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

Haibo Zhang[†] Corrado Di Maria[†] Bahar Ghezelayagh[†] Yuli Shan[‡]

January 14, 2024

Abstract

In this paper, we assess the effectiveness of early climate policy in emerging economies by causally evaluating the impact of China's 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.

Keywords: *Climate policy, China, Causal identification, Carbon emissions, Carbon intensity*

JEL Codes: *C23, P28, Q54, Q58*

*We thank Rui Xue and the participants at the SURED 2022 Conference in Ascona (CH) and the EAERE 2023 conference in Limassol (CY) for their helpful comments. We are grateful to the Associate Editor, Alexandra Brausmann, and two anonymous reviewers for their insightful and constructive comments during the review process. We also wish to thank David Jones for his generous advice on coding matters. All authors declare that no conflicts of interest or financial interests are involved with this article.

[†]School of Economics, University of East Anglia (UK). Author for correspondence: `c.di-maria@uea.ac.uk`

[‡]School of Geography, Earth and Environmental Sciences (GEES), University of Birmingham (UK).

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 paper, 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

[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 and Zhang (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.

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 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., OECD, 2000; Borck and Coglianese, 2009; André and Valenciano-Salazar, 2022). Third, we take a step forward in the 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

[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.

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, our results present a clear and robust picture whereby the LCCP is shown not to have had any significant impact either in terms of reducing per-capita emissions or carbon intensity of GDP. 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 paper develops as follows, in Section 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 3. Section 4 is devoted to the discussion of the main empirical results, their validity and some robustness checks. Section 5 discusses the potential economic mechanism and the sectoral impacts. Finally, Section 6 summarises and concludes.

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

[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.

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 and representativeness, 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 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]

[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.

[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 ve-

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 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.

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 the SCM to multiple treated units (Dube and Zipperer, 2015; Galiani and Quistorff, 2017; Donohue et al., 2019, among others). Estimating weights that minimise the average pre-treatment imbal-

hicles, 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.

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

ance 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. 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.

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 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

[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).

of the non-LCCP cities.[13]

Table 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.

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

[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.

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 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, 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 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.

4 Empirical results

We begin this section by presenting the results we obtain within the DiD framework discussed in Section 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 and Figure 1 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.

Figure 1 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

[14] The table provides the point estimates, the standard errors and the p-values for the treatment effects, as well as the number of cities in the sample and the total number of observations. In the figure, the treatment effects are normalised relative to the beginning of the corresponding treatment period, i.e. $t = 0$ represents 2010 for the first wave and 2012 for the second one.

Table 2: Staggered difference-in-differences estimation

	ATT estimate	Std. error	p-value	N/Obs.
CO_2 emissions per capita	-0.363***	0.139	0.009	245/3,673
GDP CO_2 intensity	-0.061	0.041	0.140	

Note: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

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 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 (e.g., Baker et al., 2022).

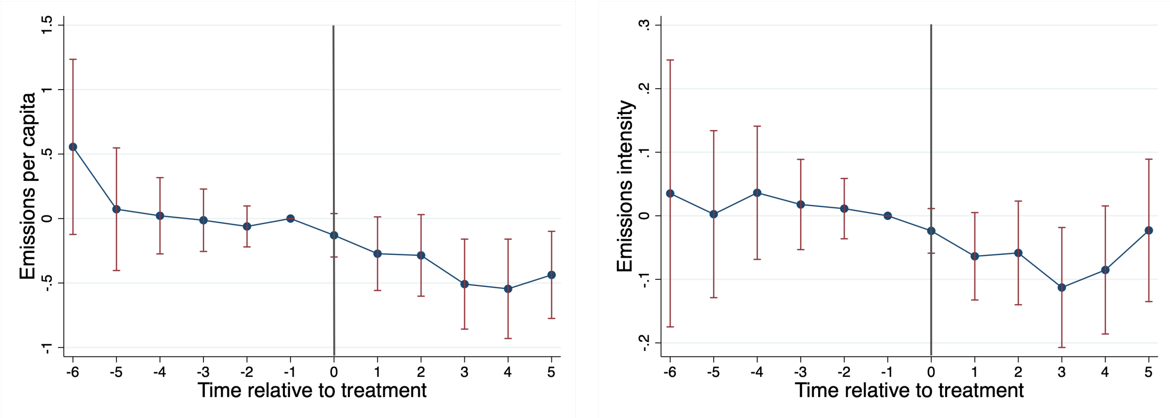
Overall, our assessment of this first set of results is that even if they represent an improvement on the current state-of-the-art, in that they at least address the potential biases in the estimated treatment effects due to the staggered nature of the treatment, they still fall short of providing a convincing identification framework for the causal effects of the LCCP. Indeed, it is clear that – as argued in Section 2 – 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 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.

The results of the partially-pooled, staggered synthetic control procedure run for the first two

[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. 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.

Figure 1: Plots of the staggered difference-in-differences effects



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.

Table 3: Staggered synthetic control estimation – baseline

	ATT estimate	Std. error	*p*-value	N/Obs.
CO_2 emissions per capita	-0.148	0.174	0.395	245/3,673
GDP CO_2 intensity	-0.065	0.077	0.399	

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

waves of the LCCP are presented in Table 3.[17] These results show that the LCCP had no statistically significant effect at conventional levels on the treated cities, relative to the non-treated ones. Figure 2 plots the estimates of the effects over time. Overall, the pre-treatment fits are satisfactory and, based on the confidence intervals plotted, the treatment effects on both measures remain insignificant throughout the treatment period.

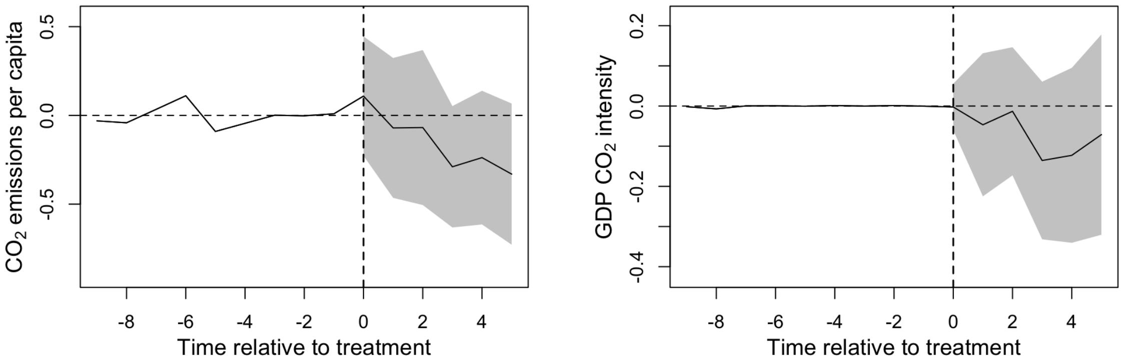
These findings are clearly at odds with the ones currently available in the literature, so in the remainder of this section, we delve deeper into the data to shore up our confidence that these results are indeed correct and robust.

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 introduc-

[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 $j = 2, ..., D$ donor units following the main specification, using the remaining $D - 2$ 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.

Figure 2: Plots of the Staggered synthetic control effects – baseline



Notes: The figures show 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.

tion 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] This poses two parallel challenges for us since, on the one hand, cities in the donor pool might be affected by the pilots leading to a potential attenuation bias; on the other hand, some of the LCCP cities might have been also included in these ETS pilot schemes, 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.

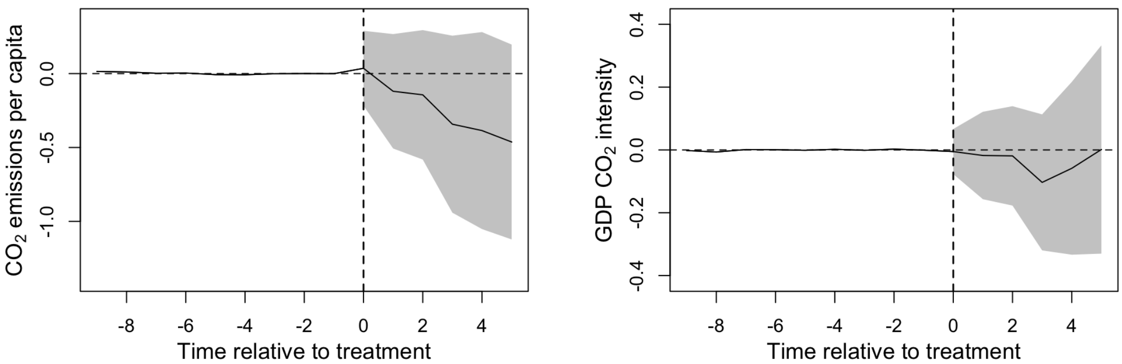
Figure 3 and Table 4 report the results of this exercise. The effect of excluding the cities treated by the ETS pilots is relatively small. Compared to the baseline discussed in Table 3, the changes in the estimated coefficients are small and they remain insignificant, with the p-value for the carbon 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 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.

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

[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).

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.

Figure 3: Staggered synthetic control estimation – controlling for policy overlap



Notes: The figures show 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.

Table 4: Controlling for policy overlap – excluding ETS cities

	ATT estimate	Std. error	*p*-value	N/Obs.
Panel A: Staggered synthetic control				
CO_2 emissions per capita	-0.236	0.230	0.304	214/3,210
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	214/3,210
GDP CO_2 intensity	-0.081	0.055	0.138	

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

As discussed in Section 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 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 5: Testing differences in reduction targets under the 12th and 13th Five-Year Plans

	Treated units Mean	Donor units Mean	Diff.	t-statistic	p-value	N/Obs.
12th Five-Year Plan (2011-15)	17.04	17.08	-0.04	-0.18	0.86	214/3,210
13th Five-Year Plan (2016-20)	18.54	18.88	-0.35	-0.96	0.34	

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

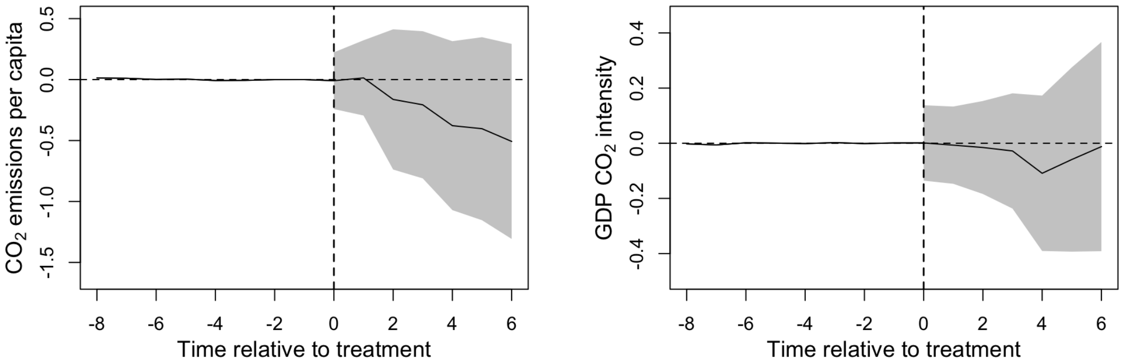
Another possible limitation of our identification strategy is that the LCCP was introduced following an earlier announcement and the selection process of suitable pilot candidates was also rather slow. From this point of view, our choice to start the treatment period from the official inception dates 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]

Figure 4 and Panel A of Table 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

[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. The full set of results is available from the authors upon request.

[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. Full details are available from the authors upon request.

Figure 4: Controlling for potential anticipation effect – alternative policy start



Notes: The figures show the results of the staggered synthetic control method on per-capita CO_2 emissions and GDP CO_2 intensity anticipating the treatment effect by one year. The effects are normalised relative to the beginning of treatment, i.e. 2009 for wave I and 2011 for Wave II.

outcomes compared to Table 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 cities treated under the ETS pilots. The results are reported in Panel B of Table 6 and in the Appendix (see Figure C.1). Again, we find no indication that an anticipation effect might have taken place.

Table 6: Controlling for potential anticipation effect – alternative policy start

	ATT estimate	Std. error	*p*-value	N/Obs.
Panel A: Baseline sample				
CO_2 emissions per capita	0.030	0.329	0.927	245/3,673
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	214/3,210
GDP CO_2 intensity	-0.033	0.110	0.766	

Notes: 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 anticipating the treatment effect by one year. *, **, *** 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.

Table 7: Controlling for treatment spillovers

	ATT estimate	Std. error	p-value	N/Obs.
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	136/2,040
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	129/1,935
GDP per capita	0.062	0.142	0.660	

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

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 re-run 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.

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

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

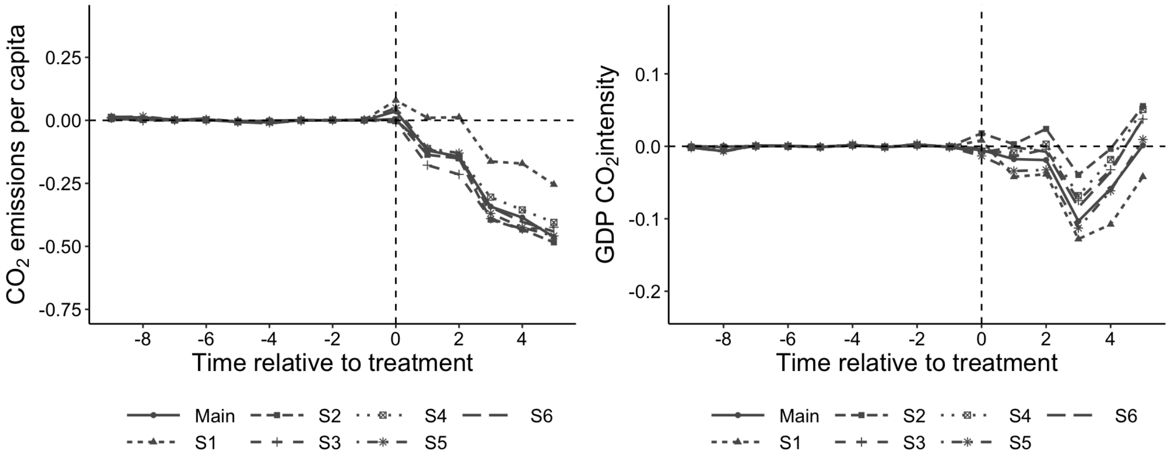
[21] Specifically, we drop all control units that share a border with a treated city.

[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.

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 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.

Figure 5: Robustness checks – changing predictor sets



Notes: The figures show 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.

Table 8 and Figure 5 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 3. We replicate our previous analysis separately for each wave. Panel A and B in Table 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. 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 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

Table 8: Robustness checks – changing predictor sets

	Baseline	S1	S2	S3	S4	S5	S6	N/Obs.
CO_2 emissions per capita	-0.236	-0.081	-0.266	-0.274	-0.221	-0.241	-0.241	214/3,210
	(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: 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, for different predictor sets. *, **, *** 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.

Assumption (SUTVA).

Table 9: Checking for robustness — results for individual waves

	ATT estimate	Std. error	*p*-value	N/Obs.
Panel A: LCCP first wave				
CO_2 emissions per capita	-0.274	0.340	0.420	192/2,880
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	183/2,745
GDP CO_2 intensity	0.010	0.114	0.930	

Notes: 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. Panel A shows the results of the first wave; panel B shows the results of the second wave. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

As a further test for the robustness of our results, we now distinguish between cities that were assigned to treatment directly (which we refer to as city-level treatment) versus cities that were 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.[23] Table 10 reports on the outcome of this test, showing that the treatment effect is insignificant, irrespective of the level of their assignment into treatment.

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

[23]We thank one anonymous reviewer for suggesting this additional test.

Table 10: Checking for robustness – heterogeneous treatment levels

	ATT estimate	Std. error	*p*-value	N/Obs.
Panel A: City-level treatment				
CO_2 emissions per capita	-0.410	0.316	0.195	187/2,805
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	188/2,820
GDP CO_2 intensity	-0.036	0.124	0.769	

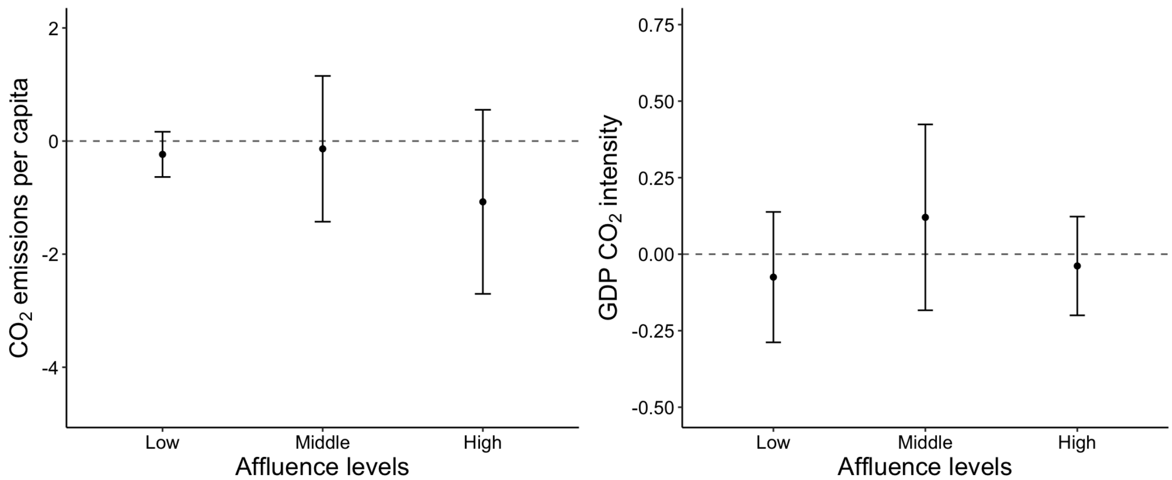
Notes: The table displays the estimates of the staggered synthetic control method on CO_2 emissions per capita and GDP CO_2 intensity for the first two waves of the LCCP. Panel A shows the results of the city-level treatment; panel B shows the results of the province-level treatment. *, **, *** indicate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

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 B.2 in Appendix B 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 6 – the estimates can be found in Panel A of Table B.1 in Appendix B – reports the results of this exercise. Once again, the treatment effects remain clearly insignificant across all groups for both outcome variables although the precision of the estimates varies greatly.

Figure 6: Checking for robustness - cities with different affluence levels

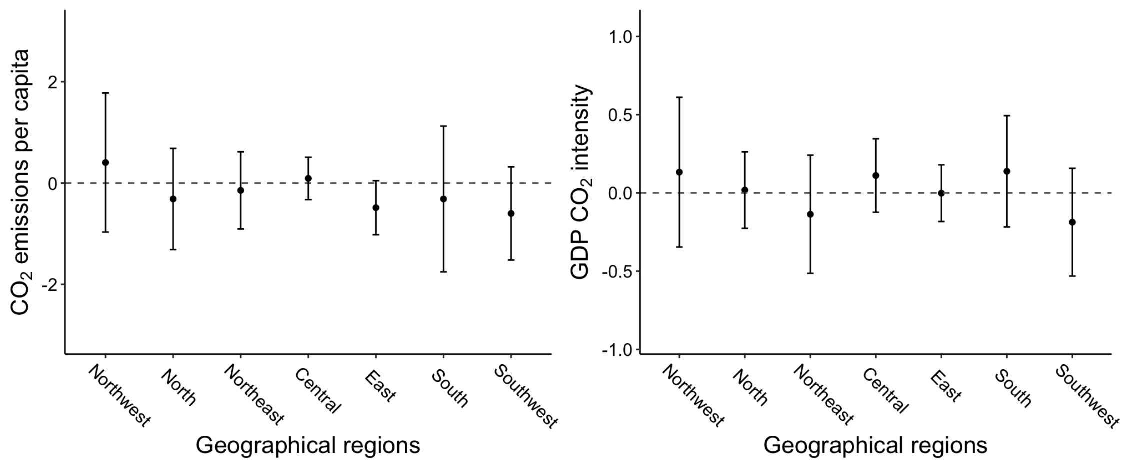


Notes: The figures show the results of differential effects in cities with different affluence levels on CO_2 emissions per capita and GDP CO_2 intensity using the staggered synthetic control method.

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.

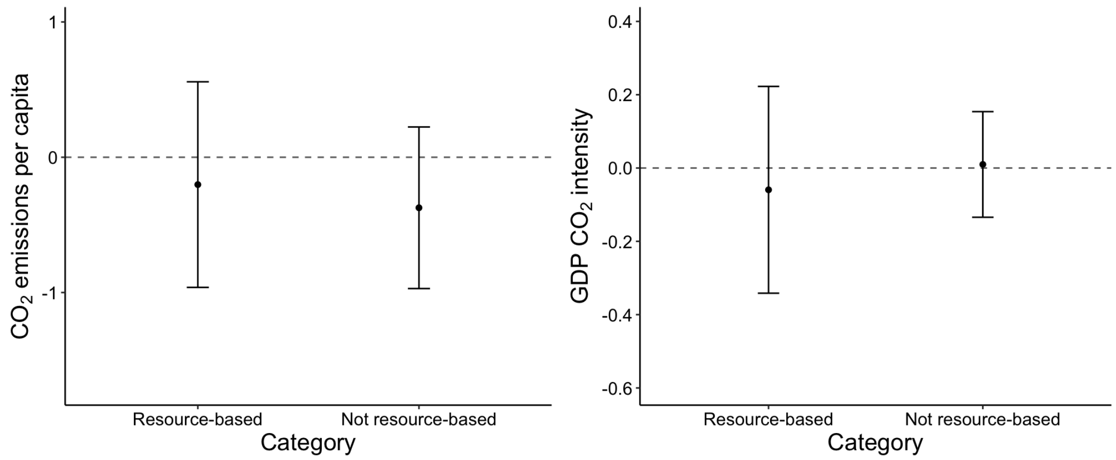
Figure 7 – and Panel B of Table B.1 in Appendix B – reports the results of this analysis. Also in this case, the results suggest that the LCCP had no significant effect on carbon intensity.

Figure 7: Checking for robustness - cities in different geographical regions



Notes: The figures show 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 8: Checking for robustness - resource-based and non-resource-based cities



Notes: The figures show 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.

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 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.[24] We perform SCMs separately for each group.

Figure 8 and Panel C of Table B.1 present the results. The estimated treatment effects are clearly insignificant for both categories and outcomes, again suggesting that the LCCP had no 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 (see Shan et al., 2018a,b, 2020, for a discussion of these issues). 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.[25] 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 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 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 pilots or whose predictors do not fall in the convex hull.[26] Figure 9 and Table 12 present the results.

[24] 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.

[25] See Appendix A for a discussion of how this inventory is constructed.

[26] We exclude 25 ETS-regulated cities and two outliers (Suzhou and Qingdao) from the sample. After the exclusion, we

Table 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.

Table 12: Checking for robustness – alternative emissions data

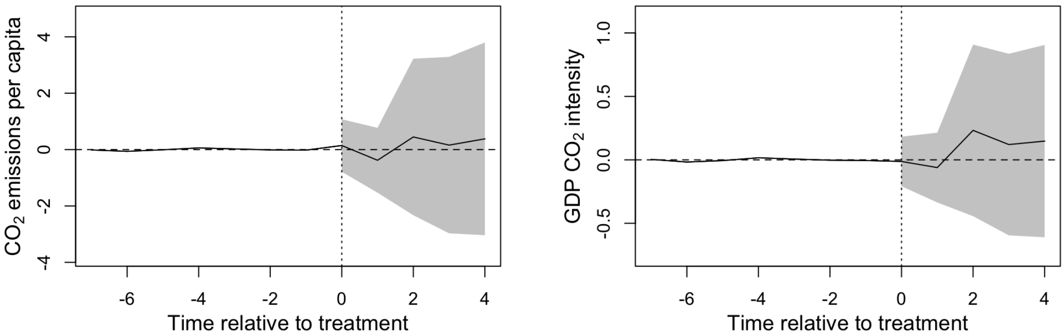
	ATT estimate	Std. error	p-value	N/Obs.
Panel A: Staggered estimation				
CO_2 emissions per capita	0.151	1.076	0.888	80/960
GDP CO_2 intensity	0.085	0.240	0.722	
Panel B: LCCP first wave				
CO_2 emissions per capita	-0.992	0.958	0.301	68/816
GDP CO_2 intensity	-0.103	0.237	0.665	
Panel C: LCCP second wave				
CO_2 emissions per capita	1.005	1.939	0.604	74/888
GDP CO_2 intensity	0.262	0.383	0.494	

Notes: 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 Low Carbon-City Pilot. The outcome variables are calculated based on the IPCC Guidelines using city-level statistics on energy use. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Despite the change in data, the results are consistent with those in Figure 3 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.

have 18 treated units and 77 donor units left.

Figure 9: Checking for robustness – alternative emissions data



Notes: The figures show 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.

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 paper 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.

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 'Investment in Science and Technology' and 'Investment in Social Fixed Capital' as proxies for the type of investment activities discussed above.[27] Table 13 reports the results of our staggered synthetic control estimations. We find no evidence that LCCP cities are investing more than other cities, at least at this level of aggregation.

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

[27] In the analysis that follows, we exclude four of the treated cities (Hangzhou, Ningbo, Qingdao, and Xiamen) from the evaluation of the impact of the LCCP on the investment in science and technology because we are unable to construct acceptable counterfactuals given our donor pool.

Table 13: Investment in Science and Technology and Social Fixed Capital

	ATT estimate	Std. error	*p*-value	N/Obs.
Investment in Science and Technology	0.005	0.032	0.880	210/3,150
Social fixed asset investment	0.551	0.502	0.272	214/3,210

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

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 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.

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 B.3 and Table B.4 in Appendix B 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.[28]

Table 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 significant 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 (Khanna et al., 2014, See also). Figure 10 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.[29]

Our next step is to focus on the cities whose agendas we used to construct Figure 10. Using this (admittedly small) set of cities, we once again drill down to the sector/fuel level. Table 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 10.

[28]For each broader sector in each year, we winsorize the observations' first differences to control for the outliers due to the potentially mis-reported data on energy consumption.

[29]As discussed in the footnote in Section 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 12^{th} FYP.

Table 14: Sectoral analysis by fuel type

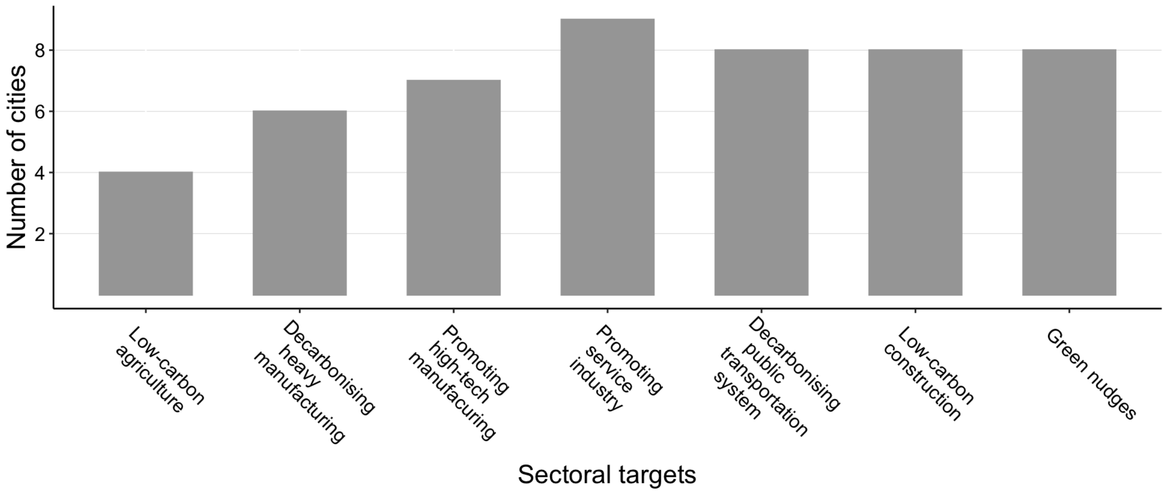
Sectors	CO_2 emissions				N/Obs.
	Total	Coal products	Gas	Oil products	
Agriculture	-0.060 (0.044)	-0.021 (0.039)	0.000 (0.001)	0.011 (0.030)	
Mining	-0.673 (0.656)	-0.790 (0.622)	0.026 (0.019)	-0.006 (0.009)	
Light manufacturing	-0.161 (0.182)	-0.078 (0.145)	-0.002 (0.036)	-0.008 (0.012)	
Heavy manufacturing	-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)	80/960
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: 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 16 show that, neither in the full sample nor among the cities

Figure 10: Distribution of cities by LCCP sectoral-level target.



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

Table 15: Sectoral analyses by fuel type

Sectors	CO_2 emissions				N/Obs.
	Total	Coal products	Gas	Oil products	
Agriculture	-0.024 (0.070)	0.022 (0.052)	0.000 (0.001)	0.031 (0.052)	66/792
Heavy Manufacturing	3.135 (2.705)	2.197 (1.928)	0.606 (0.590)	-0.220 (0.360)	69/828
High-tech Manufacturing	0.046 (0.062)	0.046 (0.057)	0.007 (0.011)	0.001 (0.007)	70/840
Construction	0.023 (0.069)	-0.005 (0.019)	-0.002 (0.002)	0.027 (0.026)	69/828
Transportation	-0.383* (0.226)	-0.020 (0.039)	0.005 (0.031)	-0.321 (0.286)	70/840
Service sector	0.100 (0.274)	0.012 (0.257)	0.060** (0.029)	0.064 (0.073)	71/852
Household usage	0.308 (0.249)	0.071 (0.167)	0.023 (0.060)	-0.026 (0.084)	70/840

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.

that have published online agendas, we can find any evidence that the LCCP made any difference to the treated cities.[30]

[30] As previously mentioned, we exclude treated cities for which we are unable to find appropriate counterfactuals. In Panel A, two cities (Hangzhou and Qingdao) are excluded when evaluating the impact of the LCCP on GDP, and one city

Table 16: Staggered synthetic control estimation – other outcomes

	ATT estimate	Std. error	p-value	N/Obs.
Panel A: Full sample				
GDP	0.641	3.558	0.857	212/3,180
GDP per capita	0.134	0.129	0.298	214/3,210
Employment	-0.039	0.056	0.488	213/3,195
Panel B: Cities with published agendas				
GDP	-0.936	4.738	0.843	
GDP per capita	0.156	0.146	0.285	190/2,850
Employment	-0.047	0.053	0.375	

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

6 Concluding remarks

In this paper, 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, instead, that the LCCP did not lead to a reduction in carbon emissions per capita, nor did it have a significant impact on the carbon intensity of GDP.

Our results contrast with the existing literature on the LCCP that emphasises a small but generally positive impact of the pilot.[31] 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

(Hangzhou) is excluded when assessing the impact on employment.

[31] 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).

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 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.

References

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

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.

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].

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.

CEADS (2020). County-level CO_2 emissions in China during 1997–2017. `https://ceads.net/` [Last accessed: March 10, 2023].

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.

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.

de Chaisemartin, C. and D'Haultfoeuille, X. (2022). Difference-in-differences estimators of intertemporal treatment effects. Working Paper 29873, National Bureau of Economic Research.

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. *IZA Discussion Paper No. 8944*, pages 1–58.

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.

Feng, T., Lin, Z., Du, H., Qiu, Y., and Zuo, J. (2021). Does low-carbon pilot city program reduce carbon intensity? evidence from chinese cities. *Research in International Business and Finance*, 58:101450.

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.

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.

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.

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

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.

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.

Liu, Z., Guan, D., Wei, W., Davis, S. J., Ciais, P., Bai, J., Peng, S., Zhang, Q., Hubacek, K., Marland, G., Andres, R. J., Crawford-Brown, D., Lin, J., Zhao, H., Hong, C., Boden, T. A., Feng, K., Peters, G. P., Xi, F., Liu, J., Li, Y., Zhao, Y., Zeng, N., and He, K. (2015). Reduced carbon emission estimates from fossil fuel combustion and cement production in China. *Nature*, 524(7565):335–338.

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.

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.

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: March 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.

OECD (2000). *Voluntary Approaches for Environmental Policy*. OECD, Paris.

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.

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].

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

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.

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.

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, 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. and Zhang, N. (2021). Low-carbon city pilot and carbon emission efficiency: Quasi-experimental evidence from China. *Energy Economics*, 96:105125.

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.

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.

A 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:

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

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:

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

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 A.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 A.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.

B Additional tables

Table B.1: Staggered synthetic control estimation - excluding ETS pilot cities

	CO_2 emissions per capita	GDP CO_2 intensity	N/Obs.
Panel A. Different affluence levels			
Low-income cities	-0.236 (0.204)	-0.075 (0.109)	195/2,925
Middle-income cities	-0.138 (0.658)	0.120 (0.155)	178/2,670
High-income cities	-1.075 (0.830)	-0.039 (0.082)	167/2,505
Panel B. Different geographical regions			
Northwest China	0.405 (0.700)	0.133 (0.244)	173/2,595
North China	-0.314 (0.510)	0.018 (0.125)	168/2,520
Northeast China	-0.145 (0.389)	-0.136 (0.193)	176/2,640
Central China	0.091 (0.213)	0.111 (0.120)	166/2,490
East China	-0.488* (0.272)	-0,002 (0.092)	169/2,535
South China	-0.315 (0.734)	0.138 (0.181)	167/2,505
Southwest China	-0.601 (0.470)	-0.187 (0.176)	173/2,595
Panel C. City category			
Resource-based cities	-0.202 (0.388)	-0.059 (0.144)	190/2,850
Non-resource-based cities	-0.374 (0.305)	0.010 (0.074)	187/2,805

Notes: 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. *, **, *** indicate 10%, 5% and 1% statistical significance, respectively.

Table B.2: Distribution of LCCP cities by affluence levels and geography

	China's Low-Carbon City Pilot	
	The first wave	The 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 B.3: 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	Energy 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	Transportation
44	Wholesale, Retail Trade and Catering Services	Service Sector
45	Other Service Sectors	
46	Urban Resident Energy Usage	Household Usage
47	Rural Resident Energy Usage	

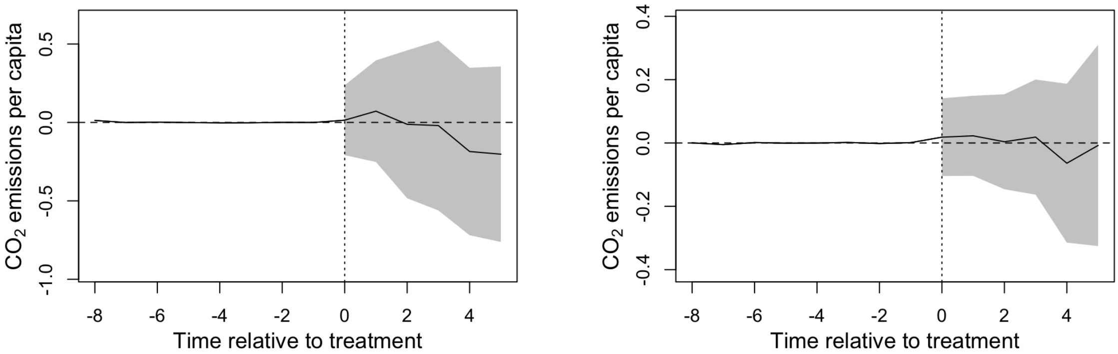
Notes: The table shows the economic sectors and categorisation. In general, we categorise 47 economic sectors into nine broader categories, partly following the suggestion in Shan et al. (2018b).

Table B.4: 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.

C Additional figures



Notes: The figures show 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. Effects are normalised relative to the beginning of treatment, i.e. 2010 for wave I and 2012 for Wave II.

Figure C.1: Staggered synthetic control estimation - excluding ETS pilot cities