Regulation, Emissions and Productivity: Evidence from China's Eleventh Five-Year Plan*

Brantly Callaway[†] Tong Li[‡] Joel Rodrigue[§] Yuya Sasaki[¶] Yong Tan[‖]

Date: August 17, 2024

Abstract

Leveraging sharp changes in environmental regulation embedded in China's eleventh Five-Year Plan (11^{th} FYP) we characterize regulation driven emissions-output tradeoffs in China. We document that China's 11^{th} FYP is associated with modest changes in average or total SO_2 emissions among manufacturing producers, but a sharp decline in the variance of the emissions intensity distribution. Extending well-known distributional estimators to characterize dynamic firm-level responses to policy change, we find large, causal emissions-intensity declines in the upper quantiles of the emissions distribution, modest evidence of increases in the lower quantiles of the emissions distribution, and no change in the middle quantiles. Policy-driven changes in firm-level emissions intensity are consistent with the differential abatement and output growth across heterogeneous producers after policy change. Interpreted through the lens of a resource misallocation framework, China's 11^{th} improved FYP increased aggregate productivity by 1.8 and 10.2 percent, respectively, through improved resource allocation. Our model suggests that uniform emissions taxes could have further increased aggregate productivity and output by 3.5 and 4.7 percent, without any increase in aggregate emissions.

Keywords: Regulation, Emissions, Productivity, China
JEL Codes: C21, D24, Q53.

*All errors are the responsibility of the authors.

[†]Department of Economics, University of Georgia. E-mail: brantly.callaway@uga.edu.
[‡]Department of Economics, Vanderbilt University, E-mail: tong.li@vanderbilt.edu.
[§]Department of Economics, Vanderbilt University, Email: joel.b.rodrigue@vanderbilt.edu.
[¶]Department of Economics, Vanderbilt University, Email: yuya.sasaki@vanderbilt.edu.
[‖]School of International Economics & Trade, NUFE, Email: yongtan_econ@163.com.

1 Introduction

> *'With a billion citizens still living on less than $4,000 per year, China's energy and environmental policies must be reconciled with the goals of maintaining economic growth and raising living standards.'* (Cao, Ho, and Jorgenson, 2013)

The aggregate production of global emissions is overwhelmingly concentrated in low- and middle-income countries. The total production of sulfur dioxide (SO_2), for example, is five times greater among developing countries than in their developed counterparts (Lin et al. (2022)). This feature of global production is hardly surprising: increasing evidence suggests that many low- and middle-income countries have experienced disproportionate growth in dirty industries.[1] It also highlights the stark trade-offs facing the developing world: while rising income, wages, and employment are strongly associated with health and well-being, the corresponding increase in pollution has a substantive countervailing impact (Ebenstein et al. (2015)). Accordingly, it is hardly surprising that substantive regulatory progress is often slow among many developing countries (UNEP (2019)).

Yet, environmental policy in developing country contexts is often fraught by uneven regulation, enforcement and outcomes. Indeed, existing evidence suggests that enforced regulation within industries and locations often varies widely across Chinese firms (Jia, 2012; Wu et al., 2013; Deng et al., 2020). If tighter environmental policy also induces resource allocation efficiency gains in developing countries it is not a foregone conclusion that pollution must rise in tandem with output growth.

Leveraging sharp changes in environmental regulation, we investigate how China's eleventh Five-Year Plan (11^{th} FYP) affected resource allocation across Chinese producers and the according implications for emissions-output tradeoffs. Specifically, we ask: how did China's 11^{th} FYP affect the nature of emissions intensity among Chinese producers? Did the consequent reallocation of resources across producers mitigate or exacerbate aggregate output losses from tighter environmental regulation? How much more output could have been produced under a uniform emission tax without increasing aggregate emissions?

We show that China's 11^{th} FYP is associated with modest changes in average or total SO_2 emissions among manufacturing producers. It does, however, coincide with a sharp decline in the variance of the emissions intensity distribution. Interpreted through the lens of a resource allocation framework, the evolution of the distribution of emissions intensity across heterogeneous producers contributed an additional 6 (20) percentage points of aggregate productivity (output) growth over the 2006-2010 period.

[1] For example studying China, see Bombardini and Li (2020) or Rodrigue et al. (2022a). Further evidence of pollution haven effects are documented in Ederington and Minier (2003), Levinson and Taylor (2008), Broner et al. (2012), Cherniwchan and Najjar (2022) and Tanaka et al. (2022) among others.

While a narrowing of the emissions distribution is consistent with a reduction in resource misallocation (Hsieh and Klenow (2009)), it is not clear that this secular change was induced by the policy itself. To flexibly characterize the causal impact of China's new regulation across the distribution of emissions intensity, we extend recent advances in the estimation of quantile treatment effects to our setting. We document that China's 11$^{\text{th}}$ FYP caused sharp *declines* in the upper tail of firm-level emissions intensity distribution. Among the upper tail of the emissions-intensity distribution, the policy induced a 38–50 percent decline in emissions intensity. In contrast, we find no evidence of any policy impact in the middle quantiles and large positive impacts among the lowest, least emissions-intensive quantiles, consistent with rapid increases in emissions intensity among firms that were initially cleanest. Indeed, we provide evidence that the comparative advantage of clean producers improved during the 11$^{\text{th}}$ FYP, enabling their disproportionate growth. Overall, China's 11$^{\text{th}}$ FYP had a substantial impact on the environmental performance of its manufacturing sector; focusing only on average treatment effects, which are generally small and insignificantly different from zero, would lead to the exact opposite conclusion.

Our empirical findings contribute to the rich literature that aims to characterize the trade-offs between environmental regulation and economic growth. Indeed, longstanding fears that tighter regulation would lead to productivity declines and job losses were central early US and EU environmental policy debates and continue to feature prominently in academic policy evaluation (Berman and Bui, 2001; Greenstone, 2002; van Soest et al., 2006; Greenstone et al., 2012; Bajari et al., 2012; Walker, 2013; Shapiro and Walker, 2018, 2024; Allcott and Greenstone, 2024). Large emerging markets face similar concerns as they seek to mitigate environmental costs while maintaining strong economic growth (Cao et al., 2013), broadly motivating research into Chinese environmental reform and the 11$^{\text{th}}$ FYP in particular. Shi and Xu (2018) and Wang and Zhang (2019) quantify the impact of China's 11$^{\text{th}}$ FYP on export sales, but do not study environmental and economic trade-offs arising from this legislation. Cao, Ho, and Garbaccio (2009) develop a dynamic computable general equilibrium (CGE) model of the Chinese economy to examine the economy-wide impact of the SO_2 policies in China's 11$^{\text{th}}$ FYP. Similarly, Cao, Ho, and Jorgenson (2009) study the costs and benefits of "green taxes" to manage Chinese air pollution, while Vennemo et al. (2009) broadly characterize trends in Chinese air and water pollution. We complement these studies by providing some of the first causal estimates of the policy change itself and characterizing the heterogeneity in firm-level responses across time and space. This paper also naturally complements the rich literature that studies the impact of Chinese environmental regulation on aggregate (Nam et al. (2014); Qi et al. (2014); Zhang et al. (2014); Zhang et al. (2016); Zhang et al. (2016)), regional (Zhang et al. (2013); Springmann et al. (2015); Kishimoto et al. (2017); Wong and Karplus (2017)), or firm-level (Cao and Karplus

(2014); Karplus and Zhang (2017)) emissions in China.[2]

In a closely related article, Chen et al. (2021) study an energy reform affecting large Chinese manufacturers which was incorporated as part of China's 11th FYP. They find that the energy conservation program caused a reallocation of output and energy consumption from more tightly regulated plants in a conglomerate to partner plants which are not affected by the policy change. They do not find any significant impact on energy efficiency across firms in a conglomerate. Building on their results, we provide complementary evidence of contemporaneous policy change across all firms. More specifically, our work differs from Chen et al. (2021) in the four following aspects. First, we focus on a key environmental outcome, emissions intensity, while Chen et al. (2021) characterize the intensity of energy consumption. We document that these objects are clearly related, but also differ due to firm-level emissions mitigation, such as the adoption of abatement technology. Second, Chen et al. (2021) estimate the average impact of more stringent energy regulation among treated firms and their related parties. We instead estimate the impact of regulation across the entire distribution of producers and characterize their heterogeneous responses to policy change. Third, Chen et al. (2021) study a very particular policy within the suite regulatory reforms embodied in China's 11th FYP. It has the advantage that the particular policy change, an energy conservation program among China's top 1000 firms, can be precisely identified in the data along with their related parties. In contrast, we use geography rather than size as a source of identification and, as such, estimate the impacts of the full suite of policy changes in China's 11th FYP firm-level emissions-intensity. Fourth, we demonstrate that the aggregate impact of the distributional evolution of firm-level responses can be interpreted through the lens of a resource allocation framework as an additional source of productivity gains, mitigating the severe firm-level emissions-output trade-offs estimated among large producers in China. Nonetheless, along the way, we document results that are broadly consistent with the primary findings from Chen et al. (2021) despite using an entirely different identification strategy; we find no impact of the policy on average energy or emissions intensity, but significant improvements in output per unit of coal consumed by the largest and most coal-intensive producers. We *also* find large, differential changes in emissions-intensity and abatement investment across the full distribution of producers, each of which contribute to industrial productivity growth and mitigate aggregate emissions-output trade-offs.

At the heart of our analysis lies the characterization of the distribution of causal, firm-level responses to policy change. Understanding the variation in response to policy change involves three fundamental challenges. First, although interactions (e.g., conditional ATT's) may provide

[2]Cao and Karplus (2014) study the drivers of energy, electricity, and carbon intensity among a sample of 800 Chinese firms, between 2005 and 2009. Our data covers a broader set of firms and allows us to investigate the policy impact of China's 11th FYP on SO_2 emissions, the primary environmental policy target.

a complementary understanding of treatment effect heterogeneity, Bitler et al. (2017) document that much heterogeneity in policy responses is not necessarily associated with observable covariates. Along these lines, we not only find significant differences in policy responses, but also document that the variation in policy responses across the distribution of heterogeneous producers remains large even within narrowly defined industries and locations. Second, characterizing the change in productivity, and thus the output-emissions trade-off, from environmental regulation requires characterizing responses across the entire distribution of firms, not just average changes among subsets of firms. Third, while it is possible to employ modern distributional estimators to capture heterogeneous responses, to our knowledge, an estimator that can be naturally applied to a setting where heterogeneous treatment effects emerge over time does not exist. While it is plausible, or even expected, that the effects of China's 11$^{\text{th}}$ FYP would appear after firms have had time to adapt production processes, install new technology, or increase abatement efforts, standard estimation approaches aim only at recovering the immediate impact of a policy change. To address this estimation problem, we extend standard quantile treatment effect estimators to an intertemporal setting. Further, we develop a series of pre-sample tests to characterize whether the suggested estimation approach is likely to violate the underlying identification assumptions in each approach. To our knowledge, these pragmatic tests are new to the quantile treatment effect literature and provide researchers with a new set of tools to help evaluate when particular distributional estimators are likely to deliver credible estimates. We find that a dynamic extension of the Callaway and Li (2019) QTT estimator performs well in our setting: there is no meaningful evidence of any significant impact of China's 11$^{\text{th}}$ FYP prior to 2007, the year after implementation, in any of our pre-trend exercises. Likewise, it outperforms leading alternatives, such as Athey and Imbens (2006) and Callaway et al. (2018).[3]

We proceed to study why the dirty producers respond to new regulation by reducing emissions, while the cleanest firms experience a relative rise in emissions intensity. Among the dirtiest firms (those in the top quantile), emissions intensity declines are largely mediated by *smokestack* abatement among coal-intensive firms with little pre-existing abatement capacity. [4] In no case do we find evidence of substitution towards cleaner energy sources across the distribution of heterogeneous producers.

At the other end of the emissions intensity distribution, initially clean firms in treated provinces experience rises in emissions intensity. We consistently find that this surprising re-

[3]Across estimators we find that a very similar qualitative pattern emerges: there is little impact of the policy on average, but there is evidence of some significant changes among pollution-intensive firms several years after implementation. That said, key differences across estimation approaches remain. We find significantly larger effects and more precise estimates using the Callaway and Li (2019) QTT estimator relative to comparable alternatives.

[4]By smokestack abatement we intend to capture all technologies to reduce emissions after coal consumption.

sult is driven smaller abatement investments among initially clean firms in treated provinces and disproportionately rapid output growth. That is, the changes in the regulatory landscape embedded in China's 11th FYP provided a competitive advantage for the initially cleanest producers. Among the cleanest producers in treatment provinces we observe disproportionately high output and emissions growth, but disproportionately small abatement. While surprising, the result that broad regulatory changes may lead a rise in emissions is not without precedent. Leslie (2018) finds that the introduction of a carbon tax to an imperfect wholesale electricity market increased carbon emissions by changing production incentives across heterogeneous firms; firm with dirtier marginal units disproportionately increased production in response to the reform. The implication of our result is similar: to the degree that the 11th FYP improved the relative productivity of clean firms in treated they were able to disproportionately benefit from the reform, even to the point of increasing emissions intensity relative to similar firms in control provinces.

While distributional impacts are often of interest in and of themselves (Sen, 1997; Carneiro et al., 2003), aggregation can be useful for linking disaggregated, case-specific outcomes (e.g., how does policy change affect firm-level emissions performance?) to broader policy objectives (e.g. how does policy change affect aggregate growth?). We quantify the benefits by marrying our intertemporal, causal estimated impact of policy change with standard models of manufacturing production and emissions (Shapiro and Walker, 2018; Forslid et al., 2018; Rodrigue et al., 2022b) through workhorse frameworks for evaluating resource allocation and productivity growth (Hsieh and Klenow, 2009; Gopinath et al., 2017). In this sense our work helps bridge the resource misallocation literature with that studying environmental regulation with market power distortions (Buchanan, 1969; Leslie, 2018; Preonas, 2023). In contrast, to preceding work differential gains arise from optimal resource allocation across heterogeneous producers rather than changes in market power. We find little evidence of any change in aggregate productivity early on in China's 11th FYP. By 2010, the decline in the dispersion of emissions intensity accounts for 2-3 percentage points of additional productivity growth, just less than half of the allocative productivity gains are attributable to the policy change.[5]

The next section presents the theoretical lens through which we link the distributional changes in emissions intensity to aggregate emissions, output and productivity. Section 3 briefly outlines the history of China's 11th FYP and documents differences in SO_2 reduction targets across provinces, while Section 4 documents key firm-level measures of environmental pollution, energy consumption, and economic performance across firms, provinces, and time. Section 5

[5]We also provide conditions under which distributional changes can be used to identify the magnitude of aggregate impacts which depend on both changes in the marginal distribution of emissions intensity and the distribution of heterogeneous productivity across firms.

describes the empirical specification, and characterizes the conditions needed to estimate quantile treatment effects over time. Section 6 documents distributional responses to policy change and investigates mechanisms driving differential firm responses, while Section 7 quantifies the aggregate implications of policy reform which differentially impacts firms across the distribution of Chinese manufacturers. Section 8 concludes.

2 Conceptual Framework

Hsieh and Klenow (2009) argue that variation in the observed value of marginal products is reflective of productivity distortions in developing countries. Blending a standard misallocation framework with workhorse models of production, abatement, and emissions (Copeland and Taylor, 2003; Shapiro and Walker, 2018; Forslid et al., 2018; Rodrigue et al., 2022b), we demonstrate that this logic extends to a structure with emissions: cross firm variation in emissions intensity is indicative of productivity reducing distortions in environmental regulation. To the degree that policy reform reduces resource misallocation across firms, it may mitigate output and emissions tradeoffs.

Specifically, consider a set of heterogeneous producers that purchase a bundle of physical inputs, V, in competitive markets and produce output Y according to the physical production function $Y_{is} = (1-\theta_s)\tilde{A}_{is}V_{is}$, where $\tilde{A}_{is}$ is firm-level TFP and θ_s captures the fraction of firm inputs that are redirected towards the reduction of firm-level emissions in sector s, E_{is}, $E_{is} = (1-\theta_s)^{1/\alpha_s}\tilde{A}_{is}V_{is} = (1-\theta_s)^{\frac{1-\alpha_s}{\alpha_s}}Y_{is}$. We can write output directly as a function of emissions in an emissions-augmented production function $Y_{is} = \tilde{A}_{is}^{1-\alpha}E_{is}^{\alpha}V_{is}^{1-\alpha_s} = A_{is}E_{is}^{\alpha_s}V_{is}^{1-\alpha_s}$, where $A_{is} = \tilde{A}_{is}^{1-\alpha_s}$ is abatement-inclusive productivity. For expositional clarity we assume that productive inputs are purchased in competitive markets at price W_s, while firm i faces the emissions tax T_{is} on each unit of emissions.

In keeping with standard market structure assumptions, we further assume firms face CES demand for their individual product, giving rise to the residual demand function $Y_{is} = (\Phi_s/P_{is})^{\sigma}$, where P_{is} is the price charged by firm i and Φ_s is a demand shifter common to all firms in sector s. Profit maximization then implies that the marginal revenue product for each input, including emissions, is inversely proportional to the intensity with which that input is employed in production

$$MPRV_{is} = (1-\alpha_s)\frac{\sigma-1}{\sigma}\frac{1}{I_{is}^{V}} = W_s \quad \text{and} \quad MRPE_{is} = \alpha_s\frac{\sigma-1}{\sigma}\frac{1}{I_{is}^{E}} = T_{is}, \tag{1}$$

where R_{is} is firm i's revenue, $R_{is} \equiv P_{is}Y_{is}$, and I_{is}^{E} and V_{is}^{E} are emissions and input intensity, $I_{is}^{E} \equiv E_{is}/R_{is}$ and $I_{is}^{V} \equiv V_{is}/R_{is}$. Measured revenue-based total factor productivity, $TFPR_{is}$, is

proportional to a geometric average of the firm's marginal revenue products of emissions and productive inputs:

$$TFPR_{is} \equiv P_{is}A_{is} = \frac{\sigma}{\sigma-1}\left(\frac{MRPE_{is}}{\alpha_s}\right)^{\alpha_s}\left(\frac{MRPV_{is}}{1-\alpha_s}\right)^{1-\alpha_s}. \tag{2}$$

Variation in emissions taxation drives variation in emissions intensity, $MRPE_{is}$, $TFPR_{is}$, and, consequently aggregate physical TFP in sector s:

$$TFP_s = \left[\sum_{i=1}^{N_s}\left(A_{is}\frac{\overline{TFPR}_s}{TFPR_{is}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}, \tag{3}$$

where $\overline{TFPR}_s$ is a geometric average of the average marginal revenue product of emissions and value added. Under the standard assumption that $\sigma > 2$, greater variance in $TFPR_{is}$ will cause TFP_s to fall. The wide dispersion of observed emissions intensities suggests that firms face very different emissions regulatory burdens. More broadly, changes in the distribution of emissions-intensities across firms directly influence aggregate TFP, aggregate output and the emissions-output tradeoff.

Consider, for example, the elasticity of the aggregate output with respect to the vector of effective firm-level regulation $\mathbf{T}_i = \{T_1, .., T_N\}$:

$$\frac{\partial \ln Y_s}{\partial \ln \mathbf{T}_{is}} = \frac{\partial \ln TFP_s}{\partial \ln \mathbf{T}_{is}} + \alpha\frac{\partial \ln E_s}{\partial \ln \mathbf{T}_{is}}$$

where we anticipate that stringent regulatory reform will aim to reduce emissions by increasing effective taxation, $\frac{\partial \ln E_s}{\partial \ln \mathbf{T}_{is}} < 0$. Any policy reform that leaves the allocation of resources (including emissions) unchanged across firms implies a stark, trade-off: given existing technology, reducing emissions can only be achieved by reducing output.[6] This implication is daunting for much of the developing world, including many Chinese provinces where income per capita remains persistently low (Ho and Li, 2008; Fan et al., 2011; Cao et al., 2013; He et al., 2018). However, policies which complement emissions reductions with aggregate productivity growth through improved resource allocation need not come at such a high cost. Indeed, equation (3) suggests that the output cost of policies that target a reduction in emissions can be offset by complementary productivity growth through improved resource allocation, $\frac{\partial \ln TFP}{\partial \ln \mathbf{T}_{is}} > 0$.[7] As we document below, the 2006-2010 enactment period of the 11$^{\text{th}}$ FYP was characterized by both a sharp increase in environmental regulation, but also marked declines in the variance of emissions

[6] For example, in our context, a policy which increased $\mathbf{T_{is}}$ by a fixed percentage would have this effect.

[7] This is consistent with the theoretical finding that it is efficient for each unit of emissions to face the same price Shapiro (2022).

intensity.

3 China's 11th Five-Year Plan

Since the early 1980s, China has sustained remarkably high rates of economic growth. This has coincided, however, with the rapid expansion of energy- and pollution-intensive industries. In the battle between economic growth and environmental sustainability, economic growth has typically carried the day. In many instances, policies aimed at reducing the health and environmental consequences of unmitigated economic growth have been either unenforced (OECD (2009)) or ineffective (Shi and Xu (2018)). More recently, the balance between economic growth and environmental security has shifted. China's 11th FYP established stringent new environmental targets, outlined location and industry-specific environmental criteria, and instituted new internal directives whereby the successful promotion of local government officials would crucially depend on meeting environmental objectives (Shi and Xu (2018)).[8]

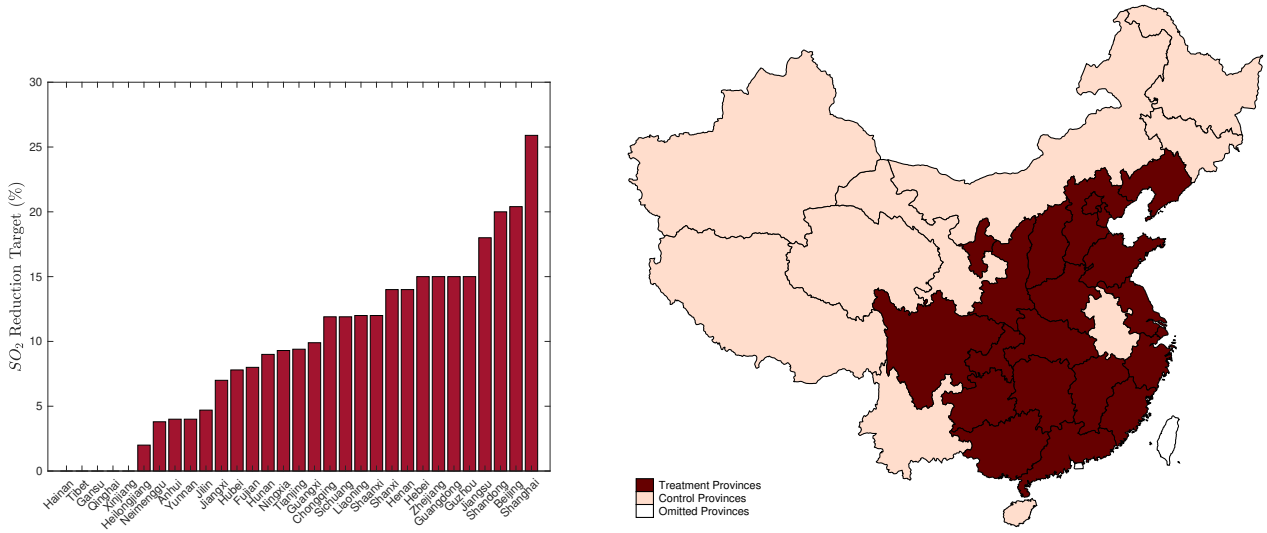
The 11th FYP distinctly emphasized the government's objective to reduce SO_2 emissions by 10 percent[9] and, in August 2006, the State Council outlined a specific SO_2 reduction target for each Chinese province. As emphasized in Xu (2011) and Shi and Xu (2018), provincial targets were the outcome of a series of complex, bilateral and secretive negotiations between the Vice-President of each provincial government and central government officials from the State Environment Protection Administration. In principle, provincial reduction targets varied with initial differences in environmental quality, environmental and abatement capacity, initial SO_2 emissions, and economic development; in practice, the opaqueness and bilateral nature of negotiations did not provide firms with a meaningful opportunity to anticipate individual outcomes.[10]

While some provinces were given very large emission reduction targets during the 11th FYP,

[8]Between 1953 and 2000, the Chinese government has proposed nine distinct Five-Year plans. There was a short gap between the second (1958–1962) and third (1966–1970) Five-Year plan. In general, economic growth was the central objective of each of these governmental directives. The 10th FYP (2001–2005) was the first to include an environmental objective: a 10 percent reduction of SO_2 emissions for the country as a whole. While the 10th FYP only stated broad principles, it did not include a clear evaluation scheme, a credible incentive program for local government officials, or sufficient institutional investment (Xu (2011)). As such, it was not surprising that it was broadly determined to be ineffective (Shi and Xu (2018)).

[9]Numerous studies have found that the decline in China's air quality has resulted in significant health and environmental costs (Ebenstein et al. (2015); Bombardini and Li (2020)) and has sparked an urgent call for stricter regulation, particularly among China's rapidly growing coastal cities (Kahn and Zheng (2016)). At the same time, Chinese industry heavily relies on coal; recent estimates suggest that the industrial sector consumes nearly 95 percent of China's coal (China Power Team (2016)). Increasing energy costs have the potential to significantly slow China's development.

[10]Moreover, Xu (2011) demonstrates that the targets formalized in the signed contracts were orthogonal to differences in firm characteristics across provinces.



(a) Provincial SO_2 Reduction Targets

(b) Provinces with High/Low Targets

Notes: The above figure reproduces SO_2 reduction targets across Chinese provinces. All provinces with a target above 5 percent are treated as having high reduction targets, while provinces with targets of 5 percent or less are treated as having low reduction targets.

Figure 1: SO_2 Reduction Targets

other, often adjacent, provinces were given very small emission reduction targets or none at all. Figure 1, retrieved directly from state documentation (State Council (2006)), illustrates the difference across Chinese provinces. Despite heterogeneity in the reduction targets across provinces, we observe a natural definition of a binary treatment variable: while 21 provinces received large reductions of nearly 10 percent or greater, 10 provinces were asked to reduce SO_2 by only a few percent or not at all. As a starting point, we consider all provinces with a target above 5 percent as having high reduction targets, while the remaining provinces are considered locations with low reduction targets.

To support the credibility of the emissions targets, in 2007 the State Council handed down three specific criteria for provincial officials to receive successful performance evaluations (State Council (2007b); Wang (2013)):

1. Attaining the quantitative SO_2 reduction targets and improving overall environmental quality;

2. The establishment and operation of institutions that set environmental objectives for major pollutants, monitor progress towards said objectives, and evaluate environmental programs;

3. Direct implementation of pollution mitigation measures including the installation and

operation of pollutant removal facilities and equipment, such as scrubbers.

Indeed, there is significant evidence that in the same year, 2007, the policy change began to affect the cost of production. Government-mandated SO_2 emissions charges for coal power plants immediately increased to \$166 USD per ton, double their preceding rate (Xu et al. (2009)). Moreover, the State Council issued notice shortly thereafter that SO_2 emissions rates among all producers would face the same increased rates by 2010 (State Council (2007a)).

On the one hand, the 11th FYP has two key characteristics that make it propitious for policy evaluation. First, to the best of our knowledge, there is no *causal* evidence of the impact of this policy change on firm-level emissions in China. This is a significant economic issue both within China, where there is growing concern regarding the health impacts of pollution, and abroad, given China's commitment to reduce its contribution to global pollution (United Nations (2015)).

Second, although the national policy required very different reduction targets across locations, the implementation began simultaneously in all provinces in 2007. Arguably, some adjustment could have taken place in 2006, the year during which the provincial targets were being set and prior to the announcement of the evaluation criteria, but most of the evidence suggests that *actual* regulatory and enforcement changes began in 2007. As such, we generally treat 2006 outcomes as preceding the policy change, but will conservatively rely on data prior to 2006 to help identify the policy impact of China's 11th FYP.

On the other hand, a policy evaluation challenge arises due to the fact that we do not directly observe the regulatory burden within or across provinces; it well known that effective environmental regulation across Chinese firms has varied widely within narrowly defined industries and locations (Jia, 2012; Wu et al., 2013; Deng et al., 2020; Wang and Lin, 2022; Yang et al., 2022). Moreover, while emissions taxes are a common regulatory tools in China, they are far from the only mechanism used to induce compliance; indeed, there is a long history of using quotas, taxes, fines, loan conditions, restricted export market access, and public pressure, among other regulatory tools to curb Chinese emissions. Accordingly, we consider an empirical strategy that flexibly quantifies the impact of the suite of policy changes embedded in China's 11th FYP on emissions intensity across firms and time. We also characterize mechanisms that drove changes in firm-level emissions-intensity, such as abatement, energy-efficiency or changes in the mix of firms energy sources, to yield insight into the key levers touched by the regulatory change.[11]

[11]This feature is in contrast to Chen et al. (2021) which characterizes the impact of a specific change in energy conversation requirements within firms in the same conglomerates. We do not have, nor do we know of, a summary measure for the full set of policy changes across firms, time and geography. However, in 2004, but no other year, we observe emissions discharge fees among manufacturing emitters. Appendix A.2 plots firm-level discharge fees per unit of emissions against firm-level emissions intensity. As expected, (i) emissions-intensive firms face a systematically lighter regulatory burden in the form of average emissions fees and (ii) among firms

4 Data

Our primary data source is China's Environmental Statistics Database (CESD), a micro-level database compiled by China's Ministry of Environmental Protection (MEP), over the 2004–2010 period.[12] Key to our study, the survey reports the total weight of SO_2 emitted by a firm over the course of a calendar year. It likewise captures firm-level energy consumption, the capacity of sulfur disposal facilities (e.g., scrubbers), and measures of firm size and revenues. Summary statistics are documented in the appendix.[13]

Although the MEP surveys the universe of polluting establishments, it has two limiting features that we note from the outset. First, the data provided to us only reports information for firms when one of its pollutants falls into the top 85 percent of the total volume emitted in a given county for at least one of the pollutants tracked by the MEP. This requirement holds across a set of pollutants and, in general, the data captures more than 85 percent of the volume emitted in any location and year (Wang et al. (2018); Zhang et al. (2018)). Second, after 2005 the reported survey data excludes all thermal power plants. While thermal power plants represent roughly half of all industrial pollution in China, data limitations prevent us from characterizing their response to the 2006 policy change. Nonetheless, there remains sufficient information to check the accuracy of our data by benchmarking it to aggregate statistics produced by the MEP.

We begin by aggregating the firm-level data for each year and comparing annual variation across data sources, with and without the contribution from thermal power plants. Figure 2a shows that the aggregated firm-level survey always lies below the aggregate emissions reported in Chinese Statistical Yearbooks. However, this difference is roughly constant over time and displays similar year-to-year changes.[14] It also documents that the early declines in SO_2 emissions were largely driven by emissions reductions among Chinese power plants. While this might suggest that the change in policy had little impact on Chinese manufacturers, we later document that there were substantive changes in emissions outcomes across the distribution of heterogeneous manufacturers.

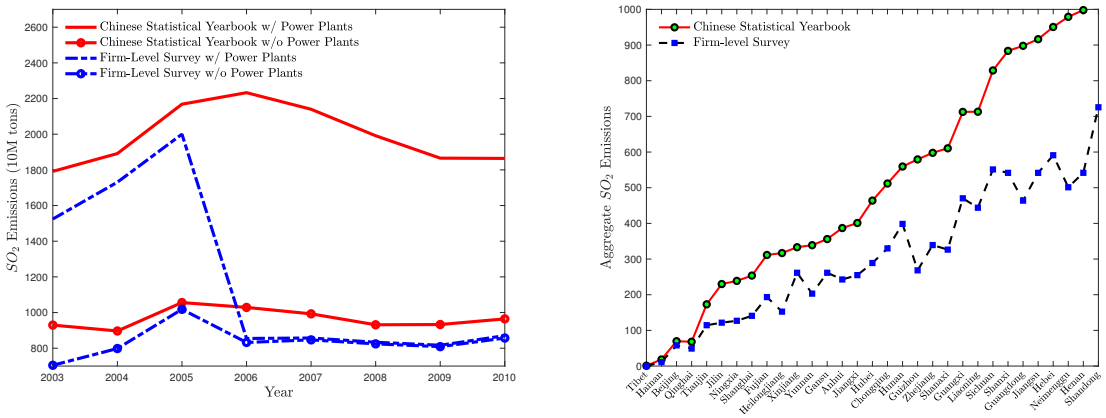
Next, we compute average provincial emissions from the firm-level data and compare them to reported provincial averages reported by the MEP over the same time frame, in Figure 2b. It

with facing the same average emissions fees there remains a wide variation in emissions intensity. Brunel and Levinson (2016) provide further discussion of the challenges associated with holistically measuring the stringency of environmental regulation.

[12] This data set has been used in numerous papers including He et al. (2020), Liu et al. (2021), He and Huang (2022) Rodrigue et al. (2022a), Rodrigue et al. (2022b), and Chen et al. (2021) among others.

[13] Appendices A.1-A.3 provide a detailed description of how emissions are measured along with a series of data validation exercises. Since these data have been used widely, we omit further discussion hereafter.

[14] The careful reader will notice that there is a slight divergence in the two aggregate SO_2 series in 2006, the year that China announced the 11^{th} FYP. This temporary change in correlation between the two series is driven entirely by a temporary change in reporting from Beijing. Our benchmark sample excludes Beijing entirely.



(a) Variation Over Time (b) Variation Across Provinces

Notes: Panel (a) plots aggregate SO_2 emissions from the Chinese Statistical Yearbook including thermal power plants (solid red line), aggregate SO_2 emissions from the Chinese Statistical Yearbook excluding thermal power plants (red solid line with dots), aggregated SO_2 emissions from the firm-level survey including thermal power plants (dashed blue line), and aggregated SO_2 emissions from the firm-level survey excluding thermal power plants (dashed blue line with dots). Panel (b) compares average provincial SO_2 emissions across provinces from the Chinese Statistical Yearbook with the aggregated firm-level survey, averaged over time.

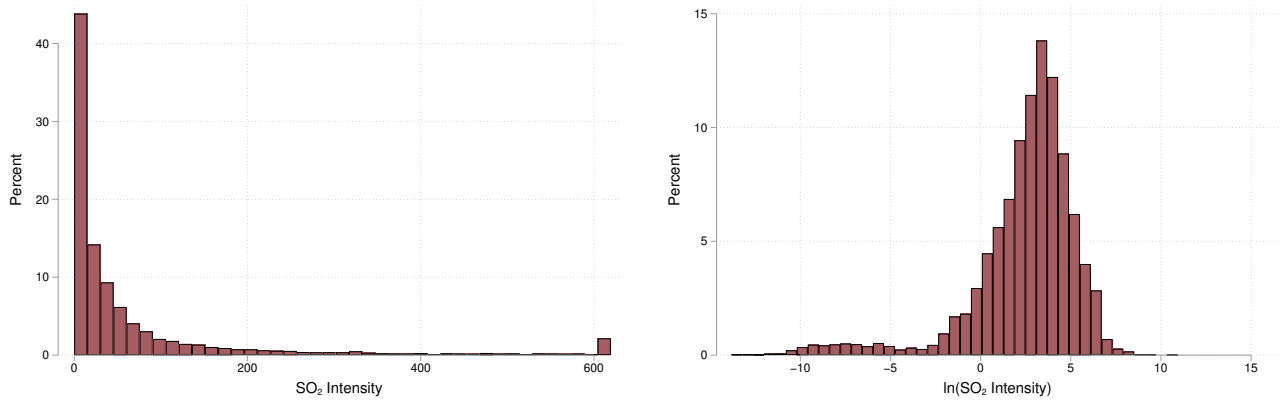
Figure 2: Data Validation, Benchmarking to Official Statistics

is clear that both data sources track each other closely and there is significant variation across provinces in the aggregate level of pollution. On average, the data captures 60 percent of total industrial emissions in a given year and captures the spatial distribution of emissions reasonably well.[15] Nonetheless, there remain significant differences in coverage rates across locations; we accordingly take this into account in our empirical specification.[16]

We are primarily interested in the impact of higher SO_2 emissions targets on firm-level emissions intensity. While large, productive firms will naturally emit more SO_2 because they produce more output, this should not necessarily imply that they are the key targets of the policy change. Rather, as documented in Rodrigue et al. (2022b), large, productive Chinese firms tend to produce fewer emissions per unit of output relative to their less productive counterparts. Should provincial policy aim to meet the nationally mandated objectives by reducing the pollution of the most intensive emitters, we expect that emissions intensity will demonstrate a greater response. Following the literature and consistent with our theoretical framework, we

[15]For data from before 2006, thermal power plants are included and, therefore, the data provides fuller coverage of industrial pollution. Using the 2003–2005 portion of our sample, our computation shows that the MEP data set captures over 90 percent of all industrial pollution. Likewise, the differences across provinces shrink once we include this information. Further details are reported in Rodrigue et al. (2022a).

[16]Further checks on emissions data quality, including comparisons with satellite measures of SO_2 emissions, can be found in the appendix. The primary analysis also relies on revenue data reported in the MEP survey. We cross-check the quality of this data series with the more commonly used ASIP (Annual Survey of Industrial Production) data set for Chinese enterprises. The cross-data source correlation coefficient was 0.99, confirming the accuracy of this series.



(a) SO_2 intensity in levels

(b) ln(SO_2 intensity)

Notes: The above figure documents the distribution of firm-level SO_2 emissions intensity. Upper values in panel (a) are top-coded for expositional purposes.

Figure 3: Distribution of SO_2 Emissions and Intensity

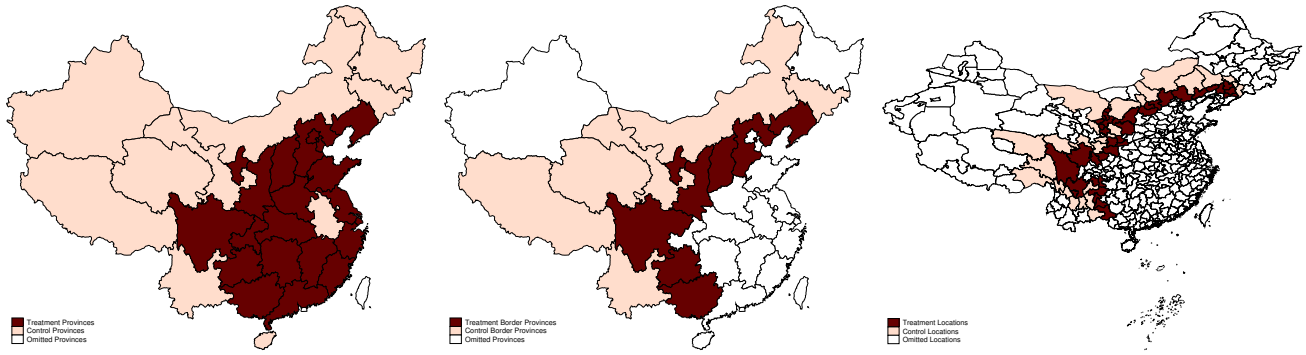
measure emissions intensity as the total quantity of emissions, E_{it}, divided by deflated firm-level revenue, R_{it}.

Figure 3 plots the firm-level distribution of SO_2 emissions intensity across sample years. Analogous to standard firm-size distributions, the firm-level emissions intensity can be broadly described as roughly following a log-normal distribution. Accordingly, we focus on the natural log of firm-level SO_2 emissions-intensity as the primary outcome variable in our empirical exercises.[17]

4.1 Policy Reform Across Chinese Provinces

To characterize the geographic variation in reduction targets, we distinguish firms located in provinces with new, stringent emissions regulation targets (treatment provinces) and those which are roughly maintaining the status quo (control provinces), where, as a starting point, we use a 5 percent threshold to determine if a province faces increasingly stringent SO_2 reduction targets. Examining the geographic distribution of emission targets, displayed in Figure 4a, two patterns are immediately evident. First, coastal provinces generally received significantly higher SO_2 reduction targets than inland provinces. Second, the large majority of non-treated (control) provinces form a contiguous border with a similar set of treated provinces. Given that the geographic landscape, industrial composition, exposure to international markets, and socioeco-

[17] Although the vast majority of firms in our sample report positive SO_2 emissions, firms occasionally report zero SO_2 emissions. Since this outcome is entirely possible we do not drop these firms from the data in our benchmark exercises, but we add the value of 1 to all observations and later investigate the degree to which our results are sensitive to this adjustment.



(a) Defn 1 (Full Sample) (b) Defn 2 (Border Provinces) (c) Defn 3 (Border Prefectures)

Notes: The above figure outlines three treatment definitions. The first definition of treatment considers all firms located in provinces with a target above 5 percent as "treated" while firms in the remaining provinces are used as control firms. The second definition of treatment requires that treated firms are located in provinces that were given a target above 5 percent and in provinces along the Northwest policy border. Likewise, control firms are those in provinces that were given targets below 5 percent and are located in provinces along the Northwest policy border. The third definition of treatment is the same as the second except that we restrict attention to cities/prefectures directly on the policy border.

Figure 4: Binary Treatment Definitions

nomic conditions are relatively similar across border provinces, our second definition of treated and control firms provides a robust comparison controlling for various sources of unobserved heterogeneity. We first restrict attention to those firms located in the border provinces prior to the policy change (Figure 4b) and employ this second definition of treatment to isolate the impact of China's 11$^{\text{th}}$ FYP on firm-level environmental performance.

On the one hand, comparing similar firms located in geographically adjacent border provinces inherently reduces potential sources of bias but omits a large number of coastal firms, arguably those most affected by the policy change. On the other hand, many of the border provinces in our second definition of treated provinces have cities (prefectures) that lie directly on the opposite side of the treatment border. Restricting our sample to these border cities provides even greater confidence that observed differences can be attributed to policy rather than unobserved differences across locations. Geographic restrictions, however, severely limit sample size. For instance, the full sample definition of treatment (Figure 4a) consists of 4.5 times more observations than the benchmark (Figure 4b), while the benchmark sample is nearly three times larger than the sample of border cities (Figure 4c). To check the consistency of our results, we study the robustness of our benchmark findings in the full and border city samples.

The data structure further allows us to evaluate both pre-treatment and post-treatment periods. We characterize the immediate impact of the change in Chinese environmental regulation, but also how this impact evolved over the course of the 11$^{\text{th}}$ FYP. Given that changes in firm capacity, energy sources, and technology likely take significant time to manifest, evaluating outcomes across years provides insight into the "time-to-adjust" necessary to evaluate the full

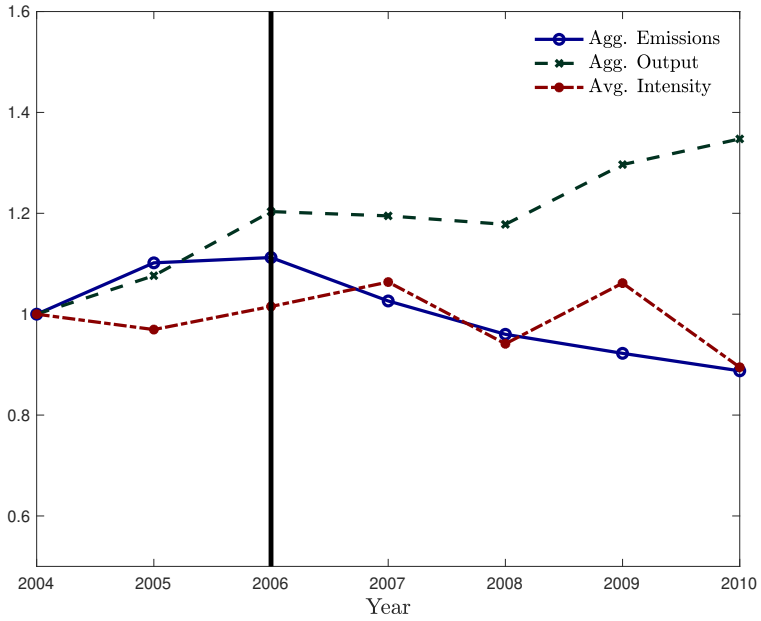
impact of the policy change. As we outline below, existing distributional estimators cannot generally recover time-varying policy impacts. We extend existing methodologies to a setting that allows for intertemporal policy impacts to characterize the distributional evolution of firm behavior in response to the 11^{th} FYP.

4.2 Mechanisms, Trade-offs, and Unobserved Heterogeneity

The data provide a series of variables that we use to study the mechanisms through which the 11^{th} FYP affects firm-level emissions outcomes. The survey collects, for each firm, the quantity of coal consumed in a calendar year, a key determinant of SO_2 emissions, along with the consumption of other energy sources (e.g. diesel fuel, petroleum). Likewise, any investment in abatement technology (e.g., the total capacity of scrubbers) is recorded for every establishment. We also observe a measure of total firm-level revenues. Changes in the distribution of revenues allow us to characterize the degree to which China's 11^{th} FYP differentially affected firm sales growth across the distribution of heterogeneous producers.

The environmental data do not, unfortunately, provide a measure of capital intensity or firm ownership. Given that large firms, foreign firms, and state-owned firms have a history of receiving special treatment from the Chinese government, we might expect that these are important determinants of the firm's response to the policy change. Similarly, firm size, emissions, and emissions intensity are likely to systematically vary across industries with very different degrees of capital intensity.

We address these data limitations with three empirical adjustments. First, the Callaway and Li (2019) approach allows for both firm-specific unobserved heterogeneity and time-fixed effects. Given that the type of output produced by the firm, its capital stock, and its ownership change slowly over time, if at all, we can control for the impact that these characteristics have on firm-level responses. Second, by focusing on border provinces, our benchmark estimates restrict attention to firms that operate under similar economic, political, and geographic constraints. To the extent that the unobserved differences co-vary with the observed differences across locations, the impact of these confounders should be minimized. Third, we develop a series of new pre-trend tests for distributional estimators. As described below, each test intuitively relates to standard DID approaches, guides model specification, and provides confidence that the underlying treatment and control observations can be employed to recover the target treatment effects.



Notes: The above figure plots aggregate manufacturing emissions (solid blue line with circles), aggregate manufacturing production (dashed green line with x's), and the average firm-level emissions intensity (dotted red line with dots) for the benchmark sample (treatment definition 2). Manufacturing production is measured as deflated revenue.

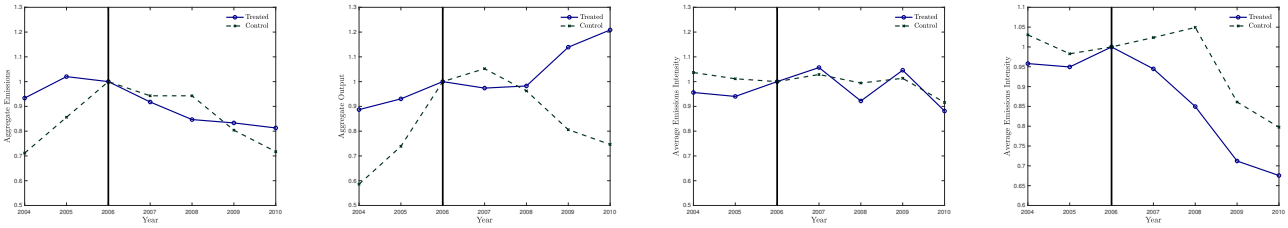
Figure 5: Emissions, Output, and Intensity Over Time

4.3 Empirical Patterns

Before describing our empirical specification, we document two empirical patterns that characterize China's 11$^{\text{th}}$ FYP across the distribution of heterogeneous producers. First, using the benchmark sample (treatment definition 2), Figure 5 documents that aggregate emissions rose between 2004 and 2006 but declined continuously thereafter; aggregate emissions were 11 percent lower in 2010 than they were in 2004. Aggregate manufacturing output, in contrast, rose strongly over the 2004–2010 period.

A common explanation for simultaneous emissions declines and output growth is secular productivity growth. Figure 5 suggests that may be the case in China: average firm-level emissions intensity declined alongside aggregate emissions, particularly over the 2007–2010 period. To the degree that treated provinces witnessed disproportionate improvements in technological performance, there may be reason to suspect the productivity growth is linked to policy change.

Figure 6 compares the evolution of emissions, output, and intensity across treatment status. Emissions demonstrate strong growth in control provinces in the pre-sample period (2004–2006) before modestly declining back to their 2004 levels. In contrast, emissions growth is more modest among the treated provinces in the pre-sample period and declines somewhat more slowly



(a) Emissions (b) Output (c) Mean Intensity (d) Standard Deviation

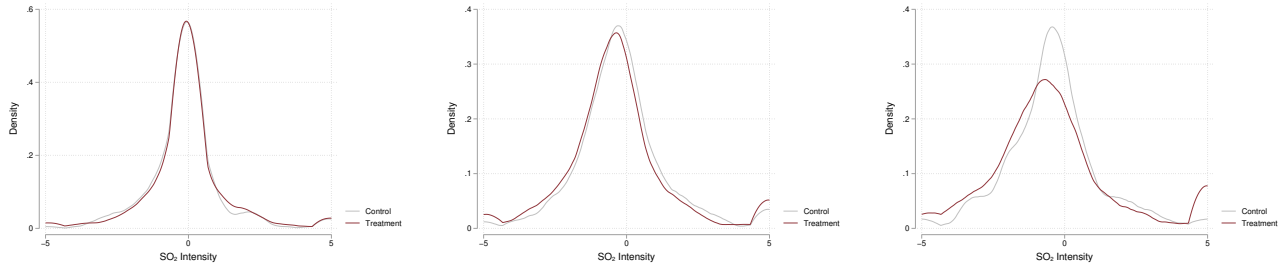
Notes: The above figure documents the path of emissions growth (panel a), output growth (panel b), average firm-level emissions intensity (panel c), and the standard deviation of emissions intensity (panel d) across treated and control firms.

Figure 6: Emissions, Output, and Intensity Across Treatment and Control Firms

thereafter. Panel (b) highlights the disproportionate output growth in control provinces between 2004 and 2006. Thereafter, output is relatively stable in both groups before disproportionately rising among *treated* firms in the latter half of the sample. This growth is surprising: firms in these provinces are subject to disproportionate regulation during China's 11$^{\text{th}}$ FYP.

Could complementary changes in technological adoption, potentially spurred on by the 11$^{\text{th}}$ FYP, help explain the differences across locations? On the one hand, panel (c) of Figure 6 suggests not: there is very little difference in average firm-level emissions intensity before and after the implementation of the policy. On the other hand, panel (d) documents a sharp, disproportionate decline in the standard deviation of emissions intensity of treated firms after 2007. This decline is consistent with a reduction in the variance of distortions across heterogeneous producers (Rodrigue et al. (2022a)) and a resource allocation driven improvement in aggregate productivity (Hsieh and Klenow (2009)).

Figure 7 further investigates relative changes in the distribution of emissions intensity *growth*, $\Delta \ln(\text{SO}_2 \text{ intensity})_{it} = \ln(\text{SO}_2 \text{ intensity})_{it} - \ln(\text{SO}_2 \text{ intensity})_{i,05}$, across treatment and control provinces in 2006, 2008 and 2010. In each panel, we provide the p-value from a Kolmogorov-Smirnov test for distributional equivalence in each year. There are three empirical features that merit comment. First, in all years and all provinces we observe a large number of firms that experience sharp declines in emissions intensity, while an almost equally large fraction of firms experience sharp increases. China's 11$^{\text{th}}$ did not result in a uniform decline in emissions intensity across all firms. Second, in 2006 the treatment and control distributions are almost indistinguishable, as confirmed by a large Kolmogorov-Smirnov p-value. Subsequently, between 2008-2010 there is a modest leftward shift in average emissions intensity growth among of the treatment province distribution relative to the control province distribution. Third, *both* tails of the emissions intensity growth distribution are larger among treatment provinces between 2008-2010. That is, there appears to be a substantial number of firms with large changes in



(a) 2006, KS p-value=0.522 (b) 2008, KS p-value=0.040 (c) 2010, KS p-value=0.019

Notes: The above figure documents the distribution of the log change in firm-level SO_2 intensity across treatment and control provinces. For expositional purposes, log changes in emissions intensity greater than 5 (less than -5) are fixed to 5 (-5). In each panel, we provide the p-value from a Kolmogorov-Smirov (KS) test for distributional equivalence across treated and control firms..

Figure 7: Distribution of $\Delta\ln(SO_2$ Intensity) Across Treatment and Time

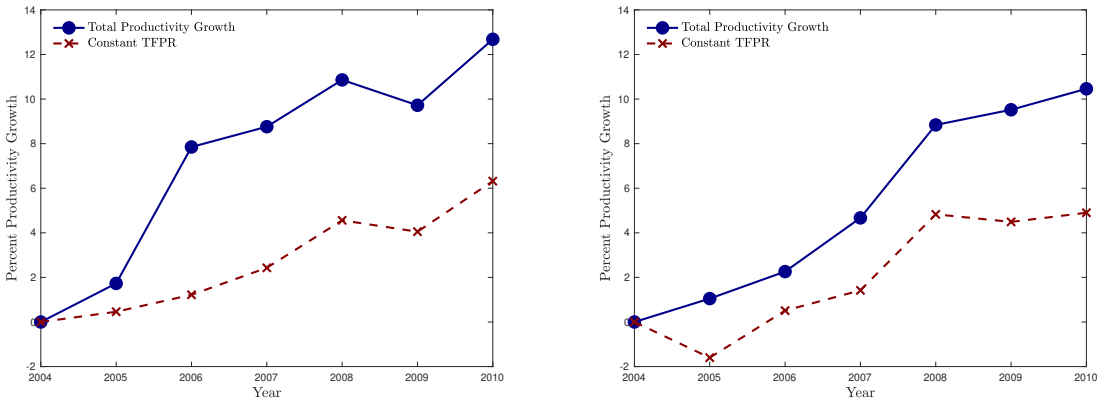
emissions intensity in treatment provinces; while many firms are experiencing sharp emissions intensity declines, other are undergoing large increases at the same time. Moreover, standard Kolmogorov-Smirnov tests between 2008-2010 suggest significant differences across treatment status in each year.

4.4 Quantifying Emissions-Intensity Driven Productivity Gains

Differential changes in firm-level emissions-intensity can be interpreted as a source of aggregate productivity gains if they induce a decline in the variance of emissions-intensity, consistent with the theoretical framework in Seciton 2. However, quantifying the aggregate gains from the evolving distribution of emissions intensity depends not only emissions-intensity itself but also requires (i) knowledge of the parameters governing emissions technology, α, and the elasticity of demand, σ. Since estimating these parameters is beyond the scope of our data, we rely on established estimates from Rodrigue et al. (2022a) to pin down α for each Chinese industry and set $\sigma = 3$, as in Gopinath et al. (2017). With these parameters in hand, it is possible to recover marginal revenue products, $MRPE_{it}$ and $MRPV_{it}$, and the firm's underlying productivity, A_{it}.[18] While we have measures of emissions and revenues for each firm and year, we can only match productive inputs for a subset of firms. As such, we fix $MPRV_{it} = 1$ for all firms and years and state plainly that this margin of adjustment is fixed in both the benchmark and all counterfactual series. As such, counterfactual aggregate productivity gains should be strictly interpreted relative to the benchmark series.

We benchmark the degree to which the changes in emissions intensity over the 2004-2010

[18]See Appendix A.6 for details.



(a) All Provinces (b) Treatment Border Provinces

Notes: The solid blue line with dots plots the observed path of the aggregate TFP, while the red line with x's captures the counterfactual path of TFP if $MRPE_{it}$ and, as a consequence, $TFPR_{it}$ are held constant at their 2004 values. Panel (a) plots the implied changes over all provinces, while panel (b) restricts attention to our benchmark sample of treatment and control provinces.

Figure 8: Emissions-Driven Productivity Growth

period affected the evolution of aggregate productivity by comparing the observed path of aggregate productivity to one which would have prevailed in the absence of changes in the emissions intensity distribution. In Figure 8 the solid blue line with dots plots the observed path of the aggregate TFP, while the red line with x's captures the counterfactual path of TFP if $MRPE_{it}$ and, as a consequence, $TFPR_{it}$ are held constant at their 2004 values. Panel (a) plots the implied changes over all provinces, while panel (b) restricts attention to our benchmark sample of treatment and control provinces.

There are two key takeaways from this exercise. First, viewed through the lens of a standard growth accounting model, the implied productivity gains from changes in the emissions-intensity distribution are economically substantive when evaluated on the basis of productive efficiency *alone*. Using the full sample we observe a 6.4 percentage point gap between the observed and counterfactual productivity emerging by 2010, while in the benchmark sample of comparison provinces the path of aggregate manufacturing TFP is 5.6 percentage points greater than that when $MPRE_{it}$ is held constant.[19] The importance of the improved productive efficiency is even more impressive recalling that emissions declined in our benchmark sample (Figure 5). That is, TFP gains did not come at the cost of greater emissions, ceteris paribus, indicating a significant decline in resource misallocation.

Second, despite the emissions-driven productivity gains implied by the quantitative model,

[19]In output terms, improved resource allocation cumulatively accounts for nearly 20 percentage points of additional aggregate production over the sample period.

it is not obvious they were driven by the 11$^{\text{th}}$ FYP. In panel (a), which uses the full sample of Chinese provinces, we observe that there is virtually no difference between the observed and counterfactual TFP in 2005, a very small difference in 2006, and a large gap thereafter. The corresponding figure for the sample of border provinces, in contrast, displays a gap that emerges immediately and grows more slowly thereafter. This may reflect that firms in coastal provinces had adjusted production in anticipation of future stringent regulation or different pre-existing trends in environmental investment across locations. Regardless, Figure 8 raises the concern that ex-ante conditions across coastal and inland provinces are widely different, even conditional on observable covariates. Attributing relative differences across firms located in coastal and inland provinces to the 11$^{\text{th}}$ FYP may lead to spurious findings that can be explained equally well by differential trends across China's vast geographic landscape. We address this concern explicitly in our empirical specification below.

5 Empirical Specification

To identify the direct effect of SO_2 reduction targets on emissions intensity across the distribution of firms, we follow a distributional DID strategy. Because no existing distributional estimator naturally allows for intertemporal treatment effects, we extend common distributional estimators to characterize the evolution of China's manufacturing emissions-intensity distribution over the course of the 11$^{\text{th}}$ FYP. Below, we focus on our preferred estimator, the quantile treatment effect on the treated, as posited in Callaway and Li (2019), and summarize its benefits relative to competing alternatives. Appendix B provides a detailed discussion of the underlying differences in estimation assumptions, along with practical examples to build intuition for observed differences in estimator performance.

Consider a setting with $\mathcal{T}$ time periods and use $t = 1, \ldots, \mathcal{T}$ to index particular time periods. We also suppose that some firms become treated in time period t^*. In our particular case, $\mathcal{T} = 7$ and $t^* = 3$, but our arguments in this section apply more generally. We denote treated potential outcomes for firm i in time period t by $Y_{it}(1)$ – this is the outcome that firm i would experience in time period t if it had been treated in period t^*. We denote untreated potential outcomes for firm i in time period t by $Y_{it}(0)$ – this is the outcome that firm i would experience in time period t if it was not treated.

Typically, the main parameter of interest in a DID setup is the average treatment effect on the treated (ATT); in our setup we are additionally interested in quantile treatment effects on the treated (QTT), which introduces additional identification challenges for recovering distributional

effects with multiple time periods. These are defined as

$$ATT_t = E[Y_t(1) - Y_t(0)|D = 1] \quad \text{and} \quad QTT_t(\tau) = Q_{Y_t(1)|D=1}(\tau) - Q_{Y_t(0)|D=1}(\tau),$$

where both parameters are indexed by the time period t. The QTT is additionally indexed by quantile $\tau \in [0, 1]$, and $Q_{Y_t(j)|D=1}(\tau) = F^{-1}_{Y_t(j)|D=1}$ for $j = \{0, 1\}$.[20] To identify the ATT_t and QTT_t in periods where $t \geq t^*$, the key identification challenge is to recover the unobserved outcome for treated firms, $E[Y_t(0)|D = 1]$ (for the ATT) and $F_{Y_t(0)|D=1}$ (for the QTT). As typical, parallel trends assumptions are helpful for recovering both the ATT_t and the QTT_t.

Assumption 1 (Parallel Trends Assumptions)

Mean Parallel Trends: $E[Y_t(0) - Y_{t^*-1}(0)|D = 1] = E[Y_t(0) - Y_{t^*-1}(0)|D = 0]$

Distributional Parallel Trends: $Y_t(0) - Y_{t*-1}(0) \perp\!\!\!\perp D$

Both versions of the parallel trends assumption in Assumption 1 say that the path of outcomes that treated firms would have experienced in the absence of participating in the treatment (this path is unobserved) is the same as the path of outcomes that untreated firms experienced (this path is observed). The mean parallel trends assumption says this holds on average; the distributional parallel trends assumption says that the entire distribution is the same (which is a stronger condition but in the same spirit). Under mean parallel trends, it immediately follows that

$$E[Y_t(0)|D = 1] = E[Y_t(0) - Y_{t^*-1}(0)|D = 0] + E[Y_{t^*-1}(0)|D = 1]$$

where all terms are identified. This step follows by (i) the mean parallel trends assumption and (ii) the linearity of expectations.

Identifying $F_{Y_t(0)|D=1}$, the distribution of outcomes that treated firms would have experienced if they had not been treated, is more complicated. The distributional parallel trends assumption implies that $F_{Y_t(0)-Y_{t^*-1}(0)|D=1} = F_{Y_t(0)-Y_{t^*-1}(0)|D=0}$, the latter of which is identified. In addition, $F_{Y_{t^*-1}(0)|D=1}$ is a directly identified distribution since $Y_{t^*-1}(0)$ are observed outcomes for firms in the treated group. However, identifying $F_{Y_t(0)|D=1}$ hinges on identifying the joint distribution of $(Y_t(0) - Y_{t^*-1}(0))|D = 1$ and $Y_{t^*-1}(0)|D = 1$; the previous arguments imply that the marginal distributions are identified. This is the important difference between identifying ATT_t and QTT_t.

[20] To be clear, the QTT is the difference between marginal distributions, which means that the *identity* of high-polluting firms could change across regimes. For example, if $QTT(0.9)$, the difference between the 90th percentiles, is negative, it would suggest that the amount of pollution coming from the highest polluting firms is lower than it would be in the absence of the policy. In addition, comparing the distribution of outcomes under the policy relative to a counterfactual distribution of outcomes in the absence of the policy is often the key ingredient for ranking different policies according to social welfare criteria (see Sen (1997) and Carneiro et al. (2003) for discussion along these lines).

Intuitively, we want to know if the biggest increases in pollution, if the policy had not been implemented, would have been by firms with initially the most pollution or least pollution, etc. The parallel trends assumption does not provide information about this additional key piece of information. To address this issue, we appeal to Sklar's theorem which states that joint distributions can be written as the copula of the marginal distributions where the copula *couples* the marginals together into a joint distribution and provides the additional information about the dependence between the marginals (Sklar (1959)). Callaway and Li (2019) and Callaway et al. (2018) replace the unknown dependence between the path of untreated potential outcomes and its level in a pre-treatment period with an observed copula for some other time period (Callaway and Li (2019)) or for some other group (Callaway et al. (2018)). Next, we state corresponding versions of these assumptions extended to the case with multiple periods.

Assumption 2 (Copula Assumption)

Copula Stability Assumption (i): $C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}(0)|D=1} = C_{Y_{t^*-1}(0)-Y_{t^*-\kappa-1}(0),Y_{t^*-\kappa-1}(0)|D=1}$ *where* $\kappa = t-(t^*-1)$*, which is the difference between period* t *and the most recent pre-treatment period.*

Copula Stability Assumption (ii): $C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}(0)|D=1} = C_{Y_{t^*-1}(0)-Y_{t^*-2}(0),Y_{t^*-2}(0)|D=1}$

Copula Invariance Assumption (iii): $C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}(0)|D=1} = C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}(0)|D=0}$

All three copula assumptions replace the unknown copula $C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}|D=1}$ with an observed copula and only differ with respect to this choice. Both of the copula stability assumptions replace the unknown copula for the treated group with one from the past. Callaway and Li (2019) provide particular advantages of this type of argument in the context of fixed effects type models, and the copula stability assumption (i) is essentially a direct analogue of theirs generalized to the case with more time periods. However, an important drawback of this approach is that it requires more and more pre-treatment time periods for values of t further away from t^*-1; this is an undesirable feature in many applications including the current one. Thus, for practical purposes we rely on copula stability assumption (ii), which replaces the copula with the copula in the immediately preceding period.

Alternatively, it is possible to extend the copula invariance assumption in Callaway et al. (2018) to identify distributional treatment effects over time. Instead of Assumption 2(ii), we can impose Assumption 2(iii) on the relationship between the unknown copula and the observed copula from the untreated group of firms. Note that the copulas on the left-hand side of Assumption 2(ii) and 2(iii) are the same. The substantive difference is that in Assumption 2(iii), the unknown copula is replaced by the observed copula for the untreated group rather than identifying the QTT based on the assumed relationship between the past copula and

Approach	QTT		QCLO		QCIC	
Year	2005	2006	2005	2006	2005	2006
	-0.14	0.04	-0.14	0.04	-0.10	0.06
	(0.10)	(0.08)	(0.10)	(0.08)	(0.11)	(0.08)

Notes: Standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2003 and 2004 are used as base periods for outcomes in 2005, while 2004 and 2005 are used as base years for outcomes in 2006.

Table 1: ATT, Pre-Trend Tests

the unknown copula in Assumption 2(ii). Appendix B provides an extended discussion of the underlying structural implications of Assumptions 2(i)-2(iii), while we next instead pursue an empirical approach to model selection.

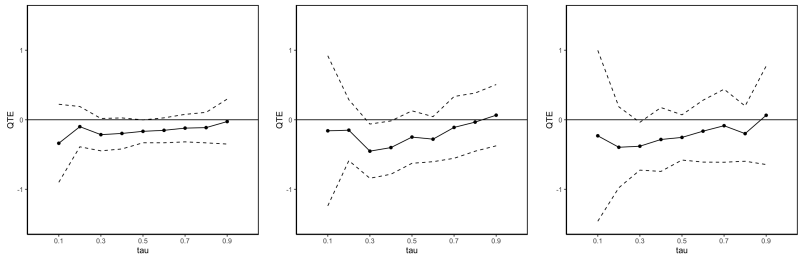
5.1 Distributional Pre-Trends: Using the Data to Help Decide

Treating Assumptions 2(ii) and 2(iii) as *reduced-form* assumptions, we propose a series of distributional pre-tests, which are new to the quantile treatment effect literature, to facilitate model selection. First, using three periods of pre-treatment data from 2003 to 2006, we estimate quantile treatment effects in pre-treatment periods and test whether the QTTs are all equal to 0, which should be the case if our assumptions are valid. We first use 2003 and 2004 as the two required pre-treatment years to identify the copula for the QTT and estimate the impact of being located in a treated province in 2005, the year prior to policy announcement.[21]

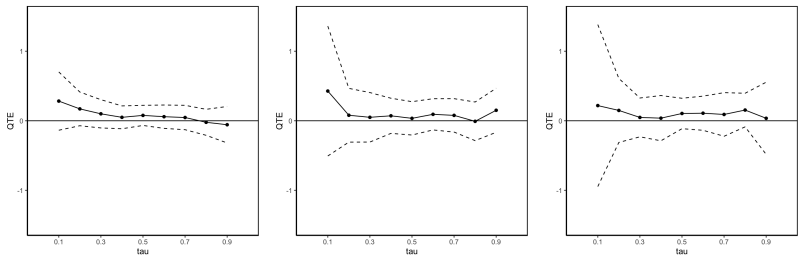
Table 1 reports the ATT for both pre-treatment years, 2005 and 2006, and for extensions of three workhorse distributional estimators. We denote quantile treatment effect on the treated estimators from Callaway and Li (2019) and Callaway et al. (2018) as the QTT and QCLO, respectively; the first employs identification assumption 2(ii), while the second relies on assumption 2(iii). We also compare performance to the well-known changes-in-changes estimator (CIC), posited in Athey and Imbens (2006). Regardless of the estimation approach and outcome variable, there are no significant ATTs in 2005 or 2006. Indeed, all of the estimates are small and not significantly different from zero.

While there is little evidence of meaningful pre-trends on average, our purposes require that

[21] On the one hand, the 2003–2005 period is particularly compelling since China's 11th FYP was only announced in late 2005 and was not ratified until mid-2006, ensuring that 2005 is truly prior to the policy announcement. On the other hand, restricting analysis to the 2003–2005 period significantly reduces the sample size because of the rapid rate of entry and exit among Chinese manufacturing firms in the early 2000s. To address this potential source of bias, we repeat our first pre-trend exercise on the 2004–2006 sample under the additional assumption that firms would not have had sufficient time to adjust to the new policy regime in 2006 given that it was not ratified until August 2006.



(a) QTT, 2005 (b) QCLO, 2005 (c) QCIC, 2005



(d) QTT, 2006 (e) QCLO, 2006 (f) QCIC, 2006

Notes: The figure contains estimates of quantile treatment effects for firms in provinces that significantly reduced SO_2 emissions. Panel (a) considers (log) SO_2 intensity in 2005 as outcome variables using the QTT estimator, while panels (d) repeat the exercise for 2006. The years 2003 and 2004 are used as base periods for outcomes in 2005, while 2004 and 2005 are used as base years for outcomes in 2006. Panels (b)–(c) and panels (e)–(f) present similar figures using the QCLO and QCIC estimators, respectively. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 9: ln(SO_2 Intensity), Pre-Trend Test

there are also no pre-trends across the distribution of outcomes. Figure 9 documents quantile point estimates for SO_2 emissions intensity across years and estimators along with 95 percent pointwise confidence intervals.[22]

With respect to the QTT estimator, there is no statistically significant evidence of pre-existing emissions intensity differences in 2005 or 2006. Figure 9 similarly indicates that Assumption 2(iii) may also be plausible in this context: In panels (b) and (e) we observe an occasional significant treatment effect for the QCLO estimator but document little systematic evidence of meaningful distribution pre-trend effects. Finally, the lack of meaningful pre-trends is also reflected in panels (c) and (f) where the CIC estimator reflects little systematic impact on any part of the distribution in 2005 or 2006, though there are multiple significant treatment effects in three of the four pre-test exercises using the CIC approach.

[22]Point estimates and pointwise standard errors are documented in Table 10 of Appendix C. We focus on pointwise confidence intervals for three reasons: (i) pointwise confidence intervals yield more conservative conclusions in our pre-trend analysis, (ii) we are primarily interested in the individual estimates of each QTT rather than joint significance, (iii) expositional parsimony. Nonetheless, we have checked our results and our broad conclusions, both in the pre-trend analysis and thereafter, are robust to using either approach to computing confidence intervals.

To help choose among estimators, we consider a summary measure of goodness of pre-trend fit for each estimator, F, where

$$F = \alpha_1 \underbrace{\sum_{\tau} (\widehat{QTT}(\tau) - 0)^2}_{\text{Bias}^2} + \alpha_2 \underbrace{\sum_{\tau} se(\widehat{QTT}(\tau))^2}_{\text{Variance}} \tag{4}$$

and where higher values of F are associated with better pre-trend fit. The first component, $\text{Bias}^2 \equiv \sum_{\tau} (\widehat{QTT}(\tau) - 0)^2$, captures the sum of squared deviations from zero across all quantiles and in each pre-trend year. The second component, $\text{Variance} \equiv \sum_{\tau} se(\widehat{QTT}(\tau))^2$, penalizes an estimator for large standard errors (se).[23] Parameters α_1 and α_2 weight the importance of each component; for transparency we set $\alpha_1 = \alpha_2 = 1$

Year	Method	Bias^2	Variance	Total Fit
2005	QTT	0.26	0.20	0.46
	QCLO	0.51	0.28	0.78
	QCIC	0.53	0.51	1.04
2006	QTT	0.14	0.12	0.26
	QCLO	0.24	0.13	0.37
	QCIC	0.13	0.25	0.39

Notes: The above table documents evidence of pre-treatment fit in each pre-treatment year. The first component of pre-treatment fit is the squared distance from zero: $\text{Bias}^2 \equiv \sum_{\tau} QTT(\tau)^2$. The second component of pre-treatment fit is the sum of squared (pointwise) standard errors: $\text{Variance} \equiv \sum_{\tau} se(QTT(\tau))^2$. Total fit is the sum of each individual component.

Table 2: Pre-Trend Fit

Table 2 reports the total value of F and the value of each subcomponent in each pre-trend year. We observe that the Callaway and Li (2019) QTT estimator generally produces the smallest deviations from zero (Bias^2) and the smallest squared error (Variance) and, consequently, has the best fit (total fit, F) in each pre-treatment year. The only slight exception is the Bias^2 in 2006; in this case, the Athey and Imbens (2006) CIC estimator produces estimates that are as close to zero, but again come at the cost of much larger standard errors. Overall, we conclude that the Callaway and Li (2019) QTT estimator has the strongest pre-treatment performance and, as such, employ it as our benchmark estimator.

[23] Strictly speaking, $\text{Variance} \equiv \sum_{\tau} se(\widehat{QTT}(\tau))^2$ does not compute the variance of the sum, as it omits covariances across τ.

	Year				
	2006	2007	2008	2009	2010
ATT	0.04	0.06	-0.12	0.12	0.15
	(0.08)	(0.11)	(0.13)	(0.14)	(0.15)

Notes: Standard errors are in parentheses. * indicates statistical significance at the 5 percent level.

Table 3: Log SO_2 Emissions Intensity, 2006–2010, ATT

6 Empirical Results

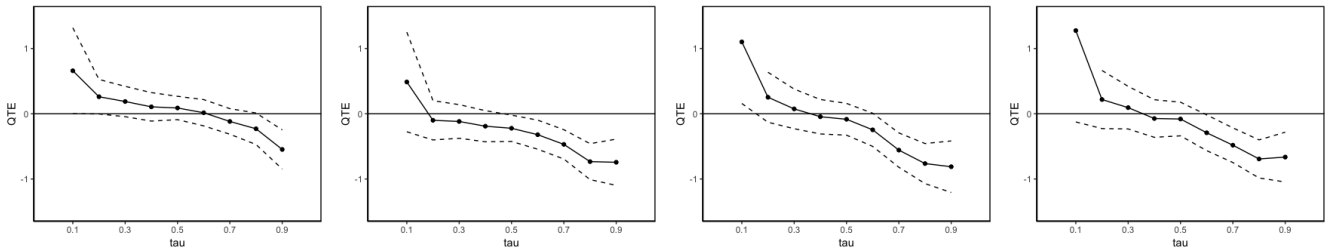
This section documents the impact of China's 11th FYP on firm-level SO_2 emissions intensity. For our benchmark results, we focus on firms located in provinces that form the policy border, as illustrated in Figure 4b. Unless otherwise noted, the 2004–2005 sample is consistently used for pre-treatment years, since these years are clearly prior to the 2006 announcement, and we vary the outcome year over the 2007–2010 period. We estimate both average responses (the average treatment effect) and the distribution of firm-level responses as they evolve over time in response to policy change.

Table 3 reports the average treatment effect on the treated firms in each year. It is immediately clear that the average treatment effects are never estimated to be significantly different from zero; indeed, the average treatment effect on treated firms is estimated to be positive in three of the four years after 2006. In this sense, the average program impact hardly suggests meaningful success in reducing firm-level emissions performance.

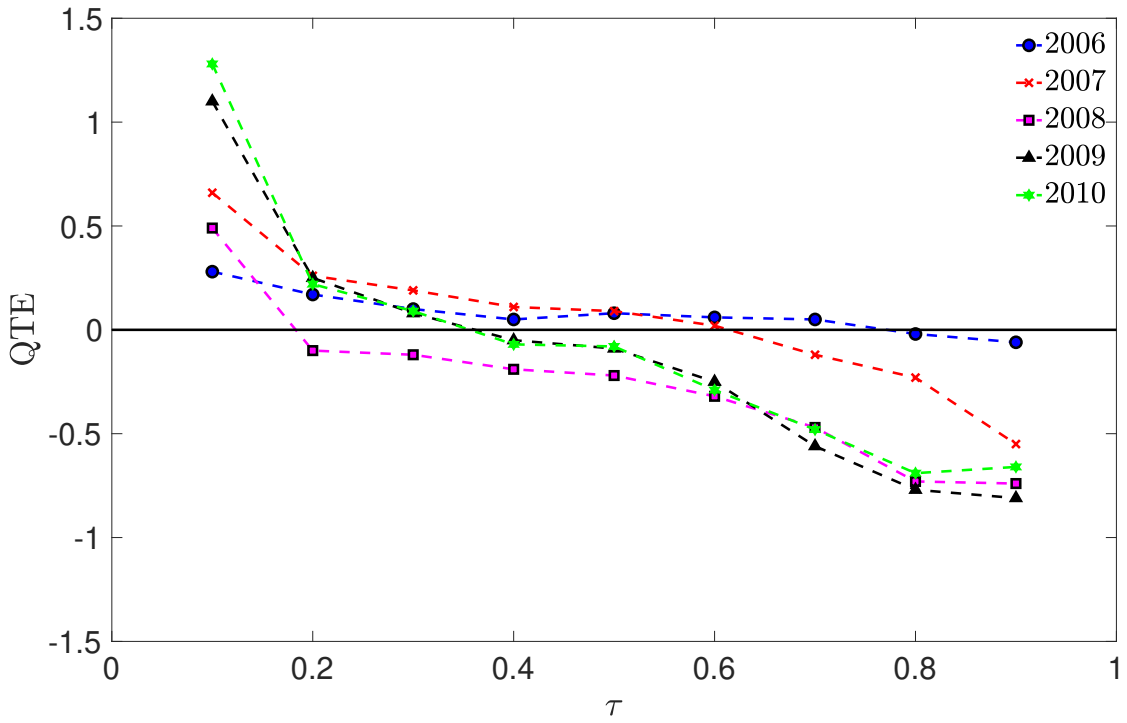
In contrast, the quantile treatment effects reveal that the policy has some bite in all years, at least for some firms. Panels (a)–(d) of Figure 10 plot the QTT estimates in each year for the distribution of Chinese producers in our benchmark sample along with bootstrapped 95 percent confidence intervals, while panel (e) collects the QTT estimates in each year (omitting the confidence intervals for expositional purposes).[24] In 2007 the uppermost decile suggests a small, but significant, decline in emissions intensity (panel (a)). By 2008 the upper four deciles of the emissions-intensity distribution significantly reduce their emissions per unit of output in response to policy change (panel (b)). For instance, while the sixth decile of the emissions distribution is estimated to reduce SO_2 pollution by 27 percent in response to the policy change, emissions intensity declined by 52 percent for the top decile. The estimated impact in each decile holds its economic and statistical significance over the remaining sample period. In this sense, it is clear that the policy had a meaningful impact on the nature of Chinese production, at least for the most emissions-intensive firms in the Chinese manufacturing sector.

While Panel (e) of Figure 10 reinforces the gradual decline of emissions intensity among the

[24]Standard errors for individual QTT estimates are reported in the appendix.



(a) 2007 (b) 2008 (c) 2009 (d) 2010



(e) 2006–2010, QTT w/o Confidence Intervals

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets.

Figure 10: Quantile Treatment Effects for Log SO_2 Intensity, by Treatment Year

upper tail of the emissions-intensity distribution, it also provides an intuitive explanation for why the average treatment effects on the treated firms are never estimated to be significantly different from zero: offsetting changes among the lower quantiles. We consistently observe rising emissions intensity among the lower quantiles over the promulgation of China's 11^{th} FYP. Although Table 11 reports that the lowest quantiles are always estimated to large and positive after 2006; moreover the policy-induced growth is estimated to be lowest quantile statistically significant in two out of four post-treatment years. While this seems counterintuitive, it is exactly what we would expect from a policy that incentivizes the flow of resources, including energy, to the cleanest producers. To the extent that China's 11^{th} FYP reduced the variance of the implied emissions regulation across firms, our findings are consistent with a *relative decline* in the regulatory burden among the initially cleanest firms. We investigate the underlying mechanisms of these distributional changes in Section 6.2.

6.1 Robustness

We conduct a series of exercises to characterize how our benchmark estimates vary with the choice of estimator, the definition of treatment, or sample selection. In the interest of parsimony, we briefly summarize our robustness findings and relegate supporting figures and tables, along with an extended discussion of each exercise, to the appendix.

Alternative Estimation Approaches: We reconsider each outcome year using the QCLO (Callaway et al. (2018)) and QCIC (Athey and Imbens (2006)) approaches. As with our benchmark findings, in each year and for each estimator the average treatment effect is never estimated to be significantly different from zero. Likewise, the quantile treatment effects confirm emissions intensity declines among the upper quantiles and increases among the lowest quantiles. However, the quantile treatment effect estimates no longer decline monotonically across the distribution and the standard errors are modestly larger. In this sense, the preferred Callaway and Li (2019) QTT estimator paints the clearest picture of the impact of the 11^{th} FYP across the distribution of Chinese producers. Indeed, the differences across estimation approaches are inline with expectations given the evidence from the pre-trend tests in Section 5.1.

Alternative Definitions of Treatment: We first consider all provinces and divide them into treatment and control locations, as illustrated in Figure 4a. This definition permits a much larger sample size and includes large, emissions-intensive coastal producers. We also consider the definition of treatment that restricts attention to firms located in border cities, as illustrated in Figure 4c. It comes at the cost of substantially shrinking our sample size. While the benchmark exercise included over 2,000 individual firms (1,500 treated firms), this last definition of treatment shrinks our sample to just over 700 firms (450 treated firms).

The estimated QTTs are qualitatively consistent with the benchmark sample: there is no impact in 2006, the largest emissions-intensity declines are initially confined to the upper quantiles, a larger fraction of quantiles experience meaningful emissions reductions in later years, and there are modest, but statistically insignificant, increases in emissions intensity among the lower quartiles (See Appendix Table 13 and Figures 21 and 22). Differences across samples are intuitive: we observe the smallest estimated changes and the smallest confidence intervals when using all provinces.[25] In contrast, the estimated QTTs from the sample of border cities are very close to those from the sample of border provinces, though the confidence intervals are substantially wider, reflecting the smaller sample size.

Balanced Panel: China's industrial sector is broadly characterized by a high degree of firm-level churning. As a consequence, our benchmark sample suffers from significant attrition: we observe nearly 6,426 firms in 2007, but only 3,663 firms in 2010, as many firms drop out over time. To address attrition, we repeat our benchmark exercise on a balanced panel of firms that are present in every year between 2004 and 2010. If anything, the average treatment effects are larger and statistically stronger in the balanced panel despite the smaller sample size (See Appendix Table 9 and Figure 23). Again, differences between the benchmark sample and the balanced panel can be explained by the fact that the balanced panel disproportionately drops small firms, which consequently, are also those most likely to be in the lower tail of our benchmark sample.

Zeros: Roughly 6 percent of firms report zero emissions in a given year. To avoid unnecessarily dropping firms, we adopt the standard practice of adding one to the value of emissions of each firm,[26] which effectively changes the lower bound of the emissions distribution and introduces a small amount of measurement error. To investigate the impact zeros have on our estimates, instead use an inverse hyperbolic sine transformation of the dependent variable. Among upper quantiles, the estimated QTTs are negative, statistically significant and take a very similar magnitude to our benchmark estimates. Among the lower quantiles, the QTTs are again positive, but the coefficients are larger and more precisely estimated than the benchmark case. Indeed, the estimated QTT for the lowest quantile is now always statistically significant after 2006 and consistently grows in magnitude over the 11^{th} FYP (See Appendix Table 15 and Figure 24).

Correlated Pollutants: SO_2 emissions-intensive firms are likely to also emit other pollutants, such as particulate matter (PM). Repeating our benchmark exercise while using PM intensity as the outcome variable, we recover similar policy-driven changes in the emissions distribution

[25] As discussed in the appendix, the differences of estimated magnitude potentially reflect the impact of unobserved heterogeneity such as preferential regulatory treatment or greater free-rider incentives in coastal locations.

[26] The median firm in 2006 emitted over 60,000 kilograms of SO_2 in a calendar year, while firms in the first percentile report emissions of 250 kilograms of SO_2 in a calendar year.

across firms and time (See Appendix Table 16 and Figure 25). In particular, there are sharp declines in the PM intensity distribution among the upper quantiles, sharp increases among the lower quantiles and little change among the middle quantiles. One notable difference: the average treatment effect in 2010 is large, positive and statistically significant. The QTT estimates reveal this is entirely driven by changes in the first decile, suggesting there may be some substitution towards other pollutants.

Finer quantiles: We explore whether a greater number of quantiles would suggest different patterns of heterogeneity across firms by estimating the QTT for 19 quantiles instead of the benchmark 9 quantiles (See Appendix Table 17 and Figures 26). Again, the highest quantiles are estimated to significantly decrease emissions intensity in response to China's FYP, while the lowest quantiles are estimated to increase emissions intensity. Both findings are again statistically significant in each year after 2006. There is again no estimated change in the middle quantiles, broadly confirming the benchmark result across the emissions intensity distribution.

Sectoral Differences: SO_2 emissions intensity naturally varies with energy intensity across industries. Although we account for persistent unobserved heterogeneity in our empirical specification, we do not characterize industrial differences in our benchmark sample. Here, we pull out the three largest industries from our benchmark sample for individual analysis: chemical manufacturing, non-metallic minerals, and metal smelting and rolling. Qualitatively, the QTT estimates in the chemicals manufacturing industry and the metal smelting and rolling industry are similar to those in the benchmark sample. Quantitatively, the estimated emissions intensity declines among the upper tail of the emissions intensity distribution are larger than those in the benchmark sample in both industries, while the rise in emissions intensity among the lower tail in the metal smelting and rolling industry is likewise larger than that estimated in the corresponding benchmark estimates.[27]

In contrast, there are relatively modest changes in emissions intensity across the entire distribution of non-metallic mineral manufacturers, though the 11^{th} FYP is associated with statistically significant emissions intensity declines in the 5^{th}, 6^{th}, and 7^{th} deciles between 2008 and 2010. We investigate whether the qualitative differences can be attributed to industrial clustering, but instead find that the differential performance of non-metallic mineral manufacturers appears to be driven by industrial heterogeneity.[28]

[27] In no case are the QTT estimates statistically significant when evaluated at conventional significance levels, though this is expected as the sample sizes are generally quite small. The sample size in chemicals manufacturing and metal smelting and rolling industries are roughly 11 and 7 percent, respectively, of the benchmark sample.

[28] To investigate the impact of industrial clustering, we measure the fraction of non-metallic mineral producers in the interquartile range during the pre-sample period. If non-metallic mineral producers are disproportionately clustered in (outside of) the middle deciles, we would expect that more (less) than 50 percent of non-metallic mineral producers will be in the pre-sample interquartile range. We find that roughly 50 percent of non-metallic

6.2 Measurement, Mechanisms, and Trade-Offs

How did environmental gains manifest themselves among Chinese producers? There are various plausible mechanisms by which China's 11th FYP may have affected firm behavior. For instance, the decline in emissions intensity may have come at the cost of a decline in output. Alternatively, if scale economies affect emissions intensity, the opposite may be true. Firms might alternatively adjust by switching energy sources, adopting in new production technology, changing their mix of products or investing in emissions abatement.

To investigate these possibilities, we start by constructing two outcome variables that capture a relevant dimension of firm-level environmental efficiency. Coal intensity, measured as coal consumed per unit of output, reflects either underlying changes in the productive efficiency of individual producers or changes in the energy sources employed by Chinese firms. Similarly, emissions inefficiency, computed as emissions per unit of coal, captures the rate at which a firm's coal consumption is eventually transformed into SO_2 emissions.[29]

The first row of Figure 11 establishes that the highest quantile of the coal-to-output ratio experienced a sharp decline over the 11th FYP, indicating efficiency improvements among the least coal-efficient producers.[30] All other quantiles, however, demonstrate almost no change over the same period. In contrast, not only do we observe a sharp fall in the emissions per unit of coal in the uppermost quantiles of the emissions-inefficiency distribution after 2007, but also among all of the top six quantiles. In this sense, our findings suggest that, outside of the top quantile, emissions intensity reductions are not fundamentally changing the relationship between coal and production (e.g. through switching energy sources), but rather are potentially mitigated through *smokestack* abatement.

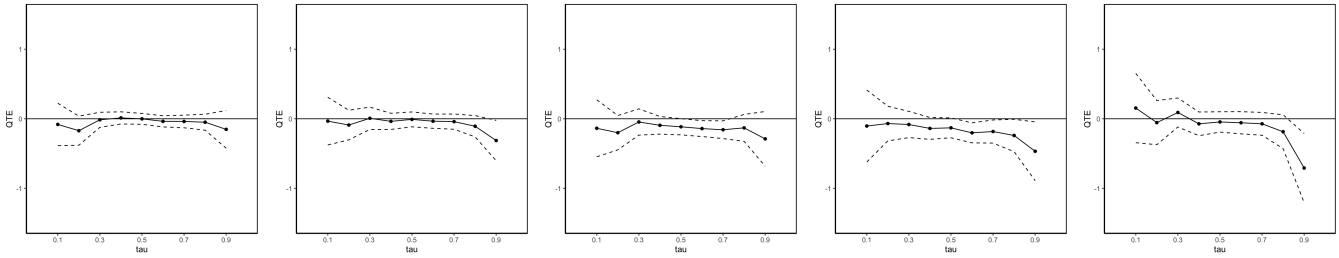
An analogous result holds for the lowest quantiles, where emissions per unit of coal in the lower quantiles rise but coal intensity is largely unchanged. While far from conclusive, the evidence in Figure 11 suggests that cleaner firms took advantage of their initial strong environmental performance to fuel future growth. Again, the results do not point to a large change in the relationship between production and coal consumption, but rather one that suggests changing rates of abatement.

We build on the indirect evidence for *smokestack* abatement by studying measures of energy sourcing, energy consumption and abatement. We start by appealing to data which measures

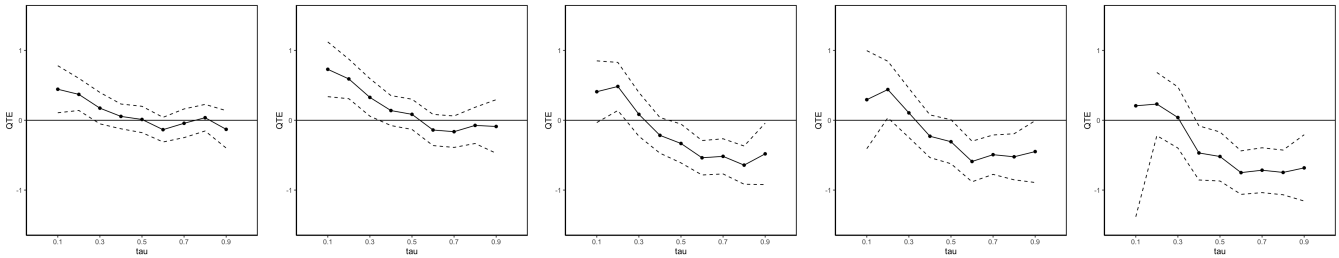
mineral manufacturers percent of producers are found in pre-sample interquartile range. Repeating this exercise for different ranges of the data, we find 80 percent non-metallic mineral producers between the 10th and 90th percentiles and 90 percent of the non-metallic minerals producers between the 5th and 95th percentiles of the benchmark, pre-sample emissions intensity distribution.

[29]Emissions per unit of coal also acts as a check on whether our benchmark results are affected by changes in the distribution of markups, as in Rodrigue et al. (2022a). As documented below, we do not find such evidence.

[30]Average treatment effects and pointwise standard errors for individual quantiles are collected in Appendix Tables 19 and 20. In all cases, there is little evidence of any change on average in any year.



(a) Coal/Output, 2006 (b) Coal/Output, 2007 (c) Coal/Output, 2008 (d) Coal/Output, 2009 (e) Coal/Output, 2010



(f) SO_2/Coal, 2006 (g) SO_2/Coal, 2007 (h) SO_2/Coal, 2008 (i) SO_2/Coal, 2009 (j) SO_2/Coal, 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations. Outcome variables are the following: the logarithm of output per unit of coal (row 1), the logarithm of SO_2 emissions per unit of coal (row 2).

Figure 11: Productivity, Technology, and Emissions

the total amount of energy consumed by each firm, by energy type (coal, diesel fuel, petroleum, etc). We convert the quantity of each fuel consumed to a quantity of embedded SO_2 emissions using conversion factors from the U.S. Energy Information Association (EIA).[31] We repeat our benchmark estimation approach on the imputed emissions intensity from energy consumption in the first row of Figure 12.

To further highlight the differences between measured emissions and implied emissions from energy consumption, the second row of Figure 11 considers the log difference between imputed and measured emissions intensity as an outcome variable. Importantly, this variable also roughly captures the difference between the amount of SO_2 *generated* and that *emitted* per unit of output; that is, it is a proxy for firm-level smokestack abatement. We normalize abatement by total emissions generated from input (energy) consumption and plot the inverse abatement intensity for facilitate comparisons with other mechanisms in the second row of Figure 12.

We compare the indirect measure of abatement to a direct measure of investment in abatement technology. Specifically, the survey records the total emissions abatement capacity for each firm (e.g. total capacity of SO_2 scrubbers). After normalizing total scrubber capacity by total firm-level output, the third row of Figure 12 reports the QTT estimates. Finally, we also

[31] U.S. EIA conversion factors can be found here: https://www.eia.gov/electricity/annual/html/epa_a_01.html

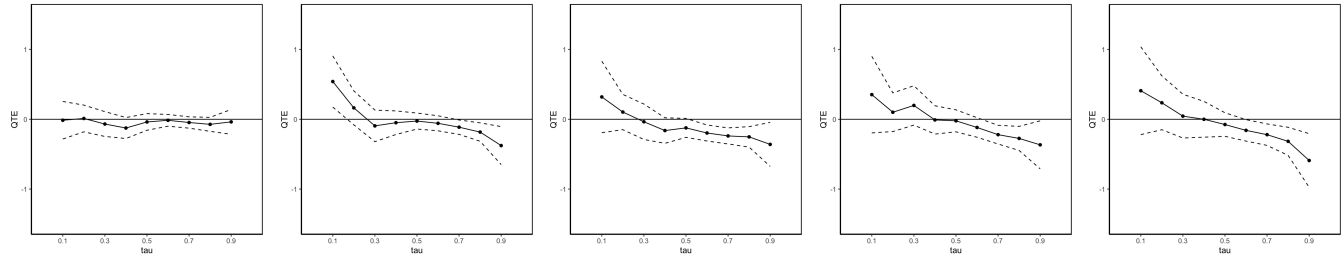
consider an alternative abatement strategy: switching energy sources. In particular, we compute the share of total emissions generated attributable to coal consumption for each firm and year. Again, we plot the inverse log coal share for comparability with other mechanisms in the fourth row of Figure 12.

Input (energy) based measurement of SO_2 emissions intensity qualitatively confirms the patterns documented in our benchmark estimates: declines among the upper quantiles and increases among the lower quantiles. Yet, there are a few differences which merit comment. The absolute magnitude of the QTT estimates are modestly smaller than our benchmark estimates in both tails, suggesting that firms actively took action to further mediate emissions. The confidence intervals are likewise slightly wider than those in the benchmark estimates, reinforcing the notion that differential firm abatement played an complementary role in the differential evolution firm-level emissions intensity across the distribution.

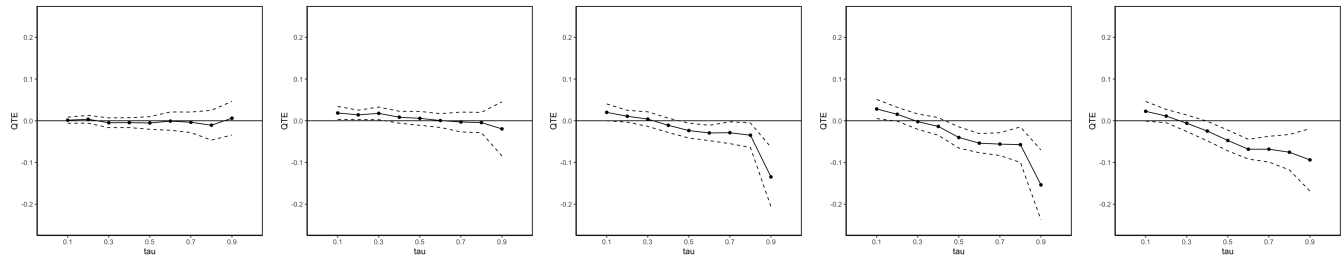
The second and third rows of Figure 12 confirm the statistically significant impact of differential abatement at both tails of the distribution. In the second row we observe that firms which initially had the smallest abatement relative to total emissions (the upper tail) experienced the largest declines: abatement rose disproportionately relative to input generated emissions by as much as 15 percent in the top quantile. At the other end of the distribution, firms with largest pre-existing abatement in treatment provinces witnessed disproportionate, statistically significant increases in input generated emissions relative to abatement after 2007 by 2-3 percent. A similar pattern emerges in the third row where we directly measure smokestack abatement with scrubber capacity per unit of output. Firms with the least scrubber capacity initially (the lower tail) experience disproportionate increases in treated provinces, while those with the largest capacity per unit of output experience declines, each of which is statistically and economically significant after 2007. In contrast, in no year and for no part of the distribution is there any evidence of meaningful change in energy sourcing; this is not surprising: coal is by far the most common source of energy for most firms in our data.

Careful comparison of figures 11 and 12 suggest each abatement mechanism plays different roles for different firms. On the one hand, firms in highest coal-intensity quantile experience sharp declines in coal-intensity after 2007, while coal-intensity declines among other upper quantiles are much more modest. On the other hand, the increase in abatement intensity is much broader based suggesting that the 11^{th} FYP induced the dirtiest firms to both adopt cleaner technology and increase smokestack abatement, while most other firms in the upper quantiles raised abatement investment alone.

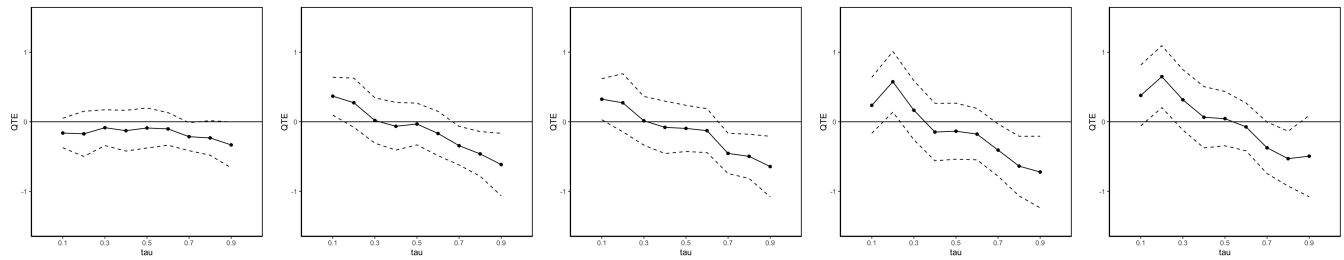
At the other end of the coal-intensity distribution there is little change over the 11^{th} FYP, while there is a consistent, relative decline in abatement investment among the firms which were initially cleanest. Conditional on firm fixed effects, clean firms in treated provinces did not



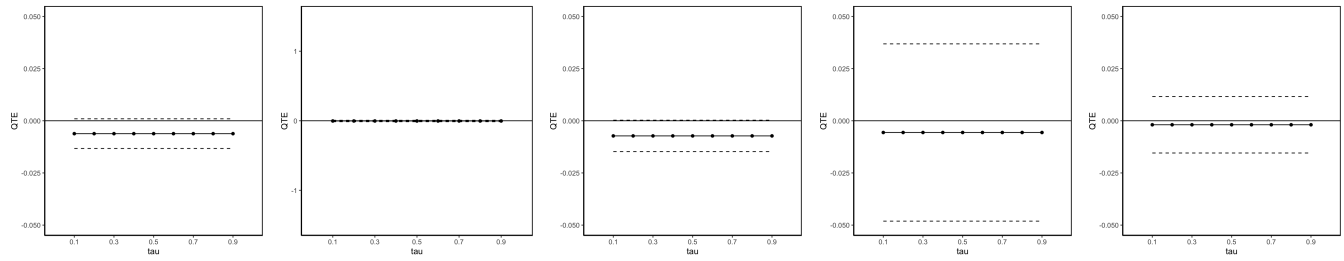
(a) Inp. SO_2 Int., 2006 (b) Inp. SO_2 Int., 2007 (c) Inp. SO_2 Int., 2008 (d) Inp. SO_2 Int., 2009 (e) Inp. SO_2 Int., 2010



(f) Inv. Abt-Int., 2006 (g) Inv. Abt-Int., 2007 (h) Inv. Abt-Int., 2008 (i) Inv. Abt-Int., 2009 (j) Inv. Abt-Int., 2010



(k) Cap./Output, 2006 (l) Cap./Output, 2007 (m) Cap./Output, 2008 (n) Cap./Output, 2009 (o) Cap./Output, 2010



(p) Ener. Switch, 2006 (q) Ener. Switch, 2007 (r) Ener. Switch, 2008 (s) Ener. Switch, 2009 (t) Ener. Switch, 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations. Outcome variables are the following: the logarithm of SO_2 intensity measured by input consumption (row 1), log inverse abatement intensity (row 2), log scrubber capacity per unit of output (row 3), log share of coal in total emissions (row 4).

Figure 12: Inputs, Abatement and Energy Switching

continue to invest in abatement at the same rate as those in control provinces. In this sense, we find that the 11$^{\text{th}}$ FYP not only encouraged the dirtiest firms in treated provinces to differentially mitigate emissions but also had the effect of increasing emissions intensity among the cleanest. This result is surprising, but remarkably robust: it is not only true in the benchmark data, but is also documented with alternative estimation approaches, treatment definitions, measures of emissions intensity, while consistently appearing when we employ finer quantiles or consider particulate matter emissions instead. It is likewise consistent with evidence for technological and abatement investment. Thus, evidence overwhelming suggests that there was some rise in emissions intensity among firms which were initially the cleanest.

There are multiple possible explanations for relative increases in emissions intensity after reform. One possibility is that a decline in coal intensity by large, emissions intensive firms through the adoption of cleaner technology depressed the (local, relative) price of coal sufficiently to induce an expansion of coal use by initially clean firms. Yet, we do not find evidence of energy switching even among firms that were initially the least coal intensive. A second is changes in relative market power, as highlighted in Buchanan (1969), Leslie (2018), and Preonas (2023), among others. Yet, most manufacturing markets are characterized by modest rates of industrial concentration and the Chinese manufacturing landscape has featured signficant entry of new firms since the year 2000, suggesting that market power was likely limited in many manufacturing industries Chor et al. (2021).

Alternatively, cleaner firms may have enjoyed stronger local comparative advantage after policy reform, inducing disproportionate growth rates. To examine this possibility we consider a series of descriptive growth regressions. In particular, we first classify firms into their initial (2004) SO_2 intensity quantiles, τ_{04}, with a series of indicator variables, $D_i^{\tau_{04}}$. We then regress (annualized) cumulative output growth over 2006-2010,

$$\Delta \ln(Output_{it}) = \frac{\ln(Output_{it} - \ln(Output_{i,06})}{t-6} \text{ for } t = 7, 8, 9, 10$$

on initial SO_2 intensity quantiles and their interactions with treatment province status, D_i:

$$\Delta \ln(Output_{it}) = \sum_{\tau_{04}=0.1}^{0.9} \gamma_{0,\tau}(1 - D_i) \times D_i^{\tau_{04}} + \sum_{\tau_{04}=0.1}^{0.9} \gamma_{1,\tau} D_i \times D_i^{\tau_{04}} + X_i'\gamma_X + \eta_t + \epsilon_i \qquad (5)$$

where X_i includes a control variable for initial (2004) firm productivity, proxied as output per unit of coal consumption, η_t is year fixed effects capturing common trends, and ϵ_i is the error term. While equation (5) does not provide causal evidence in favor or against the comparative advantage hypothesis, it does provide descriptive evidence as to whether the least emissions intensive firms in treatment provinces grew disproportionately quickly, consistent with improved

comparative advantage during the 11th FYP.

Dependent Variable: Annualized 2006-2010 Output Growth

	2004 SO_2 Intensity τ								
	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
Control Provinces	0.20*	0.21*	0.08*	0.03	0.06	-0.02	-0.13	-0.17	-0.24*
	(0.05)	(0.05)	(0.04)	(0.06)	(0.08)	(0.06)	(0.12)	(0.13)	(0.09)
Treatment Provinces	0.44*	0.19*	0.15*	0.11*	0.13*	-0.01	-0.01	0.03	-0.09*
	(0.06)	(0.04)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.06)	(0.05)
Treatment - Control	0.25*	-0.02	0.07	0.08	0.07	0.01	0.12	0.20	0.15
	(0.06)	(0.05)	(0.04)	(0.06)	(0.07)	(0.06)	(0.12)	(0.13)	(0.09)

Notes: Estimation sample: A balanced panel of 2,036 observations between 2004-2010. Co-variates: In addition to initial (2004) SO_2 intensity quantile dummies, uear fixed effects and firm-level initial productivity, proxied by the output to coal consumption ratio in 2004, included as additional co-variates. R^2 is equal to 0.2. Robust standard errors are in parentheses. * indicates statistical significance at the 5 percent level.

Table 4: Descriptive Growth Regressions

Table 4 documents evidence consistent with this interpretation. In both control and treatment provinces, firms that were the least SO_2 intensive grew disproportionately rapidly during the 11th FYP, while the most SO_2 intensive firms grew the slowest in both locations. However, the third row of Table 4 indicates that clean firms in treatment provinces experienced disproportionately large gains not only relative to other firms in the same location, but also relative to similarly clean firms in control provinces. In this sense, strong output growth, coupled with small investments in new abatement, plausibly contributed to disproportionate rises in emissions intensity among initially clean producers in response to the 11th FYP.

6.2.1 Alternative Mechanisms

We test a series of additional mechanisms through which firms may mediate emissions intensity: offshoring production (emissions leakage), product reallocation, and firm displacement. In each case, we find little evidence of a systematic impact of China's 11th FYP either on average emissions intensity or emissions intensity for any quantile. We omit further discussion and refer the interested reader to the Appendices A.5.1-A.5.3 for an extended discussion.

As discussed above, another possibility is the reallocation of output and emissions *across* plants but *within* the same firm. For example, suppose a firm has a headquarter plant located in a treatment province but associated plants in the same firm conglomerate located in control provinces. A natural consequence of the 11th FYP is that the firm might reallocate output and emissions from the headquarter plant to the control plants. However, in our data, all of the emissions and output from all plants would only be assigned to the headquarter plant

located in the treatment province. To the degree that that the distribution of plants across geographic crosses treatment and control assignment, we are naturally concerned that this may be a potential source of estimation bias. Indeed, Chen et al. (2021) suggest that regulatory changes associated with the 11th FYP did induce within-firm reallocation of output and emissions towards less regulated plants among Chinese conglomerates.

To address this challenge we take a two-pronged approach. First, we openly admit that the data does not allow us to distinguish plants within Chinese conglomerates, as in Chen et al. (2021). However, Chen et al. (2021) also robustly show that the change in policy had no significant impact on the energy efficiency (energy per unit of output) in either headquarter plants in highly regulated locations or affiliated plants in less regulated locations. As such, to the degree that the policy change induced large, dirty conglomerates chose to reallocate output and emissions to less regulated affiliates located in control provinces, our estimates of emissions declines at the upper tail of the emissions intensity distribution are decidedly conservative.

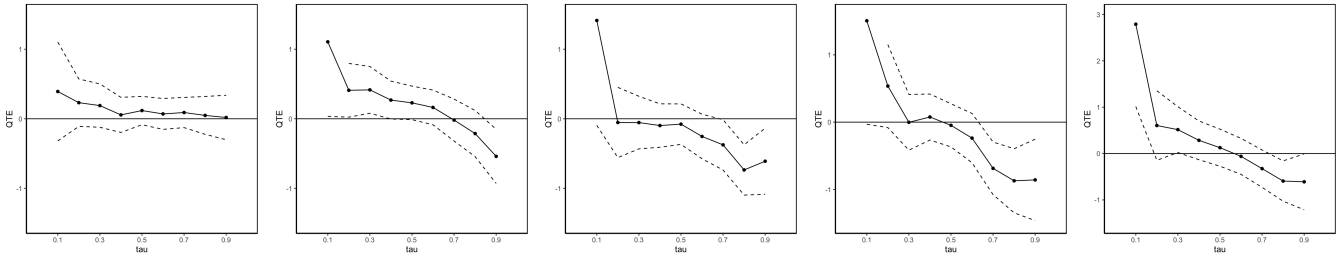
Second, Chen et al. (2021) further document that Chinese conglomerates are overwhelmingly likely to be large firms. We accordingly consider an alternative that restricts attention to small and medium sized firms since these are less likely to have multiple plants in different locations.[32] Figure 13 documents that the results from a sample of firms likely to only have a single plant mirror those from the benchmark sample: statistically significant increases in emissions intensity among the lowest quantiles, sharp declines among the highest quantiles and little change in the middle quantiles, suggesting that the key distributional outcome carries over across firm size (See Appendix Table 27). Quantitatively, the estimated declines about the upper quantiles are very close to those reported in the benchmark sample; the increases among the lowest quantiles are somewhat larger, although this also reflects increasingly large standard errors. Nonetheless, the novel patterns documented throughout this paper maintain their robustness even with the exclusion of China's largest producers. Accordingly, we next turn our attention to quantifying the aggregate productivity gains from China's 11th FYP.

7 Policy Induced Productivity Gains

Section 6 documents that China's 11th FYP contributed to the reshaping of the emission intensity distribution, but does not reveal whether these changes are meaningful in the aggregate. Incorporating the distributional QTT estimates into Section 4's aggregate misallocation framework, we quantify the aggregate TFP gains from the policy-induced changes in the emissions intensity distribution.

We first link treatment parameters to aggregate efficiency gains following Rodrigue et al.

[32]For each 2-digit industry we define a small and medium firm as a firm with below median revenues in 2005.



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for small firms in provinces that significantly increased SO_2 emissions reduction targets. A firm is defined as small if its revenues in 2004 were below the median revenues in its industry. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 13: Log SO_2 Intensity, Small Firms Only

(2024); an important difference in this context is that treatment effects vary over the distribution of heterogeneous producers. Specifically:

1. Employing the benchmark estimates from Tables 10 and 11, we recover the year-specific change in emissions intensity attributable to China's 11$^{\text{th}}$ FYP. For firm i, the estimated change in firm-level emissions intensity due to China's 11$^{\text{th}}$ FYP is computed as $\Delta EI_{it} = \theta(\tau)\ln(\text{Intensity}_{i,05}(\tau))$, where θ_{it} is the estimated QTT for quantile τ.

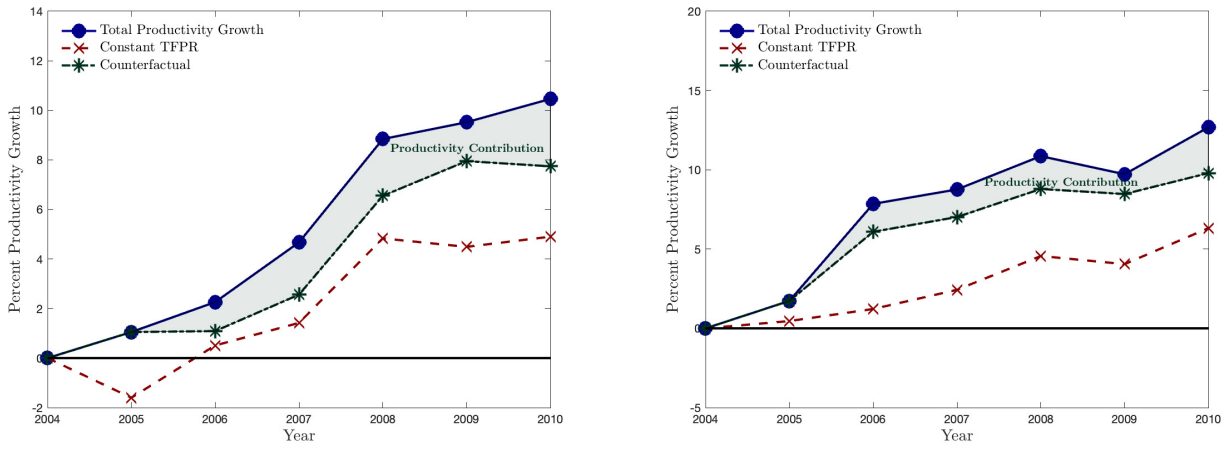
2. We add the estimated change in emissions intensity due to China's 11$^{\text{th}}$ FYP in each year back to the observed series. For firm i and year t with observed emissions intensity $\ln(\text{Intensity})_{it}$, counterfactual emissions intensity is

$$\widehat{\text{Intensity}}_{it} \equiv \frac{\widehat{E_{it}}}{R_{it}} = \exp\{\ln(\text{Intensity}_{it}) - \Delta EI_{it}\}.$$

   Note that if ΔEI_{it} is negative, we add back policy-induced emissions reductions.

3. Using the definition of $MRPE$ and equation (2), we recover the counterfactual distribution of TFPR.

The careful reader will notice that although counterfactual aggregate productivity depends on the joint distribution of firm productivity (A_{is}) and emissions intensity, the above process only returns the marginal distribution of emissions intensity. Although we recover the counterfactual marginal distribution of emissions intensity, Appendix A.8 shows that it is possible to perform standard aggregation procedures as long as the dependence (copula) between emissions intensity and firm TFP (A_{is}) remains the same. Natural intuition for this condition is that firms that have relatively high productivity under the policy would have had relatively high productivity



(a) Treatment Defn 2, Border Provinces

(b) Treatment Defn 1, All Provinces

Notes: The solid blue line with dots plots the observed path of the aggregate TFP, while the red line with x's captures the counterfactual path of TFP if $MRPE_{it}$ and, as a consequence, $TFPR_{it}$ are held constant at their 2004 values. The gap between the green dotted line and the blue solid line outlines the contribution from the implementation of the 11$^{\text{th}}$ FYP. Panel (a) uses the sample from treatment definition 2, while panel (b) uses the sample from treatment definition 1.

Figure 14: Policy-Driven Productivity Growth

absent the policy. This weaker condition allows us to quantify aggregate productivity gains implied by the estimates from the QTT estimator.

The aggregate manufacturing productivity consequences of China's 11$^{\text{th}}$ FYP are illustrated in Figure 14. As in Section 4, the solid blue line with dots plots the observed path of the aggregate TFP, while the red line with x's captures the counterfactual path of TFP if the $MRPE_{it}$ are held constant at their 2004 values. The gap between the observed path of TFP (the blue line) and green line with asterisks denotes the policy contribution of China's 11$^{\text{th}}$ FYP.

Panel (a) reexamines the benchmark sample (Treatment Definition 2). Consistent with our causal estimates, the impact of policy reform in the benchmark sample gradually increases up to 2008, after which the implementation of the 11$^{\text{th}}$ FYP is estimated to have increased aggregate manufacturing productivity growth by 2.1–2.7 percentage points alone. While these magnitudes may appear modest, the policy change accounts for 49 percent of the implied productivity growth driven by improved resource allocation.

Panel (b) considers aggregate productivity gains across all provinces (Treatment Definition 1).[33] In this case, it is potentially concerning that bias in the QTT estimates might lead to misleading aggregate productivity conclusions. Appendix A.9 documents that even if our recovered estimates are biased at every quantile, the aggregate productivity findings will be *robust* to this form of estimation bias as long as we can reasonably estimate the difference

[33]In panel (b) we use QTT estimates from Table 13 (Treatment Definition 1) in step 2 of the counterfactual analysis.

between control and treatment provinces across the distribution of heterogeneous producers. While surprising, the intuition for this result is surprisingly simple. The *level* of industrial productivity can be characterized as an aggregation of firm-level *relative* emissions intensities (and revenue based $TFPR$) across locations. As long as the QTT estimates are sufficiently flexible across the distribution of firms (e.g. we estimate a sufficiently large number of quantiles), bias in the aggregate productivity conclusions will be mitigated even if the QTT estimates do not yield a strictly causal interpretation.[34] Employing the full sample of all provinces in panel (b) of Figure 14, we find that the 11$^{\text{th}}$ FYP accounted for 1.7-2.9 percentage points of additional productivity growth, 45 percent of the implied aggregate productivity growth from improved resource allocation.

We further characterize the nature and efficiency of China's 11$^{\text{th}}$ FYP through a series of counterfactual experiments. In each case, we use the estimates of China's 11$^{\text{th}}$ FYP to counterfactually predict emissions intensity and the prevailing emissions taxes, T_{is}, in absence of the policy reform. We then consider three alternative regulatory reforms, each of which is disciplined by achieving the same decline in aggregate emissions. In the first case, we increase the implied emissions tax by the same absolute magnitude, $\hat{T}_{is} = \kappa_1 + T_{is}$, $\kappa_1 > 0$. Because all firms suffer the same percentage point increase in emissions taxation, the variance of TFPR is unchanged.[35] In this sense, the first policy experiment approximates a setting without reallocation gains relative to the 11$^{\text{th}}$ FYP, allowing us characterize the degree to which allocative gains mitigated output losses.

The second policy reform instead uniformly increases the implied emissions tax for each firm by a constant percentage, $\hat{T}_{is} = (1+\kappa_2)T_{is}$, $\kappa_2 > 0$, until predicted aggregate emissions are equal to those achieved under the 11$^{\text{th}}$ FYP. This experiment achieves the same aggregate decline in emissions, but *increases* the variance of TFPR across firms, and consequently reduces sectoral efficiency and exacerbates the emissions-output tradeoff. The third policy reform considers a uniform emissions tax across all Chinese manufacturers, $T_{is} = \bar{T}$. In this case, all firms in the same industry will have the same emissions intensity and, in our setting, the distribution of TFPR will be as narrow as possible, achieving the greatest reallocation gains.

Table 5 documents the median changes in emissions taxation, along with the aggregate output and productivity consequences of each policy option relative to China's 11$^{\text{th}}$ FYP.[36] The first two experiments suggest that modest increases in the emissions taxation lead to significantly

[34] We note that this result not only supports the counterfactual findings from the full sample of provinces where estimation bias is more likely, but also those from the benchmark exercise.

[35] A longer discussion of the impact of tax change on the distribution of emissions intensity can be found in Appendix A.10.

[36] To aggregate production across industries we apply preferences from Hsieh and Klenow (2009) so that $Y_t = \prod_{s=1}^{S}(TFP_{st}E_{st}^{\alpha_s}V_{st}^{1-\alpha_s})^{\gamma_s}$ where $E_{st} = \sum_{i=1}^{N_s} E_{ist}$, $V_{st} = \sum_{i=1}^{N_s} V_{ist}$, and $\gamma_s = \frac{P_s Y_s}{PY}$ is the revenue share of sector s.

	Percentage Change Relative to China's 11th FYP			
	Emissions Tax	Aggregate Emissions	Aggregate Output	Aggregate Productivity
Experiment 1: $\hat{T}_{is} = \kappa_1 + T_{is}$	3.7[a]	0	-10.24	-1.82
Experiment 2: $\hat{T}_{is} = (1 + \kappa_2)T_{is}$	12.6	0	-11.19	-2.97
Experiment 3: $\hat{T}_{is} = \bar{T}$	6.5[a]	0	4.73	3.51

Notes: The columns above document the percentage change in aggregate emissions, emissions taxation, aggregate output and aggregate TFP for three alternative policies relative to China's 11th FYP. (a) The percent change in emission taxation varies with the firm's initial level of emissions taxation. We report the percent change relative to the median baseline emissions tax.

Table 5: Policy Reform: Aggregate Emissions, Output and Productivity

more costly output-emissions trade-offs. In the first experiment where there are no reallocation gains, a 3.7 percent increase in emissions taxation is needed to achieve the same decline in aggregate emissions. However, aggregate output and productivity are 10.2 and 1.8 percent smaller than that achieved through the 11th FYP. Proportional increases in emissions taxation in the second experiment exacerbate these tradeoffs, with output falling by 11 percent and aggregate TFP by nearly 3 percent. Relative to these straightforward changes to baseline policy, the 11th FYP significantly mitigated the costs of reducing aggregate emissions through reallocation gains.

That said, through the lens of a workhorse model, a policy which would have equalized the price of emissions across heterogeneous producers would have led to aggregate output and productivity growth which were 4.7 and 3.5 percent greater, respectively. Although the 11th FYP was not as costly as some alternatives, regulatory variation across firms maintained an output-emissions trade-off above the productivity maximizing level.

8 Conclusion

China's 11th FYP outlined some of the most aggressive SO_2 reduction targets in Chinese history. This paper quantifies the impact that these targets had on aggregate and firm-level environmental performance. Extending distributional estimators to characterize the heterogeneous impact of China's 11th FYP across firms and time we find evidence in support of significant emissions intensity declines in the upper quantiles of the emissions distribution, modest evidence of increases in the lower quantiles of the emissions distribution, and no change in the middle quantiles. The reductions in firm-level emissions intensity are consistent with greater abatement and, among the dirtiest producers, adoption of new production technology. Rising emissions intensity in the lower quantiles, in contrast, are consistent with smaller abatement investments among the initially cleanest and disproportionate growth over the 2006-2010 period.

In aggregate, we find that the 11th FYP significantly mitigated the output-emissions trade-

off through improved resource allocation, but still fell short of uniform emissions tax. Relative to an equivalent emissions policy that did not improve resource allocation, China's 11^{th} FYP increased aggregate manufacturing productivity and output by as much as 3 and 10 percent without increasing aggregate emissions. In contrast, a uniform firm-level emissions tax achieving the same emissions gains is predicted to cause TFP to grow by an additional 3.5 percent through even greater reallocation gains.

References

Allcott, H. and M. Greenstone (2024, January). Measuring the welfare effects of residential energy efficiency programs. Working Paper 23386, National Bureau of Economic Research.

Athey, S. and G. Imbens (2006). Identification and inference in nonlinear difference-in-differences models. Econometrica 74(2), 431–497.

Bajari, P., J. C. Fruehwirth, K. i. Kim, and C. Timmins (2012). A rational expectations approach to hedonic price regression with time-varying unobserved product attributes: The price of pollution. American Economic Review 102(5), 1898–1926.

Barrows, G. and H. Ollivier (2018). Cleaner firms or cleaner products? How product mix shapes emission intensity from manufacturing. Journal of Environmental Economics and Management 88, 134–158.

Berman, E. and L. T. M. Bui (2001). Environmental regulation and productivity: Evidence from oil refineries. Review of Economics and Statistics 83(3), 498–510.

Bitler, M. P., J. B. Gelbach, and H. W. Hoynes (2017, 10). Can Variation in Subgroups' Average Treatment Effects Explain Treatment Effect Heterogeneity? Evidence from a Social Experiment. The Review of Economics and Statistics 99(4), 683–697.

Bombardini, M. and B. Li (2020). Trade, pollution and mortality in China. Journal of International Economics 125, 103321.

Brandt, L., J. V. Biesebroeck, and Y. Zhang (2012). Creative accounting or creative destruction? Firm-level productivity growth in Chinese manufacturing. Journal of Development Economics 97(2), 339–351.

Broner, F., P. Bustos, and V. M. Carvalho (2012). Sources of comparative advantage in polluting industries. Working Paper 18337, National Bureau of Economic Research.

Brunel, C. and A. Levinson (2016). Measuring the stringency of environmental regulation. Review of Environmental Economics and Policy 10(1), 47–67.

Buchanan, J. M. (1969). External diseconomies, corrective taxes, and market structure. The American Economic Review 59(1), 174–177.

Callaway, B. and T. Li (2019). Quantile treatment effects in difference in differences models with panel data. Quantitative Economics 10(4), 1579–1618.

Callaway, B., T. Li, and T. Oka (2018). Quantile treatment effects in difference in differences models under dependence restrictions and with only two time periods. Journal of Econometrics 206(2), 395–413.

Cao, J., M. S. Ho, and R. Garbaccio (2009). China's 11th five-year plan and the environment: Reducing so2 emissions. Review of Environmental Economics and Policy 3(2), 231–250.

Cao, J., M. S. Ho, and D. W. Jorgenson (2009). The local and global benefits of green tax policies in china. Review of Environmental Economics and Policy 3(2), 189–208.

Cao, J., M. S. Ho, and D. W. Jorgenson (2013). The Economics of Environmental Policies in China, pp. 329–372. Cambridge, MA: MIT Press.

Cao, J. and V. J. Karplus (2014). Firm-level determinants of energy and carbon intensity in china. Energy Policy 75, 167–178.

Carneiro, P., K. T. Hansen, and J. J. Heckman (2003). 2001 Lawrence R. Klein Lecture: Estimating distributions of treatment effects with an application to the returns to schooling and measurement of the effects of uncertainty on college choice. International Economic Review 44(2), 631–422.

Chen, Q., Z. Chen, Z. Liu, J. C. Suárez Serrato, and D. Y. Xu (2021, May). Regulating conglomerates: Evidence from an energy conservation program in China. Working Paper 29066, National Bureau of Economic Research.

Cherniwchan, J. and N. Najjar (2022). Do environmental regulations affect the decision to export? American Economic Journal: Economic Policy 14(2), 125–160.

China Power Team (2016). How is China's energy footprint changing? Technical report, China Power.

Chor, D., K. Manova, and Z. Yu (2021). Growing like China: Firm performance and global production line position. Journal of International Economics 130, 103445. NBER International Seminar on Macroeconomics 2020.

Copeland, B. R. and M. S. Taylor (2003). Trade and the Environment: Theory and Evidence. Princeton University Press.

Deng, Y., Y. Wu, and H. Xu (2020). Political connections and firm pollution behaviour: An empirical study. Environmental and Resource Economics 75(4), 867–898.

Ebenstein, A., M. Fan, M. Greenstone, G. He, P. Yin, and M. Zhou (2015). Growth, pollution, and life expectancy: China from 1991-2012. American Economic Review 105, 226–231.

Ederington, J. and J. Minier (2003). Is environmental policy a secondary trade barrier? An empirical analysis. Canadian Journal of Economics/Revue canadienne d'économique 36(1), 137–154.

Fan, S., R. Kanbur, and X. Zhang (2011, 01). China's regional disparities: experience and policy. Review of Development Finance 1(1), 47–56.

Forslid, R., T. Okubo, and K. H. Ulltveit-Moe (2018). Why are firms that export cleaner? International trade, abatement and environmental emissions. Journal of Environmental Economics and Management 91, 166–183.

Foster, L., J. Haltiwanger, and C. Syverson (2008, March). Reallocation, firm turnover, and efficiency: Selection on productivity or profitability? American Economic Review 98(1), 394–425.

Global Modeling and Assimilation Office (GMAO) (2015). Merra-2 tavgu_2d_aer_nx: 2d, diurnal, time-averaged, single-level, assimilation, aerosol diagnostics v5.12.4. Goddard earth sciences data and information services center (ges disc), Greenbelt, MD, USA.

Gopinath, G., c. Kalemli-Özcan, L. Karabarbounis, and C. Villegas-Sanchez (2017). Captial allocation and productivity in South Europe. Quarterly Journal of Economics 132(4), 1915–1967.

Greenstone, M. (2002). The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufactures. Journal of Political Economy 110(6), 1175–1219.

Greenstone, M., J. A. List, and C. Syverson (2012, September). The effects of environmental regulation on the competitiveness of u.s. manufacturing. Working Paper 18392, National Bureau of Economic Research.

He, C., Y. Yan, and D. Rigby (2018). Regional industrial evolution in China. Papers in Regional Science 97(2), 173–198.

He, G., S. Wang, and B. Zhang (2020, 06). Watering down environmental regulation in China. The Quarterly Journal of Economics 135(4), 2135–2185.

He, L.-Y. and G. Huang (2022). Are China's trade interests overestimated? Evidence from firms' importing behavior and pollution emissions. China Economic Review 71, 101738.

Ho, C.-Y. and D. Li (2008). Rising regional inequality in China: Policy regimes and structural changes*. Papers in Regional Science 87(2), 245–259.

Hsieh, C.-T. and P. J. Klenow (2009, November). Misallocation and manufacturing TFP in china and india. The Quarterly Journal of Economics 124(4), 1403–1448.

Jia, R. (2012). Pollution for promotion. Working paper, UC Sandiego.

Kahn, M. E. and S. Zheng (2016). Blue Skies over Beijing: Economic Growth and the Environment in China (1 ed.). Number 10701 in Economics Books. Princeton University Press.

Karplus, V. J. and X. Zhang (2017). Incentivizing firm compliance with China's national emissions trading system. Economics of Energy and Environmental Policy 6(2), 73–86.

Kishimoto, P. N., V. J. Karplus, M. Zhong, E. Saikawa, X. Zhang, and X. Zhang (2017). Consumption-based adjustment of emissions-intensity targets: An economic analysis for China's provinces. Transportation Research Part D: Transport and Environment 54, 30–49.

Leslie, G. (2018). Tax induced emissions? Estimating short-run emission impacts from carbon taxation under different market structures. Journal of Public Economics 167, 220–239.

Levinson, E. and M. S. Taylor (2008). Unmasking the pollution haven effect. International Economic Review 49(1), 223–254.

Lin, J., C. Zhou, L. Chen, G. Huang, J.-F. Lamarque, J. Nie, J. Yang, K. Hu, P. Liu, J. Wang, Y. Xia, Y. Yang, and Y. Hu (2022). Sulfur emissions from consumption by developed and developing countries produce comparable climate impacts. Nature Geoscience 15, 184–189.

Liu, M., R. Tan, and B. Zhang (2021). The costs of "blue sky": Environmental regulation, technology upgrading, and labor demand in China. Journal of Development Economics 150, 102610.

Ministry of Environmental Protection (2007a). Accounting methods for industrial pollution source and discharge. State council notice (guo fa [2007] no. 36), technical report.

Ministry of Environmental Protection (2007b). National calculation of emitted pollutants. State council notice (guo fa [2007] no. 36), annex.

Nam, K.-M., C. J. Waugh, S. Paltsev, J. M. Reilly, and V. J. Karplus (2014). Synergy between pollution and carbon emissions control: Comparing China and the U.S.j. Energy Economics 46, 186–201.

OECD (2009). Ensuring Environmental Compliance: Trends and Good Practices. OECD Publishing (Paris).

Ohara, T., H. Akimoto, J.-i. Kurokawal, N. Horii, K. Yamaji, X. Yan, and T. Hayasaka (2007). An Asian emission inventory of anthropogenic emission sources for the period 1980-2020. Atmospheric Chemistry and Physics 7, 4419–4444.

Preonas, L. (2023, 09). Market Power in Coal Shipping and Implications for U.S. Climate Policy. The Review of Economic Studies 91(4), 2508–2537.

Qi, T., X. Zhang, and V. J. Karplus (2014). The energy and co2 emissions impact of renewable energy development in China. Energy Policy 68, 60–69.

Rodrigue, J., D. Sheng, and Y. Tan (2022a). The curious case of the missing Chinese emissions. Journal of the Association of Environmental and Resource Economists 9(4), 755–805.

Rodrigue, J., D. Sheng, and Y. Tan (2022b, 05). Exporting, Abatement, and Firm-Level Emissions: Evidence from China's Accession to the WTO. The Review of Economics and Statistics, 1–45.

Rodrigue, J., Q. Shi, and Y. Tan (2024). Trade policy uncertainty & resource misallocation. European Economic Review 164(C).

Sen, A. K. (1997). From income inequality to economic inequality. Southern Economic Journal 64(2), 384–401.

Shapiro, J. (2022). The environmental bias of trade policy. Quarterly Journal of Economics 136(2), 831–886.

Shapiro, J. S. and R. Walker (2018, December). Why is pollution from US manufacturing declining? The roles of environmental regulation, productivity, and trade. American Economic Review 108(12), 3814–54.

Shapiro, J. S. and R. W. Walker (2024, March). Is air pollution regulation too lenient? Evidence from US offset markets. Working Paper 28199, National Bureau of Economic Research.

Shi, X. and Z. 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.

Sklar, M. (1959). Fonctions de repartition an dimensions et leurs marges. Publ. inst. statist. univ. Paris 8, 229–231.

Springmann, M., D. Zhang, and V. J. Karplus (2015). Consumption-based adjustment of emissions-intensity targets: An economic analysis for China's provinces. Environmental and Resource Economics 61(4), 615–640.

State Council (2006). Reply to pollution control plan during the eleventh five-year plan. Technical report, State Council.

State Council (2007a). Comprehensive working plan for energy conservation and pollutant emission reduction. Technical report, State Council, Beijing (In Chinese).

State Council (2007b). Implementation plans and methods of statistics, monitoring and assessment for energy conservation and pollutant emission reduction. Technical report, State Council, Beijing (In Chinese).

Streets, D. G., N. Y. Tsai, H. Akimoto, and K. Oka (2000). Sulfur dioxide emissions in Asia in the period 1985-1997. Atmospheric Environment 34, 4413–4424.

Streets, D. G. and S. T. Waldhoff (2000). Present and future emissions of air pollutants in China: SO_2, NO_x and CO. Atmospheric Environment 34, 363–374.

Tanaka, S., K. Teshima, and E. Verhoogen (2022). North-south displacement of environmental regulation: The case of battery recycling. American Economic Review: Insights 4(3), 271–288.

UNEP (2019). Environmental rule of law: First global report. Technical report, United Nations Environment Programme.

United Nations (2015). Paris agreement. United Nations treaty series, Paris.

van Soest, D. P., J. A. List, and T. Jeppesen (2006). Shadow prices, environmental stringency, and international competitiveness. European Economic Review 50(5), 1151–67.

Vennemo, H., K. Aunan, H. Lindhjem, and H. M. Seip (2009). Environmental pollution in China: Status and trends. Review of Environmental Economics and Policy 3(2), 209–230.

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

Wang, A. L. (2013). The search for sustainable legitimacy: Environmental law and bureaucracy in China. Harvard Environmental Law Review 37, 365–440.

Wang, C. and Y. Lin (2022). Does bargaining power mitigate the relationship between environmental regulation and firm performance? Evidence from china. Journal of Cleaner Production 331, 129859.

Wang, C., J. Wu, and B. Zhang (2018). Environmental regulation, emissions and productivity: Evidence from Chinese COD emitting manufacturers. Journal of Environmental Economics and Management 92, 54–73.

Wang, T. and R. Zhang (2019). Environmental regulation and performance of Chinese manufacturing firms. University of Tennessee.

Wong, C. and V. J. Karplus (2017). China's war on air pollution: Can existing governance structures support new ambitions? China Quarterly, 1–23.

Wu, J., Y. Deng, J. Huang, R. Morck, and B. Yeung (2013). Incentives and outcomes: China's environmental policy. Working Paper 18754, National Bureau of Economic Research.

Xu, Y. (2011). The use of a goal for SO_2 mitigation planning and management in China's 11^{th} five-year plan. Journal of Environmental Planning and Management 54(6), 769–783.

Xu, Y., R. H. Williams, and R. H. Socolow (2009). China's rapid deployment of SO2 scrubbers. Energy & Environmental Science 2, 59–465.

Yang, S., C. Wang, H. Zhang, T. Lu, and Y. Yi (2022). Environmental regulation, firms' bargaining power, and firms' total factor productivity: evidence from China. Environmental Science and Pollution Research 29, 9341–9353.

Zhang, B., X. Chen, and H. Guo (2018). Does central supervision enhance local environmental enforcement? Quasi-experimental evidence from China. Journal of Public Economics 164, 70–90.

Zhang, D., V. J. Karplus, C. Cassisa, and X. Zhang (2014). Emissions trading in china: Progress and prospects. Energy Policy 75, 9–16.

Zhang, D., S. Rausch, X. Zhang, and V. J. Karplus (2013). Quantifying regional economic impacts of CO2 intensity targets in China. Energy Economics 40, 687–701.

Zhang, D., M. Springmann, and V. J. Karplus (2016). Equity and emissions trading in China. Climatic Change 134(1-2), 131–146.

Zhang, X., V. J. Karplus, T. Qi, D. Zhang, and J. He (2016). CO2 emissions in china: How far can new efforts bend the curve? Energy Economics 54, 388–395.

A Supplemental Appendix (Not For Publication)

This appendix documents (i) the measurement of firm-level SO_2 emissions, (ii) a discussion of empirical robustness checks, (iii) omitted model details, (iv) a detailed description of variable measurement for the counterfactual exercises, and (v) a formal discussion of limitations and conditions needed to apply quantile treatment estimates in counterfactual analysis.

A.1 Measurement of SO_2 Emissions

The variable capturing annual firm-level SO_2 emissions measures the volume of SO_2 emissions from fuel consumption and the production process emitted from the premises of an enterprise during one calendar year. The Ministry of Environmental Protection (MEP) calculates this value using the following formula (Ministry of Environmental Protection (2007a); Ministry of Environmental Protection (2007b)):

$$SO_2 \text{ emissions} = SO_2 \text{ emissions from fuel consumption} + SO_2 \text{ emissions from production}$$

Over the course of a calendar year, each establishment is monitored across a series of k (adjacent) monitoring periods. During any particular monitoring period, equation (6) is used to calculate the aggregate emitted kilograms of SO_2 by firm i during monitoring period j:

$$E_{ij} = C_{ij} \times Q_{ij} \times F_{ij}^{-1} \times T_{ij} \times G_{ij} \times 10^{-6}, \tag{6}$$

where

- E_{ij} is the emitted kilograms of SO_2 during a monitoring period;
- C_{ij} is the average density of SO_2 over each hour during the monitoring period (milligrams/cubic metre);
- Q_{ij} is the volume of wasted gas emissions during the monitoring period (cubic metre/hour);
- F_{ij} is the production load during the monitoring period;
- T_{ij} is the number of emission hours;
- G_{ij} is the average production load of the monitored enterprise.

Aggregate annual establishment-level SO_2 pollution in kilograms, E_{it}, is calculated as

$$E_{it} = \sum_{j=1}^{k} E_{ij}, \tag{7}$$

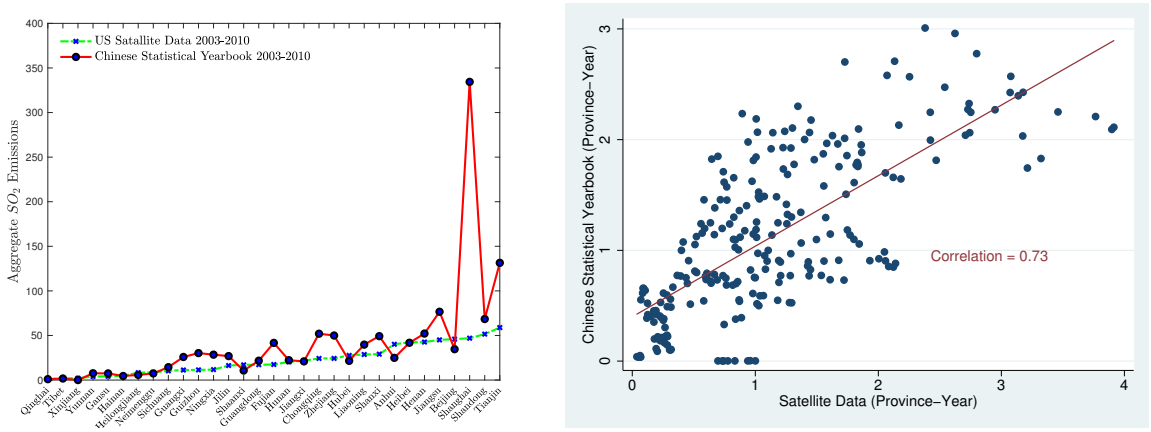
where k is the number of monitoring periods in year t. While firm-level emissions are measured at their source for each establishment, the MEE aggregates this information to the level of individual firms. Should a firm have multiple plants, only firm-level aggregate emissions are reported.

The fact that air pollution disperses rapidly over geographic space raises a concern regarding the measurement of SO_2 at a particular location. For instance, while satellite imagery may be able to provide external measures of SO_2 emissions, it cannot credibly assign pollution to any particular plant. An advantage of our data is that emissions are measured at the source establishments, ensuring accurate attribution. Nonetheless, we remain concerned that aggregation from individual plants to firm-level quantities may induce measurement error across geographic space. Likewise, given the sensitivity of environmental issues in Chinese policy circles, a skeptical reader may reasonably question whether the firm-level surveys report accurate information. To address these concerns, we investigate the quality of our data through a series of validation exercises.

A.1.1 Data Quality

We crosscheck the accuracy of the aggregate SO_2 emissions data from the MEP by comparing it with reported satellite data (Global Modeling and Assimilation Office (GMAO) (2015)) in Figure 15. Specifically, we normalize provincial SO_2 emissions by dividing each observation by the total quantity of emissions by the amount from Qinghai, the province with the smallest amount of emissions. To compute a comparable measure from satellite data, we multiply the reported density by the total area of each province and again normalize by provincial emissions from Qinghai. In Figure 15a the red line with dots reports the value reported by the MEP, while the blue line with squares represents the aggregated value constructed from U.S. satellite data. Shanghai is a clear outlier. This discrepancy plausibly reflects difference in the standard area of Shanghai we use to compute our aggregate value from the satellite data and the greater Shanghai area as classified by the MEP.

Figure 15b plots normalized province-year pairs over our sample period. We observe very strong correlation across data sources; the correlation coefficient for each province-year combination is 0.73. In sum, we find that both data sources exhibit similar empirical patterns across time and space. Although it has been previously argued that independent estimates of Chinese SO_2 emissions are likely higher than those reported in official statistics (Streets and Waldhoff (2000); Streets et al. (2000); Ohara et al. (2007); Cao et al. (2009)), our chief concern is systematic discrepancies across locations. In this sense, we do not find significant evidence of systematic reporting bias.



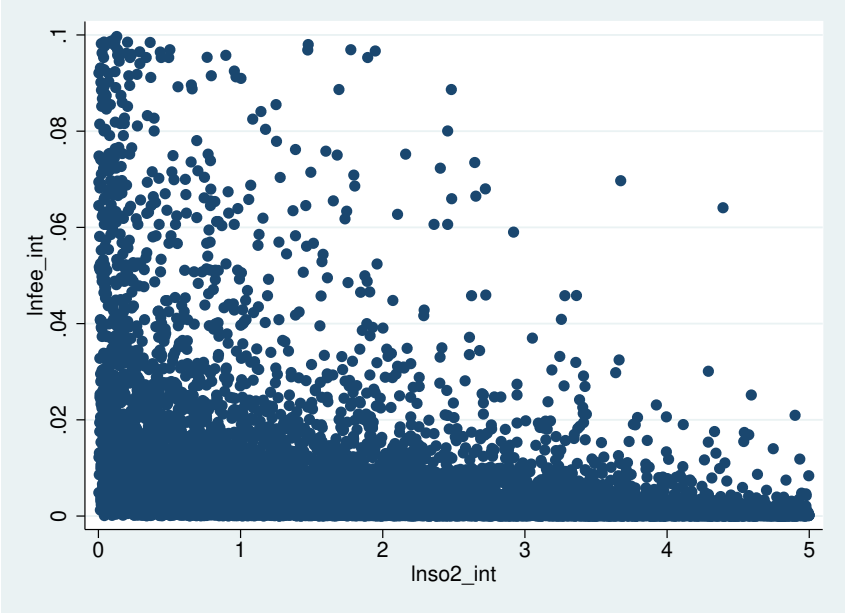
(a) Satellite Data by Province (b) Satellite Data by Province & Year

Notes: Panel (a) compares average provincial SO_2 emissions across provinces from the Chinese Statistical Yearbook with the same information from U.S. satellite data. The red line with dots represents the value reported in the Chinese Statistical Yearbook, while the blue line with squares represents the aggregated value constructed from U.S. satellite data. Panel (b) reports the same information as panel (a) but does not average the data over time. It plots province-year pairs.

Figure 15: Data Validation, Benchmarking to Satellite Data

A.2 Emission Discharge Fees and Emissions Intensity

Although emissions taxes are common regulatory tool in China, there is a wide set of additional regulatory tools which affect emissions: quotas, output taxes, fines, loan conditions, restricted export market access, and public pressure, among other regulatory tools. We do not directly observe applied regulation in any year, with one exception. In 2004, we observe emissions discharge fees among manufacturing emitters. Figure 16 plots (log) firm-level discharge fees per unit of emissions against (log) firm-level emissions intensity. As expected, (i) emissions-intensive firms face a systematically lighter regulatory burden in the form of average emissions fees and (ii) among firms with facing the same average emissions fees there remains a wide variation in emissions intensity. The correlation coefficient between discharge fees and emissions intensity is -0.19.



Notes: The above figure plots the logarithm of the average firm-level emissions discharge fee in 2004 against the logarithm of firm-level emissions intensity in 2004.

Figure 16: Emissions Discharge Fees and Emissions Intensity

A.3 Summary Statistics and Additional Pre-Trend Analysis

Table 6 documents basic summary statistics for level variables from the benchmark sample, while Table 7 conducts simple linear regressions to test for significant differences across firms in treatment and control provinces over the pre-treatment period. While we do observe a significant difference in emissions intensity in the sample that includes all Chinese provinces (Treatment Definition 1), we do not observe significant differences otherwise. As discussed in Section 3, this difference is expected given different levels of economic development across Eastern and Western China.

Variable	Year	Treatment Mean	Treatment Std. Dev.	Control Mean	Control Std. Dev.	Diff.	p-Value
		(1)	(2)	(3)	(4)	(5)	(6)
SO_2 Emissions	2004	4.06	23.77	4.60	47.07	-0.54	0.71
SO_2 Emissions	2005	4.38	26.25	5.30	48.19	-0.92	0.54
Output	2004	0.19	1.27	0.18	1.31	0.01	0.90
Output	2005	0.21	1.48	0.24	1.54	-0.03	0.67
Coal	2004	0.50	2.85	0.33	2.63	0.17	0.91
Coal	2005	0.54	3.16	0.39	2.79	0.15	0.17
Scubber Capacity	2004	2.64	40.79	0.64	7.77	2.00	0.31
Scubber Capacity	2005	2.37	42.10	0.71	6.61	1.66	0.41

Notes: Columns (1) and (2) report the mean and standard deviation for each variable among treated firms in each year. Columns (3) and (4) report the same information for untreated firms. Column (5) reports the difference in means. Column (6) reports the p-value from a t-test for statistical differences across treated and untreated firms. Output is measured as deflated revenue.

Table 6: Benchmark Sample (Defn 2) Summary Statistics

	Sample: All Provinces (Treatment Defn 1)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
	-0.468*	-0.353	0.116	-0.004
	(0.206)	(0.211)	(0.067)	(0.036)
Obs.	26,568	26,568	26,568	8,789
R^2	0.466	0.468	0.309	0.461
	Sample: Border Provinces (Treatment Defn 2)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
	0.295	0.289	-0.006	-0.086
	(0.149)	(0.149)	(0.117)	(0.115)
Obs.	7,515	7,515	7,515	3,711
R^2	0.417	0.355	0.421	0.548
	Sample: Border Cities (Treatment Defn 3)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
	0.414	0.283	-0.131	-0.061
	(0.218)	(0.246)	(0.152)	(0.070)
Obs.	2,967	2,967	2,967	1,486
R^2	0.446	0.355	0.488	0.513

Notes: Standard errors, clustered at the 2-digit industry-level, are in parentheses. * indicates statistical significance at the 5 percent level. All variables measured in log terms. Year and 4-digit industry fixed effects included in all regressions.

Table 7: 2004-2006 Balancing Table

We also conduct a series of additional tests aimed at characterizing differential average pre-trends across treated and control groups. Specifically, common investigation of differential pre-trends across treated and control groups include regression exercises that test for average differences across treatment and control groups, where the average difference in outcomes across groups in normalized to zero in the last year prior to treatment. Although we report pre-trend analysis based on the QTT estimator for the average response and for the entire distribution in Table 1 and Figure 9, we recognize that the the interpretation of the estimates is relatively uncommon. As such, we support our previous pre-trend analysis by conducting a standard difference-in-differences pre-trend regression

$$Y_{it} = \sum_{t=2004}^{2006} \alpha_t(\eta_t \times D_{it}) + \eta_t + \eta_i + \varepsilon_{it}$$

where the η_t and η_i are year and firm fixed effects, D_{it} is a treatment indicator, and ε_{it} is an error term. We report pre-treatment coefficients for our benchmark outcome variable, emissions intensity, along with emissions, output and coal consumption in Table 8. The top panel considers the full sample of all Chinese provinces (Treatment Definition 1), the middle panel reports results for the benchmark sample (Treatment Definition 2), and the bottom panel reports results for the sample of border cities (Treatment Definition 3). We find no evidence of statistically significant pre-trend differences in the primary outcome variable for any sample. However, the top panel of Table 8 reports significantly greater average output and coal consumption among treated firms when using the full sample of all provinces, while the 2004 coefficient on SO_2 emissions is also estimated to be larger in absolute magnitude and marginally significant. We take this as evidence of potentially significant underlying differences across the Chinese geographic landscape. In contrast, the point estimates in the middle and bottom panels tend to be much smaller in both 2004 and 2005, and the confidence intervals always overlap zero for all outcomes.

	Sample: All Provinces (Treatment Definition 1)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
Treatment_{2004}	0.137	0.235	0.099*	0.110*
	(0.109)	(0.119)	(0.049)	(0.029)
Treatment_{2005}	-0.032	-0.032	0.042	0.036
	(0.062)	(0.074)	(0.040)	(0.033)
Obs	26,568	26,568	26,568	8,789
R^2	0.872	0.884	0.910	0.910
	Sample: Border Provinces (Treatment Definition 2)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
Treatment_{2004}	0.069	0.129	0.060	0.101
	(0.150)	(0.112)	(0.068)	(0.058)
Treatment_{2005}	-0.043	-0.020	0.023	0.036
	(0.099)	(0.096)	(0.046)	(0.039)
Obs	7,515	7,515	7,515	3,711
R^2	0.800	0.815	0.906	0.901
	Sample: Border Cities (Treatment Definition 3)			
	SO_2 intensity	SO_2 emissions	Output	Coal Consumption
Treatment_{2004}	0.108	0.107	-0.001	-0.002
	(0.208)	(0.205)	(0.061)	(0.090)
Treatment_{2005}	0.070	0.047	-0.022	-0.026
	(0.196)	(0.180)	(0.039)	(0.059)
Obs	2,967	2,967	2,967	1,486
R^2	0.787	0.799	0.920	0.909

Notes: Standard errors, clustered at the 2-digit industry-level, are in parentheses. * indicates statistical significance at the 5 percent level. All outcome variables measured in log terms. Year and 4-digit industry fixed effects included in all regressions. All treatment parameters should be interpreted relative to 2006, which is normalized to 0.

Table 8: 2004-2006 Pre-Trend Regressions

A.4 Robustness Checks

A.4.1 Alternative Estimation Approaches

This section explores the robustness of our results using alternative methods such as quantile DID (Callaway et al. (2018), QCLO) or quantile CIC (Athey and Imbens (2006), QCIC). As noted in Section 4, Callaway et al. (2018) employ a distributional assumption where the unknown copula is replaced with the observed copula between the change and the initial level of outcomes for the untreated group. Alternatively, the QCIC model proposed by Athey and Imbens (2006) identifies the QTT under the assumption that the empirical model is monotone in a scalar unobservable. In particular, the approach of Athey and Imbens (2006) assumes that the distribution of unobservables does not change over time, although the return to unobservables may change over time. An advantage of either of these alternatives is that they do not require

two years of pre-sample data.[37] In contrast, provided that two years of pre-treatment data exist, the approach in Callaway and Li (2019) leverages the observed changes prior to treatment to identify the QTT. In this sense, the assumption in Callaway and Li (2019) is most closely related to DID assumptions, which are frequently invoked in policy analysis.

Figure 20 documents the QCLO and the QCIC estimates for the 2007–2010 period. In each year and for each estimator, the average treatment effect is never estimated to be significantly different from zero. However, the quantile treatment effect estimates no longer decline monotonically across the distribution and the standard errors are modestly larger. In contrast, the absolute magnitude of the estimates, particularly at the the upper quantiles, are much greater when we employ the Callaway and Li (2019) QTT approach and are much more consistent over time.

These latter features are both virtues and concerns. On one hand, the QTT estimates paint a much clearer picture of the distributional response to the policy change. On the other hand, the lack of consistency across estimation approaches justifies greater scrutiny of the underlying estimation assumptions. Fortunately, our pre-trend tests suggested that the estimation assumptions underlying the Callaway and Li (2019) QTT approach were arguably the most plausible in this setting.

A.4.2 Alternative Definitions of Treatment

This section examines the robustness of the benchmark estimates to alternative definitions of treatment. Instead of focusing on firms located in the treatment border provinces, we include all provinces and divide them into treatment and control locations, as illustrated in Figure 4a. This definition allows us to consider a wider range of firms, including large, emissions-intensive coastal producers. It also has the effect of substantially increasing our sample size, which may significantly improve inference across quantiles. As noted above, this approach has the disadvantage of including regions with significantly greater variation in industrial composition, exposure to international markets, and geographic determinants of pollution dispersion.

Our third definition of treatment and control firms restricts attention to firms that are located in cities that are in close proximity to the treatment border, as shown in Figure 4b. As illustrated in Figure 4c, we restrict attention to prefectures that are located on the treatment border. Since individual prefectures are typically small, this helps assure the comparability of treatment and control firms in larger provinces. Unfortunately, it also comes at the cost of substantially shrinking our sample size. While the benchmark exercise included over 2,000 individual firms (1,500 treated firms), the restricted definition of treatment shrinks our sample

[37] Consistent with the benchmark QTT estimates, we employ 2005 as the pre-treatment year for both the QCLO and QCIC estimates.

to just over 700 firms (450 treated firms).

The results from using alternative definitions of treatment are illustrated in Figures 21 and 22. Regardless of the sample under investigation, we estimate results that are qualitatively consistent with the benchmark sample: there is no impact in 2006, the largest emissions-intensity declines are initially confined to the upper quantiles, a larger fraction of quantiles experience meaningful emissions reductions in later years, and there are modest, but statistically insignificant, increases in emissions intensity among the lower quartiles. Moreover, Table 13 confirms that the average treatment effects are estimated to be negative and statistically significant in one year (2009) when using the sample comprising all provinces.

Despite the qualitative consistency of our findings, there are meaningful differences across experiments. Indeed, we observe the smallest estimated changes and the smallest confidence intervals in Figure 21. The relatively small magnitude of the changes in treatment provinces closer to the coast may reflect a number of underlying characteristics. For instance, we might expect that large, export-oriented producers would continue to receive favorable treatment from officials and, as such, may not change behavior as much. Further, to the extent that air pollution disperses more rapidly in coastal locations (Rodrigue et al. (2022b)), we might expect that there are stronger free-rider incentives among firms located in those provinces. Alternatively, given the disproportionate growth of coastal locations prior to China's 11^{th} FYP, many firms in these locations may have started adjusting to more stringent environmental regulation earlier than firms located inland.

The full sample definition of treatment (Treatment Definition 1) also returns the most precise estimates. With few exceptions, we observe much tighter confidence intervals across all quantiles. This is to be expected given that we are employing a significantly larger sample; the full sample includes nearly 9,000 firms, 4.5 times more than our benchmark setting.

The estimates presented in Figure 22 use the most restrictive definition of treatment and control firms and, consequently, report relatively wide confidence intervals. Despite the fact that the sample size shrunk by nearly two thirds relative to benchmark setting, we continue to recover QTT estimates that are similar in magnitude to those from the benchmark definition and, in most cases, maintain statistical significance at standard levels of evaluation.

A.4.3 Balanced Panel

China's industrial sector is broadly characterized by rapid growth and significant firm-level churning during our sample period. Accordingly, our benchmark samples reflect this feature in a significant degree of attrition over time. For example, our benchmark sample has nearly 6,426 firms in 2007, but only 3,663 firms in 2010, as many firms drop out over time. The change in sample composition, however, raises the complementary concern that the estimated

	Year				
	2006	2007	2008	2009	2010
ATT	-0.16	-0.16	-0.40*	-0.45*	-0.45*
	(0.14)	(0.15)	(0.18)	(0.18)	(0.18)

Notes: Standard errors are in parentheses. * indicates statistical significance at the 5 percent level.

Table 9: Log Emissions Intensity, 2006–2010, ATT, Balanced Panel

changes in quantiles over time largely reflect different sample composition rather than greater firm adjustment.

To address this concern, we repeat our benchmark exercise on a balanced panel of firms that are present in every year between 2004 and 2010. The average treatment effects are reported in Table 9 and demonstrate that, if anything, the average treatment effects are larger and statistically stronger in the balanced panel despite the smaller sample size. Among treated firms, the average firm emissions intensity is estimated to have declined by 36 percent by 2009–2010. Figure 23 confirms that this decline is again concentrated among the largest and most intensive polluters and is only present towards the end of our sample period. The difference between the benchmark sample and the balanced panel can be be explained by the fact that the balanced panel disproportionately drops the initially least emissions intensive plants, which, consequently, are also those most likely to be in the lower tail of our benchmark sample.

A.4.4 Zeros

In our benchmark results, we include all firms that meet the minimum sample criterion: each firm is present in the two pre-treatment years (2004 and 2005) and the given treatment year. This includes the roughly 6 percent of firms that report zero emissions in a given year. Although zero emissions are entirely possible in principle, it presents a dilemma when we take the natural log of emissions.

On the one hand, we do not want to arbitrarily drop these firms, particularly since they are likely to fall in the tail of the emissions distribution. On the other hand, including these firms raises two important concerns. First, in our benchmark approach we add one to the value of one to each firm. This manipulation effectively changes the lower bound of the emissions distribution and introduces a small amount of measurement error.[38] Second, these firms may be less well measured in the data in any year.

To investigate the impact zeros have on our estimates, instead use an inverse hyperbolic

[38]The median firm in 2006 emitted over 60,000 kilograms of SO_2 in a calendar year, while firms in the first percentile report emissions of 250 kilograms of SO_2 in a calendar year.

sine transformation of the dependent variable. Among upper quantiles, the estimated QTTs are negative, statistically significant and take a very similar magnitude to our benchmark estimates. Among the lower quantiles, the QTTs are again positive, but the coefficients are larger and more precisely estimated than the benchmark case. In this sense, using the inverse hyperbolic sine transformation reinforces the differential rise in emissions intensity among firms in the lowest quantiles. Indeed, the estimated QTT for the lowest quantile is now always statistically significant after 2006 and grows in magnitude over the 11^{th} FYP (See Appendix Table 15 and Figure 24).

A.4.5 Correlated Pollutants

SO_2 emissions-intensive firms are likely to also emit other pollutants, such as particulate matter (PM). While SO_2 emissions were a primary focus of the 11^{th} FYP, we would that other pollutants would be likewise affected by the regulatory change. On the one hand, the adoption of new technology or greater abatement may generate PM emissions reductions; on the other hand, substitution towards other, less regulated pollutants may do the opposite.

Repeating our benchmark exercise while using PM intensity as the outcome variable, we recover similar policy-driven changes in the emissions distribution across firms and time (See Appendix Table 16 and Figure 25). In particular, there are sharp declines in the PM intensity distribution among the upper quantiles, sharp increases among the lower quantiles and little change among the middle quantiles. One notable difference: the average treatment effect in 2010 is large, positive and statistically significant. The QTT estimates reveal this is entirely driven by differences in the first decile across pollutants, suggesting there may be some substitution towards other pollutants among these firms.

A.4.6 Finer quantiles

We explore whether a greater number of quantiles would suggest different patterns of heterogeneity across firms by estimating the QTT for 19 quantiles instead of the benchmark 9 quantiles (See Appendix Table 17 and Figures 24). Again, the highest quantiles are estimated to significantly decrease emissions intensity in response to China's FYP, while the lowest quantiles are estimated to increase emissions intensity. Both findings are again statistically significant in each year after 2006. There is again no estimated change in the middle quantiles, broadly confirming the benchmark result across the emissions intensity distribution. Moreover, estimated coefficients confirm a monotonic decline in the treatment effects across the distribution of heterogeneous producers, reinforcing this key result.

A.4.7 Sectoral Differences

While our benchmark estimates adjust for time-invariant differences across sectors, there may remain significant differences across sectors within the Chinese manufacturing sector. This approach does not allow us to characterize whether the regulatory changes induced by China's 11^{th} FYP were broad based or whether they fell disproportionately on specific industries. To investigate this possibility, we isolate three of the largest sub-sectors in our benchmark for individual analysis: (i) chemical manufacturing, (ii) non-metallic minerals and (iii) metal smelting and rolling.[39]

As documented in Table 18 and Figure 27, the QTT estimates in the chemicals manufacturing industry and the metal smelting and rolling industry are qualitatively similar to those in the benchmark sample. Quantitatively, the estimated emissions intensity declines among the upper tail of the emissions intensity distribution are larger than those in the benchmark sample in both industries, while the rise in emissions intensity among the lower tail in the metal smelting and rolling industry is likewise larger than that estimated than the corresponding benchmark estimates. That said, the QTT estimates are never statistically significant at conventional significance levels, through this is expected as the sample sizes are generally quite small. The sample size in chemicals manufacturing and metal smelting and rolling industries are 11 and 7 percent of the benchmark sample, respectively.

In the non-metallic mineral manufacturing industry, there are relatively modest changes in emissions intensity across the entire distribution producers. China's 11^{th} FYP, however, is associated with statistically significantly emissions intensity declines in the 5^{th}, 6^{th}, and 7^{th} deciles over the 2008-2010 period. To investigate the impact of industrial clustering, we measure the fraction of non-metallic mineral producers in the interquartile range during the pre-sample period. If non-metallic mineral producers are disproportionately clustered in (outside of) the middle deciles, we would expect that more (less) than 50 percent of non-metallic mineral producers will be in the pre-sample interquartile range. We find that roughly 50 percent of non-metallic mineral manufacturers percent of producers are found in pre-sample interquartile range. Repeating this exercise for different ranges of the data, we find 80 percent non-metallic mineral producers between the 10^{th} and 90^{th} percentiles and 90 percent of the non-metallic minerals producers between the 5^{th} and 95^{th} percentiles of the benchmark, pre-sample emissions intensity distribution.

[39] In all years and sectors, the ATTs are never significantly different from zero. Other sub-sectors were generally too small for analysis in our benchmark definition of treatment.

A.5 Alternative Mechanisms

A.5.1 Emissions Leakage

Firms may alternatively reduce emissions by offshoring production of intermediate inputs abroad. Indeed, Chinese importing rose rapidly in our sample period and could plausibly have been encouraged by domestic environmental regulation. To investigate this possibility, we develop a measure of firm-level emissions intensity inclusive of offshore emissions using the following four-step process:

1. Focusing on 2004 as a benchmark year, we compute the average emissions intensity for each industry in the Chinese environmental data.

2. Imported products are categorized by the HS classification system in Chinese customs data, not the Chinese industrial classification system. Using the CIC industry-HS product code concordances from Brandt et al. (2012), we map the measured emissions intensities to 6-digit HS codes.

3. We then compute embedded emissions in offshore inputs for firm i in year t as the sum of the expenditure weighted emissions intensities of each input:

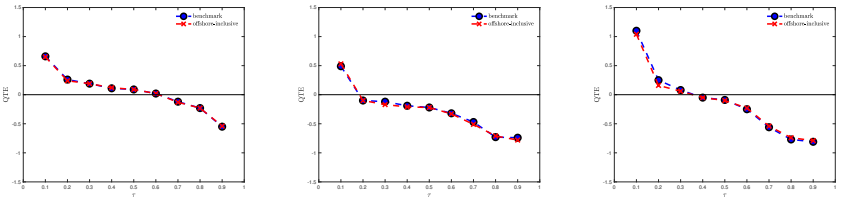
$$o_{it} = \sum_{j} \text{expenditure}_{jt} \times \text{intensity}_{j,04}.$$

4. Offshore-inclusive emissions intensity is then computed as the sum of the observed and offshore emissions divided by firm output.

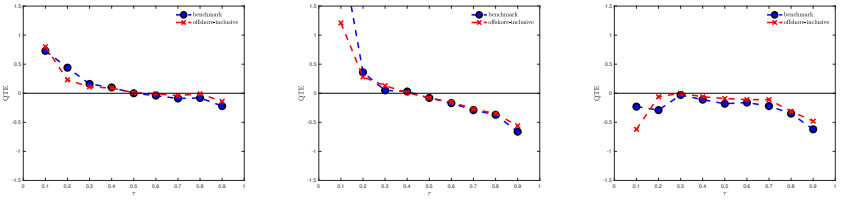
Repeating our analysis with the sample of border provinces, Table 25 and Figure 17 document nearly identical results to that reported in the benchmark exercise.[40] As documented in Figure 30, there is no meaningful change in any of the estimated quantile treatment effects. On the one hand, this suggests that environmental regulation did not lead to meaningful trade-induced pollution leakage. On the other hand, because our benchmark sample restricts attention to firms in provinces far from the eastern seaboard, we may miss leakage among firms for which this was a particularly viable option for reducing emissions.

To investigate this latter hypothesis, we repeat the exercise using the sample of firms from all Chinese provinces, including those located in coastal provinces. As documented in Figure 17 and Table 25, we observe slightly larger emissions declines in the first decile, no change in the middle deciles, and systematically smaller declines in emissions in the upper deciles

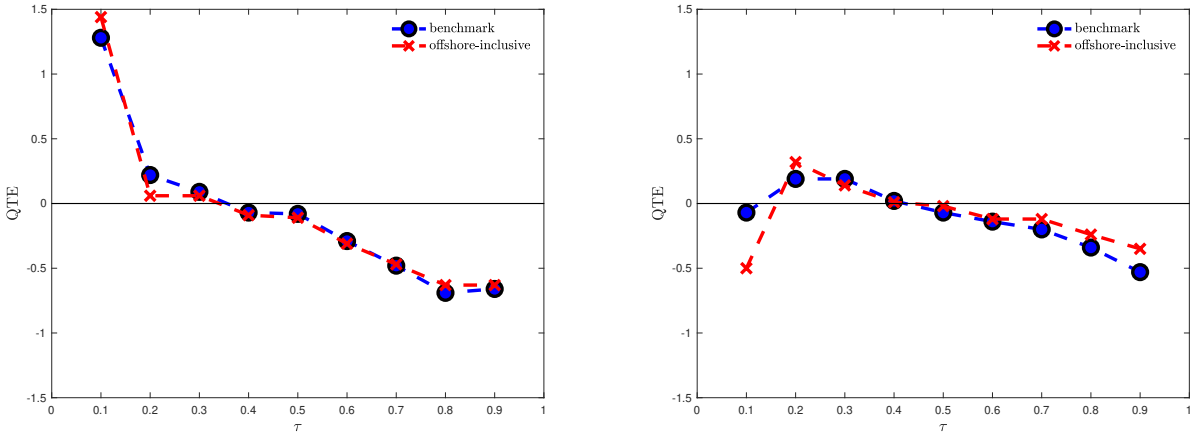
[40] For expositional clarity Figure 17 omits the confidence intervals. QTT estimates for year and definition of treatment are reported in appendix Figure 30.



(a) Border Prov, 2007 (b) Border Prov, 2008 (c) Border Prov, 2009



(d) All Prov, 2007 (e) All Prov, 2008 (f) All Prov, 2009



(g) Border Prov, 2010 (h) All Prov, 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent uniform confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 17: Emissions Leakage

once we account for offshore emissions. For the uppermost decile, our findings suggest that emissions leakage accounts for 18 percentage points, or 34 percent, of the total emissions-intensity decline in 2010. In sum, the asymmetric differences suggest that, at least for coastal producers, offshoring pollution-intensive components may have been an important channel through which trade-exposed firms adjusted to China's 11$^{\text{th}}$ FYP.

A.5.2 Product Reallocation

A third channel by which firms can reduce emissions is by shifting production towards less emissions-intensive products (Barrows and Ollivier (2018)). We investigate whether China's 11$^{\text{th}}$ FYP drove within-firm reallocation towards cleaner production by measuring whether within-firm changes in product class can explain the differential changes in the emissions-intensity distribution across Chinese provinces. To measure the product reallocation-augmented emissions intensity, we first develop a measure of firm-level emissions inclusive of those that would have been emitted if the firm had maintained the same product class over time.

1. In each year, we compute average emissions for each CIC industry using Chinese environmental data.

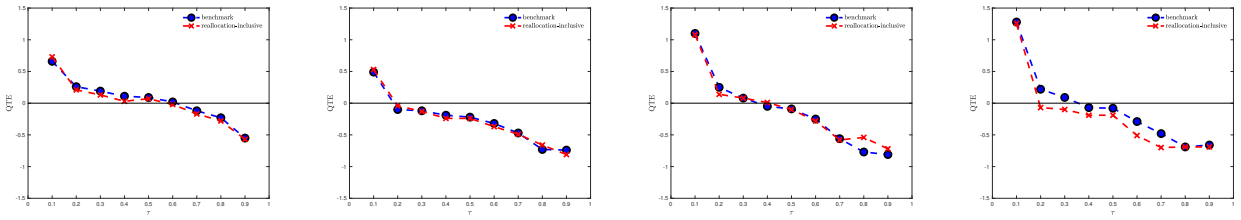
2. For each firm, we compute a product reallocation premium as the difference in emissions intensities of the firm's original and current product class:

$$\text{Product reallocation premium}_{it} = \text{Intensity}_{ot} - \text{Intensity}_{ct},$$

   where Intensity_{ot} is firm i's original product class and Intensity_{ct} is the firm's current production class in year t. The firm's original product class is based on its reported CIC code in 2004.

3. Product reallocation-inclusive emissions intensity in year t is then computed as the sum of firm i's observed emissions intensity in year t and the reallocation intensity premium in year t.

Among firms that do not change CIC codes over time, the product reallocation premium is exactly zero; this is the case for roughly 80 percent of firms in our data. Firms that start producing products that are less emissions intensive will have positive product reallocation premiums, while those that switch to producing more emissions-intensive products will have negative emissions premiums. Among the remaining 20 percent of firms that change industries over time, roughly half move to industries with cleaner emissions intensities while the other half shift production to dirtier products. If product reallocation is systematically associated



(a) 2007 (b) 2008 (c) 2009 (d) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent uniform confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 18: Emissions and Product Reallocation

with changes in emissions intensity across the distribution of Chinese producers, we expect that including the emissions premiums again would attenuate the estimated quantile treatment effects, that is, if product reallocation is an important mechanism by which firms adjust to environmental regulation, we expect that the absolute value of the estimated QTTs will be relatively small after including the product reallocation premiums.

Figure 18 reports that there is very little evidence of a strong product reallocation response before 2010. The estimated product reallocation-augmented QTTs are generally very close to those from our benchmark exercise, suggesting that procut reallocation had little impact on the policy-driven path of emissions intensity across the distribution of Chinese manufacturers. In 2010 we often recover estimates that are slightly larger in absolute magnitude than those from the benchmark exercise, the opposite of what we would have expected if product reallocation had been a key response to China's 11^{th} FYP.

This does not imply that product reallocation is not an important driver of China's emissions growth; it only indicates that we did not find evidence supporting the hypothesis that the implementation of the 11^{th} FYP induced meaningful within-firm product reallocation to mitigate emissions. China's 11^{th} FYP may have induced product reallocation effects that manifested themselves through firm-level entry and exit or through product reallocation within a narrowly defined product class, both of which are beyond of the scope of our analysis.[41]

A.5.3 Location Changes:

Firms may respond to new regulation by moving to less regulated jurisdictions. Using the addresses of the firms themselves, we verify that no treated firms in our benchmark sample move to a province with a smaller SO_2 target after the implementation of China's 11^{th} FYP. Indeed, fewer than 1 percent of all firms in any province change locations between 2007-2010.

[41] Rodrigue et al. (2022a) find that despite high rates of entry and exit, entering and exiting firms contribute relatively little to aggregate emissions growth in the Chinese manufacturing sector.

A.6 Model Details

To map the distribution of emissions intensity into a measure of aggregate productivity, we consider a simple closed-economy production environment blending Hsieh and Klenow (2009), Shapiro and Walker (2018) and Rodrigue et al. (2022a). Producers purchase physical inputs (labor, capital, materials, energy), V, in competitive markets and produce output Y according to the physical production function:

$$Y_{ist} = (1 - \theta_s)\tilde{A}_{ist}V_{ist},$$

where $\tilde{A}_{ist}$ is firm-level TFP and θ_s captures the fraction of firm inputs that are redirected towards the reduction of firm-level emissions, E_{ist}, in sector s:

$$E_{ist} = (1 - \theta_s)^{1/\alpha_s}\tilde{A}_{ist}V_{ist} = (1 - \theta_s)^{\frac{1-\alpha_s}{\alpha_s}}Y_{ist}.$$

Combining the above equations, we can write output directly as a function of emissions in an emissions-augmented production function:

$$Y_{ist} = \tilde{A}_{ist}^{1-\alpha_s}E_{ist}^{\alpha_s}V_{it}^{1-\alpha_s} = A_{ist}E_{ist}^{\alpha_s}V_{ist}^{1-\alpha_s},$$

where $A_{ist} = \tilde{A}_{ist}^{1-\alpha_s}$ is abatement-normalized productivity. Industry output Y_s is a CES aggregate of N_s varieties, $Y_s = \left(\sum_{i=1}^{N_s} Y_{is}^{\frac{\sigma-1}{\sigma}}\right)^{\frac{\sigma}{\sigma-1}}$. Within industries the CES structure gives rise to the residual demand function $Y_{ist} = (\Phi_s/P_{ist})^{\sigma}$, where P_{ist} is the price charged by firm i in year t and Φ_s is a demand shifter common to all firms in sector s. Aggregate output is a Cobb-Douglas combination of sectoral output, $Y = \prod_{s=1}^{S} Y_s^{\gamma_s}$, where $\sum_{s=1}^{S}\gamma_s = 1$.

A.6.1 Profit Maximization

Each firm chooses emissions and value added to maximize profits:

$$\max_{E_{ist},\, V_{ist}} \pi_{ist} = P_{ist}Y_{ist} - W_{st}V_{ist} - T_{ist}E_{ist},$$

where W_{st} is the price of value added common to all firms in industry s and T_{ist} is the firm-specific emissions tax incurred by firm i in year t. The first order conditions yield

$$\begin{aligned}
\left(\frac{\sigma-1}{\sigma}\right) P_{ist}A_{ist}\alpha_s \left(\frac{E_{ist}}{V_{ist}}\right)^{\alpha_s - 1} V_{ist}^{1-\alpha_s} &= T_{ist} \\
\left(\frac{\sigma-1}{\sigma}\right) P_{ist}A_{ist}(1-\alpha_s) \left(\frac{E_{ist}}{V_{ist}}\right)^{\alpha_s} V_{ist}^{1-\alpha_s} &= W_{st},
\end{aligned}$$

which in turn gives the optimal emissions-value added ratio

$$\frac{E_{ist}}{V_{ist}} = \frac{\alpha_s}{(1-\alpha_s)} \left(\frac{W_{st}}{T_{ist}} \right).$$

and the firm's pricing rule as

$$P_{ist} = \left(\frac{\sigma}{\sigma - 1} \right) \left(\frac{T_{ist}}{\alpha_s} \right)^{\alpha_s} \left(\frac{W_{st}}{1-\alpha_s} \right)^{1-\alpha_s} \frac{1}{A_{ist}},$$

which is the standard markup over the marginal cost form common to CES demand systems.

A.6.2 Marginal Revenue Products and Firm-Level Productivity

Rewriting the first-order conditions reveals that the marginal revenue products of emissions and value added are inversely proportional to measured emissions intensity and the value added-revenue ratio, respectively:

$$\begin{aligned} MRPE_{ist} &= \alpha_s \frac{\sigma - 1}{\sigma} \frac{R_{ist}}{E_{ist}} = T_{ist} \\ MRPV_{ist} &= (1-\alpha_s) \frac{\sigma - 1}{\sigma} \frac{R_{ist}}{V_{ist}} = W_{st}, \end{aligned}$$

where R_{ist} denotes firm revenue $R_{ist} = P_{ist} Y_{ist}$.

As in Foster et al. (2008), we distinguish between physical and revenue productivity but in our case focus on the emissions-adjusted version of productivity:

$$\begin{aligned} TFPQ_{ist} &= A_{ist} = \frac{Y_{ist}}{E_{ist}^{\alpha_s} V_{ist}^{(1-\alpha_s)}} \\ TFPR_{ist} &= P_{ist} A_{ist} = \frac{P_{ist} Y_{ist}}{E_{ist}^{\alpha_s} V_{ist}^{1-\alpha_s}}. \end{aligned}$$

Analogously to the Hsieh and Klenow (2009) setting, $TFPQ_{ist}$ will vary across firms in the same industry as long as there are differences in underlying physical productivity $\tilde{A}_{ist}$. In contrast, $TFPR_{ist}$ does not vary across firms within an industry unless firms face differential distortions, such as different prices for emissions. High (low) firm $TFPR$ indicates that the firm confronts barriers (subsidies) that raise (decrease) the firm's marginal products of emissions and value added, rendering the firm smaller (larger) than optimal.

Finally, it is straightforward to show that $TFPR$ is proportional to a geometric average of

the firm's marginal revenue products of emissions and value added:

$$TFPR_{ist} = \frac{\sigma}{\sigma - 1}\left(\frac{MRPE_{ist}}{\alpha_s}\right)^{\alpha_s}\left(\frac{MRPV_{ist}}{1-\alpha_s}\right)^{1-\alpha_s} \propto T_{ist} \; \forall \; i,$$

which, in turn, is proportional to emissions regulation (in the absence of other frictions).

A.6.3 Aggregate Productivity and Output

Aggregate, emissions-adjusted productivity for all N_s manufacturing firms in sector s can be written as

$$TFP_{st} = \left[\sum_{i=1}^{N_s}\left(A_{ist}\frac{\overline{TFPR}_{st}}{TFPR_{ist}}\right)^{\sigma-1}\right]^{\frac{1}{\sigma-1}}, \tag{8}$$

where $\overline{TFPR}_{st}$ is a geometric average of the average marginal revenue product of emissions and value added. Under the standard assumption that $\sigma > 2$, greater variance in $TFPR_{ist}$ will cause TFP_{st} to fall. Holding A_{ist} constant, a firm with higher (lower) $TFPR_{ist}$ has higher (lower) marginal costs and prices. This will induce the firm to produce too little (much) in the absence of distortions. The wide dispersion of observed emissions intensities suggest that firms face very different emissions regulatory burdens. Further, if $TFPR_{ist}$ and A_{ist} are positively correlated, then the distortions render firms with high physical productivity (high A_{ist}) smaller than optimal, which hurts aggregate TFP (since those firms get less weight).

Aggregate output can then be computed as

$$Y_t = \prod_{s=1}^{S}(TFP_{st}E_{st}^{\alpha_s}V_{st}^{1-\alpha_s})^{\gamma_s}$$

where $E_{st} = \sum_{i=1}^{N_s} E_{ist}$, $V_{st} = \sum_{i=1}^{N_s} V_{ist}$, and $\gamma_s = \frac{P_s Y_s}{PY}$ is the revenue share of sector s.

A.7 Parameter and Variable Measurement

Computing A_{is}, $TFPR_{is}$, and TFP requires knowledge of the emissions technology parameter α_s and the elasticity of the demand parameter, σ. For the emissions technology parameter α_s, we retrieve estimates of these parameters from Rodrigue et al. (2022a) for the year 2005. In particular, we employ the median estimated value of α_s for each industry. Since we cannot identify the elasticity of substitution, we employ a common value and set $\sigma = 3$, as established in Gopinath et al. (2017).

A.7.1 Measuring Firm Productivity

Given the model parameters (α_s, σ) and data (E_{ist}, R_{ist}), we construct the various productivity measures needed to compute TFP. The careful reader will note that we do not have access to measures of traditional physical inputs (capital, labor, materials) without matching to other data sets and losing a large fraction of observations.

Restricting attention to productivity gains *from emissions reductions alone*, we abstract from variation in $MPRV_{it}$ and impose the assumption that $MPRV_{ist}$ is constant for all firms, $MPRV_{ist} = W_{st} = 1$. This assumption is violated in practice, and its imposition eliminates the possibility of quantifying absolute productivity gains from the reductions in distortions over the sample, as in Hsieh and Klenow (2009). However, we are still able to identify the productivity gains solely due to emissions-intensity declines, which is our primary interest.

The marginal revenue product of emissions, $MPRE_{ist}$, is taken directly from the data given (α_s, σ). Moreover, given the aforementioned normalization, $MPRE_{ist}$ and $TFPR_{ist}$ are proportional to each other:

$$TFPR_{ist} = \frac{\sigma}{\sigma - 1} \left(\frac{MRPE_{ist}}{\alpha_s} \right)^{\alpha_s}. \tag{9}$$

The last object we need to measure aggregate TFP is firm-level, emissions-adjusted productivity, A_{ist}. Multiplying and dividing the emissions-augmented production function by the firm's price we can write physical productivity as

$$A_{ist} = \frac{P_{ist} Y_{ist}}{P_{ist} E_{ist}^{\alpha_s} V_{it}^{1-\alpha_s}} = \frac{1}{P_{it}} \left(\frac{R_{it}}{E_{it}} \right)^{\alpha_s} \left(\frac{R_{it}}{V_{it}} \right)^{1-\alpha_s}.$$

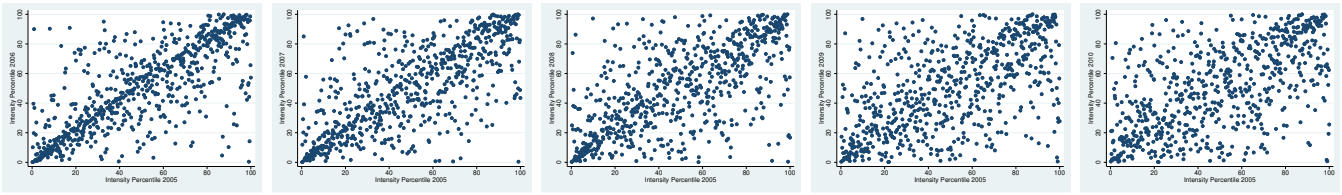
Using the definitions of $MRPE_{ist}$ and $MRPV_{ist}$ yields

$$A_{ist} = \frac{R_{it}^{\frac{1}{\sigma-1}}}{\Phi^{\frac{\sigma}{\sigma-1}}} \left(\frac{MRPE_{ist}}{\alpha_s} \right)^{\alpha_s} \left(\frac{MRPV_{ist}}{1 - \alpha_s} \right)^{1-\alpha_s} \left(\frac{\sigma}{\sigma - 1} \right),$$

where we also employ the model's implication that firm-specific revenues and prices are linked through the common demand shifter $R_{ist} = \Phi^{\sigma} P_{ist}^{1-\sigma}$. Note that $\Phi^{\frac{\sigma}{\sigma-1}}$ is common to all firms and, as such, we normalize it to one without loss of generality. Likewise, focusing on emissions-driven gains in TFP and abstracting from variation in $MRPV_{ist}$, we can measure A_{ist} as

$$A_{ist} = R_{ist}^{\frac{1}{\sigma-1}} \left(\frac{MRPE_{ist}}{\alpha_s} \right)^{\alpha_s} \left(\frac{1}{1 - \alpha_s} \right)^{1-\alpha_s} \left(\frac{\sigma}{\sigma - 1} \right), \tag{10}$$

which can be measured with the data at hand.



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The above figure presents scatter plots of the firm's percentile rank in 2005 and its percentile rank in each outcome year.

Figure 19: Emissions-Intensity Percentile Rank Correlations Over Time

A.8 Identifying Counterfactual TFP

Aggregating distributional estimators is subject to a subtle methodological challenge: while it is possible to identify the causal impact of the policy change across the distribution of emissions intensity, the estimated coefficients do not correspond to particular firms. Counterfactual aggregate productivity, however, depends on the joint distribution of firm productivity (A_{is}) and emissions intensity, through $TFPR_{is}$. In principle, the process documented in Section 7 only returns the marginal distribution of $TFPR$.[42]

To fix ideas, consider an extreme example: suppose the policy had no impact on the emissions-intensity ranking of firms. Only in this extreme case do the QTT estimates capture the causal impact of the policy change on any particular subset of firms and can be straightforwardly mapped to the existing distribution of firm TFP. On the one hand, when firms' rank does not change too much, the above concern may be overstated. Using the balanced sample, we record each firm's emissions-intensity percentile ranking in 2005 and each outcome year (2006–2010). Scatterplots, displayed in Figure 19, indicate a high degree of correlation, though it appears to decline over time. The correlation coefficients range from 0.57 in 2009 to 0.75 in 2006. The correlation is similar across treated and untreated firms, though modestly smaller among treated firms, as we would expect. On the other hand, the correlation coefficients are less than one. As such, direct inference of the firm's treatment effect is not possible.

Although we can only recover the counterfactual marginal distribution of emissions intensity, it is nonetheless possible to perform standard aggregation procedures as long as the dependence (copula) between emissions intensity and firm TFP (A_i) remains the same. Specifically, define observed sectoral TFP_s as

$$TFP_s = \int g(y,a) dF_{Y(1),A(1)}(y,a)$$

[42] While it may be tempting to rely on subset or conditioned estimates in place of QTT estimates, our empirical estimates confirm that much heterogeneity in policy responses is not necessarily associated with observable covariates, similar to that documented in Bitler et al. (2017).

where $(Y(1),\ A(1))$ is the observed joint distribution of emissions intensity and firm TFP and $g(\cdot)$ represents the function which aggregates these objects into sectoral TFP. The process outlined in Section 7 allows us to recover the counterfactual distribution of firm emissions intensity, $Y(0)$. Likewise, the assumption that the policy does not directly affect firm TFP returns the counterfactual distribution $A(0)$. However, these conditions do not generally return the counterfactual joint distribution of emissions intensity and firm TFP, $(Y(0),\ A(0))$. To conduct our the counterfactual aggregation exercises we further assume that the dependence (copula) between Y and A is the same. A natural intuition for this condition is that firms that have relatively high productivity under the policy would have had relatively high productivity absent the policy. This weaker condition allows us to quantify aggregate productivity gains implied by the estimates from the robust QTT estimator. Formally, for $j \in \{0, 1\}$, let $C_j(u, v)$ be the copula of $(Y(j),\ A(j))$. Under the above conditions we recover the joint distribution $(Y(0),\ A(0))$ (and hence $TFP(0)$) as:

$$\begin{aligned} F_{Y(0),A(0)}(y, a) &= C_0(F_{Y(0)}(y), F_{A(0)}(a)) \\ &= C_1(F_{Y(0)}(y), F_{A(0)}(a)) \\ &= P(Y(1) \leq F^{-1}_{Y(1)}(F_{Y(0)}(y)), A(1) \leq a) \\ &= P(F^{-1}_{Y(0)}(F_{Y(1)}(Y(1))) \leq y, A(1) \leq a) \end{aligned}$$

where the first equality writes the joint distribution as the copula of the marginals, the second equality employs the condition that we are holding the copula fixed in our counterfactual, and the third uses the definition of the copula and that our counterfactual holds the marginal distribution of $A(0)$ to be the same as the marginal of $A(1)$. The last equality rearranges terms. The key finding is that, although we do not impose rank invariance, it is still possible to recover aggregate counterfactual TFP.

To see how this is different than assuming rank invariance, notice that $F(Y(1))$ is uniform$[0, 1]$, so that F^{-1} applied to this variable is just a way to generate a random variable that has the same distribution as $Y(0)$. The copula assumption that we make there implies that the dependence between this random variable and $A(0)$ is the same as the observed dependence between $Y(1)$ and $A(1)$. Consider the following example. Suppose that $Y(1)$, $Y(0)$, $A(1)$, and $A(0)$ are all independent standard normal random variables. Since $Y(1)$ and $Y(0)$ are independent, this means that rank invariance between them does not hold. However, all of our other assumptions hold: (i) the distribution of $A(1)$ and $A(0)$ are the same, (ii) the copula of $(Y(1), A(1))$ is the same as the copula of $(Y(0), A(0))$ - in particular, they are both independent - and our approach recovers the joint distribution of $(Y(0), A(0))$.

A.9 Causality, Distributional Estimates & Aggregate Productivity

A natural concern with the QTT estimates from samples which might plausibly suffer from estimation bias is that they will inherently induce bias in the aggregate productivity conclusions. We argue below that these concerns are mitigated as long as the relative impact of the 11^{th} FYP is captured by QTT estimates across the distribution of heterogeneous producers. Let τ index quantiles and let N_τ be the number of quantiles we use to characterize the distribution. Likewise, let T and C denote firms assigned to treatment and control status in the experiment, respectively, so that $\Omega_{\tau,T}$ is the set of treatment firms in quantile τ. Recalling equation (3), sectoral productivity can be rewritten as

$$\begin{aligned} TFP_{st} &= \left[\prod_{\tau=1}^{N_\tau} \overline{TFPR}_{st}^{\tau,C} \left(\frac{\overline{TFPR}_{st}^{\tau,T}}{\overline{TFPR}_{st}^{\tau,C}}\right)^{\frac{N_s^{\tau,T}}{N_s}}\right] \times \\ &\quad \left\{\sum_{\tau=1}^{N_\tau}\left[\sum_{i\in\Omega_{\tau,C}}\left(\frac{A_{ist}}{TFPR_{ist}}\right)^{\sigma-1} + \sum_{i\in\Omega_{\tau,T}}\left(\frac{A_{ist}}{TFPR_{ist}}\right)^{\sigma-1}\right]\right\}^{\frac{1}{\sigma-1}} \end{aligned} \tag{11}$$

The first line in equation (11) only depends on the geometric average of TFPR among control firms, $\overline{TFPR}_{st}^{\tau,C}$, and the ratio of geometric averages across treatment and control firms, $\frac{\overline{TFPR}_{st}^{\tau,T}}{\overline{TFPR}_{st}^{\tau,C}}$. Accordingly, the first line of equation (11) does not require a causal change in $TFPR$ due to the policy on T, only a consistent estimate of the difference between treated and control locations. The same is not true in second line in equation (11). For this we will need an additional assumption.

Assumption 3(a): Assume that we know the differential impact of the policy across treatment and control firms in each quantile τ as N_τ gets arbitrarily large, $N_\tau \to \infty$.

As N_τ grows large sectoral TFP can be approximated by

$$\begin{aligned} TFP_{st} &\approx \left[\prod_{\tau=1}^{N_\tau} \overline{TFPR}_{st}^{\tau,C} \left(\frac{\overline{TFPR}_{st}^{\tau,T}}{\overline{TFPR}_{st}^{\tau,C}}\right)^{\frac{N_s^{\tau,T}}{N_s}}\right] \times \\ &\quad \left\{\sum_{\tau=1}^{N_\tau}\frac{1}{(TFPR_{st}^{\tau,C})^{\sigma-1}}\left[(A_{st}^{\tau,C})^{\sigma-1} + (A_{st}^{\tau,T})^{\sigma-1}\left(\frac{TFPR_{st}^{\tau,C}}{TFPR_{st}^{\tau,T}}\right)^{\sigma-1}\right]\right\}^{\frac{1}{\sigma-1}} \end{aligned} \tag{12}$$

where the firm index i disappears and it replaced by the pair $\frac{TFPR_{st}^{\tau,C}}{TFPR_{st}^{\tau,T}}$ since fine enough quantiles will capture pairs of treatment and control firms. The second term in equation (12) only depends

on the control value of TFPR for a particular quantile and the ratio of treatment and control values for the same quantile.

While illustrative Assumption 3(a) is not practically useful since control and treatment samples are not perfectly balanced in terms of the number of firms within each quantile as the number of quantiles grow. However, it is a useful in that it suggests a practical approach to testing whether our productivity findings are robust to unobserved reallocation. Specifically, consider the following weaker assumption.

Assumption 3(b): For each quantile τ and treatment status $J \in \{C, T\}$ let the firms be sufficiently similar such that

$$\sum_{i \in \Omega_{\tau,J}} \left(\frac{A_{ist}^{\tau,J}}{TFPR_{ist}^{\tau,J}} \right)^{\sigma-1} \approx N_{\tau,J} \left(\frac{\overline{\overline{A}}_{st}^{\tau,J}}{\overline{\overline{TFPR}}_{st}^{\tau,J}} \right)^{\sigma-1}$$

where $\overline{\overline{A}}_{st} = \frac{1}{N_{\tau,J}} \sum_{i \in \Omega_{\tau,J}} A_{ist}$ and $\overline{\overline{TFPR}}_{st} = \frac{1}{N_{\tau,J}} \sum_{i \in \Omega_{\tau,J}} TFPR_{ist}$.

Intuitively, if different *types* of firms end up in different emissions intensity quantiles aggregate TFP will only depend on the (observed) level of revenue productivities in the control provinces and ratio of average revenue productivities within quantiles across provinces. In this sense, we can approximate aggregate productivity dynamics even in the absence of strictly causal estimates measurement as long as we have a reasonable characterization of the distribution of control and treatment $TFPR$ in fine enough quantiles:

$$\begin{aligned} TFP_{st} \approx & \left[\prod_{\tau=1}^{N_\tau} \overline{TFPR}_{st}^{\tau,C} \left(\frac{\overline{TFPR}_{st}^{\tau,T}}{\overline{TFPR}_{st}^{\tau,C}} \right)^{\frac{N_s^{\tau,T}}{N_s}} \right] \times \\ & \left\{ \sum_{\tau=1}^{N_\tau} \frac{1}{(\overline{\overline{TFPR}}_{st}^{\tau,C})^{\sigma-1}} \left[(\overline{\overline{A}}_{st}^{\tau,C})^{\sigma-1} + (\overline{\overline{A}}_{st}^{\tau,T})^{\sigma-1} \left(\frac{\overline{\overline{TFPR}}_{st}^{\tau,C}}{\overline{\overline{TFPR}}_{st}^{\tau,T}} \right)^{\sigma-1} \right] \right\}^{\frac{1}{\sigma-1}} \end{aligned} \tag{13}$$

Accordlingly, this result also suggests a natural robustness check: re-estimating the benchmark sample with finer quantiles. As documented in Figure 26 and Table 17, our benchmark estimates are broadly confirmed when we increase the number of estimated quantiles, yielding further the confidence in the aggregate conclusions.

A.10 Emissions Taxes, Emissions Intensity and $TFPR$

Equation (1) establishes that firm-level $MPRE_{is}$ is proportional to firm-level emissions taxes, T_{is}, and, accordingly, equation (2) further indicates that $TFPR_{is}$ is proportional to T_{is}^{α}. Thus, policies which increase (decrease) the variance of T_{is} across firms will also increase (decrease) the variance of $MPRE_{is}$ and $TFPR_{is}$. However, because $MPRE_{is}$ is proportional to the inverse of emissions intensity, $1/I_{is}^{E} = R_{is}/E_{is}$, it is less clear than an increase (decrease) in the variance of T_{is} will induce an increase (decrease) in $MPRE_{is}$ or $TFPR_{is}$. Likewise, it is less obvious that a decline (rise) in the variance of emissions intensity inherently implies a rise (fall) in $MPRE_{is}$ and $TFPR_{is}$.

To provide intuition, we employ a Taylor series approximation to characterize the relationship between the variance of emissions intensity and the variance of emissions taxes. Specifically, equation (1) implies that emissions intensity can be written as

$$I_{is}^{E} = \alpha_s \left(\frac{\sigma - 1}{\sigma}\right) \frac{1}{T_{is}}$$

so that the variance emissions intensity, I_{is}^{E}, increases with the variance of $\frac{1}{T_{is}}$. A Taylor expansion of $Var(1/T)$ yields

$$\begin{aligned}
Var(1/T) &= Var\left[\frac{1}{\mu} + (T-\mu)\frac{\partial 1/\mu}{\partial \mu} + \frac{(T-\mu)^2}{2}\frac{\partial^2 1/\mu}{\partial \mu^2} + ...\right] \\
&= Var\left[(T-\mu)\frac{\partial 1/\mu}{\partial \mu} + \frac{(T-\mu)^2}{2}\frac{\partial^2 1/\mu}{\partial \mu^2} + ...\right] \\
&= \left(\frac{\partial 1/\mu}{\partial \mu}\right)^2 Var\left[(T-\mu)\right] + 2Cov\left[(T-\mu), \frac{(T-\mu)^2}{2}\frac{\partial^2 1/\mu}{\partial \mu^2} + ...\right] \\
&\quad + Var\left[\frac{(T-\mu)^2}{2}\frac{\partial^2 1/\mu}{\partial \mu^2}\right]
\end{aligned}$$

where μ is the mean of the $1/T$ distribution. Restricting attention to the first term we observe $Var(1/T) \approx \frac{1}{\mu^4} Var(T)$, suggesting that the variance of $1/T$, the variance of T and, consequently, the variance of emissions intensity will rise and fall together.

B Empirical Specification Details

B.1 Potential Outcomes and Distributional Estimators

In a traditional DID context the first key assumption is a parallel trends assumption that says that the "path" of outcomes that firms located in treated provinces would have experienced if they had not been treated is the same as the path of outcomes that firms located in untreated provinces actually experienced. This argument is well-known for identifying and estimating the *average* treatment effect (ATT) for the treated subpopulation. Recent work in the econometrics literature has extended these results to identify and estimate the entire distribution of potential treated and untreated outcomes in a DID setup (Callaway and Li (2019); Callaway et al. (2018)). Often the parameter of interest in this context is the *quantile* treatment effect (QTT) for the treated subpopulation. We build on this recent work and, in particular, expand the arguments in those papers to deal with the additional challenge in the current setup of having extra pre- and post-treatment time periods.

To make these arguments more concrete, recall that only one of the potential outcomes is observed for each firm in each time period. In particular, observed outcomes are given by

$$Y_{it} = \begin{cases} D_i Y_{it}(1) + (1 - D_i) Y_{it}(0) & t \geq t^* \\ Y_{it}(0) & t < t^*. \end{cases}$$

In other words, in post-treatment time periods, we observe treated potential outcomes for firms that are affected by the pollution reduction targets and untreated potential outcomes for firms that are not affected by the pollution reduction targets. In pre-treatment time periods, we observe untreated potential outcomes for all firms. Recall, that we define the Average Treatment Effect on the Treated (ATT) and Quantile Treatment Effect on the Treated accordingly as

$$ATT_t = E[Y_t(1) - Y_t(0)|D = 1] \quad \text{and} \quad QTT_t(\tau) = Q_{Y_t(1)|D=1}(\tau) - Q_{Y_t(0)|D=1}(\tau).$$

As is standard, we focus on identifying treatment effects in post-treatment periods, that is, periods where $t \geq t^*$ and the QTT is additionally indexed by $\tau \in [0, 1]$, and $Q_{Y_t(j)|D=1}(\tau) = F^{-1}_{Y_t(j)|D=1}$ for $j = \{0, 1\}$. To give some examples, $QTT_t(0.5)$ is the difference between the median outcome for firms under the pollution reduction policy and the median outcome that the same group of firms would have experienced in period t if the policy had not been implemented. $QTT_t(0.9)$ is the difference between the 90th percentiles and is informative about how pollution reduction targets affect environmental outcomes among high-polluting firms. $QTT_t(0.1)$ is the difference between the 10^{th} percentiles under the policy and, in the absence of the policy, for

the group of firms affected by the policy and is informative about how the pollution reduction targets affect environmental outcomes among low-polluting firms.

In order to identify the ATT_t and QTT_t in periods where $t \geq t^*$, the key identification challenge is on $E[Y_t(0)|D=1]$ (for the ATT) and $F_{Y_t(0)|D=1}$ (for the QTT, noting that if we can identify this distribution then we can invert it in order to obtain $Q_{Y_t(0)|D=1}$). In other words, to identify the ATT_t and QTT_t, it would be sufficient to identify $F_{Y_t(0)|D=1}$, the distribution of outcomes that treated firms would have experienced if they had not participated in the treatment. DID identification arguments proceed by noting that

$$Y_t(0) = \underbrace{(Y_t(0) - Y_{t^*-1}(0))}_{(A)} + \underbrace{Y_{t^*-1}(0)}_{(B)}. \tag{14}$$

The term on the right hand side of Equation 14 comes from adding and subtracting $Y_{t^*-1}(0)$ (which is an observed outcome because it is pre-treatment, even for firms in the treated group). Parallel trends assumptions are helpful for dealing with expressions like the one in Equation 14.

As discussed in the main text, a mean parallel trends assumption can be used to identify the ATT. Identifying the QTT requires recovering $F_{Y_t(0)|D=1}$, the distribution of outcomes that treated firms would have experienced if they had not been treated. At an intuitive level, the reason for this difference is the following: From Equation 14, unobserved untreated potential outcomes depend on (i) untreated potential outcomes in the period before a firm becomes treated (from Term (B)) and (ii) how untreated potential outcomes would have evolved between time period $t^* - 1$ and t (from Term (A)). For (i), these are observed outcomes. For (ii), we can learn about their distribution from the distribution of the path of outcomes for the untreated group. But in order to think about the distribution of unobserved untreated potential outcomes, we also need to understand the *dependence* between (i) and (ii). In other words, we need to know if the biggest increases in pollution if the policy had not been implemented would have been by firms with initially the most pollution or least pollution, etc. The parallel trends assumption does not provide information about this additional key piece of information.

B.2 Identifying Quantile Treatment Effects Over Time: Discussion

Assumption 2 in the main text, develops a series of technical conditions under which we can replace the unknown copula $C_{Y_t(0)-Y_{t^*-1}(0),Y_{t^*-1}|D=1}$ with an observed copula and only differ with respect to this choice. To get a better sense of the underlying structural implications of Assumptions 2(i)-2(iii), we consider benchmark models common to many applications including our own. While it is always possible to directly assume that the conditions for each assumption hold, this discussion is intended to provide guidance on more primitive conditions that

validate the adoption of the underlying assumptions. For instance, it is well known that the parallel trends assumption (Assumption 1) is closely related to fixed effects models for untreated potential outcomes; that is,

$$Y_{it}(0) = c_i + v_{it},$$

where c_i is time invariant unobserved heterogeneity, which can be distributed differently for the treated and untreated group, and v_{it} is a time varying unobservable.[43] As with the benchmark case, we do not need to specify a model for treated potential outcomes which, in turn, allows for general forms of selection in the treatment. Immediately, we require that $\Delta v_t \perp D$ and $\Delta_\tau v_\tau \perp D$, which guarantees that Assumption 1 will hold.

Now, for the sake of exposition, consider the following (strong) assumption.

Assumption 3 (Assumptions on Model)

(i) *The v_{it} are independent of c_i*

(ii) *The v_{it} are independent of each other*

(iii) *The v_{it} have the same distribution across time periods*

Assumption 3 is sufficient to imply that Assumption 2(i) holds. Further, for Assumption 2(ii) we have

$$\begin{aligned} P(\Delta_\tau Y_\tau(0) \leq \delta, Y_{t-1}(0)) &= P(\Delta_\tau v_\tau \leq \delta, c + v_{t-1} \leq y | D = 1) \\ &= P(v_\tau - v_{t-1} \leq \delta, c + v_{t-1} \leq y | D = 1) \\ &= P(v_{t-1} - v_{t-2} \leq \delta, c + v_{t-2} \leq y | D = 1) \\ &= P(\Delta Y_{t-1}(0) \leq \delta, \Delta Y_{t-2}(0) \leq y | D = 1), \end{aligned}$$

which implies that Assumption 2 holds. Because the entire joint distribution is the same, it follows that the copulas must also be the same. The third equality is the key step and holds because the joint distribution of $(v_{t+1}, v_{t-2}, c|D = 1)$ is the same as the joint distribution of $(v_{t-1}, v_{t-2}, c|D = 1)$. Further, this argument suggests that it is possible to relax Assumption 3(i) and rather impose that the v_{is} are independent of each other and follow the same distribution across time *conditional on c*.

In contrast, it is less obvious how to relax Assumptions 3(ii) and 3(iii). For instance, Callaway and Li (2019), who invoke Assumption 2(i), allow for serially correlated v_{it} as long as the serial correlation is the same over time. In the present setting, we require all serial correlation to

[43] It is straightforward to allow for time fixed effects, covariates, etc. As such, we focus our discussion on the simplest case.

be the result of individual fixed effects.[44] That said, in many applications, including that in Callaway and Li (2019), researchers commonly impose the assumption that the v_{it} follow the same distribution across all time periods.

Similar arguments do not apply to Assumption 2(iii) because the distribution of c is different across the treated and untreated groups. In this sense, while it may be straightforward to extend Assumption 2(ii) to allow for various forms of serial correlation, Assumption 2(i) is less likely to hold in practice.

C Additional Tables

Year	Method	ATT	τ								
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2005	QTT	-0.14	-0.31	-0.05	-0.21	-0.17	-0.16	-0.17	-0.13	-0.12	-0.06
		(0.10)	(0.29)	(0.16)	(0.12)	(0.11)	(0.09)	(0.10)	(0.10)	(0.11)	(0.16)
	QCLO	-0.14	0.01	-0.16	-0.44*	-0.37*	-0.24	-0.28*	-0.11	-0.01	0.03
		(0.10)	(0.35)	(0.15)	(0.14)	(0.14)	(0.15)	(0.10)	(0.13)	(0.14)	(0.15)
	QCIC	-0.10	-0.12	-0.40*	-0.37*	-0.30*	-0.24*	-0.16	-0.11	-0.19	0.02
		(0.11)	(0.58)	(0.16)	(0.11)	(0.14)	(0.10)	(0.13)	(0.14)	(0.13)	(0.22)
2006	QTT	0.04	0.28	0.17	0.10	0.05	0.08	0.06	0.05	-0.02	-0.06
		(0.08)	(0.20)	(0.12)	(0.10)	(0.08)	(0.07)	(0.08)	(0.09)	(0.10)	(0.14)
	QCLO	0.04	0.43	0.08	0.05	0.07	0.04	0.09	0.08	-0.01	0.15
		(0.08)	(0.36)	(0.14)	(0.13)	(0.11)	(0.10)	(0.10)	(0.09)	(0.10)	(0.11)
	QCIC	0.06	0.22	0.15	0.05	0.04	0.11	0.11	0.09	0.15	0.04
		(0.08)	(0.37)	(0.15)	(0.10)	(0.11)	(0.10)	(0.08)	(0.12)	(0.08)	(0.19)

Notes: Pointwise standard errors are in parentheses. Standard errors for QTT are computed using the empirical bootstrap with 1,000 iterations. * indicates statistical significance at the 5 percent level. The years 2003 and 2004 are used as base periods for outcomes in 2005, while 2004 and 2005 are used as base years for outcomes in 2006.

Table 10: Log SO_2 Intensity, 2005–2006, Pre-trend Tests

[44] As before, we can also condition on observable co-variates, which could potentially mitigate concerns with unobserved serial correlation.

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2007	0.08	0.66*	0.26	0.19	0.11	0.09	0.02	-0.12	-0.23	-0.55*
	(0.11)	(0.33)	(0.14)	(0.12)	(0.11)	(0.08)	(0.10)	(0.10)	(0.12)	(0.15)
2008	-0.14	0.49	-0.10	-0.12	-0.19	-0.22*	-0.32*	-0.47*	-0.73*	-0.74*
	(0.13)	(0.40)	(0.16)	(0.14)	(0.12)	(0.10)	(0.11)	(0.11)	(0.14)	(0.18)
2009	0.07	1.10*	0.25	0.08	-0.05	-0.09	-0.25*	-0.56*	-0.77*	-0.81*
	(0.14)	(0.47)	(0.19)	(0.15)	(0.13)	(0.12)	(0.12)	(0.13)	(0.15)	(0.20)
2010	0.13	1.28	0.22	0.09	-0.07	-0.08	-0.29*	-0.48*	-0.69*	-0.66*
	(0.15)	(0.74)	(0.23)	(0.16)	(0.15)	(0.13)	(0.14)	(0.14)	(0.15)	(0.20)

Notes: Pointwise standard errors are in parentheses. Standard errors for QTT are computed using the empirical bootstrap with 1,000 iterations. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 11: Log SO_2 Intensity, 2007–2010, QTT

Year	ATE	QCLO(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2007	0.08	0.17	0.18	0.15	0.11	0.12	0.23*	0.08	0.05	-0.17
	(0.10)	(0.36)	(0.20)	(0.14)	(0.13)	(0.10)	(0.10)	(0.13)	(0.12)	(0.16)
2008	-0.14	0.27	-0.33*	-0.25	-0.18	-0.06	-0.09	-0.20	-0.36*	-0.34
	(0.13)	(0.48)	(0.16)	(0.16)	(0.12)	(0.14)	(0.13)	(0.13)	(0.12)	(0.19)
2009	0.08	1.27	0.004	-0.26	-0.18	0.01	-0.24	-0.33*	-0.33*	-0.02
	(0.16)	(0.61)	(0.25)	(0.18)	(0.13)	(0.12)	(0.13)	(0.15)	(0.14)	(0.23)
2010	0.13	1.55	-0.27	-0.12	-0.29	-0.07	-0.21	-0.38*	-0.44*	-0.25
	(0.15)	(1.04)	(0.32)	(0.20)	(0.17)	(0.12)	(0.13)	(0.15)	(0.15)	(0.20)
Year	ATE	QCIC(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2007	0.13	0.24	0.26	0.13	0.12	0.29	0.17	0.20	0.17	-0.28
	(0.10)	(0.30)	(0.16)	(0.14)	(0.11)	(0.10)	(0.11)	(0.11)	(0.12)	(0.18)
2008	-0.15	0.28	-0.19	-0.12	-0.02	-0.02	-0.15	-0.15	-0.27	-0.21
	(0.14)	(0.40)	(0.17)	(0.14)	(0.12)	(0.17)	(0.13)	(0.11)	(0.19)	(0.27)
2009	0.01	0.85	0.07	-0.18	-0.03	-0.03	-0.17	-0.27*	-0.14	-0.20
	(0.19)	(0.51)	(0.25)	(0.16)	(0.10)	(0.16)	(0.12)	(0.11)	(0.10)	(0.29)
2010	0.17	1.29	-0.28	-0.20	-0.15	-0.03	-0.20	-0.33*	-0.32	-0.15
	(0.19)	(1.26)	(0.25)	(0.17)	(0.12)	(0.12)	(0.20)	(0.11)	(0.18)	(0.34)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 12: Log SO_2 Intensity, 2007–2010, Alternative Estimation Approaches

Year	ATE	Sample: All Provinces QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.07	0.05	0.38	0.20	0.12	0.05	0.04	0.001	-0.003	0.02
	(0.06)	(0.10)	(0.29)	(0.09)	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)	(0.07)
2007	0.14	0.73*	0.44*	0.16*	0.10	0.00	-0.04	-0.09	-0.08	-0.22*
	(0.08)	(0.27)	(0.16)	(0.08)	(0.07)	(0.07)	(0.06)	(0.06)	(0.06)	(0.10)
2008	0.02	2.50*	0.36*	0.05	0.03	-0.08	-0.17*	-0.29*	-0.37*	-0.66*
	(0.10)	(0.52)	(0.13)	(0.08)	(0.09)	(0.08)	(0.07)	(0.08)	(0.09)	(0.13)
2009	-0.24*	-0.23	-0.29	-0.03	-0.11	-0.18*	-0.16	-0.22*	-0.35*	-0.62*
	(0.10)	(0.16)	(0.33)	(0.12)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)	(0.13)
2010	-0.08	-0.07	0.19	0.19	0.02	-0.07	-0.14	-0.20*	-0.34*	-0.53*
	(0.11)	(0.20)	(0.42)	(0.16)	(0.11)	(0.10)	(0.09)	(0.10)	(0.11)	(0.13)
Year	ATE	Sample: Only Border Cities QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.07	-0.17	-0.07	-0.06	-0.11	0.01	-0.00	0.04	-0.08	-0.15
	(0.13)	(0.33)	(0.18)	(0.16)	(0.14)	(0.15)	(0.13)	(0.15)	(0.16)	(0.21)
2007	-0.04	0.01	0.07	0.06	0.06	0.06	0.00	-0.11	-0.18	-0.42
	(0.17)	(0.47)	(0.28)	(0.19)	(0.17)	(0.16)	(0.16)	(0.15)	(0.19)	(0.26)
2008	-0.31	-0.14	-0.27	0.07	-0.04	-0.06	-0.25	-0.45*	-0.84*	-1.21*
	(0.20)	(0.58)	(0.32)	(0.22)	(0.20)	(0.18)	(0.18)	(0.20)	(0.22)	(0.31)
2009	-0.23	-0.05	0.25	0.22	0.04	-0.09	-0.24	-0.47*	-0.81*	-1.15*
	(0.24)	(0.78)	(0.34)	(0.27)	(0.20)	(0.19)	(0.19)	(0.20)	(0.24)	(0.35)
2010	-0.06	0.25	0.20	0.27	0.15	0.01	-0.13	-0.40	-0.49*	-0.59*
	(0.25)	(1.09)	(0.42)	(0.34)	(0.24)	(0.22)	(0.22)	(0.23)	(0.24)	(0.30)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 13: Log SO_2 Intensity, 2007–2010, QTT, Alternative Samples

Year	ATE	Sample: Balanced Panel 2004–2010 QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.16	-0.25	-0.12	-0.03	-0.02	0.08	-0.04	-0.16	-0.22	-0.05
	(0.14)	(0.25)	(0.17)	(0.15)	(0.12)	(0.16)	(0.13)	(0.15)	(0.16)	(0.19)
2007	-0.16	-0.04	-0.05	0.06	0.18	0.10	-0.08	-0.30	-0.46*	-0.55*
	(0.15)	(0.27)	(0.18)	(0.18)	(0.14)	(0.15)	(0.14)	(0.16)	(0.16)	(0.21)
2008	-0.40*	-0.38	-0.14	-0.17	-0.11	-0.12	-0.36*	-0.50*	-0.74*	-0.57*
	(0.16)	(0.30)	(0.20)	(0.17)	(0.14)	(0.15)	(0.16)	(0.18)	(0.18)	(0.25)
2009	-0.45*	-0.40	-0.05	-0.12	-0.11	-0.18	-0.42*	-0.77*	-0.82*	-0.91*
	(0.18)	(0.30)	(0.24)	(0.19)	(0.17)	(0.18)	(0.18)	(0.20)	(0.22)	(0.29)
2010	-0.45*	-0.58*	-0.24	-0.29	-0.19	-0.19	-0.40*	-0.69*	-0.62*	-0.64*
	(0.18)	(0.28)	(0.21)	(0.19)	(0.17)	(0.17)	(0.17)	(0.18)	(0.20)	(0.24)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 14: Log SO_2 Intensity, 2006–2010, QTT, Balanced Panel, 2004–2010

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.05	0.12	0.15	0.10	0.07	0.08	0.06	0.05	-0.01	-0.03
	(0.05)	(0.10)	(0.11)	(0.09)	(0.07)	(0.07)	(0.07)	(0.07)	(0.08)	(0.11)
2007	0.09	0.48*	0.30*	0.25*	0.19	0.16	0.11	-0.02	-0.11	-0.44*
	(0.07)	(0.15)	(0.13)	(0.10)	(0.08)	(0.09)	(0.09)	(0.09)	(0.10)	(0.13)
2008	-0.17*	0.42*	-0.06	-0.07	-0.09	-0.16	-0.26*	-0.40*	-0.61*	-0.61*
	(0.08)	(0.17)	(0.15)	(0.12)	(0.10)	(0.09)	(0.09)	(0.09)	(0.11)	(0.13)
2009	-0.10	0.64*	0.24	0.08	-0.01	-0.02	-0.20	-0.52*	-0.70*	-0.75*
	(0.09)	(0.16)	(0.17)	(0.15)	(0.12)	(0.11)	(0.12)	(0.12)	(0.13)	(0.17)
2010	-0.14	0.56*	-0.02	-0.02	-0.11	-0.13	-0.30*	-0.51*	-0.72*	-0.69*
	(0.10)	(0.18)	(0.21)	(0.15)	(0.14)	(0.12)	(0.13)	(0.13)	(0.13)	(0.19)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 15: Log SO_2 Intensity, 2006–2010, QTT, Inverse Hyperbolic Sine Transformation

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.19	0.66	0.13	0.20	0.14	0.12	0.14	0.19	0.17	0.16
	(0.11)	(0.76)	(0.16)	(0.15)	(0.11)	(0.12)	(0.11)	(0.11)	(0.11)	(0.16)
2007	0.07	1.58*	0.46*	0.25	0.14	0.05	-0.05	-0.13	-0.38*	-0.52*
	(0.15)	(0.64)	(0.22)	(0.20)	(0.15)	(0.17)	(0.15)	(0.14)	(0.15)	(0.23)
2008	0.03	2.64*	0.43	0.10	-0.12	-0.36*	-0.33	-0.37*	-0.43*	-0.64*
	(0.19)	(0.94)	(0.28)	(0.21)	(0.18)	(0.17)	(0.17)	(0.18)	(0.20)	(0.30)
2009	0.30	3.83*	0.99*	0.46	0.07	-0.21	-0.40*	-0.53*	-0.62*	-0.75*
	(0.23)	(0.79)	(0.38)	(0.25)	(0.20)	(0.21)	(0.18)	(0.19)	(0.24)	(0.34)
2010	0.58*	4.74*	1.72*	0.82*	0.35	-0.06	-0.29	-0.51*	-0.79*	-1.08*
	(0.25)	(0.55)	(0.65)	(0.30)	(0.25)	(0.25)	(0.19)	(0.19)	(0.24)	(0.30)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 16: Log Particulate Matter Intensity, 2006–2010, QTT

Year	ATE	QTT(τ)								
		0.05	0.10	0.15	0.20	0.25	0.30	0.35	0.40	0.45
2006	0.04	-0.80	0.28	0.16	0.17	0.19	0.10	0.11	0.05	0.09
	(0.08)	(0.69)	(0.20)	(0.14)	(0.12)	(0.11)	(0.10)	(0.10)	(0.08)	(0.08)
2007	0.08	1.89*	0.66*	0.35	0.26	0.19	0.19	0.12	0.11	0.13
	(0.11)	(0.86)	(0.33)	(0.18)	(0.14)	(0.11)	(0.11)	(0.11)	(0.11)	(0.09)
2008	-0.14	2.65*	0.49	0.01	-0.10	-0.05	-0.12	-0.15	-0.19	-0.20
	(0.13)	(1.01)	(0.38)	(0.23)	(0.16)	(0.13)	(0.13)	(0.12)	(0.12)	(0.11)
2009	0.07	4.11*	1.10*	0.44	0.25	0.11	0.08	0.00	-0.05	0.03
	(0.15)	(0.92)	(0.48)	(0.29)	(0.19)	(0.16)	(0.16)	(0.14)	(0.13)	(0.13)
2010	0.13	3.76*	1.28	0.53	0.22	0.13	0.09	-0.01	-0.07	-0.00
	(0.15)	(0.79)	(0.73)	(0.30)	(0.23)	(0.20)	(0.17)	(0.16)	(0.15)	(0.15)

Year	QTT(τ)									
	0.50	0.55	0.60	0.65	0.70	0.75	0.80	0.85	0.90	0.95
2006	0.08	0.09	0.06	0.01	0.05	-0.06	-0.02	-0.10	-0.06	-0.05
	(0.07)	(0.07)	(0.08)	(0.08)	(0.09)	(0.09)	(0.09)	(0.11)	(0.13)	(0.11)
2007	0.09	0.06	0.02	-0.06	-0.12	-0.23*	-0.23	-0.40*	-0.55*	-0.55*
	(0.09)	(0.09)	(0.10)	(0.09)	(0.10)	(0.11)	(0.12)	(0.12)	(0.15)	(0.18)
2008	-0.22*	-0.24*	-0.32*	-0.48*	-0.47*	-0.64*	-0.73*	-0.83*	-0.74*	-0.80*
	(0.10)	(0.11)	(0.12)	(0.12)	(0.11)	(0.13)	(0.13)	(0.15)	(0.17)	(0.22)
2009	-0.09	-0.15	-0.25	-0.43*	-0.56*	-0.66*	-0.76*	-0.88*	-0.81*	-0.84
	(0.13)	(0.13)	(0.13)	(0.14)	(0.14)	(0.15)	(0.16)	(0.18)	(0.20)	(0.23)
2010	-0.08	-0.22	-0.29	-0.41*	-0.48*	-0.63*	-0.69*	-0.74*	-0.66*	-0.57*
	(0.13)	(0.14)	(0.15)	(0.15)	(0.14)	(0.16)	(0.15)	(0.17)	(0.20)	(0.20)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 17: Log SO_2 Intensity, 2006–2010, QTT, Finer Quantiles

Year	ATE	Sample: Chemical Manufacting QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.08	-0.23	-0.05	0.03	-0.02	0.11	0.04	0.14	-0.16	-0.25
	(0.18)	(0.40)	(0.25)	(0.21)	(0.21)	(0.20)	(0.25)	(0.21)	(0.24)	(0.36)
2007	-0.03	0.17	0.69	0.83	0.50	0.50	0.40	0.31	-0.09	-0.10
	(0.24)	(0.54)	(0.50)	(0.43)	(0.36)	(0.37)	(0.46)	(0.50)	(0.50)	(0.66)
2008	-0.05	0.24	-0.28	-0.32	-0.20	-0.50	-0.45	-0.59	-0.92*	-1.37*
	(0.28)	(0.99)	(0.36)	(0.28)	(0.30)	(0.31)	(0.30)	(0.34)	(0.39)	(0.43)
2009	0.10	0.23	-0.20	-0.15	-0.11	-0.02	-0.04	-0.41	-0.86*	-1.32*
	(0.38)	(1.14)	(0.47)	(0.40)	(0.40)	(0.37)	(0.39)	(0.43)	(0.40)	(0.58)
2010	-0.16	-0.26	-0.06	-0.44	-0.60	-0.48	-0.59	-0.39	-0.95*	-1.34*
	(0.38)	(0.91)	(0.49)	(0.37)	(0.38)	(0.35)	(0.42)	(0.44)	(0.48)	(0.47)
Year	ATE	Sample: Non-Metallic Minerals QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.03	0.21	0.12	0.03	-0.07	-0.04	-0.11	-0.03	0.07	0.10
	(0.12)	(0.22)	(0.13)	(0.10)	(0.10)	(0.10)	(0.10)	(0.11)	(0.12)	(0.11)
2007	0.09	0.28	0.15	0.11	0.08	0.04	-0.06	-0.03	-0.23	-0.25
	(0.14)	(0.22)	(0.17)	(0.11)	(0.12)	(0.10)	(0.11)	(0.13)	(0.12)	(0.15)
2008	-0.22	-0.06	-0.18	-0.24	-0.38*	-0.44*	-0.42*	-0.50*	-0.38*	-0.39*
	(0.15)	(0.26)	(0.17)	(0.13)	(0.12)	(0.12)	(0.13)	(0.13)	(0.15)	(0.18)
2009	-0.07	0.35	-0.09	-0.16	-0.26	-0.35*	-0.50*	-0.53*	-0.42*	-0.28
	(0.16)	(0.32)	(0.20)	(0.15)	(0.15)	(0.14)	(0.15)	(0.15)	(0.18)	(0.19)
2010	-0.22	-0.02	-0.20	-0.16	-0.40*	-0.38*	-0.49*	-0.47*	-0.24	-0.08
	(0.15)	(0.30)	(0.18)	(0.15)	(0.16)	(0.15)	(0.14)	(0.16)	(0.18)	(0.19)
Year	ATE	Sample: Metal Smelting and Rolling QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.23	0.28	0.87	0.66	0.55	0.17	0.02	-0.15	-0.39	-0.47
	(0.43)	(0.65)	(0.57)	(0.47)	(0.38)	(0.31)	(0.35)	(0.32)	(0.42)	(0.87)
2007	0.82	2.22	1.75*	1.19*	0.44	0.20	-0.25	-0.56	-0.69	-0.34
	(0.48)	(1.86)	(0.66)	(0.54)	(0.45)	(0.45)	(0.43)	(0.46)	(0.49)	(0.63)
2008	-0.21	0.14	-0.07	0.05	-0.52	-1.01*	-1.25*	-1.53*	-1.45*	-1.50*
	(0.54)	(1.09)	(0.70)	(0.46)	(0.50)	(0.49)	(0.54)	(0.55)	(0.62)	(0.68)
2009	-0.30	0.74	0.37	-0.17	-0.44	-0.63	-1.20	-1.50*	-1.81*	-1.94
	(0.66)	(1.15)	(0.77)	(0.63)	(0.56)	(0.60)	(0.63)	(0.68)	(0.83)	(1.46)
2010	0.50	1.34	1.42	0.08	-0.01	-0.68	-0.69	-0.89	-1.55*	-1.43
	(0.84)	(2.08)	(1.34)	(1.09)	(0.81)	(0.65)	(0.64)	(0.65)	(0.73)	(1.14)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years. The top panel restricts attention to the chemical manufacturing industry, the middle panel to the non-metallic minerals industry and bottom panel to metal smelting and rolling industry.

Table 18: Log SO_2 Intensity, 2006–2010, QTT, Industry-Specific Estimates

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.05	-0.08	-0.17	-0.01	0.01	-0.00	-0.04	-0.04	-0.05	-0.15
	(0.05)	(0.16)	(0.11)	(0.06)	(0.05)	(0.04)	(0.04)	(0.05)	(0.06)	(0.14)
2007	-0.08	-0.03	-0.09	-0.00	-0.04	-0.00	-0.04	-0.04	-0.11	-0.31*
	(0.07)	(0.18)	(0.11)	(0.08)	(0.06)	(0.06)	(0.05)	(0.06)	(0.08)	(0.15)
2008	-0.13	-0.14	-0.20	-0.05	-0.09	-0.12	-0.14*	-0.16*	-0.13	-0.29
	(0.08)	(0.20)	(0.13)	(0.10)	(0.07)	(0.06)	(0.06)	(0.07)	(0.10)	(0.20)
2009	-0.23*	-0.10	-0.07	-0.08	-0.14	-0.13	-0.20*	-0.18*	-0.24*	-0.46*
	(0.09)	(0.26)	(0.13)	(0.10)	(0.08)	(0.07)	(0.07)	(0.09)	(0.12)	(0.22)
2010	-0.09	0.15	-0.06	0.09	-0.07	-0.05	-0.06	-0.07	-0.19	-0.71*
	(0.10)	(0.25)	(0.16)	(0.11)	(0.09)	(0.08)	(0.08)	(0.08)	(0.12)	(0.25)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 19: Log Coal-Output Ratio, 2006–2010, QTT

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.13	0.44*	0.37	0.17	0.06	0.01	-0.14	-0.04	0.04	-0.13
	(0.10)	(0.17)	(0.13)	(0.12)	(0.09)	(0.10)	(0.09)	(0.11)	(0.09)	(0.14)
2007	0.32	0.73*	0.59*	0.33*	0.14	0.09	-0.14	-0.16	-0.07	-0.09
	(0.13)	(0.21)	(0.14)	(0.13)	(0.11)	(0.11)	(0.11)	(0.12)	(0.14)	(0.20)
2008	0.003	0.41*	0.48	0.09	-0.22	-0.33*	-0.54*	-0.52*	-0.64*	-0.48*
	(0.14)	(0.22)	(0.18)	(0.15)	(0.13)	(0.14)	(0.12)	(0.13)	(0.14)	(0.23)
2009	0.22	0.29	0.44*	0.11	-0.23	-0.31	-0.59*	-0.49*	-0.52*	-0.45
	(0.17)	(0.35)	(0.20)	(0.18)	(0.15)	(0.16)	(0.15)	(0.14)	(0.17)	(0.23)
2010	0.23	0.21	0.23	0.04	-0.47*	-0.52*	-0.75*	-0.71*	-0.75*	-0.68*
	(0.19)	(0.90)	(0.23)	(0.21)	(0.19)	(0.18)	(0.16)	(0.16)	(0.16)	(0.25)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 20: Log SO_2 Emissions-Coal Ratio, 2006-2010, QTT

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.11	-0.01	0.01	-0.07	-0.13	-0.04	-0.02	-0.05	-0.07	-0.04
	(0.08)	(0.14)	(0.10)	(0.09)	(0.08)	(0.06)	(0.04)	(0.04)	(0.05)	(0.09)
2007	-0.11	0.54*	0.16	-0.10	-0.05	-0.03	-0.06	-0.11*	-0.18*	-0.38*
	(0.11)	(0.19)	(0.12)	(0.12)	(0.09)	(0.06)	(0.06)	(0.05)	(0.07)	(0.14)
2008	-0.08	0.32	0.10	-0.04	-0.16	-0.12	-0.20*	-0.24*	-0.25*	-0.36*
	(0.12)	(0.26)	(0.13)	(0.13)	(0.09)	(0.07)	(0.06)	(0.06)	(0.08)	(0.16)
2009	0.04	0.35	0.10	0.20	-0.01	-0.02	-0.12	-0.22*	-0.27*	-0.37*
	(0.14)	(0.28)	(0.14)	(0.14)	(0.10)	(0.08)	(0.07)	(0.07)	(0.09)	(0.18)
2010	-0.11	0.41	0.24	0.05	0.00	-0.08	-0.16*	-0.22*	-0.32*	-0.59*
	(0.15)	(0.32)	(0.20)	(0.16)	(0.13)	(0.09)	(0.08)	(0.08)	(0.10)	(0.20)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 21: Log SO_2 Intensity, 2006–2010, QTT, Input-Based Measured Emissions

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.00	0.00	0.00	-0.01	-0.00	-0.01	-0.00	-0.00	-0.01	0.01
	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)
2007	0.01	0.02*	0.01*	0.02*	0.01*	0.01	0.00	-0.00	-0.00	-0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
2008	-0.02	0.02*	0.01	0.00	-0.01	-0.02*	-0.03*	-0.03*	-0.04*	-0.14*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)
2009	-0.03*	0.03*	0.02*	-0.00	-0.01	-0.04*	-0.05*	-0.06*	-0.06*	-0.15*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)
2010	-0.03*	0.02	0.01	-0.01	-0.03*	-0.05*	-0.07*	-0.07*	-0.08*	-0.09*
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.04)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 22: Log SO_2 Intensity, 2006–2010, QTT, Inverse Abatement Intensity

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	-0.13	-0.16	-0.17	-0.09	-0.13	-0.09	-0.10	-0.21*	-0.23	-0.33*
	(0.10)	(0.11)	(0.17)	(0.13)	(0.15)	(0.15)	(0.12)	(0.10)	(0.13)	(0.17)
2007	-0.11	0.37*	0.27	0.02	-0.07	-0.03	-0.17	-0.34*	-0.46*	-0.62*
	(0.13)	(0.14)	(0.18)	(0.17)	(0.18)	(0.15)	(0.16)	(0.14)	(0.16)	(0.22)
2008	-0.08	0.33*	0.27	0.01	-0.08	-0.10	-0.13	-0.46*	-0.50*	-0.64*
	(0.13)	(0.15)	(0.21)	(0.18)	(0.19)	(0.17)	(0.16)	(0.15)	(0.16)	(0.22)
2009	-0.06	0.24	0.57*	0.17	-0.15	-0.14	-0.18	-0.41*	-0.64*	-0.72*
	(0.15)	(0.20)	(0.22)	(0.22)	(0.21)	(0.20)	(0.19)	(0.19)	(0.22)	(0.26)
2010	0.06	0.38	0.65*	0.32	0.07	0.05	-0.07	-0.37*	-0.53*	-0.49
	(0.15)	(0.22)	(0.23)	(0.22)	(0.23)	(0.20)	(0.18)	(0.19)	(0.20)	(0.30)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 23: Log Scrubber Capacity-Output Ratio, 2006–2010, QTT

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.004	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	(0.002)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
2007	0.001	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003	-0.003
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
2008	-0.001	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007	-0.007
	(0.007)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)	(0.004)
2009	-0.003	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006	-0.006
	(0.007)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)	(0.018)
2010	0.011	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002	-0.002
	(0.014)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 24: Log Inverse Coal Share of Emissions, 2006–2010, QTT

Year	ATE	Sample: Border Provinces QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.04	0.35	0.20	0.10	0.06	0.07	0.06	-0.02	-0.07	-0.15
	(0.09)	(0.24)	(0.13)	(0.10)	(0.09)	(0.07)	(0.09)	(0.09)	(0.10)	(0.14)
2007	0.09	0.65	0.24	0.19	0.12	0.09	0.02	-0.13	-0.23	-0.54
	(0.11)	(0.34)	(0.14)	(0.12)	(0.11)	(0.09)	(0.10)	(0.10)	(0.13)	(0.15)
2008	-0.14	0.53	-0.10	-0.17	-0.21	-0.22*	-0.33*	-0.51*	-0.71*	-0.78*
	(0.13)	(0.43)	(0.16)	(0.14)	(0.12)	(0.10)	(0.11)	(0.11)	(0.14)	(0.19)
2009	0.07	1.04*	0.16	0.06	-0.05	-0.10	-0.23	-0.54*	-0.74*	-0.79*
	(0.15)	(0.47)	(0.19)	(0.17)	(0.13)	(0.13)	(0.13)	(0.13)	(0.16)	(0.20)
2010	0.13	1.44	0.06	0.06	-0.09	-0.11	-0.31*	-0.47*	-0.63*	-0.63*
	(0.15)	(0.77)	(0.24)	(0.16)	(0.15)	(0.13)	(0.15)	(0.14)	(0.15)	(0.20)
Year	ATE	Sample: All Provinces QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.07	-0.29*	0.38	0.23*	0.17*	0.10	0.09	0.03	0.11*	0.13
	(0.07)	(0.12)	(0.29)	(0.10)	(0.07)	(0.06)	(0.06)	(0.05)	(0.05)	(0.07)
2007	0.09	0.80	0.23	0.11	0.09	0.01	-0.01	-0.04	-0.01	-0.14
	(0.08)	(0.46)	(0.14)	(0.09)	(0.09)	(0.07)	(0.06)	(0.06)	(0.07)	(0.09)
2008	-0.01	1.21*	0.28	0.13	0.01	-0.08	-0.15*	-0.27*	-0.35*	-0.56*
	(0.10)	(0.47)	(0.13)	(0.09)	(0.09)	(0.08)	(0.07)	(0.08)	(0.09)	(0.12)
2009	-0.22	-0.62*	-0.06	0.00	-0.06	-0.09	-0.11	-0.11	-0.31*	-0.48*
	(0.11)	(0.24)	(0.25)	(0.13)	(0.10)	(0.10)	(0.09)	(0.09)	(0.11)	(0.14)
2010	-0.09	-0.50	0.32	0.14	0.01	-0.02	-0.12	-0.12	-0.24*	-0.35*
	(0.12)	(0.27)	(0.35)	(0.14)	(0.11)	(0.10)	(0.10)	(0.10)	(0.11)	(0.14)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

Table 25: Leakage-Augmented SO_2 Emissions Intensity, 2006–2010

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.00	0.37	0.21	0.04	-0.02	0.07	0.01	-0.03	-0.05	-0.13
	(0.08)	(0.23)	(0.12)	(0.09)	(0.09)	(0.07)	(0.08)	(0.08)	(0.09)	(0.12)
2007	0.04	0.73*	0.21	0.13	0.03	0.07	-0.02	-0.17	-0.28*	-0.57*
	(0.11)	(0.28)	(0.12)	(0.12)	(0.10)	(0.09)	(0.10)	(0.11)	(0.13)	(0.15)
2008	-0.13	0.53	-0.04	-0.13	-0.24*	-0.24*	-0.37*	-0.49*	-0.66*	-0.81*
	(0.12)	(0.40)	(0.16)	(0.12)	(0.11)	(0.10)	(0.12)	(0.12)	(0.13)	(0.16)
2009	0.12	1.08*	0.14	0.08	0.01	-0.10	-0.28	-0.58*	-0.54*	-0.72*
	(0.14)	(0.46)	(0.18)	(0.15)	(0.14)	(0.13)	(0.14)	(0.14)	(0.15)	(0.18)
2010	0.04	1.258	-0.07	-0.10	-0.19	-0.19	-0.51*	-0.70*	-0.69*	-0.69*
	(0.15)	(0.69)	(0.23)	(0.17)	(0.15)	(0.13)	(0.16)	(0.14)	(0.17)	(0.20)

Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years.

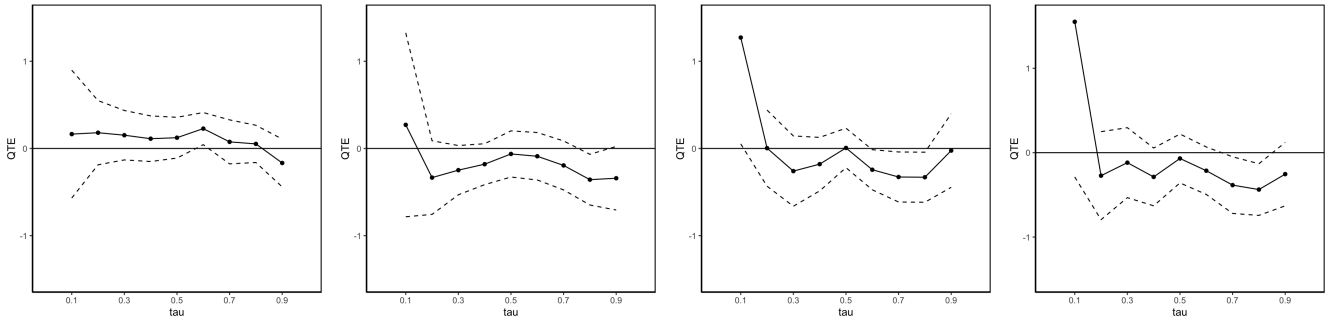
Table 26: Reallocation-Augmented SO_2 Emissions Intensity, 2006–2010

Year	ATE	QTT(τ)								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
2006	0.14	0.39	0.23	0.18	0.06	0.11	0.07	0.09	0.06	0.02
	(0.12)	(0.36)	(0.18)	(0.16)	(0.12)	(0.11)	(0.11)	(0.12)	(0.14)	(0.18)
2007	0.25	1.08*	0.39	0.40*	0.27	0.23	0.17	-0.01	-0.20	0.52*
	(0.15)	(0.55)	(0.21)	(0.18)	(0.14)	(0.12)	(0.13)	(0.15)	(0.16)	(0.19)
2008	0.03	1.44	-0.04	-0.03	-0.08	-0.07	-0.25	-0.37	-0.75*	-0.62*
	(0.18)	(0.76)	(0.26)	(0.20)	(0.16)	(0.15)	(0.16)	(0.19)	(0.17)	(0.23)
2009	0.13	1.60*	0.54	-0.01	0.07	-0.05	-0.23	-0.74*	-0.91*	-0.91*
	(0.20)	(0.80)	(0.32)	(0.21)	(0.17)	(0.16)	(0.18)	(0.19)	(0.22)	(0.28)
2010	0.37	2.72*	0.60	0.51*	0.28	0.12	-0.07	-0.32	-0.59*	-0.63*
	(0.23)	(0.93)	(0.38)	(0.25)	(0.21)	(0.20)	(0.19)	(0.20)	(0.23)	(0.30)

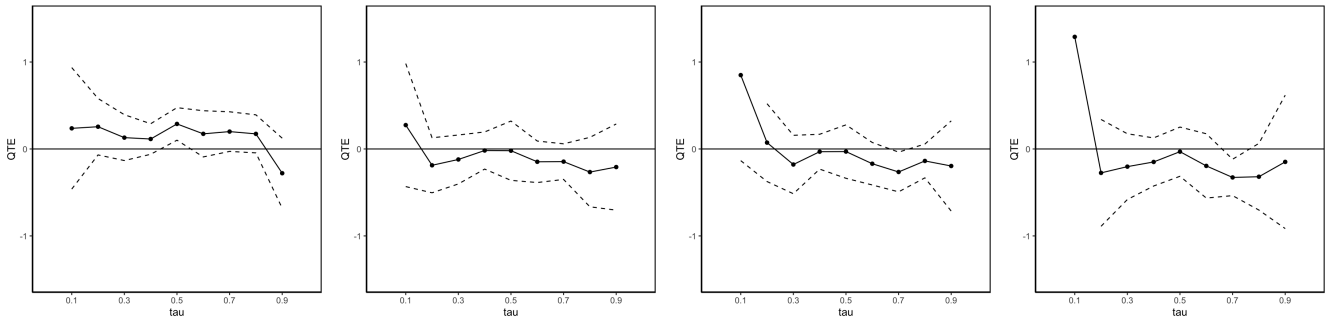
Notes: Pointwise standard errors are in parentheses. * indicates statistical significance at the 5 percent level. The years 2004 and 2005 are used as base years in all outcome years. A firm is deemed small if its 2004 revenues are below the median for its industry.

Table 27: Log SO_2 Intensity, 2006–2010, QTT, Small Firms Only

D Additional Figures



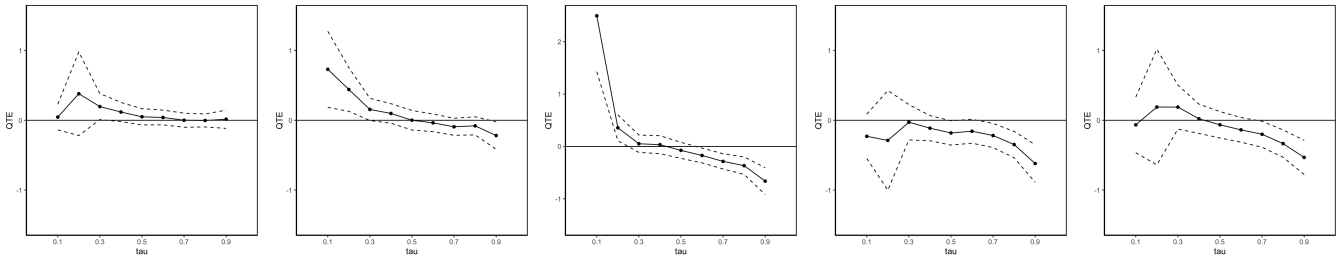
(a) QCLO, 2007 (b) QCLO, 2008 (c) QCLO, 2009 (d) QCLO, 2010



(e) QCIC, 2007 (f) QCIC, 2008 (g) QCIC, 2009 (h) QCIC, 2010

Notes: The figure contains estimates of quantile treatment effects for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations. Panels (a)–(d) report the QDID and panels (e)–(h) report the QCIC, across sample years.

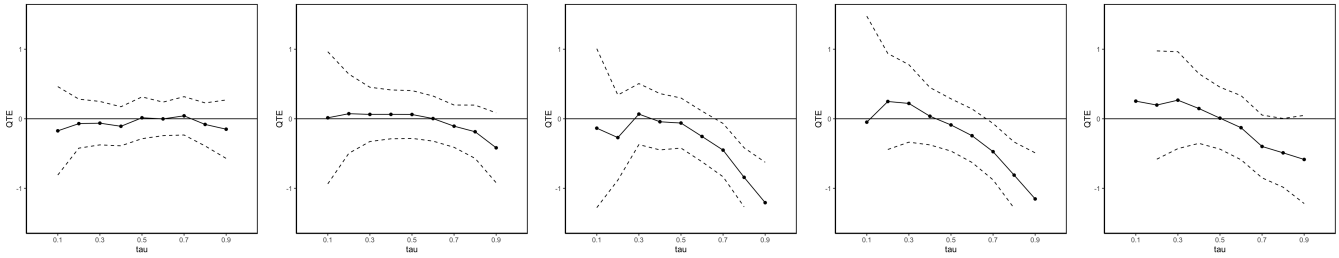
Figure 20: Log SO_2 Intensity, Alternative Estimation Approaches



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

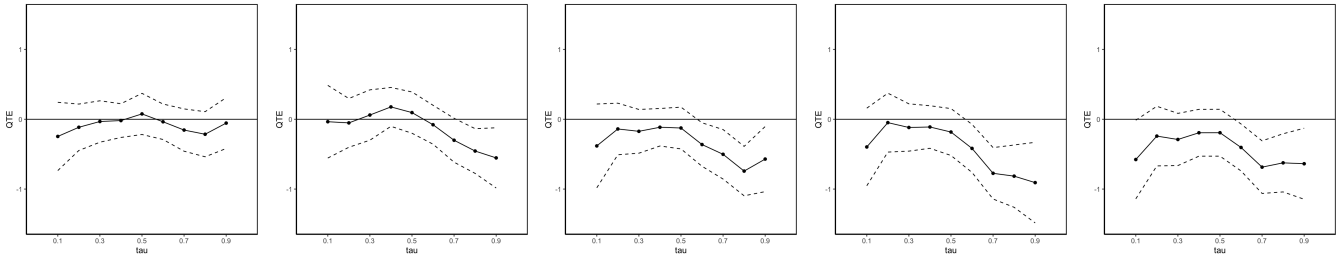
Figure 21: Log SO_2 Intensity, All Provinces



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in border cities that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

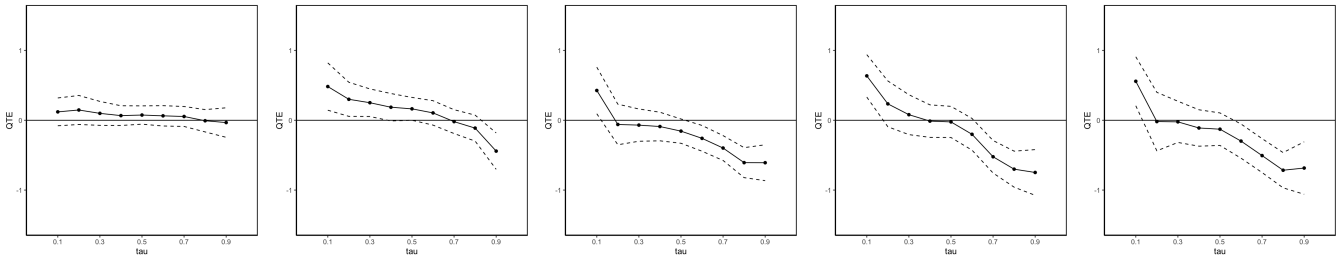
Figure 22: Log SO_2 Intensity, Border Cities



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

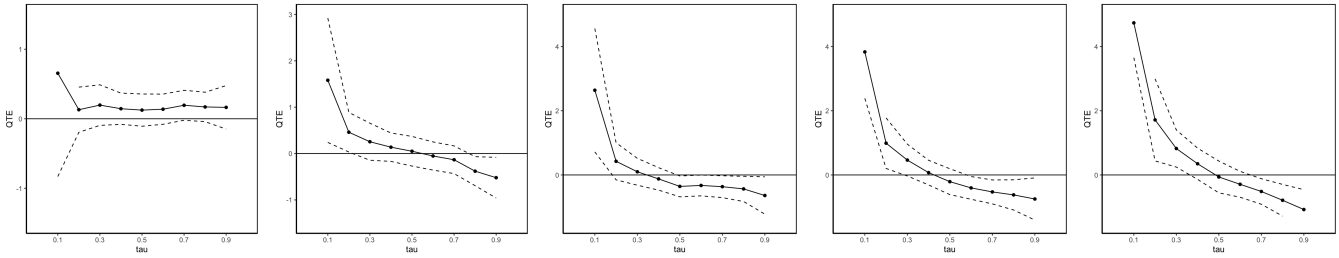
Figure 23: Log SO_2 Intensity, Balanced Panel



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

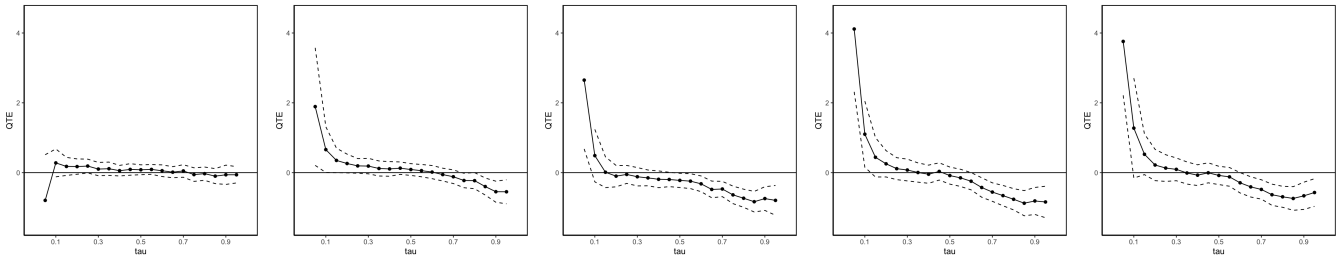
Figure 24: Inverse Hyperbolic Sine Transformation of SO_2 Intensity



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. Particulate matter (PM) intensity is used as the outcome variable. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

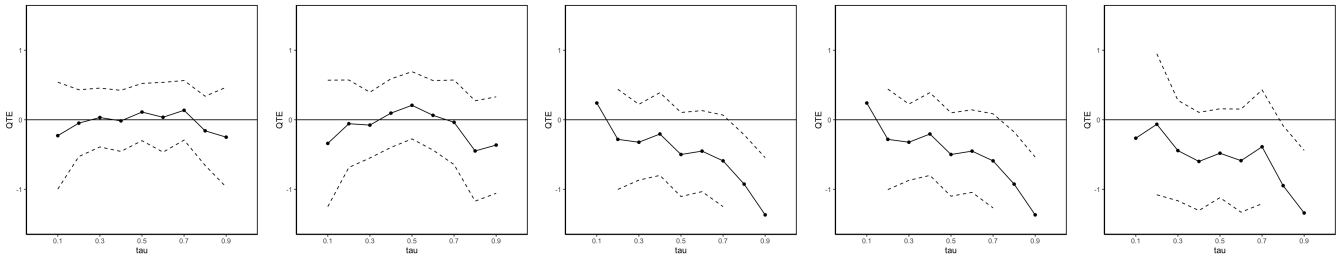
Figure 25: Log Particulate Matter Intensity



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

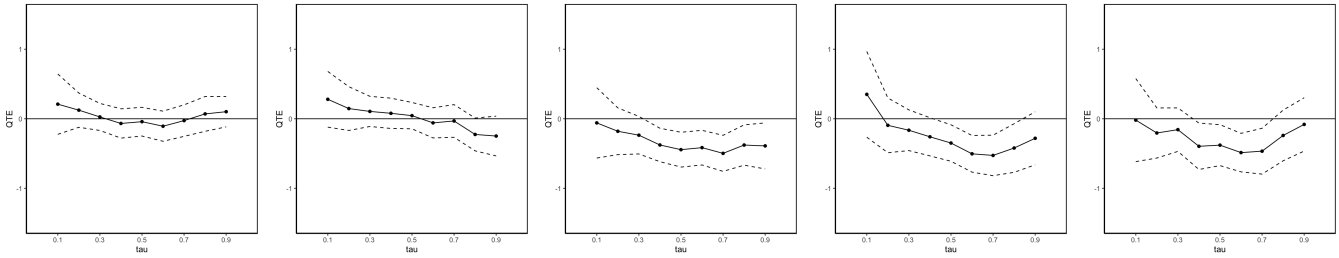
Figure 26: Log SO_2 Intensity, Finer Quantiles



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for chemical manufacturing firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

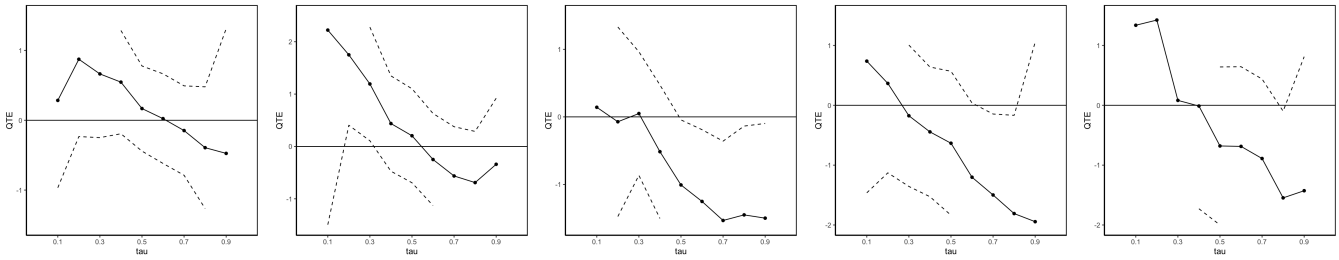
Figure 27: Log SO_2 Intensity, Chemical Manufacturing Only



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for non-metallic mineral firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

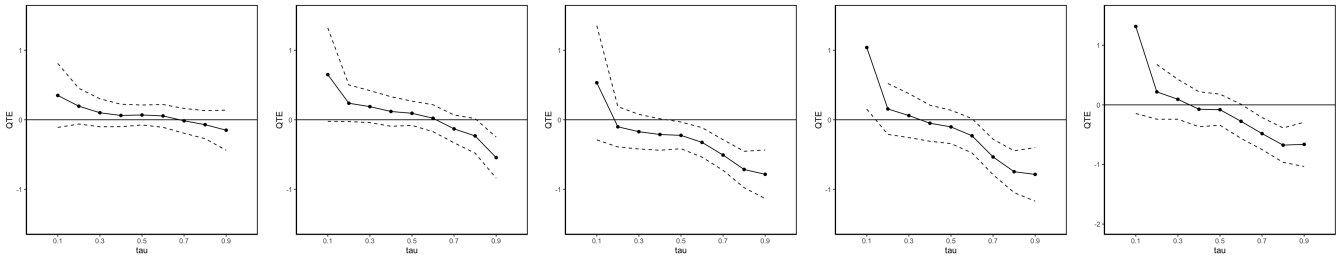
Figure 28: Log SO_2 Intensity, Non-Metallic Minerals Only



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for metal smelting and rolling firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

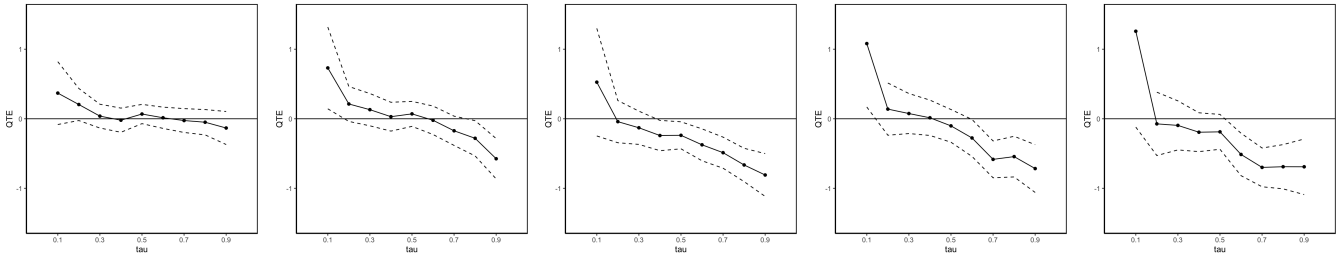
Figure 29: Log SO_2 Intensity, Metal Smelting and Rolling



(a) Benchmark, 2006 (b) Benchmark, 2007 (c) Benchmark, 2008 (d) Benchmark, 2009 (e) Benchmark, 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 30: Leakage-Augmented SO_2 Intensity, QTT



(a) 2006 (b) 2007 (c) 2008 (d) 2009 (e) 2010

Notes: The figure contains estimates of the QTT for firms in provinces that significantly increased SO_2 emissions reduction targets. 95 percent pointwise confidence intervals are computed using the empirical bootstrap with 1,000 iterations.

Figure 31: Product Reallocation-Augmented Log SO_2 Intensity, QTT