CGPA WORKING PAPER
2025-14

The Reallocation Effect of Emissions Cap-and-Trade: Evidence from China

Ohyun Kwon
Hao Zhao
Min Qiang Zhan

DREXEL UNIVERSITY
Center for
Global Policy Analysis
LeBow College of Business

The CGPA working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the CGPA Board. The CGPA does not hold any responsibility for the content and correctness of the papers in this series.

The Reallocation Effect of Emissions Cap-and-Trade: Evidence from China

Ohyun Kwon[*] Hao Zhao[†] Min Qiang Zhao[‡]
Drexel University Renmin University Xiamen University

This paper investigates the reallocation effects of emissions cap-and-trade policy, leveraging China's phased implementation of chemical oxygen demand (COD) regulations as a quasi-experiment. Our theoretical model posits that a *pro rata* emissions cap is more stringent for more productive firms, resulting in *negative reallocation*, whereas emissions trading restores efficiency through *positive reallocation* by reallocating emissions towards more productive firms. Utilizing the spatial and temporal variation in policy implementation, our empirical findings demonstrate that emission intensities of more productive firms, relative to less productive counterparts, declined after adopting the cap policy but subsequently increased with the introduction of cap-and-trade, aligning with our theoretical predictions. We also find that firms operating under a cap-and-trade regime, on average, experienced faster output growth compared to those operating under a cap-only regime, highlighting the role of emissions trading in enhancing production efficiency, even within imperfect markets.

JEL classification: L51 Q53 Q58

Keywords: emissions cap-and-trade; heterogeneous firms; imperfect competition; reallocative efficiency

Declarations of interest: none

[*]LeBow College of Business, Drexel University, US, ok85@drexel.edu.

[†]Corresponding author, School of Ecology and Environment, Renmin University of China, China, haozhao@ruc.edu.cn. Hao Zhao acknowledges financial support from the National Natural Science Foundation of China (No. 72303231).

[‡]MOE Key Laboratory of Econometrics, the Wang Yanan Institute for Studies in Economics, Xiamen University, China, kent_zhao@xmu.edu.cn. Min Qiang Zhao acknowledges financial support from the National Natural Science Foundation of China (No. 71773102).

1. Introduction

The comparison between command-and-control and market-based approaches in pollution regulation has been a central topic in environmental economics, dating back to at least Weitzman (1974). Market-based regulatory tools, such as emissions taxes and cap-and-trade programs, are widely recognized for their potential to achieve "cost-effectiveness," defined by equalizing the shadow cost of emissions across producers (Coase 1960; Crocker 1966; Montgomery 1972; Hahn and Stavins 2011). However, their impact on overall economic efficiency is less clear in markets characterized by imperfect competition (Buchanan 1969; Malueg 1990; Sartzetakis 1997), where equalized marginal costs do not necessarily translate into optimal production efficiency due to potential discrepancies in the marginal products of emissions across producers. While market imperfection may reduce some of the efficiency gains typically attributed to market-based regulations, empirical research comparing the relative efficiency of these regulatory tools in real-world settings remains limited (Carlson et al. 2000; Muller and Mendelsohn 2009; Fowlie et al. 2012). In particular, there is a scarcity of firm-level analysis on the reallocation effects of emissions trading, which could provide crucial insight into the mechanisms proposed by existing economic theory (Curtis 2018). This paper addresses this gap by empirically examining the reallocation effects of emissions cap-and-trade programs on Chinese manufacturing firms.

Our work leverages the unique policy setting of chemical oxygen demand (COD) regulation in China,[1] where COD emissions control has undergone sequential institutional changes. Prior to 2006, COD discharges in China were only subject to an unduly low emissions discharge fee. From 2006 onward, the central government's eleventh five-year (2006-2010) plan assigned a percentage-wise COD reduction target to each province, effectively instituting a *pro rata* emissions cap regulation.[2] In addition to emissions caps, a subset of provinces spontaneously implemented cap-and-trade programs.[3] This sequential implementation provides a unique opportunity to separately identify the impacts of emissions "cap" and "trade" policies, and compare their respective economic efficiencies.

To derive predictions for our empirical analysis, we develop a heterogeneous-firm model under monopolistic competition to characterize how manufacturing firms respond

[1] COD is a key indicator of water pollution and has been extensively studied in the literature (Keiser and Shapiro 2019; He et al. 2020).

[2] The emissions reduction targets ranged from 4.8 percent in Fujian province to 15.1 percent in Hebei, Jiangsu, and Zhejiang provinces. Only four provinces (Hainan, Qinghai, Tibet, and Xinjiang) were exempted from these COD reduction targets.

[3] See Zhang et al. (2016) for a review of cap-and-trade programs. To the best of our knowledge, the underlying factors that drove regional variation in the adoption of COD cap-and-trade programs remain unclear. Provinces that adopted cap-and-trade programs did so within a short timeframe, limiting the possibility of self-selection into cap-and-trade programs based on learning from peer provinces.

to changes in environmental regulations within an imperfectly competitive market. Our model highlights a novel *reallocation* channel, where the impact of emission regulations is highly heterogeneous, depending on firm productivity.[4] Specifically, when productive factors and emissions are gross substitutes,[5] a *pro rata* emissions cap results in *negative* reallocation relative to the baseline economy with a constant emissions charge. This arises because the equal percentage reduction in emissions across firms effectively imposes a higher shadow price of emissions on more productive firms. However, a cap-and-trade policy counteracts this cap effect by enabling a greater reduction in the cost of emissions for more productive firms, leading to an equilibrium shift denoted as *positive* reallocation. Our model shows that this positive reallocation effect induced by emissions trading is the key to enhancing economy-wide productivity.

Our theoretical prediction is consistent with the aggregate-level data pattern in Figure 1, which illustrates the differences in annual COD emissions and real economic output (normalized by their respective levels in 2005) between regions with and without cap-and-trade programs. While the rate of COD decline remained comparable throughout the period (solid line), the output gap (dashed line), relatively flat before 2009, widened substantially afterward, coinciding with the adoption of cap-and-trade policies. Moreover, the two regions displayed a comparable trend in SO_2 emissions, a pollutant also subject to cap and cap-and-trade policies during the sample period, mitigating concerns that the divergence in output growth was driven by concurrent environmental policies.

[Insert Figure 1 Here]

We conduct a rigorous empirical analysis to examine the reallocation effects of cap-and-trade programs. Based on the theoretical link between firm-level shadow costs of emissions and emission intensities,[6] we infer the effects of policy changes on firms' emissions costs through reversals in their emission intensities. Specifically, the negative reallocation effect of emissions cap policies manifests as relatively lower emission intensities for more productive firms following policy implementation, while the positive reallocation effect of emissions trading is associated with relatively higher emission intensities for these firms after transitioning from an emissions cap to cap-and-trade. These

[4]The existing literature comparing emissions cap and cap-and-trade policies has primarily focused on the cost-effectiveness of cap-and-trade programs (Carlson et al. 2000; Muller and Mendelsohn 2009). However, as demonstrated in Appendix Section E.2 of this paper, when productive factors and emissions are gross complements, emissions cap-and-trade can reduce overall efficiency under imperfect competition, despite being more cost-effective than emissions caps. Therefore, the reallocation effects identified in our model extend beyond the conventional focus on cost-efficiency.

[5]Following the existing literature (Copeland and Taylor 1994; Forslid et al. 2018; Shapiro and Walker 2018), we establish the equivalence between treating emissions as a by-product or an input in the production process. Therefore, consistent with the perspective of emissions as an input, productive factors and emissions can be viewed as gross substitutes.

[6]Firm-level emission intensity is defined as the mass of emissions normalized by output (Forslid et al. 2018; Shapiro and Walker 2018).

theoretical predictions motivate our triple-difference (DiDiD) design, which identifies the causal impact of environmental policies on firm-level emission intensities by comparing changes over time, across different regulatory regimes, and among firms with varying productivity levels. To address potential identification challenges arising from the staggered implementation of cap-and-trade programs, we adopt a stacked regression approach (Cengiz et al. 2019; Deshpande and Li 2019; Wing et al. 2024). Following the investigation of the reallocation mechanisms, we employ a difference-in-differences (DiD) approach to establish the causal relationship between the introduction of cap-and-trade programs and higher firm-level output, thereby providing robust micro-level support for the efficiency gains posited by our theoretical framework.

In line with the theoretical predictions, we find that more productive firms experienced a relatively greater decrease and then an increase in emission intensities vis-á-vis their less productive counterparts following the sequential implementation of cap and cap-and-trade policies, respectively. Specifically, during the emissions cap phase, a firm with 10% higher productivity experienced a 3.50% greater decline in emission intensity when located in a province with a one percentage point higher emissions reduction target. Subsequently, after the implementation of cap-and-trade programs, a firm with 10% higher productivity experienced a 4.67% greater increase in emission intensity if located in a prefecture adopting cap-and-trade. This positive reallocation effect associated with emissions trading ultimately led to an average increase in firm-level output by 10.3%.

To deepen our understanding of these empirical findings, we conduct a comprehensive set of mechanism analyses and robustness checks. First, we present direct evidence of the reallocation mechanism, demonstrating the shift of emissions and output from low-productivity to high-productivity firms following the implementation of cap-and-trade policies. Next, we perform a detailed subsample analysis, revealing that the reallocation effects are particularly pronounced for domestic firms, state-owned enterprises, more emission-intensive industries, and firms located in regions with newly appointed local governors. Finally, we assess the influence of concurrent external events, such as SO_2 regulations, the 2008 financial crisis, and changes in inclusion criteria for manufacturing firm data. Our findings remain robust across all tests, reinforcing the credibility and consistency of our main results.

This paper is closely related to two strands of literature. Its primary contribution lies in providing a micro-level examination of the reallocation effects associated with emissions cap and cap-and-trade programs. Previous studies have relied on either structural estimation of marginal abatement cost functions (Carlson et al. 2000; Muller and Mendelsohn 2009) or semi-parametric matching of firms located in cap-and-trade and cap-only zones (Fowlie et al. 2012) to explore the disparity between cap-and-trade and command-and-control policies. The sequential policy implementations in China allow us to directly assess how production and emissions decisions at the firm level adjusted before and after the policy changes, thereby uncovering the reallocation mechanisms of emissions regulations.

Importantly, our study isolates the "trade" effect from the negative production shock caused by emissions cap policies. A broader literature has focused on comparing cap-and-trade to regimes without stringent environmental policies. The findings in this literature have been mixed, due to challenges in disentangling the cost of emission reduction from the efficiency gains of trading (Oates and Strassmann 1984; Fowlie et al. 2016), or due to the "loose" design of cap-and-trade programs (Ferris et al. 2014; Martin et al. 2016; Curtis 2018; Cao et al. 2021; Goulder et al. 2022). In terms of theoretical analysis of cap-and-trade regulation, our model is closely related to Konishi and Tarui (2015) and Anouliès (2017). However, a key distinction is that we compare cap-and-trade with a *pro rata* cap policy, whereas those papers compare different cap-and-trade policy designs.

Second, this paper adds to the burgeoning economics literature on water pollution regulations (Wang et al. 2018; Keiser and Shapiro 2019; He et al. 2020). Leveraging the variation in provincial COD emissions reduction targets implemented during China's eleventh five-year plan, previous studies have documented the effectiveness of the emissions caps on firm-level emissions and the relocation of firms to regions with less stringent regulations (Wu et al. 2017; Chen et al. 2018; Fan et al. 2019). This paper also builds on the regional variation in COD regulation stringency, but we shift our focus to the reallocation effects and efficiency implications of emissions trading, a topic that has received less attention in the Chinese context.

The remainder of the paper is organized as follows. Section 2 provides the institutional background on COD emissions regulations in China. Section 3 presents a theoretical framework for evaluating emissions cap and cap-and-trade policies. Section 4 describes data sources and the construction of key variables. Section 5 introduces our empirical strategy. Section 6 presents our empirical analysis, and the final section concludes.

2. Policy Background

2.1. Discharge Fee (since 1982)

COD is an important water quality parameter included in the first national discharge standard (GBJ 4–73) promulgated in 1973 (Xu et al. 2020). Based on the "polluter pays" principle, the Chinese government published its *Interim Measures for the Collection of*

Pollutant Discharge Fees in 1982 to establish an environmental discharge fee system.[7] However, some studies have questioned the effectiveness of these discharge fees, as they were maintained at excessively low levels by local governments, who prioritized economic growth over environmental protection (Liu et al. 2017; Wang et al. 2018).

2.2. COD Emissions Cap (since 2006)

As a part of the eleventh five-year plan, the central government in China set a national target to reduce COD emissions by 10 percent below the 2005 national level by 2010.[8] This national target trickled down to province-specific targets based on factors such as past economic growth, industrial structure, current pollution intensities, and assessed potential to reduce emissions (Wu et al. 2017).[9] These mandatory reduction targets were enforced through measures such as increasing household waste treatment capacity and, more importantly, imposing emissions caps on industrial firms (Wu et al. 2017; Chen et al. 2018; Shi and Xu 2018; Fan et al. 2019).

In November 2006, the Ministry of Environmental Protection issued a guideline for distributing emission permits among industrial firms (Ministry of Environmental Protection 2006). First, the provincial industrial emissions target was allocated to each local district based on its share of COD emissions in 2005 (Chen et al. 2018). Second, local Environmental Protection Bureaus (EPBs) assessed each firm's reference emissions cap through "benchmarking." Specifically, the central government set a benchmark emission intensity standard for each product based on factors including product type, input type, production technology, and facility capacity.[10] For each manufacturing firm, the local EPB multiplied its base-year production of each product by the corresponding benchmark emission intensity and then aggregated the total emissions over all products produced by the firm to obtain its reference cap. Lastly, the local EPB allocated the total local emission permits to firms *pro rata* to their reference caps if the sum of reference caps exceeded the local emissions target. Due to the considerable heterogeneity in product-level emission intensities and firms' diverse product mixes, each firm's baseline emissions were a key determinant of its reference emissions cap.[11]

[7] The latest version of the guidance regarding discharge fees is provided on the following Chinese government's website: http://www.gov.cn/zhengce/content/2008-03/28/content_5152.htm.

[8] This 10-percent reduction target was continued by the other 10-percent industrial emissions reduction target in the twelfth five-year plan (2011-2015).

[9] Further details can be found in the "Reply to Pollution Control Plan During the Eleventh Five-Year Plan," issued by the State Council of China in 2006 (http://www.gov.cn/gongbao/content/2006/content_394866.htm).

[10] See the Handbook on Emission Coefficients of Industrial Sources of Pollution for the First National Census on Pollution Sources for the details of product-specific emission intensity standards. The Chinese government also published the standards for COD density in wastewater and wastewater discharge per unit of output at sub-industrial levels.

[11] Kwon et al. (2024) confirm that product-level COD emission intensities in China exhibit substantial variation, with a corresponding coefficient of variance being approximately 15.3.

To enforce the emissions caps, EPBs conducted site inspections and punished violators whose actual emissions exceeded their permit holdings by means of fines or the revocation of operating licenses (Kahn et al. 2015). Overall, the central government's COD reduction campaign had proven very successful by the end of the eleventh five-year plan. As shown in Figure 2, nationwide COD emissions declined by more than 1.7 million tons during the eleventh five-year plan, exceeding the goal of 1.4 million tons (or 10% reduction) set by the central government.

[Insert Figure 2 Here]

2.3. COD Emissions Cap and Trade (11 provinces, since 2009)

Following the introduction of emissions reduction targets, in 2007, the Ministry of Finance, the Ministry of Environmental Protection, and the National Development and Reform Commission approved emissions trading pilot programs in eleven provinces during the eleventh five-year plan period. As detailed in Appendix Table A1, full provincial implementation began in 2009, following initial pilot projects in select cities in 2007. The selected regions, including coal-intensive provinces like Shanxi and Inner Mongolia and industrialized areas such as Zhejiang and Jiangsu, highlight the importance of local experiences in facilitating the large-scale adoption of cap-and-trade programs.[12]

In the cap-and-trade programs, permit trading was organized through centralized exchanges or government-led auctions. In most regions, firms purchased and sold emission permits with local EPBs at a fixed price predetermined by the government. Active trading was observed in ten out of the eleven provinces, with the volume of transactions varying substantially across regions. For example, we find no trade records of COD emission permit transactions in Tianjin during our sample period, while in Chongqing, there were 236 COD emission permit transactions from December 2010 to October 2012 (Zhang et al. 2016).[13]

[12] China's experimentation with emissions trading began in the 1980s, with local governments coordinating trade between firms to meet their emission control requirements, primarily for SO_2 and COD. These early trials informed the regulatory frameworks for the later provincial-level programs (Zhang et al. 2016; Ren et al. 2020; Ye et al. 2020).

[13] Although Tianjin officially adopted a cap-and-trade program, the lack of observed trading activity leads us to classify it as a non-cap-and-trade region in our empirical examination. However, our results are robust to alternative specifications that include Tianjin as a cap-and-trade region.

3. Theoretical Framework: Heterogeneous Firms under Sequential Environmental Regulation Changes

To generate testable hypotheses that motivate our empirical analyses, we develop the following simple theoretical framework.[14]

3.1. Utility and Demand

In our economy, a representative consumer's preferences are characterized by the following quasi-linear utility function:

$$U = q_0 + E\ln(Q) - h(Z), \tag{1}$$

where q_0 denotes the consumption of a numéraire good, E is an exogenous expenditure parameter, $Q = \left[\int q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}}$ represents the aggregate consumption of differentiated varieties ω, and $h(Z)$ denotes the disutility associated with total emissions Z.

Solving the consumer's utility maximization problem yields the following demand function for each variety:

$$q(\omega) = \Phi p(\omega)^{-\sigma}, \tag{2}$$

where $\Phi \equiv EP^{\sigma-1}$ serves as a demand shifter, with $P = \left[\int p(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$ denoting the aggregate price index.

3.2. Production and Emissions in the Benchmark Equilibrium

Consider a fixed measure of infinitesimal firms, each producing a unique good under monopolistic competition. A firm's output is determined by the following simplified production function:

$$q = (1-\theta)\phi l, \tag{3}$$

where ϕ is the firm's TFP, l denotes an aggregate of productive factors (referred to as "labor" in this model), and θ reflects the share of productive factors dedicated to pollution abatement. The emissions generated post-abatement are described by:

$$z = \phi l \left[\frac{(1-\theta)^{\frac{\eta-1}{\eta}} - 1}{\alpha} + 1\right]^{\frac{\eta}{\eta-1}}, \tag{4}$$

[14] Full details of this model, including supplementary equations, underlying assumptions, formal proofs, and further discussion, are provided in the online appendix.

where $\alpha \in (0,1)$ measures the efficiency of abatement investment.

Combining equations (3) and (4) leads to a simplified CES production function:

$$q = \left[\alpha z^{\frac{\eta-1}{\eta}} + (1-\alpha)(\phi l)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}}, \tag{5}$$

where η captures the elasticity of substitution between labor and emissions.

Given the unit emissions discharge fee τ and wage w, the solution to the firm's profit maximization problem admits the following expressions for emissions z, labor l, and emission intensity $\zeta \equiv \frac{z}{q}$ in equilibrium:

$$z = \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi \alpha^{\eta} \tau^{-\eta} \left[\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}} \tag{6}$$

$$l = \frac{1}{\phi}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi\,(1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{-\eta}\left[\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}} \tag{7}$$

$$\zeta = \left[\alpha + (1-\alpha)\left(\frac{\phi\tau}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}\right]^{\frac{\eta}{1-\eta}}. \tag{8}$$

3.3. Two Environmental Policies to Curb Emissions

3.3.1. Pro Rata Emissions Cap Policy

Suppose that the government aims to reduce total emissions from the benchmark scenario by a factor of $1-\delta$ ($0<\delta<1$). Under the *pro rata* cap policy regime, the government sets the reduction factor $1-\delta$, and applies it to each firm uniformly. In particular, the level of emissions allowed by the government, z_c, is determined by each firm's benchmark emissions level z_b multiplied by δ (i.e., $z_c = \delta z_b$), where the subscripts b and c denote the benchmark and cap regimes, respectively.

The equilibrium under the cap policy can be characterized by λ_c, the shadow price of emissions under the cap, which is a hypothetical (sufficiently high) cost of emissions inducing a firm to choose an emissions level z_c in the benchmark equilibrium. λ_c is implicitly determined by the following equation,

$$\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi_c \alpha^{\eta} \lambda_c^{-\eta} \left[\alpha^{\eta}\lambda_c^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}} = \delta z_b, \tag{9}$$

where the left-hand side is the level of emissions associated with the emissions cost λ_c in the benchmark equilibrium, and the right-hand side is the mandated emissions level under the cap policy.

3.3.2. Emissions Cap-and-Trade Policy

Following the cap policy, a cap-and-trade system is introduced, allowing firms to trade emissions permits. The total unit cost of emissions, λ_t, is the sum of equilibrium permit price and the existing emissions cost. Hence, the equilibrium emissions level for a firm with productivity ϕ, $z_t(\phi)$, can be expressed similarly to equation (6) by replacing the emissions cost τ with λ_t. The emissions cost in the cap-and-trade equilibrium is determined by the market clearing condition for emission permits:

$$\int_{\underline{\phi}}^{\bar{\phi}} z_t(\phi) g(\phi) d\phi = \int_{\underline{\phi}}^{\bar{\phi}} z_c(\phi) g(\phi) d\phi, \tag{10}$$

where $g(\phi)$ is the density function of the distribution of ϕ.

3.4. A Comparison between Two Policies

To compare the emissions cap and cap-and-trade policies, we first illustrate how the emissions cap policy distorts the shadow price of emissions relative to the benchmark regime. We base our analysis on the parameter constraint $1 < \eta < \sigma$, implying that emissions and labor are gross substitutes. Empirical support for this assumption is provided in Appendix Section F.

The key result from the implementation of a emissions cap policy is summarized in the following proposition:

PROPOSITION 1. *When emissions and labor are gross substitutes ($1 < \eta < \sigma$), the pro rata emissions cap policy results in a negative reallocation of emissions. Relative to the benchmark equilibrium, the pro rata emissions cap*

1. *increases the shadow price of emissions for all firms: $\lambda_c(\phi) > \tau \, \forall \, \phi$, and this increase is higher for more productive firms: $\frac{\partial[\lambda_c(\phi) - \tau]}{\partial \phi} > 0$;*
2. *lowers firm-level emission intensity for all firms: $\zeta_c(\phi) < \zeta_b(\phi) \, \forall \, \phi$, and the percentage reduction is greater for more productive firms:*

$$\frac{\partial[\ln(\zeta_c(\phi)) - \ln(\zeta_b(\phi))]}{\partial \phi} < 0. \tag{11}$$

The key insight from Proposition 1 is that the emissions cap policy generates dispersion in the shadow price of emissions across firms, disproportionately affecting more productive firms. Consequently, this leads to negative reallocation of emissions.

The observation that a *pro rata* emissions cap has an unintended negative allocative impact has not been well recognized in the literature. This finding provides a justification for implementing a cap-and-trade policy. A key distinction between the two regimes is that the shadow price of emissions is higher for more productive firms under an emissions cap, while it is equalized across firms under a cap-and-trade policy.

PROPOSITION 2. *When emissions and labor are gross substitutes ($1 < \eta < \sigma$), the cap-and-trade policy results in a positive reallocation of emission permits. Relative to the cap regime, the cap-and-trade program*

1. *lowers the shadow price of emissions for more productive firms and raises the shadow price of emissions for less productive firms, with the decline in the shadow price being greater for more productive firms:* $\frac{\partial[\lambda_t - \lambda_c(\phi)]}{\partial \phi} < 0$;
2. *increases the emission intensity for more productive firms and lowers the emission intensity for less productive firms. The percentage increase in emission intensity is greater for more productive firms:*

$$\frac{\partial[\ln(\zeta_t(\phi)) - \ln(\zeta_c(\phi))]}{\partial \phi} > 0, \tag{12}$$

for all firms with productivity above ϕ_l, *where* ϕ_l *is characterized by*

$$\phi_l = \max\left\{\phi \middle| \alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)\left(\frac{1-\alpha}{\alpha w}\right)^{\eta-1} \phi^{\eta-1}\lambda_t^{\eta-1}\right] \phi \lambda_c^{\eta-2} \frac{\partial \lambda_c}{\partial \phi} = 0\right\}.$$

Transitioning from an emissions cap to a cap-and-trade policy generates a heterogeneous impact on emission intensities as a result of reallocating emissions from less productive to more productive firms. The overall productivity and output implications of these policies are summarized in the following proposition.

PROPOSITION 3. *When emissions and labor are gross substitutes ($1 < \eta < \sigma$), a cap-and-trade policy results in higher firm-average and aggregate output compared to a pro rata emissions cap policy:* $Q_t > Q_c$.

Proposition 3 is an intuitive result, given the negative reallocation caused by emissions caps. The cap-and-trade policy restores the distorted shadow prices of emissions and leads to positive reallocation of emission permits from less productive to more productive firms. This positive reallocation shifts production toward more productive firms, raising average productivity in the economy and ultimately resulting in higher firm-average and economy-wide output.

4. Data

To examine how firm-level emission intensities change with environmental regulation, we compile a dataset with firm-level emission intensity (ζ), productivity (ϕ), and regional cap and cap-and-trade policies from 2001 to 2013. Firm-level emission intensities are calculated using Chinese manufacturing firms' emissions and output data. The policy variables are constructed using information from official documents and the Ministry of Ecology and Environment (MEE) website, supplemented by reports from local governments and news outlets. Lastly, we construct firm-level productivity data by estimating firm-level revenue-based TFP using the method developed by Ackerberg et al. (2015).

4.1. Data Description

4.1.1. Firm-level Data

The construction of the manufacturing firm-level dataset follows the same procedure used in Kwon et al. (2023). Data on firm-level COD emissions are from China's Environmental Statistical Database (CESD), assembled by the MEE. Each year, major polluting firms that cumulatively account for 85% of COD emissions in each county are recorded in the CESD.[15] Firm-level production data are obtained from the Annual Survey of Industrial Enterprises (ASIE), collected and maintained by China's National Bureau of Statistics. This dataset covers all state-owned enterprises (SOEs) and non-SOEs with annual sales greater than 5 million RMB.[16] We merge the two datasets using firm identifiers such as names and organization codes. Our merged dataset contains 668,569 firm-year observations, accounting for approximately 75% of the total output and emissions reported in the CESD.

4.1.2. Emissions Cap

To construct the policy variables, we collect information on China's environmental regulations to identify the policy regimes that vary across regions and time. An effective emissions cap was first implemented during China's eleventh five-year plan period. The official emissions reduction mandate specified a percentage-wise emissions reduction target for each province, which serves as our measure of emissions cap stringency. Although Proposition 1 focuses on the impact of the implementation of a *pro rata* emissions cap, this binary effect can be naturally extended to a continuous effect with respect to the cap

[15] Firms have minimal incentive to manipulate their emissions data in the environmental survey, as China's Environmental Protection Law explicitly states that these survey data cannot be used to punish or regulate polluting firms (He et al. 2020).

[16] The threshold for inclusion in the ASIE was raised to 20 million RMB in 2011.

stringency:

$$\frac{\partial^2[\ln(\zeta_c(\phi)) - \ln(\zeta_b(\phi))]}{\partial\phi\partial(1-\delta)} < 0. \tag{13}$$

Namely, the cap effect intensifies with the size of the reduction target, $1 - \delta$.

While our measure of emissions caps is consistent with the existing literature (Wu et al. 2017), it is admittedly a "rough" measure due to its lack of granularity relative to our firm-year-level observations. To address this limitation, we include industry-year or prefecture-industry-year fixed effects in our regression analysis to account for potential variations in policy regulation stringency across industries and over time, reducing concerns about variations in the actual emissions cap across firms within the same region.

4.1.3. Cap-and-Trade Regions

Information about emissions trading is obtained from the official MEE website, which reported eleven provinces with a COD cap-and-trade program by 2013, the end of our sample period. We cross-checked this information with Zhang et al. (2016) to identify when the COD cap-and-trade program was implemented in each of those provinces.

Appendix Table A1 demonstrates that substantial variations in the timing of cap-and-trade implementation come from prefecture-level pilots. Therefore, in the main analysis, we define a cap-and-trade region at the prefecture level (j) as follows:

$$\text{Trade}_j = \mathbb{1}\{y_j^{trade} \leq 2013\}, \tag{14}$$

where y_j^{trade} denotes the year when prefecture j implemented cap-and-trade and Trade_j is a dummy variable which equals 1 if prefecture j implemented cap-and-trade before 2013. As shown in Appendix Figure A2, the fraction of prefectures and provinces under cap-and-trade grew over time, reaching approximately 30 percent by 2013. The fraction of firms under cap-and-trade reached about 45 percent, suggesting that cap-and-trade regions had a greater concentration of COD-emitting firms.

One potential concern regarding the identification of cap-and-trade policy effects is the thin market in some emissions trading programs. However, in our treatment provinces, all cap-and-trade programs generated a market price for COD emission permits. This effectively provided firms with a shadow price that influenced their emission decisions, consistent with our theoretical framework. Furthermore, in certain provinces, firms incurred non-trivial costs to acquire initial emission permits. Therefore, their emissions choices implicitly reflected the "market" price of emission permits, even without active participation in ongoing emissions trading. Despite the imperfections in the emission permits market, the actual presence and magnitude of cap-and-trade effects remain an empirical question, which we address in the subsequent analysis.

4.1.4. Productivity

Firm-specific productivity ϕ in our model captures how efficiently a firm produces output q using the productive factors l, which corresponds to the conventional firm-level TFP in the literature. We estimate TFP á la Ackerberg et al. (2015) (see Appendix B for details of the estimation). However, as noted by He et al. (2020), after accounting for emissions abatement, the estimated TFP is an "effective" TFP which also captures the fraction of productive units diverted to pollution abatement without generating any real output. This is consistent with our setup in equation (3), implying that our estimated TFP measure is $\varphi = (1-\theta)\phi$, a combination of abatement decision, θ, and firm-level productivity, ϕ.

To verify that the comparative statics with respect to ϕ are also valid with respect to φ, we examine whether the relationship between ϕ and φ is monotonic. Combining equations (3), (7), (6) and (8), we obtain

$$\varphi = (1-\theta)\phi = \frac{q}{l} = \frac{q}{z}\frac{z}{l} = \phi\left[1-\alpha+\alpha\left(\frac{\phi\tau}{w}\frac{1-\alpha}{\alpha}\right)^{1-\eta}\right]^{\frac{\eta}{\eta-1}}. \tag{15}$$

In the empirical analysis, we use the estimated TFP in 2005, the year before the implementation of the cap and cap-and-trade policies, to measure firm-level productivity. This fixed measure avoids any potential endogenous productivity changes induced by subsequent environmental regulations (Porter 1991). In 2005, firms were subject to a constant emissions discharge fee; hence, τ was fixed, and we can derive the relationship between φ and ϕ as follows:

$$\frac{\partial\varphi}{\partial\phi} = \left[1-\alpha+\alpha\left(\frac{\phi\tau}{w}\frac{1-\alpha}{\alpha}\right)^{1-\eta}\right]^{\frac{1}{\eta-1}}\left[1-\alpha+(1-\eta)\alpha^{\eta}(1-\alpha)^{1-\eta}\left(\frac{\tau\phi}{w}\right)^{1-\eta}\right]. \tag{16}$$

As shown in equation (16), the sign of $\frac{\partial\varphi}{\partial\phi}$ depends on the parameter values of η and α. We note that the empirical estimates of α in the literature are generally close to 0.01 (Wang et al. 2018; Shapiro and Walker 2018), whereas our estimates, shown in Appendix Section F, indicate that the η values are around 1.1 to 1.2. Consequently, the term $(1-\eta)\alpha^{\eta}(1-\alpha)^{1-\eta}$ is close to 0, resulting in an increasing relationship between ϕ and φ.[17] Therefore, all comparative statics and propositions derived with respect to ϕ remain valid for φ.

[17] This finding is consistent with He et al. (2020), who show that dirtier firms (corresponding to firms with lower ϕ in our model) experience larger efficiency losses (captured by a larger θ or smaller $\varphi = (1-\theta)\phi$ in our model).

4.2. Sample Selection and Summary Statistics

After removing firms with invalid location information, the merged CESD-ASIE dataset contains 668,569 observations with 168,196 unique firms during the sample period 2001-2013. We further limit our sample to COD-emitting firms (those emitting COD in any of the sample years), resulting in 561,921 observations with 128,569 unique firms.

We adopt two distinct empirical strategies to evaluate the emissions cap and cap-and-trade policies, respectively. To identify the cap effect, we restrict the sample period to the years before 2010 to avoid the impact of trade policy. We further restrict the sample to firms that meet the following criteria: (1) subject to positive emissions reduction targets, (2) operating both before and after the policy change, (3) remaining in the same prefecture, and (4) having a productivity level consistently falling between the 10th and 90th percentiles throughout the sample period.[18] This results in a final regression sample of 8,831 firms with 53,514 firm-year observations.[19] To identify the trade effect, we construct a stacked dataset to account for the staggered implementation of cap-and-trade programs, following Cengiz et al. (2019) and Deshpande and Li (2019). We apply the same sample selection criteria as for the cap sample before creating separate datasets for each cap-and-trade implementation year, resulting in a sample of 19,411 firms with 83,256 firm-year observations.

Table 1 reports summary statistics for the key variables in our firm-level regression sample. Panels A and B summarize the data used to examine the reallocation effects of the emissions cap and cap-and-trade policies, respectively. In Panel A, we divide the sample based on emissions reduction targets (above/below the median) and report summary statistics separately for each group. In Panel B, we present summary statistics separately for the treatment and control groups. Within each panel, there is no statistically significant difference between the two groups in terms of firm-level emission intensity and productivity, despite that firms with smaller reduction targets or those in cap-and-trade regions are on average slightly larger in terms of COD emissions and output than their counterparts.

[18] The productivity cutoff criterion, particularly the lower bound, is imposed because the effects predicted by Proposition 2 are valid only for firms above a certain productivity threshold. However, our theoretical framework does not prescribe an exact cutoff value for empirical analysis. To assess the robustness of our empirical estimates, we conduct a sensitivity analysis by varying the lower cutoff from the 5th to 15th percentiles. The results of these robustness checks for both emissions cap and cap-and-trade effects are presented in Appendix Tables A12 and A13, respectively.

[19] The majority of observations are dropped in two steps. Starting with 390,903 observations within the sample period, around 40% are dropped because these firms operated either exclusively before or after the policy. Subsequently, 75% of the remaining observations are removed due to missing productivity estimates or productivity levels falling within the bottom or top 10% quantiles. Firms that relocated between prefectures during our sample period account for only 1% of the data. Our qualitative conclusions remain unchanged even when these firms are included in the regression analysis.

[Insert Table 1 Here]

4.3. Threat to Identification and Balance Tests

The regional and temporal variations in policy adoption provide the foundation for our identification strategy. However, a potential threat to this strategy arises if regions were endogenously selected into cap-and-trade programs. While the summary statistics confirm that firms in regions with and without cap-and-trade exhibit no significant differences in key outcome variables, concerns remain that provinces underperforming economically or environmentally at the inception of the eleventh five-year plan might have disproportionately adopted cap-and-trade as a means to improve their performance, causing an omitted-variable bias.

Figure 1 suggests that this may not be the case. Regions with cap-and-trade programs experienced similar declines in emissions prior to 2009. For economic output, the annual difference curve in Figure 1 indicates that differences in output growth between regions were moderate and statistically indistinguishable before 2009 (as corroborated by Appendix Figure A1). These patterns contrast sharply with the substantial divergence in growth rates observed after 2009, coinciding with the implementation of cap-and-trade programs.

To further assess the balance between treated and control provinces, we conduct *t*-tests on key economic and environmental indicators for regions with and without cap-and-trade programs. As shown in Appendix Figure A1, there are no statistically significant differences in the growth rates of output and COD emissions over the three years preceding 2009. We also examine the stringency of COD reduction targets and provincial average productivity—two critical variables in our empirical analysis—and again find no significant differences between treatment and control regions. Based on Figures 1 and A1, we find no evidence of pre-existing trends in key variables or provincial characteristics driving the implementation of cap-and-trade programs.[20] Thus, we view the implementation of cap-and-trade programs as a plausible quasi-experiment that enables us to identify the causal effects of these policies.

[20]We also conduct sensitivity analyses to address potential violations of the parallel trends assumption when examining the impact of emissions trading on firm-level output, utilizing the methodology developed by Rambachan and Roth (2023). These analyses confirm that the estimated impact of cap-and-trade programs on firm-level output remains robust even when allowing for post-treatment violations of the parallel trends assumption up to 0.8 times the maximal pre-treatment deviation. Full results from these sensitivity analyses are available upon request.

5. Empirical Strategy

Propositions 1 and 2 in our theoretical framework demonstrate how emissions cap and cap-and-trade policies induce opposite emissions reallocation effects and how they impact firms' emission intensity differently based on their productivity levels. To examine these theoretical predictions at the firm level, we employ a DiDiD (triple-difference) method to estimate the heterogeneous impact of these environmental regulations on firm-level emission intensities. For economic efficiency, Proposition 3 predicts that transitioning from a cap to a cap-and-trade regime will increase average firm-level output. To estimate the changes in firm-level output associated with this environmental policy change, we use a DiD (difference-in-differences) method.

5.1. Empirical Model for Reallocation Effect of Emissions Cap

The simultaneous implementation of COD emissions reduction targets for each province during the eleventh five-year plan has been extensively examined in the literature using a standard DiD design (Wu et al. 2017; Chen et al. 2018; Shi and Xu 2018; Fan et al. 2019). Beyond that, our Proposition 1 suggests that the cap effect is heterogeneous across firms with different productivity levels. That motivates the following DiDiD design:

$$\ln(\zeta_{iy}) = \beta_0 + \beta_1 \cdot \text{Target}_s \times \text{Post}_y \times \ln(\varphi_i) + \beta_2 \cdot \text{Post}_y \times \ln(\varphi_i) + \Psi_i + \Psi_{jky} + \varepsilon_{iy}, \tag{17}$$

where ζ_{iy} denotes the emission intensity of firm i in year y, Target_s is the percentage-wise emissions reduction target in province s where firm i was located, φ_i is firm i's estimated TFP in 2005, and Post_y is a dummy variable which equals 1 if $y \geq 2006$, reflecting the simultaneous implementation of the emissions cap policy in 2006. Ψ_i denotes firm fixed effects and Ψ_{jky} denotes prefecture-industry-year fixed effects, where j indicates the prefecture where firm i located and k stands for 2-digit industry defined by the Bureau of Census Industry Codes (CIC). Other interaction terms, namely $\text{Target}_s \times \text{Post}_y$ and $\text{Target}_s \times \ln(\varphi_i)$, are fully absorbed by the fixed effects. ε_{iy} is the error term.

Our main coefficient of interest is β_1, which captures the average change in the cap policy effect associated with firms' productivity differentials. The inclusion of prefecture-industry-year fixed effects controls for any industry-prefecture-specific time-varying confounding factors that may lead to a correlation between productivity φ_i and the error term.[21] Furthermore, firm fixed effects capture any time-invariant confounding factors

[21]There might be unobserved variations in environmental regulation stringency (e.g., the actual size of emissions caps) at the industry-prefecture level that could simultaneously affect both productivity and firm-level emission intensity. See Wang et al. (2018), Chen et al. (2018), Zhang et al. (2018), Sun et al. (2019) and He et al. (2020) for relevant discussions on industrial-location-level variations in China's environmental

at the firm level (e.g., a firm's base-year production level that determines its reference emissions cap).

We further replace Post_y with a full set of year dummies $\{\text{Year}_y\}_{2001 \leq y \leq 2010}$ (with Year_{2005} as the reference group) to confirm the absence of pre-existing trends and evaluate the cap effect over time.

5.2. Empirical Model for Reallocation Effect of Emissions Trade

Following the introduction of emissions caps, eleven provinces implemented cap-and-trade programs during our sample period. Similar to the cap effect, Proposition 2 suggests that the reallocation effect of emissions trading is heterogeneous across firms with different productivity levels, prompting a DiDiD design to evaluate the policy impact across firms. However, a critical challenge arises due to the varying starting time of these cap-and-trade programs.

Recent econometrics literature points out that the "average" treatment effect identified by the conventional staggered DiD specification could be a non-convex combination of treatment effects for each period, potentially assigning negative weights to longer-term treatment effects, when "early-treated" units are used as control groups (Goodman-Bacon 2021). Furthermore, when treatment effects are heterogeneous across time or units, the standard staggered estimator may lead to estimation bias (de Chaisemartin and D'Haultfœuille 2020; Sun and Abraham 2021; Borusyak et al. 2024). This is a relevant concern in our context, given the substantial variation in the performance of cap-and-trade programs across different regions and years.

Several solutions to address the challenges of staggered DiD designs have been proposed in the literature, each relying on specific assumptions (See Roth et al. (2023) for a review of recent advances in the econometrics of difference-in-differences). We adopt the stacked regression approach proposed by Cengiz et al. (2019) and Deshpande and Li (2019), and further justified by Wing et al. (2024), which can be easily applied in the DiDiD setting. This stacked method involves first matching treated units to clean (not-yet-treated) controls by event year. These matched datasets are then stacked to derive an estimator, which represents a convex weighted average of the treatment effects for each event year. Compared with other methods, the stacked approach is straightforward to implement, does not require strictly balanced panel data, and is applicable for continuous treatment effects, rendering it particularly suitable for our data structure.

To perform the stacked regression, we first construct a separate dataset for each event year h. In each dataset, prefectures that implemented cap-and-trade programs in the event

regulations.

year are labeled as the treatment group, while prefectures that did not implement cap-and-trade programs between years $h + r_l$ and $h + r_h$ (i.e. within the relative time window $[r_l, r_h]$) are labeled as the control group. As illustrated in Appendix Figure A2, most cap-and-trade programs were introduced between 2008 and 2011. Given our sample range (2006 - 2013), our baseline specification employs four event years ($h = 2008, 2009, 2010, 2011$) and an associated relative time window $[r_l, r_h] = [-2, 2]$ to construct each dataset. As a robustness check, we also consider two alternative specifications: (1) event years $h = 2009, 2010, 2011$ and a time window $[r_l, r_h] = [-3, 2]$, and (2) event years $h = 2008, 2009, 2010$ and a time window $[r_l, r_h] = [-2, 3]$.

After constructing each dataset, we append them into a single dataset and estimate the following specification:

(18)

$$\ln(\zeta_{iyh}) = \beta_0 + \beta_1 \cdot \text{ Trade}_{jh} \times \text{Post}_{yh} \times \ln(\varphi_i) + \sum_{r \geq -2}^{2} \beta_3^r \cdot D^r \times \ln(\varphi_i) + \Psi_{ih} + \Psi_{jkyh} + \varepsilon_{iyh},$$

where h, the event year, indicates each separate dataset, Trade_{jh} is a treatment indicator which equals 1 if prefecture j implemented a cap-and-trade program in year h, Post_{yh} is the associated time dummy that equals 1 if year $y \geq h$, $r \in [-2, 2]$ refers to the relative year in the time window, and $D^r = \mathbb{1}\{y - h = r\}$ is an indicator function that equals 1 if year y is r years away from the event year h. Separate fixed effects Ψ_{ih} and Ψ_{jkyh} are included for each dataset h. The error term, ε_{iyh}, is also specific to each dataset.

We further replace Post_{yh} with a full set of relative year dummies D^r (using D^{-1} as the reference group) to confirm the absence of pre-existing trends and evaluate the trade effect over time.

5.3. Empirical Model for Trade Effect on Firm-level Output

Finally, we examine the impact of cap-and-trade programs on firm-level output. Proposition 3 suggests that firms located in prefectures with cap-and-trade programs should, on average, experience an increase in output compared to firms in cap-only regions. We test this prediction by estimating the following equation with the constructed stacked dataset:

$$\ln(q_{iyh}) = \beta_0 + \beta_1 \cdot \text{ Trade}_{jh} \times \text{Post}_{yh} + \Psi_{ih} + \Psi_{yh} + \varepsilon_{iyh}, \tag{19}$$

where q_{iyh} denotes firm i's output in year y included in subsample h.

6. Empirical Results

6.1. Negative Reallocation under Emissions Cap

Table 2 presents the estimation results for equation (17), which tests the negative reallocation effect predicted by Proposition 1.[22] Column (1) reports the estimation results from our baseline specification. In column (2), we replace the prefecture-industry-year fixed effects with a less restrictive set of province-industry-year fixed effects. A key threat to identifying the cap effect arises from the subsequent implementation of emissions trading, as our theory posits an opposite reallocation effect from trading. To address this, column (3) explicitly controls for the triple interaction with cap-and-trade regions. Furthermore, column (4) excludes cap-and-trade prefectures after the year when their cap-and-trade programs started; column (5) excludes those prefectures for all years.

[Insert Table 2 Here]

The coefficient estimates of the triple interaction term are negative and statistically significant across all empirical specifications. Column (5) is our preferred specification, as it imposes the most restrictive control for the potential trade effect. Its results suggest that a firm with 10% higher productivity experienced a 3.50% greater decline in emission intensity if it was located in a province with one percentage point higher emissions reduction target.

In Figure 3A, we depict the cap regression estimates using an event-study specification. This analysis serves two main purposes. First, it allows us to confirm the lack of pre-trends in the differences in emission intensities between more productive and less productive firms across regions with varying emissions caps. Second, it demonstrates the timing at which the cap effects took place. The estimated regression coefficients are consistent with the DiDiD estimates. The heterogeneous treatment effects are not statistically significant before 2005, suggesting that the cap effect did not materialize before the nationwide emissions reduction mandate (see Appendix Table A4 for details). The estimated coefficients become negative and statistically significant from 2006 onward, immediately after the implementation of the emissions cap policy. These results align with our theoretical prediction that more productive firms experience a greater reduction in emission intensities due to negative reallocation under the *pro rata* emissions cap policy.

[Insert Figure 3 Here]

[22] The average treatment effect of the emissions cap policy is reported in Appendix Table A2. F-tests on the coefficients of the interaction terms with $\ln(\varphi)$ yield p-values below 0.01 across all model specifications. This confirms that our DiDiD model, which incorporates the third dimension of difference, provides a significantly better fit to the data.

6.2. Positive Reallocation under Emissions Trade

Table 3 presents the coefficient estimates from equation (18), testing the positive reallocation effect predicted by Proposition 2.[23] Column (1) reports the estimation results from our baseline specification. Column (2) controls for the less restrictive province-industry-year fixed effects. Since the cap effect identified in the previous section may be intervened with the trade effect if cap-and-trade programs were introduced contingent on the performance of emissions caps, in column (3), we control for the impact of the cap policy by explicitly including the interaction terms specified in equation (17). Columns (4) and (5) present the estimation results of the baseline model with alternative event years and time windows.

[Insert Table 3 Here]

The coefficient estimates of the triple interaction term are all positive and statistically significant across most empirical specifications. Column (3) is our preferred specification. Its results suggest that a firm with 10% higher productivity experienced a 4.67% greater increase in emission intensity if located in a prefecture with a cap-and-trade program.

Figure 3B presents the event study estimates. We observe positive and statistically significant cap-and-trade effects, beginning in the second year of policy implementation and lasting until the end of our sample period. These positive reallocation effects are consistent across all specifications and are particularly strong in later years, as the cap-and-trade programs' impact deepened over time (see Appendix Table A5 for event-study estimates under different specifications). Moreover, the coefficient estimates are all statistically insignificant before policy implementation, confirming the absence of a pre-trend for the cap-and-trade effect.

6.3. Impact of Emissions Cap-and-Trade on Firm-level Production

Table 4 presents the regression results from estimating equation (19) and tests Proposition 3 regarding the efficiency gains from the positive reallocation of emissions trading. The model specification varies by the relative event time window used to construct the stacked dataset. The DiD estimates are positive and statistically significant across all specifications, consistent with our theoretical prediction that cap-and-trade programs increase firm-level output. The estimate in column (1) suggests that emissions trade leads to an average output increase of 10.3%.

[Insert Table 4 Here]

We also examine differential cap-and-trade effects and cap policy effects in Appendix Tables A6 and A7, respectively. We detect a positive reallocation of output from low- to

[23] The average treatment effect of the emissions trade policy is reported in Appendix Table A3.

high-productivity firms, consistent with our theoretical prediction of reallocation effects due to emissions trading. For the cap effect, we detect a negative impact on firm-level output, suggesting that some firms cut output to meet emissions reduction targets.

6.4. Additional Analysis

6.4.1. Emissions Reallocation

Proposition 2 demonstrates that the positive reallocation of emission permits under cap-and-trade leads to uneven changes in the shadow price of emissions among firms with different productivity levels. Because shadow prices are not directly observable, we infer their changes from the percentage changes in firm-level emissions intensities.

To directly test the reallocation mechanism, we examine the changes in firm-level emissions. Specifically, we replace the outcome variable in equation (18) with firm-level COD emissions and re-estimate the stacked DiDiD model. The results, presented in Table 5, show that the coefficient estimates are positive and statistically significant across all columns, confirming the reallocation of emission permits towards more productive firms.

[Insert Table 5 Here]

6.4.2. Subsample Analysis

The effectiveness of a cap-and-trade policy hinges on firms' compliance with environmental regulations. In Table 6, we examine the cap-and-trade effect across various firm subsamples to uncover potential mechanisms influencing firms' responses to the policy.

[Insert Table 6 Here]

Columns (1) and (2) compare importing firms with non-importing firms.[24] The positive reallocation effect is statistically significant only for non-importing firms. This finding aligns with the pollution offshoring hypothesis, suggesting that firms with access to foreign inputs (importing firms) may opt to import pollution-intensive intermediate inputs rather than adjust their production based on emissions permits (Cherniwchan et al. 2017; Kwon et al. 2023).

Columns (3) and (4) compare state-owned enterprises (SOEs) and non-SOEs. Contrary to previous studies indicating that SOEs are less responsive to environmental regulations (Hering and Poncet 2014; Shi and Xu 2018), our results show that the cap-and-trade effect is more pronounced among SOEs. We propose that information barriers or administrative red tape may have prevented non-SOEs from fully participating in emissions trading and

[24] In our empirical analysis, a firm is defined as an importing firm if it had positive imports in a given year.

reaping its benefits. Therefore, we suggest that emissions trading has the potential to further enhance overall productivity by increasing the participation of non-SOEs.

Columns (5) and (6) analyze firms in dirty versus clean industries, where dirty industries are defined as CIC 4-digit industries with industry-level emission intensity above the 70th percentile of the distribution.[25] The results indicate that the cap-and-trade effect is more substantial for firms in dirty industries, consistent with findings in the literature that dirty industries are more responsive to environmental regulations (Cai et al. 2016; Curtis 2018).

The final columns (7 and 8) focus on the characteristics of local party secretaries, following Shao et al. (2024). We divide the sample based on the party secretary's tenure at the time of cap-and-trade implementation, defining an official as new to the position if their tenure is below the sample median.[26] Our results suggest that positive reallocation occurs only among firms governed by a new local official, consistent with the notion that newly appointed officials are more proactive, supporting the adage "a new broom sweeps clean."

6.4.3. Concurrent Events

To ensure the robustness of our empirical findings, we conduct several robustness checks addressing potential confounding effects from concurrent events. First, we consider the simultaneous implementation of SO_2 regulations by the Chinese government, which included both cap and cap-and-trade policies similar to those for COD. To verify that the observed reallocation effects in COD emissions intensities are not driven by changes in SO_2 emissions, we conduct an analysis using firm-level SO_2 emission intensities as the dependent variable. The results, presented in Appendix Tables A8 and A9, generally indicate no statistically significant impact of the cap and cap-and-trade policies on firm-level SO_2 emission intensities, mitigating concerns about confounding effects from SO_2 regulations.

Furthermore, we address the potential influence of the 2008 financial crisis and the 2011 change in the ASIE database inclusion criteria, which raised the threshold for inclusion of non-SOEs with annual sales from 5 million to 20 million RMB. To test the robustness of our findings to these events, we conduct two additional analyses: (1) excluding the years 2008 and 2009, and (2) restricting the sample to non-SOEs with annual sales consistently above 20 million RMB throughout the study period. The estimation results, reported in Appendix Tables A10 and A11, confirm the robustness of our main results to these potential confounding factors.

[25] The estimation results are robust to using alternative distribution cutoffs to define dirtiness.

[26] The officials' data are obtained from Yao et al. (2022).

7. Conclusion

The comparison between command-and-control and market-based approaches to pollution regulation has been a central issue in environmental economics, carrying significant implications for policymakers. China's unique policy landscape, characterized by the phased implementation of COD regulations—including an initial emissions discharge fee, followed by emissions caps, and subsequently cap-and-trade programs—offers a unique quasi-experimental setting to empirically assess the relative efficiency of these regulatory tools.

Our empirical analysis demonstrates that cap-and-trade programs led to higher firm-level output, consistent with the aggregate outcomes observed during China's eleventh five-year plan, where provinces adopting cap-and-trade experienced faster economic growth while achieving comparable COD reductions. This is further supported by our mechanism analysis, which shows that cap-and-trade facilitated the reallocation of emissions and output towards more productive firms, thereby enhancing overall efficiency. These findings align with the predictions of our heterogeneous-firm model under imperfect competition, highlighting that emissions trading can promote efficiency gains even in the presence of market imperfections.

Our study yields two key policy implications. First, emissions caps disrupt allocative efficiency by disproportionately burdening more productive firms. This highlights the advantages of market-based approaches, such as cap-and-trade, particularly in sectors with substantial productivity heterogeneity. An important avenue for future research is to explore the relationship between productivity dispersion and the distortions resulting from the negative reallocation effects of cap policies. This would provide a foundation for a comprehensive cost-benefit analysis of transitioning from emissions caps to cap-and-trade systems. Second, while cap-and-trade policies mitigate these inefficiencies through permit reallocation, they fall short of achieving the first-best outcomes in the presence of imperfect competition. To enhance regulatory efficiency, cap-and-trade policies should be complemented with policy tools that promote market competition.

Tables and Figures

Table 1. Summary Statistics

Variable	Obs.	Mean	Std. Dev.	Min	Max
Panel A: Emissions Cap, 2001 - 2010					
Firms with Reduction Targets above median					
Emission intensity (ln)	21315	-10.21	3.97	-23.74	-3.71
Productivity (ln)	21315	0.65	0.09	0.40	0.89
COD emissions (ton)	21315	58.13	283.83	0.00	12,000.00
Output (million)	21315	328.36	2,136.99	1.14	133,587.73
Firms with Reduction Targets below median					
Emission intensity (ln)	20535	-11.093	4.357	-23.747	-3.709
Productivity (ln)	20535	0.674	0.121	0.234	1.082
COD emissions (ton)	20535	63.331	320.826	0	8859.375
Output (million)	20535	397.767	2935.465	0.243	220273.52
Panel B: Emissions Cap-and-Trade, 2006 - 2013					
Firms with Cap-and-Trade					
Emission intensity (ln)	50376	-10.61	3.23	-28.05	-0.04
Productivity (ln)	50376	0.70	0.16	0.32	1.31
COD emissions (ton)	50376	63.77	486.72	0.00	26452.95
Output (million)	50376	520.58	3115.65	0.01	220273.52
Firms without Cap-and-Trade					
Emission intensity (ln)	32880	-10.206	3.241	-28.884	-0.598
Productivity (ln)	32880	0.712	0.145	0.332	1.33
COD emissions (ton)	32880	57.058	347.358	0	19343.084
Output (million)	32880	489.055	3122.191	0.633	152233.56

Note: This table reports summary statistics for the key variables used in the firm-level regressions. Emission intensity and productivity are reported in log values. COD emissions are measured in tons, while output is measured in million yuan.

Table 2. The Impact of Emissions Cap on Firm-Level Emission Intensity

	ln(Emission intensity)				
	(1)	(2)	(3)	(4)	(5)
Post × Target × ln(φ)	-0.204**	-0.197***	-0.200**	-0.203**	-0.350***
	(0.0899)	(0.0715)	(0.0907)	(0.0924)	(0.109)
Post × ln(φ)	3.316***	2.743***	3.310***	3.345***	4.875***
	(1.158)	(0.900)	(1.159)	(1.174)	(1.389)
Adjusted R^2	0.65	0.60	0.65	0.65	0.65
No. of Obs.	41850	51904	41850	39025	24903
No. of Firms	7327	8580	7327	7282	4427
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Trade Effect			✓		
Sample	Full	Full	Full	No Trade	No Trade-Prefec

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap policy on firm-level COD emission intensities (2001-2010). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients.

(ii) Column (1) reports the baseline DiDiD estimates controlling for prefecture-industry-year fixed effects. Column (2) controls for province-industry-year fixed effects. Column (3) controls for the trade effect explicitly. Column (4) excludes prefectures after they started cap-and-trade programs. Column (5) excludes cap-and-trade prefectures for all years.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 3. The Impact of Emissions Trade on Firm-Level Emission Intensity

	ln(Emission Intensity)				
	(1)	(2)	(3)	(4)	(5)
Post $\times$ Trade $\times$ $\ln(\varphi)$	0.421*	0.307	0.467**	0.311	0.552**
	(0.233)	(0.210)	(0.231)	(0.268)	(0.258)
Post $\times$ Trade		-0.593***			
		(0.204)			
Adjusted R^2	0.73	0.73	0.73	0.72	0.73
No. of Obs.	113036	134519	113036	99184	102900
No. of Firms	14287	16285	14287	12919	13424
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Cap Effect			✓	✓	✓
Window	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients. Double interaction terms $D^r \times \ln(\varphi_i)$ in equation (18) are all included in the regressions.

(ii) Column (1) reports the baseline stacked DiDiD estimates with event years 2008-2011 and an event time window $[-2, +2]$. Column (2) controls for the province-industry-year fixed effects. Column (3) controls for the cap effect. Column (4) examines event years 2009-2011 and an associated time window $[-3, +2]$. Column (5) examines event years 2008-2010 and an associated time window $[-2, +3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 4. The Average Effect of Emissions Cap-and-Trade on Firm-Level Output

	ln(Output)		
	(1)	(2)	(3)
Post × Trade	0.103***	0.0986***	0.0271***
	(0.00977)	(0.00983)	(0.00896)
Adjusted R^2	0.86	0.86	0.87
No. of Obs.	191647	168431	165109
No. of Firms	22432	20388	19704
Firm FE	✓	✓	✓
Year FE	✓	✓	✓
Cap Effect	✓	✓	✓
Window	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the average impact of the emissions trade policy on firm-level output (2006-2013). We use the natural logarithm of output to facilitate the interpretation of the coefficients.

(ii) The stacked dataset in column (1) is constructed with an event time window $[-2, +2]$; columns (2) corresponds to a time window $[-3, +2]$; columns (3) corresponds to a time window $[-2, +3]$. The interaction terms between COD emissions reduction target and year dummies are included to control for the cap effect.

(iii) Standard errors clustered at the firm-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 5. Reallocation of Emissions after Emissions Trade

	ln(COD Emissions)		
	(1)	(2)	(3)
Post $\times$ Trade $\times$ ln(φ)	0.383**	0.469***	0.343**
	(0.150)	(0.157)	(0.151)
Adjusted R^2	0.86	0.85	0.85
No. of Obs.	115164	100218	104239
No. of Firms	14735	13269	13837
Firm FE	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓
Window	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emissions (2006-2013). We use the natural logarithm of emissions and productivity to facilitate the interpretation of the coefficients.

(ii) Column (1) presents the baseline stacked estimates with event years 2008-2011 and the associated time window $[-2, 2]$. Column (2) examines event years 2009-2011 and the associated time window $[-3, 2]$. Column (3) examines event years 2008-2010 and the associated time window $[-2, 3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table 6. Emissions Cap-and-Trade Effect: Subsample Analysis

	ln(Emission Intensity)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Regression Sample	Importer	Non-Imp.	SOEs	Non-SOEs	Dirty	Clean	Old	New
Post $\times$ Trade $\times$ ln(φ)	0.887	0.636**	1.393**	0.791**	1.206***	0.758*	-0.205	0.705**
	(0.777)	(0.255)	(0.594)	(0.311)	(0.375)	(0.438)	(0.806)	(0.338)
Adjusted R^2	0.70	0.74	0.72	0.74	0.75	0.71	0.74	0.73
No. of Obs.	21068	81475	26315	73807	34667	43348	30095	43560
No. of Firms	2887	11451	3966	10001	4521	6243	4618	7189
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓	✓	✓	✓	✓	✓
Control Cap	✓	✓	✓	✓	✓	✓	✓	✓
Window	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensity (2006-2013). Columns (1) and (2) compare importers and non-importers; columns (3) and (4) compare SOEs and non-SOEs; columns (5) and (6) compare dirty and clean industries; columns (7) and (8) compare firms located in prefectures with a new party secretary (tenure below sample median) to those with an old party secretary (tenure above sample median) at the time of cap-and-trade program implementation. We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients.

(ii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

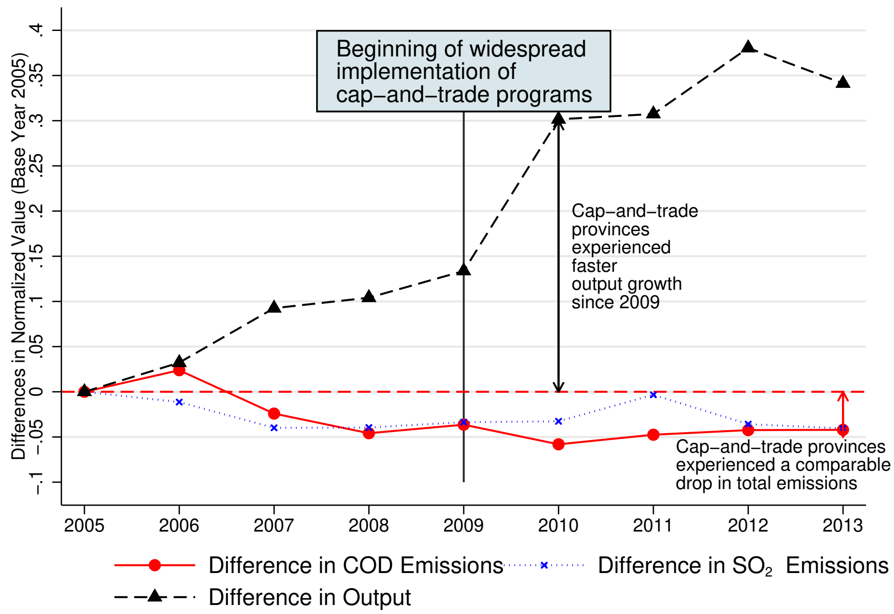


Figure 1. Differences in Output and Pollutant Emissions (COD and SO_2), Cap-and-Trade vs. Cap-Only Regions

Note: This figure shows the differences in COD and SO_2 emissions, as well as real output (normalized by their respective levels in 2005), between regions with and without emissions cap-and-trade programs. The red solid line displays the differences in annual COD emissions and the blue dotted line displays the differences in annual SO_2 emissions. The black dashed line demonstrates the differences in annual output. The vertical line indicates 2009, the year when the widespread implementation of COD cap-and-trade programs began.

Source: China's Environmental Statistical Database.

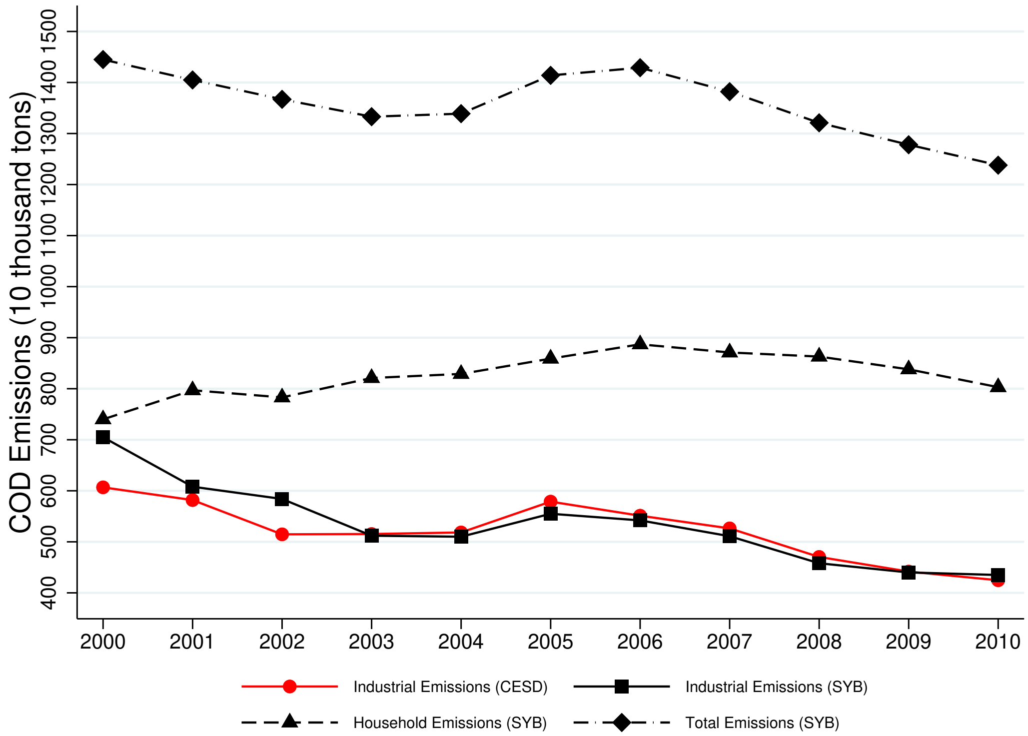
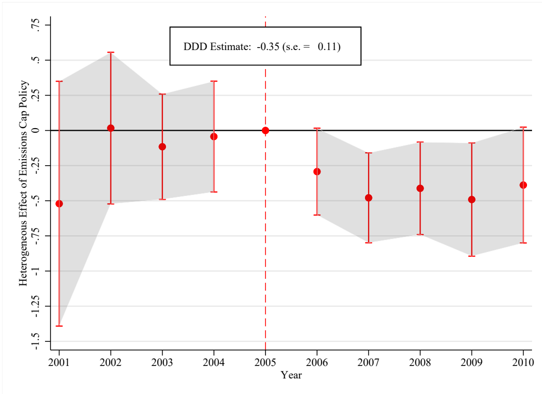


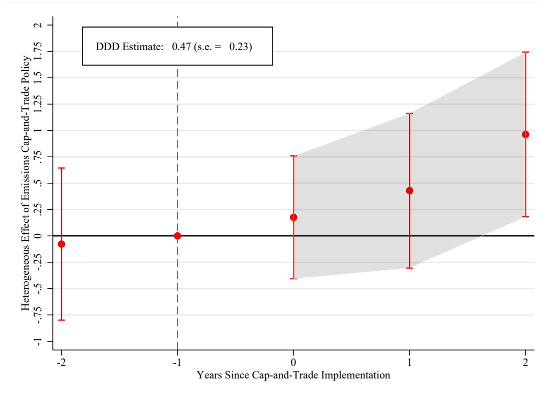
Figure 2. Annual COD Emissions 2000-2010

Note: Before 2010, total COD emissions reported in the statistical yearbook (SYB) included industrial emissions (solid line with squares) and household emissions (dashed line with triangles). Industrial emissions reported by SYB were calculated from the CESD. We find that our sum of firm-level COD emissions from the CESD (solid line with circles) closely tracks the SYB data since 2003. Household emissions were calculated by multiplying a COD emission coefficient by the size of the urban population.

Source: China's Environmental Statistical Database and Statistical Yearbook.



A Cap effect



B Trade effect

Figure 3. Event-Study Estimates of the Emissions Cap and Trade Effects

Note: This figure displays the event-study plots for the impact of emissions cap and trade policies on firm-level emission intensities. Subfigure 3A plots the estimated coefficients (red circles) and the corresponding 95% confidence intervals (bars) from column (3) of Appendix Table A4. The year before cap implementation (2005) is omitted, so the estimate is normalized to zero for that year. Subfigure 3B plots the estimated coefficients and the corresponding 95% confidence intervals from column (1) of Appendix Table A5. The year before cap-and-trade implementation is omitted, so the estimate is normalized to zero for that year.

References

Ackerberg, Daniel A, Kevin Caves, and Garth Frazer. 2015. "Identification properties of recent production function estimators." *Econometrica* 83 (6): 2411–2451.

Anouliès, Lisa. 2017. "Heterogeneous firms and the environment: a cap-and-trade program." *Journal of Environmental Economics and Management* 84: 84–101.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *The Review of Economic Studies*: 1–33.

Brandt, Loren, John Litwack, Elitza Mileva, Luhang Wang, Yifan Zhang, and Luan Zhao. 2020. "China's Productivity Slowdown and Future Growth Potential." World Bank, Washington, DC.

Brandt, Loren, Johannes Van Biesebroeck, and Yifan Zhang. 2012. "Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing." *Journal of Development Economics* 97 (2): 339–351.

Buchanan, James M. 1969. "External diseconomies, corrective taxes, and market structure." *American Economic Review* 59 (1): 174–177.

Cai, Xiqian, Yi Lu, Mingqin Wu, and Linhui Yu. 2016. "Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China." *Journal of Development Economics* 123: 73–85.

Cao, Jing, Mun S. Ho, Rong Ma, and Fei Teng. 2021. "When Carbon Emission Trading Meets a Regulated Industry: Evidence from the Electricity Sector of China." *Journal of Public Economics* 200: 104470.

Carlson, Curtis, Dallas Burtraw, Maureen Cropper, and Karen L Palmer. 2000. "Sulfur dioxide control by electric utilities: What are the gains from trade." *Journal of Political Economy* 108 (6): 1292–1326.

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer. 2019. "The Effect of Minimum Wages on Low-Wage Jobs." *The Quarterly Journal of Economics* 134 (3): 1405–1454.

Chen, Zhao, Matthew E Kahn, Yu Liu, and Zhi Wang. 2018. "The consequences of spatially differentiated water pollution regulation in China." *Journal of Environmental Economics and Management* 88: 468–485.

Cherniwchan, Jevan, Brian R Copeland, and M Scott Taylor. 2017. "Trade and the environment: New methods, measurements, and results." *Annual Review of Economics* 9: 59–85.

Coase, Ronald H. 1960. "The Problem of Social Cost." *Journal of Law and Economics* 3 (1): 2.

Copeland, Brian R., and M. Scott Taylor. 1994. "North-South Trade and the Environment." *The Quarterly Journal of Economics* 109 (3): 755–787.

Copeland, Brian R, and M Scott Taylor. 1995. "Trade and transboundary pollution." *American Economic Review*: 716–737.

Copeland, Brian R, and M Scott Taylor. 2003. *Trade and the Environment.*: Princeton University Press.

Crocker, Thomas D. 1966. "The structure of atmospheric pollution control systems." In *The Economics of Air Pollution*, edited by Harold Wolozin, 61–86. New York: W. W. Norton.

Curtis, E Mark. 2018. "Who loses under cap-and-trade programs? the labor market effects of the nox budget trading program." *Review of Economics and Statistics* 100 (1): 151–166.

de Chaisemartin, Clément, and Xavier D'Haultfœuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–2996.

Deshpande, Manasi, and Yue Li. 2019. “Who Is Screened Out? Application Costs and the Targeting of Disability Programs.” *American Economic Journal: Economic Policy* 11 (4): 213–248.

Fan, Haichao, Joshua S Graff Zivin, Zonglai Kou, Xueyue Liu, and Huanhuan Wang. 2019. “Going Green in China: Firms’ Responses to Stricter Environmental Regulations.” working paper, National Bureau of Economic Research.

Ferris, Ann E, Ronald J Shadbegian, and Ann Wolverton. 2014. “The effect of environmental regulation on power sector employment: Phase I of the title IV SO2 trading program.” *Journal of the Association of Environmental and Resource Economists* 1 (4): 521–553.

Forslid, Rikard, Toshihiro Okubo, and Karen Helene Ulltveit-Moe. 2018. “Why are firms that export cleaner? International trade, abatement and environmental emissions.” *Journal of Environmental Economics and Management* 91: 166–183.

Fowlie, Meredith, Stephen P Holland, and Erin T Mansur. 2012. “What do emissions markets deliver and to whom? Evidence from Southern California’s NOx trading program.” *American Economic Review* 102 (2): 965–93.

Fowlie, Meredith, Mar Reguant, and Stephen P Ryan. 2016. “Market-based emissions regulation and industry dynamics.” *Journal of Political Economy* 124 (1): 249–302.

Goodman-Bacon, Andrew. 2021. “Difference-in-Differences with Variation in Treatment Timing.” *Journal of Econometrics* 225 (2): 254–277.

Goulder, Lawrence H., Xianling Long, Jieyi Lu, and Richard D. Morgenstern. 2022. “China’s Unconventional Nationwide CO2 Emissions Trading System: Cost-effectiveness and Distributional Impacts.” *Journal of Environmental Economics and Management* 111: 102561.

Hahn, Robert D, and Roger G Noll. 1982. “Designing a market for tradable emissions permits.” In *Reform of Environmental Regulation,* edited by W. Magat, 119–146. Cambridge: Ballinger.

Hahn, Robert W., and Robert N. Stavins. 2011. “The Effect of Allowance Allocations on Cap-and-Trade System Performance.” *The Journal of Law and Economics* 54 (S4): S267–S294.

He, Guojun, Shaoda Wang, and Bing Zhang. 2020. “Watering down environmental regulation in China.” *The Quarterly Journal of Economics* 135 (4): 2135–2185.

Hering, Laura, and Sandra Poncet. 2014. “Environmental policy and exports: Evidence from Chinese cities.” *Journal of Environmental Economics and Management* 68 (2): 296–318.

Holladay, J. Scott. 2016. “Exporters and the Environment.” *Canadian Journal of Economics/Revue canadienne d’économique* 49 (1): 147–172.

Holz, Carsten A, and SUN Yue. 2018. “Physical capital estimates for China’s provinces, 1952–2015 and beyond.” *China Economic Review* 51: 342–357.

Kahn, Matthew E, Pei Li, and Daxuan Zhao. 2015. “Water pollution progress at borders: the role of changes in China’s political promotion incentives.” *American Economic Journal: Economic Policy* 7 (4): 223–42.

Keiser, David A, and Joseph S Shapiro. 2019. “Consequences of the Clean Water Act and the Demand for Water Quality.” *The Quarterly Journal of Economics* 134 (1): 349–396.

Khezr, Peyman, and Ian A MacKenzie. 2018. “Consignment auctions.” *Journal of Environmental Economics and management* 87 (January): 42–51.

Konishi, Yoshifumi, and Nori Tarui. 2015. “Emissions trading, firm heterogeneity, and intra-industry reallocations in the long run.” *Journal of the Association of Environmental and Resource Economists* 2 (1): 1–42.

Kwon, Ohyun, Hao Zhao, and Min Qiang Zhao. 2023. "Global Firms and Emissions: Investigating the Dual Channels of Emissions Abatement." *Journal of Environmental Economics and Management* 118: 102772.

Kwon, Ohyun, Hao Zhao, and Min Qiang Zhao. 2024. "Product level emission intensities: Measurement and application." LeBow College of Business, Drexel University.

Levinsohn, James, and Amil Petrin. 2003. "Estimating production functions using inputs to control for unobservables." *The Review of Economic Studies* 70 (2): 317–341.

Liu, Mengdi, Ronald Shadbegian, and Bing Zhang. 2017. "Does environmental regulation affect labor demand in China? Evidence from the textile printing and dyeing industry." *Journal of Environmental Economics and Management* 86: 277–294.

Malueg, David A. 1990. "Welfare consequences of emission credit trading programs." *Journal of Environmental Economics and management* 18 (1): 66–77.

Manjón, Miguel, and Juan Manez. 2016. "Production function estimation in Stata using the Ackerberg–Caves–Frazer method." *The Stata Journal* 16 (4): 900–916.

Martin, Ralf, Mirabelle Muûls, and Ulrich J Wagner. 2016. "The impact of the European Union Emissions Trading Scheme on regulated firms: what is the evidence after ten years?" *Review of Environmental Economics and Policy* 10 (1): 129–148.

Ministry of Environmental Protection. 2006. "The Guidelines for the Allocation of the Total Emissions of the Major Water Pollutants." Online.

Montgomery, W David. 1972. "Markets in licenses and efficient pollution control programs." *Journal of Economic Theory* 5 (3): 395–418.

Muller, Nicholas Z, and Robert Mendelsohn. 2009. "Efficient pollution regulation: getting the prices right." *American Economic Review* 99 (5): 1714–39.

Oates, Wallace E, and Diana L Strassmann. 1984. "Effluent fees and market structure." *Journal of Public Economics* 24 (1): 29–46.

Olley, G Steven, and Ariel Pakes. 1996. "The Dynamics of Productivity in the Telecommunications Equipment." *Econometrica* 64 (6): 1263–1297.

Porter, Michael. 1991. "America's Green Strategy." *Scientific American*: 168.

Rambachan, Ashesh, and Jonathan Roth. 2023. "A More Credible Approach to Parallel Trends." *The Review of Economic Studies* 90 (5): 2555–2591.

Ren, Shenggang, Donghua Liu, Bo Li, Yangjie Wang, and Xiaohong Chen. 2020. "Does emissions trading affect labor demand? Evidence from the mining and manufacturing industries in China." *Journal of Environmental Management* 254: 109789.

Roth, Jonathan, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. 2023. "What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature." *Journal of Econometrics* 235 (2): 2218–2244.

Sartzetakis, Eftichios S. 1997. "Tradeable emission permits regulations in the presence of imperfectly competitive product markets: welfare implications." *Environmental and Resource Economics* 9: 65–81.

Shao, Shuai, Zhenbing Yang, Le Xu, and Lili Yang. 2024. "Environmental Protection Experience of Secretaries and Effectiveness of Environmental Governance: Evidence from COD Discharge in China." *Journal of Economic Behavior & Organization* 220: 237–253.

Shapiro, Joseph S, and Reed Walker. 2018. "Why is pollution from US manufacturing declining?

The roles of environmental regulation, productivity, and trade." *American Economic Review* 108 (12): 3814–54.

Shi, Xinzheng, and Zhufeng Xu. 2018. "Environmental regulation and firm exports: Evidence from the eleventh Five-Year Plan in China." *Journal of Environmental Economics and Management* 89: 187–200.

Sun, Junxiu, Feng Wang, Haitao Yin, and Bing Zhang. 2019. "Money talks: the environmental impact of China's green credit policy." *Journal of Policy Analysis and Management* 38 (3): 653–680.

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

Wang, Chunhua, JunJie Wu, and Bing Zhang. 2018. "Environmental regulation, emissions and productivity: Evidence from Chinese COD-emitting manufacturers." *Journal of Environmental Economics and Management* 92: 54–73.

Weitzman, Martin L. 1974. "Prices vs. quantities." *The Review of Economic Studies* 41 (4): 477–491.

Wing, Coady, Seth M. Freedman, and Alex Hollingsworth. 2024. "Stacked Difference-in-Differences." National Bureau of Economic Research.

Wu, Haoyi, Huanxiu Guo, Bing Zhang, and Maoliang Bu. 2017. "Westward movement of new polluting firms in China: Pollution reduction mandates and location choice." *Journal of Comparative Economics* 45 (1): 119–138.

Xu, Ao, Yin-Hu Wu, Zhuo Chen, Guangxue Wu, Qianyuan Wu, Fangqiong Ling, Wei E Huang, and Hong-Ying Hu. 2020. "Towards the new era of wastewater treatment of China: Development history, current status, and future directions." *Water Cycle* 1: 80–87.

Yao, Yang, Lixing Li, Tianyang Xi, He Wang, Feng Wan, Qian Zhang, Songrui Liu, and Shundong Zhang. 2022. "CCER Officials Dataset." Peking University Open Research Data Platform.

Ye, Weili, Lingxuan Liu, and Bing Zhang. 2020. "Designing and Implementing Pollutant Emissions Trading Systems in China: A Twelve-Year Reflection." *Journal of Environmental Management* 261: 110207.

Zhang, Bing, Xiaolan Chen, and Huanxiu Guo. 2018. "Does central supervision enhance local environmental enforcement? Quasi-experimental evidence from China." *Journal of Public Economics* 164: 70–90.

Zhang, Bing, Hanxun Fei, Pan He, Yuan Xu, Zhanfeng Dong, and Oran R Young. 2016. "The indecisive role of the market in China's SO2 and COD emissions trading." *Environmental Politics* 25 (5): 875–898.

Online Appendix: The Reallocation Effect of Cap-and-Trade

Ohyun Kwon Hao Zhao, Min Qiang Zhao

Appendix Contents

A Additional Figures and Tables **1**

B Construction of Real Capital Stock and Firm-level TFP **14**

B.1 Capital Stock 14

B.2 TFP Measures 14

C Model: Heterogeneous Firms under Sequential Environmental Regulation Changes **16**

C.1 Utility and Demand 16

C.2 Production and Emissions in the Benchmark Equilibrium 17

C.3 Shadow Price of Emissions 19

C.4 Two Environmental Policies to Curb Emissions 19

C.5 A Comparison between Two Policies 21

C.6 Further Discussions 25

D Proofs **27**

D.1 Proof of Lemma A1 27

D.2 Proof of Lemma A2 27

D.3 Proof of Proposition A1 28

D.4 Proof of Proposition A2 29

D.5 Proof of Proposition A3 30

E Model Results under Alternative Values of η **37**

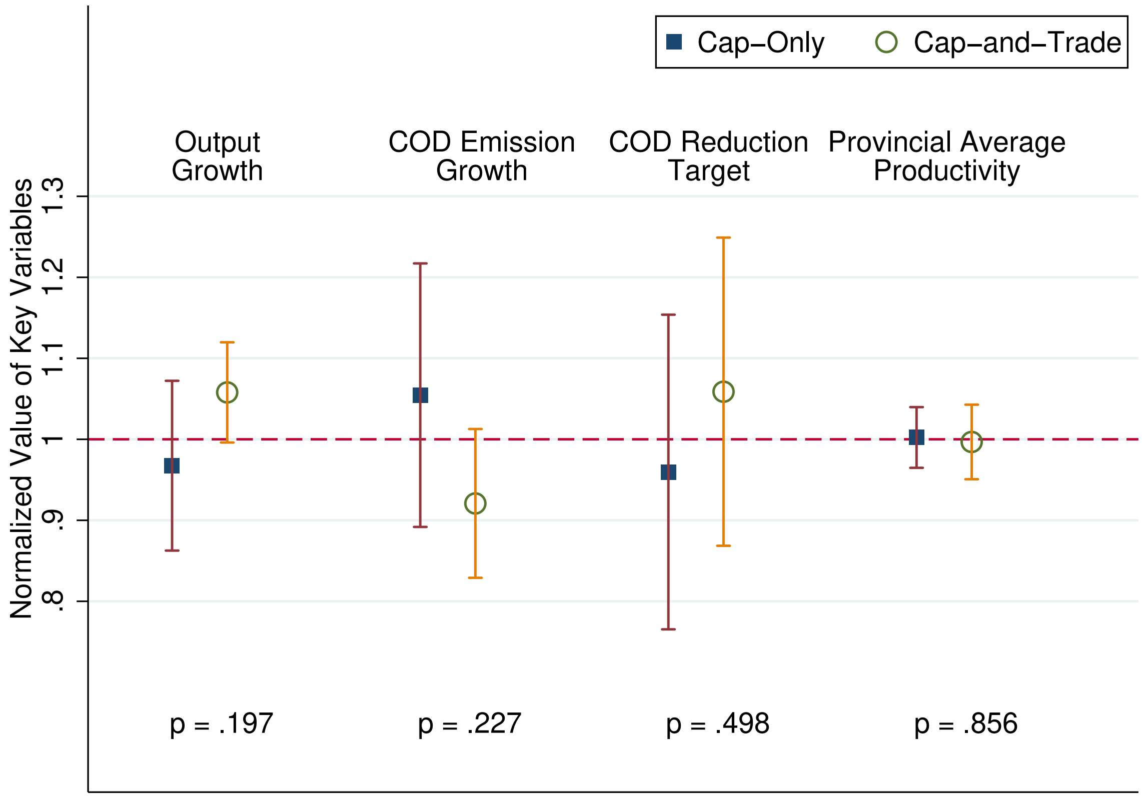
E.1 Unitary Elasticity of Substitution between Inputs 37

E.2 Inputs as Gross Complements 37

F Estimation of Elasticity of Substitution between Inputs **39**

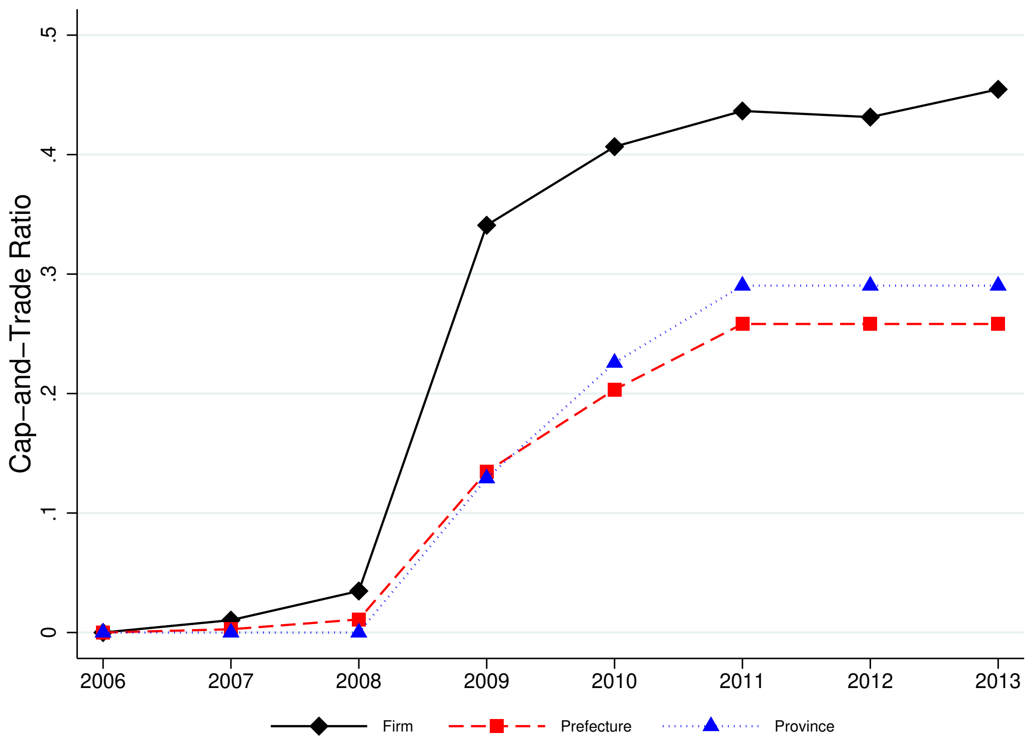
Appendix A. Additional Figures and Tables

Figure A1. Comparison of Key Variables between Cap-and-Trade and Cap-Only Regions



Note: This figure presents t-tests comparing provincial-level output growth (05-08), COD emissions growth (05-08), COD reduction targets, and average productivity (2005) between provinces with cap-and-trade programs and those without. The variable values are normalized by their respective sample means. For output growth, the t-test uses provincial output in 2005 as weights to account for different baseline output values. The lines illustrate 95% confidence intervals. Cap-only regions are marked with squares and cap-and-trade regions are marked with circles. The p-values for the t-tests are reported correspondingly at the bottom.

Figure A2. Sample Share under a Cap-and-Trade Regime



Note: This figure shows how the treatment group under a cap-and-trade regime changes over time. The reported shares include the fractions of firms, prefectures, and provinces under cap-and-trade within the entire sample. A firm is considered to be under a cap-and-trade regime if it is located in a prefecture that has implemented cap-and-trade.

Table A1. Implementation of COD Cap-and-Trade Programs

Province	Year	Pilot Prefectures	Pilot Year
Tianjin	2009		
Chongqing	2010		
Henan	-	Luoyang, Jiaozuo, Sanmenxia, Pingdingshan	2009
Hubei	2009		
Jiangsu	-	Suzhou, Wuxi, Nanjing, Changzhou, Zhenjiang	2009
Zhejiang	2009	Jiaxing	2007
		Shaoxing	2008
Shaanxi	2010		
Hebei	2011	Tangshan	2009
		Baoding	2008
Hunan	2011	Changsha	2008
Inner Mongolia	2010		
Shanxi	2011		

Note: The regions and years for COD cap-and-trade programs are mainly compiled from the website of the Ministry of Ecology and Environment (MEE, previously known as the Ministry of Environmental Protection), with supplemental details from local government websites and news outlets. Henan and Jiangsu did not implement a province-wide cap-and-trade program until 2013.

Table A2. The Average Impact of Emissions Cap on Firm-level Emission Intensities

	ln(Emission intensity)	
	(1)	(2)
Post × Target	-0.0823***	-0.0742***
	(0.0189)	(0.0176)
Adjusted R^2	0.54	0.55
No. of Obs.	208060	199146
No. of Firms	32594	31524
Firm FE	✓	✓
Year FE	✓	
Ind-Year FE		✓

Note:

(i) This table presents the estimation results for the average impact of the emissions cap policy on firm-level COD emissions (2001-2010). We use the natural logarithm of emission intensity to facilitate the interpretation of the coefficients.

(ii) Column (1) presents the baseline estimates with firm and year fixed effects. Column (2) controls for industry-year fixed effects.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A3. The Average Impact of Emissions Trade on Firm-level Emission Intensities

	ln(Emission Intensity)		
	(1)	(2)	(3)
Post × Trade	-0.199	-0.158	-0.0625
	(0.138)	(0.144)	(0.127)
Adjusted R^2	0.67	0.64	0.66
No. of Obs.	159065	140836	138559
No. of Firms	20224	18582	17954
Firm FE	✓	✓	✓
Ind-Year FE	✓	✓	✓
Window	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the average impact of the emissions trade policy on firm-level COD emissions (2006-2013). We use the natural logarithm of emission intensity to facilitate the interpretation of the coefficients.

(ii) Column (1) presents the baseline estimates with firm and year fixed effects. Column (2) controls for industry-year fixed effects.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A4. The Impact of Emissions Cap on Firm-Level Emission Intensity: Event Study

	ln(Emission intensity)		
	(1)	(2)	(3)
Target × 2001 × log(ϕ)	-0.259	-0.280	-0.521
	(0.388)	(0.389)	(0.444)
Target × 2002 × log(ϕ)	0.147	0.119	0.0168
	(0.233)	(0.239)	(0.275)
Target × 2003 × log(ϕ)	0.00975	0.00436	-0.116
	(0.172)	(0.173)	(0.191)
Target × 2004 × log(ϕ)	0.0852	0.0722	-0.0434
	(0.167)	(0.168)	(0.201)
Target × 2006 × log(ϕ)	-0.111	-0.104	-0.293*
	(0.152)	(0.148)	(0.158)
Target × 2007 × log(ϕ)	-0.240*	-0.245*	-0.479***
	(0.137)	(0.137)	(0.163)
Target × 2008 × log(ϕ)	-0.161	-0.156	-0.412**
	(0.139)	(0.143)	(0.168)
Target × 2009 × log(ϕ)	-0.224	-0.300	-0.491**
	(0.159)	(0.183)	(0.206)
Target × 2010 × log(ϕ)	-0.209	-0.240	-0.388*
	(0.165)	(0.190)	(0.210)
Adjusted R^2	0.65	0.65	0.65
No. of Obs.	41850	39025	24903
No. of Firms	7327	7282	4427
Firm FE	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓
Sample	Full	No Trade	No Trade-Prefec

Note:

(i) This table presents the event study estimates for the heterogeneous impact of the emissions cap policy on firm-level COD emission intensities (2001-2010).

(ii) Column (1) reports the baseline estimates controlling for prefecture-industry-year fixed effects. Column (2) excludes prefectures after they started cap-and-trade. Column (3) excludes cap-and-trade prefectures for all years.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A5. The Impact of Emissions Trade on Firm-Level Emission Intensity: Event Study

	ln(Emission Intensity)		
	(1)	(2)	(3)
Trade × ln(φ) × Year(-3)		0.695	
		(0.515)	
Trade × ln(φ) × Year(-2)	-0.0775	-0.167	-0.0848
	(0.368)	(0.340)	(0.402)
Trade × ln(φ) × Year(0)	0.176	0.221	0.238
	(0.297)	(0.343)	(0.339)
Trade × ln(φ) × Year(1)	0.429	0.440	0.461
	(0.375)	(0.391)	(0.406)
Trade × ln(φ) × Year(2)	0.962**	0.729*	0.870**
	(0.399)	(0.422)	(0.420)
Trade × ln(φ) × Year(3)			1.033**
			(0.512)
Adjusted R^2	0.73	0.72	0.73
No. of Obs.	113036	99184	102900
No. of Firms	14287	12919	13424
Firm FE	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓
Cap Effect	✓	✓	✓
Window	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the event study estimates for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients.

(ii) Column (1) presents the stacked estimates with event years 2008-2011 and an event time window $[-2, 2]$. Column (2) examines event years 2009-2011 and the associated time window $[-3, 2]$. Column (3) examines event years 2008-2010 and the associated time window $[-2, 3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A6. "Reallocation" of Output after Emissions Trade

	ln(Output)			Output		
	(1)	(2)	(3)	(4)	(5)	(6)
Post × Trade × ln(φ)	0.0114	0.00534	0.0271	44,907	51,496*	76,483***
	(0.0271)	(0.0271)	(0.0302)	(27,249)	(29,167)	(29,132)
Adjusted R^2	0.90	0.90	0.90	0.82	0.81	0.82
No. of Obs.	104095	91466	94337	104241	91571	94472
No. of Firms	11252	10130	10511	11263	10136	10519
Firm FE	✓	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓	✓	✓	✓
Cap Effect	✓	✓	✓	✓	✓	✓
Window	$[-2,+2]$	$[-3,+2]$	$[-2,+3]$	$[-2,+2]$	$[-3,+2]$	$[-2,+3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions trade policy on firm-level output (2006-2013). We use the natural logarithm of productivity to facilitate the interpretation of the coefficients.

(ii) Columns (1) - (3) present the estimates with the logged value of output as the dependent variable. Columns (4) - (6) present the estimates with output value as the dependent variable.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A7. The Impact of Emissions Caps on Firm-level Output

	ln(Output)	
	(1)	(2)
Post × Target	-0.0142*	-0.0116
	(0.00791)	(0.00764)
Adjusted R^2	0.85	0.85
No. of Obs.	222313	212622
No. of Firms	32878	31782
Firm FE	✓	✓
Year FE	✓	
Ind-Year FE		✓

Note:

(i) This table presents the estimation results for the average impact of the emissions cap policy on firm-level output (2001-2010). We use the natural logarithm of output to facilitate the interpretation of the coefficients.

(ii) Column (1) presents the baseline estimates with firm and year fixed effects. Column (2) controls for industry-year fixed effects.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A8. Emissions Cap Effect on SO_2 Emission Intensity

	ln(Emission intensity)				
	(1)	(2)	(3)	(4)	(5)
Post $\times$ Target $\times$ ln(φ)	-0.0275	-0.0452	-0.0287	-0.0365	0.0359
	(0.0606)	(0.0627)	(0.0610)	(0.0640)	(0.0694)
Post $\times$ ln(φ)	0.364	0.626	0.365	0.420	-0.330
	(0.707)	(0.780)	(0.707)	(0.733)	(0.792)
Adjusted R^2	0.80	0.79	0.80	0.80	0.82
No. of Obs.	38707	48635	38707	36446	23087
No. of Firms	7081	8314	7081	7017	4269
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Trade Effect			✓		
Sample	Full	Full	Full	No Trade	No Trade-Prefec

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap policy on firm-level SO_2 emission intensities (2001-2010). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients.

(ii) Column (1) reports the baseline DiDiD estimates controlling for prefecture-industry-year fixed effects. Column (2) controls for province-industry-year fixed effects. Column (3) controls for the trade effect explicitly. Column (4) excludes prefectures after they started cap-and-trade. Column (5) excludes cap-and-trade prefectures for all years.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A9. Emissions Cap-and-Trade Effect on SO_2 Emission Intensity

	ln(Emission Intensity)				
	(1)	(2)	(3)	(4)	(5)
Post $\times$ Trade $\times$ $\ln(\varphi)$	0.284	0.404**	0.192	0.331	0.0295
	(0.230)	(0.197)	(0.230)	(0.239)	(0.262)
Post $\times$ Trade		-0.328*			
		(0.185)			
Adjusted R^2	0.83	0.84	0.83	0.83	0.83
No. of Obs.	89993	110890	89993	78453	80481
No. of Firms	11182	13101	11182	10091	10316
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Cap Effect			✓	✓	✓
Window	$[-2,+2]$	$[-2,+2]$	$[-2,+2]$	$[-3,+2]$	$[-2,+3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level SO_2 emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients. Double interaction terms $D^r \times \ln(\varphi_i)$ in equation (18) are all included in the regressions.

(ii) Column (1) reports the baseline stacked DiDiD estimates with event years 2008-2011 and an event time window $[-2,+2]$. Column (2) controls for the province-industry-year fixed effects. Column (3) controls for the cap effect. Column (4) examines event years 2009-2011 and the associated time window $[-3,+2]$. Column (5) examines event years 2008-2010 and the associated time window $[-2,+3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A10. Emissions Cap-and-Trade Effect: Subsample excluding 2008 and 2009

	ln(Emission Intensity)				
	(1)	(2)	(3)	(4)	(5)
Post × Trade × ln(φ)	0.727*	0.853**	0.921**	0.445	1.104**
	(0.438)	(0.335)	(0.457)	(0.535)	(0.514)
Post × Trade		-1.262***			
		(0.247)			
Adjusted R^2	0.71	0.73	0.71	0.72	0.72
No. of Obs.	59030	79534	59030	55881	55098
No. of Firms	10517	13672	10517	9771	9935
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Cap Effect			✓	✓	✓
Window	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients. The sample excludes the years 2008 and 2009.

(ii) Column (1) reports the baseline stacked DiDiD estimates with event years 2008-2011 and an event time window $[-2, +2]$. Column (2) controls for the province-industry-year fixed effects. Column (3) controls for the cap effect. Column (4) examines event years 2009-2011 and the associated time window $[-3, +2]$. Column (5) examines event years 2008-2010 and the associated time window $[-2, +3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A11. Emissions Cap-and-Trade Effect: Subsample excluding Non-SOEs below 20 Million RMB

	ln(Emission Intensity)				
	(1)	(2)	(3)	(4)	(5)
Post × Trade × ln(φ)	0.632**	0.500**	0.793**	0.551	0.762**
	(0.309)	(0.239)	(0.317)	(0.339)	(0.347)
Post × Trade		-0.811***			
		(0.210)			
Adjusted R^2	0.73	0.72	0.73	0.72	0.72
No. of Obs.	96696	118879	96696	86009	88414
No. of Firms	12156	14351	12156	11099	11451
Firm FE	✓	✓	✓	✓	✓
Prefec-Ind-Year FE	✓		✓	✓	✓
Prov-Ind-Year FE		✓			
Cap Effect			✓	✓	✓
Window	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$	$[-3, +2]$	$[-2, +3]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients. The sample excludes non-SOEs with annual sales below 20 million RMB.

(ii) Column (1) reports the baseline stacked DiDiD estimates with event years 2008-2011 and an event time window $[-2, +2]$. Column (2) controls for the province-industry-year fixed effects. Column (3) controls for the cap effect. Column (4) examines event years 2009-2011 and the associated time window $[-3, +2]$. Column (5) examines event years 2008-2010 and the associated time window $[-2, +3]$.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A12. Emissions Cap Effect: Sensitivity of Productivity Cutoffs

	ln(Emission intensity)		
	(1)	(2)	(3)
Productivity Cutoff	5%	10%	15%
Post $\times$ Target $\times$ ln(φ)	-0.153**	-0.350***	-0.428***
	(0.0658)	(0.109)	(0.129)
Post $\times$ ln(φ)	2.451***	4.875***	6.647***
	(0.808)	(1.389)	(1.678)
Adjusted R^2	0.65	0.65	0.63
No. of Obs.	35330	24903	17077
No. of Firms	6030	4427	3130
Firm FE	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓
Sample	No Trade-Prefec	No Trade-Prefec	No Trade-Prefec

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap policy on firm-level COD emission intensities (2001-2010). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients.

(ii) The three columns report results for subsamples of firms with productivity above the 5%, 10%, and 15% quantiles of the productivity distribution, respectively, excluding the lowest-productivity firms to ensure consistency with the sample used in the cap-and-trade regression, as outlined in Proposition 2. Prefectures that have implemented cap-and-trade programs before 2010 are excluded in the regression sample.

(iii) Standard errors clustered at the province-year level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Table A13. Emissions Cap-and-Trade Effect: Sensitivity of Productivity Cutoffs

	ln(Emission intensity)		
	(1)	(2)	(3)
Productivity Cutoff	5%	10%	15%
Post × Trade × ln(φ)	0.437*	0.467**	0.630**
	(0.246)	(0.231)	(0.282)
Adjusted R^2	0.74	0.73	0.74
No. of Obs.	147854	113036	84936
No. of Firms	18018	14287	11196
Firm FE	✓	✓	✓
Prefec-Ind-Year FE	✓	✓	✓
Cap Effect	✓	✓	✓
Window	$[-2, +2]$	$[-2, +2]$	$[-2, +2]$

Note:

(i) This table presents the estimation results for the heterogeneous impact of the emissions cap-and-trade policy on firm-level COD emission intensities (2006-2013). We use the natural logarithm of emission intensity and productivity to facilitate the interpretation of the coefficients. Double interaction terms $D^r \times \ln(\varphi_i)$ in equation (18) are all included in the regressions.

(ii) The three columns report results for subsamples of firms with productivity above the 5%, 10%, and 15% quantiles of the productivity distribution, respectively, excluding the lowest-productivity firms as outlined in Proposition 2. Stacked DiDiD estimates with event years 2008-2011 and an event time window $[-2, +2]$ are reported.

(iii) Standard errors clustered at the province-year-sample level are reported in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Appendix B. Construction of Real Capital Stock and Firm-level TFP

B.1. Capital Stock

Firms report the book values of their capital stock in the ASIE dataset. These values are the sum of nominal values across years, which provide us the estimates of fixed investment. The key challenge is to compute the initial values of real capital stock that are comparable over time and across firms. We follow the procedure of Brandt et al. (2012) to construct real capital stock measures.

Our panel dataset starts in 2000. The 1993 annual enterprise survey data, kindly provided by Brandt et al. (2012), is used to compute the average growth rate of nominal capital stock between 1993 and 2000 at the two-digit industry level across provinces. Next, the growth rate data is merged with firm-level observations by industry and province. The time series of nominal values of capital stock is constructed by assuming a constant growth rate between each firm's establishment year and the year 2000. Based on the perpetual inventory method, the real capital stock for the year 2000 is calculated using these nominal data with the capital stock deflator constructed by Holz and Yue (2018). The depreciation rate is assumed to be 9%.

For the real capital stock after 2000, we use the actual changes in nominal capital stock at original purchase prices as the estimates of fixed investment. Again, the same deprecation rate and capital-stock deflator are used to compute the real capital stock after 2000. For firms that first appear in the ASIE dataset after 2000, we repeat the same procedure described above, replacing the year 2000 with the year in which the respective firm first appears in the dataset.

B.2. TFP Measures

We apply the method of Ackerberg et al. (2015) (ACF) to estimate firm-level revenue-based TFP. A major econometric challenge in the estimation of production functions is to control for the correlation between unobserved productivity shocks and factor inputs. The idea behind the methods of Olley and Pakes (1996) (OP) and Levinsohn and Petrin (2003) (LP) is to uncover unobserved productivity shocks by inverting optimal input decisions under certain theoretical and statistical assumptions. More specifically, OP identify conditions that allow using firm-level investment as a proxy variable for the firm's productivity, which can then be approximated using a low-order polynomial as a control function. LP take a similar approach but use intermediate inputs instead of investment in the control function. Both OP and LP assume that labor input is not a dynamic input, but the existence of long-term labor contracts and substantial hiring and firing costs could invalidate this assumption, raising identification issues in their two-stage approaches. ACF relax this restriction by incorporating the dynamic impacts of labor input on the firm's future profits.

Before applying the ACF's method, we deflate output and input variables using output and input deflators, respectively. We follow the procedure described in Brandt et al. (2012) to construct these deflators. In the ASIE dataset, output values at both nominal and real prices are reported before 2004. This allows us to compute output price indices as the ratio of nominal to real output and then further aggregate them at the four-digit industry level. For 2004-2013, we use the ex-factory price index at the two-digit industry level from the China Statistical Yearbooks to construct the output deflator. The construction of input deflators relies on information from the output deflators and the 2002 national input-output (IO) table. For each IO sector, the input deflator is computed as a weighted average of output deflators, using the weight variable constructed from the input-output table.

To implement the ACF's estimation procedure, we use the Stata package "acfest" developed by Manjón and Manez (2016). Investment is used as a proxy variable in the control function. The input variables in our production function include labor, capital, and the intermediate input. After 2007, the intermediate input data is not reported in the ASIE dataset, so we construct it by subtracting value added from the sum of gross revenue and value-added taxes.

Appendix C. Model: Heterogeneous Firms under Sequential Environmental Regulation Changes

In this section, we present a theoretical framework to illustrate how heterogeneous firms respond to sequential environmental regulation changes. In the benchmark case, we characterize the equilibrium subject to a constant unit emissions cost.

C.1. Utility and Demand

Consider an economy that admits a representative consumer whose preference is given by

$$U = q_0 + E\ln(Q) - h(Z), \tag{A1}$$

where $Q = \left[\int q(\omega)^{\frac{\sigma-1}{\sigma}} d\omega\right]^{\frac{\sigma}{\sigma-1}}$ is an aggregate over a continuum of differentiated varieties, indexed by ω, σ ($\sigma > 1$) is a measure of substitutability between differentiated varieties, q_0 denotes the consumption of a numéraire good (the price of which is normalized to 1), $h(Z)$ is the disutility associated with total emissions Z,[1] and E is an exogenous expenditure parameter. We adopt a quasi-linear utility function so that the consumption of the numéraire good fully absorbs the income effect, which is not central to our analysis. Firms do not internalize the negative externalities on consumers caused by their emissions Z, and thus there is an incentive for the government to curb emissions.

The numéraire good is produced with $1/w$ units of labor under constant returns to scale technology and sold in a competitive market. We also assume that the labor endowment in the economy is sufficient to ensure a positive amount of production of q_0. These two assumptions pin down the wage at w, which can be treated as a constant in the remainder of our analysis.

Let P denote the aggregate price index of differentiated goods.[2] Solving the representative consumer's utility maximization problem leads to the following variety-wise demand

$$q(\omega) = \Phi\, p(\omega)^{-\sigma}, \tag{A2}$$

where we introduce $\Phi \equiv EP^{\sigma-1}$ to denote the demand shifter for each variety.

[1] In this model, the total emissions Z is set by the government, and given this value, the disutility of emissions perceived by each consumer is exogenously determined. Hence, the theoretical results remain unchanged even if we specify $h(\cdot)$ to be individual-specific.

[2] Specifically, $P = \left[\int p(\omega)^{1-\sigma} d\omega\right]^{\frac{1}{1-\sigma}}$ where $p(\omega)$ denotes the price for variety ω.

C.2. Production and Emissions in the Benchmark Equilibrium

We assume that the economy has a fixed measure of infinitesimal firms.[3] Each firm is a monopoly of a single variety of good ω and owned by the representative consumer.

Following Forslid et al. (2018) and Shapiro and Walker (2018), a firm's final output depends on productive inputs and the fraction diverted to emissions abatement. A firm's production function can be expressed as

$$q = (1-\theta)\phi l, \tag{A3}$$

where ϕ is the firm's TFP, l denotes an aggregate of productive factors, which is referred to as "labor" in this model, and θ is the share of labor dedicated to pollution abatement.

Let z denote the firm's final emissions after abatement. We assume that z is generated with the following technology:

$$z = \phi l \left[\frac{(1-\theta)^{\frac{\eta-1}{\eta}} - 1}{\alpha} + 1 \right]^{\frac{\eta}{\eta-1}}, \tag{A4}$$

where $\alpha \in (0,1)$ measures the efficiency of abatement investment and $\eta > 0$, as will be clear later, captures the elasticity of substitution between labor and emissions in determining the optimal production. The amount of emissions is non-negative: $z \geq 0$, implying an implicit upper bound of the investment θ:

$$\theta \leq 1 - (1-\alpha)^{\frac{\eta}{\eta-1}}. \tag{A5}$$

The upper bound is the share of labor dedicated to pollution abatement such that $z = 0$.

Combining equations (A3) and (A4), we obtain a CES form of production function:

$$q = \left[\alpha z^{\frac{\eta-1}{\eta}} + (1-\alpha)(\phi l)^{\frac{\eta-1}{\eta}} \right]^{\frac{\eta}{\eta-1}}, \tag{A6}$$

where ϕ, the TFP term, turns into labor-augmenting technology, α becomes a share parameter, and η governs the substitution relationship between labor and emissions in the CES production function. After this transformation, emissions enter the production function as an input. This is consistent with a common interpretation in the literature

[3] For tractability, our theoretical model does not take into account the impact of market entry and exit, which does not deviate from the reality in China, as Brandt et al. (2020) pointed out that firm exit contributed negligibly to manufacturing productivity growth. Moreover, with a mild assumption on the productivity distribution, our model can be easily extended to a version with endogenous market access, as in Konishi and Tarui (2015), while all theoretical results still hold.

that emissions can be equivalently modeled as an input to the firm's production (Copeland and Taylor 1994, 1995, 2003; Holladay 2016; Forslid et al. 2018; Shapiro and Walker 2018).

However, our modeling strategy differs from the aforementioned papers in two main aspects. First, by departing from the widely used Cobb-Douglas production function, the CES production function provides a very flexible substitution pattern between inputs. This flexibility, as we will show, is crucial for the cap and cap-and-trade policies to generate differential impacts. Second, our model treats α, the conventional term that captures industry-specific pollution intensity, as a constant. This allows our model to identify a novel reallocation mechanism (from cap to cap-and-trade policy) associated with firm-level productivity, which is different from the conventional reallocation mechanism that hinges on sectoral dirtiness.[4]

Let π denote the operating profit of a firm, which is equal to the firm's revenue net of its variable costs

$$\pi = pq - wl - \tau z, \tag{A7}$$

where p is the output price and τ is a unit emissions discharge fee.

Let $\mu(\phi, \tau)$ denote the marginal cost of production for a firm with productivity ϕ given the emissions cost τ. $\mu(\phi, \tau)$ can be expressed as

$$\mu(\phi, \tau) = \left[\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{1}{1-\eta}}. \tag{A8}$$

The solution to the firm's profit maximization problem admits the following expression for labor l and emissions z in equilibrium:

$$l = \frac{1}{\phi}\left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\Phi\,(1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{-\eta}\left[\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}} \tag{A9}$$

$$z = \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\Phi\alpha^{\eta}\tau^{-\eta}\left[\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}}. \tag{A10}$$

We denote the equilibrium when firms are subject to a unit emissions cost τ as the *benchmark equilibrium*. A subscript b is added to denote the emissions (z_b), output (q_b) and market potential (Φ_b) in the benchmark equilibrium.

Finally, we define the firm-level emission intensity ζ as the firm's emissions per unit of

[4] In our empirical analysis, the impact of industry-specific α will be fully absorbed by time-varying industry fixed effects.

output. In the benchmark equilibrium, it is equal to

$$\zeta_b \equiv \frac{z_b}{q_b} = \left[\alpha + (1-\alpha)\left(\frac{\phi\tau}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}\right]^{\frac{\eta}{1-\eta}}. \tag{A11}$$

C.3. Shadow Price of Emissions

Let λ denote the *shadow price* of emissions, which reflects a firm's willingness to pay for an additional unit of emissions. The shadow price of emissions at emissions level $\bar{z}$ is defined as the Lagrangian multiplier of the following constrained optimization problem:

$$\begin{aligned} &\max\ pq - wl \\ &\text{s.t. } z = \bar{z}. \end{aligned} \tag{A12}$$

In the benchmark equilibrium, we can derive the shadow price of emissions by setting $\bar{z}$ equal to z_b. The following lemma characterizes λ in the benchmark equilibrium.

LEMMA A1. *In the benchmark equilibrium, the shadow price of emissions λ is equal to τ, the cost of emissions.*

Proof. See Appendix D.1 for the proof.

In the benchmark case, the firm faces a constant emissions cost without any quantity constraint on its input of emissions. Therefore, its willingness to pay for an additional unit of emissions at the equilibrium emissions level z_b is determined by the emissions cost τ. Our benchmark equilibrium serves to model the emissions discharge fee regime in China before 2006, in which firms paid a constant fee per unit of emissions. In the following subsections, we investigate the sequential implementation of *pro rata* emissions cap and cap-and-trade policies.

C.4. Two Environmental Policies to Curb Emissions

We examine two environmental policies that aim to reduce the level of emissions. The first policy is a *pro rata* emissions cap policy, a command-and-control regulation that imposes a uniform percentage-wise emissions reduction on each firm based on its benchmark emissions level. The second one is a cap-and-trade policy under which the government distributes emission permits to individual firms that can be traded in a centralized market while preserving the same economy-wide emissions reduction as the cap policy.

C.4.1. Pro Rata Emissions Cap Policy

Suppose that the government aims to reduce the level of total emissions from the benchmark scenario by a factor of $1-\delta$ ($0<\delta<1$). Under the *pro rata* cap policy regime, the government sets the reduction factor $1-\delta$ and applies it to each firm uniformly.[5] In particular, the level of emissions allowed by the government, z_c, is determined by each firm's benchmark emissions level z_b multiplied by δ (i.e., $z_c=\delta z_b$), where the subscript c denotes the cap regime.[6]

As the emissions cap restricts firm-level emissions to be lower than that under the unconstrained equilibrium, it is straightforward to see that the emissions cap is always binding.

LEMMA A2. *An emissions cap that is proportionate to each firm's benchmark emissions level binds for all firms.*

Proof. See Appendix D.2 for the proof.

Next, we characterize λ_c, the shadow price of emissions under the cap policy. With insights from Lemma A1, one can infer that λ_c is equal to a hypothetical (sufficiently high) cost of emissions that would induce the firm to choose an emissions level z_c ($\equiv \delta z_b$) in the benchmark equilibrium. With equation (A10) and the results in Lemma A2, we can formalize this idea with the following equation

$$\text{(A13)} \qquad \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\Phi_c\alpha^\eta\lambda_c^{-\eta}\left[\alpha^\eta\lambda_c^{1-\eta}+(1-\alpha)^\eta\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}}=\delta z_b,$$

where the left-hand side is the level of emissions associated with the emissions cost λ_c in the benchmark equilibrium, and the right-hand side is emissions under cap policy. Equation (A13) provides an implicit solution for λ_c, the shadow price of the emissions cap.

C.4.2. Emissions Cap-and-Trade Policy

Following the assignment of emissions caps, the government can further introduce cap-and-trade programs so that firms can trade emission permits with each other. Assume the emissions market is perfectly competitive. At the competitive equilibrium price, the initial

[5] We note that Chinese firms still have to pay the unit discharge fee τ under the cap regime. This is irrelevant in our model since only the quantity (cap) constraint binds in firms' optimization problem (Lemma A2).

[6] In a dynamic setting in which the government periodically updates the firm-level emissions caps based on the level of production in the preceding periods, firms may manipulate their production levels to obtain additional emission quotas, as discussed in Anouliès (2017). This could give firms an incentive to "overproduce" to secure more emission permits for the following year. In China, local governments set a one-time five-year emissions abatement target in 2006, which needs to be satisfied by 2010. Thus, our model does not consider the dynamic updating of emissions caps.

allocation of emission permits is equivalent to lump-sum transfers between firms that do not distort key equilibrium objects. Thus, for model tractability, we assume that the emission permits are initially entirely owned by the government and there is a Walrasian auctioneer who sets the trading price of emission permits such that the total demand for emission permits meets the environmental goal set by the government.[7]

We use t in the subscript to denote the cap-and-trade regime. The total unit cost of emissions in the cap-and-trade equilibrium is given by λ_t, which is the sum of the emission permits price and the pre-existing emissions cost τ, both of which are the same for all firms.[8] In a similar manner that we established Lemma A1, we infer that the monetary cost of emissions, λ_t, is also the shadow price of emissions under the cap-and-trade policy. Hence, the equilibrium emissions level for a firm with productivity ϕ, $z_t(\phi)$, can be expressed similarly to equation (A10) by replacing the emissions cost τ with λ_t.

To ease the comparison with the cap policy, we consider the case where economy-wide total emissions remain the same in both policy regimes. According to Lemma A2, the emissions cap binds for all firms under the cap regime, hence the total supply of emission permits under the cap-and-trade regime equals the sum of emissions caps on individual firms, expressed as $\int_{\underline{\phi}}^{\bar{\phi}} z_c(\phi)g(\phi)d\phi$, where $g(\phi)$ is the density function of the distribution of ϕ, $\underline{\phi}$ and $\bar{\phi}$ are the lower and upper bounds of ϕ, and $z_c(\phi)$ is the emissions cap for the firm with productivity ϕ. The emissions cost in the cap-and-trade equilibrium is determined by the market clearing condition for emission permits

$$\int_{\underline{\phi}}^{\bar{\phi}} z_t(\phi)g(\phi)d\phi = \int_{\underline{\phi}}^{\bar{\phi}} z_c(\phi)g(\phi)d\phi, \tag{A14}$$

where the left-hand side is the aggregate demand for emission permits in the cap-and-trade regime.

C.5. A Comparison between Two Policies

For the comparison between emissions cap and cap-and-trade policies, we first demonstrate how *pro rata* emissions cap policy distorts the shadow price of emissions from the benchmark regime. According to equation (A13), the shadow price of an emissions cap, λ_c, is an implicit function of firm-level productivity ϕ, and the correlation between λ_c and ϕ depends on the substitution parameter, η.

[7] Several mechanism designs can achieve a similar equilibrium outcome, for example, a continuous double-auction market or consignment auction (Hahn and Noll 1982; Khezr and MacKenzie 2018). In practice, many issues (hoarding of allowances, thin markets, segmentation of markets, etc.) could arise to prevent a cap-and-trade program from functioning effectively (Zhang et al. 2016).

[8] During China's eleventh five-year plan, firms were required to pay a treatment surcharge for their wastewater that contained COD. This requirement continued to exist even when the local government implemented COD emissions trading.

To identify the impact of environmental policy changes, throughout this model, we consider three sets of η values: $1 < \eta < \sigma$, $\eta = 1$, and $\eta < 1$, corresponding to the case when the two inputs are gross substitutes, of unitary elasticity of substitution, and gross complements, respectively. The analysis on the parameter range $\eta > \sigma$ is ignored as empirically η is around 1 while σ is estimated consistently above 3 (Shapiro and Walker 2018). The case in which $\eta > 1$ is our preferred scenario, since in Appendix F, our estimates of η using Chinese manufacturing micro-data range from 1.118 to 1.164. For the completeness of the discussion, we proceed in the main analysis assuming the gross-substitutability but also discuss the predictions under alternative assumptions in Appendix E.

When $1 < \eta < \sigma$, equation (A13) implies $\frac{\partial \lambda_c}{\partial \phi} > 0$, which suggests that more productive firms face a higher shadow price of emissions when all firms experience a uniform percentage-wise emissions reduction. We summarize the key results under the emissions cap policy in the following proposition.

PROPOSITION A1. *When emissions and labor are gross substitutes ($1 < \eta < \sigma$), the pro rata emissions cap policy results in negative reallocation of emissions. Relative to the benchmark equilibrium, the pro rata emissions cap*

1. *increases the shadow price of emissions for all firms: $\lambda_c(\phi) > \lambda_b(\phi)\ \forall\ \phi$, and this increase is higher for more productive firms: $\frac{\partial[\lambda_c(\phi) - \lambda_b(\phi)]}{\partial \phi} > 0$;*
2. *lowers firm-level emission intensity for all firms: $\zeta_c(\phi) < \zeta_b(\phi)\ \forall\ \phi$, and the percentage reduction is greater for more productive firms:*

$$\frac{\partial[\ln(\zeta_c(\phi)) - \ln(\zeta_b(\phi))]}{\partial \phi} < 0. \tag{A15}$$

Proof. See Appendix D.3 for the proof.

The key insight from Proposition A1 is that the emissions cap policy generates dispersion in the shadow price of emissions across firms in a way that is more disadvantageous for more productive firms. Intuitively, more productive firms have a lower emissions share in the benchmark equilibrium ($\frac{\partial \zeta_b}{\partial \phi} < 0$), and thus, a proportionate decline in the use of emissions translates into an excessive increase in the shadow price of emissions. As this schedule of the shadow price of emissions is disproportionately more restrictive to more productive firms, lowering the economy-wide average productivity, we conclude that the cap policy results in negative reallocation.

The observation that a *pro rata* emissions cap has an unintended negative allocative impact, to our knowledge, has not been well recognized in the literature, and this becomes a natural justification for the cap-and-trade policy. A key distinction between the two policy regimes is that the shadow price of emissions is higher for more productive firms under an emissions cap, whereas it is equalized under an emissions cap-and-trade policy. Due to the constraint that total emissions are equalized in both regimes, the shadow price of emissions under the emissions cap-and-trade policy, λ_t, must be less than the maximum of λ_c, the shadow prices of emissions under the emissions cap policy.[9] As illustrated in Figure A3, there exists a cutoff productivity level $\tilde{\phi}$, above which more productive firms are disadvantageous under an emissions cap regime. Conversely, firm-level emission intensities, which decrease with respect to the shadow price of emissions, exhibit an opposite relation.

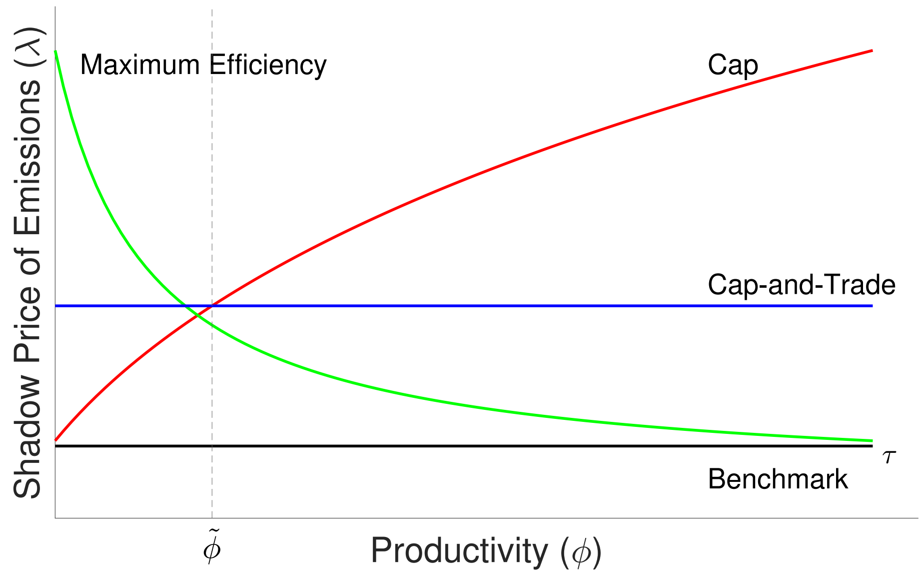


Figure A3. Shadow Prices of Emissions under Different Policy Regimes

Note: This figure illustrates the shadow prices of emissions across firms with different levels of productivity under three different policy regimes and under the maximum production efficiency when $1 < \eta < \sigma$. At the benchmark equilibrium, the shadow prices of emissions are equalized across firms at the level of the benchmark emissions cost τ. Under the emissions cap, the shadow prices of emissions are higher than the benchmark level and increase with respect to firms' productivity levels. Under the cap-and-trade regime, the shadow prices of emissions are also higher than the benchmark level but do not differ across firms. Under the maximum production efficiency, the shadow price of emissions decreases with firm-level productivity. The vertical line indicates the productivity level, $\tilde{\phi}$, where a firm faces the same shadow price under cap and cap-and-trade regimes.

We summarize these results in the following proposition.

PROPOSITION A2. *When emissions and labor are gross substitutes ($1 < \eta < \sigma$), the cap-and-trade policy results in positive reallocation of emission permits. Relative to the cap regime, the cap-and-trade program*

[9] The validity of the statement is shown in the proof of Proposition A3. If $\lambda_t > \max \lambda_c$, $p_t(\phi) > p_c(\phi)$ for all ϕ, which implies $P_t > P_c$ and $\Phi_t > \Phi_c$. However, this will lead to a contradiction $P_t < P_c$, as shown in Appendix D.5. It then follows $\lambda_t < \max \lambda_c$. Note that, it is possible $\lambda_t < \min \lambda_c$. Namely, after implementing cap-and-trade, every firm bears a lower shadow price of emissions because of the fall of the market potential Φ.

1. *lowers the shadow price of emissions for more productive firms:* $\lambda_t(\phi) < \lambda_c(\phi)$ *if* $\phi > \tilde{\phi}$, *and raises the shadow price of emissions for less productive firms:* $\lambda_t(\phi) > \lambda_c(\phi)$ *if* $\phi < \tilde{\phi}$, *with the decline in shadow price being greater for more productive firms* $\frac{\partial[\lambda_t(\phi)-\lambda_c(\phi)]}{\partial\phi} < 0$;
2. *increases the emission intensity for more productive firms and lowers the emission intensity for less productive firms. The percentage increase in emission intensity is greater for more productive firms:*

$$\frac{\partial[\ln(\zeta_t(\phi)) - \ln(\zeta_c(\phi))]}{\partial\phi} > 0, \tag{A16}$$

for all firms with productivity above ϕ_l, *where* ϕ_l *is characterized by*

$$\phi_l = \max\left\{\phi \middle| \alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)\left(\frac{1-\alpha}{\alpha w}\right)^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} = 0\right\}.$$

Proof. See Appendix D.4 for the proof.

From emissions cap to a cap-and-trade policy, there is a heterogeneous impact on emission intensities as a result of emissions reallocation from less to more productive firms. For highly productive firms, the shadow prices of emissions are lower under the cap-and-trade regime, and thus, their emission intensities increase. In our empirical analysis, we verify the relative increase in emission intensities for more productive firms as evidence for the positive reallocation of emission permits.

Given the same overall emissions target under cap and cap-and-trade policies, the aggregate output Q is a sufficient statistic for the comparison of economy-wide production efficiency between the two regimes.[10] Moreover, since the measure of firms is fixed in this model, Q also represents average firm-level "output", which could be employed to test the efficiency implication of regulation changes empirically.[11]

PROPOSITION A3. *When emissions and labor are gross substitutes* ($1 < \eta < \sigma$), *a cap-and-trade policy results in higher firm-average and aggregate output compared to a pro rata emissions cap policy:* $Q_t > Q_c$.

Proof. See Appendix D.5 for the proof.

[10] In our model, the manufacturing sector employs more labor under cap-and-trade regime than under cap regime since labor supply is inelastic. However, this is not the key reason for output growth after the introduction of emissions trading. As the proof in Appendix D.5 implies, even if firm-level labor adjustment is not allowed (hence q_0 remains unchanged before and after cap-and-trade), cap-and-trade can still realize a higher economic output by fixing the distortion in the marginal product of emissions under the cap policy.

[11] As it is derived by dividing the aggregate Q with the firm mass, the average firm-level output in this model accommodates the substitution between products. In the empirical analysis, we use firm-level real revenue (temporal industrial-level price effect absorbed by industry-year fixed effects) to measure firm-level output, which is the conventional firm-level output indicator employed to obtain the economy-level output.

Proposition A3 is an intuitive result given the negative reallocation caused by emissions caps. The cap-and-trade policy restores the distorted shadow prices of emissions and leads to positive reallocation of emission permits from less to more productive firms. The positive reallocation shifts production to more productive firms and raises average productivity in the economy, thus resulting in a higher firm-average and economy-wide output. In our empirical section, we confirm Proposition A3 by comparing average firm-level outputs.

Before moving to discussions on the theoretical findings, we derive the model prediction under alternative η values in Appendix E. As a result, the reallocation effects predicted by Propositions A1 and A2 are completely muted when the inputs demonstrate unitary elasticity of substitution, and reversed when they are gross complements, leading to ambiguous impact of cap-and-trade on the overall production efficiency.

C.6. Further Discussions

Understanding the role of gross-substitutability condition requires one to understand the intrinsic market imperfection in our baseline model. Due to monopolistic competition with a non-Hicks-neutral production function, firms' marginal cost of emissions (λ) is not aligned with their marginal "product" of emissions (in the discussion of production efficiency, the marginal "product" refers to the contribution of firm-level emissions to the aggregate product Q):

$$\text{(A17)} \qquad \frac{\partial Q}{\partial z} = Q^{\frac{1}{\sigma}} \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi \alpha^{\sigma} \lambda^{-\sigma} \left[\alpha + (1-\alpha)\left(\frac{\phi\lambda}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}\right]^{\frac{\sigma}{\eta-1}}.$$

Given total emissions Z, equalizing the latter is key to achieving the maximum economy-wide production efficiency, and this entails that more productive firms face *lower* cost of emissions (since $\frac{\partial Q^2}{\partial z \partial \phi} > 0$ and $\frac{\partial Q^2}{\partial z \partial \lambda} > 0$). In this sense, the cap-and-trade policy, which achieves "cost effectiveness" by equalizing the marginal cost of emissions across firms, also fails to achieve maximum production efficiency.[12]

Moreover, since the cap-and-trade policy cannot promise maximum production efficiency, its comparison with a cap policy hinges on the allocation under the emissions caps. As we have shown in the model, under the gross-substitutability condition, a *pro rata* emissions cap policy results in a higher cost of emissions for more productive firms,

[12] The maximum efficiency is calculated when policy tools can only affect the shadow price of emissions or the allocation of emissions. We note that even if the marginal product of emissions is equalized across firms, the economy is not at its first-best production level since the marginal product of labor is not equalized due to monopolistic competition. Policy tools, such as a firm-specific output tax/subsidy, can render the economy to achieve the first-best outcome by correcting the product market distortion. See Fowlie et al. (2016) for relevant discussions.

a cost dispersion in the opposite direction to the scenario under maximum production efficiency (See Figure A3 for a graphic illustration of shadow prices of emissions under the maximum efficiency). This unintended distortion from emissions cap policy is key for cap-and-trade to make an improvement. On the other hand, if the gross-substitutability condition fails to hold, the emissions cap policy can generate the same emissions allocation as cap-and-trade ($\eta = 1$) or an allocation where more productive firms face a lower shadow price of emissions ($\eta < 1$). Under those scenarios, a cap-and-trade policy does not always guarantee a higher production efficiency than emissions caps.

Our empirical findings in Appendix F suggest that the gross-substitutability condition holds for Chinese manufacturing firms. Appendix Tables A14 and A15 confirm that the economy-wide η is significantly higher than 1; Appendix Table A16 presents η estimates by industry and demonstrates that, despite the heterogeneity in η values across industries, all the industries have η higher than 1 except two of them (but their η estimates are not statistically significantly different from 1). Productive factors and emissions enter the production either with a unitary elasticity of substitution or as gross substitutes for Chinese manufacturing firms. Therefore, our main theoretical propositions remain valid after aggregating over industries with different η values. As a result, the reallocation effects identified in Propositions A1 and A2 should be reflected by the changes in firm-level emission intensities.

Appendix D. Proofs

D.1. Proof of Lemma A1

In the benchmark equilibrium, a firm's optimization problem defined by (A12) is

$$\max_{z,l} pq - wl \text{ s.t. } z = z_b$$

where p and q are both functions of z and l.

The Lagrangian of this problem is

$$L_b = pq - wl - \lambda(z - z_b).$$

The firm's original profit maximization problem is

$$\max_{z,l} pq - wl - \tau z$$

with the following Lagrangian:

$$L = pq - wl - \tau z.$$

Since z_b is a constant, the two Lagrangians, L_b and L, have the same solution when $\lambda = \tau$.

D.2. Proof of Lemma A2

Under the emissions cap policy, an additional resource constraint, $z \leq z_c$, will be added to the firm's cost minimization problem. Specifically, to produce an output q, the firm searches for the optimal bundle l and z to minimize the production cost:

$$\min_{l,z} wl + \tau z \text{ s.t. } \left[\alpha z^{\frac{\eta-1}{\eta}} + (1-\alpha)(\phi l)^{\frac{\eta-1}{\eta}}\right]^{\frac{\eta}{\eta-1}} = q \text{ and } z \leq z_c.$$

According to the derivation of benchmark equilibrium, without the resource constraint, or when the resource constraint does not bind, the firm's cost minimization problem yields a constant marginal cost of production, μ_b, and subsequent emissions to output ratio (emission intensity), ζ_b. Therefore, if the optimal output level $q < \frac{z_c}{\zeta_b}$, the resource constraint does not bind and the optimal emissions to produce q take an interior solution

$$z = \zeta_b q.$$

Otherwise, when $q \geq \frac{z_c}{\zeta_b}$, the emissions constraint binds and the optimal emissions take a corner solution

$$z = z_c.$$

Therefore, to prove Lemma A2, it suffices to show that the firm will always produce $q \geq \frac{z_c}{\zeta_b}$. We prove this claim by contradiction. Suppose some firms produce $q < \frac{z_c}{\zeta_b}$. The optimization problem takes an interior solution for those firms. It is easy to show that the aggregate price index $P_c \geq P_b$ and market potential $\Phi_c \geq \Phi_b$.[13]

At the interior solution, a firm's marginal product revenue is equal to $(1-\frac{1}{\sigma})\Phi_c^{\frac{1}{\sigma}} q^{-\frac{1}{\sigma}}$, derived from the demand function (A2). Meanwhile, at the interior solution, its marginal cost of production $\mu_c = \mu_b = (1-\frac{1}{\sigma})\Phi_b^{\frac{1}{\sigma}} q_b^{-\frac{1}{\sigma}}$. Since $\Phi_c \geq \Phi_b$, it implies that $q \geq q_b = \frac{z_b}{\zeta_b}$. However, by the definition of emissions cap, $z_b > z_c$, which leads to a contradiction as we assume $q < \frac{z_c}{\zeta_b}$ to begin with. As a result, the interior solution can be ruled out, and all firms will emit at the $z = z_c$ level under the emissions cap policy.

D.3. Proof of Proposition A1

We first confirm the changes in shadow prices. According to equation (A10), $\frac{\partial z}{\partial \tau} < 0$ and $\frac{\partial z}{\partial \Phi} > 0$. Lemma A1 implies that the firm's shadow price under an emissions cap policy is equivalent to the emissions cost such that benchmark emissions equal z_c. Appendix D.2 shows that $\Phi_c \geq \Phi_b$. Since $z_c < z_b$, $\lambda_c > \tau$ must hold for all firms. Moreover, equation (A13) can be rearranged as follows:

$$\phi = \left[\frac{(\delta\tau^{-\eta}\frac{\Phi_b}{\Phi_c}\lambda_c^{\eta})^{\frac{\eta-1}{\sigma-\eta}}\alpha^{\eta}\tau^{1-\eta} - \alpha^{\eta}\lambda_c^{1-\eta}}{(1-\alpha)^{\eta} w^{1-\eta}\left[1-(\delta\tau^{-\eta}\frac{\Phi_b}{\Phi_c}\lambda_c^{\eta})^{\frac{\eta-1}{\sigma-\eta}}\right]} \right]^{\frac{1}{\eta-1}}. \tag{A18}$$

The equation implies $\frac{\partial \phi}{\partial \lambda_c} > 0$, and hence, $\frac{\partial \lambda_c}{\partial \phi} > 0$.

For the change in emission intensities, equation (A11) suggests $\frac{\partial \zeta}{\partial \tau} < 0$. Since $\lambda_c > \tau$ for all firms, it must hold $\zeta_c < \zeta_b$ for all firms.

The percentage reduction in firm-level emission intensity can be expressed as follows:

$$\ln(\zeta_c) - \ln(\zeta_b) \quad = \quad \ln(\frac{\zeta_c}{\zeta_b})$$

[13]For a firm with the interior solution, its marginal cost of production $\mu_c = \mu_b$, where μ_b is given by equation (A8). Note that, for a firm with the corner solution, the firm is constrained by the supply of z and therefore its willingness to pay for z is higher than the cost τ and its marginal cost of production in the equilibrium $\mu_c \geq \mu_b$. With the iso-elastic demand function (A2), a firm's optimal price is $p = \frac{\sigma}{\sigma-1}\mu$. Hence, firm-level prices $p_c \geq p_b$ for all firms.

$$= \ln \left(\frac{\left[\alpha + (1-\alpha)\left(\frac{\phi\lambda_c}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}\right]^{\frac{\eta}{1-\eta}}}{\left[\alpha + (1-\alpha)\left(\frac{\phi\tau}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}\right]^{\frac{\eta}{1-\eta}}} \right) \qquad \text{from equation (A11)}$$

$$= \ln \left((\frac{\tau}{\lambda_c})^{\eta} \left[\frac{\alpha^{\eta}\lambda_c^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}}{\alpha^{\eta}\tau^{1-\eta} + (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}} \right]^{\frac{\eta}{1-\eta}} \right)$$

$$= \ln \left((\frac{\tau}{\lambda_c})^{\eta} \left[\left(\frac{\delta\Phi_b\tau^{-\eta}}{\Phi_c\lambda_c^{-\eta}} \right)^{\frac{\eta-1}{\sigma-\eta}} \right]^{\frac{\eta}{1-\eta}} \right) \qquad \text{from equation (A13)}$$

$$= \ln[(\frac{\tau}{\lambda_c})^{\frac{\eta\sigma}{\sigma-\eta}}(\frac{\Phi_c}{\delta\Phi_b})^{\frac{\eta}{\sigma-\eta}}]. \tag{A19}$$

Given $\sigma > \eta$, equation (A19) shows that the percentage reduction in the firm-level emission intensity is larger for more productive firms (corresponding to a larger value of λ_c).

D.4. Proof of Proposition A2

When transitioning from the emissions cap regime to the emissions cap-and-trade regime, the percentage change in the emission intensity varies across firms with different productivity:

$$\ln(\zeta_t) - \ln(\zeta_c) = \ln \left[\frac{\alpha + (1-\alpha)\left(\frac{\phi\lambda_t}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}}{\alpha + (1-\alpha)\left(\frac{\phi\lambda_c}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}} \right]^{\frac{\eta}{1-\eta}}.$$

Its derivative with respect to ϕ is

$$\frac{\partial[\ln(\zeta_t) - \ln(\zeta_c)]}{\partial\phi} = \frac{\eta}{1-\eta} \frac{\alpha + (1-\alpha)\left(\frac{\phi\lambda_c}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}}{\alpha + (1-\alpha)\left(\frac{\phi\lambda_t}{w}\frac{1-\alpha}{\alpha}\right)^{\eta-1}} \frac{(1-\alpha)(\frac{1-\alpha}{w\alpha})^{\eta-1}(\eta-1)\phi^{\eta-2}}{\left[\alpha + (1-\alpha)(\frac{1-\alpha}{w\alpha})^{\eta-1}\phi^{\eta-1}\lambda_c^{\eta-1}\right]^2} \cdot$$
$$\left\{ \alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)(\frac{1-\alpha}{\alpha w})^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} \right\}.$$

Note that the product of the terms outside the braces is negative. The first term inside the braces is negative when $\lambda_t < \lambda_c$ (equivalently, $\phi > \tilde{\phi}$) and is positive when $\lambda_t > \lambda_c$ ($\phi < \tilde{\phi}$). The second term inside the braces is always negative since $\frac{\partial\lambda_c}{\partial\phi} > 0$.

As a result, for $\phi > \tilde{\phi}$, $\frac{\partial[\ln(\zeta_t)-\ln(\zeta_c)]}{\partial\phi} > 0$, namely, the rise in emission intensity becomes even larger for firms with higher productivity. For $\phi < \tilde{\phi}$, the sign of $\frac{\partial[\ln(\zeta_t)-\ln(\zeta_c)]}{\partial\phi}$ depends on the absolute values of the two terms in the braces. Note that, $\alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) = 0$ and $-\left[\alpha + (1-\alpha)(\frac{1-\alpha}{\alpha w})^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} < 0$ when $\phi = \tilde{\phi}$. Moreover, the functions above are continuous in ϕ. Therefore, for ϕ close to $\tilde{\phi}$, the absolute value of the second term is larger, and hence, $\frac{\partial[\ln(\zeta_t)-\ln(\zeta_c)]}{\partial\phi} > 0$.

Since the terms inside the braces are negative when $\phi \geq \tilde{\phi}$, we can identify the threshold ϕ_l for ϕ such that the negative sign holds for all $\phi \geq \phi_l$. Denote $\{\phi|\alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)(\frac{1-\alpha}{\alpha w})^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} = 0\}$ the set of solutions to

$$\text{(A20)} \qquad \alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)(\frac{1-\alpha}{\alpha w})^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} = 0.$$

Then, $\phi_l = \max\{\phi|\alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)(\frac{1-\alpha}{\alpha w})^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} = 0\}$.

The existence of ϕ_l can be demonstrated as follows. When ϕ approaches 0, equation (A13) shows that λ_c will approach $(\frac{\Phi_c}{\delta\Phi_b})^{\frac{1}{\sigma}}\tau$.[14] Then, according to Appendix D.3, $\frac{\partial\lambda_c}{\partial\phi}$ will also approach a finite number. In other words, as ϕ approaches 0, the terms inside the braces will approach $\alpha[\lambda_t^{\eta-1} - (\frac{\Phi_c\tau^\sigma}{\delta\Phi_b})^{\frac{\eta-1}{\sigma}}]$, which should be positive, and hence, $\frac{\partial[\ln(\zeta_t)-\ln(\zeta_c)]}{\partial\phi} < 0$. Since $\frac{\partial[\ln(\zeta_t)-\ln(\zeta_c)]}{\partial\phi} > 0$ when $\phi = \tilde{\phi}$, by the intermediate value theorem, there exists a value of ϕ_l that satisfies equation (A20).

D.5. Proof of Proposition A3

Note that in this model, the total expenditure E is an exogenous parameter. Since $E = PQ$, the comparison of the aggregate output under the two regimes hinges on the comparison of the aggregate price index, P_c and P_t.

Recalling the results from profit maximization, the firm-level price and emissions in equilibrium are functions of emissions cost τ, technology ϕ and market potential Φ:

$$\text{(A21)} \qquad p(\phi,\tau) = \frac{\sigma}{\sigma-1}\left[\alpha^\eta\tau^{1-\eta} + (1-\alpha)^\eta\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{1}{1-\eta}}$$

$$\text{(A22)} \qquad z(\phi,\tau,\Phi) = \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma}\Phi\alpha^\eta\tau^{-\eta}\left[\alpha^\eta\tau^{1-\eta} + (1-\alpha)^\eta\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}}.$$

Under the emissions cap policy, $p_c = p(\phi,\lambda_c)$ and $z_c = z(\phi,\lambda_c,\Phi_c)$ denote the equilibrium price and emissions, respectively. Under the cap-and-trade policy, $p_t = p(\phi,\lambda_t)$ and $z_t = z(\phi,\lambda_t,\Phi_t)$.

[14] For $\phi < \underline{\phi}$, equation (A13) can still be applied by fixing Φ_b and Φ_c at their equilibrium values.

In the cap-and-trade regime, total emissions are fixed at Z such that

$$\begin{aligned} Z &= \int_\phi z(\phi, \lambda_t, \Phi_t) g(\phi) d\phi \\ &= \int_\phi \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi_t \alpha^\eta \lambda_t^{-\eta} \left[\alpha^\eta \lambda_t^{1-\eta} + (1-\alpha)^\eta \left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma-\eta}{\eta-1}} g(\phi) d\phi. \end{aligned} \tag{A23}$$

We note that $\Phi_t = EP_t^{\sigma-1}$ and $P_t^{1-\sigma} = \int_\phi p(\phi, \lambda_t)^{1-\sigma} g(\phi) d\phi$. Therefore, according to (A21), we have

$$E = \Phi_t \int_\phi (\frac{\sigma}{\sigma-1})^{1-\sigma} \left[\alpha^\eta \lambda_t^{1-\eta} + (1-\alpha)^\eta \left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{1-\sigma}{1-\eta}} g(\phi) d\phi. \tag{A24}$$

The two equations (A23) and (A24) pin down the equilibrium levels of emissions price λ_t and market potential Φ_t in the cap-and-trade regime.

Since $\Phi = EP^{\sigma-1}$, proving $P_t < P_c$ is equivalent to proving $\Phi_t < \Phi_c$. We do this in two steps. First, we prove that there is a unique $\Phi_t \in (0, \infty)$ that satisfies equations (A23) and (A24). Second, we show that the equilibrium $\Phi_t \in (0, \Phi_c)$.

Step 1: Uniqueness of Φ_t

Denote $f(\lambda_t, \phi) = \alpha^\eta \lambda_t^{1-\eta} + (1-\alpha)^\eta \left(\frac{w}{\phi}\right)^{1-\eta}$. Combining equations (A23) and (A24), we have

$$\frac{\int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}} (1-\alpha)^\eta \left(\frac{w}{\phi}\right)^{1-\eta} g(\phi) d\phi}{\int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}} g(\phi) d\phi} = \frac{(\frac{\sigma-1}{\sigma})E - Z\lambda_t}{Z} \alpha^\eta \lambda_t^{-\eta}. \tag{A25}$$

The RHS of equation (A25) strictly decreases with λ_t. The derivative of the LHS with respect to λ_t is

$$\begin{aligned} &\frac{\sigma-\eta}{\eta-1} \frac{f_1(\lambda_t, \phi)}{[\int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}} g(\phi) d\phi]^2} \\ &\cdot \left[(\int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}} g(\phi) d\phi)^2 - \int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}+1} g(\phi) d\phi \int_\phi f(\lambda_t, \phi)^{\frac{\sigma-\eta}{\eta-1}-1} g(\phi) d\phi \right] \end{aligned}$$

Since $f_1(\lambda_t, \phi) = \frac{\partial f(\lambda_t, \phi)}{\partial \lambda_t} < 0$ and the term in the second line is negative, the LHS strictly increases with λ_t. As a result, there is a unique solution to λ_t for $\lambda_t > 0$. According to equation (A23), Φ_t increases with λ_t. Therefore, there is a unique Φ_t in the cap-and-trade equilibrium.

Step 2: $\Phi_t < \Phi_c$

From equation (A23), we write $\lambda_t = \lambda_t(\Phi_t)$. Denote

$$h(\Phi_t) = EP_t^{\sigma-1} = \frac{E}{\int_\phi p(\phi, \lambda_t(\Phi_t))^{1-\sigma} g(\phi) d\phi} \tag{A26}$$

The equilibrium Φ_t is the solution to $h(\Phi_t) = \Phi_t$. It is easy to verify that $h'(\Phi_t) > 0$ and $h'(\Phi_t) \to \infty$ when $\Phi_t \to 0$. It remains to show $h(\Phi_c) < \Phi_c$. If that inequality holds, the equilibrium $\Phi_t < \Phi_c$ since $h(\Phi_t)$ is continuous in Φ_t. Note that to show $h(\Phi_c) < \Phi_c$, it suffices to show $P_t = (\int_\phi p(\phi, \lambda_t(\Phi_t))^{1-\sigma} g(\phi) d\phi)^{\frac{1}{1-\sigma}} < P_c$ where $\Phi_t = \Phi_c$.

Given $\sigma > 1$, to demonstrate $P_t < P_c$, it suffices to show

$$\int p_t(\phi)^{1-\sigma} g(\phi) d\phi > \int p_c(\phi)^{1-\sigma} g(\phi) d\phi. \tag{A27}$$

Note that $p(\phi) > 0$ for all ϕ under both environmental regulating regimes. To facilitate our proof, we make use of the following claim:

CLAIM A1. *Given $p(\phi) > 0$ for all ϕ, $\int p_t(\phi)^{1-\sigma} g(\phi) d\phi > \int p_c(\phi)^{1-\sigma} g(\phi) d\phi$ if*

1. *The aggregation over p_t is smaller than that over p_c*

$$\int_{\underline{\phi}}^{\bar{\phi}} p_t(\phi) g(\phi) d\phi < \int_{\underline{\phi}}^{\bar{\phi}} p_c(\phi) g(\phi) d\phi;$$

2. *$p(\phi)$ decreases with ϕ;*
3. *There exists a $\tilde{\phi}$ such that for all $\phi < \tilde{\phi}$, $p_t(\phi) > p_c(\phi)$, and for all $\phi > \tilde{\phi}$, $p_t(\phi) < p_c(\phi)$.*

The validity of the claim can be shown as follows:

Proof. To simplify the notation, we denote $\rho = 1 - \sigma$ and write $p_t(\phi)$ as p_t and $p_c(\phi)$ as p_c.

For $\phi > \tilde{\phi}$, $p_t < p_c$. We make the following decomposition:

$$p_t^\rho - p_c^\rho = (p_c - p_t) \frac{p_t^\rho - p_c^\rho}{p_c - p_t}.$$

Define $f(p_t, p_c) = \frac{p_t^\rho - p_c^\rho}{p_c - p_t}$, $f_1 = \frac{\partial f(p_t, p_c)}{\partial p_t}$ and $f_2 = \frac{\partial f(p_t, p_c)}{\partial p_c}$. We have

$$f_1 = \frac{\rho p_t^{\rho-1} p_c - (\rho - 1) p_t^\rho - p_c^\rho}{(p_c - p_t)^2}.$$

Denote the numerator as a function of p_t, $\nu(p_t) = \rho p_t^{\rho-1} p_c - (\rho-1)p_t^\rho - p_c^\rho$. It is easy to see that $\nu(p_t) = 0$ when $p_t = p_c$. Moreover,

$$\nu'(p_t) = \rho(\rho-1)p_t^{\rho-2}(p_c - p_t) > 0.$$

Therefore, for all $p_t < p_c$, $\nu(p_t) < 0$, and consequently, $f_1 < 0$. By the same token, $f_2 < 0$ for all $p_t < p_c$. Since $p(\phi)$ decreases with ϕ, for $\phi > \tilde{\phi}$, we have $p_t(\phi) < p_t(\tilde{\phi})$ and $p_c(\phi) < p_c(\tilde{\phi})$. We also note that $p_t(\phi) < p_c(\phi)$ for $\phi > \tilde{\phi}$. As a result,

$$\frac{p_t(\phi)^\rho - p_c(\phi)^\rho}{p_c(\phi) - p_t(\phi)} > \frac{p_t(\phi)^\rho - p_c(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^+ - p_t(\phi)} > \frac{p_t(\tilde{\phi})^{+\rho} - p_c(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^+ - p_t(\tilde{\phi})^+}$$

where $p(\tilde{\phi})^+ = \lim_{\phi\to\tilde{\phi}^+} p(\phi)$.

For $\phi < \tilde{\phi}$, $p_t > p_c$. We make the following decomposition:

$$p_c^\rho - p_t^\rho = (p_t - p_c)\frac{p_t^\rho - p_c^\rho}{p_c - p_t}.$$

Following the same procedure, we can show that for $\phi < \tilde{\phi}$,

$$\frac{p_t(\phi)^\rho - p_c(\phi)^\rho}{p_c(\phi) - p_t(\phi)} < \frac{p_t(\tilde{\phi})^{-\rho} - p_c(\tilde{\phi})^{-\rho}}{p_c(\tilde{\phi})^- - p_t(\tilde{\phi})^-}$$

where $p(\tilde{\phi})^- = \lim_{\phi\to\tilde{\phi}^-} p(\phi)$.

We also note that

$$\begin{aligned}
& \frac{p_t(\tilde{\phi})^{+\rho} - p_c(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^+ - p_t(\tilde{\phi})^+} \\
= \quad & -\frac{[p_t(\tilde{\phi})^+ + p_c(\tilde{\phi})^+ - p_t(\tilde{\phi})^+]^\rho - p_t(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^+ - p_t(\tilde{\phi})^+} \\
= \quad & \lim_{\Delta\to 0} -\frac{[p_t(\tilde{\phi})^+ + \Delta]^\rho - p_t(\tilde{\phi})^{+\rho}}{\Delta} \\
= \quad & -\rho p_t(\tilde{\phi})^{+\rho-1}
\end{aligned}$$

where $\Delta = p_c(\tilde{\phi})^+ - p_t(\tilde{\phi})^+$ approaches 0 since $p_t(\tilde{\phi}) = p_c(\tilde{\phi})$.

Similarly,

$$\frac{p_t(\tilde{\phi})^{-\rho} - p_c(\tilde{\phi})^{-\rho}}{p_c(\tilde{\phi})^- - p_t(\tilde{\phi})^-} = -\rho p_t(\tilde{\phi})^{-\rho-1}.$$

Since $p(\phi)$ is continuous in ϕ, we have

$$\frac{p_t(\tilde{\phi})^{+\rho} - p_c(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^{+} - p_t(\tilde{\phi})^{+}} = \frac{p_t(\tilde{\phi})^{-\rho} - p_c(\tilde{\phi})^{-\rho}}{p_c(\tilde{\phi})^{-} - p_t(\tilde{\phi})^{-}} = -\rho p_t(\tilde{\phi})^{\rho-1}.$$

Finally,

$$\begin{aligned}
& \int p_t(\phi)^\rho g(\phi)d\phi - \int p_c(\phi)^\rho g(\phi)d\phi \\
= & \int_{\tilde{\phi}}^{\bar{\phi}} [p_t(\phi)^\rho - p_c(\phi)^\rho] g(\phi)d\phi - \int_{\underline{\phi}}^{\tilde{\phi}} [p_c(\phi)^\rho - p_t(\phi)^\rho] g(\phi)d\phi \\
= & \int_{\tilde{\phi}}^{\bar{\phi}} [p_c(\phi) - p_t(\phi)] \frac{p_t(\phi)^\rho - p_c(\phi)^\rho}{p_c(\phi) - p_t(\phi)} g(\phi)d\phi - \int_{\underline{\phi}}^{\tilde{\phi}} [p_t(\phi) - p_c(\phi)] \frac{p_t(\phi)^\rho - p_c(\phi)^\rho}{p_c(\phi) - p_t(\phi)} g(\phi)d\phi \\
> & \int_{\tilde{\phi}}^{\bar{\phi}} [p_c(\phi) - p_t(\phi)] \frac{p_t(\tilde{\phi})^{+\rho} - p_c(\tilde{\phi})^{+\rho}}{p_c(\tilde{\phi})^{+} - p_t(\tilde{\phi})^{+}} g(\phi)d\phi - \int_{\underline{\phi}}^{\tilde{\phi}} [p_t(\phi) - p_c(\phi)] \frac{p_t(\tilde{\phi})^{-\rho} - p_c(\tilde{\phi})^{-\rho}}{p_c(\tilde{\phi})^{-} - p_t(\tilde{\phi})^{-}} g(\phi)d\phi \\
= & \int [p_c(\phi) - p_t(\phi)][-\rho p_t(\tilde{\phi})^{\rho-1}] g(\phi)d\phi \\
= & -\rho p_t(\tilde{\phi})^{\rho-1} \int [p_c(\phi) - p_t(\phi)] g(\phi)d\phi \\
> & \; 0.
\end{aligned}$$

□

Our next step is to validate the three conditions in the claim. For p_t, the validity of condition 2 directly follows equation (A21) since λ_t does not vary across firms. For p_c, since z_b (under the benchmark scenario) increases with ϕ, z_c should also increase with ϕ, as does q_c. From the demand function, $q_c = \Phi_c p_c^{-\sigma}$, it is easy to show that p_c decreases with ϕ.

Equation (A21) also implies that conditional on the firm's productivity, ϕ, its output price p increases with emissions cost τ. According to Proposition A1, the shadow price under the emissions cap policy increases with firm-level productivity. As the shadow price under cap-and-trade policy is the same across all firms, there must exist a cutoff productivity such that condition 3 holds. Note that, since Φ is different under the two regimes, the cutoff $\tilde{\phi}$ may take a value lower than $\underline{\phi}$ or higher than $\bar{\phi}$, but this will not affect the validity of our proof.

To verify condition 1, we first combine equations (A21) and (A22) to obtain z as a function

of ϕ, p, and Φ:
(A28)

$$z(\phi, p, \Phi) = \left(\frac{\sigma}{\sigma-1}\right)^{-\sigma} \Phi\alpha^{\eta} \left[\frac{\left(p \cdot \frac{\sigma-1}{\sigma}\right)^{1-\eta} - (1-\alpha)^{\eta}\left(\frac{w}{\phi}\right)^{1-\eta}}{\alpha^{\eta}}\right]^{\frac{\eta}{\eta-1}} \left(p \cdot \frac{\sigma-1}{\sigma}\right)^{\eta-\sigma}.$$

Consequently, p is an implicit function of z, $p = p(\phi, z, \Phi)$. Next, we further expand $p(\phi, z, \Phi)$ as follows:

$$p(\phi, z, \Phi) = \int_0^z \frac{\partial p(\phi, \nu, \Phi)}{\partial \nu} d\nu + \lim_{\nu \to 0} p(\phi, \nu, \Phi).$$

Note that since $\frac{\eta}{\eta-1} > 1$ and $\eta < \sigma$, $z(\phi, p, \Phi) = 0$ when $p = \frac{\sigma}{\sigma-1}(1-\alpha)^{\frac{\eta}{1-\eta}} \frac{w}{\phi}$, which does not depend on the value of Φ. Therefore, condition 1 is equivalent to

$$\int_{\underline{\phi}}^{\bar{\phi}} \int_0^{z_t(\phi)} \frac{\partial p(\phi, \nu, \Phi_t)}{\partial \nu} d\tau g(\phi) d\phi - \int_{\underline{\phi}}^{\bar{\phi}} \int_0^{z_c(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\tau g(\phi) d\phi < 0$$

where $z_t(\phi) = z(\phi, \lambda_t, \Phi_t)$ and $z_c(\phi) = z(\phi, \lambda_c(\phi), \Phi_c)$.

Denote $\hat{\phi}$ as the productivity cutoff such that for a firm with $\hat{\phi}$, its emissions are the same under both environmental regulating regimes. $\hat{\phi}$ is implicitly defined by the equation $z(\hat{\phi}, \lambda_t, \Phi_t) = z(\hat{\phi}, \lambda_c(\hat{\phi}), \Phi_c)$, where the z function is given by equation (A22). Note that if the market potential Φ differs under the two regimes, $\hat{\phi}$ is different from $\tilde{\phi}$.

When $\Phi_t = \Phi_c$, condition 1 can be rearranged as follows:

$$\begin{aligned}
& \int_{\underline{\phi}}^{\bar{\phi}} \int_0^{z_t(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi - \int_{\underline{\phi}}^{\bar{\phi}} \int_0^{z_c(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi \\
= & \int_{\underline{\phi}}^{\hat{\phi}} \int_0^{z_t(\phi)} \left[\frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} - \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu}\right] d\nu g(\phi) d\phi \\
& + \int_{\hat{\phi}}^{\bar{\phi}} \int_0^{z_c(\phi)} \left[\frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} - \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu}\right] d\nu g(\phi) d\phi \\
& + \int_{\hat{\phi}}^{\bar{\phi}} \int_{z_c(\phi)}^{z_t(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi - \int_{\underline{\phi}}^{\hat{\phi}} \int_{z_t(\phi)}^{z_c(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi \\
= & \int_{\hat{\phi}}^{\bar{\phi}} \int_{z_c(\phi)}^{z_t(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi - \int_{\underline{\phi}}^{\hat{\phi}} \int_{z_t(\phi)}^{z_c(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi) d\phi \\
< & \ 0.
\end{aligned}$$

From equation (A28), we have

$$\frac{\partial^2 p(\phi, z, \Phi)}{\partial z^2} = -\frac{\frac{\partial^2 z(\phi,p,\Phi)}{\partial p^2}}{(\frac{\partial z(\phi,p,\Phi)}{\partial p})^3} > 0. \tag{A29}$$

We also have

$$\begin{aligned}\frac{\partial p(\phi, z, \Phi)}{\partial z} &= \frac{\partial p(\phi, \tau)}{\partial \tau}\frac{\partial \tau(\phi, z, \Phi)}{\partial z} \\ &= -\frac{\left(\frac{\sigma}{\sigma-1}\right)^{1+\sigma}\Phi^{-1}}{\frac{\eta}{t}\left[\alpha^\eta\tau^{1-\eta}+(1-\alpha)^\eta\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma}{\eta-1}}+\frac{(\sigma-\eta)\alpha^\eta}{t^\eta}\left[\alpha^\eta\tau^{1-\eta}+(1-\alpha)^\eta\left(\frac{w}{\phi}\right)^{1-\eta}\right]^{\frac{\sigma}{\eta-1}-1}}\end{aligned} \tag{A30}$$

where $\tau = \tau(\phi, z, \Phi)$. Therefore,

$$\frac{\partial \frac{\partial p(\phi,z,\Phi)}{\partial z}|_{z=z_t(\phi)}}{\partial \phi} = \frac{\partial \frac{\partial p(\phi,z,\Phi)}{\partial z}|_{\tau=\lambda_t}}{\partial \phi} < 0 \tag{A31}$$

Finally, note that $\int_{\underline{\phi}}^{\hat{\phi}}\int_{z_t(\phi)}^{z_c(\phi)} dzg(\phi)d\phi = \int_{\hat{\phi}}^{\bar{\phi}}\int_{z_c(\phi)}^{z_t(\phi)} dzg(\phi)d\phi$ since $\hat{\phi}$ is the productivity cutoff for a firm with equal emissions under the two regimes and the aggregate level of emissions is the same under both regimes.

As a result, the inequality in condition 1 can be verified as

$$\begin{aligned}
& \int_{\hat{\phi}}^{\bar{\phi}}\int_{z_c(\phi)}^{z_t(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi)d\phi - \int_{\underline{\phi}}^{\hat{\phi}}\int_{z_t(\phi)}^{z_c(\phi)} \frac{\partial p(\phi, \nu, \Phi_c)}{\partial \nu} d\nu g(\phi)d\phi \\
< & \int_{\hat{\phi}}^{\bar{\phi}} [z_t(\phi) - z_c(\phi)]\frac{\partial p(\phi, z, \Phi_t)}{\partial z}|_{z=z_t(\phi)} g(\phi)d\phi - \int_{\underline{\phi}}^{\hat{\phi}} [z_c(\phi) - z_t(\phi)]\frac{\partial p(\phi, z, \Phi_t)}{\partial z}|_{z=z_t(\phi)} g(\phi)d\phi \\
& (\text{since } \frac{\partial^2 p(\phi, z, \Phi)}{\partial z^2} > 0) \\
< & \int_{\hat{\phi}}^{\bar{\phi}} [z_t(\phi) - z_c(\phi)]\frac{\partial p(\phi, z, \Phi_t)}{\partial z}|_{z=z_t(\phi)} g(\phi)d\phi - \int_{\hat{\phi}}^{\bar{\phi}} [z_t(\phi) - z_c(\phi)]\frac{\partial p(\phi, z, \Phi_t)}{\partial z}|_{z=z_t(\phi)} g(\phi)d\phi \\
& (\text{since } \frac{\partial \frac{\partial p(\phi,z,\Phi)}{\partial z}|_{z=z_t(\phi)}}{\partial \phi} < 0 \text{ and } \int_{\underline{\phi}}^{\hat{\phi}}\int_{z_t(\phi)}^{z_c(\phi)} dzg(\phi)d\phi = \int_{\hat{\phi}}^{\bar{\phi}}\int_{z_c(\phi)}^{z_t(\phi)} dzg(\phi)d\phi) \\
= & \ 0.
\end{aligned}$$

Finally, since the unique equilibrium $\Phi_t < \Phi_c$ and $\Phi = EP^{\sigma-1}$, we reach the conclusion that the aggregate price index in the cap-and-trade regime is smaller than that in the cap regime: $P_t < P_c$. As a result, $Q_t > Q_c$.

Appendix E. Model Results under Alternative Values of η

E.1. Unitary Elasticity of Substitution between Inputs

The production function takes a Cobb-Douglas form: $q = z^{\alpha}(\phi l)^{1-\alpha}$, when $\eta = 1$. In that case, equation (A13) reduces to $\Phi_c \lambda_c^{\alpha(\sigma-1)-1} = \delta \Phi_b \tau^{\alpha(\sigma-1)-1}$, with productivity ϕ on both sides canceled out, resulting in a constant shadow price of emissions cap, λ_c, for all firms. Since total emissions are the same under emissions cap and cap-and-trade regimes, the shadow price of emissions will remain the same after the regulation change ($\lambda_t = \lambda_c$), and thus firm-level emissions remain unchanged for all firms. As a result, there is no incentive for firms with different productivity levels to trade emission permits.[15]

COROLLARY A1. *When the elasticity of substitution between emissions and labor is equal to one ($\eta = 1$), pro rata emissions cap and cap-and-trade equilibria are identical.*

E.2. Inputs as Gross Complements

We have shown that when emissions and labor are gross substitutes, emissions cap-and-trade policy restores the distortion caused by the *pro rata* emissions cap and leads to higher production efficiency in the economy. Under the alternative assumption that the inputs are gross complements, the model generates the opposite reallocation effects and an ambiguous implication on economic efficiency. We present formal theoretical results in the following corollaries.

COROLLARY A2. *When emissions and labor are gross complements ($\eta < 1$), relative to the benchmark equilibrium, the pro rata emissions cap*

1. *increases the shadow price of emissions for all firms: $\lambda_c(\phi) > \lambda_b(\phi)\ \forall\ \phi$, and this increase is higher for less productive firms: $\frac{\partial[\lambda_c(\phi)-\lambda_b(\phi)]}{\partial\phi} < 0$;*
2. *lowers firm-level emission intensity for all firms: $\zeta_c(\phi) < \zeta_b(\phi)\ \forall\ \phi$, and the percentage reduction is greater for less productive firms:*

$$\frac{\partial[\ln(\zeta_b(\phi)) - \ln(\zeta_c(\phi))]}{\partial\phi} < 0. \tag{A32}$$

Corollary A2 can be proved by showing $\frac{\partial\lambda_c}{\partial\phi} < 0$ when $\eta < 1$ using equation (A18). It identifies an opposite reallocation effect to Proposition A1 when the economy transitions from benchmark to the emissions cap regime. When the inputs are complements, the uniformly proportional emissions cap policy is disproportionately more restrictive for less

[15] If α varies across firms as in Copeland and Taylor (1995), there is emissions reallocation from cap to cap-and-trade regime associated with firms' "dirtiness", as is widely studied in the literature. However, the reallocation effect has an ambiguous prediction on production efficiency as we don't know how α is correlated with the distribution of productivity ϕ.

productive firms. As a result, more productive firms face relatively lower shadow prices of emissions and feature relatively higher emission intensities after the implementation of the cap policy.

As the cap effect goes in the opposite direction, the reallocation from cap to cap-and-trade also exhibits the opposite effects to those identified in Proposition A2. The following corollary is an analogue to that proposition. Its proof can be obtained by re-examining the conditions in Appendix D.4.

COROLLARY A3. *When emissions and labor are gross complements ($\eta < 1$), relative to the cap regime, the cap-and-trade policy*

1. *lowers the shadow price of emissions for less productive firms: $\lambda_t(\phi) < \lambda_c(\phi)$ if $\phi < \tilde{\phi}$, and raises the shadow price of emissions for more productive firms: $\lambda_t(\phi) > \lambda_c(\phi)$ if $\phi > \tilde{\phi}$, with $\frac{\partial[\lambda_c(\phi)-\lambda_t(\phi)]}{\partial\phi} < 0$;*
2. *increases the emission intensity for less productive firms and lowers the emission intensity for more productive firms. The percentage rise in emission intensity is greater for less productive firms:*

$$\frac{\partial[\ln(\zeta_t(\phi)) - \ln(\zeta_c(\phi))]}{\partial\phi} < 0, \tag{A33}$$

for all firms with productivity below ϕ_h, where ϕ_h is characterized by:

$$\phi_h = \min\left\{\phi \middle| \alpha(\lambda_t^{\eta-1} - \lambda_c^{\eta-1}) - \left[\alpha + (1-\alpha)\left(\frac{1-\alpha}{\alpha w}\right)^{\eta-1}\phi^{\eta-1}\lambda_t^{\eta-1}\right]\phi\lambda_c^{\eta-2}\frac{\partial\lambda_c}{\partial\phi} = 0\right\}$$

Therefore, the cap-and-trade policy leads to reallocation of emission permits from more to less productive firms. This reallocation effect has an ambiguous implication on the change of economy-wide production efficiency. On the one hand, it achieves cost-effectiveness by equalizing the marginal cost of emissions across all firms; on the other hand, it allows less productive firms to expand at the cost of more productive firms.[16] A comparison of Proposition 2 and Corollary 3 clearly shows that the effect of emissions trading hinges critically on the substitution pattern between labor and emissions. This is a key insight from our paper that has been neglected in the literature.

[16] One can easily show that both $Q_t < Q_c$ and $Q_t > Q_c$ are possible when $\eta < 1$ using numerical simulations.

Appendix F. Estimation of Elasticity of Substitution between Inputs

Our theoretical prediction hinges on the value of η, the elasticity of substitution between emissions z and labor l. In this section, we describe our methodology to estimate η from the data.

Solving the firm's cost-minimization problem leads to the following equation:

$$\frac{\tau}{w} = \frac{l}{z}\left(\frac{\alpha}{1-\alpha}\right)^{\eta}\left(\frac{\tau}{w/\phi}\right)^{1-\eta}. \tag{A34}$$

Without reliable tax data on emissions, it is not possible to estimate equation (A34) directly. Instead, we use equation (A34) to substitute for τ/w in equation (A11) as follows:

$$\zeta = \left[\alpha + (1-\alpha)\left(\frac{z}{l}\frac{1}{\phi}\right)^{\frac{1-\eta}{\eta}}\right]^{\frac{\eta}{1-\eta}} = \left[\alpha + (1-\alpha)e^{\frac{1-\eta}{\eta}[\ln(\frac{z}{l})-\ln(\phi)]}\right]^{\frac{\eta}{1-\eta}}. \tag{A35}$$

To enable the estimation of equation (A35), we first take the natural logarithm on both sides and then add an intercept term and an error term to obtain the regression model below:

$$\ln(\zeta_i) = c + \frac{\eta}{1-\eta}\ln\left[\alpha + (1-\alpha)e^{\frac{1-\eta}{\eta}[\ln(\frac{z_i}{l_i})-D_i]}\right] + \epsilon_i, \tag{A36}$$

where the subscript i denotes a firm, ϵ_i is the error term, c is the intercept term, and D_i includes a set of fixed effects that capture the impact of $\ln(\phi)$. We estimate equation (A36) using nonlinear least square. Column (1) of Table A14 reports the estimate of η when we only include province and industry fixed effects in equation (A36). Column (2) of Table A14 adds a set of ownership fixed effects, and Column (3) further controls for the firm age fixed effects. As shown in Table A14, the estimated values of η range from 1.14 to 1.17. All of them are statistically significant above 1.

We conduct two robustness checks to ensure the results reported in Table A14 are robust with additional controls. First, the TFP term, estimated using the method of Ackerberg et al. (2015), is added to equation (A36) to further control for the firm-specific productivity:

$$\ln(\zeta_i) = c + \frac{\eta}{1-\eta}\ln\left[\alpha + (1-\alpha)e^{\frac{1-\eta}{\eta}[\ln(\frac{z_i}{l_i})-\ln(TFP_i)-D_i]}\right] + \epsilon_i. \tag{A37}$$

Second, to better connect theory to empirical data, we expand the productive factors, l,

Table A14. Estimation of η

	Dependent Variable: ln(Emission Intensity)		
	(1)	(2)	(3)
η	1.170***	1.164***	1.137***
	(0.0512)	(0.0467)	(0.0431)
Province FE	✓	✓	✓
Industry FE	✓	✓	✓
Ownership FE		✓	✓
Firm Age FE			✓
Obs.	340535	340535	210794
Firms	104688	104688	64823

Note:
(i) This table reports the estimation results of equation (A36). The sample includes all COD-emitting firms.
(ii) Standard errors clustered at the provincial level are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

into a CES composite term of real capital (K), labor (L) and intermediate inputs (M):
(A38)

$$\ln(\zeta_i) = c + \frac{\eta}{1-\eta}\ln\left[\alpha + (1-\alpha)e^{\frac{1-\eta}{\eta}\left[\ln\left(\frac{z_i}{\left(\beta_k K_i^{\gamma} + \beta_l L_i^{\gamma} + (1-\beta_k-\beta_l)M_i^{\gamma}\right)^{\frac{1}{\gamma}}}\right) - \ln(TFP_i) - D_i\right]}\right] + \epsilon_i.$$

The estimation results of equations (A37) and (A38) are reported in Columns (1)-(3) and (4)-(6) of Table A15, respectively. All the η estimates reported in Table A15 are statistically significant above 1 and comparable to the ones reported in Table A14.

Table A15. Estimation of η – Robustness Check

	Dependent Variable: ln(Emission Intensity)					
	(1)	(2)	(3)	(4)	(5)	(6)
η	1.161***	1.153***	1.129***	1.118***	1.123***	1.144***
	(0.0553)	(0.0495)	(0.0463)	(0.0266)	(0.0254)	(0.0345)
Province FE	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓
Ownership FE		✓	✓		✓	✓
Firm Age FE			✓			✓
Obs.	340535	340535	210794	340535	340535	210794
Firms	104688	104688	64823	104688	104688	64823

Note:
(i) Columns (1)-(3) report the estimation results of equation (A37), and Columns (4)-(6) report the estimation of equation (A38). The sample includes all COD-emitting firms.
(ii) Standard errors clustered at the provincial level are shown in parentheses. *** $p < 0.01$, ** $p < 0.05$, and * $p < 0.1$.

Lastly, we estimate η values at the industry level (CIC 2-digit) with the specification in Column (3) of Appendix Table A14. The estimation results are reported in Appendix Table A16. The estimates are obtained for each manufacturing industry except for two with less than a thousand observations. Besides, we group industries from the following three sectors: mining, food and beverage, and energy and water, to avoid the small sample issue. Overall, the industrial-level η estimates are consistent with our assumption that productive factors and emissions are gross substitutes in the production function.

Table A16. Estimation of η – By Industry

Industry	η	p-value($\eta > 1$)	Obs.
Mining	1.56	.03	3278
Food and Beverage	1.03	.771	29058
Manufacture of Textile	1.89	.004	24829
Manufacture of Textile Wearing Apparel, Footware, and Caps	1.33	.129	3452
Manufacture of Leather, Fur, Feather and Related Products	1.08	.164	4080
Processing of Timber, Manufacture of Wood, Bamboo, etc.	1.08	.793	3220
Manufacture of Paper and Paper Products	1.01	.974	14300
Printing,Reproduction of Recording Media	1.03	.926	1590
Manufacture of Articles For Culture, Education and Sport Activity	3.31	.248	1045
Processing of Petroleum, Coking, Processing of Nuclear Fuel	1.11	.535	1460
Manufacture of Raw Chemical Materials and Chemical Products	1.52	.071	28123
Manufacture of Medicines	.97	.751	9015
Manufacture of Chemical Fibers	2.01	.361	1574
Manufacture of Rubber	1.39	.195	2548
Manufacture of Plastics	.95	.749	2960
Manufacture of Non-metallic Mineral Products	2.1	.486	16686
Smelting and Pressing of Ferrous Metals	1.17	.345	5444
Smelting and Pressing of Non-ferrous Metals	2.74	.302	4293
Manufacture of Metal Products	1.74	.201	9202
Manufacture of General Purpose Machinery	1.97	.26	9140
Manufacture of Special Purpose Machinery	1.04	.835	3853
Manufacture of Transport Equipment	1.41	.003	8756
Manufacture of Electrical Machinery and Equipment	1.34	.096	6261
Manufacture of Communication Equipment, Computers and Other	3.43	.099	6916
Manufacture of Measuring, Cultural and Office Instruments	1.64	.001	1457
Manufacture of Artwork and Other Manufacturing	1.46	.256	3517
Energy and Water	1.01	.929	3424

Note: This table reports η estimates by industry (CIC 2-digit) using the specification of Column (3) in Appendix Table A14. Manufacturing industries with less than a thousand observations are excluded.