

# Effectiveness and cost of air pollution control in China

**Thomas Stoerk**

**November 2018**

**Grantham Research Institute on  
Climate Change and the Environment  
Working Paper No. 273**

**ISSN 2515-5717 (Online)**

**The Grantham Research Institute on Climate Change and the Environment** was established by the London School of Economics and Political Science in 2008 to bring together international expertise on economics, finance, geography, the environment, international development and political economy to create a world-leading centre for policy-relevant research and training. The Institute is funded by the Grantham Foundation for the Protection of the Environment and the Global Green Growth Institute. It has six research themes:

1. Sustainable development
2. Finance, investment and insurance
3. Changing behaviours
4. Growth and innovation
5. Policy design and evaluation
6. Governance and legislation

More information: [www.lse.ac.uk/GranthamInstitute](http://www.lse.ac.uk/GranthamInstitute)

**Suggested citation:**

Stoerk T (2018) *Effectiveness and cost of air pollution control in China*. Grantham Research Institute on Climate Change and the Environment Working Paper 273. London: London School of Economics and Political Science

This working paper is intended to stimulate discussion within the research community and among users of research, and its content may have been submitted for publication in academic journals. It has been reviewed by at least one internal referee before publication. The views expressed in this paper represent those of the authors and do not necessarily represent those of the host institutions or funders.

# EFFECTIVENESS AND COST OF AIR POLLUTION CONTROL IN CHINA\*

Thomas Stoerk<sup>†</sup>

November 21, 2018

## Abstract

I evaluate the effectiveness and cost of China's first serious air pollution control policy. Using both official, misreporting-prone data as well as NASA satellite data in a differences-in-differences strategy that exploits variation in reduction targets, I find that the policy reduced air pollution by 11% as intended. Compliance was initially rhetorical but later real, and did not differ by intensity of enforcement. I construct marginal abatement cost curves for  $SO_2$  for each province in China to calculate the cost of a counterfactual market-based policy instrument compared to the command-and-control policy that China used. I find that the market-based policy instrument would increase average (marginal) efficiency by 25% (49%). I further provide cost estimates for the total cost of a one unit decrease in  $PM_{2.5}$  concentrations in China to complement recent WTP estimates.

**JEL Codes:** Q52, Q53, H11

**Keywords:** Air pollution, China, abatement cost, instrument choice

---

\*I am indebted to Antoine Dechezleprêtre and Humberto Llavador as well as Geoff Barrows, Tim Besley, Robin Burgess, François Cohen, Dan Dudek, Denny Ellerman, Ruben Enikolopov, Carolyn Fischer, Roger Fouquet, Friedrich Geiecke, Matthieu Glachant, Ashley Gorst, Stephen Hansen, Martin Heger, Matthew Kahn, Nat Keohane, Felix Koenig, Gianmarco León, Kyle Meng, Niclas Moneke, Juan-Pablo Montero, Gerard Padró i Miquel, Steve Pischke, Giacomo Ponzetto, Alessandro Tavoni, Jeroen van den Bergh, Reed Walker, Guo Xu and Yimei Zou for valuable comments and conversations. Jia Yang provided outstanding research assistance. I thank Janusz Cofala and Robert Sander at IIASA, Nickolay Krotkov, Can Li and Chris McLinden at NASA, Qin Hu at EDF, and Litao Wang at Hebei University of Engineering for generous help in data access. I have further benefitted from comments by audiences in a variety of seminars and conferences. All errors are my own. This research presents my personal views and does not necessarily reflect the view of the European Commission.

<sup>†</sup>European Commission, Directorate-General for Climate Action, and London School of Economics and Political Science, Grantham Research Institute on Climate Change and the Environment. Contact: [t.a.stoerk@lse.ac.uk](mailto:t.a.stoerk@lse.ac.uk)

# 1 Introduction

Effective design and implementation of environmental regulation is crucial for correcting environmental externalities. Traditionally, economists have analyzed environmental regulation in developed countries where technical expertise, appropriate monitoring of pollution and rule of law often allowed successful cost-effective implementation of regulation. Recently, attention has turned to environmental regulation in developing countries and how the cost of regulation interacts with imperfect institutions (Duflo et al., 2013; Oliva, 2015). This shift in research focus is timely, as developing countries are often more severely affected by the most important environmental externalities such as air pollution. According to the latest WHO estimates, air pollution is responsible for one in eight global deaths, or 7 million deaths a year (WHO, 2014). One country which is particularly struck by air pollution is China. As development has soared, so has air pollution. Economic research on optimal air pollution control in China, however, is still in its infancy.

This study is the first to provide evidence on the effectiveness and cost of air pollution control in China. An active literature has recently provided comprehensive willingness to pay (WTP) estimates for reduced air pollution in China (Barwick et al., 2018; Freeman et al., 2017; Ito and Zhang, 2016). However, full benefit-cost analysis is hampered by a lack of comparable estimates on the cost of reducing air pollution.

To provide such estimates, I evaluate China's first serious air pollution control policy, a total emissions control target in the 11th Five-Year Plan (FYP) from 2006 to 2010. In an effort to bring down air pollution, the Chinese government decided to limit the total emissions of sulphur dioxide ( $SO_2$ ) by 10% relative to 2005. The national limit was later assigned by command-and-control into widely varying reduction targets for each province. This research uses a combination of unique datasets, microeconometrics and detailed marginal abatement cost curves to provide a comprehensive evaluation of the  $SO_2$  reduction policy along four margins: First, did the policy improve  $SO_2$  pollution outcomes? Second, how did the regulated provincial governments comply? Third, how costly was the policy and how efficient was it compared to a counterfactual market-based policy instrument? Fourth, what is the actual cost of a one unit decrease in  $PM_{2.5}$  concentrations in China?.

Greenstone and Jack (2015) suggest that one explanation for high pollution levels in developing countries might be the high cost of improving environmental quality at the margin. My setting is particularly relevant to investigate this conjecture. The 11th

FYP marks a turning point in environmental policy-making in China; it is considered 'the most environmentally ambitious document in the history of the Communist Party' (Watts, 2011). However, when the policy was passed in 2005, China's regulatory agency was the weak State Environmental Protection Administration (SEPA). SEPA did not have access to reliable  $SO_2$  pollution data in 2005 and had to implement the regulation based on limited information from  $SO_2$  emission statistics. This situation changed in 2008, when the central government upgraded SEPA to become the Ministry of Environmental Protection (MEP) allowed it to track  $SO_2$  pollution independent of provincial governments (State Council, 2007).

This empirical setting is insightful for several reasons. My setting is unique because of the availability of real pollution data in the period before the Chinese government could monitor it. This is due to coincidence: in late 2004, just before the start of the policy, NASA launched the *EOS-Aura* satellite that provides an independent and reliable data source for  $SO_2$  pollution in China. My setting also allows to study the cost of the policy in detail due to the availability of micro-level data on the cost of  $SO_2$  abatement in each province. I use these data to construct marginal abatement cost curves at the province level, allowing me to construct a detailed estimate for the cost of air pollution control in China.

This paper proceeds in two steps. First, I evaluate whether the  $SO_2$  control policy actually improved pollution outcomes despite the lack of regulatory capacity at the start of the 11th FYP. Exploiting variation across provinces and prefectures in a differences-in-differences (DID) specification, I recover the causal effect of the  $SO_2$  control policy on real pollution, measured through NASA satellite data. I then study whether the effect of the  $SO_2$  reduction target differs at the county level according to the initial distribution of pollution within the province. Finally, I investigate whether enforcement interacts with the targets.

I find that the policy was a success: a one-standard deviation increase in the stringency of the reduction target leads to a statistically significant 11% decrease in  $SO_2$  emissions as measured by the NASA satellite. Within a province, the estimated effect is stronger for counties that were initially more polluted. Combining a subsample of hand-collected prefecture-level data covering one third of China with data on the number of environmental enforcement officials, I find no evidence for heterogeneous treatment effects by

intensity of enforcement.

The second step of this research combines my empirical findings with detailed marginal abatement cost curves for each province in China, allowing me to estimate the actual cost of reducing air pollution in China by one unit. To further ask whether lower abatement costs are possible, I evaluate the efficiency of the policy design and quantify the gains from trade across different policy instruments. These curves show the large heterogeneity in  $SO_2$  abatement cost across the provinces of China. Based on the MAC curves, I find that command-and-control policy did not equate marginal abatement cost across space. Instead, the Chinese government favored reductions in coastal provinces in the East where abatement costs are higher. Using the MAC curves, I construct the counterfactual market-based allocation of  $SO_2$  reduction targets across provinces needed to achieve the 10%  $SO_2$  reduction target. This allows me to study the gains from trade from moving from a command-and-control regulation to the allocation of  $SO_2$  reduction targets that would result from a stylised emissions trading scheme across provinces. I find that the market-based allocation would increase efficiency by 25%, lowering the average abatement cost from \$437/t $SO_2$  to \$323/t $SO_2$ . At the margin, efficiency would rise by 49%, lowering marginal abatement cost from \$816/t $SO_2$  to \$419/t $SO_2$ . Combining my empirical and cost findings with an ex ante study in atmospheric science (Wang, Jang, et al., 2010b), I find that the cost of a 1  $\mu g/m^3$  reduction in  $PM_{2.5}$  concentrations is \$217,100, or 25% lower at \$161,997 using a market-based policy.

The paper is organised as follows: Section 2 discusses the related literature. Section 3 describes the policy setting. Section 4 explains my data sources, while Section 5 contains the empirical analysis. Section 6 constructs detailed marginal abatement cost curves at the province level to assess the cost of air pollution control and to compute the gains from trade across different policy instruments. Section 7 concludes.

## 2 Related literature

Air pollution in China is rampant, and it is man-made. The enormous health costs of air pollution in China are well documented by literatures in both economics and health. Chen, Ebenstein, et al. (2013), for instance, use a natural experiment to find that one coal-subsidy alone led to the loss of 2.5 billion life-years in Northern China. Epidemio-

logical studies summarized in Yang et al. (2013) give the same sense of magnitude: they find air pollution to be the fourth most important health burden in China. In monetized terms, the health cost amount to 1.2 to 3.8% of GDP (World Bank and State Environmental Protection Administration, 2007). Air pollution furthermore induces losses in productivity (Chang et al., 2016; Fu, Viard, and Zhang, 2018; He, Liu, and Salvo, 2018) and cognitive performance (Zhang, Chen, and Zhang, 2018). At the same time, Jia (2014) has shown in a convincing causal setting that pollution is a side effect of political incentives. A large literature in urban economics, political economy and environmental law backs this conclusion (Almond et al., 2009; Wang, 2013; Zheng and Kahn, 2013; Zheng, Sun, et al., 2014). Air pollution in China, therefore, is a problem that can in principle be solved through the right combination of policies and incentives. How to do so in practice, however, is far from resolved.

My study is the first to provide a causal empirical evaluation of the total emissions control policy in the 11th FYP. Despite the huge burden from air pollution in China, there has been no empirical evaluation of China's flagship air pollution control policy. Evaluation so far has come in one of two guises: through detailed narrative accounts of the changes (Cao, Garbaccio, and Ho, 2009; Hao et al., 2007; Schreifels, Fu, and Wilson, 2012) or through model-based studies in atmospheric science (Lu et al., 2010; Wang, Jang, et al., 2010a,b). Additionally, I am amongst the first to evaluate any environmental policy in China. The main other study I am aware of is Kahn, Li, and Zhao (2015), who analyze water pollution regulation.

This research also contributes to three distinct literatures in environmental economics. The first literature estimates the willingness to pay (WTP) for clean air in China. WTP estimates for environmental quality have traditionally focused on the U.S. (Chay and Greenstone, 2005; Deschenes, Greenstone, and Shapiro, 2017). Recently, the focus of WTP for clean air research has shifted to China, for which comprehensive WTP estimates now exist: Ito and Zhang (2016) exploit a policy discontinuity across space to recover WTP estimates for clean air from defensive investments in air purifiers. Freeman et al. (2017), by contrast, use an instrumental variable strategy to recover WTP estimates through a residential sorting model. Lastly, Barwick et al. (2018) recover a lower bound for consumer WTP for clean air in China based on healthcare spending, while Mu and Zhang (2017) estimate a lower bound based on defensive investment in facemasks.

Estimates for the cost of abating air pollution in China, however, are still lacking. This lack matters for policy: Greenstone and Jack (2015) hypothesize that one of the reasons environmental quality in developing countries is low is the high cost of improvements in environmental quality at the margin. By providing data on the actual cost of air pollution control, my research complements the WTP estimates from the literature to allow for credible cost-benefit analysis of air pollution control in China.

I further ask whether the cost of air pollution control in China could be reduced by better policy design. This question contributes to a second literature on the design of environmental regulation and the efficiency of command-and-control versus market-based policy instruments. This literature centers on air pollution control regulation in the U.S. (Carlson et al., 2000; Ellerman et al., 2000; Keohane, 2003, 2006; Oates, Portney, and McGartland, 1989; Schmalensee et al., 1998; Stavins, 1998). Carlson et al. (2000) compute marginal abatement cost curves for  $SO_2$  for the electricity sector in the U.S. to quantify the efficiency gains from trade of moving from command-and-control regulation to  $SO_2$  emissions trading. While those cost savings are large, at a 43% efficiency gain from trading, they are surprisingly lower than anticipated ex ante. Ellerman et al. (2000) find similar efficiency gains of 50%, while Keohane (2003) estimates only 16% to 25%. Another closely related paper is Oates, Portney, and McGartland (1989), who compare the efficiency of incentive-based regulation against command-and-control regulation to control air pollution in Baltimore. They find that a well designed command-and-control regulation can deliver pollution reductions at a welfare cost that can be lower than a comparable incentive-based regulation. Taken together, these studies show that while moving from a command-and-control regulation to a market-based regulation is generally seen as increasing the efficiency of the regulation (Schultze, 1977), whether this is so is an empirical question that depends on the particular case of the regulation under consideration.

I estimate detailed marginal abatement cost curves for  $SO_2$  for each province in China in 2005, contributing to the few studies that estimate full marginal abatement cost curves in environmental economics in general (Gollop and Roberts, 1985; Carlson et al., 2000 and Abito, 2012)<sup>1</sup>. In particular, this study is the first to derive comprehensive marginal

---

<sup>1</sup>Partial estimates of marginal abatement cost curves for compliance with air pollution control regulation are further reported in Hartman, Wheeler, and Singh (1997), Becker and Henderson (2001), Keller and Levinson (2002) and Becker (2005).



abatement cost curves at the province level in China. Two earlier contributions by Tu and Shen (2014) and Li, Wu, and Zhang (2015) provide interesting analysis in this direction but both studies rely on modelling in addition to microdata and only compute partial snapshots rather than full MAC curves. I use the marginal abatement cost curves to predict the counterfactual allocation of  $SO_2$  reduction targets across provinces in China that would result from an emissions trading scheme across provinces. This allows me to quantify the efficiency gains from trade from moving from the actual command-and-control allocation of  $SO_2$  reduction targets to a market-based allocation.

My research lastly complements a nascent literature that studies environmental regulation in developing countries (Duflo et al., 2013; Hansman, Hjort, and Leon, 2015; Oliva, 2015). My research sheds light on whether a simple command-and-control policy can be effective in a setting in which the government has a very low regulatory capacity initially. Due to the rich data availability, I can study the behaviour of regulated agents in terms of rhetorical and real compliance. I find that provincial governments strongly adjusted their rhetoric to the political goals from the center: a one-standard deviation increase in the province's  $SO_2$  reduction target leads to a 30% increase in political statements on air pollution, mainly driven by mentions of sulfur. Ultimately, however, compliance became real, mainly through the shutdown of small, inefficient power plants. I furthermore to the existing literature on the effectiveness of air pollution control, which has focused almost exclusively on developed countries such as the U.S. (Auffhammer and Kellogg, 2011; Chay and Greenstone, 2005; Chay, Dobkin, and Greenstone, 2003; Henderson, 1996)<sup>2</sup>. Additionally, while research on air pollution regulation has led to clear findings for pollutants such as TSP, the results for  $SO_2$  are more mixed (Greenstone, 2004; Greenstone and Hanna, 2014).

### 3 The $SO_2$ reduction policy in context

This section provides the context around the 11th Five-Year Plan (FYP) and China's flagship air pollution control policy that is the focus of this research. Environmental governance in China has undergone a rapid transformation in the last two decades. Until 2005, economic growth was the defining development paradigm. Environmental policies,

---

<sup>2</sup>One exception are Greenstone and Hanna (2014), who investigate the effect of regulation on air and water pollution in India.

where they existed, were paper tigers: they lacked political support from the central government and were rarely enforced. 2005 marks the turning point with a Five-Year Plan that 'was the most environmentally ambitious document in the history of the Communist Party' (Watts, 2011). The following paragraphs sketch how this change can best be seen as a change in political will rather than a change in formal laws.

**Before the policy** Laws regulating  $SO_2$  emissions have existed in China since 1998, when the State Council approved the establishment of the 'Two Control Zones', a policy to address acid rain and  $SO_2$  emissions (McElwee, 2011). Enforcement of this policy intensified in 2000, but has remained constant since. Implementation of  $SO_2$  policies, however, still encountered great difficulties (Gao et al., 2009), as China's overall development strategy remained firmly rooted in economic growth. Existing environmental policies overall, for instance on energy efficiency, were left underfunded by the central government (Gao et al., 2009; Lin, 2007). And while the 10th Five-Year Plan included a nationwide goal to reduce  $SO_2$  emissions by 10%, it did not have political backing, and failed to induce  $SO_2$  emissions reductions (Schreifels, Fu, and Wilson, 2012), possibly because of a lack of incentives for meeting the targets (Wang, 2013).

First change in the political outlook of the central government came in 2003, when President Hu - a hydroengineer - and Prime Minister Wen - a geologist - took power. The 'scientific development' paradigm, which emphasized environmental protection alongside economic growth, started to substitute for economic growth. Environmental governance, however, was still weak.

**2005: The 11th Five-Year Plan (2006-2010)** Amidst the increasing political will to implement and enforce environmental policies, the general directions of the 11th FYP started being discussed as early as mid-2003, and probably ended by 2004 (Xu, 2011). During the National People's Congress in March 2006, the 11th FYP was presented in its final form and included emissions control targets for air pollution ( $SO_2$ ) and water pollution (chemical oxygen demand, or COD) as well as a target on energy efficiency. Concurrently, environmental governance started being taken seriously, when the once powerless SEPA successfully stopped hundreds of billions of Yuan of industrial investment on environmental grounds at the beginning of 2005. This radical action came as a surprise to Chinese society (Gao et al., 2009). In March 2008, the SEPA's new political

authority was formalized when SEPA obtained full rank in the State Council and received ministry status as the MEP (McElwee, 2011).

The air pollution control target consisted of a 10%  $SO_2$  emissions control target for China as a whole. This reduction target was handed down to the provinces in May 2006 at the latest, when SEPA - with high-level political backing - signed formal, binding reduction targets with the provincial governments (Gao et al., 2009; Xu, 2011)<sup>3</sup>. These reduction targets were given the highest political priority, paralleled only by mandates on growth, social stability and the one-child policy (Wang, 2013). Provincial  $SO_2$  reduction targets, in particular, were made a veto target: failure to comply would nullify all other performance achievements of a provincial leader (Kahn, Li, and Zhao, 2015; Xu, 2011). The reduction targets specified reductions in  $SO_2$  emissions from 2006 to 2010 with 2005 as the baseline. These are the reduction targets that I use in this study. Figure 1 shows that these targets vary considerably, mandating reductions from 0% to more than 25%.

[INSERT FIGURE 1 ABOUT HERE]

$SO_2$  emissions data in China in 2005 were of bad quality, and misreporting-prone. A province's  $SO_2$  emissions were calculated by aggregating the physical quantity of coal used in a province in a given year, and by then multiplying this quantity with  $SO_2$  emissions factors depending on the sulfur-intensity of the type of coal used. The political authority to compile these data rested with the provincial governments. While this procedure yields only coarse aggregate data at best (Guan et al., 2012), it can also be corrupted to produce the desired data. Anecdotal evidence and extensive field work in several provinces confirm that misreporting of  $SO_2$  emissions data was prevalent in the first 2 years of the  $SO_2$  control policy (Song et al., 2015).

Up to 2005, provincial governments in China had little incentive to control air pollution. A province in China is governed by a pair of provincial leaders, the governor and the party secretary. Provincial leaders are career officials who are appointed in a top-down manner. As career officials, they are often positioned outside their native provinces and move frequently as a result of promotion or demotion. Cadre regulations stipulate a maximum term length of 5 years for both governors and party secretaries (Kahn, Li, and Zhao, 2015), and this rule is mostly respected. As a consequence, provincial leaders are prone to short-termism that furthers their own career. Following the regulations in the

---

<sup>3</sup>The 31 province-level administrative units include the four municipalities directly under the central government (Beijing, Chongqing, Shanghai and Tianjin) and the five autonomous regions (Guangxi, Neimongol, Ningxia, Tibet and Xinjiang).

past, this meant reliance on pollution-intensive, quick-and-dirty GDP growth (Chen, Li, and Zhou, 2005; Jia, 2014; Li and Zhou, 2005). Incentives for provincial leaders to fix air pollution are further weakened by vested interests in polluting enterprises, of which local governments are often major shareholders (Gao et al., 2009). As provincial leaders got promoted to posts elsewhere, air pollution therefore remained unfixed.

**2008: Changes to  $SO_2$  monitoring** In 2007, the State Council passed a law that fundamentally changed the regulator’s capacity to monitor  $SO_2$  pollution (the law is known as the ‘Reduction of the Three Ways’). The core of the law was twofold: to change the politics of  $SO_2$  data collection to avoid tampering with data at the political level, and to install appropriate monitoring equipment and ensure frequent statistical inspections on the ground (State Council, 2007).

On the political side, reporting was taken from provincial governments and put directly under the political control of the MEP. The MEP, in turn, directly reports to the State Council. On the ground,  $SO_2$  measurement stations were built in pollution hotspots and the number of environmental monitoring officials was increased by 17% (Song et al., 2015). Key industrial polluters for each prefecture had their  $SO_2$  emissions tracked on site. By May 2008, uninterrupted automatic monitoring devices with data feeds directly into the local environmental agency were used for this purpose (McElwee, 2011; MEP, 2008a,b). All changes became effective in July 2008 at the latest (Song et al., 2015).

## 4 Data

I compile a unique dataset that allows me to study the effect of the  $SO_2$  control policy along two margins: First, to evaluate the effect of the policy on pollution outcomes, I use two different data sources: (i) the official, misreporting prone  $SO_2$  emissions indicator and (ii) independent satellite data from NASA. Second, I evaluate the reactions by the regulated provincial governments. I divide these reactions into rhetorical compliance and real compliance. To measure rhetorical compliance, I build a novel dataset of political reports, which I quantify. Behaviour related to real compliance is based on official data sources on  $SO_2$  abatement measures. Finally, I add data on the number of enforcement officials per province to measure the effect of enforcement. I first describe the pollution data, then the data related to the behaviour of the regulated provincial governments, and finally the enforcement official data.

## 4.1 Data on pollution outcomes

**Official  $SO_2$  data** The  $SO_2$  control policy relies on  $SO_2$  *emissions* as the official indicator. Note that this is a proxy indicator for the ultimate goal of reducing pollutant *concentrations*. It is likely that the central government chose this indicator due to a combination of a legacy of state-planning that focused on total emissions control and a lack of suitable  $SO_2$  concentrations data from *in situ* measurement stations. I use the data reported in the *China Energy Databook* (Fridley, Romankiewicz, and Fino-Chen, 2013), who compile the official  $SO_2$  emissions data from the statistical yearbooks.

The official  $SO_2$  emissions data gives at best a noisy picture of true  $SO_2$  emissions (Guan et al., 2012), and anecdotal evidence and research based on fieldwork suggest that the incentives before the 2008 policy changes led to severe misreporting (Song et al., 2015). The literature has also noted more general misreporting of air pollution data in China (Chen, Jin, et al., 2012; Ghanem and Zhang, 2014; Karplus, Zhang, and Almond, 2018; Stoerk, 2016).

**$SO_2$  satellite data** To measure real  $SO_2$  pollution, I make use of the uniqueness of my empirical setting. In August 2004 - just before the start of the  $SO_2$  control policy - NASA launched a satellite with the Ozone Monitoring Instrument (OMI). In lay terms, OMI is an instrument that captures images of Earth from space at different wavelengths. Post-processing through extraction algorithms produces  $SO_2$  vertical columns of high precision that became available in 2014<sup>4</sup>. It is unlikely that the Chinese government would have had access to the satellite data during the policy, and these data were not used in the official evaluation of the policy. Dates with cloud cover can lead to missing values over individual pixels, but this is not generally considered a first-order problem (Krotkov et al., 2016)<sup>5</sup>. All in all, the NASA  $SO_2$  satellite data are a very good proxy for ground-level  $SO_2$  emissions. To illustrate, Figure 2 shows the cross-section in January 2006.

[INSERT FIGURE 2 ABOUT HERE]

---

<sup>4</sup>The new algorithm significantly improved the precision of the extracted vertical column densities and removed a number of biases compared to the earlier data product that was available from OMI (Krotkov et al., 2016). Note further that the OMI data itself have a detection threshold that is two magnitudes smaller than earlier satellite data and can thus enable the detection of  $SO_2$  pollution from human activity in the lowest part of the atmosphere (NASA, 2014).

<sup>5</sup>Furthermore, I aggregate daily values to at least monthly, and pixels of not bigger than 0.25 degrees by 0.25 degrees latitude-longitude to counties and provinces, further reducing the magnitude of the potential problem due to clouds.

**Relation between both data sources**  $SO_2$  emissions decay in a span of 4-36 hours (Fioletov et al., 2015; He, 2012). Since the satellite data capture Earth daily, they represent a snapshot of  $SO_2$  pollution on that day. Given the quick decay, this prevents leakage from confounding the outcome and allows me to capture local rather than transported  $SO_2$  emissions. NASA’s retrieval algorithm produces four different data products, each of which corresponds to  $SO_2$  pollution at different levels of altitude in the atmosphere (NASA, 2014). For this analysis, I use the lowest level at an altitude of 900m above ground, for two reasons: first, because this is the best proxy for anthropogenic emissions sources, and secondly, because a lower altitude further minimizes transportation. Second, the precision of the satellite images is high enough to identify individual sources of pollution that produce as little as 30kt of  $SO_2$  annually (Fioletov et al., 2015). I aggregate these daily cross-sections to the province-month and the province-year levels. Annual changes in the  $SO_2$  satellite data can therefore be expected to be mimicked closely in the official statistics<sup>6</sup>.

## 4.2 Data on the number of environmental enforcement officials in a province

This section describes the data on the number of enforcement officials for each province in China. The dataset comes from an NGO that has trained 34,887 enforcement officials since 2006 for the Ministry of Environmental Protection (since renamed to Ministry of Ecological Environment). Overall, this represents 44% of the China total. Since further information on the province of the remaining enforcement officials is unavailable, I assume that the enforcement official data I have is an accurate proxy for the number of enforcement officials in a province at a given time. This assumption is reasonable in practice because the NGO knew the total number of officials in a province and used this criterion to avoid unduly influencing the cross-province distribution of officials. My data show that training of enforcement officials was scarce in 2006 and 2007, and picked up from 2008 on, concurrent with the upgrade of the Ministry and the rollout of monitoring technology to measure  $SO_2$ .

Given the large differences in the size of China’s provinces, I compute a normalized number of enforcement officials by dividing the number of officials by the province area to account for the time cost of travelling. This density of enforcement officials is my main

---

<sup>6</sup>A cross-check of known ground-level emission sources with  $SO_2$  OMI satellite data revealed a correlation of 0.91 (Fioletov et al., 2015).

measure for enforcement<sup>7</sup>.

## 5 Empirical analysis

### 5.1 Empirical strategy

**Baseline specification** I use the following difference-in-differences model to investigate whether a higher  $SO_2$  reduction target led to a relatively stronger decrease in  $SO_2$  emissions:

$$y_{pt} = \beta_0 + \beta_1 \text{Reductiontarget}_p \times D(\text{Post})_t + \beta_2 \text{Reductiontarget}_p + \sum_{t=1}^T \beta_{3t} \gamma_t + \alpha_p + u_{pt} \quad (1)$$

The outcome  $y_{pt}$  for province  $p$  at time period  $t$  is either the official  $SO_2$  emissions data that was used by the central government to assess the policy or the independent satellite  $SO_2$  data. The variable  $\text{Reductiontarget}_p$  is the provincial  $SO_2$  reduction target and captures the cross-sectional variation shown in Figure 1. Policy variation over time is captured in the indicator  $D(\text{Post})$  that takes on the value of 1 from 2006 onwards.  $\alpha_p$  are province fixed effects. The estimate for  $\beta_1$  gives the causal effect of an increase in the target stringency by one unit for the whole period of the 11th Five-Year Plan, from 2006 to 2010.

**Identification** The DID specification exploits cross-province variation in the stringency of the  $SO_2$  reduction target to estimate the causal effect of the pollution control policy. Identification relies on a combination of three factors: (i) common trends in the outcome variables prior to the  $SO_2$  control policy, (ii) a sharp deviation from those trends following the policy changes, and (iii) the absence of forward-looking considerations that would also explain  $SO_2$  abatement efforts by the provincial governments in 2006.

Provincial governments can rely on three main channels to bring down pollution. All of which are quick to implement, in particular for a country like China: fuel-switching to higher quality coal with a lower sulfur content, installation of desulfurisation devices, and the shutdown of small, inefficient thermal units. It is therefore reasonable to expect an immediate effect of the  $SO_2$  reduction policy on pollution outcomes. Lu et al. (2010) provide evidence that abatement measures started immediately: power plants already started to switch to better quality coal in 2005 (with a sulfur content reduced by about

---

<sup>7</sup>Alternative density measures - such as the number of enforcement officials per tonne of  $SO_2$  emissions - correlate strongly with area, and yields similar results.



20% compared to the preceding year), and flue-gas desulfurization technology doubled from below 10% to more than 20% of all operating power plant capacity.

Common pre-trends and the timing of effects are empirically testable, and I show below that these conditions are fulfilled. Consideration (iii) is not directly testable, but supporting evidence shows that it is likely fulfilled. While the exact algorithm used by the Chinese government to allocate the targets is unknown, and it is unlikely that the targets were distributed randomly, random allocation is not needed for my identification. Instead, I only need that the allocation of  $SO_2$  reduction targets across provinces was independent of forward-looking considerations that would explain  $SO_2$  pollution reductions by a province independent of the  $SO_2$  reduction targets.

The official statement by the State Council on how the target distribution would have taken place mentions a whole array of factors that were used to determine the allocation of targets for a province (State Council, 2006): (i) environmental quality and environmental capacity, (ii) current pollution levels, (iii) level of economic development, (iv)  $SO_2$  mitigation capabilities and (v) regional differentiation (Eastern, Central, Western). Xu (2011) finds that the allocation of targets does not correlate with either of those factors (barring a correlation of non-power  $SO_2$  emissions divided by the area of a province). Therefore, there is no evidence that any of these factors drove the target allocation, which suggests that the allocation of targets followed guidelines orthogonal to changes in  $SO_2$  emissions at the turn of the 11th FYP. I also rule out the possibility that the allocation of  $SO_2$  reduction targets followed the cost of abatement, both by itself and net of benefits. If that were the case, my empirical strategy would pick up the compound effect of an  $SO_2$  reduction target and a cost advantage. In Section 6, I derive detailed marginal abatement cost (MAC) curves and combine those with a measure of marginal abatement benefits. I find that neither the MAC nor the marginal welfare impacts correlate with the target allocation, lending further credibility to my empirical strategy. Appendix A.3 validates this claim in more detail.

**Inference** I follow Bertrand, Duflo, and Mullainathan (2004) and compute standard errors that are heteroskedasticity-robust and clustered at the province level. Statistical inference based on these standard errors, however, could still be incorrect if the number of clusters is too small, as the required asymptotic results might not apply. China has 31 provinces, yielding substantially fewer than 50 clusters, the usual rule of thumb.



While reporting the clustered standard errors, I therefore base my statistical inference on p-values derived from the wild bootstrap method described in Cameron, Gelbach, and Miller (2008). This is common practice in applied research on China (e.g., Martinez-Bravo et al., 2017).

## 5.2 The effect of the $SO_2$ control policy on $SO_2$ pollution

**Baseline results** Table 1 provides the summary statistics for my sample. The first two columns of Table 2 show the results from estimation Equation (1) for the effect of the policy for the full 11th FYP. A first glance reveals that the policy had a different effect depending on the indicator used for evaluation. According to the official  $SO_2$  emissions data, the policy was a success:  $SO_2$  emissions decrease in response to the target, with the estimated magnitude being a 5.8% decrease for a one-standard deviation increase in target stringency. The satellite data, by contrast, do not show a significant relationship between the targets and the  $SO_2$  pollution outcomes. The sign of the estimated coefficient is in the same direction as with the official indicator, and the estimated magnitude is even higher, but it is not statistically different from zero.

To improve the precision of the estimates, I increase statistical power by focussing on polluted cities only. Given the same amount of noise in the data, a higher absolute effect can be expected to be more easily detected in this sample. The sample of polluted cities is built by taking the location of each *in situ* measurement stations run by the MEP<sup>8</sup>. The point estimate for the effects of a higher reduction target is nearly identical to the overall sample in percentage terms, but the coefficient turns significant because the higher absolute effect in the polluted sample increases statistical power (Table 2). All baseline estimates taken together, I find that that the  $SO_2$  control policy was effective in reducing air pollution over the whole sample period.

**Dynamic treatment effects** To take a closer look at what might explain these differing findings, I estimate yearly versions of the above DID specification. Instead of collapsing the time periods into before and after the policy as in Equation (1), I interact a

---

<sup>8</sup>I select the 25 nearest pixels up to a distance of at most 25 kilometers around the city centroids, and compute the province-level observations only based on those pixels. Results from this sample are robust to changes in the 25 kilometers cutoff, as most pixels are closer than 15 kilometers from the city centroid.

dummy for each time period with the reduction target. This yields the following equation:

$$y_{pt} = \beta_0 + \sum_{t=1}^T \beta_{1t} \text{Reductiontarget}_p \times \gamma_t + \beta_2 \text{Reductiontarget}_p + \sum_{t=1}^T \beta_{3t} \gamma_t + \alpha_p + u_{pt} \quad (2)$$

As before,  $y_{pt}$  are the  $SO_2$  pollution outcomes for province  $p$  in time period  $t$ . In this specification, the  $SO_2$  reduction targets are interacted with each time period  $\gamma_t$  to estimate the differential trajectory of  $SO_2$  pollution for provinces with different  $SO_2$  reduction targets. These effects are captured in the point estimates for  $\beta_{1t}$  for periods 2 through  $T$ , where  $T$  is the last observation for 2010.

Figure 3 plots the estimate of the interaction coefficients  $\sum_{t=1}^T \beta_{1t}$  in equation (2) and show that the identifying assumption of common pre-trends is satisfied. Both  $SO_2$  data sources show a common trend before the start of the policy in 2006. However, there is only one year of satellite data before the start of the policy, because NASA only launched the satellite in late 2004. Figure 4 zooms in on the satellite data and plots the estimates for  $\sum_{t=1}^T \beta_{1t}$  estimated on monthly satellite data. While there is more noise in the monthly data, there are common pre-trends before the start of the policy, as illustrated through the vertical grey lines.

[INSERT FIGURES 3 and 4 ABOUT HERE]

The  $SO_2$  satellite data paints a clear picture: the effect of a higher reduction target on  $SO_2$  emissions only sets in once the central government gains the ability to monitor pollution in 2008. Appendix A.1 collects data on how provincial governments comply. I show that compliance was initially rhetorical and turned real once provincial governments shut down small, inefficient power plants.

**Heterogeneous treatment effects based on initial pollution levels** Where do the improvements in air pollution take place? As shown in Figure 2,  $SO_2$  pollution differs strongly across space. Sichuan, for instance, suffers from a pollution hotspot towards the East, and enjoys comparatively lesser pollution in the Western part of the province. The NASA satellite data allow me to exploit this heterogeneity to ask whether the effect of the  $SO_2$  pollution reduction target in the province differs depending on the initial level of pollution. To answer this question, I map the  $SO_2$  satellite data to the county-level, yielding 2,638 cross-sectional units. For each county, I measure its mean  $SO_2$  pollution for 2005 relative to all other counties within the same province. This information is

captured in a variable that takes the value of 1 for counties in the lowest quartile of initial pollution within their province, up to a value of 4 for counties that have the highest initial pollution. I then re-estimate the DID specification from Equation (1) at the county-yearmonth level on each subsample along the distribution of initial pollution.  $\sum_{t=1}^T \gamma_t$  therefore are year-month fixed effects in this subsample. As above, statistical inference relies on heteroskedasticity standard errors at the province-level and t-statistics from a wild bootstrap procedure with 1000 repetitions. I expect to find no effects for the lowest quartile, because initial pollution in these counties is below 0.2 Dobson Units on average, making further air quality improvements unlikely. Furthermore, for the same precision of data, a nominally smaller effect is harder to detect statistically, making it more likely to find a significant effect the higher the initial level of pollution.

The results in Table 4 confirm that this is the case. As expected,  $SO_2$  reductions in response to the targets mainly take place in the higher two quartiles of the initial distribution, though only the most polluted quartile is statistically significant. This reproduces the baseline results, where the effect only turned statistically significant for the sample of polluted cities due to noise. The effect size for the two highest quartiles is a 9.4% and 9.9% decrease in  $SO_2$  pollution per one standard-deviation increase in the stringency of the reduction target, respectively.

### 5.3 The effect of enforcement

Figure 3 showed that the effect of the targets sets in from 2008 on, after the Central Government gained the ability to accurately track pollution. Research from the U.S. has shown that enforcement can have an important influence on how polluters respond to air pollution limits (Blundell, 2017). In this section, I therefore ask whether the effect of a reduction target on pollution outcomes depends on the level of enforcement in a province. In other words, I estimate heterogeneous treatment effects by level of enforcement.

**Selection of the prefecture subsample** To overcome the low number of cross-sectional units in my province-level sample, I estimate the role of enforcement amplifying the effects of the target in a new sample at the prefecture level. The prefecture is China's second administrative level after the province. Compared to the province level with its 31 units, the prefecture level features 334 units.

My sample contains hand-collected reduction targets for 123 different prefectures, a

little over a third of China’s prefectures. This sample was collected from a variety of sources such as provincial government webpages through a significant data collection effort<sup>9</sup>. I am confident that this sample is as comprehensive as can be collected: The four municipalities directly under the Central Government do not report prefecture-level targets. Additionally, five of China’s provinces containing another 124 prefectures were given a reduction target of zero (Gansu, Hainan, Qinghai, Tibet, Xinjiang), and therefore did not use prefecture-level targets. Figure 5 shows the full prefecture-level sample.

To gauge how representative this sample is of the general sample, it is useful to ask: what are the incentives to report a prefecture-level target publicly? This depends on who reports the target: the prefecture itself would likely self-select to publicize its target only if it has been able to comply. A province, on the other hand, would publicize the targets of all its constituent prefectures simply to show the Central Government and the public that it is taking air pollution seriously. Luckily, the large majority of my subsample data comes from province-wide lists, and is thus unlikely to involve a selection issue.

To substantiate this claim, I compare the distribution at the prefecture level to the distribution of the province-level targets. Both distributions are very similar in shape, as well as quantitatively: the subsample has a mean reduction target of 14.43 (SD: 10.90) compared to the full sample’s mean reduction target of 9.65 (SD: 6.71). The reason for the slightly higher mean in the subsample is that some prefectures received targets up to 53.6% compared to at most 25.9% for a province (see Figure A.1 in the Appendix for a plot of both histograms). Both distributions, therefore, are very similar.

**Estimation** To analyze the effect of enforcement, I use the density of enforcement officials for each province in the subsample shown in Figure A.2 in the Appendix. I then split the prefecture-level sample into quartiles according to the enforcement official density. As can be seen from the distribution, the upper quartile shows a comparatively large standard deviation of enforcement official density while the lower three quartiles are more homogeneous in their enforcement official density.

The rollout of  $SO_2$  monitoring devices and the large-scale employment of enforcement officials started at the same time, from 2008 on. I exploit the timing of this change together with the cross-sectional variation in the number of enforcement officials (normalized by province area). In this way, I can study heterogeneous treatment effects depending on the

---

<sup>9</sup>A full list of primary sources is available upon request.

number of enforcement officials in a province. For each quartile, I re-estimate a prefecture-level version of equation 1:

$$y_{ipt} = \beta_0 + \beta_1 \text{Reductiontarget}_i \times D(\text{Post1})_t + \beta_2 \text{Reductiontarget}_i \times D(\text{Post2})_t + \beta_3 \text{Reductiontarget}_i + \sum_{t=1}^T \beta_{4t} \gamma_t + \alpha_i + u_{it} \quad (3)$$

where  $i$  denotes a prefecture in province  $p$  in year  $t$ .  $D(\text{Post1})_t$  is a dummy for the beginning of the 11th Five-Year Plan, 2006 to 2007, whereas  $D(\text{Post2})_t$  is a dummy that captures the rollout of monitoring devices and enforcement officials from 2008 onwards. Apart from these ex ante considerations, the reason to study the effect of enforcement officials separately from 2008 onwards are my dynamic estimates of the effects of the targets. As shown in table 3, the effect of the reduction target on pollution outcomes sets in only from 2008 onwards. To study how the presence of enforcement officials influences the effect of a reduction target on pollution outcomes, it is thus useful to study the compliance period by itself.

**Results** Table 5 shows the results from estimating equation 3. Overall, the evidence is mixed. The prior on enforcement officials is that they are useful. That is, ex ante one would expect that the higher the density of enforcement officials, the larger the reduction effect of a target. The point estimates for the compliance period for the lower three quartiles point in this direction: the effect size for the prefectures with the lowest enforcement density is 0.5% of the pre-treatment mean, or close to zero. The point estimates for the second and third quartiles are  $-6.4\%$  and  $-8.4\%$ , respectively. Only the estimate for the third quartile is statistically different from zero. The results for the fourth quartile show the opposite, with a 7.6% increase in pollution, with the point estimate marginally significant. Overall, I can find no evidence for heterogeneous treatment effects by differences in enforcement.

## 6 Instrument choice and gains from trade

In this section I evaluate two questions of policy design for the  $SO_2$  reduction targets in the 11th FYP: what allocation of targets across provinces would a market-based policy instrument have created? And: what are efficiency gains from trade associated with moving from the command-and-control allocation to the counterfactual market-based policy?

Both questions bear enormous policy importance for China.  $SO_2$  trading was initially discussed as an alternative to a command-and-control policy in the 11th FYP. Though ultimately discarded, China has now embraced emissions trading for  $CO_2$  in its 8 pilot markets and in the upcoming national carbon market (Stoerk, Dudek, and Yang, 2018).

Moreover, both types of questions have been studied for the U.S. (Carlson et al., 2000; Oates, Portney, and McGartland, 1989). This research has shown that the efficiency of different policy instruments is an empirical question. It will depend, among other things, on whether the command-and-control regulation is designed in an enlightened way that takes the cost of pollution abatement into account. Whether the efficiency of China's flagship air pollution control regulation could have been increased through the use of a different policy instrument is therefore of academic interest.

In this section, I construct detailed marginal abatement cost (MAC) curves for  $SO_2$  at the province level for China. I use these  $SO_2$  MAC curves to predict the counterfactual market-based allocation. In this way, I can assess whether the actual command-and-control regulation was enlightened and took into account the cost of abatement. Furthermore, I can quantify the efficiency gains from trade from moving from the command-and-control allocation of reduction targets compared to the market-based allocation. As noted by Stavins (2003), cost might be the best measure to assess policy efficiency. Finally, I compute a back-of-the-envelope measure for the marginal benefits of  $SO_2$  abatement at the province level to study whether the gains from trade based on cost alone are driven by omitting the benefits of reducing air pollution.

## 6.1 Construction of the marginal abatement cost curves

**Data** I use a rich set of micro data on  $SO_2$  emissions and abatement costs in each province in China at a very detailed level. These data are compiled by the IIASA research institute for use as input into their GAINS model for China (on the model, see IIASA, 2010a, IIASA, 2010b, Klimont et al., 2009). The data rely on a variety of sources of two kinds: *common data*, that are used across different countries and rely on the assumption of free international markets in abatement equipment, and *country-specific data*. Common data include the unit investment cost for technologies, fixed costs of operation, and the amount of input factors needed for some of the variable cost components. These data have been compiled and updated by IIASA for several decades on the basis of expert meetings at the UN (AP EnvEcon, 2010).

Country-specific data, on the other hand, include a detailed breakdown of China’s industrial structure: the type and size of polluting installations, the facility operating conditions, national fuel consumption data, local input prices (labour, electricity, fuel cost) as well as unabated emission factors and removal efficiencies (AP EnvEcon, 2010). The local data are compiled by IIASA experts in collaboration with local experts from the Chinese Energy Research Institute in Beijing and Tsinghua (Purohit et al., 2010). These data are combined into unit cost estimates per technology, as well as abatement potential<sup>10</sup>.

The data give a detailed breakdown of all  $SO_2$  emission sources for each province, split into different sectors (such as the combustion of coal) that use different fuels (such as gas or low-sulfur coal) and each of which has different abatement technologies at its disposal (such as limestone injection). Overall, the dataset contains 2170 distinct sector-fuel combinations, and 5680 different abatement technologies. Each of the abatement technologies is characterised by a unit cost of abatement (in  $\$/tSO_2$ )<sup>11</sup> and its abatement potential (i.e. how much would the emissions factor of the current sector-fuel combination be lowered when switching to the abatement technology). To illustrate, one abatement technology would be the use of limestone-injection (*abatement technology*) in a modern coal-fired power plant (*fuel* and *sector*) at the cost of  $\$639.07/tSO_2$  (*unitcost of abatement*).

My data also allow me to look at sectors individually. The main  $SO_2$ -emitting sectors are power plants (63.4% of China’s  $SO_2$  emissions), industry (24.8%) and the residential sector (12.2%). The power sector, in particular, is of interest because it allows for an additional robustness check for the quality of the abatement cost data. Figure A.6 in the

---

<sup>10</sup>While the data on the cost of abatement are ex ante cost estimates for China, they are not pure engineering estimates. Instead, the common part of the abatement cost data rely on the ex post experience of other countries. The data are likely to be an upper bound for the ex post abatement cost: cheaper local inputs and China’s capacity for scale likely allowed for cheaper abatement in practice. The use of identified ex post MAC estimates from revealed-preference settings such as Meng (2017) and Gosnell, List, and Metcalfe (2016) would be desirable, but no such estimates are available for China or at the required level of comprehensiveness. My ex ante estimates, by contrast, are consistent across provinces and thus allow for cross-province comparisons. To err on the conservative side for abatement cost estimates also has a second advantage: it avoids underestimating the cost side in the cost-benefit calculus for the overall regulation. Regulatory practice, such as the EPA’s subsequent assessments of the Clean Power Plan, supports this.

<sup>11</sup>The original datasource is in €, which I convert into \$ at the contemporaneous exchange rate of 1.24.

Appendix depicts the power sector’s MAC curve in China to show that over 90.8% of the  $SO_2$  emissions from China’s power sector can be abated before cost become prohibitive. This level of possible abatement, as well as the general shape of the MAC curve, reflect the earlier  $SO_2$  abatement experience in the U.S. (Ellerman et al., 2000, Figure 9.1).

**Construction of the MAC curves** I first construct two simplified versions of the full MAC curve: a MAC curve that only uses the least cost abatement technology per province, fuel and sector (*least cost*), and a MAC curve that solely relies on the abatement technology per province, fuel and sector that offers the largest emissions abatement (*highest abatement*). To construct the *least cost* MAC curve for one province, I follow a two-step procedure. First, I rank each abatement option by the unit cost of abatement within each sector and fuel. Second, I abate  $SO_2$  emissions within each sector and fuel in increasing cost order across the province. I follow an analogous procedure to construct the *highest abatement* MAC curve.

In principle, construction of the full MAC curve for a given province requires an optimization that trades off both dimensions (see Appendix A.4). In practice, the difference between just using the least cost MAC and the optimised MAC is negligible for the comparatively low levels of abatement that were part of the 11th Five-Year Plan. I therefore take a computational shortcut to draw the full MACs: I construct it as the lower envelope of the *least cost* and the *highest abatement* curves. For low levels of abatement, the least cost curve dominates, whereas higher levels of abatement can be reached using the highest abatement technologies.

Figure 6 shows two examples for full MAC curves, Gansu and Sichuan, to illustrate the large heterogeneity in marginal abatement cost across provinces (Appendix A.5 contains the full MAC curves for each province). In contrast to many MAC studies, I do not rely on a top-down model but base the cost estimates entirely on data, and I am the first to provide complete MAC curves on Chinese provinces in this way.

[INSERT FIGURE 6 ABOUT HERE]

**Consistency checks** To check that my data represent the official Chinese data well, Figure 7 shows the average ratio between the GAINS  $SO_2$  emissions and the official MEP  $SO_2$  emissions as a function of the province’s level of emissions: the overall fit between my data and the MEP data is good (correlation: 85.9%) and fairly stable across provinces,



except for two outliers with very low emissions: Tibet and Hainan. Overall, the GAINS data report higher emissions for all provinces than do the official data. This is in line with the literature in atmospheric science that has found that GAINS data tend to be more comprehensive and slightly overpredict official  $SO_2$  data sources, which is possibly due to differing assumptions on the distribution of fuel consumption across sectors (Klimont et al., 2009) and the fact that official MEP statistics lack rural pollution sources and biofuels (Lu et al., 2010).

[INSERT FIGURE 7 ABOUT HERE]

To gauge whether my abatement cost data accurate, I compare my dataset to a firm-level analysis from China in 2005. It turns out that the use of scrubbers in the power sector, for instance, cost between \$84 and \$916 per ton of abated  $SO_2$  in 2005, depending on the quality of the coal used Wang, Pan, and Peng (2005). This is in line with abatement cost estimates for the U.S. from Ellerman et al. (2000), who find an average cost of scrubbing of \$265 per ton of abated  $SO_2$ . It is certainly possible that the cost of abatement came down over the course of the 11th Five-Year Plan through innovation and economies of scale from wider deployment. The magnitude of my ex ante cost curves, however, reflects the reality on the ground.

Additionally, there is a drawback in using detailed  $SO_2$  emissions data at the microlevel: the detailed breakdown across sectors does not allow me to compute the cost of moving activity across sectors (such as moving electricity generation from coal-fired power plants to wind turbines). In other words, these data limitations make it necessary to assume scrappage cost and imperfect substitution across sectors in the short run. Given the 5-year horizon of the policy, however, I feel that this is a reasonable assumption to make.

## 6.2 The counterfactual market-based allocation of $SO_2$ reduction targets

In this subsection, I compute the counterfactual allocation of  $SO_2$  reduction targets across provinces for the market-based policy. I compute the market-based policy-instrument as the cost-optimal allocation of targets to provinces given the national target. The market-based policy instrument thus represents a simplified emissions trading scheme in which provinces trade emissions allowances until their marginal abatement cost are equalized. To compute the counterfactual allocation for the market-based policy instrument, I pool

the province-level microdata and construct the  $SO_2$  marginal abatement cost curve at the national level for China. As shown in Figure 8, marginal abatement cost is relatively flat until an abatement level of 35% of 2005  $SO_2$  emissions. Up until this abatement level, the cost of  $SO_2$  abatement at the margin is less than \$500/t $SO_2$ . Beyond ca. 70% of 2005  $SO_2$  emissions, however, the picture changes and the cost of abatement rises rapidly. Beyond 75%, abatement becomes prohibitively costly.

[INSERT FIGURE 8 ABOUT HERE]

Figure 8 also shows the counterfactual marginal abatement cost for the national  $SO_2$  reduction target of 10%, given by the intersection between the vertical line at 10% and the MAC curve. At an abatement level of 10%, the market-based allocation of  $SO_2$  reduction targets would have led to a marginal abatement cost of \$419/t $SO_2$ . This Figure is the counterfactual marginal cost of the Chinese government's  $SO_2$  emissions control strategy had the central government distributed the provincial reduction targets in a cost-optimal way.

The intersection of the marginal abatement cost curve for China as a whole with the 10% national abatement level also produces the market-based allocation of  $SO_2$  reduction targets across provinces. Figure 9 shows that this allocation is far more skewed than the distribution of targets that was actually used in the 11th FYP. The command-and-control allocation already ranges from targets of 0% to 25.9%, but is approximately uniformly distributed within this range. The market-based allocation, by contrast, is more unequal. A small number of provinces would bear most of the reductions. These provinces are Sichuan, Shandong, and Zhejiang. The industrial structure of these provinces allows for comparatively cheap installation of wet flue-gas desulfurisation in industry and power plants and the use of more efficient combustion processes in refineries and steel sintering.

[INSERT FIGURE 9 ABOUT HERE]

Based on the MAC curves, I find that the Chinese government did not equate marginal abatement cost across space. Instead, the reduction targets targeted coastal provinces in the East even though abatement costs are higher at the margin (shown in Table A.2 and Figure A.7 in the Appendix). The actual command-and-control allocation is consistent with a tale-of-two-cities story, in which China would develop amenity-based consumer cities along the coast, while maintaining a base of polluting manufacturing in its interior (Kahn, 2006; Zheng and Kahn, 2013). Shanghai is a prime example for this: under the

cost-efficient allocation, Shanghai would have received an  $SO_2$  reduction target of 6.2% on 2005 levels. Under command-and-control, however, Shanghai received a reduction target of 25.9%, or more than four times the cost-efficient target.

My findings show that actual command-and-control regulation that China used to control  $SO_2$  pollution in the 11th FYP was not cost-optimal. As suggested by Oates, Portney, and McGartland (1989), command-and-control regulation will only be efficient if it is designed in an enlightened fashion by keeping an eye on the cost of abatement. Figure 10 shows that this was not done by the Chinese government in 2005. There is no statistically significant relationship between the  $SO_2$  reduction target a province received under the 11th FYP and its abatement cost at the margin.

[INSERT FIGURE 10 ABOUT HERE]

### 6.3 A measure for marginal abatement benefits

Because China's provinces differ markedly with respect to income levels, population densities and initial pollution, I include a measure for the benefits of air pollution abatement as a robustness check. I use the method employed by Oliva (2015) to construct a back-of-the-envelope measure for the marginal abatement benefits of reducing  $SO_2$  pollution at the province level. This method proceeds in 3 steps: (i) how does the  $SO_2$  control policy change pollutant concentrations?, (ii) what health effects do the changes in pollutant concentrations cause? and (iii) what is the monetary value of those health effects?

**(i) Changes in pollutant concentrations** I use the results from Wang, Jang, et al. (2010b), who use the CMAP modelling system (maintained by the U.S. Environmental Protection Agency) to simulate the effects of the 11th FYP's  $SO_2$  reduction policy on concentrations of  $SO_2$  and  $PM_{2.5}$  under the assumption of full compliance. My empirical analysis shows that this assumption is well-founded. The measurements from Wang, Jang, et al. (2010b) allow me to attribute changes in pollutant concentrations of  $SO_2$  and  $PM_{2.5}$  to the  $SO_2$  emissions reduction target of each province. Their estimates are net of spatial spillovers.

**(ii) Health effects** To convert the changes pollutant concentrations to changes in health outcomes, I use dose-response estimates for  $PM_{2.5}$  and  $SO_2$  from Bombardini and

Li (2016). They estimate an elasticity of infant mortality rates of 0.9 to  $SO_2$  and of 2.2 for  $PM_{2.5}$ . I combine those estimates with data on  $SO_2$  and  $PM_{2.5}$  levels in 2005 from MEP and the *China Energy Databook* (Fridley, Romankiewicz, and Fino-Chen, 2013) to approximate a linear dose-response function for each pollutant.

I use the estimates from Bombardini and Li (2016) for two reasons: first, they use an instrumental strategy approach to estimate a dose-response function for the health effects of air pollution, thus correcting downward bias from OLS estimates (due to migration, income effects and avoidance behaviour). Additionally, their study is from China, from a recent period, and includes consistent estimates for both  $SO_2$  and  $PM_{2.5}$ , which are the main pollutants that are affected by the  $SO_2$  reduction policy. The downside is that this restricts my focus on infant mortality when calculating the benefits from reducing air pollution.

Infant mortality, however, is likely to capture a first-order welfare effect. Matus et al. (2012) calculate that 71.4% of all air pollution costs in China are health costs, and that mortality captures over 85% of those health costs. Chen, Ebenstein, et al. (2013), in turn, show that mortality impacts from TSP are strongest in infants. Evidence from Indonesian wildfires in 1997 also points to large infant mortality effects from exposure to particulates, mostly driven by prenatal exposure (Jayachandran, 2009). Greenstone and Hanna (2014) also focus on infant mortality to evaluate the effect of pollution. Furthermore, data on infant mortality is of high quality and available for all regions of China, allowing me to answer questions of allocative efficiency across space<sup>12</sup>.

**(iii) Valuation** To convert the health damages into monetary values, I use a baseline value of a statistical life (VSL) of 1 million yuan from World Bank and State Environmental Protection Administration (2007). This value is a midpoint between the VSL estimates from the reviewed studies ranging from 0.24 to 1.7 million yuan<sup>13</sup>. Following Hammitt and Robinson (2011), I account for the income heterogeneity across Chinese

---

<sup>12</sup>Many more benefits exist to reducing air pollution, e.g. reduced adult morbidity and mortality (Barwick et al., 2018), increased productivity (Chang et al., 2016; Fu, Viard, and Zhang, 2018; He, Liu, and Salvo, 2018), as well as improvements in cognitive performance (Zhang, Chen, and Zhang, 2018) and mental health (Chen, Oliva, and Zhang, 2018). The literature is active, however, and the verdict on the full social cost of air pollution is still out. To be able to make statements about the optimal allocation of reduction targets across space, however, data on the full benefits is not needed. A shortcut is to assume that the damages from air pollution are proportional to population, which is an acceptable simplifying assumption given that all studies focus on effects in humans.

<sup>13</sup>VSL estimates for China are available for more recent periods (Ito and Zhang, 2016), but to evaluate a policy from 2005 I prefer to use estimates from that period.

provinces by adjusting the central VSL estimate according to the income level in each province using an income elasticity of VSL of one. This yields VSL estimates from 360,000 yuan (Guizhou) to 3,000,000 yuan (Shanghai).

Taking into account the reductions in concentrations of both  $SO_2$  and  $PM_{2.5}$ , I multiply these VSL estimates by the mortality numbers to compute the benefit of reducing  $SO_2$  emissions by 1kt for each province. Since I employ a linear approximation, the marginal benefits of abatement is constant<sup>14</sup>.

Then, in a similar vein to the MAC data, I can use the combined data on both marginal cost and benefit to construct marginal welfare impact curves. I combine the marginal abatement benefit data with the marginal abatement cost data by dividing the marginal abatement cost by the marginal abatement benefit to obtain a measure of the marginal welfare impact of abatement. When this ratio is below one, benefits are larger than cost. Once this ratio exceeds one, costs are higher than benefits. Figure 11 pools the province-level data to obtain the marginal welfare impacts for China as a whole, where the cost of abatement is normalized by the benefits within each province before pooling the data. The vertical dashed line at 0.1 marks the 10% national  $SO_2$  emissions control target of the 11th FYP. The marginal benefits of abatement exceed the marginal cost of abatement by 12 times at this abatement level, and this calculus abstracts from a number of benefits from reductions in air pollution.

**Caveats** I rely on an overly conservative measure of benefits by including only infant mortality which may underestimate the true benefits and lead to an underestimation of the welfare-optimizing level of  $SO_2$  abatement for China. However, this concern is muted in practice since even my lower bound benefit measure suggests  $SO_2$  abatement up until the prohibitive marginal abatement cost ranges. Additional abatement benefits at the margin would thus only have a negligible effect on the welfare-optimizing abatement level. The advantage is that the dose-response functions used are more precise for infants because low migration translates into better knowledge of lifetime exposure to pollution, improving the consistency of estimates *across* provinces.

---

<sup>14</sup>For those provinces with a reduction target of zero, I find that the benefits are 0 or nearly 0, too. This is due to low initial pollution, low VSL estimates, and low population numbers. The exception is Gansu, for which I compute the marginal benefits as the average of its 5 nearest neighbours with respect to initial pollution, VSL estimates, and population numbers.

## 6.4 Gains from trade

Finally, I study the gains from trade across the two different policy instruments. Gains from trade will be possible if the actual command-and-control regulation used to allocate the  $SO_2$  reduction targets in the 11th FYP was not done optimally with respect to the cost of abatement. As I have shown above, this is the case. This subsection quantifies these efficiency gains in terms of cost-efficiency. To further take into account the heterogeneities across Chinese provinces, I construct a back-of-the-envelope measure for the marginal benefits of air pollution abatement to study the robustness of the findings based on pure cost-efficiency.

Table 6 shows the gains from trade that are possible. Firstly, I follow the literature (e.g. Stavins, 2003) and assess the gains from trade using the cost efficiency measure. I find that moving from the command-and-control allocation to the market-based allocation would decrease abatement cost by 49% at the margin ( $\$816/tSO_2$  to  $\$419/tSO_2$ ). Secondly, I find that this conclusion is robust to taking into account the benefit side. Adding the benefit measure, I find efficiency improvements of 45% (from a welfare ratio of 2.19 to a welfare ratio of 1.2).

Overall efficiency would increase by 25%, lowering average abatement cost from  $\$437$  per ton of  $SO_2$  for the command-and-control policy to  $\$323$  per ton of  $SO_2$  for the market-based policy. Efficiency gains of 25% might seem low compared to the literature from the U.S. that finds efficiency gains up to 50% (Carlson et al., 2000; Ellerman et al., 2000; Keohane, 2003). However, a crucial difference in the analyses is that my market-based policy instrument focuses on the province rather than the firm level. The potential inefficiency compared to using plant-level data is therefore lower, and a 25% efficiency gain considerable.

## 6.5 Abatement cost expressed per unit of pollution

For policy, the cost of achieving a certain pollution outcome rather than a change in pollution inputs often matter. This indicator is particularly useful to compare the cost of reducing air pollution to willingness to pay (WTP) estimates. The WTP for clean air in China, in particular, is an active research topic, with important contributions by Ito and Zhang (2016), Freeman et al. (2017) and Barwick et al. (2018) and Mu and Zhang (2017). These estimates are typically expressed as a per unit decline in concentrations of

particulate matter.

This indicator reflects the typical policy goal, to reduce *concentrations* of pollutants rather than their *emissions*. To go from reductions in emissions to reductions in concentrations, I combine my empirical findings with an ex ante study in atmospheric science (Wang, Jang, et al., 2010b). Assuming full compliance with the 11th FYP's  $SO_2$  reduction targets, Wang, Jang, et al. (2010b) use an atmospheric science model from the U.S. EPA to compute the resulting improvements in pollutant concentrations across China, net of transport and spatial spillovers. Given that I find near universal compliance in my empirical evaluation, their estimates are likely to reflect the true reduction in pollutant concentrations as a result of the  $SO_2$  reduction policy I analyze in this research.

Wang, Jang, et al. (2010b) find that the policy reduced  $PM_{2.5}$  concentrations by 3 to 15  $\mu g/m^3$ , or 9  $\mu g/m^3$  on average. These changes are calculated for the most populated areas in China. Based on my MAC curves, the total cost of the  $SO_2$  abatement in the 11th FYP was \$1,953,898. The cost of a 1  $\mu g/m^3$  reduction in  $PM_{2.5}$  concentrations is, therefore, \$217,100. Note that these figures abstract from co-benefits such as lowered concentrations in pollutants other than particulate matter. Had China used a market-based policy instrument with a total cost of \$1,457,971, the cost of a 1  $\mu g/m^3$  reduction in  $PM_{2.5}$  concentrations would have been 161,997, or 25% lower. These figures can be used directly to address whether environmental quality in developing countries is low because of high marginal cost for improvements, as posited by Greenstone and Jack (2015).

## 7 Concluding remarks

I evaluate the effectiveness and cost of China's first serious air pollution control policy. Using both official, misreporting-prone data as well as NASA satellite data in a differences-in-differences strategy that exploits variation in reduction targets, I find that the policy reduced air pollution by 11% as intended. Compliance was initially rhetorical but later real in the form of increased shutdowns of small, inefficient thermal units. Combining a subsample of hand-collected prefecture-level data covering one third of China with data on the number of environmental enforcement officials, I find no evidence for heterogeneous treatment effects by intensity of enforcement. Compliance, however, only started when the central government upgraded its pollution monitoring capacity in 2008.

I further estimate the cost of a 1  $\mu g/m^3$  reduction in  $PM_{2.5}$  concentrations for China at \$217,100, or 25% less using a market-based policy instrument (49% less at the margin).



Such cost estimates are crucial for complementing the recent literature on WTP for clean air in China (Ito and Zhang, 2016; Freeman et al., 2017; Barwick et al., 2018). The air pollution abatement cost estimate based on my empirical analysis shows that high marginal abatement cost are unlikely to explain the high levels of air pollution in China. Contrary to Greenstone and Jack (2015)’s conjecture, therefore, the reason for high levels of pollution appears to be a lack of ambition in the design and implementation of policy. As its biggest environmental policy ambition to date, China has now started to build a national carbon market. Abatement cost loom amongst the biggest concerns in practice.

## References

- Abito, J. M. (2012): “Welfare Gains from Optimal Pollution Regulation”, Job Market Paper.
- Almond, D., Y. Chen, M. Greenstone, and H. Li (2009): “Winter Heating or Clean Air? Unintended Impacts of China’s Huai River Policy”, *American Economic Review (Papers and Proceedings)* 99(2), pp. 184–190.
- AP EnvEcon (2010): *GAINS Abatement Cost Guide. Version 1: A synthesis of information relating to the estimation, handling and use of abatement cost variables in the GAINS model framework.*
- Auffhammer, M. and R. Kellogg (2011): “Clearing the Air? The Effects of Gasoline Content Regulation on Air Quality”, *American Economic Review* 101, pp. 2687–2722.
- Barwick, P. J., S. Li, D. Rao, and N. B. Zahur (2018): “The Morbidity Cost of Air Pollution: Evidence from Consumer Spending in China”, NBER Working Paper No. 24688.
- Becker, R. A. (2005): “Air pollution abatement costs under the Clean Air Act: evidence from the PACE survey”, *Journal of Environmental Economics and Management* 50(1), pp. 144–169.
- Becker, R. A. and J. V. Henderson (2001): “Cost of Air Quality Regulation”, *Behavioral and Distributional Effects of Environmental Policy*, ed. by C. Carraro and G. E. Metcalf, Chicago: University of Chicago Press, pp. 159–186.
- Bertrand, M., E. Duflo, and S. Mullainathan (2004): “How Much Should We Trust Differences-in-Differences Estimates?”, *Quarterly Journal of Economics* 119(1), pp. 249–275.
- Blundell, W. (2017): “When Threats Become Credible: A Natural Experiment of Environmental Enforcement from Florida”, Mimeograph.
- Bombardini, M. and B. Li (2016): “Trade, Pollution and Mortality in China”, NBER Working Paper No. 22804.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008): “Bootstrap-Based Improvements for Inference with Clustered Errors”, *Review of Economics and Statistics* 90(3), pp. 414–427.
- Cao, J., R. Garbaccio, and M. S. Ho (2009): “China’s 11th Five-Year Plan and the Environment: Reducing  $SO_2$  Emissions”, *Review of Environmental Economics and Policy* 3(2), pp. 231–250.
- Carlson, C., D. Burtraw, M. Cropper, and K. L. Palmer (2000): “Sulfur Dioxide Control by Electric Utilities: What Are the Gains from Trade?”, *Journal of Political Economy* 108(6), pp. 1292–1326.
- Chang, T., J. Graff Zivin, T. Gross, and M. Neidell (2016): “The Effect of Pollution on Worker Productivity: Evidence from Call-Center Workers in China”, NBER Working Paper No. 22328.
- Chay, K. Y. and M. Greenstone (2005): “Does Air Quality Matter? Evidence from the Housing Market”, *Journal of Political Economy* 113(2), pp. 376–424.



- Chay, K., C. Dobkin, and M. Greenstone (2003): “The Clean Air Act of 1970 and adult mortality”, *Journal of Risk and Uncertainty* 27(3), pp. 279–300.
- Chen, S., P. Oliva, and P. Zhang (2018): “Air Pollution and Mental Health: Evidence from China”, NBER Working Paper No. 24686.
- Chen, Y., H. Li, and L.-A. Zhou (2005): “Relative performance evaluation and the turnover of provincial leaders in China”, *Economics Letters* 88(3), pp. 421–425.
- Chen, Y., A. Ebenstein, M. Greenstone, and H. Li (2013): “Evidence on the impact of sustained exposure to air pollution on life expectancy from China’s Huai River policy”, *Proceedings of the National Academy of Sciences* 110(32), pp. 12936–12941.
- Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2012): “Gaming in Air Pollution Data? Lessons from China”, *B.E. Journal of Economic Analysis & Policy* 12(3).
- Deschenes, O., M. Greenstone, and J. S. Shapiro (2017): “Defensive Investments and the Demand for Air Quality: Evidence from the NOx Budget Program”, *American Economic Review* 107(10), pp. 2958–89.
- Dufo, E., M. Greenstone, R. Pande, and N. Ryan (2013): “Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India”, *Quarterly Journal of Economics* 128(4), pp. 1499–1545.
- Ellerman, A. D., P. L. Joskow, R. Schmalensee, J.-P. Montero, and E. M. Bailey (2000): *Markets for Clean Air. The U.S. Acid Rain Program*, Cambridge, UK: Cambridge University Press.
- Fioletov, V. E., C. A. McLinden, N. Krotkov, and C. Li (2015): “Lifetimes and emissions of SO<sub>2</sub> from point sources estimated from OMI”, *Geophysical Research Letters* 42(6), pp. 1969–1976.
- Freeman, R., W. Liang, R. Song, and C. Timmins (2017): “Willingness to Pay for Clean Air in China”, NBER Working Paper No. 24157.
- Fridley, D., J. Romankiewicz, and C. Fino-Chen (2013): *China Energy Databook Version 8.0*, Lawrence Berkeley National Laboratory.
- Fu, S., V. B. Viard, and P. Zhang (2018): “Air Pollution and Manufacturing Firm Productivity: Nationwide Estimates for China”, SSRN Working Paper.
- Gao, C., H. Yin, N. Ai, and Z. Huang (2009): “Historical Analysis of SO<sub>2</sub> Pollution Control Policies in China”, *Environmental Management* 43, pp. 447–457.
- Ghanem, D. and J. Zhang (2014): “‘Effortless Perfection:’ Do Chinese cities manipulate air pollution data?”, *Journal of Environmental Economics and Management* 68(2), pp. 203–225.
- Gollop, F. M. and M. J. Roberts (1985): “Cost-Minimizing Regulation of Sulfur Emissions: Regional Gains in Electric Power”, *Review of Economics and Statistics* 67(1), p. 81.
- Gosnell, G. K., J. A. List, and R. Metcalfe (2016): “A New Approach to an Age-Old Problem: Solving Externalities by Incenting Workers Directly”, NBER Working Paper No. 22316.
- Greenstone, M. (2004): “Did the Clean Air Act cause the remarkable decline in sulfur dioxide concentrations?”, *Journal of Environmental Economics and Management* 47(3), pp. 585–611.
- Greenstone, M. and R. Hanna (2014): “Environmental Regulations, Air and Water Pollution, and Infant Mortality in India”, *American Economic Review* 104(10), pp. 3038–3072.
- Greenstone, M. and B. K. Jack (2015): “Envirodevonomics: A Research Agenda for an Emerging Field”, *Journal of Economic Literature* 53(1), pp. 5–42.
- GTZ (2009): “Air Quality Management”, *Sustainable Transport: A Sourcebook for Policy-Makers in Developing Cities*, Eschborn, Germany: Gesellschaft fuer Technische Zusammenarbeit.
- Guan, D., Z. Liu, Y. Geng, S. Lindner, and K. Hubacek (2012): “The gigatonne gap in China’s carbon dioxide inventories”, *Nature Climate Change* 2(9), pp. 672–675.

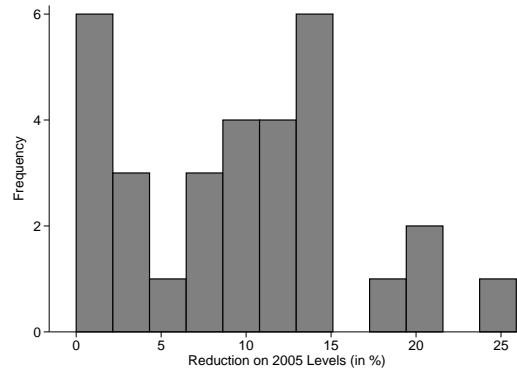
- Hammitt, J. K. and L. A. Robinson (2011): “The Income Elasticity of the Value per Statistical Life: Transferring Estimates between High and Low Income Populations”, *Journal of Benefit-Cost Analysis* 2(1), pp. 1–27.
- Hansman, C., J. Hjort, and G. Leon (2015): “Interlinked Firms and the Consequences of Piece-meal Regulation”, Mimeograph.
- Hao, J., K. He, L. Duan, J. Li, and L. Wang (2007): “Air pollution and its control in China”, *Frontiers of Environmental Science & Engineering in China* 1(2), pp. 129–142.
- Hartman, R. S., D. Wheeler, and M. Singh (1997): “The cost of air pollution abatement”, *Applied Economics* 29(6), pp. 759–774.
- He, H. (2012): “Air Pollutant Concentrations and Trends over the Eastern U.S. and China: Aircraft Measurements and Numerical Simulations”, Doctoral Dissertation, University of Maryland, College Park.
- He, J., H. Liu, and A. Salvo (2018): “Severe Air Pollution and Labor Productivity: Evidence from Industrial Towns in China”, *American Economic Journal: Applied Economics*.
- Henderson, V. (1996): “Effects of Air Quality Regulation”, *American Economic Review* 86(4), pp. 789–813.
- IIASA (2010a): “GAINS Asia: A Tool to Combat Air Pollution and Climate Change Simultaneously”.
- (2010b): “GAINS Asia: Scenarios for Cost-Effective Control of Air Pollution and Greenhouse Gases”.
- Ito, K. and S. Zhang (2016): “Willingness to Pay for Clean Air: Evidence from Air Purifier Markets”, NBER Working Paper No. 22367.
- Jayachandran, S. (2009): “Air quality and early-life mortality evidence from Indonesia’s wild-fires”, *Journal of Human Resources* 44(4), pp. 916–954.
- Jia, R. (2014): “Pollution for Promotion”, Job Market Paper.
- Kahn, M. E. (2006): *Green Cities: Urban Growth and the Environment*, Washington D.C.: Brookings Institution Press.
- Kahn, M. E., P. Li, and D. Zhao (2015): “Water Pollution Progress at Borders: The Role of Changes in China’s Political Promotion Incentives”, *American Economic Journal: Economic Policy* 7(4), pp. 223–242.
- Karplus, V. J., S. Zhang, and D. Almond (2018): “Quantifying coal power plant responses to tighter SO<sub>2</sub> emissions standards in China”, *Proceedings of the National Academy of Sciences* 115(27), pp. 7004–7009.
- Keller, W. and A. Levinson (2002): “Pollution abatement costs and foreign direct investment inflows to US states”, *Review of Economics and Statistics* 84(4), pp. 691–703.
- Keohane, N. O. (2003): “What Did the Market Buy? Cost Savings Under the U.S. Tradeable Permits Program for Sulfur Dioxide”, Yale School of Management Working Paper ES 33.
- (2006): “Cost savings from allowance trading in the 1990 Clean Air Act: Estimates from a choice-based model”, *Moving to markets in environmental regulation: Lessons from twenty years of experience*, ed. by C. E. Kolstad and J. Freeman, New York: Oxford University Press, pp. 194–229.
- Klimont, Z. et al. (2009): “Projections of SO<sub>2</sub>, NO<sub>x</sub> and carbonaceous aerosols emissions in Asia”, *Tellus B* 61(4), pp. 602–617.
- Krotkov, N. A. et al. (2016): “Aura OMI observations of regional SO<sub>2</sub> and NO<sub>2</sub> pollution changes from 2005 to 2015”, *Atmospheric Chemistry and Physics* 16(7), pp. 4605–4629.
- Li, H. and L.-A. Zhou (2005): “Political turnover and economic performance: the incentive role of personnel control in China”, *Journal of Public Economics* 89(9-10), pp. 1743–1762.
- Li, X., X. Wu, and F. Zhang (2015): “A method for analyzing pollution control policies: Application to SO<sub>2</sub> emissions in China”, *Energy Economics* 49, pp. 451–459.
- Lin, J. (2007): “Energy conservation investments: A comparison between China and the US”, *Energy Policy* 35, pp. 916–924.

- Lu, Z., D. G. Streets, Q. Zhang, S. Wang, G. R. Carmichael, Y. F. Cheng, C. Wei, M. Chin, T. Diehl, and Q. Tan (2010): “Sulfur dioxide emissions in China and sulfur trends in East Asia since 2000”, *Atmospheric Chemistry and Physics* 10(13), pp. 6311–6331.
- Martinez-Bravo, M., G. Padró i Miquel, N. Qian, and Y. Yao (2017): “The Rise and Fall of Local Elections in China: Theory and Empirical Evidence on the Autocrat’s Trade-Off”, NBER Working Paper 24032.
- Matus, K., K.-M. Nam, N. E. Selin, L. N. Lamsal, J. M. Reilly, and S. Paltsev (2012): “Health damages from air pollution in China”, *Global Environmental Change* 22(1), pp. 55–66.
- McElwee, C. R. (2011): *Environmental Law in China*, Oxford; New York: Oxford University Press.
- Meng, K. C. (2017): “Using a Free Permit Rule to Forecast the Marginal Abatement Cost of Proposed Climate Policy”, *American Economic Review* 107(3), pp. 748–784.
- MEP (2008a): “Issuance of Automatic Pollution Source Monitoring”, Original in Chinese.
- (2008b): “Management Measures for the Operation of Automatic Monitoring Installations for Pollution Sources”, Original in Chinese.
- Mu, Q. and J. Zhang (2017): “Air pollution and defensive expenditures: evidence from particulate-filtering face masks”, *Journal of Environmental Economics and Management*.
- NASA (2014): “OMSO2 Readme File v1.2.0 Released Feb 26, 2008 Updated: September 26, 2014”.
- Oates, W. E., P. R. Portney, and A. M. McGartland (1989): “The net benefits of incentive-based regulation: a case study of environmental standard setting”, *American Economic Review* 79(5), pp. 1233–1242.
- Oliva, P. (2015): “Environmental Regulations and Corruption: Automobile Emissions in Mexico City”, *Journal of Political Economy* 123(3), pp. 686–724.
- Purohit, P., M. Amann, R. Mathur, I. Gupta, S. Marwah, V. Verma, I. Bertok, J. Borken-Kleefeld, A. Chambers, J. Cofala, et al. (2010): “Gains Asia: Scenarios for cost-effective control of air pollution and greenhouse gases in India”.
- Schmalensee, R., P. L. Joskow, A. D. Ellerman, J. P. Montero, and E. M. Bailey (1998): “An Interim Evaluation of Sulfur Dioxide Emissions Trading”, *Journal of Economic Perspectives* 12(3), pp. 53–68.
- Schreifels, J. J., Y. Fu, and E. J. Wilson (2012): “Sulfur dioxide control in China”, *Energy Policy* 48(2012), pp. 779–789.
- Schultze, C. R. (1977): *The Public Use of Private Interest*, Washington D.C.: The Brookings Institution.
- Song, X.-l., L.-y. Zhang, M.-j. Gong, X. Wang, and Y. Qi (2015): “China’s Emission Data Accounting System of the Total Emission Control During the 11th Five-Year Plan Period”, *China Population, Resources and Environment* 25(1), pp. 46–54.
- State Council (2006): “Major Pollutant Emissions Control in the 11th Five-Year Plan Period”, Original in Chinese.
- (2007): “State Council approved the transfer of energy - saving emission reduction statistics. Monitoring and assessment of the implementation of programs and methods of notification”, Original in Chinese.
- Stavins, R. N. (1998): “What Can We Learn from the Grand Policy Experiment? Lessons from SO<sub>2</sub> Allowance Trading”, *Journal of Economic Perspectives* 12(3), pp. 69–88.
- (2003): “Experience With Market-Based Environmental Policy Instruments”, *Handbook of Environmental Economics*, ed. by K.-G. M’aler and J. Vincent, vol. Volume 1, Elsevier Science B.V.
- Stoerk, T. (2016): “Statistical corruption in Beijing’s air quality data has likely ended in 2012”, *Atmospheric Environment* 127, pp. 365–371.

- Stoerk, T., D. J. Dudek, and J. Yang (2018): “China’s National Carbon Emissions Trading Scheme: Lessons learned from the pilot ETSSs, the academic literature, and the known policy details”, Mimeograph.
- Tu, Z. and R. Shen (2014): “Can China’s Industrial SO<sub>2</sub> Emissions Trading Pilot Scheme Reduce Pollution Abatement Costs?”, *Sustainability* 6(11), pp. 7621–7645.
- Wang, A. L. (2013): “The search for sustainable legitimacy: Environmental law and bureaucracy in China”, *Harvard Environmental Law Review* 37, pp. 367–440.
- Wang, L., C. Jang, et al. (2010a): “Assessment of air quality benefits from national air pollution control policies in China. Part I: Background, emission scenarios and evaluation of meteorological predictions”, *Atmospheric Environment* 44(28), pp. 3442–3448.
- (2010b): “Assessment of air quality benefits from national air pollution control policies in China. Part II: Evaluation of air quality predictions and air quality benefits assessment”, *Atmospheric Environment* 44(28), pp. 3449–3457.
- Wang, Z.-x., L. Pan, and J. Peng (2005): “Analysis of Current Status, Cost and Policies of Power Sector SO<sub>2</sub> Emission Control”, *Research of Environmental Sciences* 18(4), Original in Chinese, pp. 11–20.
- Watts, J. (2011): *When a Billion Chinese Jump*, London: Faber and Faber Ltd.
- WHO (2014): *7 million premature deaths annually linked to air pollution*.
- World Bank and State Environmental Protection Administration (2007): *Cost of Pollution in China: Economic Estimates of Physical Damages*, Conference Edition.
- Xu, Y. (2011): “The use of a goal for SO<sub>2</sub> mitigation planning and management in China’s 11th Five-Year Plan”, *Journal of Environmental Planning and Management* 54(6), pp. 769–783.
- Yang, G., Y. Wang, Y. Zeng, G. F. Gao, X. Liang, M. Zhou, X. Wan, S. Yu, Y. Jiang, M. Naghavi, et al. (2013): “Rapid health transition in China, 1990-2010: findings from the Global Burden of Disease Study 2010”, *The Lancet* 381(9882), pp. 1987–2015.
- Zhang, X., X. Chen, and X. Zhang (2018): “The impact of exposure to air pollution on cognitive performance”, *Proceedings of the National Academy of Sciences*.
- Zheng, S. and M. E. Kahn (2013): “Understanding China’s Urban Pollution Dynamics”, *Journal of Economic Literature* 51(3), pp. 731–772.
- Zheng, S., C. Sun, Y. Qi, and M. E. Kahn (2014): “The Evolving Geography of China’s Industrial Production: Implications for Pollution Dynamics and Urban Quality of Life”, *Journal of Economic Surveys* 28(4), pp. 709–724.

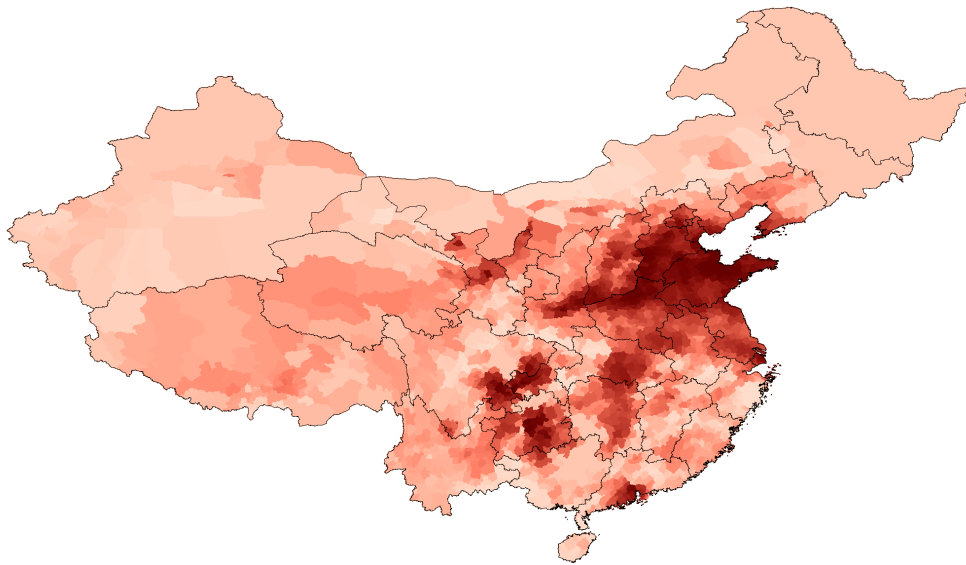
# Figures and tables

Figure 1: THE  $SO_2$  EMISSION CONTROL TARGETS



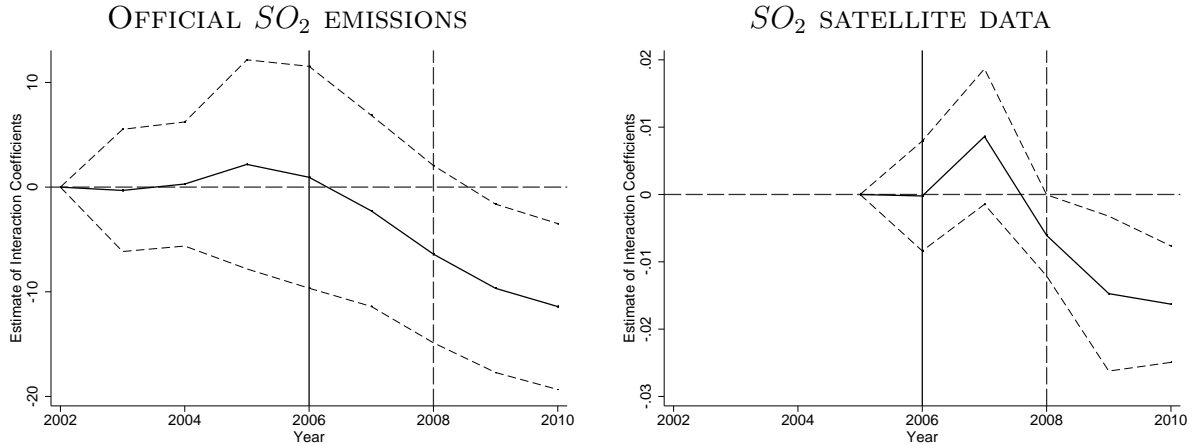
**Notes:** This figure shows the variation in the  $SO_2$  emissions reduction targets across the 31 provinces. Targets are shown as percentage reduction on 2005  $SO_2$  emission baselines. The mean of the distribution is 9.4% and the standard deviation is 6.8 percentage points. Data source: State Council (2006).

Figure 2:  $SO_2$  POLLUTION IN CHINA IN JANUARY 2006 BASED ON NASA SATELLITE DATA



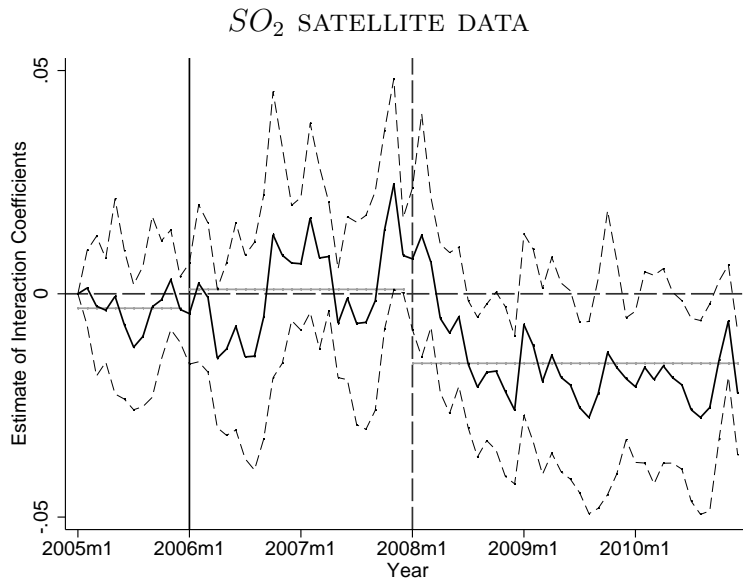
**Notes:** This figure shows the cross-section of the  $SO_2$  satellite data based on the NASA OMI  $SO_2$  data product for January 2006 mapped to the county-level.

Figure 3: DYNAMIC TREATMENT EFFECTS



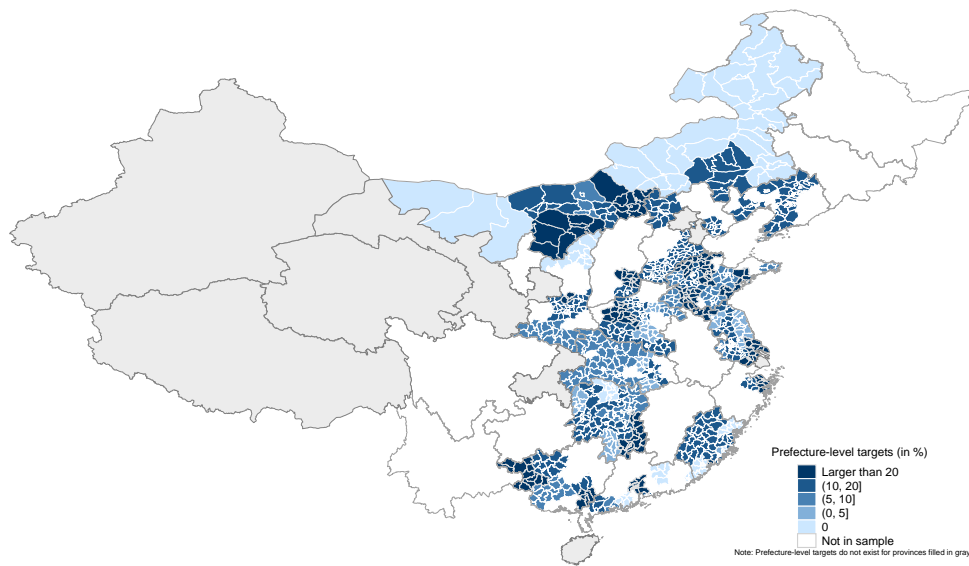
**Notes:** The solid line plots the point estimate for yearly coefficient estimates of the interaction coefficients  $\sum_{t=1}^T \beta_{1t}$  in Equation (2) for the official  $SO_2$  emissions data (*left panel*) and the NASA  $SO_2$  satellite data (*right panel*). The dashed lines show the 95% confidence bands, based on heteroskedasticity-robust standard errors clustered at the province-level. The excluded time period  $t = 1$  is the year 2002 (*left panel*) and the year 2005 (*right panel*). The vertical line before 2006 marks the start of the policy and the vertical line before 2008 marks the start of  $SO_2$  monitoring.

Figure 4: DYNAMIC TREATMENT EFFECTS: MONTHLY



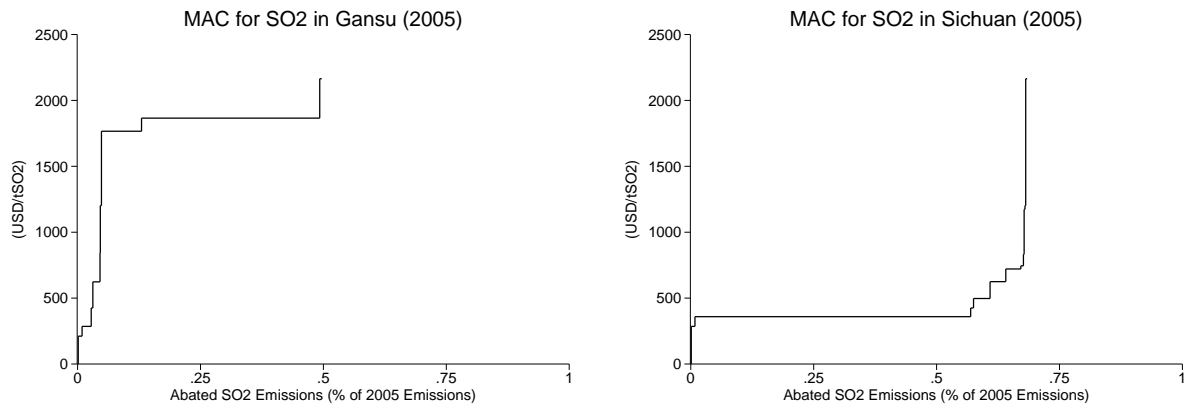
**Notes:** The solid line plots the point estimate for monthly coefficient estimates of the interaction coefficients  $\sum_{t=1}^T \beta_{1t}$  in Equation (2) for the NASA  $SO_2$  satellite data. The dashed lines show the 95% confidence bands, based on heteroskedasticity-robust standard errors clustered at the province-level. The horizontal grey lines plot the average of the point estimates of the interaction coefficient in the pre-period (2005), the period without monitoring (2006-2007) and the period with monitoring (2008-2010). The excluded time period  $t = 1$  is January 2005. The vertical line at 2006 marks the start of the policy and the vertical line at 2008 marks the start of  $SO_2$  monitoring.

Figure 5: SAMPLE OF PREFECTURE-LEVEL REDUCTION TARGETS



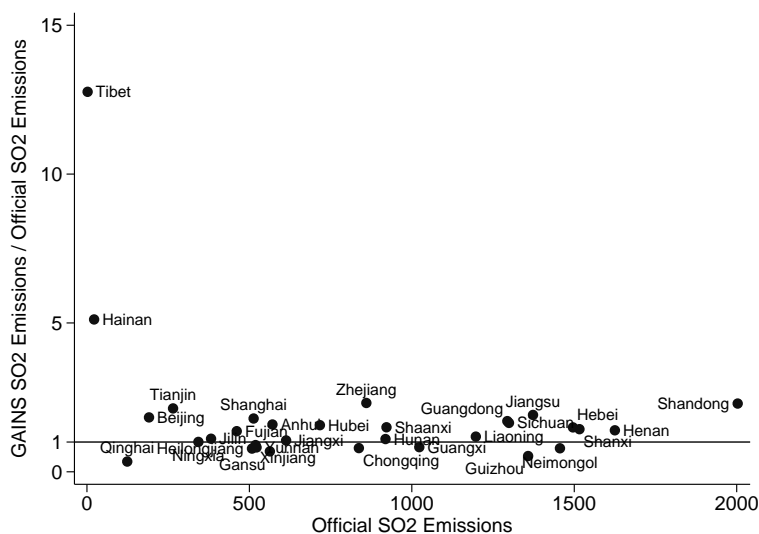
**Notes:** The graph shows the sample of prefecture-level  $SO_2$  reduction targets used to estimate heterogeneous treatment effects. Provinces filled in gray do not report prefecture-level targets.

Figure 6: MARGINAL ABATEMENT COST CURVE EXAMPLES



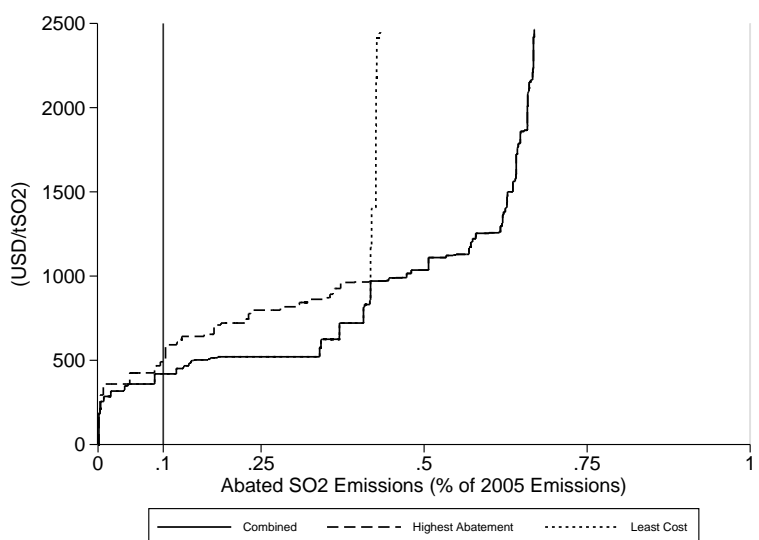
**Notes:** This graph shows the marginal  $SO_2$  abatement cost curves for Gansu and Sichuan in 2005. The horizontal axis lists the abatement intensity relative to each province's 2005  $SO_2$  emissions level.

Figure 7: CONSISTENCY OF  $SO_2$  MARGINAL ABATEMENT COST DATA AND OFFICIAL DATA



**Notes:** This graph shows the ratio of the  $SO_2$  emissions data underlying the construction of the marginal abatement cost curve (from IIASA’s GAINS model) and the official  $SO_2$  emissions data from the Ministry of Environmental Protection for the year 2005 on the vertical axis. The datapoints are ordered according to the official  $SO_2$  emissions level on the horizontal axis. The raw correlation between both data sources is 85.9%.

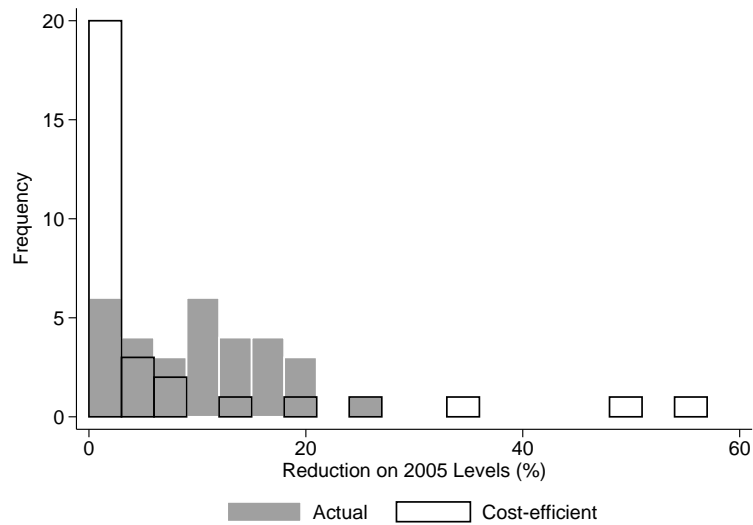
Figure 8:  $SO_2$  MARGINAL ABATEMENT COST CURVE FOR CHINA



**Notes:** This graph shows the marginal  $SO_2$  abatement cost curve for China in 2005. The horizontal axis plots the abatement intensity relative to the 2005  $SO_2$  emissions level. The vertical line at an abatement level of 10% illustrates the 10% national  $SO_2$  emissions control target for China as a whole. Its intersection with the marginal abatement cost curves shows that the counterfactual marginal abatement cost for this abatement level is \$419/t $SO_2$ .

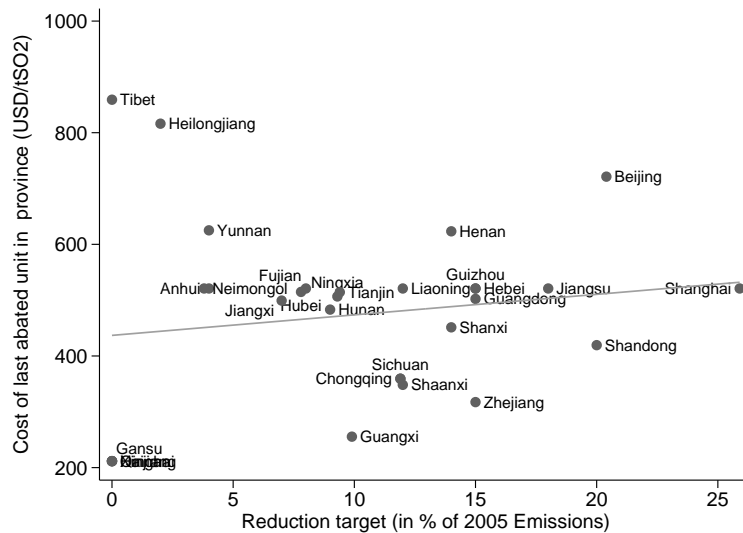


Figure 9: THE COST-EFFICIENT ALLOCATION OF  $SO_2$  REDUCTION TARGETS



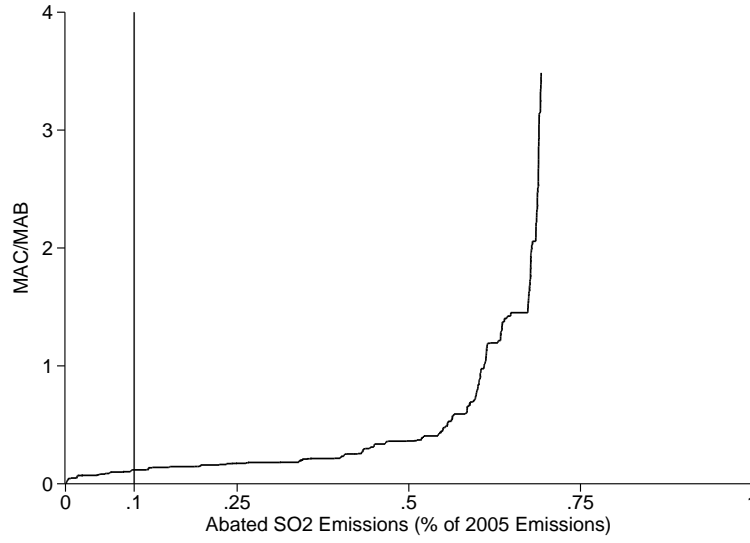
**Notes:** This graph shows the counterfactual market-based allocation of  $SO_2$  reduction targets under the 10%  $SO_2$  total control target of the 11th Five-Year Plan (2006-2010), shown in white boxes. These targets are based on equating MAC across provinces. The three provinces with the highest targets under the counterfactual market-based allocation are Sichuan, Shandong and Zhejiang. Solid grey boxes show the actual command-and-control allocation of reduction targets in comparison.

Figure 10: RELATION BETWEEN MAC AND  $SO_2$  REDUCTION TARGETS



**Notes:** This graph shows the lack of correlation between a province's command-and-control  $SO_2$  reduction target in the 11th Five-Year Plan (2006-2010) and its marginal abatement cost at the level of the target. The solid line fits a linear regression with slope parameter  $b=2.96$  and  $p=0.54$  computed from standard errors clustered at the province level.

Figure 11: MARGINAL WELFARE IMPACT AND  $SO_2$  REDUCTION TARGETS



**Notes:** This graph shows the ratio of marginal abatement cost to marginal abatement benefits for  $SO_2$  for China in 2005. The horizontal axis plots the abatement intensity relative to the 2005  $SO_2$  emissions level. The horizontal, longdashed line at 1 marks the welfare-optimal level of  $SO_2$  abatement (40.3%). The vertical, shortdashed line at 0.1 marks the ratio of marginal abatement cost to marginal abatement benefit for the 10%  $SO_2$  reduction target (0.07). Prohibitive cost ranges beyond a MAC/MAB ratio of 3.5 not shown.

Table 1: SUMMARY STATISTICS FOR MAIN VARIABLES

	<u>Mean</u>	<u>Std. Dev.</u>	<u>Min.</u>	<u>Max.</u>	<u>Time Period</u>
SO2 Emissions (kt)	739	471	1	2003	2002-2010
SO2 Satellite (DU)	0.37	0.39	-0.02	1.93	2005-2010
Selected SO2 Sat. (DU)	0.72	0.53	-0.03	2.57	2005-2010
SO2 Reduction Target (% on 2005 Baseline)	9.65	6.71	0	25.9	-

*Note:* Satellite data is measured in Dobson Units (DU). Selected  $SO_2$  Sat. is the sample of polluted cities. Satellite  $SO_2$  measurements below 0.2 Dobson Units are generally considered as clean air and negative values are likely noise from measurement error. Replacing negative values as either 0s or missing does not change the subsequent results.

Table 2: THE EFFECT OF THE POLICY FOR THE WHOLE PERIOD (2006-2010)

	<u>SO2 Emissions</u> (Kt)	<u>Satellite SO2</u> (Dobson Units)	<u>Selected Sat. SO2</u> (Dobson Units)
Reductiontarget × D(Post)	-6.30*** (2.22) [0.00]	-0.0057 (0.0040) [0.24]	-0.0115** (0.0048) [0.05]
Reductiontarget	0.36 (1.23) [0.81]	0.0214*** (0.0034) [0.00]	0.0251*** (0.0040) [0.00]
Year FE	✓	✓	✓
Province FE	✓	✓	✓
Effect size (% of mean/ $\sigma$ )	-5.8%	-10.5%	-10.9%
Observations	279	186	186
Provinces	31	31	31
$R^2$	0.98	0.94	0.91

*Note:* Heteroskedasticity-robust standard errors clustered at the province level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the province level are shown in square brackets. The effect size gives the estimated coefficient of the interaction term  $\beta_1$  in Equation (1) as percentage of the mean of the dependent variable for a one standard deviation( $\sigma$ )-increase in the  $SO_2$  reduction target. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 3: THE EFFECT OF THE POLICY FOR EACH YEAR OF THE 11TH FIVE-YEAR PLAN (2006-2010)

	<u>SO2 Emissions</u> (Kt)	<u>Satellite SO2</u> (Dobson Units)	<u>Selected Sat. SO2</u> (Dobson Units)
2002× Reductiontarget	Excluded	-	-
2003× Reductiontarget	-0.32 (3.93) [0.93]	-	-
2004× Reductiontarget	0.30 (3.64) [0.93]	-	-
2005× Reductiontarget	2.17 (4.04) [0.66]	Excluded	Excluded
2006× Reductiontarget	0.94 (4.08) [0.86]	-0.00 (0.00) [0.93]	0.00 (0.01) [0.94]
2007× Reductiontarget	-2.27 (3.70) [0.62]	0.01 (0.00) [0.21]	0.01 (0.01) [0.29]
2008× Reductiontarget	-6.40 (3.69) [0.14]	-0.01 (0.00) [0.11]	-0.01*** (0.01) [0.00]
2009× Reductiontarget	-9.67** (3.79) [0.02]	-0.01** (0.00) [0.02]	-0.02** (0.00) [0.01]
2010× Reductiontarget	-11.41*** (3.89) [0.00]	-0.02*** (0.00) [0.00]	-0.03*** (0.00) [0.00]
Year FE	✓	✓	✓
Province FE	✓	✓	✓
Mean dep. var.	739	0.37	0.72
Observations	279	186	186
Provinces	31	31	31
$R^2$	0.98	0.96	0.93

*Note:* Heteroskedasticity-robust standard errors clustered at the province level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the province level are shown in square brackets. The yearly interaction coefficients are estimates for  $\sum_{t=1}^T \beta_{1t}$  in Equation (2), while 'Excluded' is the omitted time period. Note that the  $SO_2$  reduction started in 2006 while government monitoring of  $SO_2$  pollution became effective in 2008. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 4: HETEROGENEOUS TREATMENT EFFECTS DEPENDING ON INITIAL POLLUTION LEVELS

	SO2 Satellite Data			
	Quartile 1	Quartile 2	Quartile 3	Quartile 4
Reductiontarget	0.0004	-0.0032	-0.0082	-0.0119*
× D(Post)	(0.0028)	(0.0030)	(0.0055)	(0.0053)
	[0.93]	[0.41]	[0.26]	[0.07]
Reductiontarget	0.0311***	0.0248***	0.0224***	0.0210***
	(0.0023)	(0.0025)	(0.0046)	(0.0044)
	[0.00]	[0.00]	[0.00]	[0.00]
Year-Month FE	✓	✓	✓	✓
Province FE	✓	✓	✓	✓
Effect size	1.5%	-5.8%	-9.9%	-9.4%
(% of mean/ $\sigma$ )				
Observations	47,376	47,592	47,016	48,096
Provinces	31	31	31	31
$R^2$	0.59	0.66	0.66	0.67

*Note:* The table reports estimates for Equation (1) at the county-yearmonth level for different subsamples. Quartile marks the quartile of initial pollution based on its 2005  $SO_2$  pollution, calculated from  $SO_2$  satellite data, relative to the mean  $SO_2$  pollution within the same province. Heteroskedasticity-robust standard errors clustered at the province level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the province level are shown in square brackets. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 5: THE EFFECT OF ENFORCEMENT

	<u>SO2 Satellite Data</u>			
	Enforcement officials density			
	<u>Quartile 1</u>	<u>Quartile 2</u>	<u>Quartile 3</u>	<u>Quartile 4</u>
Reductiontarget × D(Post1)	0.1855 (0.1367) [0.20]	0.1153 (0.2392) [0.66]	0.1068 (0.1640) [0.58]	1.2713 (0.6337) [0.15]
Reductiontarget × D(Post2)	0.0429 (0.1405) [0.68]	-0.3351 (0.2232) [0.14]	-0.4061*** (0.1408) [0.006]	1.0497* (0.4632) [0.09]
Year FE	✓	✓	✓	✓
Prefecture FE	✓	✓	✓	✓
Effect size (% of mean/ $\sigma$ )	0.5%	-6.4%	-8.4%	7.6%
Observations	2,016	2,088	2,448	2,304
Prefectures	28	29	34	32
$R^2$	0.80	0.73	0.57	0.58

*Note:* The table shows the results from estimating equation 3 at the prefecture level. Quartile marks the quartile of the number of enforcement officials per area at the end of the 11th Five-Year Plan. Heteroskedasticity-robust standard errors clustered at the prefecture level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the prefecture level are shown in square brackets. Effect size expresses the point estimate for the period in which there was enforcement (Reductiontarget × D(Post2)) as per cent change in pollution per standard deviation increase in reduction target stringency relative to each subsample's 2005 mean of SO2 pollution. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Table 6: Efficiency gains from trade

	<u>Efficiency measure</u>		
<u>Allocation</u>	Average abatement cost (in $\$/tSO_2$ )	Marginal abatement cost (in $\$/tSO_2$ )	Marginal welfare impact (MAC/MAB)
Command-and-control	436.78	816.2	2.2
Cost-efficient	323.45	419.4	1.2

*Note:* The average abatement cost measure reports the total cost divided by the total number of abated units of  $SO_2$  for the actual command-and-control allocation from the 11th Five-Year Plan (2006-2010) and the counterfactual market-based allocation reported in Table A.2. The marginal efficiency measures report the highest cost of all last abated units across provinces for either allocation. Estimates for the cost of the actual allocation exclude outliers. The welfare impact at the margin is the ratio between marginal abatement cost and marginal abatement benefits.

# Appendix

## A.1 Compliance: rhetorical and real

How did the regulated provincial governments react to the sudden air pollution targets? On the one hand, related research shows that regulated local governments likely misreport desired environmental data when the central government cannot monitor them (Ghanem and Zhang, 2014; Stoerk, 2016). On the other hand, my analysis shows that the  $SO_2$  reduction control policy was ultimately successful in reducing air pollution. Provincial governments could thus have reacted along two margins: (i) rhetorical compliance and (ii) real compliance. I show that their reaction was initially rhetorical, but ultimately real. I first discuss the data sources, followed by a discussion of my empirical findings.

### A.1.1 Data on reactions by the regulated provincial governments

**Rhetorical compliance** I build a comprehensive dataset of political statements by each provincial government to study whether the regulated provincial governments respond to the  $SO_2$  reduction targets by mimicking the central government's rhetoric. In China, each provincial government has the obligation to issue a government work report every year. This report is publicly delivered by one of the two highest ranked officials in the province, the party secretary or the governor. Each report contains information on the provincial government's activities and achievements. These reports are divided into two parts: Part 1 discusses the provincial government's work and achievements in the preceding period, while Part 2 discusses the work in the period to come. This unique setting allows me not only to investigate political rhetoric in general, but to also specifically investigate rhetorical responses relating to past and future achievements in response to the  $SO_2$  reduction targets.

To measure the provincial government's political attention towards air pollution for a given province-year, I scan each province's government work report for the years 2002-2010 and construct a variable that is equal to the number of occurrences of keywords related to air pollution<sup>15</sup>. Keywords were chosen from a technical document on urban air pollution in developing countries (GTZ, 2009) as well as from China-specific air pollution articles, webpage entries and blog posts in March 2014 from *China Daily*, *Global Times*, *Beijing Review* and *Jinyang Yangcheng Evening News*. To further rule out cherry-picking of keywords, I have defined the list of keywords as widely as possible. Figure A.3 contains

---

<sup>15</sup>6 reports are missing in 2002 and 1 report is missing in 2003.

the raw count of keywords over time, offering two take-aways. First, there is a distinct increase in air pollution related keywords: Mentions of air pollution increase by more than 400% during the 11th Five-Year Plan (2006-2010). Second, one keyword drives this increase: the specific keyword 'sulfur'. This keyword is directly related to the provincial  $SO_2$  reduction targets from the 11th Five-Year Plan. The outcome variable therefore appears to capture relevant political statements.

[INSERT FIGURE A.3 ABOUT HERE]

**Real compliance** Based on my own calculations of the marginal cost of  $SO_2$  emissions abatement (see Section 6), the installation of desulfurization devices in existing industrial and power plants (scrubbers), fuel-switching to better quality coal and the shutdown of small, inefficient thermal units are the main margins by which the provincial governments could reduce  $SO_2$  emissions over the relatively short time horizon of 5 years. I collect data on both the timing and the quantity of installation of desulfurization devices at the province level during the 11th Five-Year Plan (2006-2010). Furthermore, I compute the number and capacity of small thermal units that were shut down by 2010 in each province. The latter are based on a planning document from 18th January 2008, in which the MEP asked the provincial governments to submit a concrete proposal for the thermal units to be shut down over the following two years. All data were compiled from sources available through the data center of the MEP (datacenter.mep.gov.cn). I do not analyze the sulfur content of coal used at the province-year level since reliable data is unavailable.

### A.1.2 Findings on rhetorical and real compliance

**Rhetorical compliance** This subsection investigates whether the provincial governments changed their political rhetoric in response to the  $SO_2$  reduction control targets. To do so, I estimate versions of Equation (2), using different counts of keywords as outcome variables. First, I use the overall number of keywords related to air pollution in each government work report. The left panel in Figure A.4 plots the estimates for the yearly interaction coefficients  $\beta_{1t}$ . It shows that provinces that received a higher  $SO_2$  reduction target show a distinct increase in their political rhetoric on air pollution.

[INSERT FIGURE A.4 ABOUT HERE]



Next, I zoom in and split political statements into those on work done in the preceding year and those related to projects for the year to come. The central panel in Figure A.4 shows that the increase is most pronounced in the period before the central government had the capacity to monitor. Provincial governments that received a higher  $SO_2$  reduction target claim past work on projects related to air pollution in 2007, thus exploiting the central government’s inability to monitor  $SO_2$  initially. That is, a higher  $SO_2$  reduction target induces provinces to claim work on air pollution in 2006. Political statements relative to future work on air pollution also increase with the  $SO_2$  reduction targets, although the evidence is less stark as shown in the right panel of Figure A.4. Finally, Table A.1 summarizes these results based on estimating Equation (1) on the political attention variables to find that: (i) political attention to air pollution increases by 30% of the pre-treatment mean per standard deviation increase in target stringency, and (ii) statements about past work on air pollution peak in the period in which the government could not monitor  $SO_2$ <sup>16</sup>. As shown in the analysis of NASA satellite data, however,  $SO_2$  pollution during that period did not improve. My empirical findings therefore suggest that the provincial governments exploit the central government’s inability to monitor pollution by only adjusting their rhetorical compliance through public political statements.

**Real compliance** As shown above, the  $SO_2$  reduction targets worked in reducing  $SO_2$  pollution significantly from 2008 onwards. Based on my own calculations of the marginal cost of  $SO_2$  emissions abatement (see Section 6), the installation of desulfurization devices in existing industrial and power plants, fuel-switching to better quality coal, and the shutdown of small, inefficient thermal units are the main margins by which the provincial governments could reduce  $SO_2$  emissions over relatively short time horizon of 5 years. I

---

<sup>16</sup>Table A.1 includes a robustness check to further show that my estimates on rhetorical compliance are meaningful. This test involves estimating the effect of the  $SO_2$  reduction targets on closely related, yet different placebo outcomes that should not be affected by the  $SO_2$  reduction targets. The 11th Five-Year Plan (2006-2010) included goals to increase China’s forest cover from 18.2% to 20% and to extend the coverage of rural medical care from 23.5% to 80%. To measure the political attention towards these policies, I use the count of ‘forest’ and of ‘medical care’ in the government work reports as dependent variables for the falsification test. Results based on these outcomes using the specification in Equation (1) are reported in the columns ‘Placebo Outcomes’ in Table A.1. As in the case of the keywords related to air pollution, both keywords are mentioned more often during the 11th Five-Year Plan (2006-2010) than before. Provinces that received a higher  $SO_2$  reduction target, however, do not talk more about either topic. This strongly suggests that governments of provinces with higher reduction targets do not mimic the Central government’s political agenda in their own statements in general. Instead, they specifically change their political communications in response to the  $SO_2$  reduction targets.

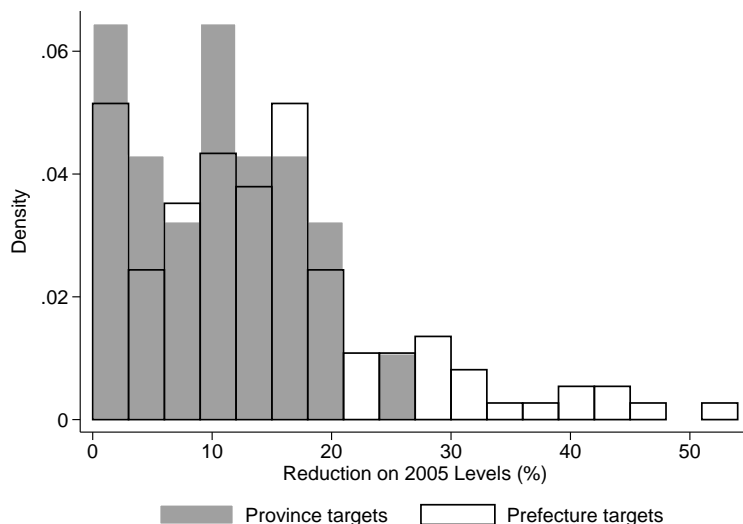
provide evidence on the last two channels.

The data on the *installation of scrubbers* include both the capacity of the desulfurization equipment as well as the timing of its installation in each province. The right panel of Figure A.5 shows that while provinces with a higher  $SO_2$  reduction targets installed more desulfurization devices on average, that effect is not statistically different from zero. The central panel of Figure A.5 computes the skewness of the timing of the installation of the scrubbers between the years 2006 and 2010 and correlates it to the  $SO_2$  reduction targets at the province level. As can be seen, provinces with a higher  $SO_2$  reduction target did not install scrubbers earlier than other provinces. The data on the *shutdown of small, inefficient thermal units* shows a much clearer picture. The right panel of Figure A.5 shows that the higher the reduction target, the higher the capacity of small thermal units shut down by 2010. Shanghai and Beijing are the exception to this rule, most likely because they had already shut down inefficient plants in the past.

[INSERT FIGURE A.5 ABOUT HERE]

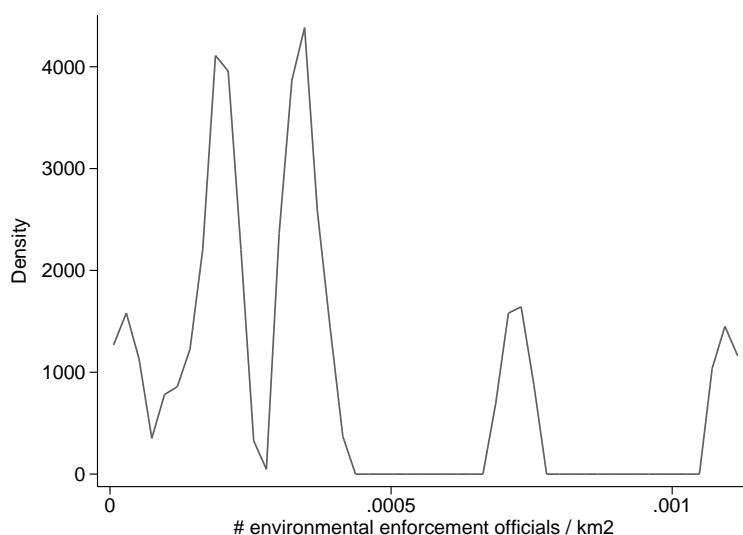
## A.2 Appendix Figures and Tables

Figure A.1: REPRESENTATIVENESS OF THE PREFECTURE SUBSAMPLE



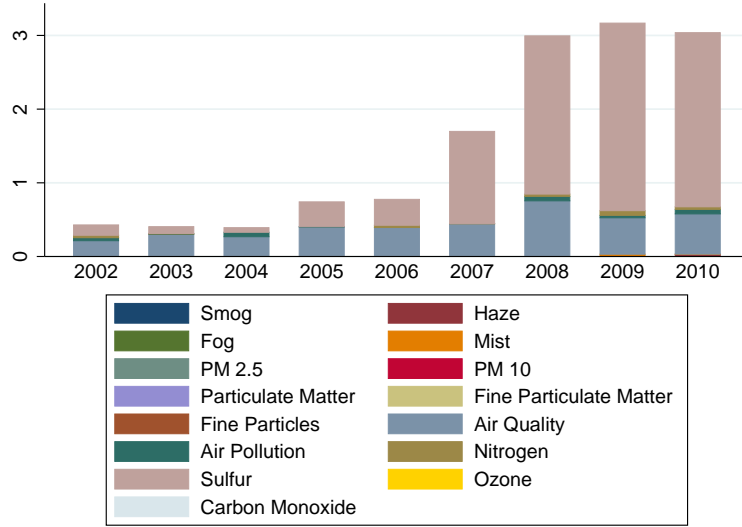
**Notes:** This graph overlays the  $SO_2$  reduction targets for the provinces and the prefectures in the subsample used to study the effect of enforcement. The province targets have a mean of 9.65 (standard deviation: 6.71) while the prefecture targets have a mean of 14.43 (standard deviation: 10.90).

Figure A.2: DISTRIBUTION OF ENFORCEMENT OFFICIAL DENSITY



**Notes:** The figure shows the kernel density measure for the number of environmental enforcement officials per square kilometer for the provinces in my prefecture-level sample (Fujian, Guangdong, Guangxi, Hebei, Henan, Hubei, Hunan, Jiangsu, Liaoning, Neimenggu, Shaanxi, Shandong, Shanxi and Zhejiang). The kernel function is an Epanechnikov kernel with half-width  $2.064e - 05$ .

Figure A.3: POLITICAL ATTENTION TO AIR POLLUTION OVER TIME



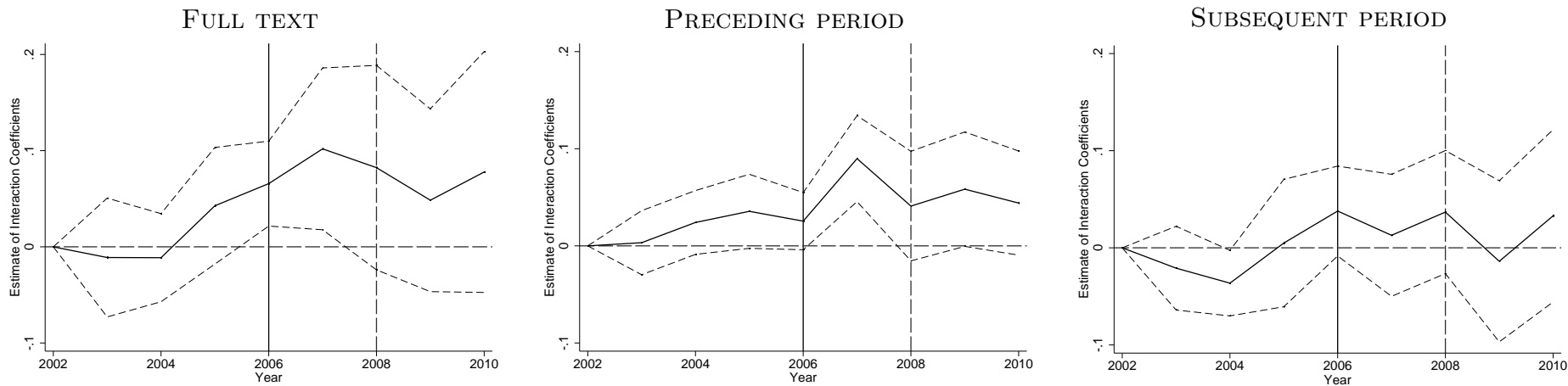
**Notes:** The graph shows the mean count of keywords related to air pollution for all provincial government work reports in a given year from 2002 to 2010.

Table A.1: RHETORICAL COMPLIANCE

	Political attention to air pollution			Placebo outcomes	
	All statements	Past period	Future period	'Forest'	'Medical care'
Reductiontarget × D(Post)	0.07** (0.03) [0.02]	0.03** (0.01) [0.02]	0.04** (0.02) [0.03]	-0.12 (0.12) [0.36]	0.02 (0.09) [0.73]
Reductiontarget	-0.01 (0.02) [0.75]	-0.00 (0.01) [0.70]	-0.00 (0.01) [0.69]	-0.07 (0.07) [0.34]	0.34*** (0.05) [0.00]
Year FE	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓
Effect size (% of mean/ $\sigma$ )	30.9%	29.6%	31.5%	-7.2%	1.2%
Observations	274	273	273	274	274
Provinces	31	31	31	31	31
$R^2$	0.66	0.64	0.54	0.72	0.58

*Note:* Heteroskedasticity-robust standard errors clustered at the province level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the province level are shown in square brackets. The effect size gives the estimated coefficient of the interaction term  $\beta_1$  in Equation (1) as percentage of the mean of the dependent variable for a one standard deviation( $\sigma$ )-increase in the  $SO_2$  reduction target. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

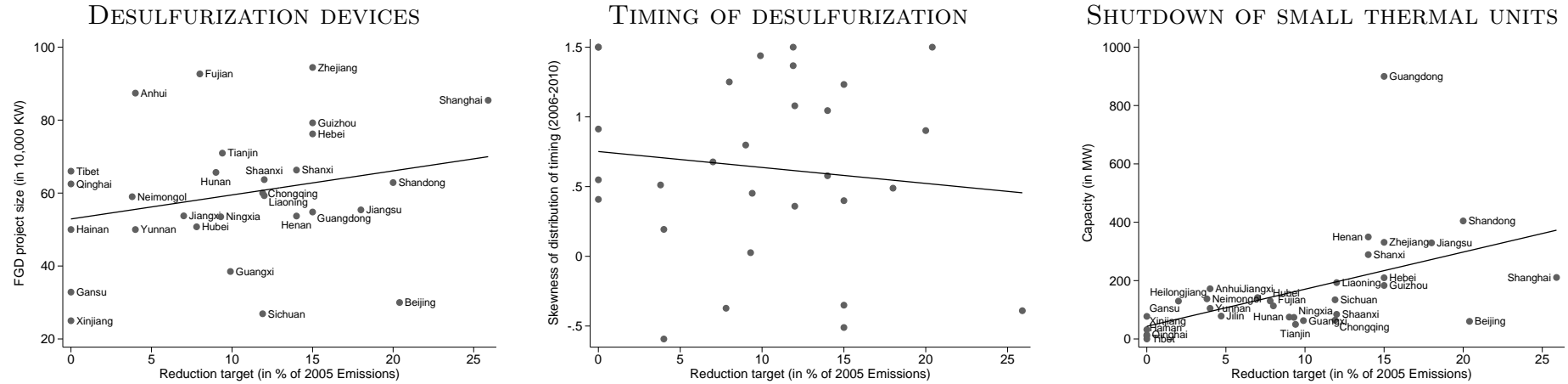
Figure A.4: RHETORICAL COMPLIANCE



50

**Notes:** The solid line plots the point estimate for yearly coefficient estimates of the interaction coefficients  $\sum_{i=1}^T \beta_{1t}$  in Equation (2) for the number of keywords related to air pollution in a province-year government work report. *Full text* refers to the entire document, whereas *preceding period* only analyses sections on word done in the period preceding the report. Likewise, *Subsequent period* refers to work announced for the period following the report. The dashed lines show the 95% confidence bands, based on heteroskedasticity-robust standard errors clustered at the province-level. The excluded time period  $t = 1$  is the year 2002. The vertical line at 2006 marks the start of the policy and the vertical line at 2008 marks the start of  $SO_2$  monitoring.

Figure A.5: REAL COMPLIANCE



15

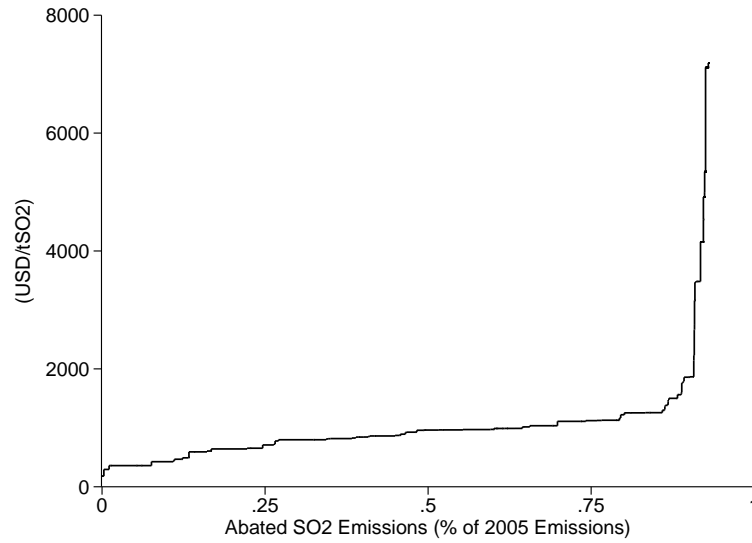
**Notes:**

*Left panel:* This graph plots the relationship between the  $SO_2$  reduction target and the planned installation of desulfurization devices at the province level. The solid line fits a linear regression with slope parameter  $b=0.66$  and  $p=0.22$  computed from standard errors clustered at the province level.

*Central panel:* This graph tests whether provinces with a higher reduction target installed the planned desulfurization devices earlier. The vertical axis shows the skewness for each province of the 5 yearly observations from 2006-2010, where the weight is the capacity (in 10,000 KW) of planned desulfurization devices in each year. The solid line fits a linear regression with slope parameter  $b=-0.01$  and  $p=0.56$  computed from standard errors clustered at the province level.

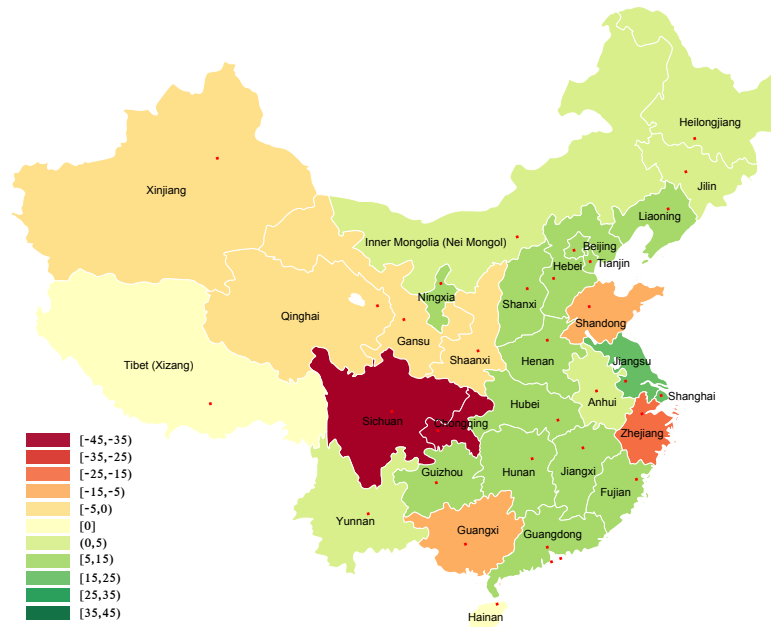
*Right panel:* This graph plots the relationship between the  $SO_2$  reduction target and the decommissioning of small thermal units at the province level. The solid line fits a linear regression with slope parameter  $b=12.72^{***}$  and  $p=0.004$  computed from standard errors clustered at the province level.

Figure A.6:  $SO_2$  MAC IN CHINA'S POWER SECTOR



**Notes:** This graph shows the marginal abatement cost curve for  $SO_2$  abatement in China's power sector in 2005. The curve is constructed by using the abatement options that offer the highest level of abatement.

Figure A.7: DIFFERENCE BETWEEN COST-EFFICIENT AND ACTUAL ALLOCATION



**Notes:** The graph shows the difference in percentage points between the  $SO_2$  reduction target of each province under the command and control minus the market-based allocation (data from Table A.2).

Table A.2: ACTUAL AND COUNTERFACTUAL  $SO_2$  TARGET ALLOCATIONS

	SO2 reduction target under different allocations		Difference between target allocations
	(in % of 2005 emissions)		(in percentage points)
	Market-based	Actual	Actual - Market-based
Anhui	1.5	4	2.5
Beijing	7.3	20.4	13.1
Chongqing	50.5	11.9	-38.6
Fujian	0.7	8	7.3
Gansu	2.5	0	-2.5
Guangdong	0.7	15	14.3
Guangxi	20	9.9	-10.1
Guizhou	5.8	15	9.2
Hainan	0	0	0
Hebei	1.7	15	13.3
Heilongjiang	1.7	2	0.3
Henan	0.6	14	13.4
Hubei	2.5	7.8	5.3
Hunan	1	9	8
Jiangsu	0.9	18	17.1
Jiangxi	1.6	7	5.4
Jilin	1.6	4.7	3.1
Liaoning	4.2	12	7.8
Neimongol	1.1	3.8	2.7
Ningxia	0.1	9.3	9.2
Qinghai	3.5	0	-3.5
Shaanxi	13.9	12	-1.9
Shandong	25.5	20	-5.5
Shanghai	6.2	25.9	19.7
Shanxi	0.6	14	13.4
Sichuan	56.2	11.9	-44.3
Tianjin	2.2	9.4	7.2
Tibet	0	0	0
Xinjiang	1.6	0	-1.6
Yunnan	1.3	4	2.7
Zhejiang	35.4	15	-20.4

*Note:* This table shows the  $SO_2$  emissions reduction targets under each policy regime. 'Market-based' are the counterfactual targets for achieving the counterfactual market-based allocation given the national 10%  $SO_2$  reduction target, and 'Actual' are the actual command-and-control provincial reduction targets used in the 11th Five-Year Plan (2006-2010).



### A.3 Further robustness checks for identification

This section further substantiates my identification strategy by showing that neither the cost nor the benefit of reducing air pollution nor the Great Recession correlates with the  $SO_2$  reduction targets.

Firstly, I show that the neither the cost nor the welfare impact at the margin correlates with the  $SO_2$  reduction targets, using the following empirical specification:

$$y_{pt} = \beta_0 + \beta_1 \text{Reductiontarget}_p \times D(\text{Post})_t + \beta_2 X_p \times D(\text{Post})_t + \beta_3 \text{Reductiontarget}_p + \beta_4 X_p + \sum_{t=1}^T \beta_{5t} \gamma_t + \alpha_p + u_{pt} \quad (4)$$

where  $X_p$  is either a province-level measure of (i) the marginal abatement cost or (ii) the ratio of marginal abatement benefits to marginal abatement cost. Specific details for the construction of those measures are provided in Section 6. In a nutshell, my approach is this: I compute detailed marginal abatement cost curves for  $SO_2$  for each province in China, based on a reliable set of micro data on the cost and abatement potential of fine-grained polluting activities. The marginal benefits from reducing air pollution are based on a back-of-the-envelope calculation that follows Oliva (2015) and evaluates health improvements based on the value of a statistical life.

Table A.3 shows that the estimates for  $\beta_2$  are 0 and that the estimates for  $\beta_1$  are nearly identical to those in Table 2.

Secondly, I find that there is no relationship between the economic downturn in the Great Recession and the stringency of a province's  $SO_2$  reduction targets. If this were the case, my results could be confounded because a slowdown in economic activity could go hand in hand with a reduction in  $SO_2$  pollution. It is important to note that even though China experienced a slowdown in growth in 2009, the slowdown was comparatively mild. Even in 2009, there are only 3 provinces with a growth rate below 5%, and the growth rate for those provinces is still positive. I compute the magnitude of the downturn as the deviation in the average of the growth rates for 2002-2007 for each province. I find that there is no relationship between the magnitude of the recession and the reduction targets in a province. This is because different provinces with similar reduction targets experienced rather different deviations from their long term growth rates in 2009, and overall there is no statistically significant correlation between  $SO_2$  reduction targets and the downturn. As shown in Figure A.8, a linear regression with standard errors clustered

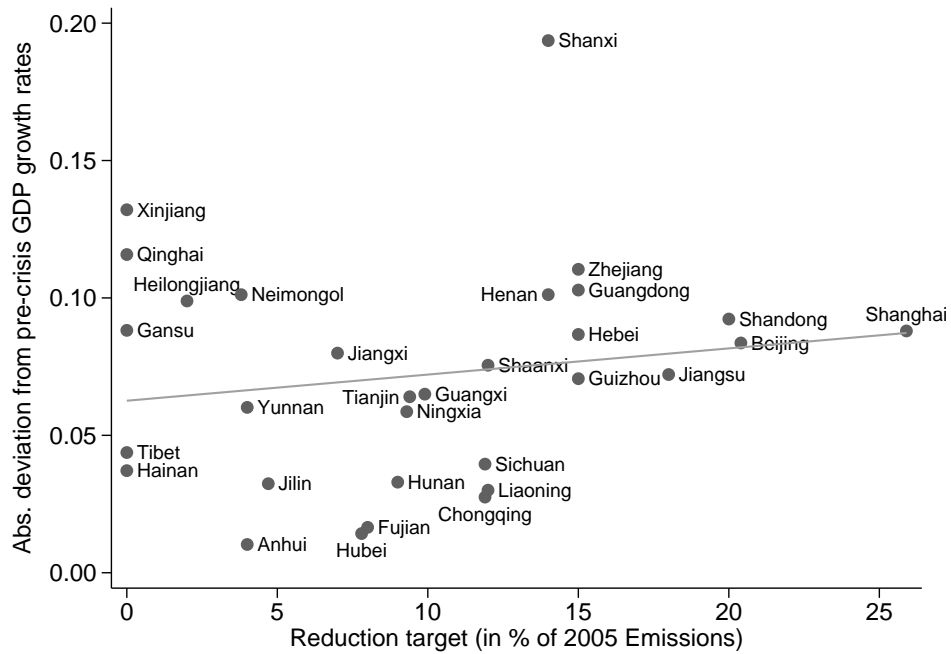
at the province level finds a best fit with a p-value of 0.31. The Great Recession can therefore not explain the decrease in  $SO_2$  pollution from 2008 onwards.

Table A.3: CONTROLLING FOR MARGINAL ABATEMENT COST AND MARGINAL WELFARE IMPACT

	<u>SO2 Emissions</u> (Kt)		<u>SO2 Satellite</u> (Dobson Units)		<u>Sel. Satellite SO2</u> (Dobson Units)	
Reductiontarget × D(Post)	-6.26*** (2.23) [0.01]	-6.15*** (2.22) [0.00]	-0.0057 (0.0040) [0.20]	-0.0058 (0.0041) [0.27]	-0.0116** (0.0048) [0.05]	-0.0117* (0.0048) [0.05]
Reductiontarget	0.33 (1.24) [0.80]	1.33 (1.23) [0.28]	0.0214*** (0.0034) [0.00]	0.0364*** (0.0034) [0.00]	0.0252*** (0.0040) [0.00]	0.0461*** (0.0040) [0.00]
MAC × D(Post)	0.17 (0.026) [0.45]		-0.0000 (0.0000) [0.87]		-0.0000 (0.0000) [0.20]	
MAC	-0.12*** (0.014) [0.00]		-0.0003*** (0.0000) [0.00]		-0.0003*** (0.000) [0.00]	
MAC/MAB × D(Post)		-0.59 (0.76) [0.43]		0.0001 (0.0013) [0.90]		0.0009 (0.0011) [0.34]
MAC/MAB		0.71* (0.42) [0.09]		0.0053*** (0.0010) [0.00]		0.0068*** (0.0009) [0.00]
Year FE	✓	✓	✓	✓	✓	✓
Province FE	✓	✓	✓	✓	✓	✓
Observations	279	279	186	186	186	186
Provinces	31	31	31	31	31	31
$R^2$	0.98	0.98	0.94	0.94	0.91	0.91

*Note:* Heteroskedasticity-robust standard errors clustered at the province level are shown in parentheses. Wild bootstrap p-values from 1000 repetitions clustered at the province level are shown in square brackets. The table reports the results from estimating Equation (4). MAC refers to the marginal abatement cost given the actual command-and-control  $SO_2$  reduction target and MAC/MAB is the ratio of the marginal abatement benefits to the marginal abatement cost. \*\*\* $p < 0.01$ , \*\* $p < 0.05$ , \* $p < 0.1$ .

Figure A.8: THE GREAT RECESSION AND  $SO_2$  REDUCTION TARGETS



**Notes:** This graph plots the  $SO_2$  reduction targets on the horizontal axis against the absolute deviation from the pre-crisis (2002-2007) GDP growth rate for each province in China in 2009, the year China was struck by the Great Recession. The solid line fits a linear regression with slope parameter  $b=0.001$  and  $p=0.305$  computed from standard errors clustered at the province level.

## A.4 Construction of optimized marginal abatement cost curves

In principle, construction of the full MAC curve for a given province requires an optimization that trades off the unitcost and the abatement potential of different abatement technologies to reach a given abatement level  $\bar{x}$ . Figure A.9 illustrates this point: for most of the lower abatement levels, up to ca. 40%, the *least cost* MAC would be picked as the optimal MAC (black line). At one point, however, the *least cost* crosses the *highest abatement* curves (dashed line), offering the chance for cheaper abatement by switching to the *highest abatement* cost curve.

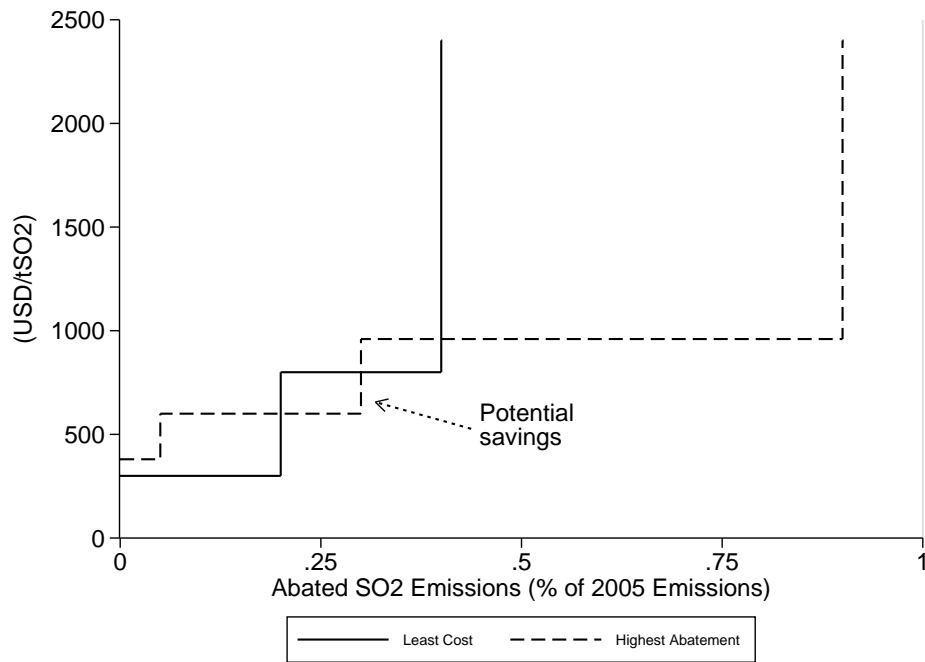
[INSERT FIGURE A.9 ABOUT HERE]

For example, pick  $\bar{x}$  to be an abatement level of 25% of the baseline, or 0.25. Here, the last unit of abatement would be cheaper than on the least cost curve. However, to be able to use the highest abatement technology at that point, one would have to forego the use of cheaper abatement technologies from the least curve for earlier units. To calculate the marginal abatement cost for abatement level  $\bar{x}$ , one would therefore have to compute the total abatement cost for abatement level  $\bar{x} - \epsilon$  for both the least cost and the highest abatement MACs plus the cost of the marginal unit along each curve. In other words, the cost for abatement at the margin depends on the history of abatement for units that were abated earlier.

The following algorithm would be required fully calculate the optimized MAC: (i) Compute all possible combinations of abatement technologies for each province, (ii) for each discrete abatement level  $\bar{x}$ ,  $\bar{x} \in (0, 100)$ , calculate the total cost of abatement along all combinations of abatement technologies, (iii) record the cost at the margin for abatement level  $\bar{x}$  by subtracting the same cost for abatement level  $\bar{x} - \epsilon$ . Do this calculation for infinitesimally small  $\Delta\bar{x}$ .

Computational considerations make these calculations overly costly for my application. More importantly, however, they are not needed: for the relatively low levels of reduction considered in this research, the least cost MAC is nearly always below the highest abatement MAC. In the relevant range, the fully optimized MAC would be nearly identical to the least cost MAC .

Figure A.9: EXAMPLE TO ILLUSTRATE MAC CURVE CONSTRUCTION



**Notes:** This graph is an example for illustrative purposes, based on simulated data. It shows that the *least cost* MAC curve could, in principle, overestimate the cost of abatement for the abatement levels for which the *highest abatement* MAC offers lower cost (area marked as 'Potential savings').

## A.5 Marginal $SO_2$ abatement cost curves for all provinces in China

