Emissions Trading Scheme and Directed Technological Change: Evidence from China

Ruijie Tian*

October 29, 2020

Job Market Paper

Latest version available at this link

Abstract

This paper examines the impact of carbon emissions trading schemes (ETS) on technical change proxied by the number of green patents in the context of the pilot ETS in China. I find a small increase of 0.16 patents per firm and year. A 10 percent increase in carbon prices increases green patents by 2 percent. The strongest effects are for the two regions in the upper range of carbon prices and for more productive firms. However, there are contrasting patterns at the extensive and intensive margins of green innovation: the pilot ETS reduces entry into green innovative activities but increases levels of innovating for firms that were innovative before they were regulated by ETS, especially for the more productive firms. This indicates that an important policy challenge is to encourage the firms covered by ETS to start innovation in green technologies; this applies particularly to the larger and more productive firms.

JEL Classification: Q54, Q55, O44, O33

Keywords: Carbon Pricing; Directed Technological Change; Innovation; Heterogeneous Firms.

*Department of Economics, University of Gothenburg. Vasagatan 1, 40530 Gothenburg, Sweden. E-mail: ruijie.tian@economics.gu.se. I am indebted to Inge van den Bijgaart, Thomas Sterner and Aico van Vuuren for their invaluable guidance and encouraging support. I am deeply grateful to Antoine Dechezleprêtre for an insightful discussion on an earlier version of this paper. I thank Jintao Xu for providing data. I would also like to thank Fredrik Carlsson, Li Chen, Håkan Eggert, Randi Hjalmarsson, Jimmy Karlsson, Elin Lokrantz, Samson Mukanjari, Sugandha Srivastav, Thomas Stoerk, Chiman Yip and seminar participants at Helsinki GSE, Stockholm University, University of Gothenburg, the 2019 EAERE Conference, the 26th Ulvön Conference, the 3rd NAERE Workshop, and the 2020 EAERE-ETH Winter School for helpful comments.

1 Introduction

The past decade witnessed a take-off of large-scale CO2 emissions reduction policies, including emissions trading schemes (ETS) that started to play a promising role in combating climate change.[1] One of the most notable ETS developments in recent years has been the implementation of pilot schemes in China. These schemes currently cover 11 percent of Chinese CO2 emissions. It is expected that the Chinese pilot schemes will be integrated into a nation-wide emissions trading scheme in the future. An integrated scheme would cover more than a third of Chinese emissions (about 10 percent of global carbon emissions), making it the largest ETS globally. The effect of an ETS is to put a price on carbon emissions, with the purpose of achieving environmental goals in an efficient manner. The introduction of an emission price provides a continuous incentive for adoption and innovation of emission-reducing technologies (Baranzini et al., 2017). [2] In this paper, I empirically identify the causal effect of emission pricing on innovation in the context of the Chinese emissions trading pilots. I construct a unique Chinese firm-level panel dataset, using yearly patent counts as a measure of innovation. The dataset contains detailed information on firm characteristics, including patent activity and regulatory status (whether or not the firm is covered by ETS).

The empirical identification of the ETS effect on innovation is based on a differences-in-differences estimation, using a zero-inflated Poisson model. The sources of variation are the years of implementation of the pilot ETS in different pilot regions with both regulated firms and non-regulated firms in each region. Ideally, one would either compare firms that are identical in all aspects except for treatment status (being regulated or not), or exploit a random assignment of the treatment to firms. However, in the Chinese pilot ETS, only firms with yearly carbon emissions above a certain threshold are regulated. Hence, estimates from simply comparing the patent counts between treated and control firms before and after the

[1]The European Union ETS (EU ETS), set up in 2005, is the world's first carbon emissions trading system and currently operates in 28 EU member states, plus Iceland, Liechtenstein and Norway. Subsequently, ETS have been established in California and 10 states in the US (RGGI), with further implementation scheduled in Japan and more states in the US, among others.

[2]For the literature on the the role of environmental regulation in firm innovation, see e.g., Fischer et al. (2003), Biglaiser and Horowitz (1994) Requate and Unold (2003), Di Maria and Smulders (2017) and Requate (2005).

implementation of the regulation would be biased. I address this issue by matching regulated firms with non-regulated firms on a vector of pre-treatment variables, such that firms in the two groups are balanced on the observable variables.

Applying my estimation strategy to the data, I find a statistically significant effect of the pilot ETS on *green patenting*. I show that the pilot ETS increased the firm average annual number of green patents by 0.16. This increase amounts to 11.7 percent of the yearly average green patents in the pre-treatment period (2007-2012) and 2.8 percent in the post-treatment period (2013-2016). In addition, I estimate the carbon price elasticity: a 10 percent increase in carbon price increases green patents by 2.3 percent. I find no evidence that this increase leads to crowding out of non-green patents. I then show that the effects are heterogeneous across both pilot regions and firms, with the strongest effects for the two regions that have some of the highest carbon prices (Beijing and Shanghai) and, at the intensive margin, for the firms that are at the higher end of worker productivity and thus are initially more competitive.

This paper contributes to the literature that analyzes the impact of environmental policies on innovation. The three papers most closely related to this study are Calel and Dechezleprêtre (2016), Zhu et al. (2019) and Cui et al. (2018).[3] Calel and Dechezleprêtre (2016) evaluate the causal effect of the EU ETS on low-carbon innovation, proxied by the number of patents filed by firms. They use a matched differences-in-differences estimator, and find a small but positive effect of the EU ETS on firms' innovation. Further, Zhu et al. (2019) and Cui et al. (2018) study the impact of the pilot ETS on innovation in China. They both find increases in green patenting induced by the pilot ETS.

This paper extends the literature in four principal ways. The first is the focus on heterogeneity across firms and pilot regions, unlike previous studies, which have estimated the average treatment effects of carbon pricing on green innovation. My analysis of heterogeneity provides new evidence on what might be driving the significant effects found in previous studies. I show that the effectiveness of the pilot ETS differs across the pilot regions. A possible explanation is the regional differences in the policy design, such as allowance allocation,

[3]Other related empirical studies evaluate impacts of ETS on firms' investment strategy and carbon leakage (aus dem Moore et al., 2019, Fell and Maniloff, 2018), productivity and competitiveness (Bushnell et al., 2013, Chan et al., 2013) and emission abatement (Anderson and Di Maria, 2011, Petrick and Wagner, 2014)

coverage threshold, sectors regulated, and costs of non-compliance; these lead to substantially different emission prices across the regions. I also find that the increase in green innovation is primarily driven by intensive margin decisions by regulated firms that already have high output per worker (and therefore higher productivity and/or more capital). This provides evidence on characteristics of firms that may make them more likely to respond to ETS with green innovation.

Second, I estimate carbon price elasticity for green patents as an indicator of the continuous incentives for innovation. The pilot ETS in China is an ideal setting to estimate this because of the substantial variation in carbon prices. The various pilot schemes provide considerable heterogeneity across regions because of the decentralized manner in which they were introduced: each local government designs its own rules. (See Section 2.)

The third contribution is a more precise measure of the outcome variable - the number of green patents - which has the advantage of reducing potential measurement error. The policy effect is more precisely estimated in this study, compared to the two earlier studies on the Chinese pilot ETS effect on green innovation, because I only focus on the type of patents that are more valuable (invention patents)[4] and the patents that are directly impacted by the regulation (low-carbon patents). The patents in the invention category[5] need to pass through a thorough examination for novelty, and therefore are more likely to be radical innovations. I also exclude from the sample all patents that are either carbon-intensive, such as technologies for gas-turbine plants and cremation furnaces, or not directly related to low-carbon innovation, such as innovation in agricultural technologies.

Lastly, this paper separately identifies the effects of the ETS on green innovation at the extensive and intensive margins, i.e., both the likelihood of entry into green innovation and the amount of such innovation. I find contrasting patterns at the two margins: the pilot ETS

[4]There are three categories of patents in the Chinese patenting system, namely invention, utility and design. Utility and design patents require no substantive examination and reflect only incremental innovation (Hu et al., 2017). Applications for invention patents need to pass through an examination for novelty and non-obviousness. Because the other two types of patents are not subject to examination, they are particularly vulnerable to the abuses of the patenting system to preempt competition from foreign firms (Hu and Jefferson, 2009).

[5]This is a common practice in the existing literature related to studies on Chinese patenting. To list a few, depending on the type of questions answered, the literature either categorizes the patenting variables by the type of innovation (Liu and Qiu 2016 and Hu et al. 2017), or only focuses on the invention patent category (Bombardini et al. 2017, Li 2012, and Dang and Motohashi 2015).

reduces entry but increases levels of green patents for innovating firms, especially at the upper range of output per worker distribution.

The Chinese ETS pilots are of particular interest for three reasons. First, China contributes over a quarter of global carbon emissions (Le Quéré et al., 2017). Even though this paper focuses on regional pilot implementation of ETS, even a partial policy response can have large cumulative effects on global emission trajectories. Second, China is moving towards integrating the separate emissions trading pilots; as a first step, they launched a national trading scheme in December 2017. Even though the national scheme covers only the electricity sector at present, it already comprises the world's largest carbon market by covering over 30% of Chinese emissions (ICAP, 2018). A greater understanding of the industry responses to the pilot schemes will allow policymakers to better anticipate the impacts of the national ETS. Third, the Chinese context distinctly differs from the developed country context of most existing ETS: China is a transitional economy with a number of institutional and historical differences from the European and US economies. Hence, it is not obvious whether one can simply extrapolate results from the latter context to the Chinese ETS. By considering the Chinese case specifically, this paper assesses whether past research on European and North American environmental regulation generalizes to the Chinese context.

I find positive and significant effects of the pilot ETS on firms' innovation, which is in line with the existing literature on the effect of environmental regulation on innovation and technology adoption. For instance, Gray and Shadbegian (1998) find that new plants in states in the US with more stringent environmental regulation are less likely to adopt dirtier production technologies. Popp (2003) explores the effect of the US Clean Air Act (CAA) of 1990 on innovations in pollution control for power plants, and finds that innovation occurring after passage of the CAA was more environmentally friendly. Brunnermeier and Cohen (2003) find that increases in pollution abatement expenditures are associated with a small but statistically significant increase in environmental innovation. Tang (2015) studies the impact of Cleaner Production Audit (CPA) programs on innovation in Chinese listed companies, and confirms a positive effect. In summary, findings from these studies conclude that there is a positive link between environmental regulation and innovation. Beyond these substantive findings, this paper points the way forward in learning the effect of carbon pricing on green innovation.

The remainder of the paper proceeds as follows. Section 2 provides some additional institutional background by reviewing the main characteristics of the Chinese ETS pilot schemes. The data used in the empirical analysis are described in Section 3, while Section 4 lays out the empirical strategy. Results are presented in Section 5, and Section 6 concludes.

2 Pilot Emissions Trading Schemes in China

In recent decades, China has adopted several market mechanisms to combat climate change. With the target of efficiently reducing greenhouse gas emissions by 2020, the Chinese National Development and Reform Commission (NDRC) approved the implementation of pilot emissions trading schemes (ETS) in 2011. Seven provinces, municipalities and regions were selected as "*pilot regions*".[6] The aim of these pilot regions is to reduce CO2 emissions, learn about the effects of the program, and ease the transition towards country-wide, market-based environmental regulation. Beijing, Shanghai, Tianjin and Guangdong released individual plans and implemented pilot ETS at the end of 2013, while Shenzhen implemented its pilot ETS in June 2013. Hubei and Chongqing initiated pilot ETS in April and June 2014, respectively. Lastly, on 22 September 2016, Fujian Province voluntarily opted in and released a conditional announcement of the introduction of China's eighth pilot scheme.

The China pilot ETS are designed as trading systems based on either an absolute cap or an intensity target. In all pilots, the large majority of firms receive grandfathered emission allowances. Firms that emit less than their allowances can sell excess allowances at the market price. Conversely, if emissions exceed the initial allowance, additional allowances have to be purchased to ensure compliance. Below, I discuss several additional key aspects of the Chinese ETS, including the regulated sectors and the coverage threshold that determines which firms are regulated. Further details about these are presented in Appendix A.

[6] These are four municipalities (Beijing, Tianjin, Shanghai and Chongqing), one special economic zone (Shenzhen), and two provinces (Hubei and Guangdong).

2.1 Allowances Allocation

There are two approaches to the allocation of emissions allowances: they are either freely allocated or sold by auction. In China, the allowances are freely allocated in all the pilot regions except for Guangdong, where at most 5% of the total amount of allowances are auctioned. Two ways of allocating allowances freely are grandfathering and benchmarking, which are commonly used in China.[7]

All eight pilot regions determined the total allowances based on the emissions mitigation targets in the 13th Five Year Plan (period 2015-2020). For instance, the target for Beijing is to rigorously control total carbon emissions and meanwhile reduce carbon emissions intensity, while Hubei aims to reduce the emissions intensity annually, without controlling for total carbon emissions. These intensity reduction targets differ slightly in a majority of the pilot regions, ranging from a 19 percent to 22 percent reduction by 2020 compared to the intensity in 2015.

2.2 Coverage Thresholds

Unlike the thresholds in the EU ETS, which are determined at the plant level, the thresholds in the pilot ETS in China are determined at the firm level and differ across the pilot regions. The threshold is highest in Hubei at over 100,000 tons of annual CO2 emissions over the period 2013-2015, and lowest in Shenzhen at 3,000 tons of annual CO2 emissions. Since 2016, the thresholds dropped in Beijing, Shanghai and Hubei by over 50 percent on average. In contrast, Shenzhen, Chongqing, Tianjin and Guangdong have not reduced the thresholds.

2.3 Regulated Sectors

Apart from the thresholds, a firm's sector might determine whether a firm is regulated or not. In Tianjin, for instance, firms in the transportation sector are exempted from the regulation, regardless of emissions, while in Beijing, the threshold is the sole determinant of whether a

[7] With grandfathering, regulated firms receive free allowances initially according to their historical emissions in a base period; with benchmarking, the firms receive allowances according to performance indicators, such as firms' annual production and emissions relative to an industry or a sector.

firm is part of the ETS. In Guangdong, more sectors, i.e., the paper and aviation industries, are included in the ETS. Over time, the coverage of the regulation has become broader and more sectors and firms are being regulated.

Due to differences in total allowable emissions, coverage thresholds and the sectors subject to the ETS, equilibrium prices for the emission allowances differ across the eight regions. The monthly average allowance price ranges from 87 Yuan (13 US dollars) in the Beijing pilot to 1.61 Yuan (0.24 US dollars) in the Chongqing pilot. This heterogeneity in allowance prices implies that firms' costs of compliance, and thereby the incentive to innovate, in CO2-reducing technologies differ across regions.

3 Data

In this section I describe the data used for the analysis. The data originate from three different sources: the regulatory status from local Development and Reform Commissions, patent application data from the State Intellectual Property Office, and firm characteristics from the Annual Survey of Manufacturing Enterprises (ASME).

3.1 Regulatory Status

Information on the regulatory status of firms is obtained through municipal and provincial development and reform commissions (DRCs). As the Chongqing DRC does not publish the list of regulated firms, it is excluded from this study. The number of regulated firms is summarized in Table 1.[8] Specifically, it lists the number of regulated firms in each pilot region and each year from 2013 to 2016. Most notable from Table 1 is the rapid increase in the number of regulated firms in Beijing, Shanghai and Shenzhen in 2016, caused by the downward adjustments in coverage thresholds.

[8]Regulated firms in this paper refer to those that are part of the pilot ETS regulation and hence are in the treatment group. Non-regulated firms are those that are not regulated by the pilot ETS and hence are in the control group.

Table 1: Number of Entities Regulated in China Pilot ETS

Pilot	Year			
	2013	2014	2015	2016
Beijing	450	543	551	947
Shanghai	197	197	197	310
Shenzhen	639	636	635	824
Tianjin	114	112	109	109
Hubei	NA	138	167	236
Guangdong	184	194	186	244
Fujian	NA	NA	NA	277

3.2 Patent Data

The annual number of patent applications is used as a proxy for firms' innovation activities.[9] Patent data come from the system of Patent Search and Analysis, which is hosted by the State Intellectual Property Office (SIPO) of China.[10]

All patents in China are categorized based on the International Patent Classification (IPC). The IPC provides a universal language for the classification of patents according to the different technology areas to which they pertain. Because the interest of this study is to explore the effect of CO2 regulation on the firms' green innovation activity, I consider a subset called the "IPC Green Inventory" between 2007 and 2016. These are the patents related to so-called Environmentally Sound Technologies (EST, henceforth green patents) (IPC Committee, 2017), as listed by the United Nations Framework Convention on Climate Change. I use the patent classification codes for technologies on alternative energy production, transportation, energy conservation, waste management, nuclear power generation and administrative, regulatory or design aspects to select the green patents, with technologies on agriculture excluded from

[9]An alternative measure of innovation in the literature is RD expense. Though patent data is broadly accessible in China, RD expenses of firms for consecutive years is limited, making it infeasible in the current context. Using patent data to proxy for innovation is a common approach in empirical studies, such as Hu and Jefferson (2009), Dang and Motohashi (2015), Bombardini et al. (2017) and Liu and Qiu (2016).

[10]SIPO was renamed the China National Intellectual Property Administration (CNIPA), on 28 August 2018. The data are accessible through the URL http://www.pss-system.gov.cn/sipopublicsearch/portal/uiIndex.shtml (first accessed December 2017 with subsequent access in July, 2018). I collected the data using web-scraping. There is a time lag between publication date and application date. Some patents applied before 2018 might not be published by the date of access. The average time lag between 2007 and 2012 is 400 days, and the median is 230 days. More than 75 percents of the filed patent are published after 540 days (around one and half years) of the application date. Therefore, by the date of access, the patent data could well represent the population of patent applications, at least for patents filed before 2017.

the category because these technologies are not directly related to low-carbon technology. In addition, following Dechezleprêtre et al. (2020), I exclude from the IPC green inventory patents in carbon-intensive technologies such as gas-turbine plants, cremation furnaces, and steam-engine plants.

In order to estimate the ETS effect on the direction of the technological change, and whether the ETS increases the green patents at the cost of dirty patents, I rely on Dechezleprêtre et al. (2020) to identify the patent classification codes on the dirty technologies. These mainly include patents on electricity generation technologies and technologies in the automobile industry.

For each individual patent, the dataset contains information on the IPCs, the name of the invention, application number and date, publication number and date, applicants, address of applicants, and whether or not an application is approved.[11]

I use this dataset to construct the number of patent applications at the firm-year level.[12] Figures 1 and 2 show the numbers and shares of green and dirty patent applications for regulated and non-regulated firms from 2007 to 2016. Figure 1 presents both the total and weighted number of green patents, where in the latter case a $1/n$ share of the patent is assigned to each applicant firm, with n the number of applying firms. As such, the weighted patents avoid double-counting when the patent is filed by several co-applicants.

The vertical dashed lines in the figures indicate the years that ETS pilots were announced (2011) and implemented (2013). As shown in Figure 1, the total number of green patent applications by regulated firms did not grow as fast as those by non-regulated firms. Meanwhile, the shares of green patent applications for regulated and non-regulated firms increased nearly parallel to each other before 2011 (Figure 2). Since 2011, the share for regulated firms has increased rapidly, while the share for non-regulated firms has been rather flat. The trends in the unweighted green patents are similar to the weighted ones both for regulated and non-regulated firms, indicating that the average number of applicants per patent does not noticeably vary across firm types and over time. The shares of dirty patents have been flat both for

[11]Contrary to patent data hosted by the European Patent Office, SIPO does not include information on citation, which is commonly used as a measure on patent quality.

[12]Details about merging and constructing the dataset are in Appendix B.

Figure 1: Number of green patents 2007-2016, weighted and unweighted

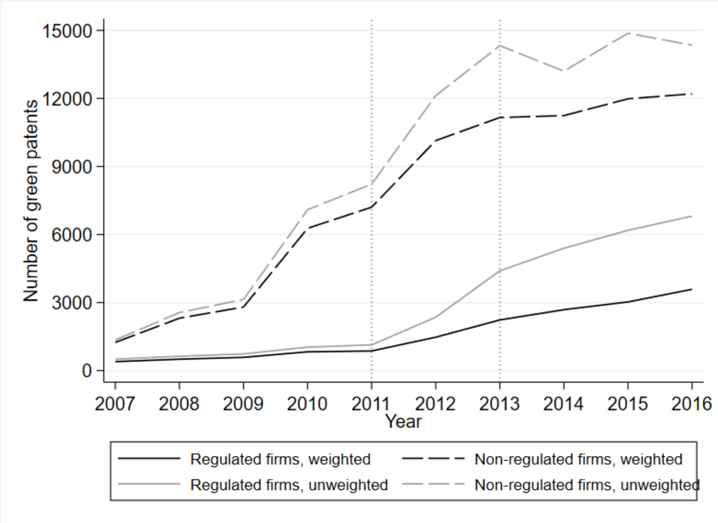
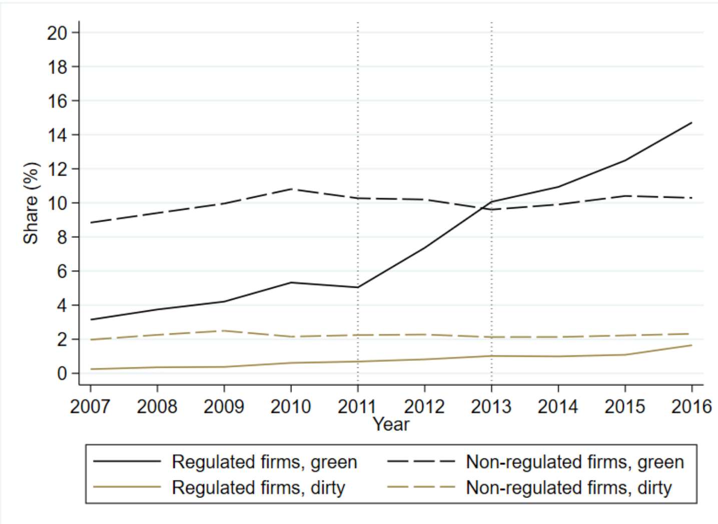


Figure 2: Share of green and dirty patents 2007-2016, weighted



regulated and non-regulated firms.[13] The figures suggest that following the implementation of the ETS pilots, regulated firms have shifted towards "greener" innovation. Such a shift is not apparent for non-regulated firms.

3.3 Firm-level production data

The firm-level production data, Annual Survey of Manufacturing Enterprises, are collected on an annual basis by China's National Bureau of Statistics (NBS). All industrial firms above a given size of annual sales are surveyed. This includes all state-owned firms, as well as non-

[13]The shares of green and dirty patents are calculated as the weighted patent counts in each respective category divided by the sum of all the weighted patent counts in one year.

state owned firms with sales exceeding 5 million Yuan.[14] In 2011, the designated size increased from 5 million to 20 million Yuan for all surveyed firms.[15]

The manufacturing data used in this study spans 2007 until 2013. I do not use the 2010 data due to data quality concerns,[16] and no data is available after 2013. The dataset includes basic information such as firm name, location and the number of employees. Almost all of the entries in a balance sheet and an income statement are included in most of the census years, such as sales revenue, total assets, output and costs.

Table 2 presents the summary statistics. In the table, the "pilot regions" refer to the provinces or municipalities that implemented the pilot ETS, as introduced in Section 2. The "non-pilot regions" include all other regions in mainland China. Table 2 shows that, compared to those in non-pilot regions (column 2), firms in pilot regions (column 3) are slightly larger: on average, they have higher employment, greater sales, produce more output and hold more assets and capital. In pilot regions (columns 4-5), employment in regulated firms is on average six times the employment in non-regulated firms; sales, output and assets are more than ten times larger.

Table 3 presents the summary statistics for patent applications. On average, firms in pilot regions file more patents, and especially more green patents, both before and after 2013 (columns 3-6). It is noteworthy that from year to year, for regulated firms, the average number of green patent applications more than quadrupled from 1.37 to 5.76 (weighted counts, columns 9 and 10), while the increase for non-regulated firms in the pilot regions is rather modest (columns 7 and 8). The number of dirty patents has also tripled, both for regulated and non-regulated firms.

The dataset presented above is constructed by first of all merging the two sources of the data, regulatory status and patent data, which gives a sample with 370,267 non-regulated

[14]This is equivalent to about 740,000 US dollars.

[15]For further characteristics and caveats of this dataset, see Brandt et al. (2014).

[16]Concerns have been raised about the quality of this data after 2008. For instance, Chen et al. (2019) find that investments, net exports and value-added of sectors are largely discrepant between local and national statistics. In another study, Chen (2018) discusses several issues to which the user should pay attention when using these data and suggests a method for validating the authenticity of the main variables in the survey data. Using their method, I find that the 2010 is likely problematic, while the data quality is good in other years. For this reason, I do not use the 2010 data. Cai and Liu (2009) and Feenstra et al. (2014) additionally point out potential misreporting due to administrative errors. To address this, I follow their suggested approach to clean the data and drop firms with fewer than 8 employees.

Table 2: Summary Statistics 2007-2012

	(1) All	(2) Non-pilot regions	(3) Pilot regions	(4) Pilot regions Non-regulated firms	(5) Pilot regions Regulated firms
Employment	638.35	635.33	647.22	483.14	2,994.96
	(2,976.84)	(2,935.17)	(3,096.12)	(1,768.13)	(9,805.52)
Total assets	660.43	619.14	781.80	407.63	6,135.72
	(6,155.51)	(4,165.89)	(9,911.92)	(4,232.05)	(34,892.08)
Current assets	300.41	286.93	340.04	199.61	2,349.43
	(1,998.00)	(1,722.86)	(2,645.71)	(1,149.76)	(9,162.34)
Sales	630.51	613.92	679.27	387.03	4,861.03
	(4,600.52)	(4,251.69)	(5,499.13)	(3,386.16)	(16,740.77)
Cost of sales	526.63	511.50	571.12	321.66	4,140.62
	(3,929.21)	(3,603.69)	(4,758.75)	(3,072.66)	(14,071.74)
Output	607.77	589.86	660.41	382.03	4,643.73
	(4,149.04)	(3,728.56)	(5,191.30)	(3,244.61)	(15,653.41)
Capital	134.21	114.98	190.76	114.17	1,286.73
	(3,231.56)	(2,780.51)	(4,290.85)	(3,723.94)	(9,064.54)
Observations	191143	142629	48514	45345	3169

This table presents means and standard errors for each variable. Standard errors are in parentheses. All variables except for employment are in million Yuan.

All the statistics are based on data between 2007-2012, with the data in 2010 excluded because it is not validated, as discussed in this section.

Table 3: Summary statistics: number of patents, full sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	2007-2012	2013-2016	2007-2012	2013-2016	2007-2012	2013-2016	2007-2012	2013-2016	2007-2012	2013-2016
All patents	2.13	6.78	1.46	5.26	4.02	11.02	2.13	6.52	31.53	62.71
	(38.53)	(87.57)	(7.43)	(21.43)	(74.19)	(166.63)	(19.15)	(65.27)	(281.91)	(543.24)
Green patents	0.20	0.86	0.15	0.58	0.35	1.64	0.24	0.77	1.88	11.63
	(3.68)	(29.91)	(1.56)	(6.29)	(6.69)	(57.28)	(4.19)	(10.44)	(20.94)	(199.11)
Dirty patents	0.04	0.14	0.04	0.13	0.04	0.16	0.03	0.10	0.18	0.86
	(0.46)	(1.78)	(0.42)	(0.97)	(0.55)	(3.07)	(0.39)	(0.70)	(1.56)	(10.58)
All patents, weighted	1.92	5.83	1.36	4.81	3.48	8.67	1.85	5.55	27.23	44.62
	(35.28)	(53.13)	(6.90)	(16.49)	(67.92)	(99.66)	(12.71)	(48.15)	(262.29)	(309.94)
Green patents, weighted	0.17	0.62	0.14	0.48	0.27	1.02	0.20	0.61	1.37	5.76
	(2.21)	(12.02)	(1.38)	(3.44)	(3.64)	(22.68)	(2.41)	(6.22)	(10.95)	(77.21)
Dirty patents, weighted	0.04	0.12	0.03	0.12	0.04	0.12	0.03	0.08	0.15	0.55
	(0.42)	(1.13)	(0.41)	(0.89)	(0.46)	(1.62)	(0.37)	(0.59)	(1.15)	(5.33)
Observations	202086	114237	149126	84105	52960	30132	49554	27720	3406	2412
Sample	All	All	Non-pilot regions	Non-pilot regions	Pilot regions	Pilot regions	Non-Rregulated firms	Non-Rregulated firms	Regulated firms	Regulated firms

This table presents means and standard errors for each variable on the full sample. Standard errors are in parentheses.

firms and 1,495 regulated firms. Then I exclude all the firms in the service sector, i.e., all the universities, government agencies, and restaurants and hotels, because these entities are not likely incentivized to innovate on their own, but rather adopt abatement technologies to reduce the marginal cost of abatement. Next, I merge the data with the firm-level production data[17], the Annual Survey of Manufacturing Enterprises, which further reduces the sample size and gives a sample with 61,358 non-regulated firms and 1,081 regulated firms. Then I drop the firms that do not contain information on industry classification, sales and labor, which leads to 56,335 non-regulated firms and 784 firms respectively. This is less than the actual number of regulated firms (2,621) for the following two reasons.

First of all, there are 1,495 regulated firms that filed at least one patent between 2007 and 2016 (regardless of being 'green' innovation or not), while there are 1,126 that never filed a patent in this period, which are excluded from the sample. These excluded firms filed no patents either before or after the implementation and hence do not respond to the policy by innovating more. Secondly, in ASME, only manufacturing firms with annual sales above a certain threshold are surveyed, as introduced in Section 3.3. Therefore, regulated firms that do not reach this threshold, or reach this threshold but are not manufacturing firms, such as firms in the transportation sector, would not be surveyed. In other words, the further reduction of the number of regulated firms when merging three sources of data is because those firms were not surveyed, because they did not achieve high enough annual sales.

4 Empirical Strategy

Section 3 documented that regulated firms and non-regulated firms are different in observable characteristics. This section introduces the empirical framework, which relies on a count data model with a matched dataset. The motivation for matching is also discussed in this section.

[17]See Appendix B for the steps of the data construction.

4.1 Empirical Model

The empirical identification of the effect of the pilot ETS on green innovation by regulated firms is based on the variations in regulatory status across firms, as well as differences in the regulation of the pilot ETS across pilot regions. I adopt a differences-in-differences design to estimate the effect of the ETS pilots on firm-level innovation.

A main challenge of empirically identifying the causal effect of the pilot ETS on innovation is the non-random assignment of the treatment due to the regulation threshold introduced in Section 2.2. If I know carbon emissions intensity (emissions per unit of output) of the population of firms, I could compare the green patenting of regulated firms with that of the non-regulated firms that have exactly the same emission intensity as the regulated firms before and after the implementation of the regulation. An alternative would be to include a vector of control variables that correlate with firms' emissions and therefore the treatment status, if I had data on full sets of control variables in both pre- and post-treatment periods – in other words, all the data on ASME between 2007 and 2016. Then I could obtain an unbiased estimation on the effect of the regulation on the number of patent applications. However, due to the lack of data availability after 2013, as discussed in Section 3.3, this is not feasible. To address the issue, I first pre-process the dataset using matching methods. Then I estimate the regression equations on the matched dataset. Matching is favourable as it requires only the data in the pre-treatment period and hence the matched data have better balance between the treatment group and the control group. The related matching methods are described in detail in Section 4.2 and Appendix C.

Because the dependent variable of interest, the number of green patents, is a numerical count, I use a count data model to estimate the effect of pilot ETS. Specifically, I adopt a zero-inflated Poisson (ZIP) regression model, as proposed by Lambert (1992).[18] This model allows me to deal with the zero patent applications observed for a substantial number of firms, and allows for greater flexibility in the distributions of zeros and strictly positive applications. The firms that file a positive number of green patents likely have a different data generating

[18]This model is commonly applied in patenting studies. To give a few examples, Hu and Jefferson (2009) use ZIP regression to analyse the factors that led to a patenting surge in China; Noailly and Smeets (2015) study the driving forces of innovation on renewable and fossil-fuel energy in the electricity generation sector in Europe.

process of patent counts than those with zero counts. Hence it is intuitive to use two-part models to allow for flexible specification of the distributions of zeros and positives, as proposed by Mullahy (1986).[19] Such a two-step process allows for an analysis of multiple margins of decision-making: an extensive margin decision of whether green patenting is worthwhile to the firm, followed by an intensive margin decision of how many green patents to file.

The basic idea behind ZIP is as follows. The firms are categorized as two types: firms that invest in R&D to innovate green technology (henceforth innovators), and firms that do not make any investments in green technology (henceforth non-innovators). The probabilities of being an innovator and a non-innovator are $1-\pi$ and π respectively. In turn, for an innovating firm i, the distribution of patent counts in year t is Poisson with mean λ_{it}. This then gives the baseline regression specification:

$$f(y_{it}) = e^{-\lambda_{it}} \lambda_{it}^{y_{it}} / y_{it}!, \tag{1}$$

where

$$\lambda_{it} = \mathbf{E}[y_{it}] = exp(\beta_1 regulated_i \times post_t + \beta_2 regulated_i + \gamma_{i,o} + \delta_{i,size} + \alpha_t + \eta_l). \tag{2}$$

In the above equation, y_{it} denotes the count of green patents that innovator firm i filed in year t. The primary variable of interest, the interaction term $regulated_i \times post_t$, is an indicator equal to one if, in year t, firm i is regulated in the carbon market. That is, the treatment indicator, $regulated_i \times post_t$, turns on for firms included in the pilot trading scheme; for control group firms, this interaction term does not change over time and equals zero. I control for year fixed effects (α_t), which account for the time-variant changes that affect all firms similarly. I include the region dummy η_l to account for time-invariant green patenting difference across regions. This dummy controls for region-level institutional differences, such as province-level patent subsidy programs.[20] In addition, the specification also includes a vector of ownership

[19]This is important for the following reasons. First, there is a significant proportion of zeros in the number of filed patent applications. Second, there are very large counts of filed patents that contribute substantially to overdispersion. See also Figure 9 in Appendix D.

[20]However, the effects of the pilot ETS on green patenting are not biased by these regional policy initiatives, because 29 out of 31 provinces and municipalities in mainland China had a patent subsidy program in place by

dummies $\gamma_{i,o}$ to account for differences in patenting behavior between state-owned and non-state-owned firms,[21] and size dummies $\delta_{i,size}$ to take into consideration different patenting ability for firms with different size.[22]

The ZIP model therefore specifies

$$Pr(greenpat_{it} = y_{it}) = \begin{cases} \pi_{it} + (1 - \pi_{it}) f(0; \lambda_{it}) & \text{if } y_{it} = 0, \\ (1 - \pi_{it}) f(y_{it}; \lambda_{it}) & \text{if } y_{it} = 1, 2, 3, 4, \ldots \end{cases} \tag{3}$$

Here, $greenpat_{it}$ is the number of green patents filed by firm i in year t. Note that the large number of zero counts of patents may occur for two different reasons. The first reason is that firms do not find it profitable to innovate regardless of the regulation or fail to innovate and therefore file no patents (non-innovator). The second reason for zeros is that firms do innovate but do not use patents as a way of protecting their intellectual property, or are incapable of filing a patent (potential innovator). These two different sources of zeros in patenting data are characterized by π_{it} and $(1 - \pi_{it}) f(0; \lambda_{it})$ respectively. As noted above, π_{it} is the probability of being a non-innovator for firm i in year t; $(1 - \pi_{it}) f(0; \lambda_{it})$ is the probability of being a potential innovator with zero patents filed. At the extensive margin, the firm decides whether or not to be an actual innovator with positive applications, which is captured by the following logit regression, as in Lambert (1992),

$$logit(\pi_{it}) = log(\pi_{it}/(1 - \pi_{it})) = X'_{it}\beta. \tag{4}$$

Hence the likelihood of not being an innovator is estimated via logistic regression

$$\pi_{it} = \frac{e^{\mu_{it}}}{1 + e^{\mu_{it}}}, \tag{5}$$

where $\mu_{it} = log(\lambda_{it})$ in Equation 2 influences the extensive margin of patenting, i.e., whether

the end of 2007 (Li, 2012).

[21] The results by Hu and Jefferson (2009) indicate that non-state-owned firms may be more keen to seek patent protection.

[22] I categorize firms as large, medium, small and miniature firms based on sales and labor according to the firm size measure by the National Bureau of Statistics. For details see Appendix C.

or not the firm files patents. In summary, in the first regression, a logit model estimates the probability of filing green patents with an outcome of zero or one (extensive margin). In the second regression, a count data model estimates the patent count using a Poisson model for firms with at least one green patent filed (intensive margin).

A large variation in the carbon prices across different pilot regions in China provides a chance for me to look directly at the continuous treatment effect of the pilot ETS on firms' green innovation. Fell and Maniloff (2018) and Calel and Dechezleprêtre (2016) study the effect of the U.S. Regional Greenhouse Gas Initiative (RGGI) and the effect of the EU ETS. In these two studies, they estimate the discrete treatment effects instead of the continuous effects that would be captured by the carbon prices, which is due to little variation in the carbon prices in the RGGI states and EU ETS countries during the period studied. Complementary to their studies, I study the effect of carbon pricing on the number of green patents using the following regression specification

$$y_{it} = exp(\beta_3 price_{t+g,l} \times regulated_i \times post_t + \beta_4 regulated_i + \gamma_{i,o} + \delta_{i,size} + \alpha_t + \eta_l) + \epsilon_{it}. \tag{6}$$

Here $price_{t,l}$ is the logarithm of the yearly average carbon price in region l in year t. Carbon prices are strictly positive for regulated firms after the implementation of the pilot ETS, and are zero for all non-regulated firms and regulated firms before the implementation of the pilot ETS. The coefficient β_3 is the parameter of interest that captures the average change of green patents as carbon price increases by one percent. Assuming that on average current carbon prices are the best predictor of future carbon prices, I use the current carbon prices in the baseline regression.[23]

One complexity arises from the possible firm heterogeneity that influences firms' patenting ability, which is not accounted for by matching. There is a rich literature on the econometric techniques to account for firm-level fixed effects in Poisson models, primarily Blundell

[23] The additional results on the estimations with different leads of carbon prices ranging from 1 to 3 are presented in Appendix D.1 to take into consideration that firms decide whether to innovate based on their expectation of carbon prices in the future. Here I assume that firms are informed and are able to fully anticipate the carbon price level in the future.

et al. (1995), Blundell et al. (1999), Blundell et al. (2002) and Hausman et al. (1984). The first three papers by Blundell et al. propose that time-invariant firm heterogeneity could be accounted for using pre-sample mean of patent count, and a dummy equal to one if the firm innovated in the pre-sample period.[24] However this would require a long pre-sample history of the dependent variable to proxy the firm fixed effects, which is not feasible in this study due to lack of data in the pre-sample period. Hausman et al. (1984) developed a conditional maximum likelihood estimator which can be applied to count data of a panel nature to capture the persistent firm fixed effects. They suggest an estimator conditioning on the total sum of outcomes over the observed years to proxy the fixed effects.

The proxied firm-fixed effects in Hausman et al. (1984) require strict exogeneity, i.e., that the firm-specific effect is uncorrelated with the explanatory variables. This would be violated if firms have strong innovation ability in the pre-treatment period, and hence are able to reduce the carbon emissions below the regulatory threshold. The firm-specific effect might therefore be negatively correlated with the treatment dummy. Therefore, the proxies of firm fixed effects using data in either pre-sample or in-sample period are infeasible. An alternative is to assume that the zero counts and non-zero counts have the same data-generating process without explicitly considering the probability of a regulated firm switching from a non-innovator to an innovator. Under such an assumption, I can then estimate a fixed effects Poisson model. I discuss the potential issue with this model in Section 5.4.4.

The remaining issue relates to the estimation of standard errors. Across specifications, I cluster the standard errors at the four-digit sector level, because the regulations differ in different sectors. For instance, different sectors might be subject to different coverage threshold and rules of allowances allocation, as introduced in Section 2.[25]

[24]Building on Blundell et al., Aghion et al. (2016) derive a similar approach using the post-sample mean and dummy to capture such firm heterogeneity.

[25]See Appendix A for a detailed review on the difference of the regulation in different pilot regions. Ideally, I would adjust standard errors for clustering at region level to allow for serial correlation within a region across years. However, with six clustering units, standard errors would be underestimated, which leads to an inference problem. (Bertrand et al., 2004)

4.2 Matching

One complexity of this study arises from the lack of data on the Annual Survey of Manufacturing Enterprises (ASME) in the post-treatment period. Matching could address this by only using the data in the pre-treatment period, so that treatment and control groups are better balanced on a vector of control variables. To control for the confounding influence of pre-treatment control variables, I match regulated and non-regulated firms in the same 2-digit sector, region, as well as on labor and sales revenue, and whether filing at least one patent in the pre-treatment period, number of green patent applications and number of all patent applications. That is, I first of all implement exact matching for firms on a 2-digit sector and province or municipality and a dummy equal to one if a firm filed at least one patent before 2013. The firms in the non-pilot region are thus dropped from the baseline sample. I then match firms on labor, sales revenue and number of patents with measures of tolerable distance between regulated and non-regulated firms, which I discuss below. The first two are selected to capture firms' size and profitability.[26] The last two variables control for firms' pre-treatment innovation ability.

The key goal of matching is to prune observations from the data so that the remaining data have better balance between the treated and control groups, meaning that the empirical distributions of the covariates in the groups are more similar (Iacus et al., 2012).[27] I use coarsened exact matching (CEM), as proposed by Iacus et al. (2012), in combination with genetic matching (GM), proposed by Diamond and Sekhon (2013). The intuition and the technical details of matching are presented in Appendix C.

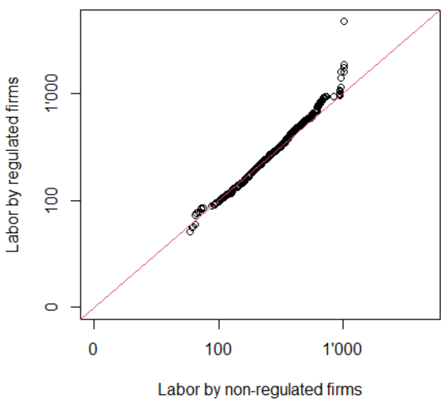
Figure 3 shows the quantile-quantile plots for the matched variables, average employment, average sales, and the numbers of all patents and green patents between 2007 and 2012. The points on the plots fall reasonably on the 45 degree straight line. Of course, matching only on the selective subset of the variables might not capture all these dimensions. I thus show in Figure 4 the quantile-quantile plots for the matched sample on variables that are not used for

[26]The other reason for choosing these variables is that the information on these two variables is always reported across years.

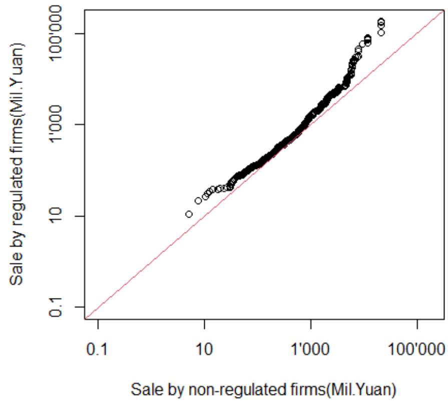
[27]Due to the large size of the control group compared to the size of the treatment group, I could identify a sub-group of non-regulated firms which are comparable with regulated firms with matching. For a useful review and practical guidance on matching methods, see Stuart (2010)

Figure 3: Quantile-quantile plots on matched sample, matching variables

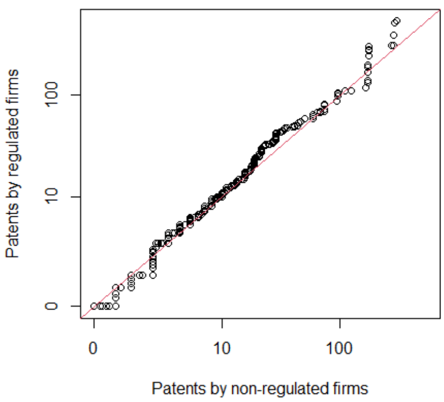
(a) Employment



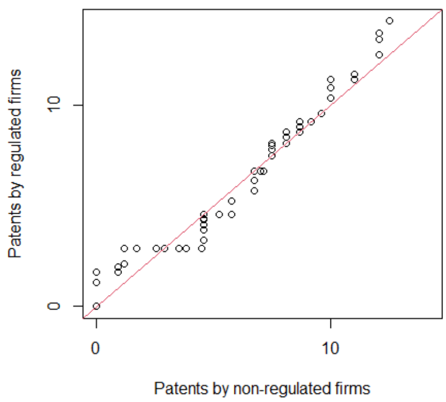
(b) Sales



(c) Number of all patents



(d) Number of green patents

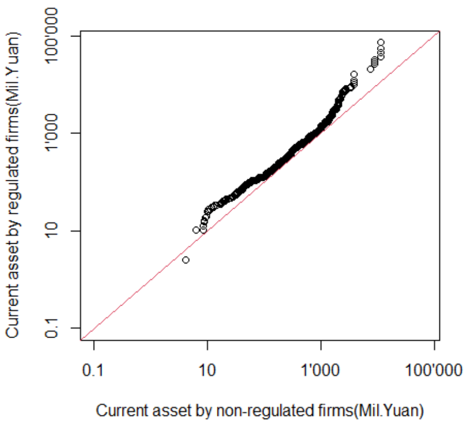


matching, including current assets, output, operating cost and total assets. As Figure 4 shows, the empirical distribution of the non-matching variables of the regulated and non-regulated firms are very similar.

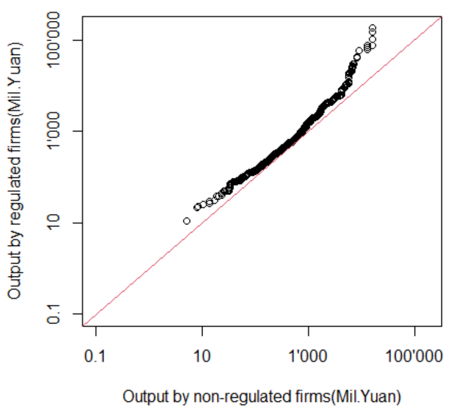
Ideally, I would have a group of unregulated firms that is exactly the same as the group of regulated firms in every aspect, especially those influencing their green innovation ability, except for the regulatory status. A related concern is that, even though the empirical distributions of the matched regulated and non-regulated firms are very similar in the variables shown in Figures 3 and 4, they might have very distinct emissions intensity of production, and therefore might not be comparable with each other. However, due to the general lack of availability of firm-level carbon emissions data in the pilot regions, it is not feasible to directly

Figure 4: Quantile-quantile plots on matched sample, non-matching variables

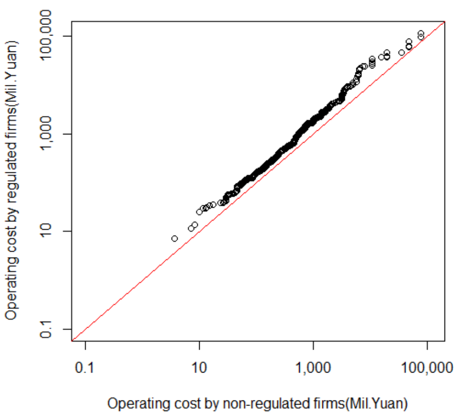
(a) Current asset



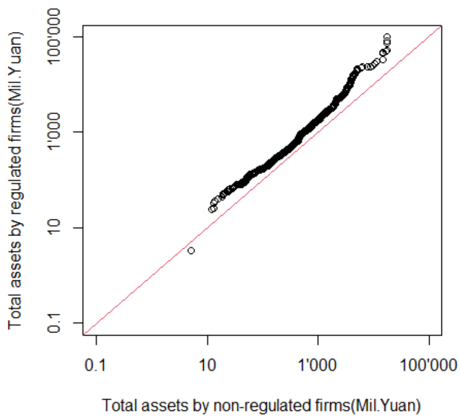
(b) Output



(c) Operating cost

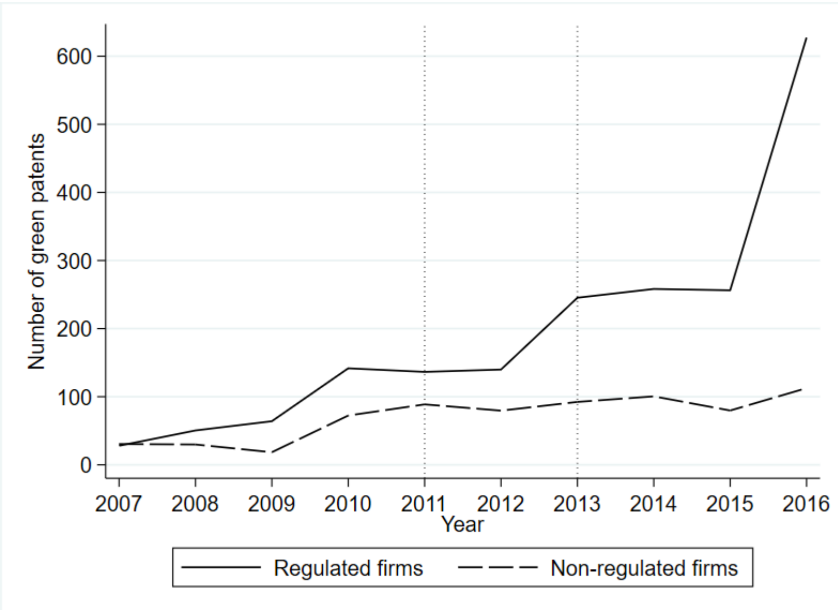


(d) Total assets



compare firms with the same emission intensity. Imagine that the matched regulated firms have far higher emissions intensity than the matched non-regulated firms. This case could be due to, for instance, the regulated firms using more carbon-intensive energy or dirtier technology for their output. However, as Figure 5 shows, the number of green patents of the regulated and non-regulated firms before 2013 is very similar. This provides some confidence that the regulated firms' emissions intensity is not substantially higher than the non-regulated firms' emissions intensity.[28] Figure 5 is also suggestive of parallel pre-regulation trends. Table 4 presents summary statistics for the number of patents on the matched (columns 1-4) and non-matched firms (columns 5-8) in the pilot regions before and after the implementation of the pilot ETS regulation. Comparing columns (7) and (3), the regulated firms that are relatively more innovative are not matched with any of the unregulated firms.

Figure 5: Number of green patents 2007-2016, matched sample



[28] Additionally, as Figure 11 in Appendix D shows, the means of the number of green patents on the matched sample are similar in the pilot regions. This provides reassuring evidence that the production techniques should not be largely different and therefore the emissions intensity of matched regulated and non-regulated firms should be similar.

Table 4: Summary statistics: number of patents, matched and non-matched samples

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Pre	Post	Pre	Post	Pre	Post	Pre	Post
All patents	3.15	7.26	4.24	14.00	2.07	6.80	187.37	304.88
	(9.27)	(15.43)	(15.89)	(52.16)	(19.13)	(68.53)	(708.34)	(1,293.47)
Green patents	0.19	0.59	0.23	1.05	0.24	0.82	11.28	64.08
	(0.75)	(1.99)	(0.96)	(7.22)	(4.21)	(10.96)	(53.14)	(482.13)
Dirty patents	0.05	0.10	0.06	0.28	0.03	0.10	0.91	3.76
	(0.29)	(0.58)	(0.43)	(3.75)	(0.39)	(0.72)	(3.81)	(24.31)
All patents, weighted	2.77	6.05	3.44	11.18	1.80	5.78	162.99	211.11
	(8.18)	(12.65)	(11.19)	(39.41)	(12.67)	(50.53)	(661.71)	(728.72)
Green patents, weighted	0.17	0.46	0.19	0.90	0.20	0.65	8.07	29.91
	(0.68)	(1.35)	(0.81)	(7.03)	(2.42)	(6.53)	(27.29)	(185.88)
Dirty patents, weighted	0.04	0.08	0.05	0.25	0.03	0.09	0.73	2.05
	(0.28)	(0.55)	(0.42)	(3.74)	(0.37)	(0.61)	(2.72)	(10.04)
Observations	1864	1076	3005	1897	49249	25077	510	406
Sample	Non-regulated firms	Non-regulated firms	Regulated firms	Regulated firms	Non-regulated firms	Non-regulated firms	Regulated firms	Regulated firms
Matched	Yes	Yes	Yes	Yes	No	No	No	No

This table presents means and standard errors for each variable of firms in the pilot regions. Standard errors are in parentheses.

5 Results

5.1 The Impact of the Pilot ETS: Main Results

The first column in Table 5 present the Poisson estimations while the rest of the columns are estimations from the zero-inflated Poisson (ZIP) regression. Columns (2)-(5) compare results from estimations of Equation 2 with ownership, pilot region, and firm size dummies added. Column (6) presents results from estimations of Equation 6.[29] Column (7) shows the estimations of Equation 2 using the weighted approved green patent counts as an outcome variable. All models include a full set of year dummies (not reported). ZIP is more flexible than the Poisson regression, because it relaxes the assumption that data are equi-dispersed, i.e., the variance of count data conditional on a vector of regressors x equals the conditional mean. Meanwhile, ZIP enables me to model zero green patenting by innovator and non-innovator differently, which better captures the data generating process. Therefore, I use the ZIP regression model as my baseline specification.

For columns (2)-(7), the top part of the table presents the estimations from the Poisson regression for the number of green patents, whereas the bottom part of the table presents the estimations of the logit model in the inflation equation discussed in Section 4.1. The coefficient estimations in the inflation equation assess the likelihood of inflated zeros, i.e., the likelihood of being a non-innovator. Therefore, a negative (positive) coefficient is interpreted as a positive (negative) effect on the likelihood of being an innovator. The estimates in columns (2)-(5) compare the effects of adding pilot region dummies, the ownership dummies, and the firm size dummies. The estimates reveal significant effects for green patenting, while the size of the regulation effect differs. Also, the Akaike information criteria (AIC), shown as AIC divided by the number of observations at the bottom of the table, is decreased by adding the three sets of dummies. This reveals the importance of including these dummies in the regressions.[30] Therefore, I add the ownership dummies, pilot region dummies and firm size dummies in all

[29] The results using unmatched data are shown in appendix D. Generally speaking, the signs of the estimations are the same as the estimations from the matched sample, but with higher magnitude.

[30] A joint hypothesis test also rejects the null hypothesis that the coefficients on the pilot region dummies, the ownership dummies, and the firm size dummies are zero, with a p-value equal to zero.

the following regressions (not reported).

The estimations in column (5) suggest that, compared to the non-regulated firms, the regulated firms respond to ETS by increasing the number of green patents. The average marginal effect of ETS is 0.16, i.e., the number of green patents for regulated firms increased on average by 0.16 (standard error$=0.08$, $p=0.051$).[31] This is equivalent to 11.68 percent and 2.78 percent of the average number of green patents in the pre-treatment period (2007-2012) and post-treatment period (2013-2016), respectively. For large firms, the average marginal effect is 0.20 (standard error$=0.09$, $p=0.03$). The magnitude of the effects decreases as the firm size becomes smaller. For small and medium-size firms, the average marginal effects are 0.15 (standard error$=0.09$, $p=0.08$) and 0.06 (standard error$=0.03$, $p=0.06$) respectively. In the extensive margin, the effects for the regulated firms are all positive, suggesting that the pilot ETS decreases the probability of being an innovator, at least for some regulated firms.[32] However, no significant effects of the pilot ETS in the extensive margin are observed in the data. Therefore, firms respond to the pilot ETS significantly only in the intensive margin.

The estimation in column (6) yields the elasticity of carbon prices on the number of green patents. I assume that on average current carbon prices are the best predictor of future carbon prices. Qualitatively, a higher carbon price leads to more green patents for innovators (at the intensive margin) on average. The elasticity of patents with respect to the carbon price is 0.23. This means that a 10 percent increase in the carbon price will increase green patents produced by 2.3 percent. There could also be forward-looking effects, since innovation requires a stream of investment for a period and will potentially generate returns in the future. Assuming that the firms can perfectly anticipate the future carbon price, I use the carbon price with leads up to three years to take into account the firms' expectation on carbon prices. The results are shown in Appendix D.1. The one-year lead effects of the carbon prices are significant with a magnitude similar to the estimations based on the current price. No significant effects with two- and three-year leads can be observed in the data. This could be because the firms are able to anticipate the carbon price one year ahead and respond to it accordingly, but not beyond

[31] Because the magnitudes for the estimations using ZIP regression are not directly interpretable, I use the Stata built-in command `margins` to get the marginal effect of the regulation on green innovation.

[32] Recall that the coefficient in the logit regression captures the probability of inflated zeros, and a positive coefficient is interpreted as a negative effect on the likelihood of being an innovator.

that.

It is essential to mention one characteristic of patent application data from SIPO: SIPO does not record citations, which is typically used as a measure of patent quality in the literature. This is a common issue in studies of the development of innovation in China using data from SIPO. Thus, granting rate is usually used as an alternative measure for patent quality (Dang and Motohashi, 2015). However, patent granting takes on average 3.87 years after filing a patent with SIPO. Therefore, using the patent granting rate of firms to account for patent quality would not be sufficiently informative in this study, as the policy was implemented in 2013.[33] Still, I report the estimation of the effects using the number of granted patent counts as an outcome variable in column (7) to compare whether the policy has similar effects on the number of approved patents and the number of filed patents.[34] There is a measurement error in this outcome variable in that many of the patents might not yet have been granted at the time of accessing the data. Though the positive sign remains, the effect is underestimated.[35]

The quantitative result on the carbon price elasticity of 0.23 should be interpreted cautiously for two reasons. First, this is an average effect of an increase in carbon prices on the number of green patents. However, if the carbon price is not above a certain level, as in Tianjin (TJ), the pilot regulation would not be effective in terms of inducing green innovation despite the increase in carbon price.[36] Second, the carbon price in Beijing is the highest among all the pilot regions. Hence, for regulated firms in Beijing, a one percent increase in the carbon price will influence the firms more significantly than those located in Tianjin. So, the 0.23 estimated elasticity of carbon prices on the number of green patents implies the average value across all six pilot regions, and applies exclusively to the regulation within the period studied.

[33]One might be concerned that the estimation also captures the anticipation effect, as the policy was announced two years before 2013. However, this is not likely because the list of regulated firms and crucial rules, i.e., the coverage threshold and the allowances allocation, were not released in 2011. Therefore, firms could not predict their regulatory status precisely.

[34]The trends in the means of the number of granted green patents for regulated and non-regulated firms are presented in Figure 10 in Appendix D, which suggests the parallel pre-regulation trends. The approval year is usually not the same as the filing year. The data is compiled based on years that patents are filed.

[35]In another regression, I use the patent grant rate (the number of granted green patents divided by the number of filed green patents) as an outcome variable and estimate the ETS effect on this grant rate using an OLS regression with and without firm fixed effects included. I restrict my sample to a subsample of firms that have at least one green patent filed in each year. The estimations are 0.03 (without firm fixed effects) and 0.05 (with firm fixed effects) respectively. However they are not precisely estimated. The results are not reported.

[36]The carbon price in Tianjin is the lowest among all the six pilot regions explored in this study. See Figure 13 in Appendix D.

Table 5: Emissions trading scheme and innovation

	(1) Poisson	(2) ZIP	(3) ZIP	(4) ZIP	(5) ZIP	(6) ZIP	(7) ZIP
main							
regulated*post	0.49**	0.67**	0.63**	0.65**	0.75**		0.67
	(0.21)	(0.31)	(0.32)	(0.31)	(0.32)		(0.50)
regulated	0.20	0.10	0.21	0.34	0.27	0.23	0.24
	(0.13)	(0.26)	(0.23)	(0.25)	(0.24)	(0.23)	(0.28)
Logarithm carbon price						0.23**	
						(0.10)	
inflate							
regulated*post		0.21	0.28	0.36	0.49		0.58
		(0.25)	(0.27)	(0.29)	(0.30)		(0.48)
regulated		0.01	0.09	0.13	0.13	0.07	0.21
		(0.18)	(0.18)	(0.20)	(0.20)	(0.19)	(0.25)
Logarithm carbon price						0.16*	
						(0.09)	
Observations	7829	7842	7842	7829	7829	7829	7829
Mean dependent var.	0.39	0.40	0.40	0.39	0.39	0.39	0.14
Sd. of dependent var.	3.56	3.56	3.56	3.56	3.56	3.56	0.81
Pilot dummy	Yes	No	Yes	Yes	Yes	Yes	Yes
Ownership dummy	Yes	No	No	Yes	Yes	Yes	Yes
Size dummy	Yes	No	No	No	Yes	Yes	Yes
R-squared	0.17						
log likelihood	-8240.01	-6964.46	-6784.96	-6537.58	-6431.86	-6424.40	-2896.48
AIC/N	2.11	1.78	1.74	1.68	1.66	1.65	0.75

This table reports OLS and maximum likelihood estimators using a count data model for the sample processed using matching. Column (1) shows the results from the Poisson regression; columns (2)-(7) show the results from the zero-inflated Poisson regression. Columns (2)-(5) show the results for estimating the overall effect of the pilot ETS on innovation. Column (6) shows the estimations on the carbon price elasticity on number of green patents. Column (7) shows the estimations using the number of approved green patents as an outcome variable. Standard errors are clustered at 4-digit sector level, with 268 clusters. Specifications in all the columns include year fixed effects.

* p <0.1, ** p<0.05, *** p<0.01

Therefore, I next present the estimation of the pilot heterogeneity effects using sub-samples of each pilot region.

5.2 The Impact of the Pilot ETS: Heterogeneity and the Direction of Technical Change

Pilot heterogeneity. As described in Section 2, the ETS regulation differs across pilot regions: each of the local Development and Reform Commissions (DRC) decides on its own allowances allocation, the coverage threshold, and which sectors are part of the pilot system. For this reason, effects of the pilots are likely heterogeneous across regions. To assess whether this is the case, I estimate the average treatment effect (ATE) with the baseline regression in Equation 2 for each subsample corresponding to each of the six pilot regions. The Tianjin and Guangdong pilots, however, have relatively few firms, which limits statistical power. For this reason, I additionally estimate specification 7 below, using the full sample. This specification adds a vector of pilot region dummies interacted with the treatment interaction term to 2, which capture any heterogeneity in the effect of the pilot on firm innovation across regions.

$$
\begin{aligned}
y_{it} = exp(\sum_{l=1}^{6} \beta_{1l} \times pilot_l \times regulated_i \times post_t + \sum_{l=1}^{6} \beta_{2l} \times pilot_l \times regulated_i + \\
\sum_{l=1}^{6} \beta_{3l} \times pilot_l \times post_t + \gamma_{i,o} + \delta_{i,size} + \alpha_t + \eta_l) + \epsilon_{it}.
\end{aligned} \tag{7}
$$

In the above specification, $pilot_l$ is the pilot region dummy that equals 1 if a firm i is located in pilot region l.[37] In this regression, β_{1l} is the parameter of interest, representing the regulation effects in region l after the pilot ETS is implemented; β_{2l} captures the average differences among pilot regions of green patent counts between regulated and non-regulated firms; β_{3l} captures the average differences of green patent counts before and after the regulation implementation among pilot regions.

Column (1) in Table 6 reports the estimations of Equation 7 for the heterogeneity effects

[37] Recall that I exclude two pilot regions from this study. This is due to lack of data availability on firms' regulatory status in Chongqing, and late implementation of the regulation in Fujian. Thus the pilot region in this study includes Beijing (BJ), Tianjin (TJ), Shanghai (SH), Hubei (HB), Guangdong (GD), and Shenzhen (SZ).

and columns (2)-(5) report the estimations of the baseline regression (Equation 2) using different sub-samples of pilot regions Beijing, Shanghai, Hubei and Shenzhen. Estimating the effects using the pilot subsamples of Tianjin and Guangdong results in lack of statistical power and low numbers of clusters (390 and 411 observations, and 29 and 26 clusters in the subsamples of Tianjin and Guangdong respectively), I therefore estimate the pilot heterogeneity effects in these two regions using Equation 7 on the full sample.[38] The estimates in column (1) reveal significant effects for green patenting in only one pilot region, Beijing. The estimations in columns (2)-(5) are qualitatively similar to the estimations in column (1) on each respective pilot region, with differing magnitudes.

To better understand the implications of the econometric results for pilot heterogeneity effects in Table 6, I present the marginal effects of the regulation in each of the regions in Figure 6.[39] The marginal effects are positive and significant at the 5% significance level in one region, Beijing, equal to 0.21 more green patents (standard error$=$ 0.1), and marginally significant in Shanghai, equal to 0.23 (standard error$=$ 0.12). One of the reasons for the significant effects is the carbon price: Beijing and Shanghai have the highest and the third highest average carbon prices among all the regions. Although Shenzhen has the second highest average carbon prices, the effect in Shenzhen is not significant.

Next, I estimate continuous treatment effects by the subsamples of pilot regions. Table 7 shows the results. Consistent with the results in Table 6, the increase of carbon prices increases the number of green patents significantly only in Beijing and Shanghai. On average, a 10 percent increase in carbon price is associated with about 4 percent more green innovation both in Beijing and Shanghai. Again, the insignificant estimations of the carbon price elasticity in the extensive margin suggest that only firms in the intensive margin respond to the variation of carbon prices. The effect of carbon pricing on the rest of the pilot regions (Hubei and Shenzhen) is less precisely estimated. The coefficients are positive but not statistically significant; thus it is possible that some regulated firms in these two regions were induced to file more green patents.

[38] The results using the pilot subsamples of Tianjin and Guangdong are not significant and not reported.

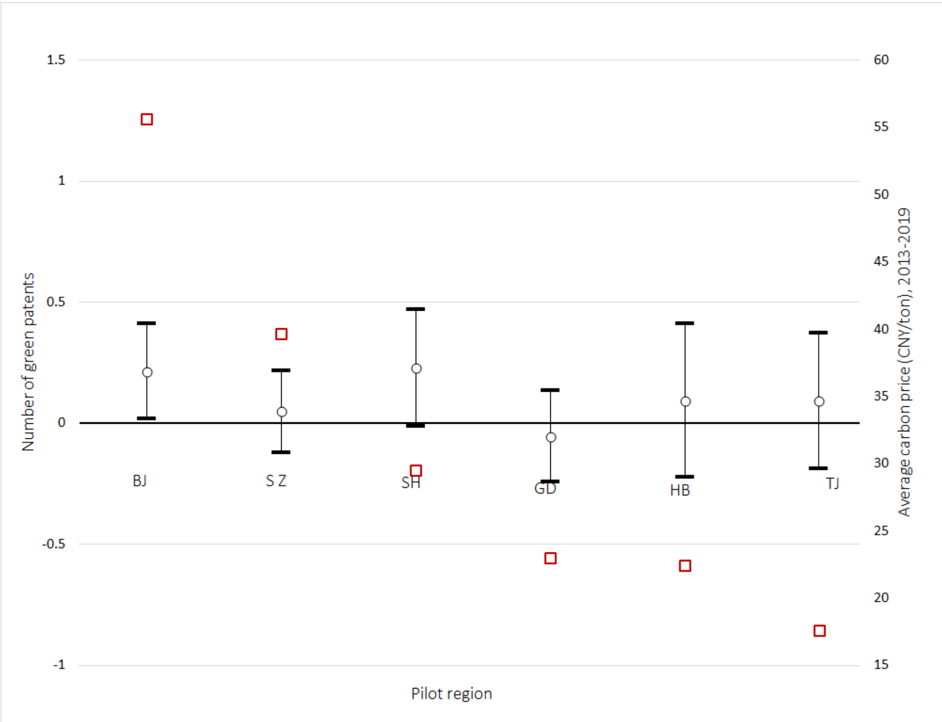
[39] Again, the marginal effects of the regulation on green innovation are calculated by the Stata built-in command `margins`. The marginal effects in Tianjin and Guangdong are obtained using the estimations in column (1).

Table 6: Effect of pilot ETS on green patenting using matched sample, by pilot regions

	(1)	(2)	(3)	(4)	(5)
Green patents, weighted					
regulated*post in BJ	1.72**				
	(0.79)				
regulated*post in TJ	2.29*				
	(1.22)				
regulated*post in SH	1.05				
	(0.76)				
regulated*post in HB	0.42				
	(0.60)				
regulated*post in GD	-0.46				
	(1.55)				
regulated*post in SZ	0.30				
	(0.41)				
regulated*post		1.60**	1.34**	0.47	0.37
		(0.66)	(0.66)	(0.74)	(0.45)
regulated		0.44	-0.82**	-0.35	0.35
		(0.38)	(0.38)	(0.43)	(0.27)
inflate					
regulated*post in BJ	0.78				
	(0.64)				
regulated*post in TJ	2.19				
	(4.11)				
regulated*post in SH	0.21				
	(0.64)				
regulated*post in HB	1.07				
	(0.90)				
regulated*post in GD	0.16				
	(1.61)				
regulated*post in SZ	0.26				
	(0.36)				
regulated*post		1.09	0.60	1.94	0.28
		(0.75)	(0.55)	(1.19)	(0.39)
regulated		0.52	-0.78*	-2.22**	0.29
		(0.52)	(0.46)	(1.04)	(0.32)
Observations	7829	1203	1638	1066	3121
Mean dependent var.	0.39	0.56	0.35	0.20	0.48
Sd. of dependent var.	3.56	7.17	1.94	0.81	3.10
Pilot	Full sample	Beijing	Shanghai	Hubei	Shenzhen
log likelihood	-6425.57	-1087.74	-1176.92	-474.60	-2873.91
AIC/N	1.66	1.88	1.49	0.97	1.87

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample processed using matching. Column (1) shows the results for estimating Equation 7. Columns (2)-(5) show the results for estimating the pilot heterogeneity effects using the sub-samples by regions. Standard errors are clustered at 4-digit sector level, with 268, 93, 111, 88, and 143 clusters respectively in columns (1)-(5). Specifications in all the columns include year fixed effects, ownership dummies and firm size dummies.
* p <0.1, ** p<0.05, *** p<0.01

Figure 6: The ETS heterogeneity effects in pilot regions



Note: The primary vertical axis stands for the effect of ETS on the number of green patents, and the secondary vertical axis is the average carbon price in each pilot region in 2013-2019 with units of Chinese Yuan (CNY)/ton. Along the horizontal axis, from left to right, each point represents one pilot region, with the order of the regions from the highest to the lowest average carbon price in 2013-2019, i.e., BJ for Beijing, SZ for Shenzhen, SH for Shanghai, GD for Guangdong, HB for Hubei, TJ for Tianjin. The lines vertical to the horizontal axis at each of the pilot regions present the regulation marginal effects in different regions respectively, from the estimations in Table 6 with 95% confidence intervals of the marginal effects presented simultaneously. The square markers show the average carbon prices in each of the pilot regions.

Table 7: Effect of pilot ETS on green patenting using matched sample, carbon price elasticity by pilot regions

	(1)	(2)	(3)	(4)
Green patents, weighted				
Logarithm carbon price	0.40**	0.45**	0.17	0.09
	(0.17)	(0.18)	(0.24)	(0.11)
regulated	0.46	-0.91**	-0.36	0.37
	(0.38)	(0.37)	(0.41)	(0.26)
inflate				
Logarithm carbon price	0.28	0.23	0.66	0.08
	(0.19)	(0.16)	(0.43)	(0.10)
regulated	0.53	-0.90**	-2.15**	0.29
	(0.52)	(0.46)	(1.00)	(0.31)
Observations	1203	1638	1066	3121
Mean dependent var.	0.56	0.35	0.20	0.48
Sd. of dependent var.	7.17	1.94	0.81	3.10
Pilot	Beijing	Shanghai	Hubei	Shenzhen
log likelihood	-1088.34	-1173.84	-474.77	-2874.39
AIC/N	1.88	1.48	0.97	1.87

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample processed using matching. Columns (1)-(4) report the estimations on the carbon price elasticity on number of green patents by pilot regions using the carbon price in the same year. Standard errors are clustered at 4-digit sector level, with 93, 111, 88, and 143 clusters respectively in columns (1)-(4). Specifications in all the columns include year fixed effects, ownership dummies and firm size dummies.
* p <0.1, ** p<0.05, *** p<0.01

Firm heterogeneity. Another source of heterogeneity comes from firms that potentially respond differently to the regulation because they have different quantities of inputs available with which to produce innovation. For instance, firms with more capital are able to produce more output and therefore generate more revenue, which leads to more investment, including the R&D investment that is likely to produce more innovation. To capture such a potential indirect effect of the regulation, I use output per worker as a proxy for firms' available inputs on R&D. Output per worker correlates with the capital-labor ratio, which is used as an input in R&D. The output per worker also correlates with firms' productivity, which is largely influenced by technological development. Firms that were already productive before the treatment might continue to have a stronger ability to innovate and be more likely to respond to the regulation.

To test this hypothesis, I add a vector of interaction terms between the firms' output per worker and the regulation dummy in Equation 8. The interaction captures the different patenting ability of firms with different output per worker. I use the data on output and labor in 2012, the year before the implementation of the ETS regulation, to generate the output per worker measure. For firms with missing data in 2012, I use the data from the year between 2007 and 2011 that is closest to 2012. Because output per worker varies greatly by sectors[40], it is more reasonable to compare firms in the same sector. I therefore assign an index from 1 to 4 to all firms based on the output per worker relative to the 4-digit sector average. I then run a ZIP regression with the following specification at the intensive margin:

$$
\begin{aligned}
y_{it} = exp(\sum_{q=1}^{4} \beta_{1q} \times Q_{ij}^{q} \times regulated_i \times post_t + \sum_{l=2}^{4} \beta_{2q} \times Q_{ij}^{q} \times regulated_i \\
+ \sum_{l=1}^{4} \beta_{3l} \times Q_{ij}^{q} \times post_t + \sum_{2}^{4} Q_{ij}^{q} + \beta_5 regulated_i + \gamma_{i,o} + \delta_{i,size} + \alpha_t + \eta_l) + \epsilon_{it}.
\end{aligned}
\tag{8}
$$

In the above specification, q indexes each of the four quartiles of output per worker distribution and Q_{ij}^{q} equals one if firm i in 4-digit industry j belongs to quartile q. The coefficient β_{1q} measures the effect of different quartiles of output per worker on regulated firms.

Estimation of Equation 8 is reported in the first columns of Table 8. The coefficients in column (1) estimated from the ZIP regression imply the following quantitative response in the number of green patents to the pilot ETS: the pilot ETS induces a statistically significant increase in green innovation only in the fourth quartile of the output per worker distribution.[41] Figure 7 presents the average marginal effects of the pilot ETS regulation evaluated for large, medium and small firms[42] and different quartiles of the output per worker distribution. The average marginal effects have higher magnitudes for firms with larger size and yet the effects are significant at the 10 percent significance level only for large firms at the fourth quartile. For a regulated large firm at the fourth quartile of output per worker, the regulation on average increases the number of green patents by 0.34 (standard error= 0.20). However, for a regulated firm at the top quartile of output per worker distribution that files no patents,

[40] For instance, in 2012, the mean of output per worker in the water supply industry was 986 thousand Yuan, while the means in heating supply and electricity supply industries are 5230 and 252,668 thousand Yuan respectively.

[41] As a robustness test, I assign a quintile index instead and find that the effects are significant only in the top quintile of the output per worker distribution, with the coefficient equal to 1.74 and standard error of 0.49. The estimations are not reported.

[42] The firms with miniature size are not considered because there are no regulated miniature firms in the sample.

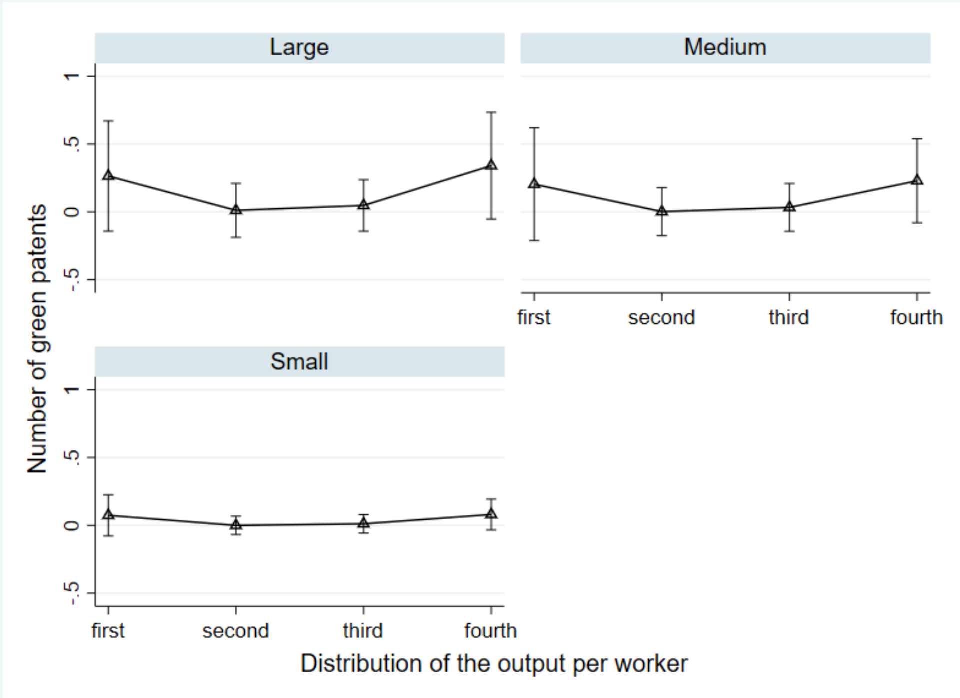
Table 8: Effect of pilot ETS on green patenting and dirty patenting using matched sample

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
main									
first quartile × regulated*post=1	0.73								
	(0.52)								
second quartile × regulated*post=1	0.12								
	(0.35)								
third quartile × regulated*post=1	0.26								
	(0.38)								
fourth quartile × regulated*post=1	1.47***								
	(0.50)								
regulated*post			-0.01	-0.04	-0.01	-0.02*	-0.02	-0.17	0.33*
			(0.02)	(0.04)	(0.01)	(0.01)	(0.01)	(0.56)	(0.17)
regulated	-0.17	-0.63			0.00			0.74	0.16
	(0.31)	(0.45)			(0.01)			(0.57)	(0.16)
first quartile × Logarithm carbon price		0.09							
		(0.13)							
second quartile × Logarithm carbon price		0.06							
		(0.09)							
third quartile × Logarithm carbon price		0.02							
		(0.14)							
fourth quartile × Logarithm carbon price		0.58***							
		(0.21)							
inflate									
first quartile × regulated*post=1	0.51								
	(0.71)								
second quartile × regulated*post=1	0.16								
	(0.34)								
third quartile × regulated*post=1	0.18								
	(0.46)								
fourth quartile × regulated*post=1	1.24**								
	(0.53)								
regulated*post								-0.49	-0.11
								(0.55)	(0.14)
regulated	0.23	-0.54						0.56	0.19**
	(0.30)	(0.45)						(0.57)	(0.09)
first quartile × Logarithm carbon price		-0.02							
		(0.19)							
second quartile × Logarithm carbon price		0.07							
		(0.09)							
third quartile × Logarithm carbon price		0.14							
		(0.15)							
fourth quartile × Logarithm carbon price		0.45**							
		(0.18)							
Observations	7829	7829	7828	1249	7829	7828	4922	7829	7829
Mean dependent var.	0.39	0.39	0.15	0.81	0.06	0.06	0.09	0.10	5.10
Sd. of dependent var.	3.56	3.56	0.35	0.36	0.19	0.19	0.22	1.88	19.83
R-squared			0.30	0.51	0.03	0.25	0.35		
log likelihood	-6323.67	-6291.49						-2087.56	-52978.94
AIC/N	1.64	1.63						0.55	13.55

This table reports maximum likelihood estimators using a zero-inflated Poisson model (columns (1), (2), (8) and (9)), and OLS estimations (columns (3)-(7)) for the sample processed using matching. The columns (1) and (2) show the results for estimating the pilot ETS effects by quartile of firms' output per worker distribution. Columns (3)-(7) show the results from OLS with firm fixed effects (columns (3)-(4), and (6)-(7)) and without (column (5)). The outcome variables are the ratio between the number of green patents and the sum of the numbers of green and dirty patents (columns (3) and (4)), and the ratio between the number of green patents and the number of all the patents (columns (5)-(7)), with (columns (3), (5) and (6)) and without 10^{-6} added (columns (4) and (7)) in the denominator. Column (8) presents the effect of the pilot ETS on dirty patenting. Column (9) presents the effect on the number of patents excluding the green patents. Standard errors are clustered at 4-digit sector level, with 266, 266, 268, 131, 268, 268, 241, 268 and 268 clusters in the eight columns respectively. Specifications in all the columns include year fixed effects; specifications in columns (1), (2), (8) and (9) include pilot fixed effects, firm size dummies, and the ownership dummies.

* p <0.1, ** p<0.05, *** p<0.01

Figure 7: Marginal effects of pilot ETS on green patenting, by firm size



the pilot ETS is associated with a reduction in the likelihood of entry into green technology innovation.[43]

The indirect effect of carbon prices on heterogeneous firms. To capture the indirect effect of carbon prices on firms at different output per worker quartiles, I add an interaction between carbon prices and the quartiles. The intuition is that, for regulated firms in the same pilot region facing identical carbon prices, the firms with distinct output per worker might respond to the regulation differently. To assess this relationship, I replace the discrete treatment dummy with the logarithm carbon prices in year t in the above specification (Equation 8) to allow for heterogeneous effects of carbon price changes on firms at different quartiles. Column (2) presents the indirect effect of output per worker on carbon prices. The estimations address the following response of regulated firms by the number of green patents: for firms located

[43] A related concern is that, for firms in the top quartile of output per worker, the significant increase in the number of green patents in the intensive margin is because the pilot ETS forces these firms that have relatively small amount of green innovation stop innovating. Accordingly, the regulation appears to, on average, increase the number of green patents for firms in the top quartile, and meanwhile decrease the likelihood of entering into green innovation. To address this issue, I drop from the sample the firms in the top quartile that exited green innovation after the implementation of the pilot ETS (about 4 percent of the sample). Then, I estimate the heterogeneity effects using the same regression (Equation 8) on this sample. The results are robust to such an exercise. (not reported) Therefore, the increase of green innovation in the intensive margin is not due to the decrease in the likelihood of entering into green innovation.

in the same pilot region and thus facing the same carbon price level, only firms in the fourth quartile of the output per worker distribution respond to the carbon price increase, which is consistent with what the estimations in column (1) imply. The elasticity of green patents to the carbon price for firms in the fourth quartiles of the output per worker distribution is 0.58. This means that a 10 percent increase in the carbon price will increase the green patents by 5.8 percent for firms in the top quartile. However, in the extensive margin, the increase in carbon prices reduces the likelihood of technological entry into green innovation, especially for firms in the upper range of the output-per-worker distribution.

The direction of technical change. One related question is about the direction of the technological change. Carbon pricing imposes a cost to pollute on the regulated firms, which in turn increases the value of innovation in clean technology. Firms might shift their innovation activities from dirty fossil fuel technology to clean low-carbon technology. To test whether the regulated firms file more green patents at a cost of reducing dirty innovation, I use the share of green patents as an outcome variable, calculated as the ratio between the number of green patents and the sum of the numbers of green and dirty patents, and estimate the ETS effect using the following regression specification:

$$share_{it} = \beta_5 regulated_i \times post_t + \alpha_t + \alpha_i + \epsilon_{it}. \tag{9}$$

In the above specification, $share_{it}$ is the share of green patents. I control for the firm fixed effects α_i and year fixed effects α_t. Around 85 percent of the observations in the sample file neither green nor dirty patents; these need be dropped from the sample, which might potentially leads to a sample selection problem. I therefore add a small number $10^{(-6)}$ to the sum of the green patent counts and dirty patent counts to keep all the observations. Columns (3) and (4) in Table 8 compare whether adding this small number affects the results in a significant way. The insignificant estimations in the two columns suggest that the pilot ETS does not significantly induce the development of technology to a "greener" direction.

Because the pilot ETS increases green innovation without shifting technology in a greener direction, one of the immediate concerns is that the regulation might meanwhile increase the number of dirty patents. Therefore, I estimate the effect of the pilot ETS on the number of dirty patent applications. Column (8) reports the estimations from the ZIP regression. No significant effects of the pilot ETS on dirty innovation are observed in the data. Then, a related concern is that the discrepancy between the insignificant effects on the number of dirty patents and the share of green patents, and the significant effects on the number of green patents, might be driven by time-invariant unobservable firm heterogeneity, which is not accounted for in the ZIP regression. I address this concern by showing in Section 5.4.4 that the estimations on the policy effects are robust to different model specifications including Poisson and OLS regressions with firm fixed effects.

Assessing the crowding-out effect. Another question is whether the regulation might increase green innovation and meanwhile crowd out patents which do not belong to the classification of green patents (non-green patents). To test whether the regulated firms increase green innovation at a cost of other types of innovation, I use the ratio between the number of green patents and the number of all patents filed by a firm in a year as an outcome variable, and estimate the effect on this ratio using regression 9. Columns (6) and (7) show the results. Similarly a small number (10^{-6}) is added to the number of all patents in the ratio in column (6) to avoid dropping observations with zero patents filed in certain years. The estimations are not affected by adding the number and both are negative.[44] To further address the concern about firm specific effects, I compare the estimations on the effects of this share with (column 6) and without firm fixed effects (column 5). The estimation with firm fixed effects is slightly lower; however it is not statistically different from the one without firm fixed effects ($p = 0.58$). Column (9) presents the estimation on the policy impact on the number of patents excluding the green patents. The estimation is positive and significant at the 10 percent significance level. This could be because, for instance, some patents are somewhat related to low-carbon innovation but not counted in the outcome.

5.3 Event-Study Test of Parallel Trends Assumption

The key identifying assumption for the above estimates is that there are parallel pre-regulation trends in the number of green patents for regulated and non-regulated firms. I test this as-

[44] Because of rounding, both estimations seem to be significant. The more precise rounding estimations are -0.016 (standard error=0.009) and -0.017 (standard error=0.011) for specifications (6) and (7), respectively.

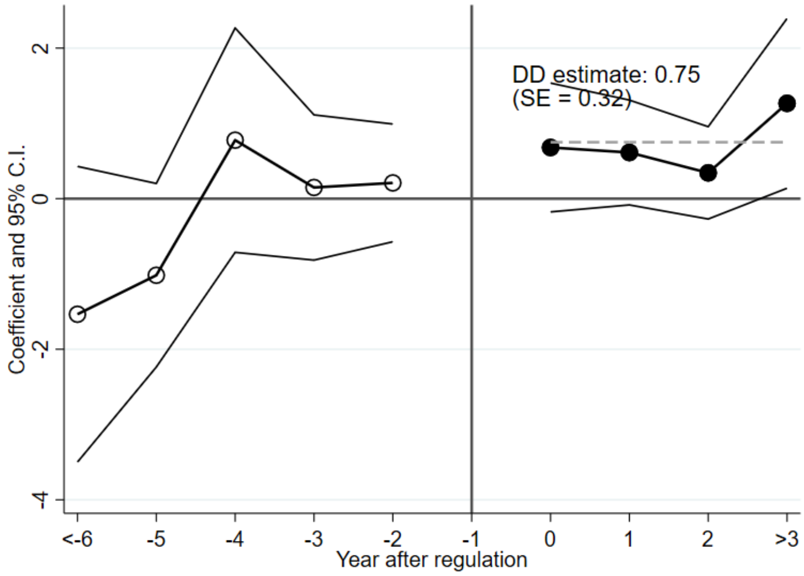
sumption in an event-study specification. That is, I interact the treatment indicator with year dummies leading up to and following the pilot ETS regulation, going from six years before to four years after the regulation. The omitted year is 2012, the year before the first implementation of the regulation. As Figure 8 shows, I do not find any differential green-patenting behavior for regulated versus non-regulated firms in the years leading up to the regulation. The estimations on the leads in the event-study specification are never significantly different from zero. This supports the assumption regarding parallel pre-regulation trends. In addition, this supports the assumption of no anticipatory effects. That is, even though the regulation was announced in 2011, firms did not respond to the regulation as of the year of implementation of the regulation. I discuss this issue in Section 5.4.1. It is worth noting that there is a delay in the policy effect: the effect is only significant in 2016, three years after the implementation of the regulation. This pattern is reasonable because innovation is an ongoing process which requires continuous inputs and has some possibility of failure. One cannot expect an immediate reaction from the regulated firms to the pilot ETS regulation.

As discussed in Section 4.2, one of the caveats of this study is that firms are matched at firm level, not at installation level, because the regulatory threshold is determined at the firm level, as introduced in Section 2.2. A parallel pre-trend reassures to some extent that the matched regulated and non-regulated firms are not systematically different from each other, including their innovating ability and the firms' emissions intensity.

5.4 Robustness Analysis

The baseline results suggest that the regulated firms overall respond to the ETS by innovating slightly more. In addition, I show that the effects are heterogeneous across both pilot regions and firms. The main findings are robust to various specifications. In this section, I report a number of robustness tests. I consider mainly whether the results are driven by self-selection into non-treatment, and whether they are driven by the measurement of the outcome variable. I also consider whether or not there are spillover effects of the regulation, and whether time-invariant firm heterogeneity drives the estimations in a significant way.

Figure 8: Event study of the implementation of the pilot ETS



5.4.1 Are the Results Driven by Self-Selection?

My identifying assumption relies on the fact that the firms cannot self select their regulatory status. As discussed before, one of the main concerns is whether firms are able to influence whether they are regulated. For example, if the cost of abating by reducing productivity is lower than the cost of investing in abatement technology, firms would reduce productivity to comply rather than innovating. Hence, regulated and non-regulated firms would be systematically different from each other. In this case, the estimates would be biased. However, there is little evidence that firms have this power. Since the pilot ETS regulation was announced in 2011 but the coverage threshold was not announced at that time, firms could not obtain information in advance on how the threshold would be set. In other words, firms could not take regulation into consideration when they made decisions on productivity and hence emissions before 2013. Therefore, they could not adopt precautionary measures to strategically avoid being regulated. Moreover, the regulation came into effect in 2013 and remained unchanged until 2016. In 2015, local DRC, except for Tianjin, lowered the coverage threshold significantly for the following years. If regulated, firms just above this threshold before 2016 would behave strategically in order to not be regulated in the following years. They would have to greatly reduce their production, at a cost of losing market share and annual sales. However, purchasing carbon emissions permits from the local carbon market would by no means become a large cost share for regulated firms compared to the cost of reducing productivity, because carbon price in these pilot regions are currently not high. In addition, Figure 8 in Section 5.1 provides the evidence that there is no pre-regulation trend. Therefore, the evidence of having such a self-selection issue is weak.

5.4.2 Are the Results Driven by the Measurement of the Outcome Variable?

All the specifications I show in Section 5.1 use the patent counts weighted by the number of co-applicants on each of the filed patents. My results could be driven by the re-weighting of the patent counts. If the regulated firms co-apply more (less) compared to the non-regulated firms after the implementation of the regulation, my estimation using the weighted patent counts would be lower (higher) than the estimations using the unweighted patent counts. Table 9 presents all the related estimations using the unweighted patent counts as an outcome variable. Column (1) presents the estimation of the overall policy impact. The average marginal effect of ETS on the unweighted number of green patents is 0.17 (standard error= $0.09, p = 0.068$), which is close to the estimation of the effect on the weighted green patent counts. Columns (2)-(5) show the carbon price elasticity on the number of green patents using different carbon price leads. The elasticity of the green patents to the current carbon price is 0.26, which is comparable to the main estimation on the carbon price elasticity of 0.23. The elasticities to carbon prices with leads one to three are less precisely estimated and are all qualitatively comparable to the estimations using the weighted patent counts as an outcome. In summary, the magnitudes of the estimations using the unweighted patent counts are generally slightly higher than the estimations using the weighted counts, but they do not differ significantly. Therefore, the results discussed above are robust to the re-weighting.

Column (6) presents the indirect policy effect through the output per worker. The effects are significant for regulated firms in the first quartile of output per worker, but not for the firms with higher output per worker. The average marginal effects for the firms in the first quartile is 0.38 (standard error= $0.18, p = 0.04$). One potential explanation on the difference of the effects on the weighted and unweighted green patent counts is that the regulated firms in the first quartile co-apply more after the implementation of the regulation, and there is no such an effect for firms with higher output per worker. Columns (7)-(9) show the estimations on the effects of the direction of the technical change. I again use the ratio of the number of green patents and the sum of green and dirty patents as an outcome variable and estimate a fixed-effects OLS model. Columns (7) and (8) present the results. Similarly, in column (7), I add a small number (10^{-6}) to the sum of the counts of the green and dirty patents to avoid dropping the observations that file neither green nor dirty patents. The estimations in columns (7) and (8) are not significantly different and therefore the results are not driven by dropping the observations that filed neither green nor dirty patents. I then estimate the effect on dirty patents with a ZIP and column (9) shows the result. There is no significant effect on dirty patents, though the sign becomes positive. However, the average marginal effects on the weighted and unweighted dirty patent counts are similar at 0.019 (standard error= 0.023) and 0.015 (standard error= 0.026) respectively.

Again, the key identifying assumption is the parallel pre-regulation trends in the unweighted number of green patents for the regulated and non-regulated firms. Figure 12 in Appendix D shows the means of the weighted number of green patents in 2007-2016 by the pilot regions on the matched sample. There is little to no difference in the means between the regulated and non-regulated firms before 2013.

5.4.3 Are There Any Spillover Effects?

In the main analysis, I match the regulated firms and non-regulated firms in the same pilot region. The effects might be under- or over-estimated if the non-regulated firms in the pilot regions also respond to the regulation – for example, to avoid being regulated in the future. To test whether there are such spillover effects of the regulation, I match regulated firms with non-regulated firms outside pilot regions on variables introduced above. If there is no significant difference between the estimations using this sample and the ones in my baseline estimations, I could conclude that non-regulated firms in the pilot regions are not responding to the regulation and the estimations are not biased by spillover effects. Otherwise, if the new estimation results in a higher point estimator, I could conclude that non-regulated firms

Table 9: Effect of pilot ETS on unweighted green patenting using matched sample, count data model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
main									
regulated*post	0.86***						-0.02	-0.06	0.24
	(0.33)						(0.02)	(0.05)	(0.56)
regulated	0.09	0.03	0.17	0.36	0.22	-0.11			0.63
	(0.29)	(0.31)	(0.34)	(0.46)	(0.50)	(0.50)			(0.61)
Logarithm carbon price		0.26**							
		(0.12)							
Logarithm carbon price T+1			0.21*						
			(0.11)						
Logarithm carbon price T+2				0.15					
				(0.12)					
Logarithm carbon price T+3					0.20				
					(0.17)				
first × regulated*post=1						0.74**			
						(0.29)			
second × regulated*post=1						1.11*			
						(0.58)			
third × regulated*post=1						0.98			
						(0.60)			
fourth × regulated*post=1						1.29			
						(0.85)			
inflate									
regulated*post	0.52**								0.07
	(0.26)								(0.50)
regulated	-0.01	-0.12	-0.02	0.15	0.06	-0.17			0.14
	(0.18)	(0.19)	(0.22)	(0.26)	(0.29)	(0.41)			(0.51)
Logarithm carbon price		0.19**							
		(0.08)							
Logarithm carbon price T+1			0.14						
			(0.08)						
Logarithm carbon price T+2				0.10					
				(0.09)					
Logarithm carbon price T+3					0.13				
					(0.12)				
first × regulated*post=1						0.07			
						(0.31)			
second × regulated*post=1						1.18**			
						(0.55)			
third × regulated*post=1						1.06*			
						(0.60)			
fourth × regulated*post=1						0.66			
						(0.69)			
Observations	7129	7129	7129	7129	7129	7129	7129	899	7129
Mean dependent var.	0.36	0.36	0.36	0.36	0.36	0.36	0.12	0.79	0.11
Sd. of dependent var.	3.74	3.74	3.74	3.74	3.74	3.74	0.32	0.38	1.96
R-squared							0.25	0.51	
log likelihood	-5367.54	-5360.02	-5450.49	-5379.60	-5370.85	-5193.10			-1803.54
AIC/N	1.52	1.52	1.54	1.52	1.52	1.49			0.52

This table reports the effect of the pilot ETS on green patenting using the patent counts which are not weighted by the number of co-applicants on each patent. Columns (1)-(6) and (9) show the results from the zero-inflated Poisson regression, with the outcome variables as the green patent counts in columns (1)-(6) and the dirty patent counts in column (9); columns (7) and (8) show the results from OLS regression with the outcome variable as the share of the green patent counts. Column (1) shows the overall effect of the regulation on green patenting; columns (2)-(5) show the the estimations on the carbon price elasticity on number of green patents, with different price leads; column (6) shows the results for estimating the pilot ETS effects by quartile of firms' output per worker distribution; columns (7) and (8) present the estimations of the ETS effects on the share of green patenting; column (9) shows the estimations on the ETS effects on dirty patenting. Standard errors are clustered at 4-digit sector level, with 118 clusters in column (8) and 270 clusters in all the other columns. Specifications in all the columns include year fixed effects; specifications in columns (1)-(6) and (9) include pilot fixed effects, firm size dummies and the ownership dummies. (not reported)

* $p < 0.1$, ** $p<0.05$, *** $p<0.01$

located in a pilot region innovate more than non-regulated firms outside of pilot regions and, therefore, the effects in my baseline estimation are underestimated. By contrast, if the new estimation is lower, I could conclude that non-regulated firms in pilot regions innovate less than firms outside of the pilot regions and thus the effects in my baseline estimation are overestimated.

Columns (1)-(5) in Table 10 report the estimations on the effects of the pilot ETS and the carbon price elasticities with different price leads. The estimations become more precisely estimated in these columns compared to the estimations in Section 5.1, with the magnitudes higher in the estimations in the first three columns. Because I include the region fixed effects to control for the regional unobserved heterogeneity that influence firms' green patenting, I rule out the possibility that the firms in the pilot region systematically file more green patents than the non-pilot regions. The estimations with higher magnitude therefore suggest that the non-regulated firms within the pilot regions also respond somewhat positively to the ETS regulation. A possible explanation is that, to avoid being regulated in the future, given the full information on the regulatory threshold after 2013, the non-regulated firms that have carbon emissions close to the threshold and therefore are more likely to be regulated also increase their green innovation, which potentially helps mitigating carbon emissions.[45] This suggests an underestimation of the policy effects discussed in Section 5.1. I can therefore interpret my estimations as a lower bound of the policy effects. Column (6) presents the estimations on the effect of the pilot ETS on each quartile of output per worker distribution. The pilot ETS induces a statistically significant increase in the number of filed green patents only in the third quartile of the output per worker distribution. The effect on the rest of the quartiles is positive but not statistically significant. This does not necessarily suggest that the result is inconsistent with the baseline estimation, because the matched non-regulated firms which are outside the pilot regions do not belong to exactly the same industries as the matched non-regulated firms which locate in the pilot regions, and the technology development might differ across industries.[46] Columns (7)-(9) present the estimations on the effects on the share of

[45]However this is not empirically testable because of lack of availability on firm-level carbon emissions data.

[46]43 percent of the matched non-regulated firms are in Jiangsu and Zhejiang, and 57 percent in the other 18 provinces. The matched sectors are different across provinces. For instance, 24 percent and 36 percent of the matched non-regulated firms in the chemistry industry and the computer and telecommunications industry are

green patenting and on the number of dirty patents, which are statistically indistinguishable from the respective estimations in columns (3)-(4) and (7) in Table 8.

5.4.4 Are the Results Robust to Controls for Firm-Fixed Effects?

Because there is no standard routine available for estimating the ZIP with fixed effects, as discussed in Section 4.1, the most common practice is to include the pre-sample, post-sample, or in-sample sum of the patent counts as a proxy for the unobserved firm heterogeneity which correlates with firms' innovation ability. This type of method requires either a long pre-sample or post-sample period, or an assumption on the strict exogeneity of the firm-specific effect, which are ruled out because of lack of data or unfulfilled assumptions. To compare whether the unobserved firm heterogeneity drives the results, I use an OLS with firm-fixed effects (FE) as a baseline reference and compare it without firm-fixed effects:

$$y_{it} = \beta_6 regulated_i \times post_t + \alpha_t + \alpha_i + \epsilon_{it}. \tag{10}$$

If, for instance, the estimations with and without FE differ significantly, then the unobserved firm heterogeneity might drive the results upward or downward depending on the difference between the two estimations. Table 15 in Appendix D.2 presents the results. Columns (9) and (10) compare the effects on the number of dirty patents with and without firm fixed effects, and they are not statistically significantly different ($p = 0.30$). However, compared to the estimations on the effects on the number of green patents with firm fixed effects (column (1)), the magnitude of the estimations without such effects (column (7)) is inflated moderately and they are significantly different ($p = 0.05$). This seemingly suggests that the unobserved firm heterogeneity correlates with the ETS effects positively and not accounting for it might lead to an overestimation of the marginal effects of the pilot ETS on green innovation. I therefore estimate a fixed-effects Poisson regression; Table 16 in Appendix D.3 shows the results. Column (2) shows the estimations of the ETS effect from a fixed-effects Poisson model. This estimation suggests that, on average, the pilot ETS increases the number of filed green patents by 0.28, which is higher than the baseline estimation of the average marginal effect

located in Jiangsu.

Table 10: Effect of pilot ETS on green patenting, regulated firms matched with firms outside the pilot regions

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
main									
regulated*post	1.31***						-0.00	-0.03	-0.45
	(0.44)						(0.02)	(0.05)	(0.48)
regulated	-0.87	-1.14**	-0.86*	-0.43	-0.68	-0.68			2.37**
	(0.55)	(0.52)	(0.47)	(0.56)	(0.59)	(1.16)			(1.20)
Logarithm carbon price		0.41***							
		(0.11)							
Logarithm carbon price T+1			0.33***						
			(0.10)						
Logarithm carbon price T+2				0.22*					
				(0.12)					
Logarithm carbon price T+3					0.27**				
					(0.13)				
first quartile × regulated*post=1						0.92			
						(0.69)			
second quartile × regulated*post=1						0.28			
						(0.57)			
third quartile × regulated*post=1						2.17***			
						(0.60)			
fourth quartile × regulated*post=1						1.31			
						(1.16)			
inflate									
regulated*post	1.89**								0.29
	(0.87)								(0.47)
regulated	-0.18	-0.57	-0.16	0.40	0.05	0.60			5.61
	(0.98)	(0.95)	(0.88)	(1.04)	(1.08)	(1.58)			(15.50)
Logarithm carbon price		0.58***							
		(0.21)							
Logarithm carbon price T+1			0.48***						
			(0.17)						
Logarithm carbon price T+2				0.35*					
				(0.20)					
Logarithm carbon price T+3					0.42*				
					(0.22)				
first quartile × regulated*post=1						1.74			
						(1.46)			
second quartile × regulated*post=1						1.46**			
						(0.72)			
third quartile × regulated*post=1						3.96*			
						(2.31)			
fourth quartile × regulated*post=1						1.33			
						(1.48)			
Observations	11985	11985	11985	11985	11985	11980	11966	1223	11985
Mean dependent var.	0.17	0.17	0.17	0.17	0.17	0.17	0.09	0.72	0.09
Sd. of dependent var.	0.95	0.95	0.95	0.95	0.95	0.95	0.29	0.42	0.76
R-squared							0.22	0.47	
log likelihood	-5012.93	-5004.32	-5014.87	-5025.36	-5020.48	-4935.47			-2607.79
AIC/N	0.85	0.85	0.85	0.85	0.85	0.85			0.45

This table reports the effects of the pilot ETS on green patenting using the sample that the regulated firms matched with the non-regulated firms outside the pilot regions. Columns (1)-(6) and (9) show the results from the zero-inflated Poisson regression, with the outcome variables as the green patent counts in columns (1)-(6) and the dirty patent counts in column (9); columns (7) and (8) show the OLS estimators with firm fixed effects, and the outcome variable is the share of the green patent counts. Column (1) shows the overall effect of the regulation on green patenting. Columns (2)-(5) show the the estimations on the carbon price elasticity on number of green patents, with different price leads. Column (6) shows the results for estimating the pilot ETS effects by quartile of firms' output per worker distribution. Columns (7) and (8) present the estimations on the direction of the technological change; column (9) shows the estimations on the ETS effects on dirty patenting. Standard errors are clustered at 4-digit sector level, with 147 clusters in column (8) and 342 clusters in all the other columns. Specifications in all the columns include year fixed effects; specifications in columns (1)-(6) and (9) include pilot fixed effects, firm size dummies and the ownership dummies. (not reported)

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

of 0.16 from the ZIP model. One caveat of this model is that all the firms that have a constant amount of innovation between 2007-2016 are dropped because they are not informative in estimating the model. This is not ideal because half of the matched firms are dropped, which might introduce selection bias. This can potentially lead to an overestimation of the policy effect if, for instance, the comparable treatment firms that file no green patents over time are dropped. Moreover, in Table 15, I present the OLS estimations with firm fixed effects using the same subsample of firms used in the fixed effects Poisson model (column (8)). If the unobserved firm heterogeneity accounted for in the fixed-effects Poisson model indeed drives the estimation, I expect that this estimation that uses partial information (column (8)) differs from the estimation that does not account for the unobserved firm heterogeneity but uses full information (column (7)). Because the two estimations do not differ significantly, I have some confidence that accounting for this unobserved heterogeneity but using partial information is at least not superior than not accounting for the firm heterogeneity but using full information.

6 Conclusion

In this paper, I study the impact of an environmental regulation on technological change in the context of a transitional economy. Specifically, I estimate the effect of China's pilot ETS on firms' green innovation, measured by the number of green patent applications. My main contribution is to study the heterogeneity across regions and firms in the factors that induce technological change. I take into consideration that innovator firms may or may not file patents and therefore distinguish between zero patent counts from innovators and non-innovators. Additionally, I consider innovation decisions at both the intensive margin, i.e., the level of green innovation, and the extensive margin, i.e., whether firms enter into green innovation. Using a zero-inflated Poisson estimation on a uniquely constructed dataset, I find that the ETS regulation induces a small but positive effect on green innovation in those two pilot regions with sufficiently high carbon price, with an upward trend, but no significant effects in the other regions. The effect is most pronounced for large firms and firms in the top quartile of the output per worker distribution. I also estimate a carbon emission price elasticity, showing that a 10 percent increase in the carbon price is associated with a 2.3 percent increase in the number of filed green patents.

These estimation results lead to two main implications. First, this finding adds to the debate on the effectiveness of the pilot ETS in China. Overall, the regulation works effectively in terms of inducing technological change through green innovation. However, the effects are not significant in all pilot regions. One possible explanation is the varying carbon emission prices. Varying prices between different pilot regions reflect regional differences in policy designs, such as allowances allocation, coverage threshold, the sectors being regulated, and the cost of non-compliance (i.e., enforcement and penalties). I show that, on average, the higher the carbon prices, the more green innovation is induced by the pilot ETS. Second, the pilot ETS is advantageous in the intensive margin to the regulated firms that already have high output per worker (and therefore higher productivity and/or more capital) and are likely to be more competitive initially. However, the firms in the top quartile of output per labor are less likely to enter into green innovation if they previously had zero knowledge stock of green innovation. The policy challenge thus is to encourage the regulated firms to start innovation in green technologies, and this is especially important for firms that are larger and more productive. Once they actually start and continue with conducting green innovation, they can potentially be the firms that are the most promising in green technologies.

A major objective of environmental regulation is to reduce pollution at a reasonable cost. The goal can be achieved in several ways, such as fuel-switching, technology diffusion and adoption, or innovation. Further research could explore the policy effects of the spread and adoption of new technology. Also, future research should explore more directly the short-term effectiveness of the pilot ETS, using firm-level carbon emissions data as an outcome variable.

References

Aghion, P., A. Dechezleprêtre, D. Hemous, R. Martin, and J. Van Reenen (2016). Carbon taxes, path dependency, and directed technical change: Evidence from the auto industry. *Journal of Political Economy 124*(1), 1–51.

Anderson, B. and C. Di Maria (2011). Abatement and allocation in the pilot phase of the EU ETS. *Environmental and Resource Economics 48*(1), 83–103.

aus dem Moore, N., P. Großkurth, and M. Themann (2019). Multinational corporations and the EU Emissions Trading System: The specter of asset erosion and creeping deindustrialization. *Journal of Environmental Economics and Management 94*, 1–26.

Baranzini, A., J. C. Van den Bergh, S. Carattini, R. B. Howarth, E. Padilla, and J. Roca (2017). Carbon pricing in climate policy: seven reasons, complementary instruments, and political economy considerations. *Wiley Interdisciplinary Reviews: Climate Change 8*(4), e462.

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

Biglaiser, G. and J. K. Horowitz (1994). Pollution regulation and incentives for pollution-control research. *Journal of Economics & Management Strategy 3*(4), 663–684.

Blundell, R., R. Griffith, and J. V. Reenen (1995). Dynamic count data models of technological innovation. *The Economic Journal 105*(429), 333–344.

Blundell, R., R. Griffith, and J. Van Reenen (1999). Market share, market value and innovation in a panel of British manufacturing firms. *The Review of Economic Studies 66*(3), 529–554.

Blundell, R., R. Griffith, and F. Windmeijer (2002). Individual effects and dynamics in count data models. *Journal of Econometrics 108*(1), 113–131.

Bombardini, M., B. Li, and R. Wang (2017). Import competition and innovation: Evidence from China.

Brandt, L., J. Van Biesebroeck, and Y. Zhang (2014). Challenges of working with the Chinese NBS firm-level data. *China Economic Review 30*, 339–352.

Brunnermeier, S. B. and M. A. Cohen (2003). Determinants of environmental innovation in US manufacturing industries. *Journal of Environmental Economics and Management 45*(2), 278–293.

Bushnell, J. B., H. Chong, and E. T. Mansur (2013). Profiting from regulation: Evidence from the European carbon market. *American Economic Journal: Economic Policy 5*(4), 78–106.

Cai, H. and Q. Liu (2009). Competition and corporate tax avoidance: Evidence from Chinese industrial firms. *The Economic Journal 119*(537), 764–795.

Calel, R. and A. Dechezleprêtre (2016). Environmental policy and directed technological change: evidence from the European carbon market. *Review of Economics and Statistics 98*(1), 173–191.

Chan, H. S. R., S. Li, and F. Zhang (2013). Firm competitiveness and the European Union Emissions Trading Scheme. *Energy Policy 63*, 1056–1064.

Chen, L. (2018). Re-exploring the usage of China's industrial enterprise database (in Chinese). *Economic Review 6 (2018)*(6), 140–153.

Chen, W., X. Chen, C.-T. Hsieh, and Z. Song (2019). A forensic examination of China's national accounts. Technical report, National Bureau of Economic Research.

Cui, J., J. Zhang, and Y. Zheng (2018). Carbon pricing induces innovation: Evidence from China's regional carbon market pilots. In *AEA Papers and Proceedings*, Volume 108, pp. 453–57.

Dang, J. and K. Motohashi (2015). Patent statistics: A good indicator for innovation in China? Patent subsidy program impacts on patent quality. *China Economic Review 35*, 137–155.

Dechezleprêtre, A., C. B. Muckley, and P. Neelakantan (2020). Is firm-level clean or dirty innovation valued more? *The European Journal of Finance 0*(0), 1–31.

Di Maria, C. and S. Smulders (2017). A paler shade of green: Environmental policy under induced technical change. *European Economic Review 99*, 151–169.

Diamond, A. and J. S. Sekhon (2013). Genetic matching for estimating causal effects: A general multivariate matching method for achieving balance in observational studies. *Review of Economics and Statistics 95*(3), 932–945.

Feenstra, R. C., Z. Li, and M. Yu (2014). Exports and credit constraints under incomplete information: Theory and evidence from china. *Review of Economics and Statistics 96*(4), 729–744.

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

Fischer, C., I. W. Parry, and W. A. Pizer (2003). Instrument choice for environmental protection when technological innovation is endogenous. *Journal of Environmental Economics and Management 45*(3), 523–545.

Gray, W. B. and R. J. Shadbegian (1998). Environmental regulation, investment timing, and technology choice. *The Journal of Industrial Economics 46*(2), 235–256.

Hausman, J., B. H. Hall, and Z. Griliches (1984). Econometric models for count data with an application to the patents-R&D relationship. *Econometrica 52*(4), 909–938.

He, Z.-L., T. W. Tong, Y. Zhang, and W. He (2018). A database linking Chinese patents to China's census firms. *Scientific Data 5*, 180042.

Hu, A. G. and G. H. Jefferson (2009). A great wall of patents: What is behind China's recent patent explosion? *Journal of Development Economics 90*(1), 57–68.

Hu, A. G., P. Zhang, and L. Zhao (2017). China as number one? Evidence from China's most recent patenting surge. *Journal of Development Economics 124*, 107–119.

Iacus, S. M., G. King, and G. Porro (2012). Causal inference without balance checking: Coarsened exact matching. *Political Analysis 20*(1), 1–24.

ICAP (2018). ICAP ETS Map. Technical report, ICAP.

IPC Committee (2017). IPC Green Inventory. Technical report, IPC Committee.

King, G. and R. Nielsen (2015). Why propensity scores should not be used for matching. *Copy at http://jmp/1sexgVw Export BibTex Tagged XML Download Paper 481*.

Lambert, D. (1992). Zero-inflated Poisson regression, with an application to defects in manufacturing. *Technometrics 34*(1), 1–14.

Le Quéré, C., R. M. Andrew, P. Friedlingstein, S. Sitch, J. Pongratz, A. C. Manning, J. I. Korsbakken, G. P. Peters, J. G. Canadell, R. B. Jackson, et al. (2017). Global carbon budget 2017. *Earth System Science Data Discussions*, 1–79.

Li, X. (2012). Behind the recent surge of chinese patenting: An institutional view. *Research Policy 41*(1), 236–249.

Liu, Q. and L. D. Qiu (2016). Intermediate input imports and innovations: Evidence from Chinese firms' patent filings. *Journal of International Economics 103*, 166–183.

Mielke, P. W. and K. J. Berry (2007). *Permutation methods: a distance function approach.* Springer Science & Business Media.

Mullahy, J. (1986). Specification and testing of some modified count data models. *Journal of Econometrics 33*(3), 341–365.

Noailly, J. and R. Smeets (2015). Directing technical change from fossil-fuel to renewable energy innovation: An application using firm-level patent data. *Journal of Environmental Economics and Management 72*, 15–37.

Petrick, S. and U. J. Wagner (2014). The impact of carbon trading on industry: Evidence from German manufacturing firms. *Available at SSRN 2389800.*

Popp, D. (2003). Pollution control innovations and the Clean Air Act of 1990. *Journal of Policy Analysis and Management 22*(4), 641–660.

Requate, T. (2005). Dynamic incentives by environmental policy instruments—a survey. *Ecological Economics 54*(2-3), 175–195.

Requate, T. and W. Unold (2003). Environmental policy incentives to adopt advanced abatement technology: Will the true ranking please stand up? *European Economic Review 47*(1), 125–146.

Stuart, E. A. (2010). Matching methods for causal inference: A review and a look forward. *Statistical science: a review journal of the Institute of Mathematical Statistics 25*(1), 1.

Tang, R. (2015). Environmental regulation and firm innovation: Evidence from China. Technical report, Working Paper, University of Alberta.

Zhu, J., Y. Fan, X. Deng, and L. Xue (2019). Low-carbon innovation induced by emissions trading in China. *Nature Communications 10*(1), 1–8.

Appendix

A Additional Institutional Detail

A.1 Allowances Allocation

Grandfathering refers to a practice whereby future free emission allowances are dependent on past emissions or emission intensity. Specifically, grandfathering emission intensity determines the allowances in such a way that future allowances are in proportion to the emission intensity of an entity, while grandfathering emission requires that future allowances are in proportion to average yearly emission of an entity in a certain period. Benchmarking determines the allowances based on an emission benchmark of an industry, as well as firms' annual production. The difference of allowances allocation matters because allocating allowances overly generously dampen firms' incentives to adjust production plans to adapt to the regulation, and thus offers little incentive to innovate.

- Beijing: For heating companies and thermal power companies, allowances are allocated based on grandfathering emission intensity; for firms in industries other than heating and thermal power, allowances are allocated based on grandfathering yearly average emission between 2009 and 2012.

- Shanghai: Allowances are allocated based on benchmarking for the power and heating industries. For industries such as aviation, ports and waterway transportation, grandfathering is based on emission intensity. For those in commercial industries, hotels and airports, and firms for which it is hard to measure production, it is difficult to use industry benchmarking or emission intensity grandfathering, and therefore grandfathering based on historical emissions is adopted.

- Shenzhen: Government can repurchase allowances, at most 10% of total allowances, to stabilize the market price. Taking into consideration the annual decrease rate of carbon intensity, allowances are allocated based on grandfathering emission intensity for all firms regardless of industry. This annual decrease rate is formulated by Shenzhen DRC.

- Chongqing: Annual allowances are the same as reported emissions (RE) if total RE is smaller than an upper limit of total allowances. The upper limit of allowances is determined by the maximum yearly carbon emissions (YCE) between 2008 and 2012. Before 2015, this is decreased by 4.13% yearly; after 2015, this is determined by the central government's mitigation goal. If total RE is larger than the upper limit of allowances, allowances are allocated based on both reported emissions and historical maximum emissions between 2008 and 2012. [47]

- Tianjin: Allowances are allocated mainly for free through grandfathering based on emissions from 2009 to 2012 or emission intensity. Benchmarking is adopted for new entrants and expanding capacity. Auction or purchasing at fixed prices may be implemented to stabilize the allowance price in case of acute fluctuations in market prices. Tianjin DRC did not publish clear guideline for how each industry's allowances would be allocated.

- Hubei: For firms in the power-generation industry, allowances are allocated using benchmarking; for firms in industries other than power generation, allowances are allocated using grandfathering based on average emissions of the last three years. The allowances allocation method in 2017 has been changed. Allowances for firms in the cement, power generation, and heating industries are allocated using benchmarking, while for firms in the paper, glass and ceramic industries, allowances are allocated using grandfathering

[47] See Chongqing DRC for more details.http://www.cqdpc.gov.cn/c/2014-05-29/521437.shtml

based on emission intensity. Allowances for all the other regulated firms are allocated using grandfathering based on emission intensity.

- Guangdong: 95% of the allowances are allocated for free to the power-generation industry, while 97% of the allowances are allocated for free to the other industries. Benchmarking is adopted for firms in the coal-fired and gas-fired generation, cement, steel, paper and aviation industries. For other industries, the allowances are based on grandfathering and reduction of historical emission intensity.
- Fujian: All allowances are allocated for free in the first year. The threshold of being regulated is decreasing to 5,000 tons yearly carbon emissions (YCE) gradually. Meanwhile, aiming to introduce an allowances auction over time, the share of free allowances will be reduced. Similar to the allowances allocation method in Shanghai, benchmarking, grandfathering based on emissions, or emissions intensity are adopted for different sectors and industries.

A.2 Coverage Threshold

- Beijing: On November, 20th, 2013, Beijing Municipal DRC announced that entities with YCE higher than 10,000 tons (including both direct and indirect emissions [48]) are regulated in the scheme. On December, 16, 2015, this threshold was adjusted to target those with YCE higher than or equal to 5,000 tons. Emission-reporting entities are those who have consumed more than 2,000 tons coal equivalent (tce) of energy.
- Shanghai: For the first period (2013-2015), firms in industrial sectors such as iron and steel, petrochemicals, chemicals, non-ferrous metals, power generation, building materials, and paper, textiles, rubber and chemical fiber with YCE higher than or equal to 20,000 tons in either 2010 or 2011 are regulated in the pilot scheme. In contrast, firms in non-industrial sectors such as aviation, ports, airports, railways, commercial sectors, hotels and finance with YCE higher than or equal to 10,000 tons in either 2010 or 2011

[48] Direct emissions refer to the emissions generated during the production process by burning fossil fuels. Indirect emissions refer to emissions related to the use of purchased electricity and heating. Direct and indirect emissions are counted in all eight pilot regions.

are regulated in the pilot scheme [49].

As to the second period (2016-now), firms in industrial sectors that were not regulated in the first period, with YCE higher than or equal to 20,000 tons, are added to list of regulated firms. In addition, firms regulated from 2013 to 2015 with YCE higher than 10,000 tons are covered in the second period. Transportation sectors such as ports and aviation with YCE higher than 10,000 tons and waterway transportation with YCE higher than 100,000 tons are regulated in the pilot ETS.

- Shenzhen: Firms with YCE higher than 3,000 tons in any year are regulated. Firms with YCE higher than 1,000 tons and lower than 3,000 tons in any year are responsible for reporting carbon emissions annually. Moreover, Shenzhen Municipality also requires that owners of buildings for public affairs and national authority offices with area exceeding 10,000 square meters should be regulated in the pilot ETS as well.

- Chongqing: Before 2015, industrial firms with CO_2 equivalent higher than 20,000 tons in years between 2008 and 2012 are regulated. It is noteworthy that the Chongqing ETS is the only pilot that covers six greenhouse gases (GHGs) including CO_2, CH_4, N_2O, $HFCs$, $PFCs$ and SF_4. All other seven pilot regions only regulate firms on CO_2 emissions.

- Tianjin: Firms in steel, power generation, heating, petrochemical, oil and gas exploration and construction industries with YCE higher than 20,000 tons in years between 2009 and 2012 are regulated.

- Hubei: Industrial firms with energy consumption exceeding 60,000 tons coal equivalent (tce) in either 2010 or 2011 are regulated in 2014, the starting year of the pilot ETS in Hubei. In contrast, in 2015, this time horizon changed to any year between 2009 and 2014. In 2016, the coverage became broader. Firms in the "seven industries"[50] with yearly energy consumption higher than 10,000 tce, as well as industrial firms in an

[49] See Municipal Government's Opinions on Pilot ETS in Shanghai published by Shanghai Municipal People's Government in July, 3, 2012. URL http://www.shanghai.gov.cn/nw2/nw2314/nw2319/nw10800/nw11407/nw29273/u26aw32789.html (in Chinese).

[50] These are the iron and steel, petrochemicals, chemicals, non-ferrous metals, power generation, building materials, and paper industries.

industry other than the "seven industries" with yearly energy consumption exceeding 60,000 tce, in any year between 2013 and 2015, are regulated in the Hubei ETS. The coverage threshold is even stricter in 2017. All industrial firms with energy consumption exceeding 10,000 tce in any year between 2014 and 2016 are regulated.

- Guangdong: At the beginning of the pilot ETS, firms or entities in the industries of power generation, cement and petrochemical, with YCE higher than 20,000 tons, are regulated in the pilot; firms in the above industries with YCE higher than 10,000 tons are defined as emission-reporting entities. Starting from 2016, firms or entities in the paper and aviation industries satisfying the above coverage criteria are regulated as well [51].

- Fujian: Firms or entities with total energy consumption higher than 10,000 tce in any year between 2013 and 2015, in the seven industries and industries of aviation and ceramics, are regulated in the Fujian pilot ETS.

  Fujian is famous for both the productivity and quality of ceramics and this industry contributes a large share of CO_2 emissions. Therefore, the Fujian DRC regulates the ceramics industry. There are 119 firms in ceramics among the 277 regulated firms.

A.3 Punishment

If the cost of non-compliance is lower than the cost of technology development, emission reduction and purchasing allowances, firms tend to disregard the mitigation responsibility and to not take carbon emission into consideration in production planning. In this case, it is necessary to increase the cost of non-compliance. Pilot firms are punished if they emit more than the verified allocated allowances. Specifically,

- Beijing: Firms are fined for excess emissions at three to five times the average allowance price for the past year.

- Shanghai: Firms are fined at between 50,000 Yuan and at the highest 100,000 Yuan.

[51] There were 4 firms in the aviation industry and 51 firms in paper industries newly added as regulated firms. See Summary of Allowances Allocation Method for Aviation and Paper Industries in Guangdong ETS http://www.gddrc.gov.cn/zwgk/zcwj/zcjd/201712/t20171229_458124.shtml for more details (in Chinese).

- Shenzhen: The amount of excess emissions is deducted from next year's allowance; firms are fined for excess emissions at three times the average allowance price for the last six months.

- Chongqing: Firms are not allowed to receive subsidies for energy-saving and climate-change related projects for three years. For state-owned companies, the irregularities are recorded in the Performance Appraisal System for State-owned Enterprise Leaders.

- Tianjin: Firms can not be financially supported for the next three years.

- Hubei: Two times the excess emissions are deducted from next year's allowances; firms are fined for excess emissions up to three times the average allowance price for the last year (however, no more than 150,000 Yuan).

- Guangdong: Two times the excess emissions are deducted from next year's allowances; firms are fined 50,000 Yuan.

- Fujian: Two times the excess emissions are deducted from next year's allowances; firms are fined for excess emissions up to three times the average allowance price for the last year (however, no more than 30,000 Yuan).

The stringency of punishment varies to a certain degree among different pilot regions. For instance, there is an upper limit of fines in Fujian, Guangdong, Shanghai and Hubei, which makes the punishment less harsh. Meanwhile, there is an allowance deduction for the next years in different degrees in Fujian, Guangdong, Hubei and Shenzhen if firms fail to stay within their allowances. In contrast, in Tianjin and Chongqing, firms are punished only by not being able to get subsidies or financial support for their projects. Notably, there is no upper limit of the fine in Shenzhen and Beijing, signalling that punishment is harsher than in cities that determine a specific upper limit on the fine [52].

[52]For example, two firms in Shenzhen failed to achieve their 2016 mitigation liability; therefore they were fined 1,540,000 Yuan and 1,220,000 Yuan respectively. See http://www.szpb.gov.cn/xxgk/qt/tzgg/201709/t20170914_8689504.htmand http://www.szpb.gov.cn/xxgk/qt/tzgg/201709/t20170914_8689503.htm for details of penalty decisions (in Chinese).

Table 11: Government Plans and Interim Measures in Eight Pilot ETSs

Pilot	Document	Time
Beijing	Implementation Plan of Beijing ETS (Beijing DRC)	20 November 2013
	Threshold Adjustment on Beijing ETS (Beijing Municipal People's Government)	28 December 2015
	2017 Beijing ETS Plan (Beijing DRC)	15 December 2016
Shanghai	Implementation Plan of Shanghai ETS (Shanghai Municipal People's Government)	3 July 2012
	Interim Management Measures of Shanghai ETS (Shanghai Municipal People's Government)	18 November 2013
	Allowances Allocation Plans (Shanghai DRC, 2016 and 2017)	10 November 2016 and 20 December 2017
Shenzhen	Interim Management Measures of Shenzhen ETS (Shenzhen Municipal People's Government)	19 March 2014
Chongqing	Interim Management Measures of Choingqing ETS (Chongqing Municipal People's Government)	26 April 2014
	Allowance Allocation Plans (Chongqing DRC)	29 May 2014
	Interim Rules of Carbon Emission Verification (Chongqing DRC)	29 May 2014
Tianjin	Implementation Plan of Tianjin ETS (Tianjin Municipal People's Government)	5 February 2013
	Interim Management Measures of Tianjin ETS (Tianjin Municipal People's Government, 2013 and 2016)	20 December 2013 and 3 March 2016
	Tianjin ETS China Certified Emission Reduction (CCER) (Tianjin DRC)	9 July 2015
Hubei	Implementation Plan of Hubei ETS (Hubei Provincial People's Government)	18 February 2013
	Allowances Allocation Plans (Hubei DRC, 2014, 2015, 2016 and 2017)	14 April 2014, 25 November 2015, 3 Jaunuary 2017 and 10 January 2018
	Interim Rules for Allowances Launch and Repurchase (Hubei DRC)	29 September 2015
	Hubei ETS CCER (Hubei DRC, 2016 and 2017)	8 July 2016 and 13 June 2017
Guangdong	Interim Management Measures of Guangdong ETS (Guangdong Provincial People's Government)	1 March 2014
	Allowance Allocation Plans (Guangdong DRC, 2014, 2015, 2016 and 2017)	8 August 2014, 10 July 2015, 8 July 2016 and 25 August 2017
	Allowances Verification and Compliance (Guangdong DRC, 2015, 2016 and 2017)	18 February 2016, 22 February 2017 and 12 February 2018
Fujian	Interim Management Measures of Fujian ETS (Fujian Provincial People's Government)	22 September 2016
	Interim Management Measures of Fujian GHGs Reporting (Fujian DRC)	30 November 2016
	Interim Implementation Measures of Fujian ETS (Fujian DRC)	2 December 2016

A.4 Measures and Plans

Principles of determining coverage threshold and punishment measures are from government official plans and plans. I will summarize the measures and plans in this section [53]. As listed in Table 11, government and DRC in some pilot regions, such as Hubei and Guangdong, release Allowances Allocation Plans annually, while DRC in some pilot regions, such as Chongqing and Shenzhen, merely released the plans once, when they were about to implement pilot ETS.

B Steps of Merging Datasets

He et al. (2018) discuss technical details on merging each firms' data in ASME with patent data in SIPO. I follow the main steps discussed in that paper but the principles used for merging the two datasets are more comprehensive; for example, I use a broader list to define a company as a patent applicant. In order to construct the dataset, I first of all merge firms' information in ASME with each of the patent applications in SIPO using the entity name in ASME and the applicant in SIPO, and then merge the dataset with regulatory status.

[53] Reference URLs for all the plans and measures in Table 11 can be given upon request.

B.1 Preparation of SIPO and ASME

First of all, I drop those patents applied for by individuals. There are two related caveats about SIPO. As this dataset is accessed via web-scraping, and there is often at least one co-applicant for each patent application, the information about applicants is scraped as one column regardless of the number of applicants for each application. Moreover, patent applications in SIPO are not specifically categorized as different types of applicants, for example, individual, educational institution and company. Therefore, I first of all need to drop those applied for by individuals and only keep applications for which there is at least one applicant which is non-individual.

I determine whether the applicant is an individual or a company based on three criteria. First, if the length of the name of the applicant is no longer than two Chinese characters, the applicant is an individual, as it is not likely that a firm's name only has two Chinese characters. This criteria rules out applications with a single individual applicant if the applicants' name is at most two Chinese characters. Next, patent applicants contain at least one non-individual applicant if the applicant ends with characters [54]such as station, plant, bureau, department, or school. Moreover, if applicants contain the characters[55], for instance, university, academy, laboratory, hospital, headquarter, park, supermarket, trading, organization, committee, or group, they are considered as non-individual.

After dropping all patents applied for by individuals, I duplicate each patent such that one applicant takes one row. For example, if there are five applicants for a certain patent, there are five observations for the same patent application.

The only identifier available for merging the two datasets is the firm's name. I use stem name in two datasets to do the matching. First of all, I remove all the punctuation[56] in firms' name for both datasets. Then, words specifying firms' type and ownership are removed, for example, group, board, branch, limited company.

Because of administrative error which leads to misreporting by firms, there are potential measurement errors in the variables in this dataset. Following Cai and Liu (2009) and Feenstra

[54]Full list contains 19 different characters. This can be given upon request.
[55]The full list contains more than 200 key words.
[56]Punctuation includes parentheses, brackets, slash, comma, and space, as well as some other marks.

et al. (2014), I clean the data using the following criteria to obtain a clean sample. First, the total assets must be higher than the current assets and fixed assets, as well as the net value of the fixed assets; second, the year of incorporation must be earlier than the year that the data were surveyed, and the opening month must be between 1 and 12; third, the interest expenses must be non-negative; fourth, the firms that have fewer than 8 employees are dropped.

B.2 Merging and Post-Merging Validation

In order to not to lose information, I apply ever-matching, which only requires the ASME firm name and the patent applicants' name to be matched, irrespective of the year in which the firm appears in the ASME database or is filed with the SIPO. Moreover, firm name and patent applicants are matched as long as the ASME firm name is a left-aligned strict substring of the patent applicant's name.

Thereafter, I conducted a post-merging check to validate whether or not matched pairs are true matches [57]. Post-matching validation was checked in Python. There were 1,628,058 true matches after running the checking algorithm, while 383,058 matches required manual checking. I compiled them into 20,444 unique pairs of firm name and patent applicant's name. Then I checked these pairs manually. If the applicant is a subsidiary of the firm in ASME, it is considered as a true match. The match is also considered as a true match if there is no obvious reason indicating that they are not. By manually checking the matches, I find that 129,225 matches are not true matches, while 253,833 pairs are true matches.

C Coarsened Exact Matching and Genetic Matching

Rather than adopting the more commonly used and more conventional propensity score matching (PSM) techniques, I use coarsened exact matching (CEM), in combination with genetic matching (GM). One of the main caveats of PSM is that by projecting a number of covariates to a scalar propensity score and trying many models before choosing one to present, the data generation process is rarely known. Hence PSM potentially increases model dependence and imbalance on matching variables (King and Nielsen, 2015). In contrast, CEM operates on the

[57] The method is discussed in He et al. (2018).

same metric as the original data and thus obeys the congruence principle.[58] Methods violating this principle lead to less robust inferences (Mielke and Berry, 2007).

The intuition of CEM is that, by choosing a certain value for each matching variable (defined as the coarsening in CEM, which describes how rough the matching is), observations are assigned with the same numerical value of strata if they are in the same coarsened strata. In other words, for a certain variable, the coarsening splits the variable into several intervals. Then observations in the same interval are assigned with the same numerical value of strata. Hence, CEM defines a number of strata based on the coarsening, and the observations in the same strata are grouped together. Then, matches are determined by exact matching on the numerical value of the strata. Using CEM, I can match based on the distribution of matching variables rather than the absolute distance defined by a certain caliper. Moreover, CEM is particularly appropriate in this study because the distributions of the matching variables, such as the number of patent applications, are highly right-skewed. Also, CEM prunes few unmatched treated units if there is a large number of control units in the dataset (Iacus et al., 2012). In addition, with CEM it is guaranteed that all variables are balanced on all higher order moments and interactions. Therefore, unlike propensity score matching, this method requires no checks on the balance of interactions on matching variables.[59]

CEM has the many advantages discussed above only if the coarsening is chosen based on substantive criteria. However, concerns can arise if the coarsening is set more arbitrarily. A reasonable argument with meaningful economic sense is thus important for the choice of coarsening. I address this potential threat in two aspects. First of all, I define the coarsening according to the statistical size of firms in China announced in *the Measures for Classification of Large, Medium, Small and Miniature Enterprises* by the National Bureau of Statistics.[60] With the

[58]Methods that violate the congruence principle include, for instance, propensity score matching and Mahalanobis distance matching. Both methods project the covariates from the k-dimensional space in the original data to one space defined by propensity score or Mahalanobis distance metrics. (Iacus et al., 2012)

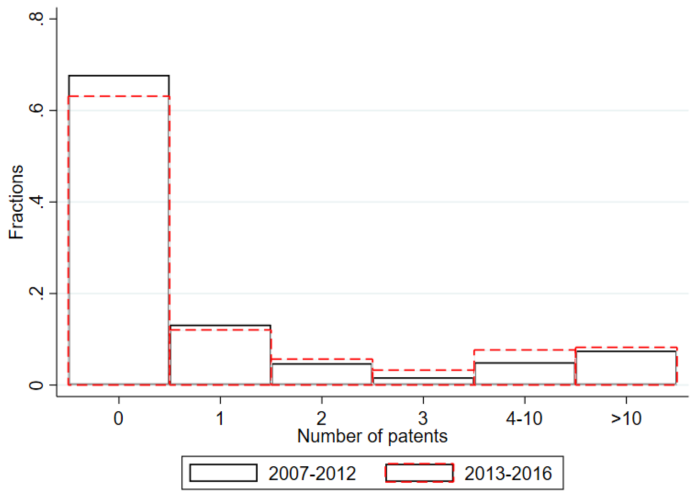
[59]CEM also reduces model dependence and is computationally efficient. For a detailed discussion, see Iacus et al. (2012).

[60]Large firms have annual sales higher than 400,000 thousand yuan and more than 1,000 employees; firms with sales between 20,000 thousand yuan and 400,000 thousand yuan and between 300 and 1,000 employees are medium-size; small firms have between 20 and 300 employees and sales between 3,000 and 20,000 thousand yuan; miniature firms have fewer than 20 employees and 3,000 thousand yuan. All the large, medium and small firms must fulfill both criteria for sales and employees; otherwise, the firms would be classified as one level lower. See http://www.stats.gov.cn/tjsj/tjbz/201109/t20110909_8669.html for a full list of classifications.

defined coarsening, firms in the same pilot region and sector with the same statistical size are assigned to the same CEM stratum. Next, I complement the CEM with genetic matching (GM) to improve the balance between regulated and non-regulated firms in the pre-treatment period and to reduce the model dependence. That is, I run GM within each CEM stratum to assure that both the congruence principle and the monotonic imbalance bounding are satisfied. Firms within the same CEM strata are matched on the number of filed green patents and the number of all filed patents between 2007 and 2012, the dummy for whether or not a firm filed at least one patent before 2013, and average sales and the employment between 2007 and 2012.[61] The main advantage of GM is that it directly optimizes covariate balance and avoids iterative manual checking on the estimated propensity score.

D Additional Empirical Results

Figure 9: Number of green patents: fraction distribution



[61] Again, for sales and employment, data in 2010 are excluded from the study due to poor quality of the data in this year, as discussed in Section 3.

Figure 10: Averages of weighted granted green patents 2007-2016, matched sample

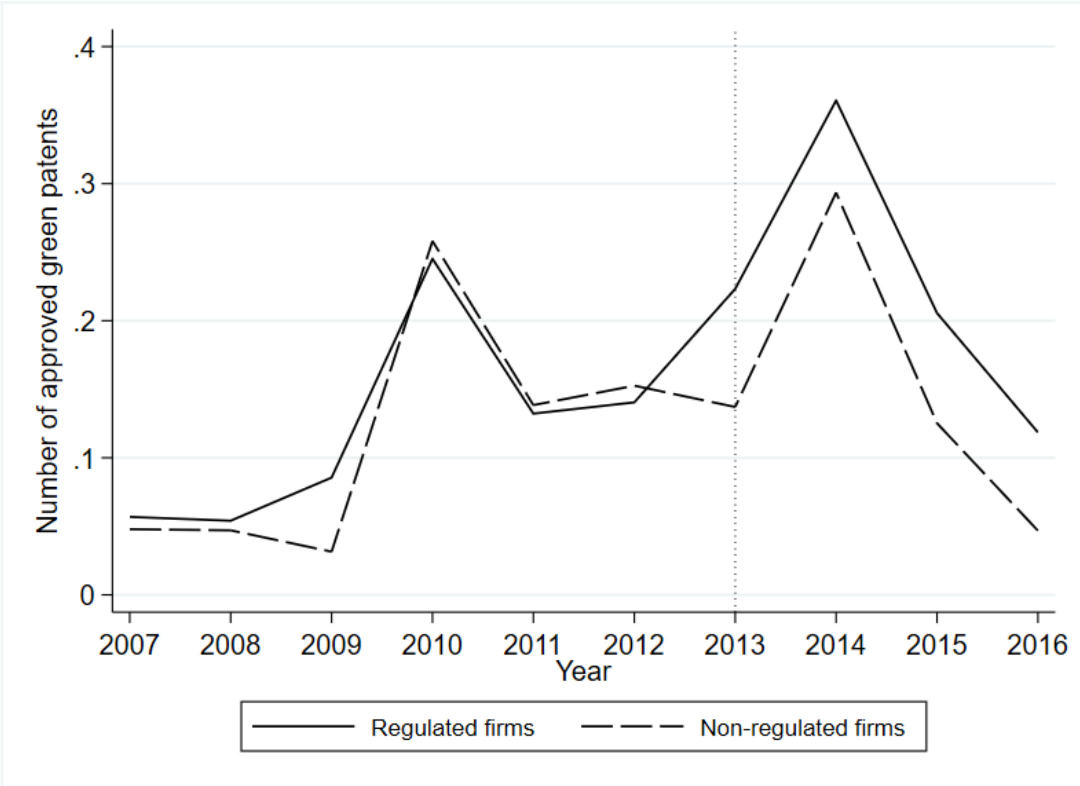


Figure 11: Averages of weighted green patents 2007-2016 by pilot region, matched sample

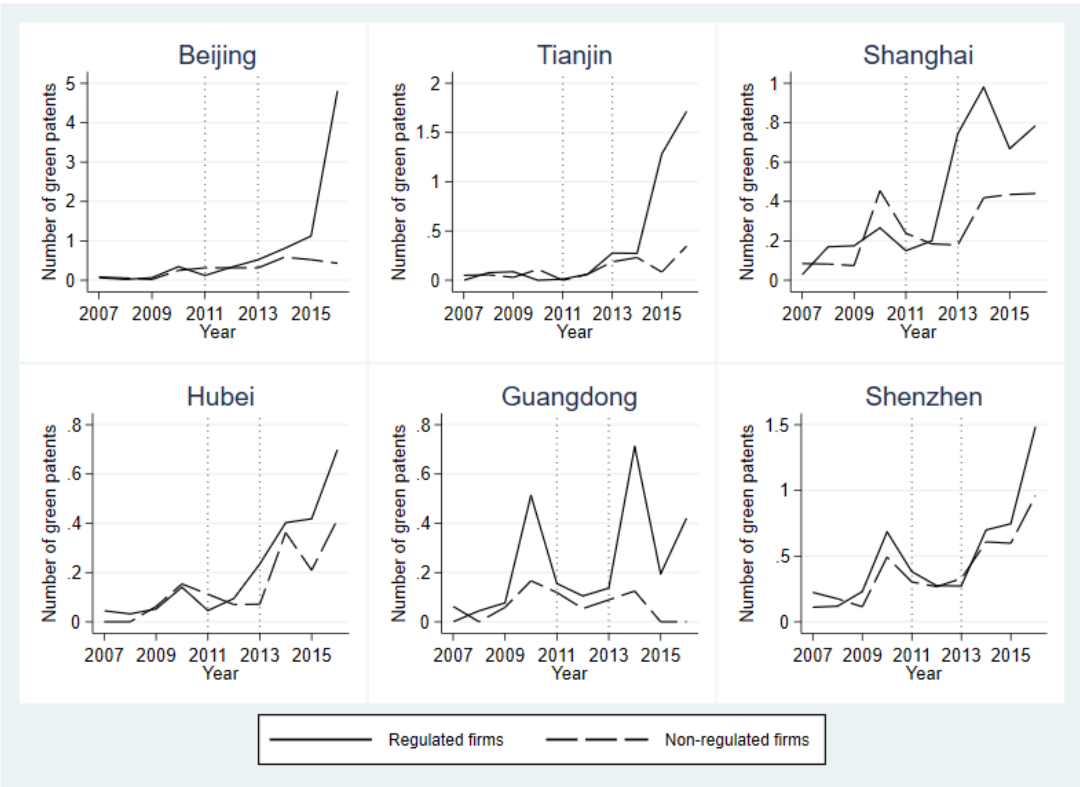


Figure 12: Averages of unweighted green patents 2007-2016 by pilot region, matched sample

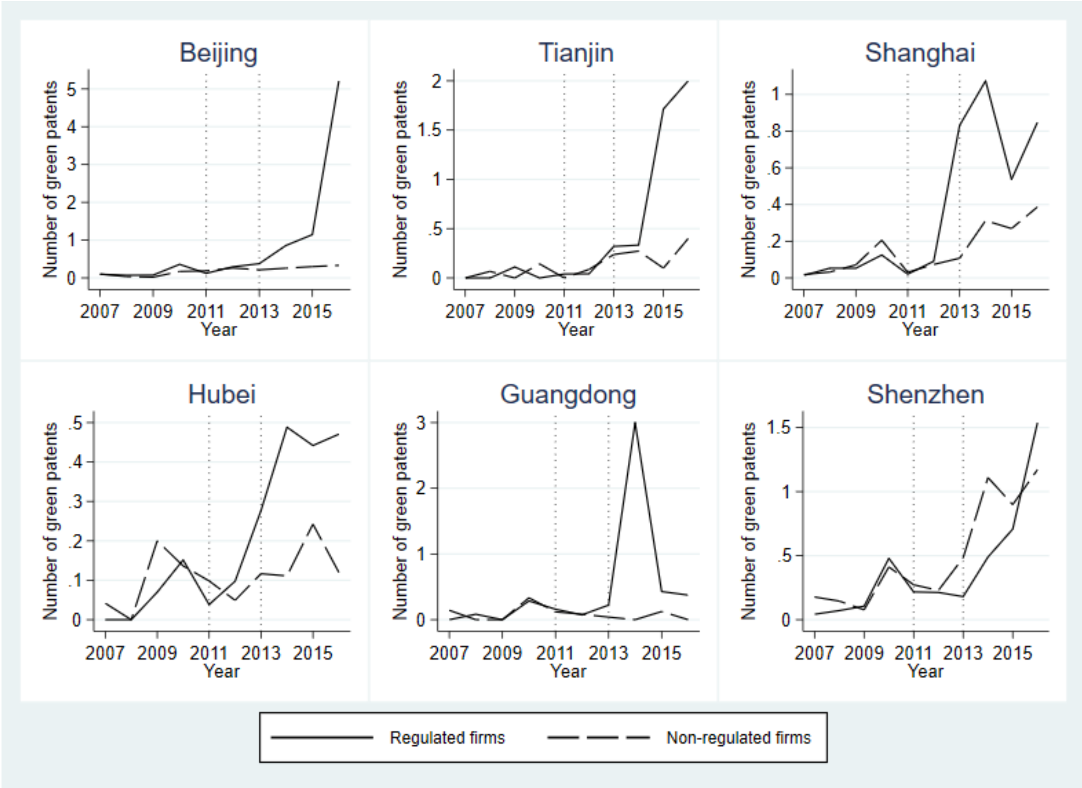
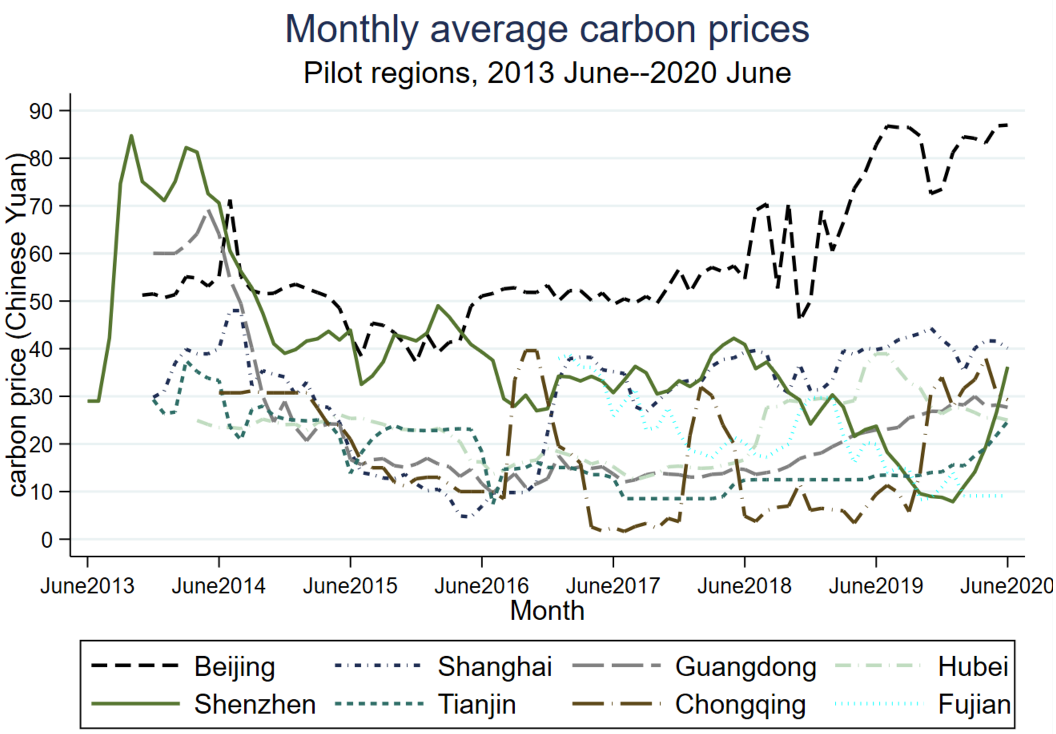


Figure 13: Monthly average carbon price in pilot regions



D.1 The carbon price elasticity using different leads

Table 12: Effect of pilot ETS on green patenting using matched sample, carbon price elasticity by pilot regions

	(1)	(2)	(3)	(4)	(5)
Green patents, weighted					
Logarithm carbon price T+1	0.19**	0.40**	0.41**	0.11	0.09
	(0.10)	(0.17)	(0.18)	(0.28)	(0.12)
regulated	0.31	0.44	-0.83**	-0.23	0.37
	(0.25)	(0.38)	(0.37)	(0.49)	(0.26)
inflate					
Logarithm carbon price T+1	0.13	0.28	0.18	0.58	0.08
	(0.09)	(0.19)	(0.15)	(0.45)	(0.11)
regulated	0.14	0.52	-0.79*	-2.04*	0.29
	(0.19)	(0.52)	(0.44)	(1.14)	(0.31)
Observations	7829	1203	1638	1066	3121
Mean dependent var.	0.39	0.56	0.35	0.20	0.48
Sd. of dependent var.	3.56	7.17	1.94	0.81	3.10
Pilot	Full sample	Beijing	Shanghai	Hubei	Shenzhen
log likelihood	-6433.02	-1087.85	-1176.05	-474.91	-2874.39
AIC/N	1.66	1.88	1.49	0.97	1.87

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample processed using matching. Columns (1)-(6) report the estimations on the carbon price elasticity on number of green patents by pilot regions using the carbon price with one-year ahead. Standard errors are clustered at 4-digit sector level, with 93, 29, 111, 88, 26, and 143 clusters respectively in columns (1)-(6). Specifications in all the columns include year fixed effects.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 13: Effect of pilot ETS on green patenting using matched sample, carbon price elasticity by pilot regions

	(1)	(2)	(3)	(4)	(5)
Green patents, weighted					
Logarithm carbon price T+2	0.26	0.39**	0.35*	0.12	0.10
	(0.16)	(0.17)	(0.20)	(0.25)	(0.12)
regulated	0.15	0.46	-0.65*	-0.30	0.37
	(0.37)	(0.38)	(0.38)	(0.48)	(0.25)
inflate					
Logarithm carbon price T+2	0.14	0.26	0.13	0.61	0.08
	(0.12)	(0.19)	(0.17)	(0.39)	(0.11)
regulated	0.03	0.54	-0.65	-2.23**	0.29
	(0.24)	(0.52)	(0.46)	(1.12)	(0.30)
Observations	7829	1203	1638	1066	3121
Mean dependent var.	0.39	0.56	0.35	0.20	0.48
Sd. of dependent var.	3.56	7.17	1.94	0.81	3.10
Pilot	Full sample	Beijing	Shanghai	Hubei	Shenzhen
log likelihood	-6547.22	-1088.12	-1180.14	-474.51	-2874.44
AIC/N	1.68	1.88	1.49	0.97	1.87

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample processed using matching. Columns (1)-(4) report the estimations on the carbon price elasticity on number of green patents by pilot regions using the carbon price with two years ahead. Standard errors are clustered at 4-digit sector level, with 93, 111, 88, and 143 clusters respectively in columns (1)-(4). Specifications in all the columns include year fixed effects.

* p <0.1, ** p<0.05, *** p<0.01

Table 14: Effect of pilot ETS on green patenting using matched sample, carbon price elasticity by pilot regions

	(1)	(2)	(3)	(4)	(5)
Green patents, weighted					
Logarithm carbon price T+3	0.32	0.40**	0.35*	0.17	0.10
	(0.23)	(0.17)	(0.19)	(0.21)	(0.12)
regulated	0.02	0.38	-0.70*	-0.44	0.39
	(0.49)	(0.39)	(0.40)	(0.42)	(0.26)
inflate					
Logarithm carbon price T+3	0.20	0.27	0.18	0.64*	0.09
	(0.17)	(0.19)	(0.17)	(0.33)	(0.12)
regulated	-0.07	0.48	-0.74	-2.35**	0.30
	(0.31)	(0.53)	(0.48)	(0.99)	(0.30)
Observations	7829	1203	1638	1066	3121
Mean dependent var.	0.39	0.56	0.35	0.20	0.48
Sd. of dependent var.	3.56	7.17	1.94	0.81	3.10
Pilot	Full sample	Beijing	Shanghai	Hubei	Shenzhen
log likelihood	-6523.06	-1085.94	-1179.44	-474.22	-2874.81
AIC/N	1.68	1.88	1.49	0.97	1.87

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample processed using matching. Columns (1)-(6) report the estimations on the carbon price elasticity on number of green patents by pilot regions using the carbon price with three years ahead. Standard errors are clustered at 4-digit sector level, with 93, 29, 111, 88, 26, and 143 clusters respectively in columns (1)-(6). Specifications in all the columns include year fixed effects.

* p <0.1, ** p<0.05, *** p<0.01

D.2 OLS estimations

Table 15: Effect of pilot ETS on green patenting using matched sample, OLS estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
regulated*post	0.23*						0.36*	0.34	0.11	0.14
	(0.14)						(0.19)	(0.23)	(0.09)	(0.09)
Logarithm carbon price		0.08								
		(0.05)								
Logarithm carbon price T+1			0.07							
			(0.05)							
Logarithm carbon price T+2				0.06						
				(0.05)						
Logarithm carbon price T+3					0.07					
					(0.06)					
first quartile × regulated*post=1						0.41				
						(0.35)				
second quartile × regulated*post=1						-0.07				
						(0.11)				
third quartile × regulated*post=1						0.06				
						(0.14)				
fourth quartile × regulated*post=1						0.44				
						(0.40)				
regulated							0.06			0.03
							(0.06)			(0.06)
Observations	7828	7828	7828	7828	7828	7828	7829	3882	7828	7829
Mean dependent var.	0.40	0.40	0.40	0.40	0.40	0.40	0.39	0.80	0.10	0.10
Sd. of dependent var.	3.56	3.56	3.56	3.56	3.56	3.56	3.56	5.03	1.88	1.88
Adjusted R-squared	0.10	0.10	0.10	0.10	0.10	0.10	0.01	0.11	0.11	0.00

This table reports the OLS estimations for the sample processed using matching. Column (1) and (9) show the overall effects of the regulation on green patenting; columns (2)-(5) show the the estimations on the carbon price elasticity on number of green patents, with different price leads; column (6) shows the results for estimating the pilot ETS effects by quartile of firms' output per worker distribution; columns (7) and (10) show the effects on dirty patenting; column (8) shows the effects on the share of green patents calculated as the ratio between the number of green patents and the sum of the numbers of green and dirty patents. Standard errors are clustered at 4-digit sector level, with 266 clusters in column (6) and 268 clusters in the rest of the columns. Specifications in columns (1)-(7) include year fixed effects and firm fixed effects; specifications in columns (8)-(10) include year fixed effects, the pilot region dummies, the ownership dummies and the firm size dummies.
* p <0.1, ** p<0.05, *** p<0.01

D.3 Fixed-effect Poisson estimations

Table 16: Effect of pilot ETS on green patenting using matched sample, fixed-effect Poisson estimations

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
main											
regulated*post	0.49**	0.28***	0.96**	0.93	0.47	-0.17	-0.03	0.06			0.48
	(0.23)	(0.09)	(0.44)	(0.77)	(0.35)	(0.49)	(0.65)	(0.33)			(0.35)
regulated	0.20										
	(0.23)										
Logarithm carbon price									0.09		
									(0.06)		
first quartile × regulated*post=1										0.15	
										(0.40)	
second quartile × regulated*post=1										-0.06	
										(0.33)	
third quartile × regulated*post=1										-0.03	
										(0.30)	
fourth quartile × regulated*post=1										0.60	
										(0.43)	
Observations	7829	3882	586	157	827	551	174	1587	3882	3882	1584
Mean dependent var.	0.39	0.80	1.15	0.46	0.71	0.38	0.32	0.94	0.80	0.80	0.50
Sd. of dependent var.	3.56	5.03	10.25	1.46	2.68	1.09	0.75	4.30	5.03	5.03	4.15
Pilot			Beijing	Tianjin	Shanghai	Hubei	Guangdong	Shenzhen			
Pseudo R-squared	0.17										
log likelihood	-8240.01	-3068.36	-532.83	-68.50	-576.16	-261.40	-67.76	-1384.34	-3066.96	-3040.07	-842.06
AIC/N											

This table presents estimations from the Poisson regression with firm fixed effects using the matched sample. Column (1) shows the results for estimating the overall ETS effects without firm fixed effects, while column (2) shows the effects with firm fixed effects included. Columns (3)-(8) show the results for estimating the pilot heterogeneity effects using sub-samples by pilot regions, i.e., the effects of regulation in different municipalities or provinces with the implementation of emissions trading scheme (ETS). Standard errors are clustered at the 4-digit sector level in column (1), with 268 clusters. Robust standard errors are reported in columns (2)-(10). Specifications in all the columns include year fixed effects. Specifications in column (1) include the ownership dummies, the pilot region dummies and the firm size dummies.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

D.4 Estimations using the non-matched sample

Table 17: Effect of pilot ETS on green patenting using non-matched sample, count data model

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
main									
regulated*post	0.65						-0.00	-0.02	0.54
	(0.47)						(0.01)	(0.02)	(0.34)
regulated	0.42	0.26	0.24	0.18	0.13	0.02			0.59*
	(0.27)	(0.27)	(0.28)	(0.30)	(0.32)	(0.49)			(0.35)
Logarithm carbon price		0.24							
		(0.15)							
Logarithm carbon price T+1			0.25						
			(0.16)						
Logarithm carbon price T+2				0.28*					
				(0.17)					
Logarithm carbon price T+3					0.29*				
					(0.17)				
first quartile × regulated*post=1						0.93*			
						(0.48)			
second quartile × regulated*post=1						-0.03			
						(0.21)			
third quartile × regulated*post=1						0.26			
						(0.19)			
fourth quartile × regulated*post=1						-0.03			
						(0.34)			
inflate									
regulated*post	0.27***								0.43**
	(0.09)								(0.20)
regulated	-0.09	-0.13	-0.14	-0.16	-0.18*	-0.01			-0.47*
	(0.13)	(0.12)	(0.11)	(0.11)	(0.11)	(0.20)			(0.26)
Logarithm carbon price		0.09***							
		(0.03)							
Logarithm carbon price T+1			0.09***						
			(0.03)						
Logarithm carbon price T+2				0.10***					
				(0.03)					
Logarithm carbon price T+3					0.11***				
					(0.04)				
first quartile × regulated*post=1						0.17			
						(0.13)			
second quartile × regulated*post=1						0.40*			
						(0.23)			
third quartile × regulated*post=1						0.41**			
						(0.20)			
fourth quartile × regulated*post=1						0.17			
						(0.15)			
Observations	83050	83050	83050	83050	83050	83046	82686	9202	83050
Mean dependent var.	0.55	0.55	0.55	0.55	0.55	0.55	0.12	0.82	0.07
Sd. of dependent var.	13.97	13.97	13.97	13.97	13.97	13.97	0.32	0.36	1.04
R-squared							0.33	0.65	
log likelihood	-103555.85	-103177.58	-103106.69	-102912.47	-102772.60	-98399.51			-17017.89
AIC/N	2.50	2.49	2.48	2.48	2.48	2.37			0.41

This table reports maximum likelihood estimators using a zero-inflated Poisson model for the sample of all firms locating in the six pilot regions processed without matching. Columns (1)-(6) and (9) show the results from the zero-inflated Poisson regression, with the outcome variables as the green patent counts in columns (1)-(6) and the dirty patent counts in column (9); columns (7) and (8) show the results from OLS regression with the outcome variable as the share of the green patent counts. Column (1) shows the overall effect of the regulation on green patenting; columns (2)-(5) show the estimations of the carbon price elasticity on number of green patents, with different price leads; column (6) shows the results for estimating the pilot ETS effects by quartile of firms' output per worker distribution; columns (7) and (8) present the estimations of the ETS effects on the share of green patenting; Column (9) shows the estimations on the ETS effects on dirty patenting. Standard errors are clustered at 4-digit sector level, with 330 clusters in column (8) and from 532 to 536 clusters in the other columns. Specifications in all the columns include year fixed effects.

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$