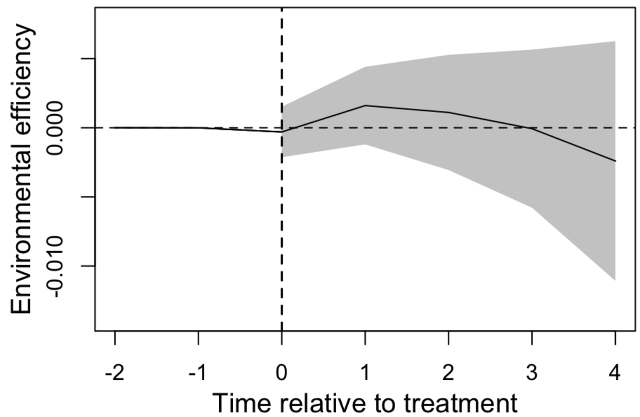
Notes: The figure shows the effect of the LCCP first wave on middle-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.6: Heterogeneity analysis – LCCP first wave on middle-income cities



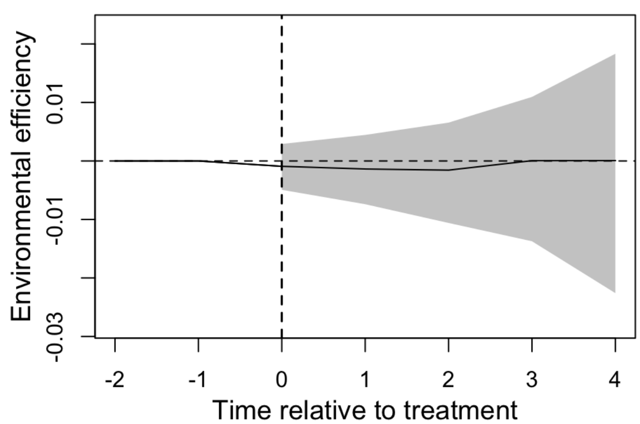
Notes: The figure shows the effect of the LCCP first wave on high-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.7: Heterogeneity analysis – LCCP first wave on high-income cities



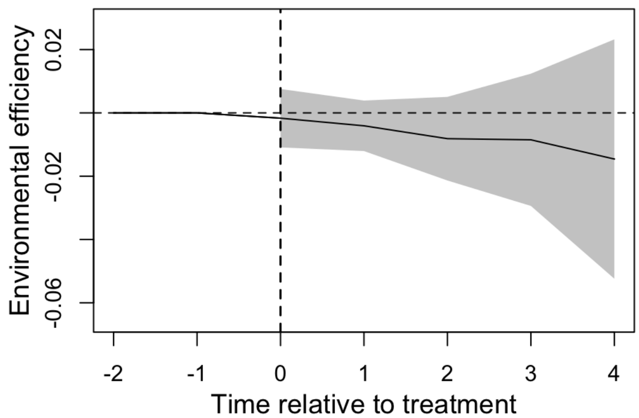
Notes: The figure shows the effect of the LCCP second wave on low-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.8: Heterogeneity analysis – LCCP second wave on low-income cities



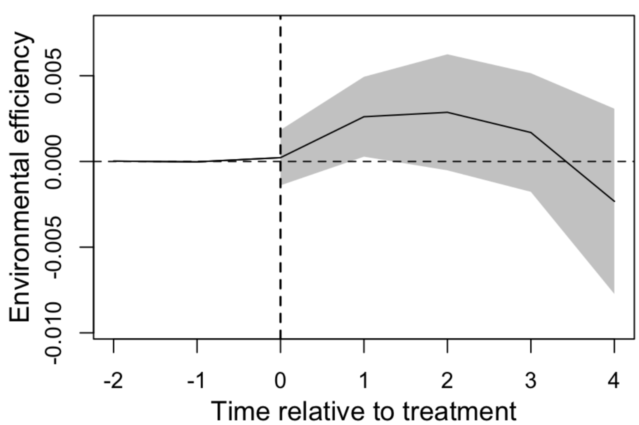
Notes: The figure shows the effect of the LCCP second wave on middle-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.9: Heterogeneity analysis – LCCP second wave on middle-income cities



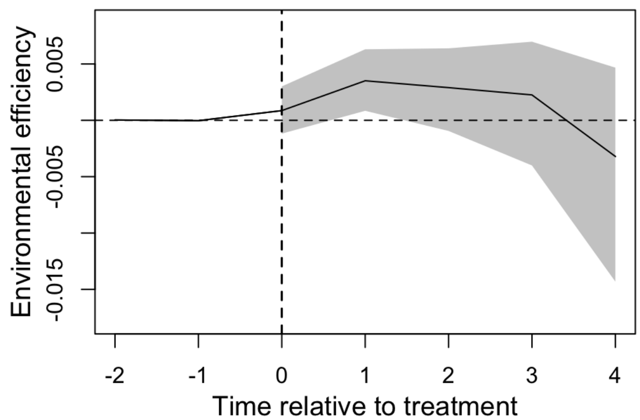
Notes: The figure shows the effect of the LCCP second wave on high-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.10: Heterogeneity analysis – LCCP second wave on high-income cities



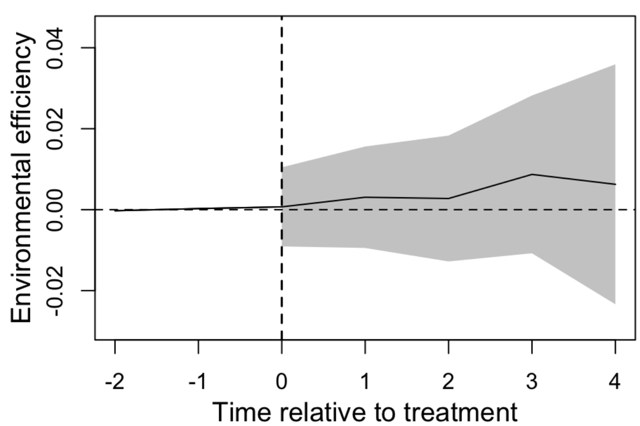
Notes: The figure shows the learning effect on low-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.11: Heterogeneity analysis – learning effect on low-income cities



Notes: The figure shows the learning effect on middle-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.12: Heterogeneity analysis – learning effect on middle-income cities

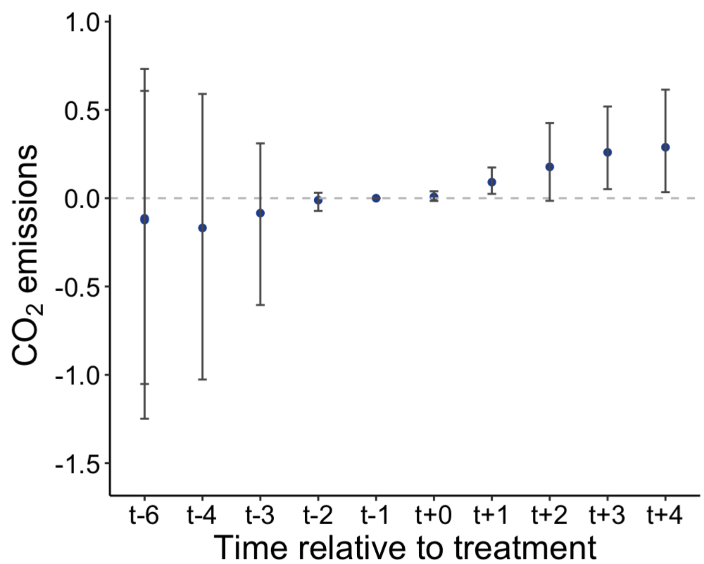


Notes: The figure shows the learning effect on high-income cities' environmental efficiency using the *partially pooled* synthetic control method.

FIGURE E.13: Heterogeneity analysis – learning effect on high-income cities

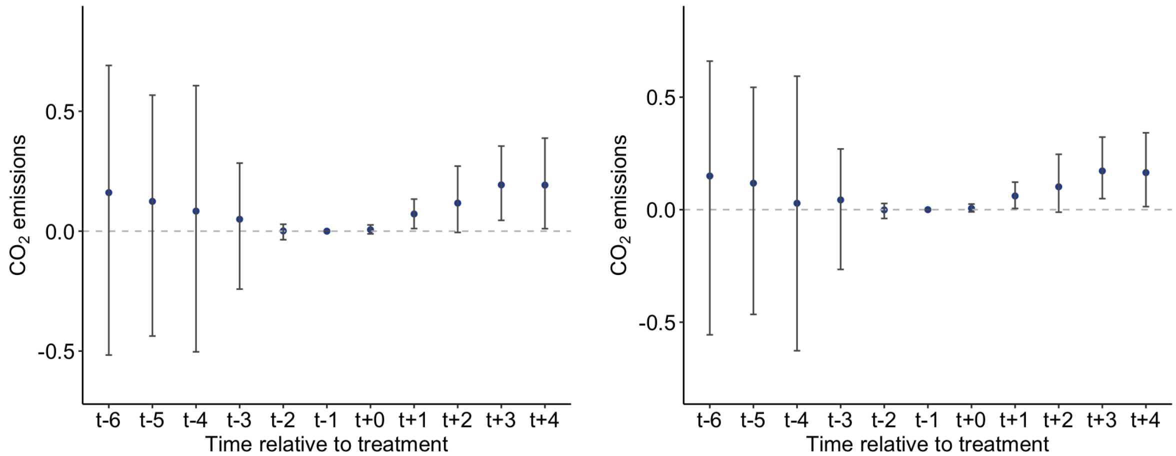
Appendix F

Additional figures and tables for Chapter 4



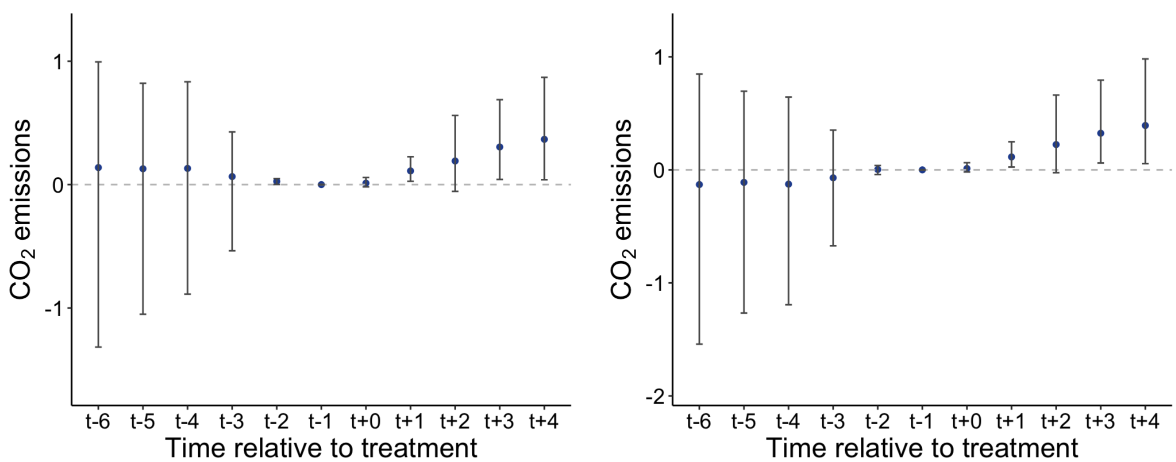
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions using as control variables the significant variables identified by the dynamic DiD approach. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.1: Different specifications – FE covariates



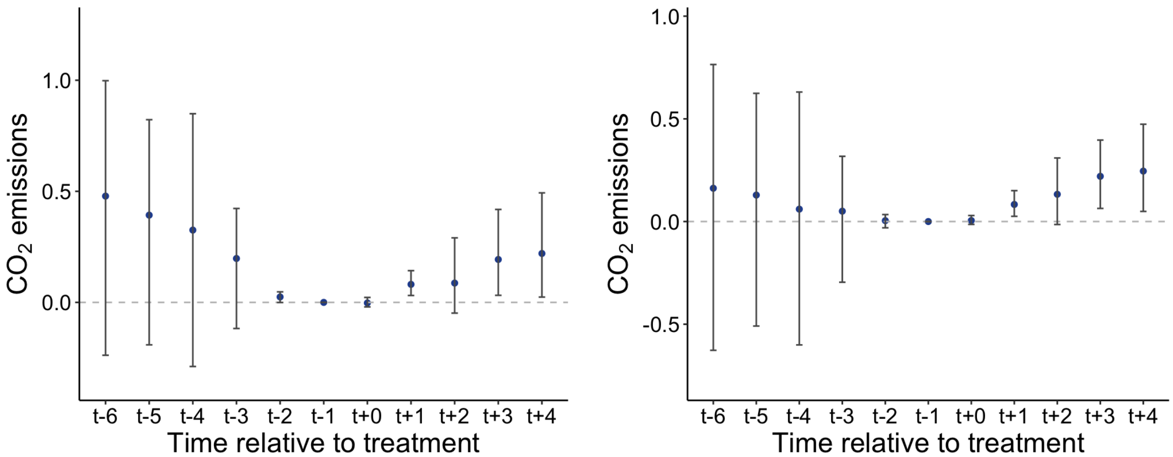
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions using different number of match size. The left panel displays the results where the match size increases to two, and the right panel displays the results where the match size increases to three. Treatment effects are normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.2: Different specifications – alternative match size



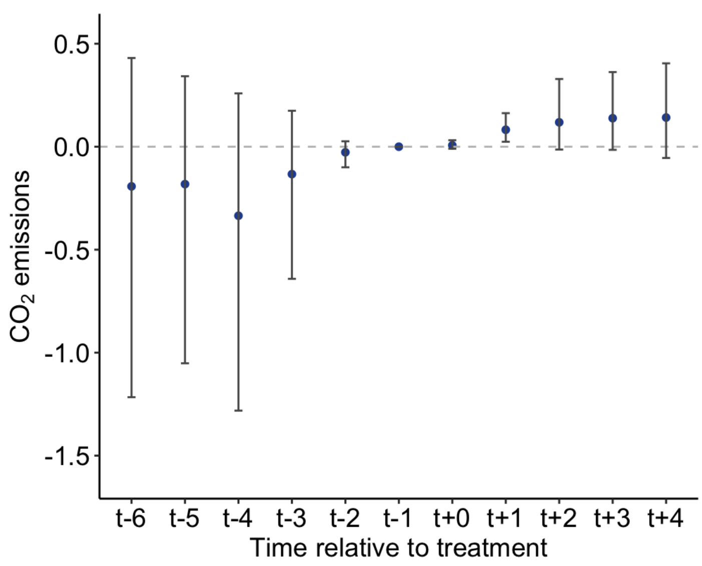
Notes: The figure reports the results of applying alternative matching method on county-level CO_2 emissions. The left panel displays the results of applying propensity score matching, and the right panel displays the results of applying propensity score weighting. Treatment effects are normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.3: Different specifications – propensity score matching and weighting



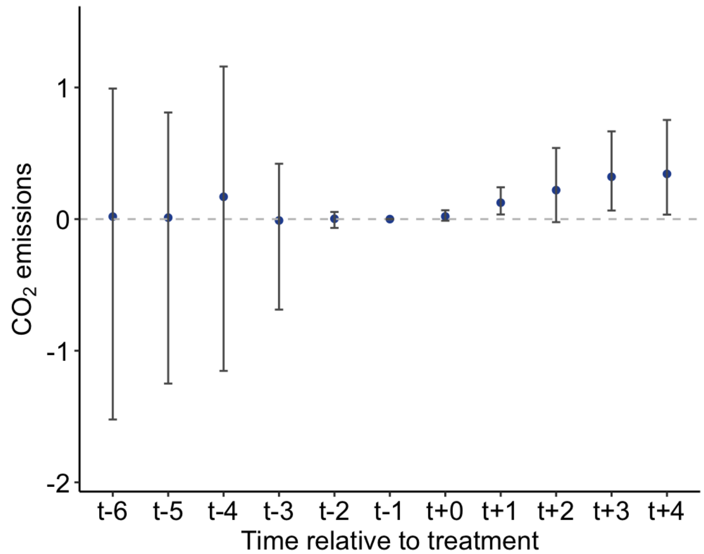
Notes: The figure reports the results of applying alternative matching method on county-level CO_2 emissions. The left panel displays the results of applying covariates balance propensity score matching, and the right panel displays the results of applying covariates balance propensity score weighting. Treatment effects are normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.4: Different specifications – covariates balance propensity score matching and weighting



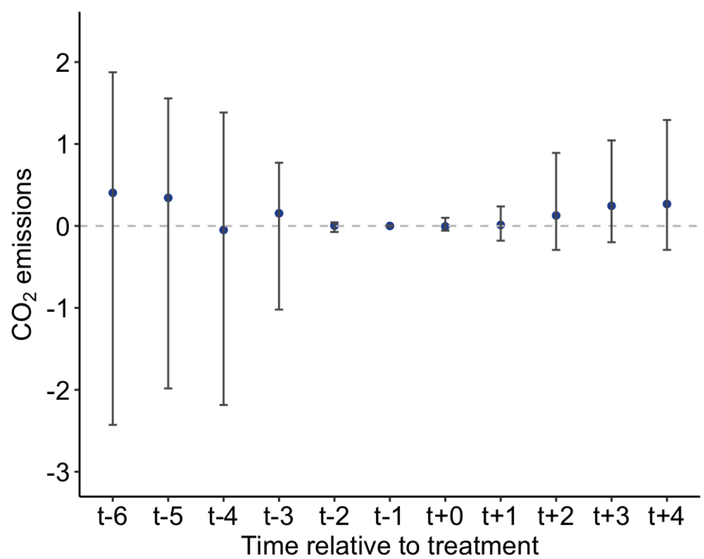
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the low-income counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.5: Heterogeneity analysis – low-income counties



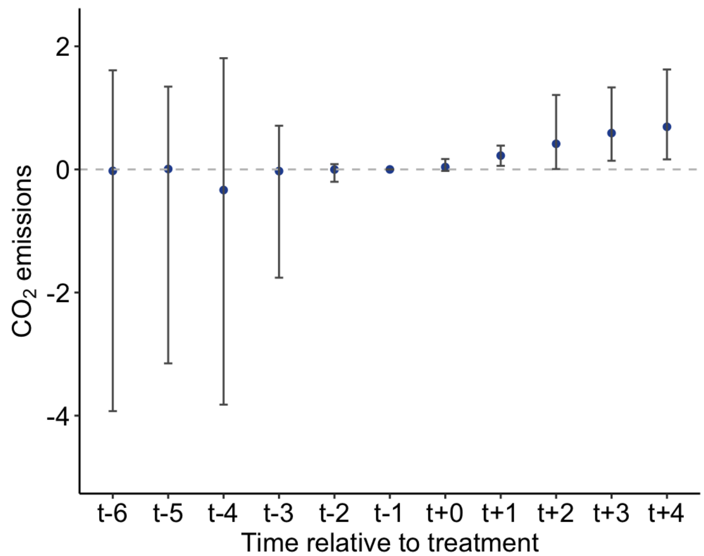
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the middle-income counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.6: Heterogeneity analysis – middle-income counties



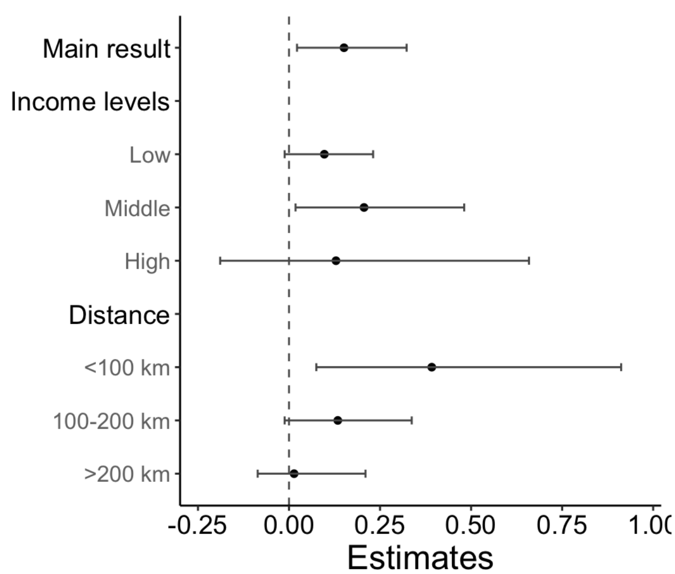
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the high-income counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.7: Heterogeneity analysis – high-income counties



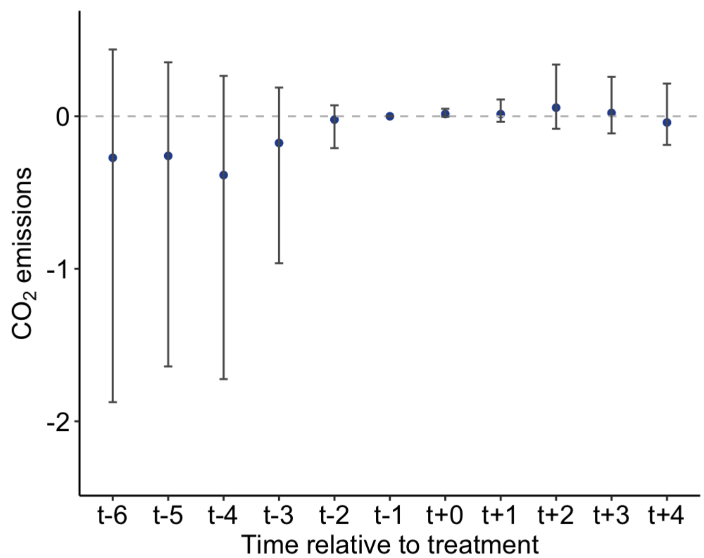
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the close counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.8: Heterogeneity analysis – close counties



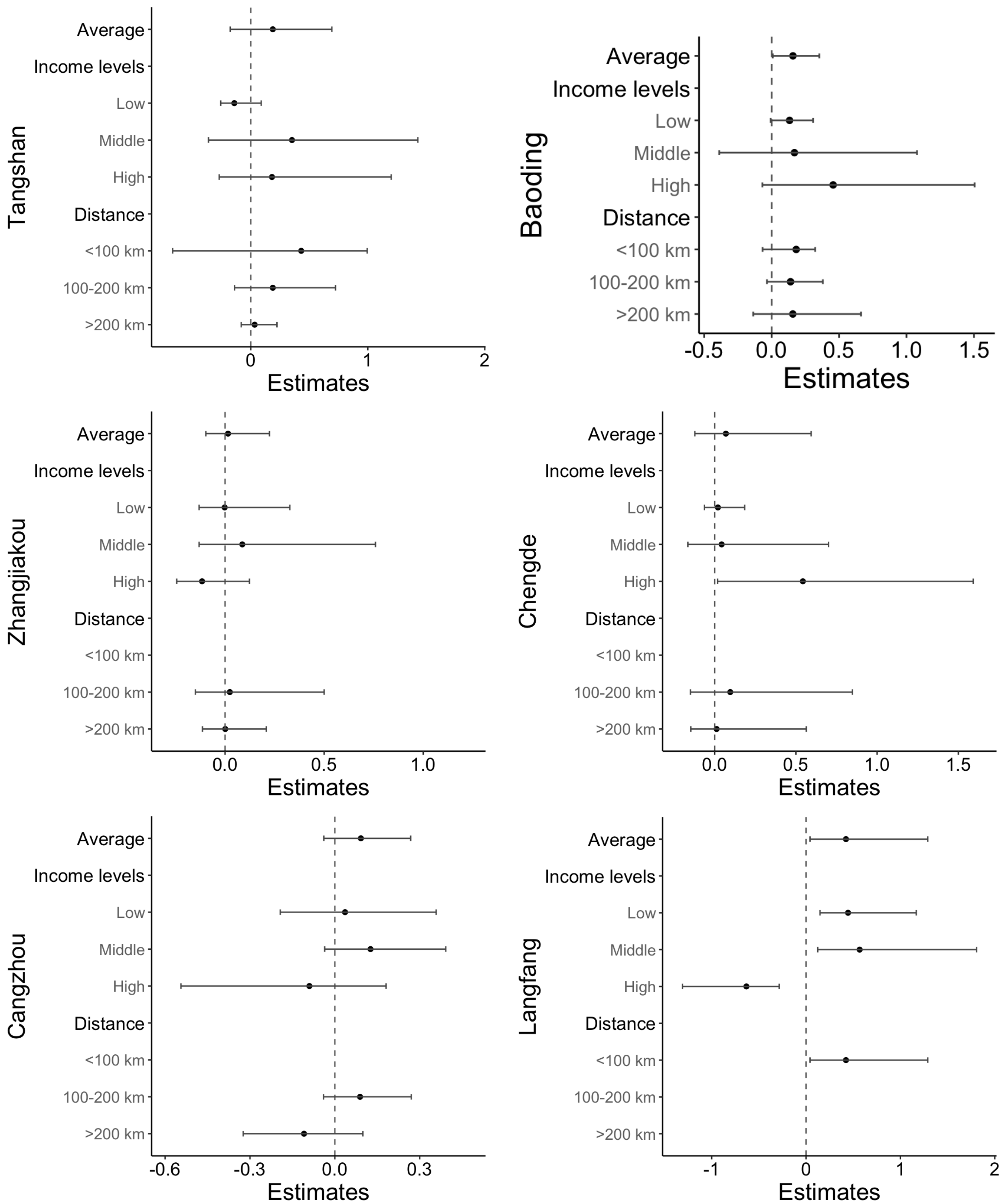
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the nearby counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.9: Heterogeneity analysis – nearby counties



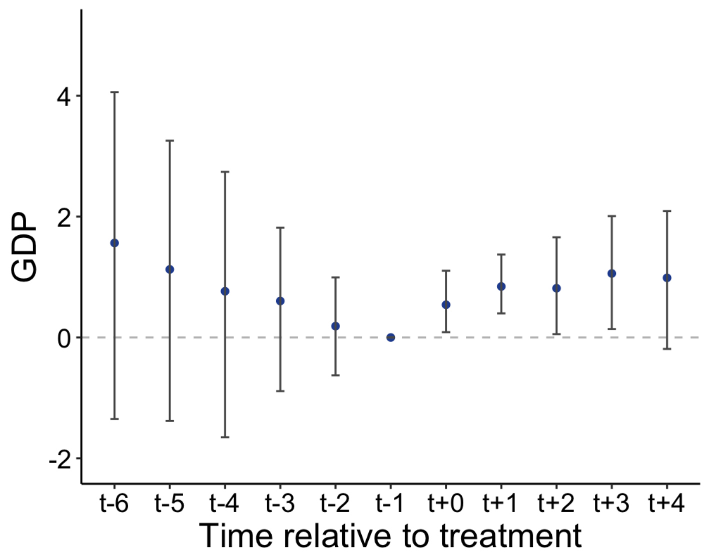
Notes: The figure reports the results of the nearest neighbour matching on county-level CO_2 emissions for the distant counties. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.10: Heterogeneity analysis – distant counties



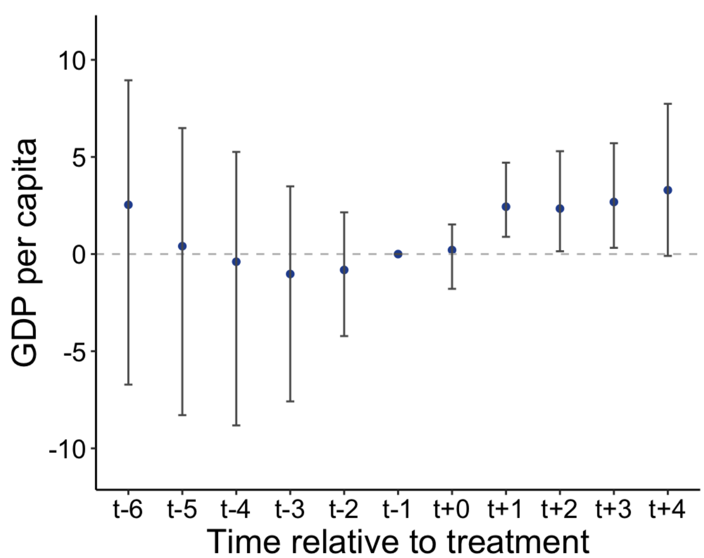
Notes: (i) The figure reports the results of heterogeneity analysis for each neighbouring city. For convenient comparison, we display the city-specific average treatment effects at the top. (ii) We are unable to perform estimations for some categories, since there is no available neighbouring counties due to the county-specific geographical characteristics. (iii) The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.11: Heterogeneity analysis for each neighbouring city



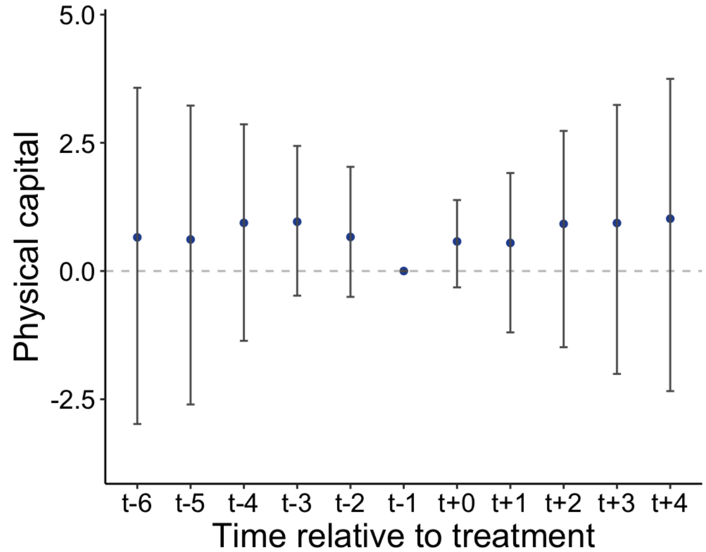
Notes: The figure reports the results of the nearest neighbour matching on GDP. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.12: Further results – GDP



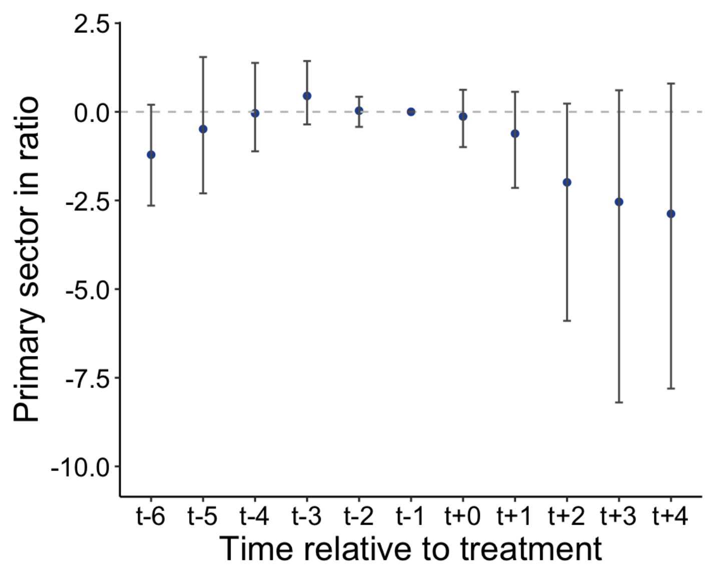
Notes: The figure reports the results of the nearest neighbour matching on GDP per capita. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.13: Further results – GDP per capita



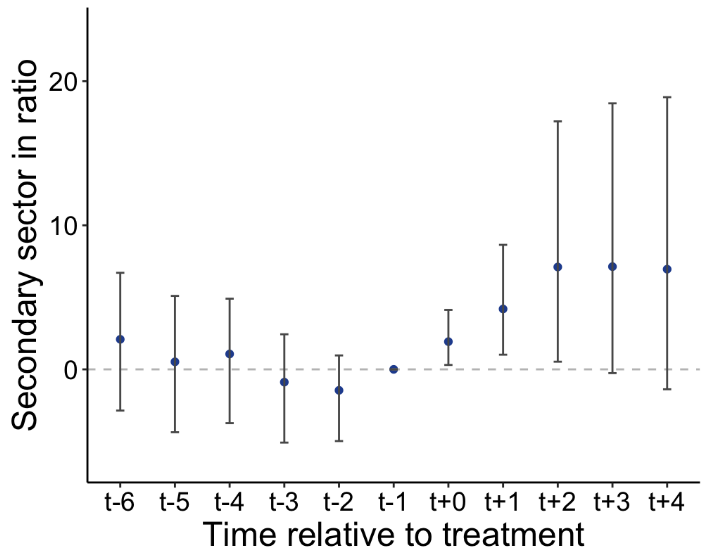
Notes: The figure reports the results of the nearest neighbour matching on county-level physical capital. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.14: Further results – physical capital



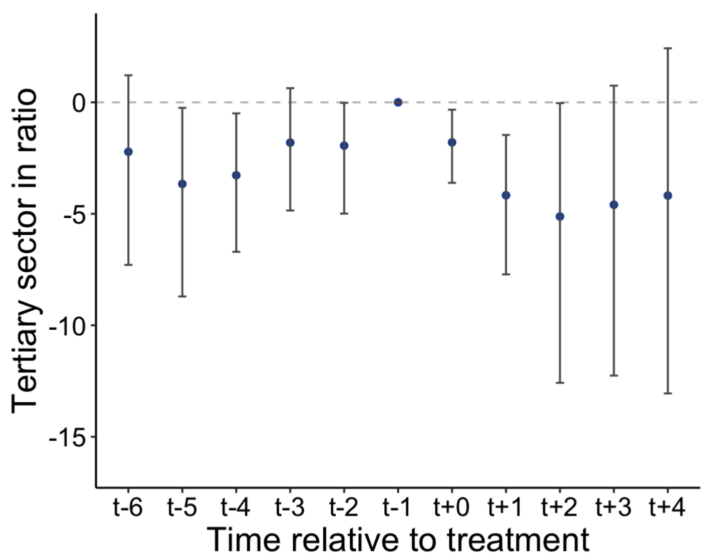
Notes: The figure reports the results of the nearest neighbour matching on share of GDP taken by primary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.15: Further results – share of GDP taken by primary sector



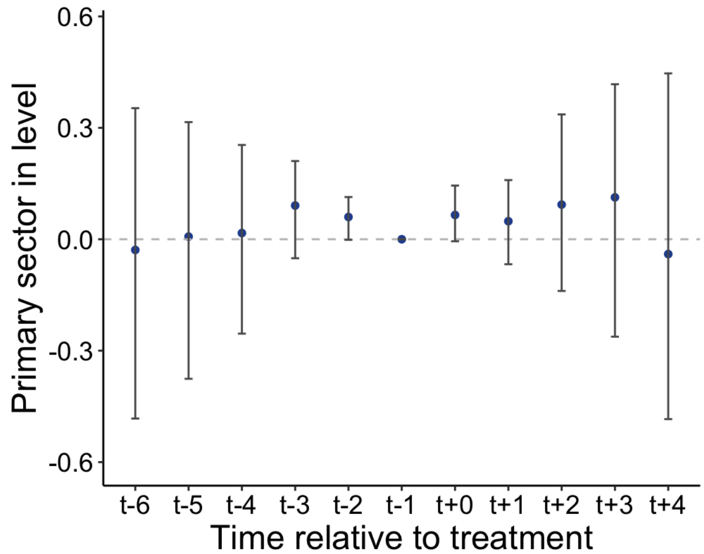
Notes: The figure reports the results of the nearest neighbour matching on share of GDP taken by secondary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.16: Further results – share of GDP taken by secondary sector



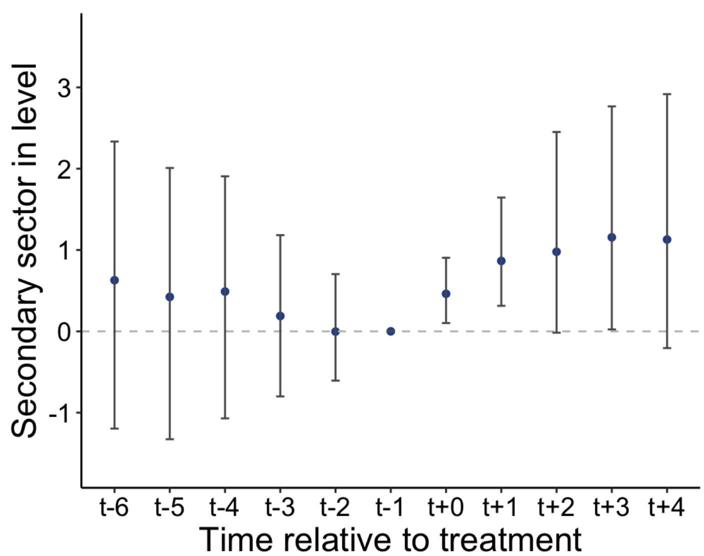
Notes: The figure reports the results of the nearest neighbour matching on share of GDP taken by tertiary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.17: Further results – share of GDP taken by tertiary sector



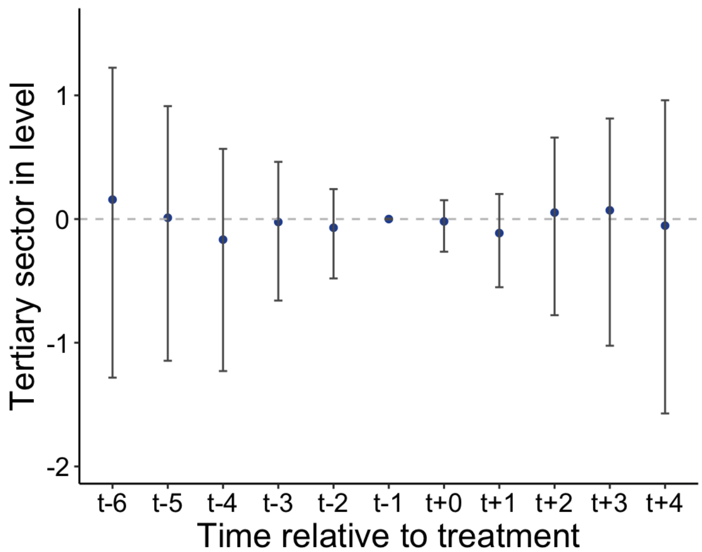
Notes: The figure reports the results of the nearest neighbour matching on GDP taken by primary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.18: Further results – GDP taken by primary sector



Notes: The figure reports the results of the nearest neighbour matching on GDP taken by secondary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.19: Further results – GDP taken by secondary sector



Notes: The figure reports the results of the nearest neighbour matching on GDP taken by tertiary sector. Treatment effect is normalised relative to the beginning of the treatment. The error bars are constructed by the 95% quantiles of the bootstrapped estimates.

FIGURE F.20: Further results – GDP taken by tertiary sector

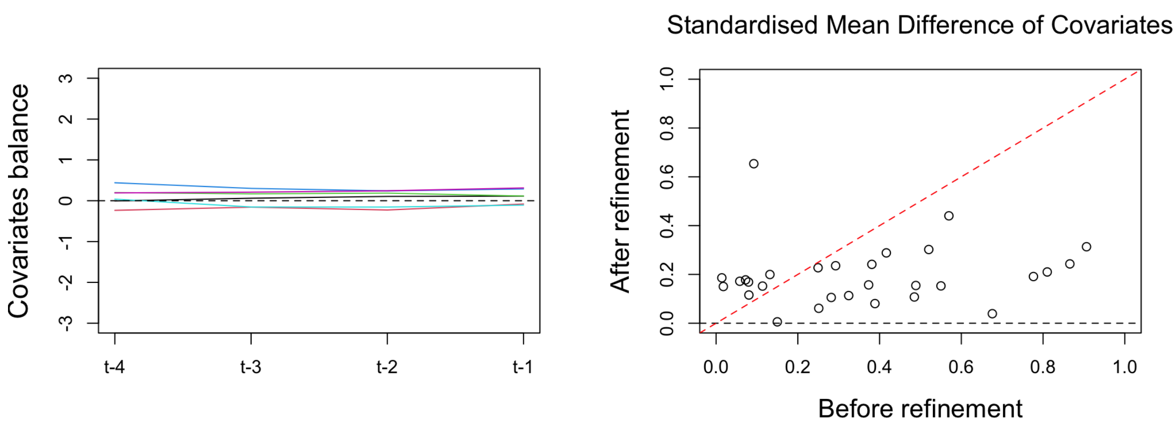
TABLE F.1: Descriptive statistics, 2007-2017

	Mean	SD	Min	Max	N
Panel A. Neighbouring counties					
Outcome variable					
CO_2 emissions (million tonnes)	3.86	2.73	0.21	13.65	1,012
Economic indicators					
GDP per capita (thousand CNY)	30.35	20.54	5.04	142.71	989
Share of GDP taken by secondary sector (%)	46.94	14.42	13.75	90.56	992
Share of GDP taken by social fixed asset investment (%)	83.75	41.01	1.10	261.63	983
Share of GDP taken by fiscal expenditure (%)	14.91	9.28	1.88	88.60	990
Infrastructure indicators					
Highway per land area (km)	1.08	0.48	0.11	3.09	821
Telephones (fixed and mobile) per capita (unit)	0.15	0.08	0.00	0.61	818
Beds in health care institutions per thousands (unit)	3.39	2.32	0.85	25.59	838
Panel B. Counties in peripheries					
Outcome variable					
CO_2 emissions (million tonnes)	3.45	2.39	0.58	14.16	869
Socioeconomic indicators					
GDP per capita (thousand CNY)	25.92	13.79	5.55	111.11	828
Share of GDP taken by secondary sector (%)	47.68	12.90	9.73	84.72	779
Share of GDP taken by social fixed asset investment (%)	88.00	36.18	17.08	223.67	828
Share of GDP taken by fiscal expenditure (%)	12.73	6.74	1.19	41.06	828
Infrastructure indicators					
Highway per land area (km)	1.43	0.47	0.34	3.94	712
Telephones (fixed and mobile) per capita (unit)	0.12	0.09	0.01	0.89	703
Beds in health care institutions per thousands (unit)	3.22	2.08	0.91	17.18	733

Notes: The table shows means, standard deviations, minimum values, maximum values and the number of observations. Panel A displays the statistics for the neighbouring counties, and Panel B displays the statistics for those in the peripheries.

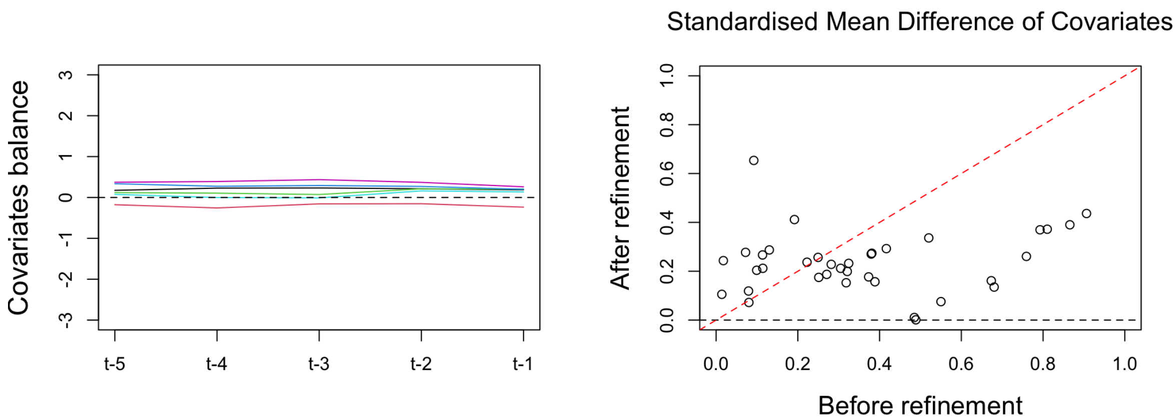
Appendix G

Supplementary figures for Chapter 4



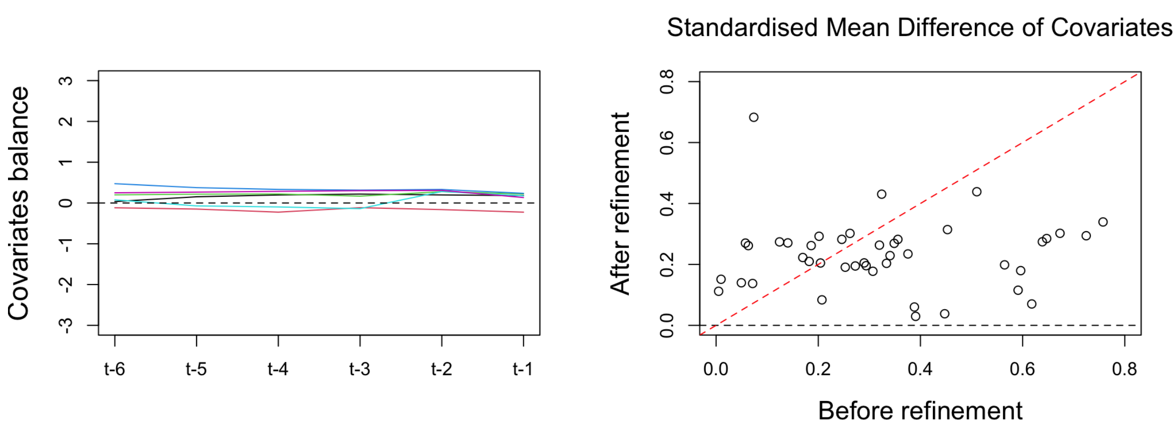
Notes: (i) The figure shows the matching quality for the examination of the Beijing's and Tianjin's ETS pilots. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.1: Covariates balance for the examination of the Beijing's and Tianjin's ETS pilots, alternative start at 2011



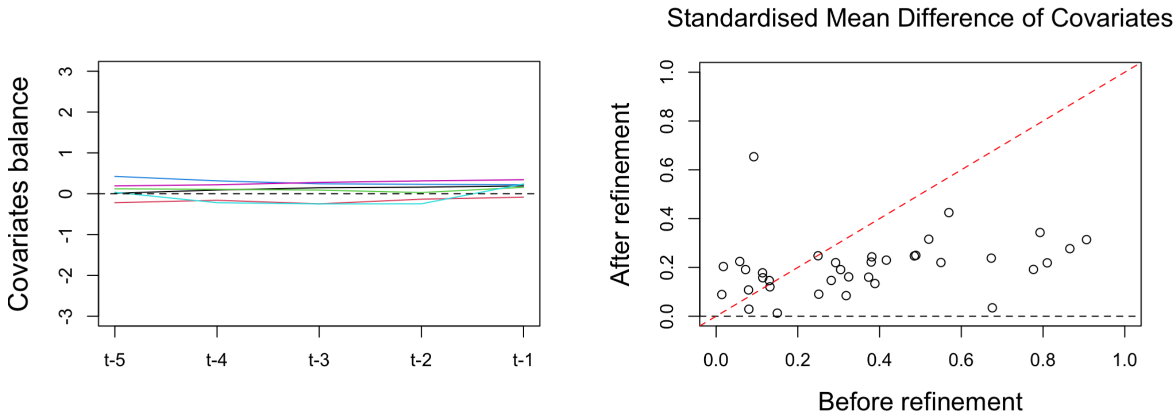
Notes: (i) The figure shows the matching quality for the examination of the Beijing's and Tianjin's ETS pilots. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.2: Covariates balance for the examination of the Beijing's and Tianjin's ETS pilots, alternative start at 2014



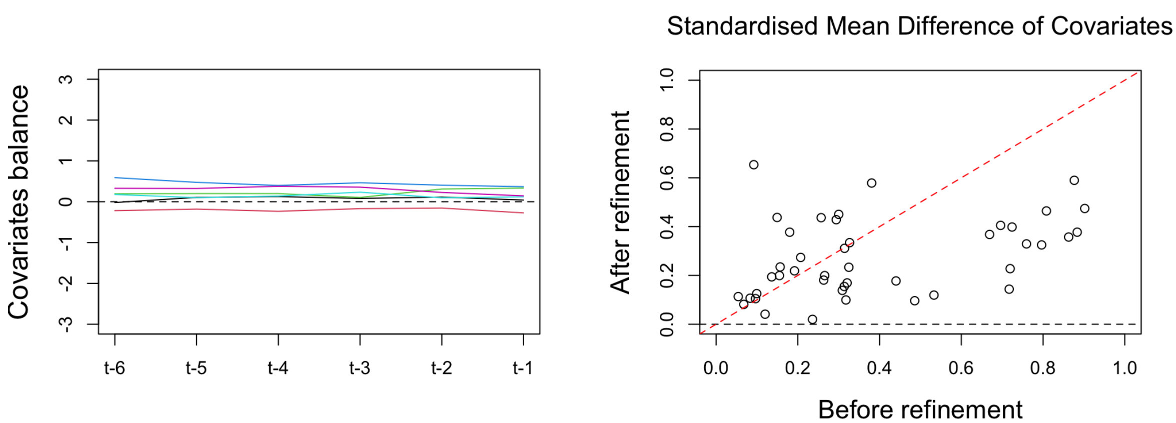
Notes: (i) The figure shows the matching quality for the examination of the administrative adjustments. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.3: Covariates balance for the examination of the administrative adjustments



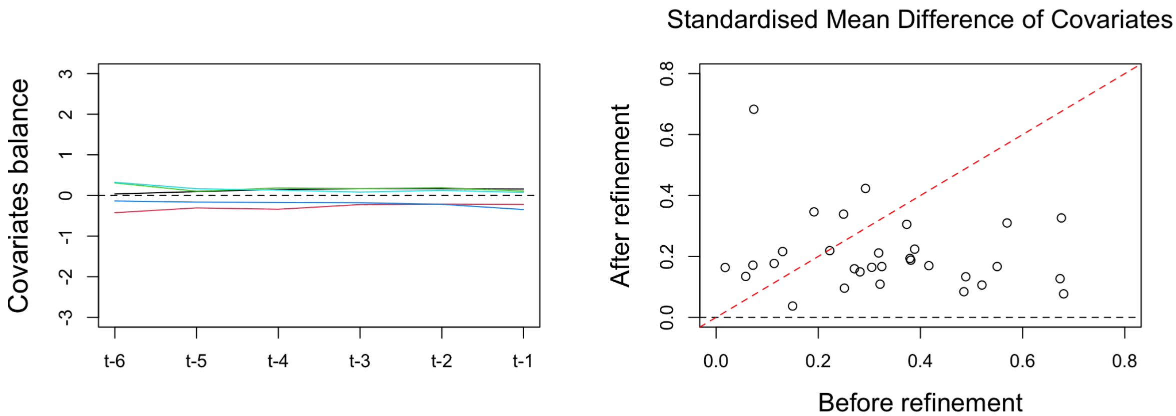
Notes: (i) The figure shows the matching quality for the examination of the announcement effect. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.4: Covariates balance for the examination of the announcement effect



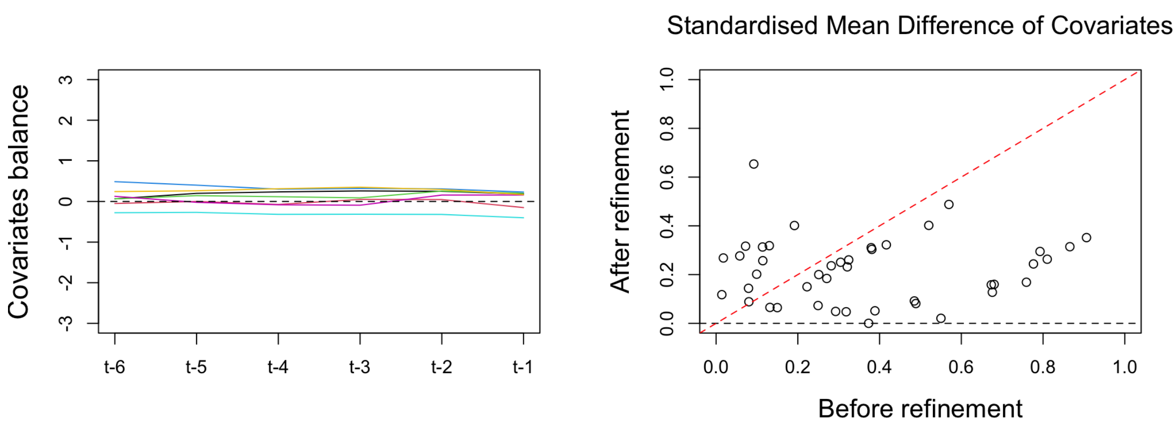
Notes: (i) The figure shows the matching quality for the examination of re-defining the neighbouring cities. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.5: Covariates balance for re-defining the neighbouring cities



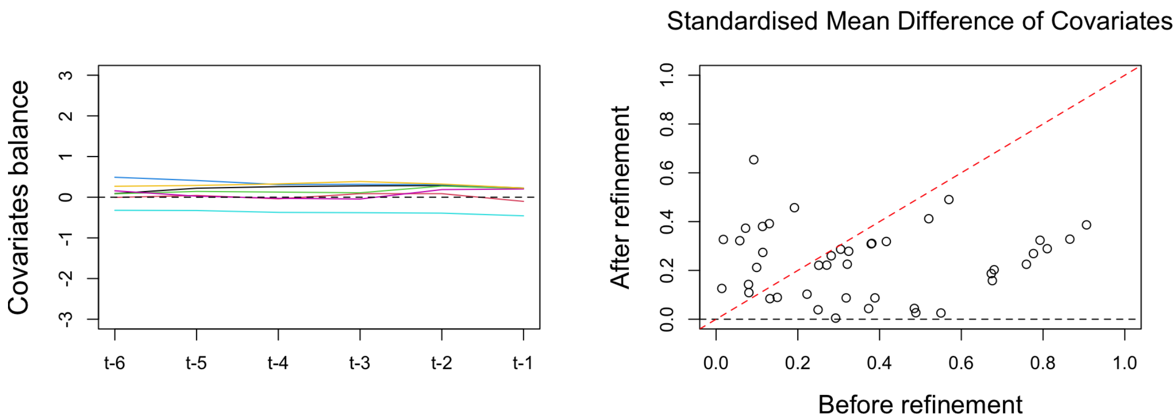
Notes: (i) The figure shows the matching quality for the examination of using as control variables the significant variables identified by the dynamic DiD approach. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.6: Covariates balance for the examination of the announcement effect



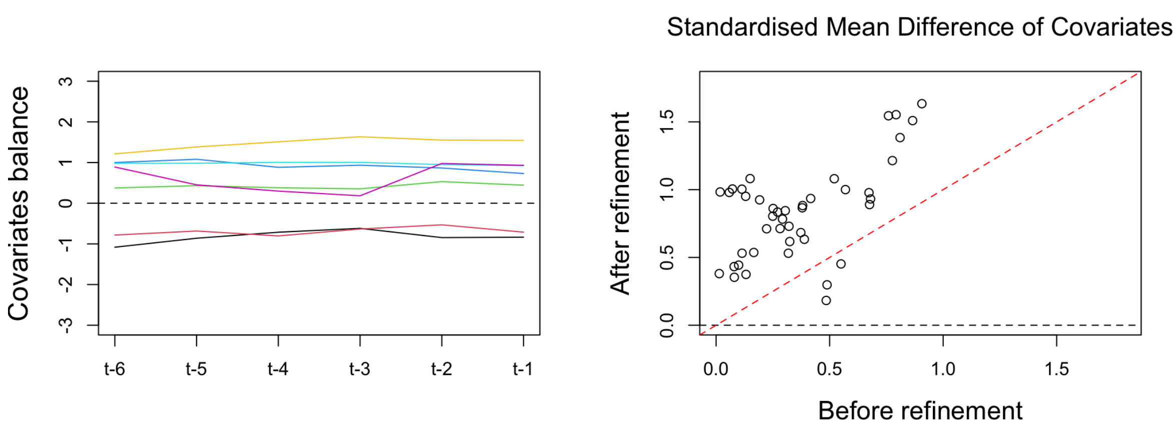
Notes: (i) The figure shows the matching quality for the examination of increasing the match size to two. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.7: Covariates balance for the examination of increasing the match size to two



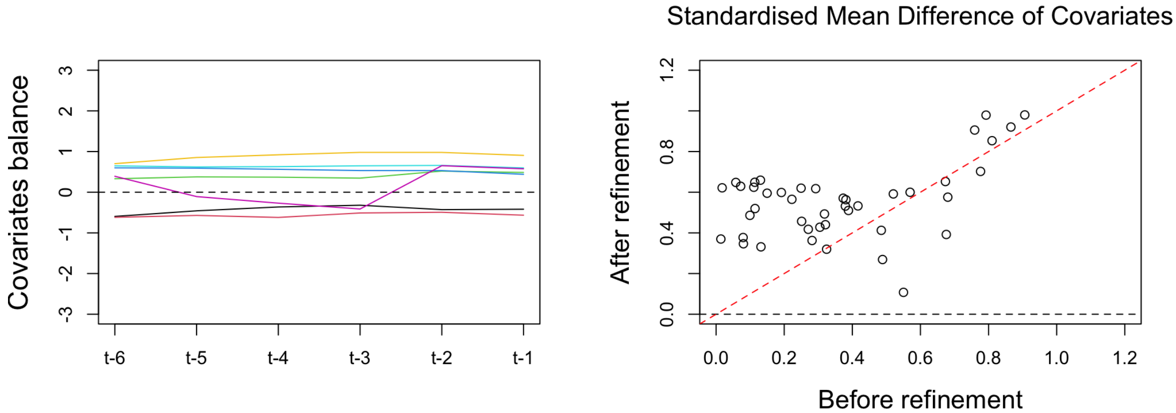
Notes: (i) The figure shows the matching quality for the examination of increasing the match size to three. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.8: Covariates balance for the examination of increasing the match size to three



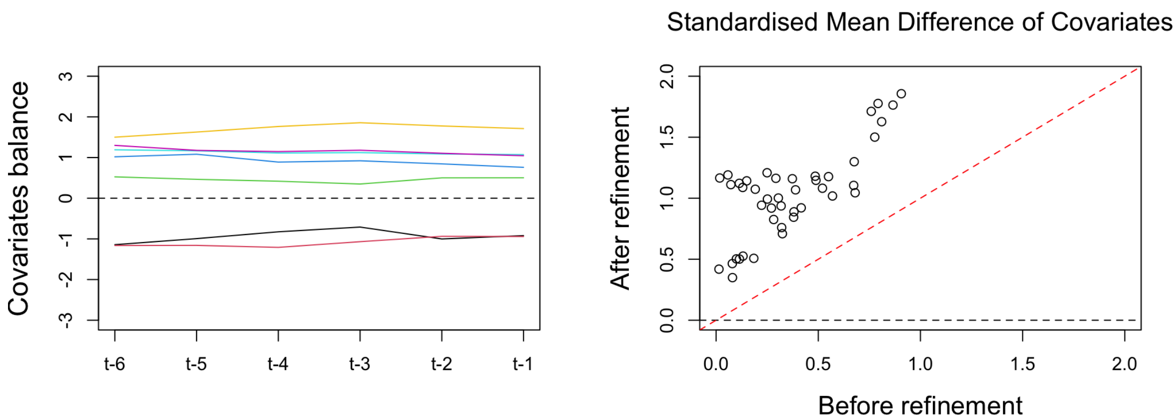
Notes: (i) The figure shows the matching quality for the examination of using propensity score matching. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.9: Covariates balance for the examination of using propensity score matching



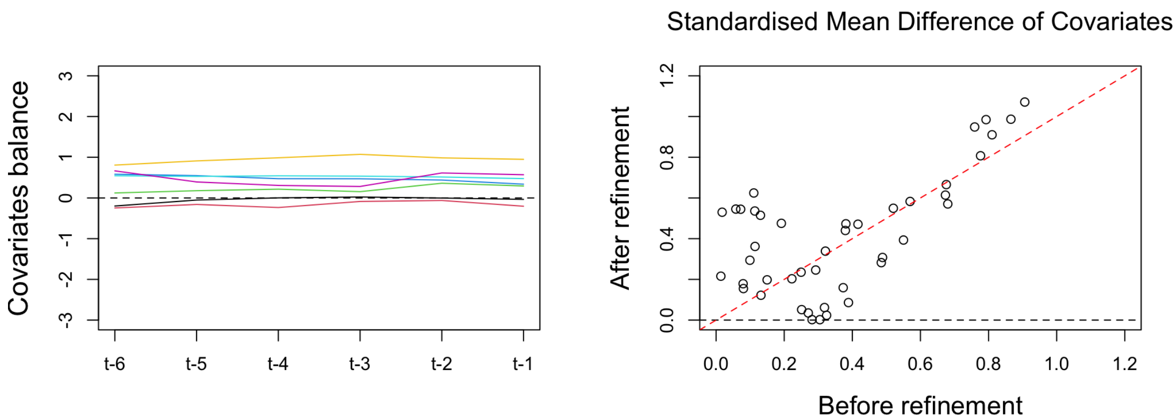
Notes: (i) The figure shows the quality for the examination of using propensity score weighting. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.10: Covariates balance for the examination of using propensity score weighting



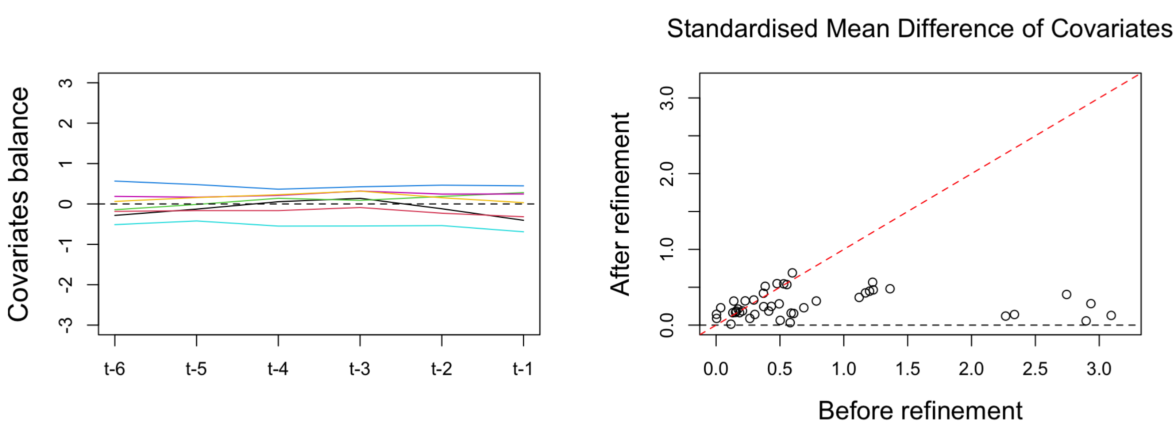
Notes: (i) The figure shows the matching quality for the examination of using covariates balance propensity score matching. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.11: Covariates balance for the examination of using covariates balance propensity score matching



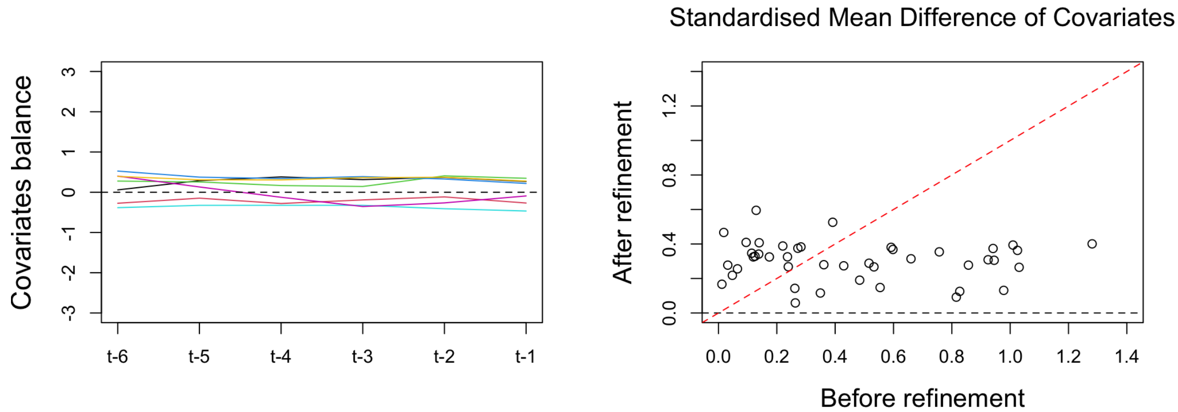
Notes: (i) The figure shows the quality for the examination of using covariates balance propensity score weighting. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.12: Covariates balance for the examination of using covariates balance propensity score weighting



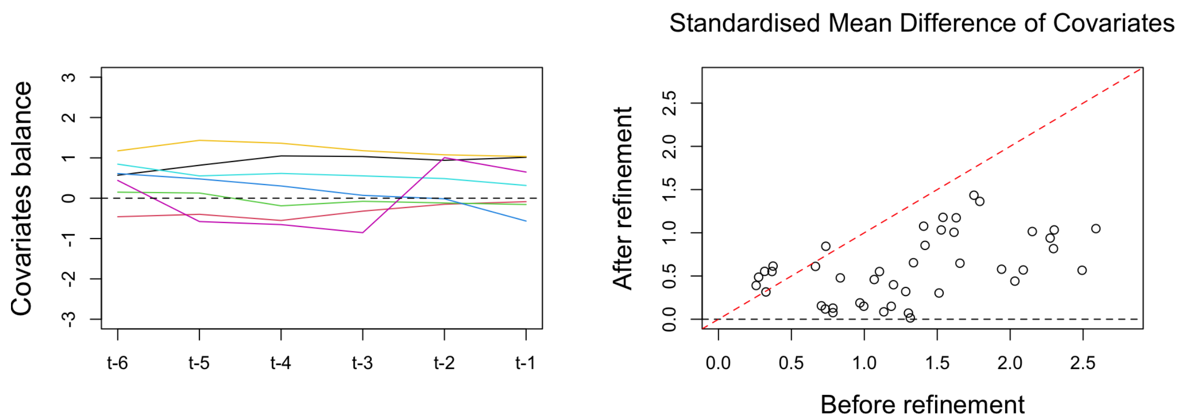
Notes: (i) The figure shows the quality for the heterogeneity analysis for the low-income counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.13: Covariates balance for the heterogeneity analysis for the low-income counties



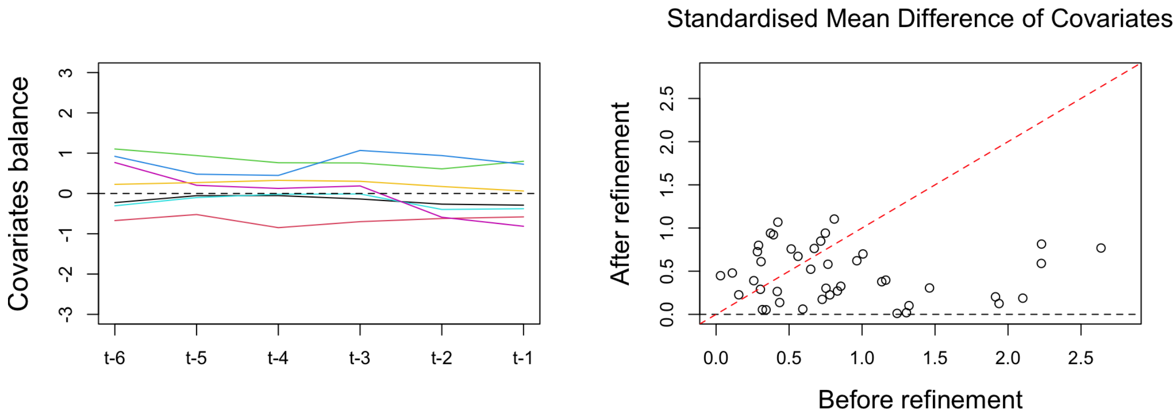
Notes: (i) The figure shows the quality for the heterogeneity analysis for the middle-income counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.14: Covariates balance for the heterogeneity analysis for the middle-income counties



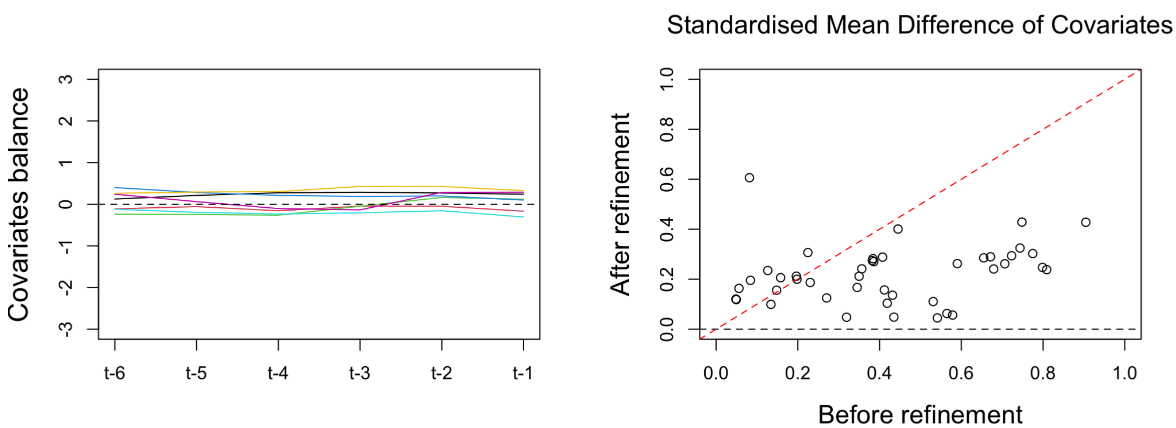
Notes: (i) The figure shows the quality for the heterogeneity analysis for the high-income counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.15: Covariates balance for the heterogeneity analysis for the high-income counties



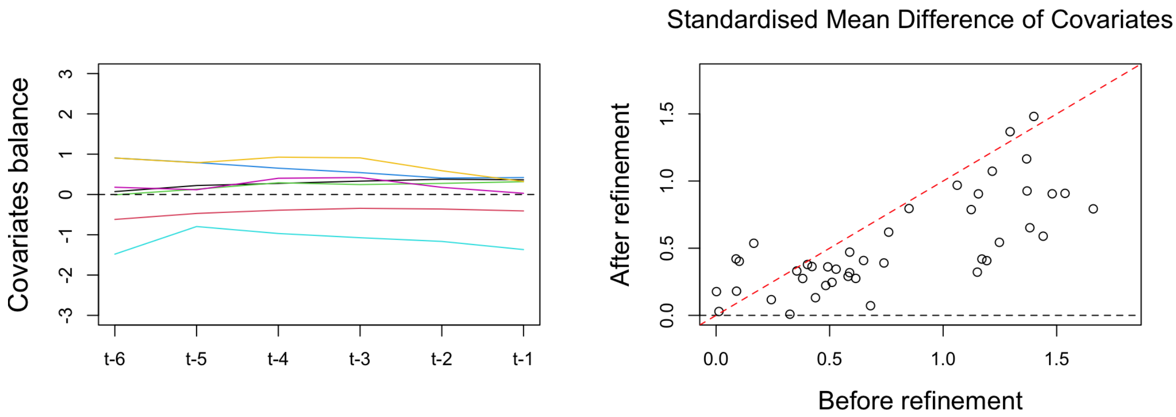
Notes: (i) The figure shows the quality for the heterogeneity analysis for the close counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.16: Covariates balance for the heterogeneity analysis for the close counties



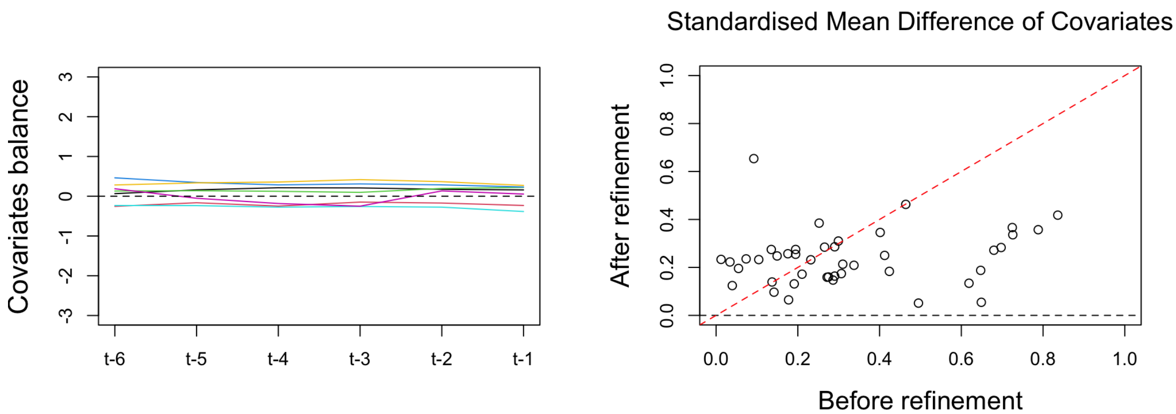
Notes: (i) The figure shows the quality for the heterogeneity analysis for the nearby counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.17: Covariates balance for the heterogeneity analysis for the nearby counties



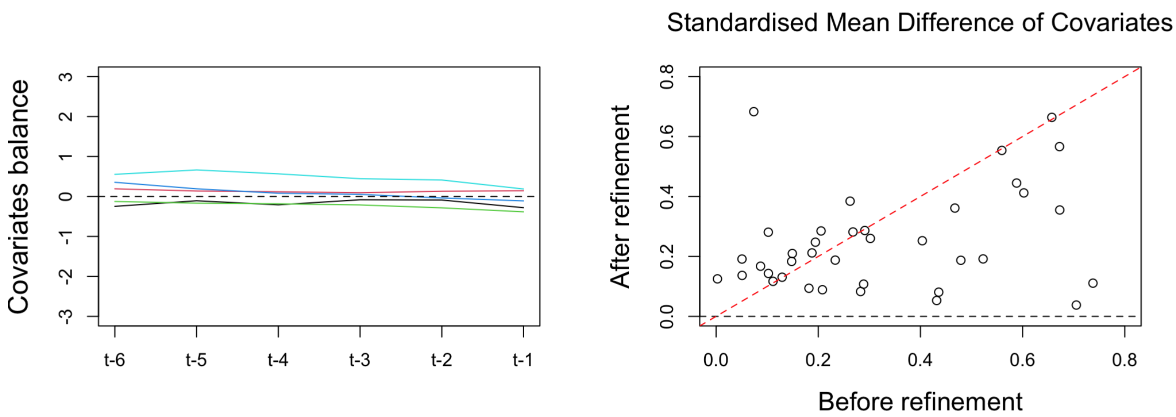
Notes: (i) The figure shows the quality for the heterogeneity analysis for the distant counties. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.18: Covariates balance for the heterogeneity analysis for the distant counties



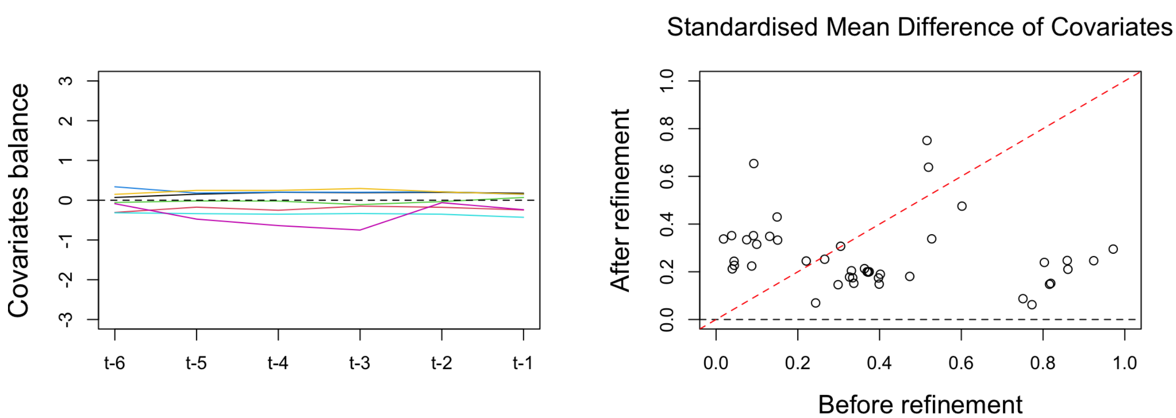
Notes: (i) The figure shows the quality for the results on GDP. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.19: Covariates balance for the results on GDP



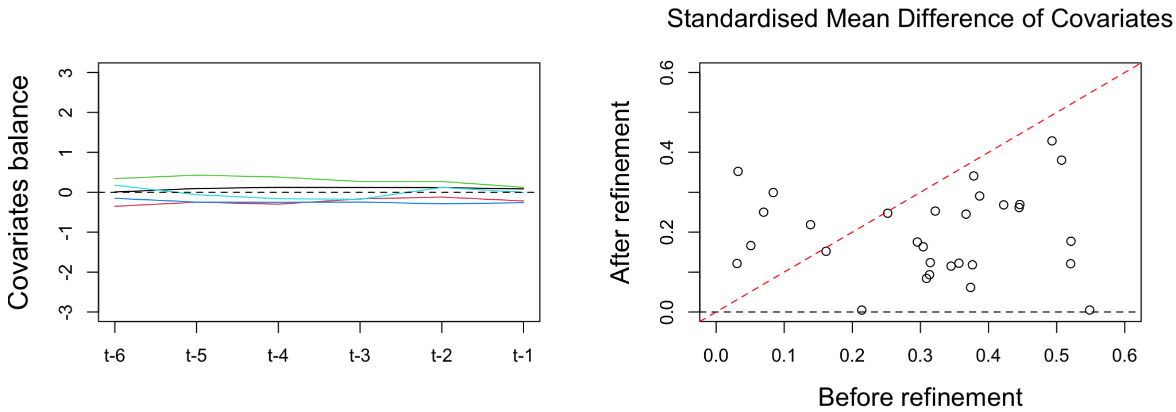
Notes: (i) The figure shows the quality for the results on GDP per capita. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.20: Covariates balance for the results on GDP per capita



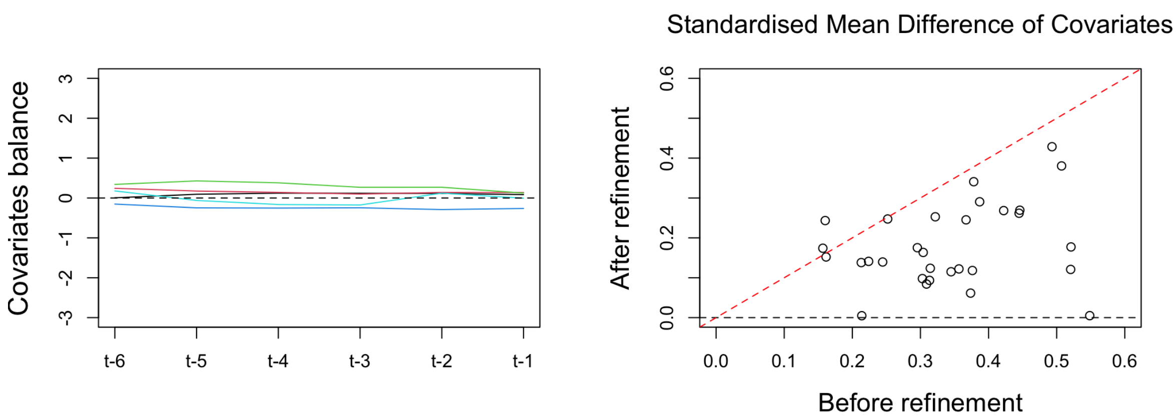
Notes: (i) The figure shows the quality for the results on physical capital. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.21: Covariates balance for the results on physical capital



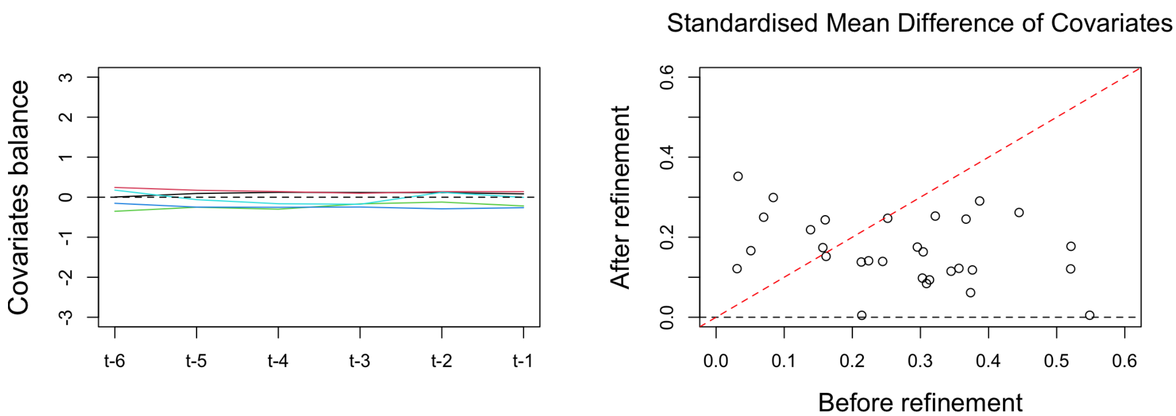
Notes: (i) The figure shows the quality for the results on share of GDP taken by primary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.22: Covariates balance for the results on share of GDP taken by primary sector



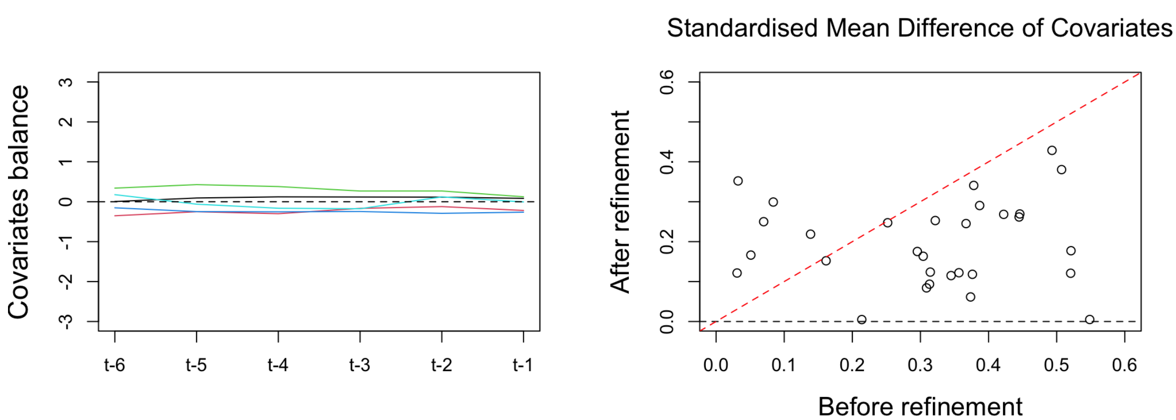
Notes: (i) The figure shows the quality for the results on share of GDP taken by secondary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.23: Covariates balance for the results on share of GDP taken by secondary sector



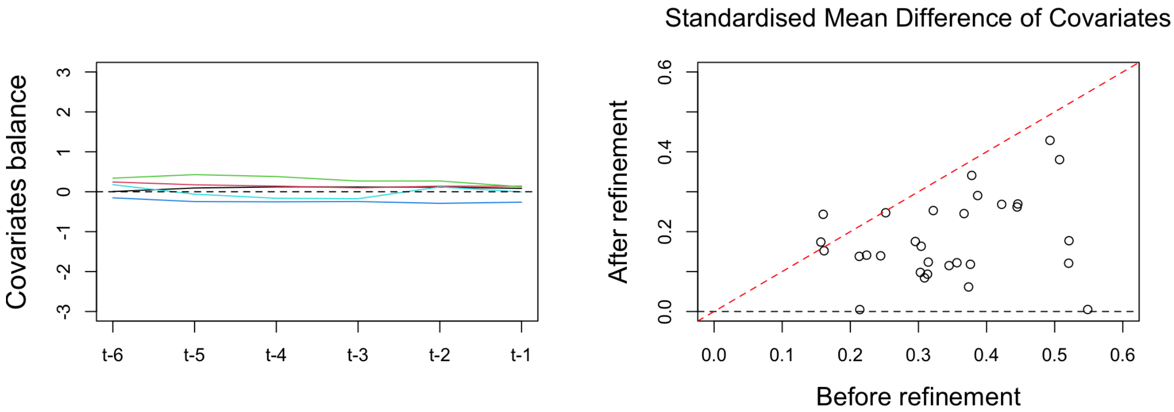
Notes: (i) The figure shows the quality for the results on share of GDP taken by tertiary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.24: Covariates balance for the results on share of GDP taken by tertiary sector



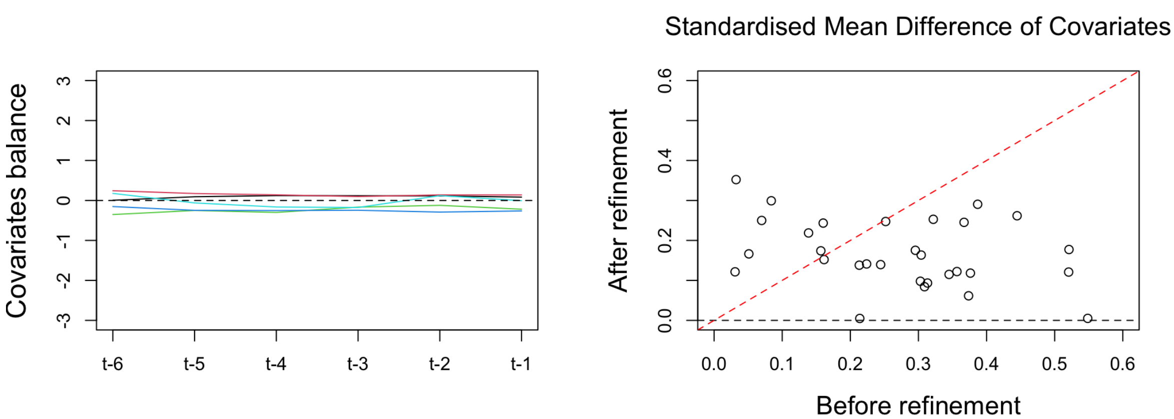
Notes: (i) The figure shows the quality for the results on GDP taken by primary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.25: Covariates balance for the results on GDP taken by primary sector



Notes: (i) The figure shows the quality for the results on GDP taken by secondary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.26: Covariates balance for the results on GDP taken by secondary sector



Notes: (i) The figure shows the quality for the results on GDP taken by tertiary sector. The left panel displays the covariates balance, and the right panel displays the improvement of the balance before and after the refinement. (ii) The mean difference between the treated and their counterfactuals are standardised by the standard deviation.

FIGURE G.27: Covariates balance for the results on GDP taken by tertiary sector