Career concerns and multitasking local bureaucrats: Evidence of a target-based performance evaluation system in China

Abstract

This paper examines whether a target-based performance evaluation system can properly motivate local bureaucrats to implement an environmental regulation policy at the cost of slow economic growth. In late 2005, the Chinese central government made cutting sulfur dioxide (SO_2) emissions the main performance evaluation criterion for prefectural city mayors and Party secretaries. Using a difference-in-differences framework, we find that the performance evaluation system caused a significant decrease in SO_2 emissions, as well as in the GDP growth rate. Our mechanism analyses further corroborate that local bureaucrats in Two Control Zone cities were willing to trade off GDP growth to achieve the more stringent emissions reductions goal. Our findings contribute to the understanding of multitasking agency problems for government agents.

1. Introduction

Economic growth has many implications for the environment and sustainability. Long-run growth depends not only on emission flows but also on the assimilative capacity of the environment. The importance of environmental protection has brought emissions reductions to the awareness of policy makers in many developing countries. However, emissions reductions are usually tied to short-run growth cuts. In developing countries, where economic growth has been the focus, how can local bureaucrats be motivated to implement an emission control policy, and how would it affect economic growth?

This paper investigates how local bureaucrats respond to a performance evaluation system targeting emissions control in China. China has been the fastest-growing major economy over the past three decades, but this growth has come at the cost of heavy environmental pollution.[1] The central government realized the pressing environmental issues and proposed a major policy initiative in 1998 to establish control zones in regions with the highest levels of sulfur emissions or acid rain. The main objective of the Two

[1] Twelve of the twenty most polluted cities in the world are located in China, and only 1% of the country's 560 million city dwellers breathe air considered safe by the European Union (World Bank, 2007). The total economic cost of China's air pollution was assessed at 4.1% of the GDP in 2002 (Wen and Chen, 2008), and the total economic loss of water and air pollution was evaluated at nearly 6.9% of the GDP in 2003 based on the willingness to pay approach (World Bank, 2007). China consistently ranks in the bottom quintile for ambient air quality according to the World Health Organization.

Control Zones (hereafter, TCZ) policy was to reduce SO_2 emissions, which have long been a major contributor to China's ambient air pollution.[2]

Contrary to the initial policy design, the effect of the TCZ policy on SO_2 emissions reductions was small and temporary. After a short dip in 1998-1999, SO_2 emissions continued to rise and reached a historical high in 2005-2006 (see Fig. 1).[3] The unsuccessful implementation of the TCZ policy led the central government to reconsider its strategy. In 2005, new administrative regulations were imposed, and the emissions quota was included in the performance evaluation system for top local bureaucrats (city mayors and Party secretaries), making them subject to administrative sanction or removal from office if the quota was not met. This new administrative enforcement was the first time that an emissions quota was included in the performance evaluation system and brought significant changes to the lives of these local bureaucrats.

[Insert Figure 1 here]

For empirical analysis, we collected a rich set of data on 286 prefectural cities, among which 162 are TCZ cities and the rest are non-TCZ cities, and implement a difference-in-differences (DD) identification framework. We find that TCZ prefectural cities exhibited a similar pattern of reduction in SO_2 emissions before 2005 compared to the control group (non-TCZ cities). After SO_2 reductions became part of the performance measures in 2005, however, we find that TCZ cities experience larger emissions reductions as they face more stringent reduction goals. In the meantime, we also find that the gross domestic product (GDP) growth rates in TCZ cities significantly decreased compared to that of the control group after the policy change. These results imply that the new evaluation system in 2005 introduced a substantial trade-off between SO_2 emissions and GDP growth. A back of envelope calculation shows that for every 1 tonne of reduction in industrial SO_2 emissions at city level, it's GDP would be reduced by 190 thousand Chinese Yuan.[4] A further welfare analysis suggests that the focal policy led to 833 lives saved, which is valued at 487.3 million in 1997 Chinese Yuan based on the World Bank's valuation of a statistical life. The fact that the cost of the program exceeds its benefit suggests that the Chinese government is willing to pay a premium for immediate and sizable reductions in

[2]The ambient concentration of SO_2 in many Chinese cities was among the highest in the world when China's coal consumption peaked in 1995. The damage from acid rain in the early 1990s was estimated to be about 0.7% of the GDP (World Bank, 1997).

[3]In Appendix Figure A.1, we further plot the time trends of SO_2 emissions for TCZ and non-TCZ cities, respectively. However, due to the data availability, we can only start the analysis from 2001 onward (for data details, see Section 3.2). We find similar upward increasing trends of SO_2 emissions for TCZ and non-TCZ cities before 2005. After 2005, SO_2 emissions stagnated in non-TCZ cities but declined in TCZ cities.

[4]To put this number into perspective, in 1999 the average emission charge on industrial SO_2 at the national level was 46.37 Chinese Yuan per tonne (He, 2005).

SO_2.

Our results are robust to a battery of specification checks, such as a placebo test with a random assignment of TCZ cities, an alternative measurement of pollution reduction, a check on spatial spillover, the inclusion of a regional aggregate linear trend, a combination of DD and propensity score matching (PSM), the inclusion of time-varying controls, and the exclusion of large cities (such as Beijing, Shanghai, etc.), as well as the 2008 Summer Olympic Games venue cities.

To shed light on the substitution of efforts between pollution reduction and economic development, we analyze the channels through which the policy was at work. In particular, we closely examine the actual policies implemented by local bureaucrats in order to reduce SO_2 emissions. By using detailed plant-level data from the electric power sector, we find that most of the measures for reducing SO_2 emissions were enforced by local bureaucrats, which explains the significant drop in SO_2 emissions in TCZ cities after 2005 compared with non-TCZ cities. Meanwhile, some of these measures (such as the closure and establishment restrictions on thermal power plants, control of growth in dirty sectors, etc.) had negative implications for economic development.[5] We then conduct a thorough content analysis of each city government's *Annual Work of the Government Report* from 2001 to 2015 to investigate how local bureaucrats revealed their efforts for pollution reduction and economic growth. We find that TCZ city governments became more concerned about environmental protection but had lower target GDP growth rates after 2005 than non-TCZ city governments. These results further indicate that local bureaucrats in TCZ cities were willing to trade off GDP growth to achieve the more stringent emissions reductions goal.

Our findings contribute to the literature in several ways. First, our paper adds new empirical evidence to the economics literature on incentivizing bureaucrats with concrete performance targets. Public management literature has long been trying to understand the effect of a concrete performance target on the behaviors of local bureaucrats (OECD, 1994; Hood, 2007). In the economics literature, mixed evidence has been generated in previous studies. Lockwood and Porcelli (2013) show that the comprehensive performance assessment system in England has increased public service quality but has no significant effect on efficiency overall. Recent studies in Nigeria (Rasul and Rogger, 2015, 2017) suggest that concrete targets tend to decrease public service delivery.[6] Our paper

[5] A dirty sector is defined by matching the 2-digit industry classification code to that of major SO_2 pollution sectors identified by the *China Pollution Source Census 2007*. The 2007 census, the only pollution source census in China up to this date, examined 5.92 million pollution sources across all provinces in China. The final report identified six industries (as shown in Appendix Table A.1) that rank the highest in terms of SO_2 emissions share in total industrial SO_2 emissions. These six industries comprise 88.5% of total SO_2 industrial emissions.

[6] The failure of performance-based incentives is usually attributed to the multi-task nature of bureaucratic jobs and bureaucrats' intrinsic motives (Dixit, 2002; Alesina and Tabellini, 2008). As bureaucrats

provides empirical evidence for the effectiveness of a concrete performance target in an environmental regulation setting. In particular, our results suggest that a target-based evaluation system can be effective in motivating bureaucrats to shift their efforts toward the tasks that are more heavily weighted by the evaluators and seek to achieve the targeted policy output. The magnitude of our estimates also suggests that bureaucrats appear to be willing to trade off substantial GDP growth for SO_2 emissions reduction.

Our paper adds to the understanding of multitasking agency problems for government agents. Career concern theories suggest that top bureaucrats are largely driven by the outcomes of their mandated tasks (Holmstrom and Milgrom, 1991; Dewatripont et al., 1999; Alesina and Tabellini, 2007, 2008). Under this framework, bureaucrats choose their effort levels and their distribution of efforts across tasks to maximize their signaled capability to the market (or the principal). Specifically, our results speak to two theoretical predictions made by Wilson (1989), Dewatripont et al. (1999) and Dewatripont et al. (2000). The first prediction is that bureaucrats' accountability increases as their mission becomes more "focused"and "clear". Compared to the initial establishment of the TCZ in 1998, the new administrative policy in 2005 brought more clarity to the local bureaucrats' mission agenda. And this narrowly defined goal increases efforts exerted by bureaucrats for controlling emissions. Our results support this prediction by showing that local bureaucrats implemented a series of policies that specifically targeted SO_2 emissions reduction after the administrative enforcement. The second prediction that relates to our results is that when effort is encouraged on one task, efforts on other tasks would decrease. Our results show that compared to non-TCZ cities, real GDP growth significantly decreases in TCZ cities after the new administrative policy was in place in 2005, indicating a shift of attention from GDP growth to emissions control by local bureaucrats.

Our paper addresses an important yet relatively neglected topic in the environmental economics literature for developing countries - the empirical relationship between stringent environmental regulations and GDP growth. The effect of environmental regulations on growth mainly stems from three channels. The first is the direct cost of regulations. For example, in 2015, the total government expenditure on environmental protection among the European Union member countries amounted to 0.4% of the GDP (EuroStat, 2017), and the Chinese government designated 7.5 billion U.S. dollars for environmental protection in 2017 (MEPC, 2017). Diverting government resources towards enforcement and compliance of environmental regulations would potentially hurt economic growth. Secondly, regulations might lead to shutdowns of certain plants and thus unemployment.

usually juggle multiple tasks or multiple dimensions of a task, concrete performance targets potentially cause social distortion by requiring bureaucrats to exert effort on less productive outcomes (Rasul and Rogger, 2017). Emphasizing output as a performance measure would also likely crowd out the intrinsic motivations for bureaucrats (Benabou and Tirole, 2003).

Empirical evidence in the U.S. suggests that stringent regulations result in significant job loss in both the short and long run (Greenstone, 2012; Walker, 2013). The third channel is firm competitiveness. Conventional economic theories suggest that, at the firm level, stringent environmental regulations result in the relocation of resources from more productive activities to pollution abatement, and they decrease productivity (Xepapadeas, 2005). The opposite view, as suggested by Porter and van der Linde (1995), argues that in the long run, improved efficiency and innovation might increase productivity due to technological enhancements.[7] Empirical literature on the relationship between environmental regulations and economic growth is rather limited in developing countries. Existing studies usually focus on estimating regulations' impacts on firm competitiveness (Yuan et al., 2017; Wang and Shen, 2016; Zhao et al., 2015). At the aggregate level, less is known about the impact on economic growth. Our results address this gap by directly estimating the effect of a stringent environmental regulation on GDP growth. We show that bringing a SO_2 emissions cut to the performance evaluation system leads to a 0.846 percentage point drop in GDP growth per annum.[8] Our results indicate that the impact of environmental regulations on GDP growth can be both significant and long lasting.

Our findings also contribute to the small but growing literature on what motivates local governments to adopt more stringent environmental policies. Rigor and willingness to implement environmental regulation at the local government level are usually associated with political connections, promotion probability and other political concerns (Zheng et al., 2014; Jin and Lin, 2014; Kahn et al., 2015).[9] Our paper shows that the readiness of the central government to reward and punish local officials based on their environmental performance could also motivate local government officials to reach the green goal.[10]

The remainder of the paper is organized as follows. Section 2 provides a review of the policy background on the TCZ policy and performance evaluation system. We

[7]Empirical studies have provided inclusive evidence for the Porter Hypothesis. For example, a recent meta-analysis, which was conducted by Cohen and Tubb (2017), showed that environmental regulations display heterogeneous impact on productivity at the firm and industrial level. The impact size was significantly correlated with the types of regulations imposed.

[8]To put this number into perspective, Jorgenson and Wilcoxen (1990) showed that the aggregate impact of environmental regulation on U.S. economic growth in the 1980s was about 0.191 percentage point.

[9]Zheng et al. (2014) show that a city mayor's environmental performance is related to his or her promotion probability. Jin and Lin (2014) focus on explaining the relationship between political connections and environmental pollution. She suggests that local governments are possibly motivated by their environmental concerns. Kahn et al. (2015) study the behavior of local governments when they are incentivized to reduce water pollution and provide a mechanism that can resolve cross-border pollution. Our paper focuses on evaluating the impact of a public management incentive system. Although all these papers address environmental pollution and promotion in China, the research questions and focus are very different.

[10]Our results complement and echo the literature evaluating the target-based vertical control system (see Xu (2011) for a review).

then present our empirical strategy, data descriptions and results in sections 3 and 4. Mechanisms and welfare analyses are presented in sections 5 and 6.

2. Policy background

2.1. Administrative Enforcement of SO_2 Reductions Targets

SO_2 emissions have long been a major contributor to China's ambient air pollution, and reducing SO_2 emissions has been a priority for China's environmental authorities since the 1990s. Beginning with the 9th Five-Year Plan (i.e., from 1996 to 2000) in 1996, China's central government began to set limits on total SO_2 emissions. In 1998, the State Council instituted an SO_2 reduction program (henceforth referred to as the 1998 policy) to limit ambient SO_2 pollution and to curb the growing incidence of acid rain. The 1998 policy encompassed two target policy areas: the SO_2 Control Zone, covering cities in North China, and the Acid Rain Control Zone, covering cities in South China. In aggregate, the TCZ regions cover 11% of China's territory (see Fig. 2) and are responsible for more than 60% of China's total SO_2 emissions.[11] The 1998 policy also laid out long-term SO_2 emissions reduction goals. Specifically, it stated that by 2000, SO_2 emissions from TCZ cities should be under the quota set by the central government.[12] By 2010, SO_2 emissions should not be higher than 2000 levels, and air quality in major TCZ cities needs to meet the national standard by 2010 (SO_2 air concentration below 60 $\mu g/m^3$). To reach the total SO_2 emissions reduction targets (also known as the green goal), the Chinese central government imposed nationwide emissions quotas in 2000, with more stringent measures applied to TCZ cities. A total emissions control policy was incorporated into China's 10th Five-Year Plan (i.e., from 2001 to 2005) in 2001 to work in concert with the TCZ policy.[13]

[Insert Figure 2 here]

However, the implementation of the TCZ policy was not successful. The effect of the TCZ policy on SO_2 emissions reductions was small and temporary. Although total SO_2 emissions leveled off between 1997 and 2002, they quickly picked up speed afterward. By 2005, the total emissions level had not declined but instead, had increased by 28% from the 2000 level (see Fig. 1).

To force localities to adhere to the national policy, administrative enforcement was put into place by the end of 2005 (henceforth also referred to as the 2005 policy) as part of the 11th Five-Year Plan (i.e., from 2006 to 2010). Starting in Dec 2005, local

[11] The relevant Chinese document for the TCZ policy is available at the following government website: `http://www.mep.gov.cn/gkml/zj/wj/200910/t20091022_172231.htm`

[12] The actual SO_2 emissions in 1995 was 23.7 million tons. The target SO_2 emissions for 2000 set by the central government was 24.6 million tons.

[13] The relevant Chinese document for the 10th 5-Year Plan guideline for the TCZ policy is available at the following government website: http://www.ynepb.gov.cn/xxgk/read.aspx?newsid=11012.

government leaders (mayors and Party secretaries) would be held accountable for reaching the environmental protection goals set by the central government in their administrative region, including the SO_2 reduction goal.[14] The top-down target-based responsibility scheme ensures that local officials are tied to satisfying higher-level mandates for career advancement and legitimacy.[15] This was the first time that an emissions quota targeting a specific pollutant was introduced into their performance evaluation system. A follow-up announcement was made by the central government in 2007 specifically about the emissions reductions targets, which emphasized that the SO_2 emissions reduction targets to be attained by the end of 2010 would be incorporated into the responsibility contracts signed with upper-level governments.

After the central government imposed the administrative enforcement at the end of 2005, the central government's consistent emphasis on the need to reduce SO_2 emissions has made it incrementally more difficult for local governments and polluters to drag their feet in carrying out the central government's SO_2 emissions reduction policy. By 2010, the SO_2 emissions reduction targets were finally met, and an overall 14.29% reduction was achieved nationwide (see Fig. 1). The target-based performance system was carried through the 12th Five-Year Plan (i.e., 2011-2015) and by 2015, a further national total SO_2 emissions reduction target of 8% was also achieved.

2.2. SO_2 emissions reductions and career concerns

In this subsection, we discuss a conceptual framework and present descriptive evidence for the link between the target-based performance evaluation system and emissions reductions. Following the seminal work by Dewatripont et al. (1999) and Alesina and Tabellini (2008), bureaucrats are partly motivated by their career concerns. Specifically, career concern theories suggest that top bureaucrats are largely driven by the outcomes of their mandated tasks. Under this framework, bureaucrats choose their effort levels and their distribution of efforts across tasks to maximize their signaled capability to the non-public sectors. In addition to career motives, studies show that bureaucrats are motivated by peer pressure, social norms and bureaucrats' own *glow of warmth* (Wilson, 1989; Maskin and Tirole, 2004).

In China, the most important career concern for bureaucrats is to move up the hierar-

[14] The relevant Chinese document is available at the following government website: http://www.gov.cn/zwgk/2005-12/13/content_125680.htm

[15] In the meantime, the central government banned new construction projects in regions with unsatisfactory SO_2 emissions reductions, in an effort to align incentives for economic growth with local governments' achievement of the green goal. Given that the achievement rate of the SO_2 emissions target was very high (i.e., 98% for the TCZ cities; see Table 1), the policy was not binding. Nonetheless, to alleviate the concern that our estimates may be caused by the banning of new construction projects, we conduct a robustness check by including the logarithm of fixed investment per capita in Appendix Table A.2, column 1. We find that our estimates are robust to this control, further suggesting that our results are not affected by the banning policy.

chical leadership ladder. To signal their ability to the upper level supervisors, bureaucrats need to know the most important task and weights evaluated by the supervisor. Since the economic reform in the 1970s, the central government has been incentivizing local governments to “get rich first.” The GDP growth has been the most important evaluation criterion for local bureaucrats in China. Empirical evidence suggests that the promotions and bonuses of local bureaucrats are highly correlated with regional GDP growth (e.g., Li and Zhou, 2005).

The 2005 policy was the first time that a pollution emissions quota was explicitly mentioned in the performance evaluation criteria. Given that this policy contained significant measures for promotion, SO_2 emissions reductions became one of the most important factors in the assessment of local cadres.[16] Local government leaders who fail to meet the annual SO_2 emissions reduction targets would be penalized with an administrative sanction or removal from office.[17] The administrative enforcement ensures the effectiveness of the TCZ policy across the chessboard of China's territorial administration, and further ties local cadres to their fulfillment of the target SO_2 emissions reductions.

Table 1 presents the emissions reductions goals for the 11th (i.e., from 2006 to 2010) and 12th Five-Year Plans (i.e., from 2011 to 2015), as well as the average achievement rate. We present the numbers for TCZ and non-TCZ cities, separately. The information was compiled by the authors from various government documents to cover as many cities in our sample as possible. Note that although the reduction goals for TCZ cities are much more aggressive than those for non-TCZ cities, the achievement rates are much higher in TCZ cities. The average achievement rate for TCZ cities has remained at 98% since 2005.

We cannot provide direct evidence for the link between removal from office with failing to achieve emissions reduction goals, as local bureaucrats could be removed from office for multiple reasons (corruption, scandals etc.). However, anecdotal evidence suggests that the veto power mentioned in the policy was effective. For example, during a press release conference held by the State Council Information Office, a government official mentioned that the veto power is a signal from the central government that pollution control is no longer a hand-waving gesture but something the government really cares about. It is also

[16]Other important performance measures for local bureaucrats include GDP growth, implementation of the One Child Policy and fatal accidents etc.. GDP growth has been shown to be highly correlated with promotions and bonuses for local bureaucrats (Chen et al., 2005). Fatal accidents as a measure of performance were first brought up in a provision issued by the central government in 2001, which declared that “when an extremely fatal accident happens, the local government officials would be penalized based on the severity of the accident.”This measure is important because it is tied to the economic development and stability of society that the Party has been trying to achieve. The One Child Policy has been declared as the “basic state policy” since the late 1970s and was closely tied to performance evaluations of local bureaucrats (Feng et al., 2013).

[17]The relevant Chinese document is available at the following government website: http://www.gov.cn/jrzg/2007-06/03/content_634545.htm

mentioned that if local bureaucrats fail to achieve the reductions goals, the bureaucrats would not be eligible to run for the outstanding performance award among their peers.[18]

[Insert Table 1 here]

To summarize, our evidence suggests that the policy change in 2005 affected the behavior of local bureaucrats. To achieve the emissions reductions goals set by the Five-Year plans, local bureaucrats weigh the pollution reduction against GDP growth. In the following sections, we formally investigate and quantify the effect of the target-based performance evaluation system in the 2005 policy on emissions reductions and economic growth. Meanwhile, we also conduct a welfare analysis of the policy.

3. Estimation strategy

3.1. Estimation framework

To investigate whether the performance-based evaluation system based on SO_2 emissions reduction introduced in 2005 affects government officials' behavior in implementing environmental regulations and promoting economic growth, we conduct a DD analysis by comparing outcomes in TCZ cities (those cities with strict environmental regulations) with those of non-TCZ cities before and after 2005. Specifically, our regression equation is

$$y_{ct} = \beta TCZ_c * Post05_t + \delta_c + \gamma_t + \mathbf{Z_c} * \mathbf{f(t)} + \varepsilon_{ct}, \quad (1)$$

where y_{ct} contains our outcomes of concern (logarithm of the SO_2 emissions and the GDP growth rate) in city c at year t; TCZ_c equals 1 if city c was designated a TCZ city in 1998 and 0 otherwise; $Post05_t$ takes a value of 1 if $t > 2005$ and 0 otherwise; δ_c is the city fixed effect, capturing all time-invariant differences across cities; γ_t is year fixed effect, capturing all yearly factors that are common to cities such as macro-level shocks; $\mathbf{Z_c}$ denotes the determinants of TCZ selection measured in the pre-treatment period and $\mathbf{f(t)}$ is a third-order polynomial function (both explained later); and ε_{ct} is the error term. To accommodate potential heteroskadesticity and serial correlation, we cluster the standard errors at the city level, following the suggestion by Bertrand et al. (2004).

We expect β to be negative for the outcome of SO_2 emissions (in the logarithm) if the performance-based evaluation system is effective, as TCZ cities were subject to a larger emissions reductions target in 2000. Thus the new evaluation system introduced in 2005 would induce more reductions in TCZ cities. Meanwhile, we also expect β to be negative for the GDP growth rate as after the 2005 policy, local bureaucrats would reallocate efforts

[18] The relevant interview information is available at `http://www.scio.gov.cn/xwfbh/xwbfbh/wqfbh/2007/1129/Document/324843/324843.htm`

from promoting growth to reducing pollution given the substantial career consequences of the policy.[19]

However, the unbiased estimation of β requires that the logarithm of SO_2 emissions and GDP growth in TCZ cities would have followed the same time trend as those in non-TCZ cities after 2005 if a performance-based evaluation system had not been imposed by the central government in 2005. A primary threat to our identifying assumption is that cities were not randomly selected into TCZ and non-TCZ groups in 1998. This non-random selection implies that TCZ and non-TCZ cities could have been systematically different before 1998. If the values of these unobserved differences changed in 2005, then our estimate would be biased. To address this concern, we follow an approach used by Gentzkow (2006), in which key determinants of the TCZ selection are first identified and then post-treatment variations in the outcome variable generated by these key determinants are flexibly controlled in the regression. The premise of this approach is that conditional on these key determinants, the central government did not select TCZ cities based on remaining unobservables with a perspective that there would be a new performance-based evaluation system eight years later and these remaining unobservables would behave differently for TCZ and non-TCZ cities at that time.

Given that the designation of TCZ cities was initiated by the central government in the mid-1990s, we look at government policy documents to understand the factors that shaped the government's decision-making processes for selecting TCZ cities. Specifically, according to the Air Pollution Prevention and Control Law (APPCL) amended in 1995, based on conditions in the atmosphere, terrain and soil, cities could be designated as either an SO_2 Control Zone or an Acid Rain Control Zone. In addition, the 1998 TCZ policy classified those cities that were already seriously polluted by SO_2 or acid rain as TCZ cities. We collected two sets of variables as our TCZ selection variables. The first set of variables reflects the conditions of the natural environment based on the 1995 APPCL amendment: *Roughness* (the standard deviation of slope), *Elevation* (average elevation in kilometers), *Wind Speed* (annual average wind speed in 1990-1995), *Precipitation* (annual average precipitation in 1990-1995) and *Soil pH* (average pH level of topsoil). The second set of variables reflects pre-policy SO_2 pollution levels, which include *Coldness* (percentage of days with a temperature of 5^oC or below in 1990-1995) and the total employment in dirty sectors in 1996.[20] To flexibly control for trends in SO_2 emissions

[19] Policies implemented that target SO_2 emissions reduction would also induce a mechanical trade-off between emissions and GDP growth.

[20] Ideally we should include pre-policy SO_2 levels as a determinant for TCZ designation for each city, but unfortunately, the data are not available. Therefore, we use pollutants that are known to be highly correlated with SO_2 and employees in the dirty sector as a proxy for baseline SO_2 (see footnote 5 for the definition of dirty sectors.). When the daily outdoor temperature drops to 5°C or below for a few days, the northern region of China enters the heating period. As a result, coal combustion in boilers is associated with the release of air pollutants that are highly correlated with SO_2.

generated by these selection variables over time, we interact them with a third-order polynomial function of time $\mathbf{Z_c} * \mathbf{f(t)}$.

It is possible that the implementation of the TCZ policy itself could have altered the characteristics of the treatment and control groups so that they became incomparable when the target-based evaluation system was in place by the end of 2005. Specifically, existing studies have shown that the implementation of the TCZ policy in 1998 affected foreign investment (Cai et al., 2016), exports (Hering and Poncet, 2014) and health (Tanaka, 2015). To address this concern, we add a set of variables that are known in the literature to be affected by the TCZ policy: specifically, city-level foreign direct investment (FDI), mortality and export sales data, interacted with *Post05*.[21]

Appendix Table A.3 shows the balancing between TCZ and non-TCZ cities to shed light on whether controlling for TCZ selection variables improves the matching between our treatment and control groups. Panel A shows the seven selection variables. We find that TCZ and non-TCZ cities differed significantly in many of these variables, such as the percentage of days with a temperature of 5^oC or below in 1990-1995, annual wind speed in 1990-1995, total employment in dirty sectors in 1996, and annual average precipitation in 1990-1995. Panel B compares TCZ and non-TCZ cities on various economic and social development variables in the first year in our sample (i.e., 2001).[22] We find that without controlling for TCZ selection variables, there were significant differences in initial conditions between TCZ and non-TCZ cities. However, once conditional on TCZ selection variables, the differences becomes statistically and economically insignificant. These results indicate that our treatment and control groups are balanced with the controls of TCZ selection variables, which is crucial for our identification.

As further checks on our identifying assumption in equation 1, we conduct several sensitivity exercises, including an event study of whether the outcome variables had similar pre-treatment trends between TCZ and non-TCZ cities, a placebo test with the random assignment of TCZ cities, an alternative measurement of pollution reduction, a check on spatial spillover, the inclusion of the regional aggregate linear trend, a combination of DD and PSM, the inclusion of time-varying controls, and the exclusion of large cities (such as Beijing, Shanghai, etc.) as well as the 2008 Summer Olympic Games venue cities. For details on these robustness checks, see Section 4.2.

3.2. Data and variables

Data for the empirical analyses presented in this section were collected by the authors from various official statistical publications and public databases. Using these data sources, we construct a data set containing environmental, socioeconomic and meteoro-

[21] All are averaged between 2001 and 2005

[22] See footnote 5 for the definition of dirty sectors.

logical conditions of each city spanning the period from 2001 to 2015, if not otherwise mentioned. Detailed variable definition and summary statistics are reported in Table 2.

[Insert Table 2 here]

City-level SO_2 emissions and GDP growth. The city-level industrial SO_2 emissions data were collected from the provincial-level *China Statistical Yearbooks*, supplemented by the *China City Statistical Yearbooks 2004-2016* and *China Environmental Yearbooks 2002-2016*.[23] We use emissions data between 2001 and 2015 for the analyses presented in this paper. Data before 2001 were subject to significant missing values, and thus were not included in the study. The city-level annual industrial SO_2 emissions were calculated by summing up emissions from all reporting firms in a given year. The State Environmental Protection Agency (SEPA) records each reporting firm's SO_2 emissions through direct monitoring. If direct monitoring is not feasible, the SEPA uses reverse engineering measures computed from the industrial waste reported by the firm. The list of reporting firms is determined by the SEPA and covers more than 85% of the major polluting firms in the city, and is subject to minor revisions each year. The monitoring process follows a strict protocol and is subject to internal auditing processes. It is very unlikely that the data were systemically manipulated via collusion between SEPA officials and local bureaucrats. Data on city-level real annual GDP growth rate were collected from the *China Statistical Yearbooks for Regional Economy*, supplemented by the provincial-level *China City Statistical Yearbooks 2004-2016* and the *China Statistical Yearbooks*.

City-level control variables. The city slope and elevation were extracted from the *Shuttle Radar Topographic Mission 90m Digital Elevation Model* data. We collected city wind speed, temperature and precipitation data from *the China Meteorological Data Sharing Service System* (http://cdc.cma.gov.cn/). Soil pH data were extracted from the grid data of the *Global Dataset of Derived Soil Properties* at 0.5 degree by 0.5 degree. Data on total employees in the dirty sectors are from the *China Establishment Census 1996*. We aggregate the firm-level data by sector and city.[24]

Export sales data are collected from China's customs data 2001-2005. The dataset covers the universe of all export transactions by Chinese exporters. Specifically, it includes product information (at the HS 8-digit level), trade value, identity of Chinese exporters, and export destinations. We first match the HS classification to the Chinese Industry Classification using the concordance table from the National Bureau of Statistics of China. We then aggregate the export value from the 8-digit HS-product level to the 2-digit industry and city levels.

[23] According to the *China Pollution Source Census 2007*, industrial activity is responsible for 91.4% of total SO_2 emissions. The remaining 8.6% is contributed by domestic activity.

[24] See footnote 5 for definition of dirty sectors.

Other city-level control variables, including GDP per capita, investment per capita, FDI, mortality, manufacturing share of total GDP, trade intensity (ratio of imports plus exports to GDP) and coal abundance were collected from the *China Statistical Yearbooks for Regional Economy*, supplemented by the *China City Statistical Yearbooks 2004-2016* and the provincial-level *China Statistical Yearbooks*.

4. Main results

4.1. Baseline estimates

Table 3 reports the main results from equation 1. Column 1 presents results from a simple DD estimation for SO_2 emissions with only the city and year fixed effects included. The interaction between TCZ_c and $Post05_t$ is negative and statistically significant, suggesting that SO_2 emissions fell more significantly in TCZ cities than in non-TCZ cities after 2005. Given that the TCZ policy imposed stricter regulations on SO_2 emissions in TCZ cities than in non-TCZ cities, and that a performance-based evaluation system was imposed in 2005, this result implies that the new evaluation system has disciplined or incentivized bureaucrats to enforce the environmental regulations.

[Insert Table 3 here]

To alleviate the concern that our estimate is biased due to the non-random selection of TCZ cities in 1998, we include interactions between determinants of TCZ selection and a third-order polynomial function of time in column 2. We continue to find a negative and statistically significant estimate of $TCZ_c * Post05_t$. In column 3, we further include three interactions between $Year05_t$ and FDI, mortality and dirty export sales, respectively, to address the concern that the TCZ policy in 1998 made TCZ and non-TCZ cities non-comparable before 2005. We continue to find a negative and statistically significant coefficient. In Appendix Table A.2, column 2, we estimate the policy effect on the level of SO_2 emissions, and uncover a negative coefficient, despite the lack of statistical significance due to the large standard errors.

In columns 4 to 6, Table 3, we examine the effect of the new performance evaluation system on GDP growth. Column 4 reports estimation from the baseline DD with only the city and year fixed effects. Columns 5 and 6 report results when more controls are included. We find negative and statistically significant estimated coefficients of the interaction between TCZ_c and $Post05_t$ across all specifications. These results indicate that the new evaluation system has led to a significant decrease in economic growth.

Economic magnitude. To gauge the economic significance of our estimates, we report the control means in Table 3. For SO_2 emissions, our estimate in column 3 shows that the 2005 policy reduces the logarithm of SO_2 emissions by 0.146, which corresponds to a reduction in the level of SO_2 emissions by 14.6%. Given that the control mean of the logarithm of SO_2 emissions is 1.338 (which implies a level of SO_2 emissions as 3.811 with a unit of 10,000 tonnes), our estimate suggests that the policy leads to a decrease in SO_2 emissions by $3.811 \times 0.146 \times 10000 = 5,564$ tonnes.

For the GDP growth rate, our estimate in column 6 shows that the policy reduces the GDP growth rate by 0.846 percentage point. With the control mean of annual growth rate at 12.17%, this number suggests that the 2005 policy reduced the annual GDP growth rate by $0.846/12.17 = 6.9\%$.

4.2. Robustness checks

Our estimates hinge on the identifying assumption that cities designated to be a TCZ in 1997 would not have had differential trends in SO_2 emissions or GDP growth starting in 2005. To further verify that our identifying assumption is sustained, in this section, we report a battery of robustness checks.

Time trends before the treatment. A necessary condition for satisfying our identifying assumption is that TCZ and non-TCZ cities have similar time trends in the outcomes before the treatment. Specifically, if similar time trends existed for a long period before the treatment and diverged after the treatment, it may indicate that the differential trends may be caused by the treatment. To this end, we apply an event study approach along the lines described by Jacobson et al. (1993). Specifically, we estimate the following equation

$$y_{ct} = \sum_{t=2001}^{2015} \beta_t TCZ_c * \gamma_t + \delta_c + \gamma_t + \mathbf{Z_c} * \mathbf{f(t)} + \varepsilon_{ct}, \qquad (2)$$

in which β_t is a series of estimates from 2001 to 2015. In particular, with the omitted time category being 2005, $\{\beta_t\}$ captures the differences in outcomes between TCZ and non-TCZ cities relative to the difference just before the policy launch.

Fig. 3a plots the estimates for the outcome of SO_2 reduction along with the 95% confidence interval. We find that the estimates are positive and are stable in the magnitudes from 2001 to 2004, suggesting a common trend between TCZ and non-TCZ cities before the treatment. The estimates for the GDP growth rate are plotted in Fig. 3b. Despite some fluctuations, all the estimates were positive from 2001 to 2004 without a clear trend.

[Insert Figures 3a and 3b here]

Figs 3a and 3b also allow us to visualize the marginal effect of the 2005 policy by year. In Fig 3a, the difference dropped to negative immediately after the adoption of the treatment in 2005 and continued to be negative with increasing magnitude. These results are consistent with the hypothesis that the policy change in 2005 generated the differential trends in SO_2 emissions between TCZ and non-TCZ cities. In Fig. 3b, interestingly, the estimates for 2006 and 2007 were also positive and similar in magnitude to those in the pre-treatment period, and they became negative since 2008. These results suggest some lagging effects of the 2005 policy change on the GDP growth rate.

A placebo test with randomization of the treatment. As a second robustness check, we randomly draw 162 cities out of the total of 286 cities in our sample and assign TCZ status TCZ_c^{false} to these cities. Similarly, we randomly assign the timing of the adoption of the new performance-based evaluation system $Post_t^{false}$ to these cities, and then construct a new regressor of interest $TCZ_c^{false} \times Post_t^{false}$. Given this random data-generating process, a necessary condition for satisfying our identifying assumption is that $TCZ_c^{false} \times Post_t^{false}$ shall produce zero policy effect on the SO_2 emissions and the GDP growth rate; otherwise, it indicates that our estimation equation is misspecified. To increase the power of this test, we conduct the random assignment 1,000 times. The distributions of 1,000 estimated coefficients for the SO_2 emissions and the GDP growth rate are plotted in Appendix Figures A.2(a) and A.2(b), respectively. Both distributions are narrowly centered around zero, and less than 1% of the estimates are more negative than our true baseline estimates in Table 3. These graphs provide further support for our identification strategy.[25]

Inclusion of regional linear time trend. To control for the aggregate time trend, we insert the regional linear time trend in our baseline equation (1), i.e., $\lambda_r \times t$, where the regions include Northeast, East, Central and West China. Estimation results are reported in Table 4, column 1, with panel A for SO_2 emissions and panel B for the GDP growth rate. We find consistent estimates, suggesting that our estimates are not affected by the aggregate trends.

[Insert Table 4 here]

Exclusion of large cities and Summer Olympic Games venue cities. Cities in China differ in hierarchical significance. In particular, our regression sample includes four municipalities directly under central government control (i.e., Beijing, Shanghai, Tianjin and Chongqin). Meanwhile, the 2008 Summer Olympic Games were held in China, which had significant impacts on the economy and society. To examine whether our estimates are driven by these municipalities and the 2008 Summer Olympic Games venue cities, we exclude them and re-estimate the equation. Estimation results are reported in Table 4, column 2. We again find similar results, suggesting that our results are not driven by influential cities.

Spatial spillovers. One concern for identification is that non-TCZ cities bordering

[25] See Chetty et al. (2009) and La Ferrara et al. (2012) for a review and application of the method.

TCZ cities would also be affected by the new evaluation system, as polluting firms could be easily moved from TCZ to non-TCZ cities that border each other. To rule out this possibility, we exclude non-TCZ cities bordering TCZ cities in the analysis. As shown in Table 4, column 3, the estimated coefficients for SO_2 emissions and GDP growth rate remain negative and statistically significant. These results provide evidence that our results are not driven by bordering cities.

An alternative measurement of pollution reduction. Note that SO_2 emissions data were monitored and collected by the SEPA. The whole process followed a strict protocol and was subject to internal auditing, making it difficult for local bureaucrats to manipulate the data. However, one may still be concerned that there might be some collusion between SEPA officials and local bureaucrats, and the SO_2 emissions data were inaccurate. To address this concern, we use PM2.5 as an alternative measure of SO_2 emissions. PM2.5 particles can be directly emitted or produced from emitted precursors. Pollutants such as SO_2, nitrate and volatile organic compounds (VOCs) can react in the atmosphere to produce PM2.5. Although changes in PM2.5 can be a complicated process, PM2.5 can be used as a good alternative measure for SO_2 emissions (Hoden and Barnard, 2015). In our setting, we find a correlation coefficient of 0.22 between the logarithm of the SO_2 emissions and the logarithm of PM2.5.

The PM2.5 data were extracted from the grid data of *Global Annual PM2.5 Grids* from Moderate Resolution Imaging Spectroradiometer (MODIS) and Multi-angle Imaging Spectro Radiometer (MISR) Aerosol Optical Depth, which provides a continuous surface of concentrations of particulate matter of 2.5 micrometers or smaller.[26] The grid data were derived by researchers in the U.S. from satellite data provided by NASA; thus, the source of the PM2.5 data is completely independent of the SO_2 emissions levels reported by the SEPA. Estimation results of the policy effect on PM2.5 particles are reported in Table 4, column 4. The estimated coefficient is negative and statistically significant. A decrease in PM2.5 as a result of the 2005 policy, in the absence of regulations for the other reactors, supports our hypothesis that the policy effectively reduced SO_2 emissions.

Inclusion of additional time-varying city characteristics. The inclusion of $Z_c * f(t)$ in Table 3, column 2 shows a significant drop in the estimated magnitude (i.e., about 36%). As $Z_c * f(t)$ controls for the flexible time trends in outcomes between TCZ and non-TCZ cities due to the differences in initial conditions, the drop in magnitude may imply that the time-varying characteristics of the TCZ and non-TCZ cities differ. To address this concern, we further include several time-varying city characteristics such as log GDP per capita, share of manufacturing sectors, coal abundance and trade intensity.

[26]The raster grids have a grid cell resolution of 30 arc-minutes and cover the world from 70°N to 60°S latitude.

We first replace the cubic time trends of the initial conditions with these time-varying controls in Appendix Table A.2, column 3. We continue to find a negative and statistically significant estimate, with the decrease in magnitude about 24%. These results confirm that our selection of time-varying city characteristics capture important differential trends in outcomes between TCZ and non-TCZ cities. We then further add these time-varying controls to our baseline regression shown equation (1) (that is, along with the cubic time trends of the initial conditions). Estimation results are reported in Table 4, column 5. Our results remain robust to these additional controls, lending further support to our identification.

Matched DD estimation. As a further robustness check, we combine our DD estimation with a PSM strategy. Specifically, we match for each TCZ city, to a non-TCZ city based on seven key criteria that the central government used to determine TCZ cities. These variables include roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment.[27] The estimation results are reported in Table 4, column 6. We continue to find negative and statistically significant estimates for the SO_2 emissions and the GDP growth rate from this combined DD and PSM estimation framework, implying the robustness of our findings.

4.3. Non-targeted pollutants

Our results show that the new evaluation system in 2005 introduced a substantial trade-off between SO_2 emissions and GDP growth. One possible explanation for the observed policy effect is that the new system made bureaucrats more concerned about the reduction in SO_2 emissions (the policy-targeted pollutant) and shifted their efforts away from economic development. It would then be interesting to investigate when the efforts were redistributed into environmental protection in general, or were they placed on the policy-targeted pollutant only. To this end, we examine the top pollutants outlined by the *China Pollution Source Census 2007.* Specifically, according to the Census, the four major air pollutants in China are SO_2, soot, NO_x and dust. As no reliable national data are available for the latter two, we include the analysis of the 2005 policy effect on soot. We further estimate the policy effect on other two other non-target pollutants: CO_2 and waster water. If the policy works by incentivizing bureaucrats to target at SO_2 emissions only, then should we not expect to see any policy effect on these non-targeted pollutants.

Specifically, data for waste water discharge were collected from the annual statistical yearbooks of the Chinese provinces, supplemented by the *Chinese City Statistical Yearbooks 2004-2014* and *Chinese Environmental Yearbooks 2002-2014*. Data on soot were

[27]See footnote 5 for the definition of dirty sectors.

collected from the *Chinese City Statistical Yearbooks.* Data on CO_2 emissions were collected from the Global Change Simulation Data Center at National Science & Technology Infrastructure of China. CO_2 data were available only for years 2003-2008, which we used for our analysis. We match the CO_2 data with the city boundaries map and extract the corresponding CO_2 concentration.

Estimation results are reported in Table 5. None of the three estimated coefficients are statistically significant. We further test whether the regression coefficients reported in Table 5 are equal to those reported in column 3, Table 3. Test statistics are reported in Panel B, Table 5. We cannot reject the null hypothesis that the policy impacts on Log(Soot) are equal to those on Log(SO_2). However, our results do suggest that the policy impacts on Log(CO_2) and Log(wastewater) are different from those on Log(SO_2), as shown in columns 2 and 3 in Panel B, Table 5. Note that the polluting sources for SO_2 are very different from CO_2 and wastewater but similar to soot, particularly for industrial SO_2 and soot emissions. Thus, our findings suggest that local bureaucrats, who are motivated by their career concerns, care only about policy-targeted pollution reductions instead of the overall environmental improvement.

[Insert Table 5 here]

5. Mechanisms

Our analysis shows that the 2005 policy (which included the SO_2 emissions reduction in the bureaucrat promotion consideration) caused a significant fall in the SO_2 emissions and in the GDP growth rate. In this section, we explore mechanisms through which this policy was at work to shed light on the substitution of efforts between pollution reduction and economic development.

5.1. Policy measures

The 2005 policy included a series of measures to meet the emissions reduction targets in the TCZ cities, such as the closure of high SO_2 concentration coal mines, the increasing control of SO_2 emissions from thermal power plants and industrial boilers, and the introduction of more energy-saving and low pollution technologies (Cao et al., 2009). Local government officials in China have administrative powers over state-owned industrial sectors. Per the second *China National Economic Census* of 2008, over 51.1% of power plants are owned by the state. These power plants comprise more than 77.8% of workers in the entire power plant sector and 81.4% of the sector's total capital. For state-owned power plants, city mayors and secretaries determine whether or not to operate these plants as well as plant locations and operating capacity. In the remainder of this section, we will analyse five policy measures that are related to the operations of power plants. If it is shown that these measures were more effectively enforced by local bureaucrats in TCZ cities, it will suggest that the new administrative policy induces local bureacrats to exert more effort on SO_2 emissions reductions.[28]

Installation of desulfurization technologies in thermal power plants. One measure is to install desulfurization facilities in the thermal power plants to contain their emissions, which are one of the major sources of SO_2 emissions. To this end, we collect desulfurization data for thermal power plants from the *Compilation of Statistical Materials of Electric Power Industry* 2001-2010. This dataset covers all thermal power plants with an installed capacity of 6,000 KW and above, and accounts for more than 99% of all fossil-fuel power electricity generation capacity in China.[29] There are a total of 2210 power plants in our sample, averaged at 7.7 plants per city.[30] Firm-level estimation results are reported in Table 6, column 1. We find a positive and statistically significant coefficient, indicating that after the new evaluation system in 2005, thermal power plants increased their desulfurization rates (measured as the rate of desulfurization facilities over total facilities). These results suggest that this measure for reducing SO_2 emissions was implemented by local bureaucrats.

[Insert Table 6 here]

Exit and entry of thermal power plants. To reduce SO_2 emissions, the 2005 policy also required TCZ cities to shut down small thermal power plants and contain the new establishment of thermal power plants. To check the enforcement of these measures, we construct the entry and exit of thermal power plants at the city level using the same plant-level data mentioned above. Estimation results are reported in Table 6, columns 2 and 3. Our estimates show that the 2005 policy caused more exits and fewer entries of thermal power plants, confirming that the measures were in effect. Given that thermal power plants were the major source of electricity generation in China, these data may imply a slowdown in the economic development.

Controlling the growth of dirty sectors.[31] Another important measure is to contain

[28]The plant-level data also help address the issue of gaming brought up in previous literature. Under the target-based system, firms or individuals would be incentivized to manipulate data and reporting to achieve the pre-set goal (Bevan and Hood, 2006; Chen et al., 2012; Gao, 2015). Plant-level data serve as a crosscheck for city-level SO_2 emission levels and provide further support for our main results.

[29]The plant level data can allow us to investigate the heterogeneity of responses across firms. In particular, one potential future work is to collect data on connections between local officials and firms, and examine how the burdens of regulation differ across connected and disconnected firms. We thank a reviewer for this suggestion.

[30]The authors identify locations for all thermal power plants mentioned in the *Compilation* by manually searching for their physical addresses. We then match the plant-level desulfurization progress data released by the Ministry of Environment Protection to our sample. Plant entry and exit data were obtained from the list released annually by the National Development and Reform Commission.

[31]See footnote 5 for the definition of dirty sectors.

the fast growth of dirty sectors (i.e., SO_2 intensive sectors listed in Appendix Table A.1) in the TCZ cities. To check the implementation of this measure, we collect data from the *Annual Surveys of Manufacturing Firms* from 2000 to 2008, which are the most comprehensive firm data in China (including all state-owned enterprises, and non-state-owned enterprise with annual sales above about US$ 74,000). Specifically, we compare the firm entry and exit pattern in the dirty sectors between TCZ and non-TCZ cities during this sample period. Estimation results are reported in Table 6, columns 4 and 5. We find a negative but not statistically significant estimated coefficient of firm entry, and a positive and statistically significant estimated coefficient of firm exit. These results imply that after the 2005 policy, there were more firm exits in the dirty sectors in the TCZ cities compared with the non-TCZ ones, though the two groups of cities had very similar firm entry patterns in the dirty sectors.

Improving the SO_2 removal rate. One policy measure requested the improvement of the SO_2 removal rate, especially in the dirty sectors. SO_2 removal rates are the percentage of total amount of SO_2 emissions reduced by various desulfurization technologies. To examine this measure, we calculate the city-level SO_2 removal rate as the ratio of the total amount of city SO_2 removals over the sum of city SO_2 removed and city SO_2 emissions. Data were collected from the *China City Statistical Yearbooks 2004-2016*. Estimation results are reported in Table 6, column 6. We find a positive and statistically significant estimated coefficient, suggesting the improving SO_2 removal rate in the TCZ cities after the 2005 policy.

Restructuring of the energy composition. To reduce SO_2 emissions, the 2005 policy also suggested an improvement in the energy structure, that is, a shift toward clean energies (such as hydropower, wind power, and nuclear power). To examine the policy effect on the energy structure, we construct a city-level share of clean energies, that is, the share of electricity-generating capacity from hydropower, wind power, and nuclear power over the total capacity. Data for the share of clean energy were collected from the *Compilation of Statistical Materials of Electric Power Industry* 2001-2010. Estimation results are reported in Table 6, column 7. The estimated coefficient is small and not statistically significant, suggesting no shift toward clean energies after the 2005 policy.

In sum, our analyses show that most of the measures for reducing SO_2 emissions in the 2005 policy were enforced by local bureaucrats, which explains the significant drop in SO_2 emissions in TCZ cities after 2005 compared with non-TCZ cities. Meanwhile, some of these measures (such as the closure and establishment restriction of thermal power plants, controlling of growth in dirty sectors, etc.) had negative implications for economic development.

5.2. Revealed preference by local bureaucrats

As the second mechanism analysis, we investigate how local bureaucrats revealed their efforts for pollution reduction and economic growth. To this end, we conduct a thorough content analysis of each city government's *Annual Work of the Government Report* from 2001 to 2015. The report is usually presented by the city mayor at the city-level National People's Congress early in the year. The report sets the goals and objectives for city governments' work for the upcoming year and is meant to be supervised by local residents. Each report outlines the target GDP growth rate for the year, as well as other policy objectives for the city government. As the reports are not available in any existing database or statistical yearbooks, we collected information by manually searching 4288 reports via the *China Statistical Yearbooks* at city level, city governments' websites and local newspapers.

For pollution reduction, we search for the keywords "environmental protection" (*huan bao* or *huanjing baohu* in Chinese) and "emissions reduction" (*jian pai* in Chinese), and calculate the ratio of the number of these mentions to be the total number of words in the report. We then examine whether the government reports in TCZ cities had more coverage of environmental protection than in non-TCZ cities after 2005. Regression results are reported in Table 7, column 1. We find that $TCZ_c * Post05_t$ is positive and statistically significant, suggesting that TCZ city governments became more concerned about environmental protection after 2005 than non-TCZ city governments.

[Insert Table 7 here]

In the report, the governments also report their target GDP growth rate in the coming year, which allows us to capture each city government's emphasis on economic growth. To this end, we estimate our baseline specification using the target GDP growth rate as the outcome. Estimation results are reported in Table 7, column 2. We find that after 2005, TCZ cities had lower target GDP growth rates than non-TCZ ones.

Combined, these results suggest that local bureaucrats in TCZ cities were willing to trade off GDP growth to achieve the more stringent emissions reductions goal.

5.3. Signaling effect

While we focus on the career concerns explanation of our findings, another potential channel for the policy to operate is to send stronger information signals to local bureaucrats about the preferences of the central government and the Communist Party. Specifically, the veto power of emissions quota in the 2005 policy could strongly suggest the determination of the central government to combat the increasing SO_2 emissions. Thus, our findings of SO_2 reductions after the 2005 policy could be an improvement in the information received by local bureaucrats, instead of the motivation of career concerns.

To examine whether our findings reflect the signaling argument, we show that there is a clustering of cities at the target SO_2 reduction level. To this end, we use our manually collected city SO_2 emissions quota by the end of the 11th Five-Year Plan (i.e., 2010), which allows us to construct the difference between the realized and target SO_2 reductions for 158 out of the 286 cities in our sample.[32] We plot the kernel density distributions of the differences for TCZ and non-TCZ cities in Fig. 4. We find a clear clustering at the point 0, suggesting that local bureaucrats managed to just complete the pollution reduction requirement. Meanwhile, for TCZ cities, there is a much sharper concentration of SO_2 reductions just at the target level. These results further suggest the relevance of the career concerns argument for our findings.

[Insert Figure 4 here]

6. Welfare analyses

In this section, we present a welfare analysis of the 2005 policy. Specifically, we first focus on the trade-offs between SO_2 emissions reductions and GDP growth, and then compute the health damage and value of statistical lives saved due to the policy.

6.1. Trade-offs between SO_2 emissions reductions and GDP growth

As shown in Table 3, columns 3 and 4, the 2005 policy led to an average reduction in SO_2 emissions by 5564 tonnes at the city level, and a slowdown in the GDP growth rate by 0.846 percentage point. The average city-level GDP in our sample period between 2001 and 2015 is 126.9×10^9 in Chinese Yuan (CNY).[33] Due to the spatial difference in prices, we follow Brandt and Holz (2006) and adjust all GDP values using a provincial price deflator with the price of Beijing in 1999 as the baseline. Therefore, a 0.846 percentage point reduction in GDP growth means a loss of 1.07 billion CNY in output ($126.9 \times 10^9 \times 0.846\% = 1.07 \times 10^9$). In other words, one tonne of reduction in industrial SO_2 emissions would cost 192,950 CNY ($1.07 \times 10^9/5564 = 192,950$). Our estimate shows that local bureaucrats made a decision to forgo a large GDP in exchange for a fast and sizable reduction in SO_2 emissions.

Welfare benefits of emissions reductions, such as mortality and health consequences, are usually computed based on pollutant concentration ($\mu g/m^3$) in the literature. To be consistent with existing literature, we also compute the marginal cost per $\mu g/m^3$ reduction in SO_2 concentration. Specifically, we first convert SO_2 emissions reductions

[32] Note that the sample size we use in this analysis is smaller than that in Table 2. This is because in constructing Table 1, we need to know only whether a city has achieved its targeted reduction or not. For the analysis presented here, we need to know the target and actual reductions in numbers.

[33] Authors' own calculation using data from all 286 cities in our sample. We take the average city-level GDP output across all cities between 2001 and 2015.

from weights to reductions in the concentration in the air. Appendix A outlines the detailed method we use to do the conversion. Using the method, a reduction of 5564 tonnes in SO_2 emissions weight is equivalent to a reduction of 4.1 $\mu g/m^3$ in SO_2 air concentration at city level. Therefore, based on our estimates in Table 3, every 1 $\mu g/m^3$ reduction in SO_2 concentration would cost 261 million CNY ($1.07 \times 10^9/(4.1\mu g/m^3) = 261 \times 10^6$).

In the next section, we use the converted reduction in SO_2 air concentration to compute value of statistical lives saved by the policy.

6.2. Health damage and value of statistical lives

As an alternative method for welfare analysis, we deviate from the SO_2-GDP trade-off, and calculate health damage of pollution to evaluate our estimated effect of the policy.

According to World Bank (1997), there is no direct mortality risk associated with SO_2 exposure. The report also suggests that the disease incidences associated with per million per $\mu g/m^3$ increase in SO_2 concentration for chest discomfort and respiratory symptoms (child) are 10,000 cases and 5 cases per million people, respectively. Using the valuation given in Ho and Jorgenson (2003), the cost per case is 9.8 in 1997 CNY.[34] According to the government documentation, TCZ cities cover about 39% of population in China.[35] Assume that China has a population of 1.4 billion, and the 2005 policy reduced the SO_2 air concentration by 4.1 $\mu g/m^3$, the estimated policy effect on reduction in disease incidences for TCZ cities therefore has a monetary value of 219.5 million (1997 CNY) ($4.1\times$ 39% $\times$ $1400\times$ (10,000$\times$ 9.8 +5$\times$ 9.8) = 219.5 million). Note that the cost of SO_2 reduction (in terms of GDP) presented in the previous section is at the city level, while the social benefit of disease incidences reduction is at the national level. To make the numbers comparable, we need to multiply the cost of SO_2 reduction at city level by the number of TCZ cities in our sample (i.e. 1.07×162 =173 billion).

While there is no direct mortality risk of SO_2 as suggested by the World Bank report, there would be secondary mortality risks if SO_2 is converted to PM2.5. Our analysis indicates that the 2005 policy reduced PM2.5 concentration by 1.5% (Table 4, Column 4). Based on Chen et al. (2017), the national average mortality risk associated with every 10 $\mu g/m^3$ reduction in PM2.5 is about 17.55 thousand lives. The average PM2.5 level in our sample city is 31.66 $\mu g/m^3$.[36] Therefore, a 1.5% reduction in PM2.5 concentration would lead to 833 ($31.66 \times 1.5\% \times 17550/10 = 833$) lives saved. Based on World Bank's valuation, the estimated value of a statistical life is about 585,000 in 1997 CNY.[37] Thus,

[34] In later updates made by Ho and Jorgenson (2007), the authors adjust the valuation per case to be 6.2 (low scenario) and 10 (high scenario), both in 1997 CNY.

[35] The relevant Chinese document for the TCZ policy is available at the following government website: `http://www.mep.gov.cn/gkml/zj/wj/200910/t20091022_172231.htm`

[36] See Table 2.

[37] The estimated value of statistical life in the literature ranges from 230,000 - 1.2 million in 1997 CNY.

the total value of statistical lives saved due to the 2005 policy is about 487.3 million in 1997 CNY ($833 \times 585{,}000 = 487.3 \times 10^6$).

In summary, our welfare analysis suggests that the cost of the program far exceeds its benefit. As the program is able to achieve the targeted emissions cut in a very short period of time, our calculation suggests that the Chinese government is willing to pay a premium for immediate and sizable pollution reductions.

7. Conclusion

In this paper, we examine the effect of a target-based performance evaluation system on SO_2 emissions reductions. We find that when SO_2 emissions quotas were built into the evaluation system for local bureaucrats, SO_2 emissions were significantly reduced. In the meantime, real GDP growth rates also decreased significantly. Our analysis suggests that the new evaluation system induced local bureaucrats to exert more effort on environmental issues related to SO_2 emissions. We find no spillover effects of the policy on other non-targeted pollutants such as CO_2 or wastewater. Using a comprehensive dataset that we constructed, we observe that after the new administrative system was in place, local bureaucrats are more likely to implement policies that specifically targeted SO_2 emissions reduction, such as shutting down power plants and installing desulfurization technologies. Implementation of these policies indicates that local bureaucrats are exerting more effort on achieving the target set for decreasing emissions by the central government.

Our finding adds to the empirical debate on the effectiveness of target-based performance evaluation systems for public sector workers, particularly for management-level bureaucrats. As the earlier literature has highlighted, the addition of a concrete performance target into an evaluation system might affect the overall comprehensive output as bureaucrats will divert their attention to the output that is more heavily weighted by their supervisors. Our analysis shows a clear pattern that the local bureaucrats made a conscious trade-off between growth and SO_2 emissions when emissions targets are brought into the picture.

Our results indicate that setting a concrete target-based performance evaluation system for environmental policies can be effective in ensuring that bureaucrats work to curb pollution and promote environmental protection. The effect is immediate, and the impact is significant. Developing countries facing similar pressing environmental issues might want to consider the implementation of such a policy.

Ho and Jorgenson (2007) Table 9.9 provides a summary of various valuations and their sources.

Appendix A

We use a method proposed by Ho and Jorgenson (2003) to convert reductions in emissions in weight to reductions in the SO_2 ambient air concentration.

Specifically, total SO_2 emissions at height c can be expressed as

$$E_c = \Sigma_{j \in c} E_j \tag{3}$$

where c = low, medium and high. High-emissions industries include all types of electricity power sectors. Medium-emissions industries include coal mining, natural gas and metal ming sectors etc. Low-emissions industries include agriculture and all tertiary sectors. [38]

To convert the total emissions E_c into the air concentration, we follow the linear reduced form suggested by Ho and Jorgenson (2002) and Garbaccio et al. (2000) [39] :

$$C = \gamma_{low} E_{low} + \gamma_{medium} E_{medium} + \gamma_{high} E_{high}$$

where $\gamma_{low}, \gamma_{medium}$ and γ_{high} are conversion coefficient computed by the authors based on Lvovsky and Hughes (1997). For urban SO_2 emissions, $\gamma_{low} = 0.03364$, $\gamma_{medium} = 0.00607$ and $\gamma_{high} = 0.00096$. These coefficients are computed for 11 Chinese cities using a emission-concentration dispersion model. Ho and Jorgenson (2003) reweigh these coefficients to arrive at the conversion coefficient γ for all urban China. We use a similar strategy in Ho and Jorgenson (2003) and reweigh their coefficients proportionately so that it matches the average SO_2 concentration level in our sample period. We then compute the reduction of the SO_2 concentration as a result of the 2005 policy change. Our method can be outlined in the following steps:

Step 1: Recalibrate the conversion coefficients γ_{low}, γ_{medium} and γ_{high}. The original conversion coefficients computed by Lvovsky and Hughes (1997) are $\gamma_{low} = 0.03364$, $\gamma_{medium} = 0.00607$ and $\gamma_{high} = 0.00096$, but the coefficients capture only the 11 Chinese cities in their sample. In order to compute a national conversion coefficient, we need to calibrate these numbers so that they match the national average SO_2 concentration in our sample period. There are no SO_2 air concentration values for cities in our dataset; therefore we resort to other resources. The Ambient Air Quality Standard in China classifies cities with average SO_2 concentration below 60 $\mu g/m^3$ as healthy and below 10 $\mu g/m^3$ as good. According to the annual report published by the Ministry of Environmental Protection in China, in our sample period more than 80% of cities has an average SO_2

[38] Please see Ho and Jorgenson (2003) for a complete list of industry emissions height classifications.

[39] In Ho and Jorgenson (2007), the authors discuss the difference between the emissions height method and a method based on intake fractions. We are not able to implement the intake fractions method as we do not know the exact breakdown of the emissions industries (we know only the broad categories as indicated in Table 1). Thus, for our purpose, the method presented in this section is the most feasible.

concentrations between 10 $\mu g/m^3$ and 60 $\mu g/m^3$.[40] When the initial TCZ policy was designed in 1998, all cities with annual SO_2 concentrations higher than 60 $\mu g/m^3$ were classified as TCZ cities. Therefore, we would calibrate the coefficients against 60 $\mu g/m^3$ for our analysis. The average city- level annual industrial emissions in our control sample is 38110 tonnes. According to the *China Pollution Source Census 2007*, industrial emissions usually consist of 91.4% of the total emissions, while the remaining 8.6% is non-industrial emissions. Therefore, we assume the average total emissions are $38110/0.914 = 41696$ tonnes (38110 industrial emissions and 3586 non-industrial emissions). Following Ho and Jorgenson (2003), we assume the coefficients are being re-scaled proportionately. Let φ be the scaling factor. Then the following conditions have to be satisfied:

$$
\begin{aligned}
&60\mu g/m^3 = 0.03364\varphi E_{low} + 0.00607\varphi E_{medium} + 0.00096\varphi E_{high} \\
&E_{low} + E_{medium} + E_{high} = 41696
\end{aligned} \tag{4}
$$

To determine emissions at different heights, i.e., E_{low}, E_{medium} and E_{high}, we first match the emitting industries in Table 1 to the emission height class data in Ho and Jorgenson (2003). Among the industrial sectors mentioned in Table 1, the production and supply of electricity power and heat is considered high while everything else is in the medium category. Note that Table 1 covers only 88.5% of all industrial SO_2 emissions. For simplicity, we assume the missing 12.5% industrial emissions are proportionately distributed between the high and medium categories. In other words, out of the total 38110 tonnes industrial emissions, 56.95% (50.4%/88.5%) would be in the high emissions category, and the rest (43.05%) would be in the medium category. Therefore, $E_{high} = 38110 \times 56.95\% = 21704$ tonne and $E_{medium} = 38110 \times 43.05\% = 16406$ tonnes. Non-industrial emissions come from either the agriculture sector or households; both are classified as the low emissions category. Therefore, $E_{low} = 3586$ tonnes. Substitute the emissions number E_{low},E_{medium} and E_{high} back in equation 4, and we get $\varphi = 0.25$. After calibration, the conversion coefficients, therefore, are $\gamma_{low} = 0.00841, \gamma_{medium} = 0.00152$ and $\gamma_{high} = 0.00024$.

Step 2: In this step, we compute the reduction in the SO_2 concentration. We have computed that the policy reduces total industrial emissions by 5564 tonnes. We assume the reductions are distributed proportionately across sectors in Table 1 and the policy has no impact on non-industrial sectors. Thus, 56.95% of the emissions reductions happens in the high emissions category while 43.05% happens in the low emissions category. This means after reduction $E_{high} = 21704 - 5,564 \times 56.95\% = 18535$ tonnes and $E_{medium} = 16406 - 5564 \times 43.05\% = 14011$ tonnes. Therefore, after reduction, the SO_2 concentration

[40] For relevant government documents, please see the Ministry of Environmental Protection's website at http://www.mep.gov.cn/

becomes $0.00841 \times 3586 + 0.00152 \times 14011 + 0.00024 \times 18535 = 55.9 \mu g/m^3$

Compared to the baseline SO_2 concentration of 60 $\mu g/m^3$, this means a reduction in concentration of 4.10 $\mu g/m^3$, or 6.8%.

References

A. Alesina and G. Tabellini. Bureaucrats or politicians? Part i: A single policy task. *American Economic Review*, 97:169–179, 2007.

A. Alesina and G. Tabellini. Bureaucrats or politicians? Part ii: Multiple policy tasks. *Journal of Public Economics*, 92:426–447, 2008.

R. Benabou and J. Tirole. Intrinsic and extrinsic motivation. *Review of Economic Studies*, 70:489–520, 2003.

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

G. Bevan and C. Hood. What's measured is what matters: Targets and gaming in the English public health care system. *Public Administration*, 84(3):517–538, 2006.

L. Brandt and C. A. Holz. Spatial price differences in China: Estimates and implications. *Economic Development and Cultural Change*, 55(1):43–86, 2006.

X. Cai, Y. Lu, M. Q. Wu, and L. H. Yu. Does environmental regulation drive away inbound foreign direct investment? Evidence from a quasi-natural experiment in China. *Journal of Development Economics*, 2016.

J. Cao, R. Garbaccio, and M. S. Ho. China's 11th five-year plan and the environment: Reducing SO_2 emissions. *Review of Environmental Economics and Policy*, 3(2):231–250, summer 2009.

R. Chen, P. Yin, X. Meng, C. Liu, L. Wang, X. Xu, J. A. Ross, L. A. Tse, Z. Zhao, H. Kan, and M. Zhou. Fine particulate air pollution and daily mortality: A nationwide analysis in 272 Chinese cities. *American Journal of Respiratory and Critical Care Medicine*, 196(1):73–81, 2017.

Y. Chen, H. B. Li, and L. A. Zhou. Relative performance evaluation and the turnover of provincial leaders in China. *Economics Letters*, 88:421–425, 2005.

Y. Y. Chen, G. Z. Jin, N. Kumar, and G. Shi. Gaming in air pollution data? Lessons from China. *The B.E. Journal of Economic Analysis & Policy*, 12(3):1–43, 2012.

R. Chetty, A. Looney, and K. Kroft. Salience and taxation: Theory and evidence. *American Economic Review*, 99(4):1145–1177, 2009.

M. A. Cohen and A. Tubb. The impact of environmental regulation on firm and country competitiveness: A meta-analysis of the porter hypothesis. *Journal of the Association of Environmental and Resource Economics*, 2017. Accepted.

M. Dewatripont, I. Jewitt, and J. Tirole. The economics of career concerns, Part ii: Application to missions and accountability of government agencies. *Review of Economic Studies*, 66:199–217, 1999.

M. Dewatripont, I. Jewiit, and J. Tirole. Multitask agency problems: Focus and task clustering. *European Economic Review*, 44:869–877, 2000.

A. Dixit. Incentives and organizations in the public sector: An interpretative review. *Journal of Human Resources*, 37(4):696–727, 2002.

EuroStat. Government expenditure on environmental protection. Online, Mar. 2017. URL `http://ec.europa.eu/eurostat/statistics-explained/index.php/Government_expenditure_on_environmental_protection`.

W. Feng, Y. Cai, and B. Gu. Population, policy, and politics: How will history judge China's one-child policy? *Population and Development Review*, 38(S1):115–129, 2013.

J. Gao. Pernicious manipulation of performance measures in China's cadre evaluation system. *China Quarterly*, 223:618–637, 2015.

R. F. Garbaccio, M. S. Ho, and D. W. Jorgenson, editors. *The Health Benefits of ContControl Carbon Emissions in China*, Ancillary Benefits and Costs of Greenhouse Gas Mitigation, Proceedings of an IPCC Co-Sponsored Workshop, Washington, D.C., March 2000. Organization for Economic Co-operation and Development.

M. Gentzkow. Television and voter turnout. *Quarterly Journal of Economics*, 121:931–972, 2006.

M. Greenstone. The impacts of environmental regulations on industrial activity: Evidence from the 1970 and 1977 clean air act amendments and the census of manufacturers. *Journal of Political Economy*, 110(6):1175–1219, 2012.

J. He. Estimating the economic cost of China's new desulfur policy during her gradual accession to WTO: The case of industrial SO_2 emission. *China Economic Review*, 16: 364–402, 2005.

L. Hering and S. Poncet. Environmental policy and exports: Evidence from Chinese cities. *Journal of Environmental Economics and Management*, 68:296–318, 2014.

M. S. Ho and D. W. Jorgenson. *Pollution Taxes and Public Health*, pages 165–200. In Economics of the Environment in China. Boyds: Aileen International, Maryland, 2002. China Council for International Cooperation on Environment and Development Environmental Economics Working Group.

M. S. Ho and D. W. Jorgenson. Air pollution in China: Sector allocation of emissions and health damage. Report, China Council for International Cooperation on Environment and Development, 2003.

M. S. Ho and D. W. Jorgenson. Sector allocation of emissions and damage. In M. S. Ho and C. P. Nielsen, editors, *Clearing the air : the health and economic damages of air pollution in China*, chapter 9, pages 279–330. The MIT Press, Cambridge, Massachusetts, 2007.

W. M. Hoden and W. R. Barnard. Evaluating the contribution of PM2.5 precursor gases and re-entrained road emissions to mobile source PM2.5 particulate matter emissions. Technical report, U.S. Environmental Protection Agency, 2015.

B. Holmstrom and P. Milgrom. Multitask principal-agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, & Organization*, 7:24–52, 1991.

C. Hood. Public service management by numbers: why does it vary? where has it come from? what are the gaps and the puzzles? *Public Money and Management*, 27(2): 95–102, 2007.

Y. Jin and L. Lin. China provincial industrial pollution: The role of technical efficiency, pollution levy and pollution quantity control. *Environmental and Development Economics*, 19:111–132, 2014.

D. W. Jorgenson and P. J. Wilcoxen. Environmental regulation and u.s. economic growth. *The RAND Journal of Economics*, 21(2):314–340, 1990.

M. E. Kahn, P. Li, and D. X. Zhao. Water pollution progress at borders: The role of changes in China's political promotion incentives. *American Economic Journal: Economic Policy*, 7(4):223–42, 2015.

E. La Ferrara, E. A. Chong, and S. Duryea. Soap operas and fertility: Evidence from Brazil. *American Economic Journal: Applied Economics*, 4(4):1–31, 2012.

B. Lockwood and F. Porcelli. Incentive schemes for local government: Theory and evidence from comprehensive performance assessment in England. *American Economic Journal: Economic Policy*, 5(3):254–286, 2013.

K. Lvovsky and G. Hughes. An approach to projecting ambient concentrations of SO_2 and PM-10. Unpublished Annex 3.2 to World Bank, Clear water, blue skies: China's environment in the new century. Washington, D.C: World Bank, 1997.

E. Maskin and J. Tirole. The politician and the judge: Accountability in government. *American Economic Review*, 94(4):1034–1054, 2004.

MEPC. China earmarks 7.5 billion dollars for environment protection. Online, Nov. 2017. URL `http://english.mep.gov.cn/News_service/media_news/201711/t20171124_426867.shtml`.

OECD. Performance management in government: Performance measurement and results-oriented measurement. Technical report, OECD, Paris, 1994.

M. E. Porter and C. van der Linde. Toward a new conception of the environment-competitiveness relationship. *Journal of Economic Perspectives*, 9(4):970118, 1995.

I. Rasul and D. Rogger. The impact of ethnic diversity in bureaucracies: Evidence from the Nigerian civil service. *American Economic Review: Papers & Proceedings*, 105(5): 457–461, 2015.

I. Rasul and D. Rogger. Management of bureaucrats and public service delivery: Evidence from the Nigerian civil service. *The Economic Journal*, 2017. Forthcoming.

S. Tanaka. Environmental regulations on air pollution in China and their impact on infant mortality. *Journal of Health Economics*, 42:90–103, 2015.

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

Y. Wang and N. Shen. Environmental regulation and environmental productivity: The case of china. *Renewable and Sustainable Energy Reviews*, 62:758–766, 2016.

Z. G. Wen and J. N. Chen. A cost-benefit analysis for the economic growth in China. *Ecological Economics*, 65(2):356–366, 2008.

J. Q. Wilson. *Bureaucracy: What Government Agencies Do and Why They Do It.* New York: Basic Books, Inc., New York, 1989.

World Bank. Clear water, blue skies: China's environment in the new century. Technical report, Washington D.C:, 1997.

World Bank. Cost of pollution in China. Technical report, Washington, DC: World Bank, East Asia and Pacific Region, 2007.

A. Xepapadeas. *Economic growth and the environment*, volume 3, chapter 23, pages 1220–1266. Elsevier, 2005.

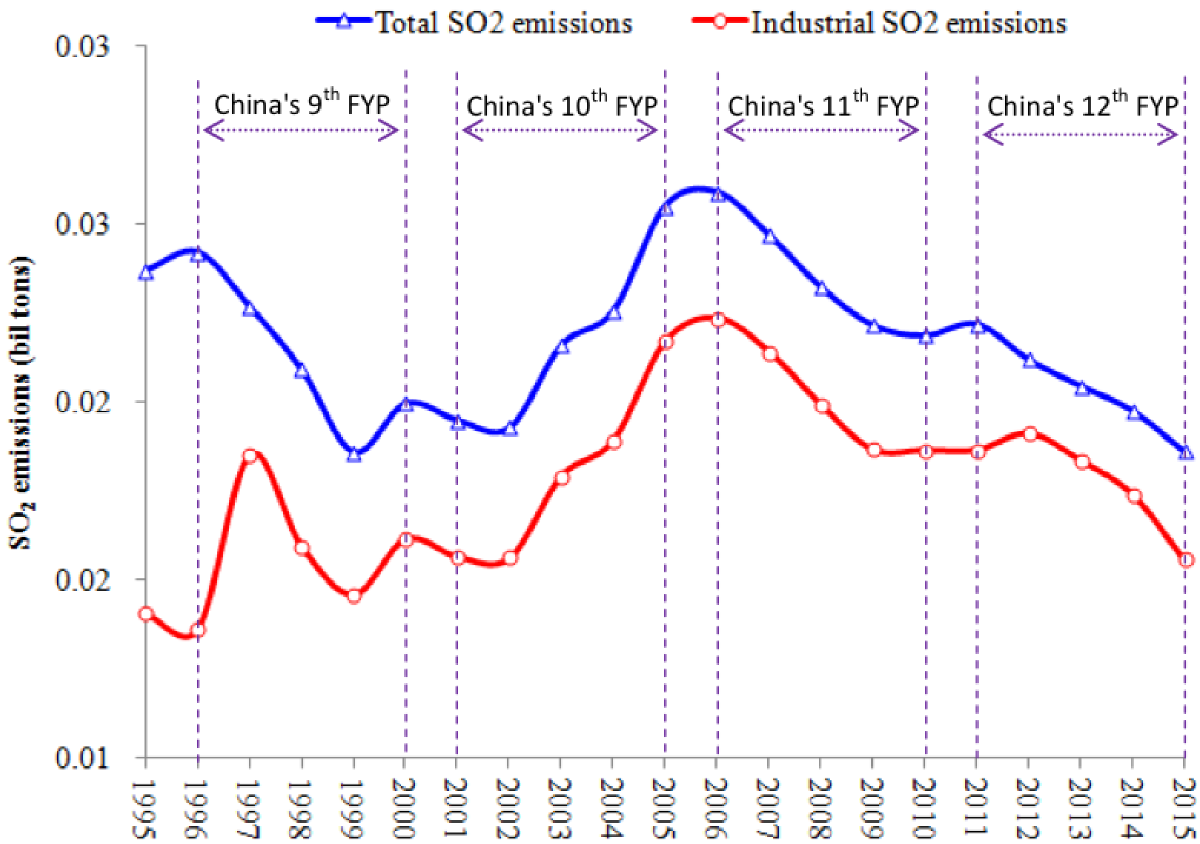
C. G. Xu. The fundamental institutions of China's reforms and development. *Journal of Economic Literature*, 49:1076–1151, 2011.

B. Yuan, S. Ren, and X. Chen. Can environmental regulation promote the coordinated development of economy and environment in china's manufacturing industry? *Journal of Cleaner Production*, 149:11–24, 2017.

X. Zhao, Y. Zhao, S. Zeng, and S. Zhang. Corporate behavior and competitiveness: Impact of environmental regulation on chinese firms. *Journal of Cleaner Production*, 86:311–322, 2015.

S. Q. Zheng, M. E. Kahn, W. Z. Sun, and D. L. Luo. Incentives for China's urban mayors to mitigate pollution externalities: The role of the central government and public environmentalism. *Regional Science and Urban Economics*, 47:61–71, 2014.

Figure 1: SO_2 Emission Trend 1995-2015



Data Source: China Statistical Yearbooks in various years

Figure 2: Two Control Zone

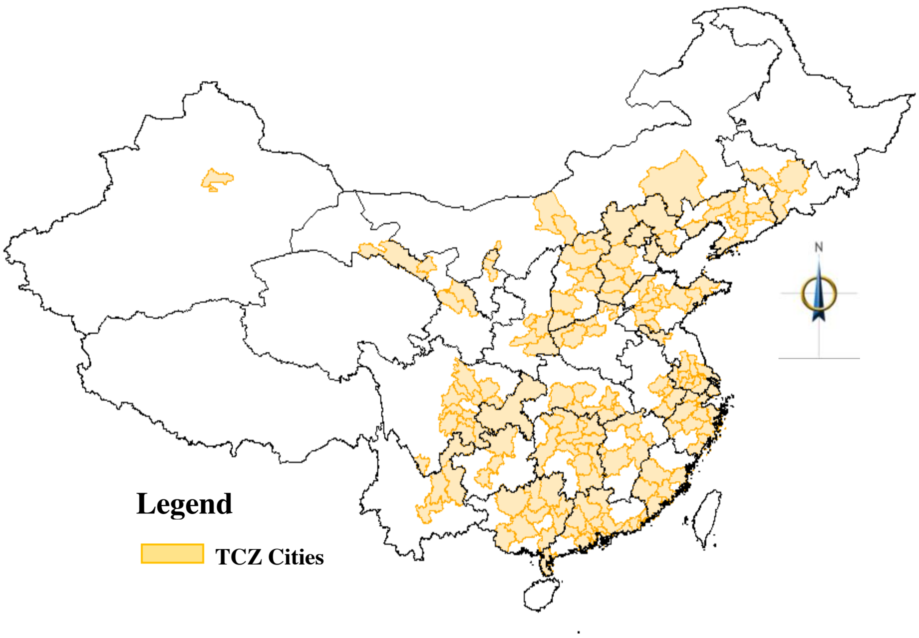
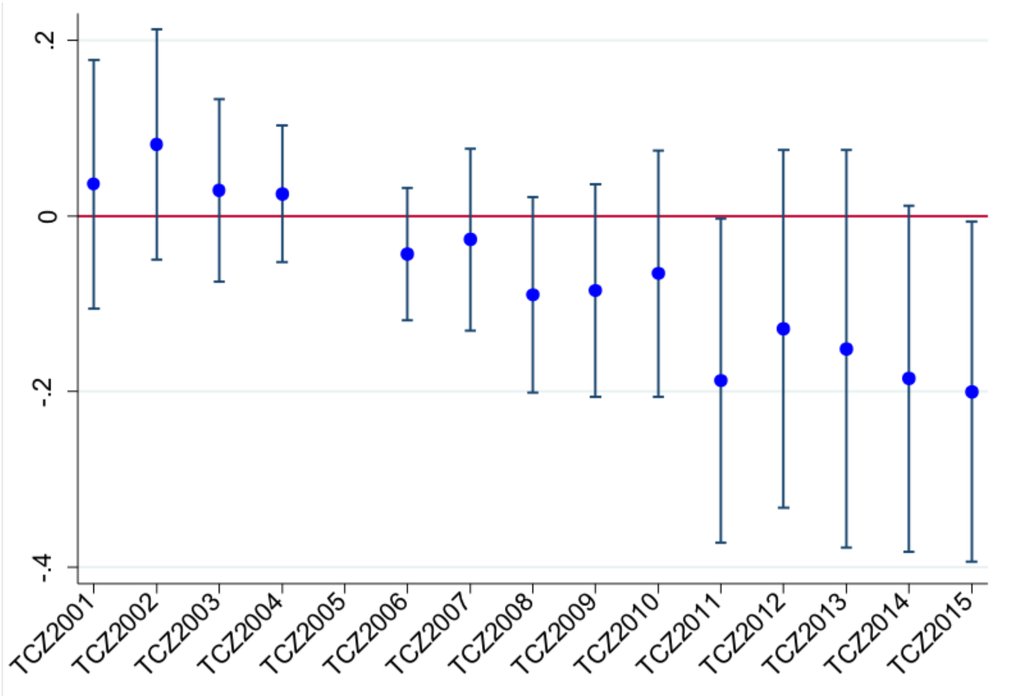


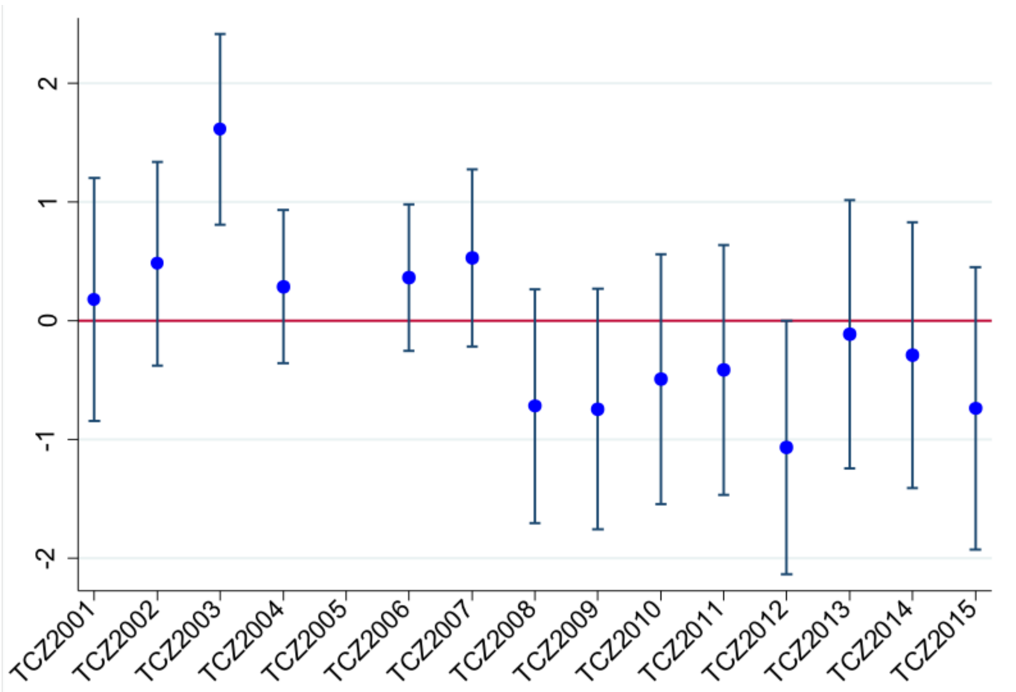
Figure 3: Estimated coefficients

(a) Log SO_2



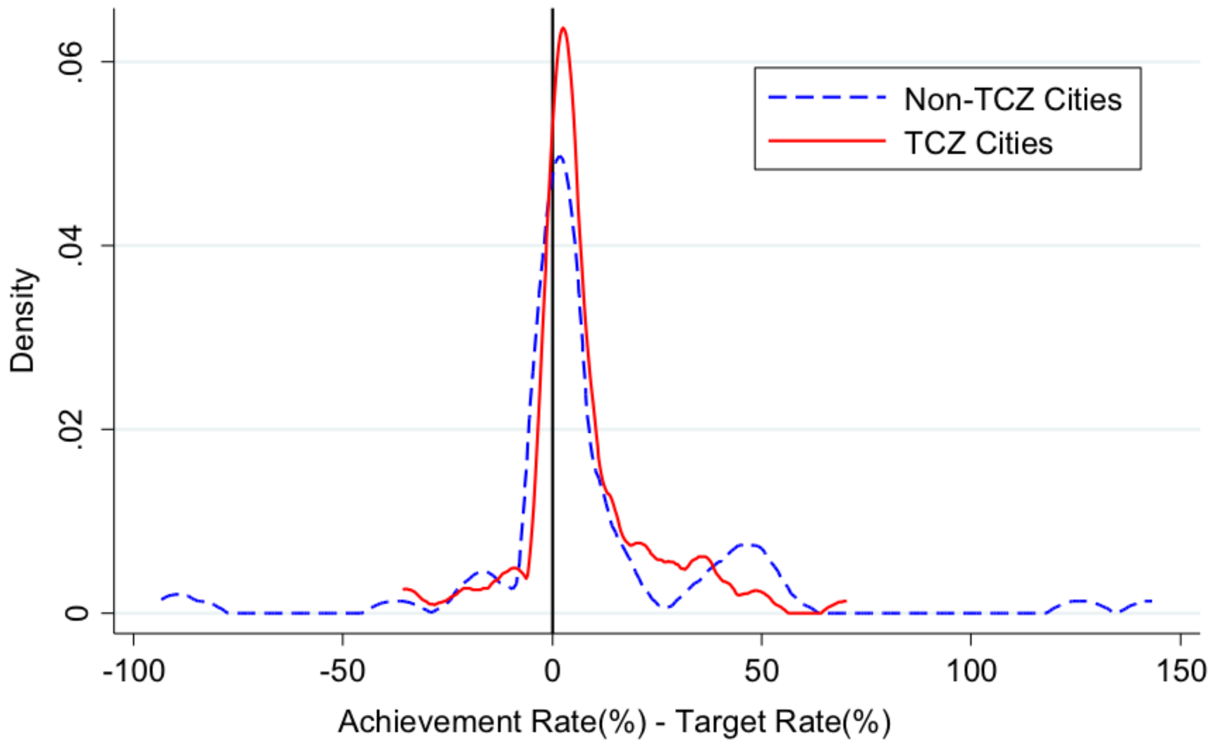
Note: The figure illustrates the time trends of industrial SO_2 emissions (in logarithm) between TCZ cities and that of non-TCZ cities.

(b) Real GDP Growth



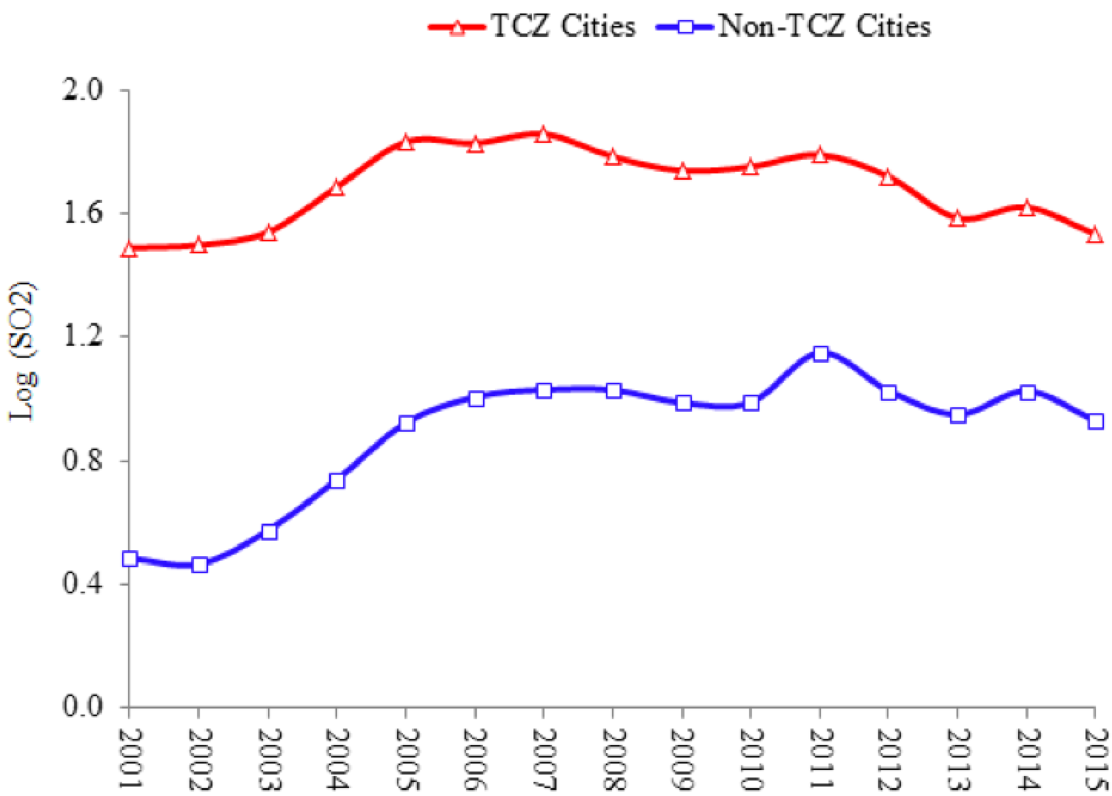
Note: The figure illustrates the time trends of real GDP growth rates between TCZ cities and that of non-TCZ cities.

Figure 4: Distribution of Achievement Rate for the 11th 5-Year Plan



Note: Data on achievement rate and target rate at city level were collected by the authors from various government documents and online sources. City level differences are calculated at the end of the 11th 5-Year Plan (e.g., 2010).

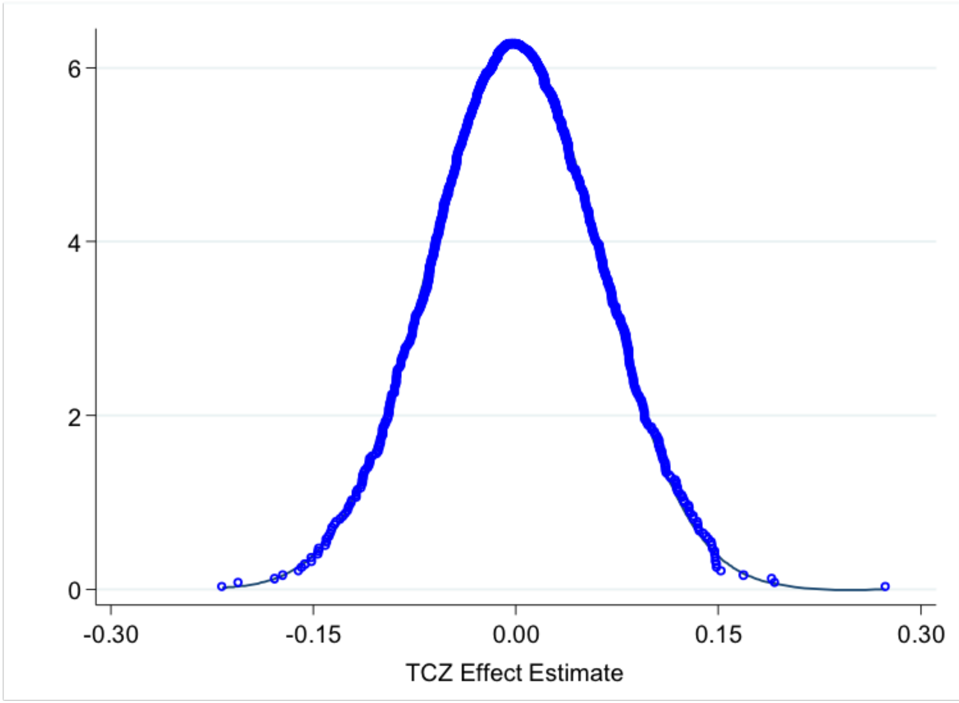
Figure A.1: Industrial SO_2 Emissions Trend 2001-2015



Note: The figure illustrates the time trends of total industrial SO_2 emissions (in logarithm) of TCZ and non-TCZ cities.

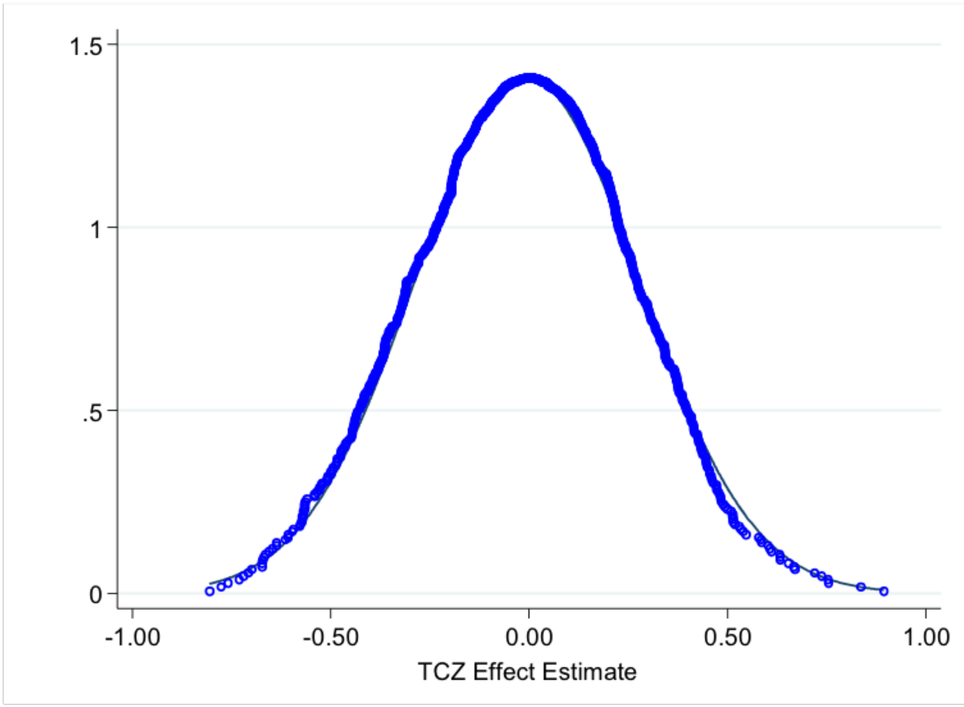
Figure A.2: Placebo Test

(a) Log SO_2



Note: The figure shows the probability distribution of the estimated coefficients from 1000 simulations randomly assigning the TCZ status to cities.

(b) Real GDP Growth



Note: The figure shows the probability distribution of the estimated coefficients from 1000 simulations randomly assigning the TCZ status to cities.

Table 1

SO_2 Emissions Reductions Goal and Achievement Rate

	(1)	(2)
	2006-2010 (11th Five-Year Plan)	2011-2015 (12th Five-Year Plan)
TCZ		
5-Year Reduction Goal	13.55%	12.19%
Achievement Rate	97.53%	98.14%
Non-TCZ		
5-Year Reduction Goal	4.04%	5.39%
Achievement Rate	83.87%	87.90%
All		
5-Year Reduction Goal	9.57%	9.22%
Achievement Rate	91.61%	93.68%

Note: Data compiled from the following sources by the authors: *China City Statistical Yearbooks*; City Level Environmental Protection 11th and 12th Five-Year Plan Document; City Level Annual Work of the Government Report; City Level 11th and 12th Five-Year Plan Report. Due to data limitation, out of the 286 prefectual cities in our sample, reduction goals were found for 206 and 213 cities, for the 11th and 12th Five-Year Plan respectively.

Table 2
Summary Statistics

Variables	Definition	Mean	S.D.	No. of Observations
Key Outcome Variables				
TCZ	=1 if a city is a TCZ city; =0 otherwise	0.57	0.50	4,288
Emissions /Pollutants				
SO_2	Total industrial SO_2 emissions (10^4 tons)	5.95	5.91	4,288
$PM_{2.5}$	Concentrations of $PM_{2.5}$ (micrograms/m^3)	31.66	14.11	2,860
Soot	Total industrial soot emissions (10^4 tons)	3.15	8.65	4,285
CO_2	Concentrations of CO_2 (100 ppm)	3.84	0.04	1,716
Wastewater	Total industrial waste water discharges (10^7 tons)	7.60	10.09	4,285
GDP Related Outcome Variables				
Target_GDP_Growth	Annual GDP growth target (%)	11.82	2.89	4,284
GDP_Growth	Real annual GDP growth (%)	12.17	3.76	4,288
Plant Level Varibles				
Desulfurization_Rate	A fossil-fuel power plant's desulfurization rate (%)	22.62	40.04	22,100
Firm_Size	A fossil-fuel power plant's electricity generation capacity (10^4 KW)	19.11	41.17	22,100
Firm_Age	A fossil-fuel power plant's age (year)	21.54	19.03	22,100
Control Variables for TCZ Selection				
Roughness	Standard deviation of land gradient	19.99	12.04	4,288
Elevation	Average elevation (km)	0.52	0.59	4,288
Wind_Speed	Annual average wind speed 1990-1995 (0.1m/s)	22.83	9.49	4,288
Coldness	Percent of days with temperatures of 5^o or below in one year 1990-1995 (%)	18.89	15.56	4,288
Precipitation	Annual average precipitation 1990-1995 (m)	9.86	4.98	4,288
Soil_PH	Average soil PH level of the topsoil (0-30 cm)	3.29	1.04	4,288
Dirty_Labor	No. of employees in dirty sector in 1996 (10,000 persons)	10.50	9.78	4,288
Other City Level Variables				
Government_EcoIndex	Count of "green" keywords in the government work report divided by total word count of the report in this year ($\times 10^3$)	0.42	0.29	4,288
FDI	Average inward FDI flows over 2001-2005 (million $)	0.02	0.06	4,288
Dirty_Export	Average ratio of export sales in dirty sectors to total export sales over 2001-2005	0.33	0.29	4,288
Mortality	Average mortality rate over 2001-2015 (‰)	5.39	1.05	4,288
Investment_PerCapita	Fixed Invetment per capita (10,000 Yuan)	1.82	2.10	4,285
GDP_PerCapita	GDP per capita (10,000 Yuan)	2.47	2.97	4,285
Manu_GDPShare	Secondary sector GDP share (%)	43.71	12.87	4,285
Coal_Abundance	Ratio of net coal imported from other provinces and countries to total coal supply in its province	-0.20	0.92	4,285
Trade_Intensity	Ratio of imports plus exports to GDP	0.22	0.48	4,285
SO2_Remove_Rate	Industrial SO_2 emission removal rate (%)	31.26	24.61	2,288
Clean_Energy_Share	Share of electricity generated by clean energy sources in total electricity generated (%)	31.68	34.25	2,860
Thermal Power Plant				
Entry	New fossil-fuel power plant electricity generation capacity (10^4 KW)	45.05	177.00	2,860
Exit	Closed fossil-fuel power plant electricity generation capacity (10^4 KW)	0.20	0.76	2,574
Dirty Sector				
Entry	No. of new firms in SO_2 intensive sectors	8.10	12.08	2,288
Exit	No. of exited firms in SO_2 intensive sectors	28.55	41.69	2,288

Note: "Green" keywords include "environmental protection" (*huan-jing-bao-hu* or *huan-bao*) and "pollutant emissions reduction" (*jian-pai*). See appendix table A1 for definitions of dirty sectors. Data sources are described in full in section 3.2 and section 5.

Table 3
Baseline Estimates

Dependent variable	Log (SO2)			GDP_Growth		
	(1)	(2)	(3)	(4)	(5)	(6)
TCZ × Post2005	-0.260***	-0.164**	-0.146**	-0.797**	-0.949***	-0.846***
	(0.065)	(0.066)	(0.068)	(0.317)	(0.328)	(0.328)
Dependent variable mean	1.338	1.338	1.338	12.170	12.170	12.170
City dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
TCZ Controls × f(t)		Yes	Yes		Yes	Yes
Other Controls × Post2005			Yes			Yes
No. of observations	4,283	4,283	4,283	4,288	4,288	4,288
Year coverage	2001-2015	2001-2015	2001-2015	2001-2015	2001-2015	2001-2015
Adjusted R-squared	0.837	0.846	0.847	0.499	0.543	0.546
No. of clusters	286	286	286	286	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level.

Table 4

Robustness Tests

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A		Log (SO_2)		Log($PM_{2.5}$)	Log (SO_2)	
TCZ × Post2005	-0.153**	-0.137*	-0.195**	-0.015**	-0.138**	-0.181*
	(0.068)	(0.070)	(0.096)	(0.007)	(0.066)	(0.094)
Dependent variable mean	1.338	1.306	1.467	3.334	1.338	1.424
No. of observations	4,283	4,178	2,963	2,860	4,283	3,306
Adjusted R-squared	0.847	0.841	0.855	0.966	0.848	0.848
Panel B			GDP_Growth			
TCZ × Post2005	-0.662**	-0.738**	-1.803***		-0.778**	-1.530***
	(0.316)	(0.336)	(0.488)		(0.322)	(0.380)
Dependent variable mean	12.170	12.176	12.081		12.170	12.112
No. of observations	4,288	4,183	2,968		4,285	3,309
Adjusted R-squared	0.553	0.544	0.548		0.556	0.518
City dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
TCZ Controls × f(t)	Yes	Yes	Yes	Yes	Yes	
Other Controls × Post2005	Yes	Yes	Yes	Yes	Yes	Yes
Region_Trend	Yes					
Log(GDP_PerCapita)					Yes	
Manu_GDPShare					Yes	
Coal_Abundance					Yes	
Trade_Intensity					Yes	
Largest&Olympic venue cities		Drop				
Neighoring cities			Drop			
PMS+DID						Yes
Year coverage	2001-2015	2001-2015	2001-2015	2001-2010	2001-2015	2001-2015
Adjusted R-squared	0.847	0.841	0.855	0.966	0.848	0.848
No. of clusters	286	279	198	286	286	220

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level. Variables used for matching in the PMS exercise in column 6 are the seven key criteria that the central government used to choose TCZ cities, which include roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment.

Table 5

Untargeted Pollutants

Dependent variable	Log (Soot)	Log(CO_2)	Log(Wastewater)
	(1)	(2)	(3)
Panel A			
TCZ × Post2005	-0.123	0.001	-0.023
	(0.086)	(0.001)	(0.059)
Panel B: Different from Log(SO_2)?			
TCZ × Post2005	-0.023	-0.147***	-0.123***
	(0.050)	(0.031)	(0.042)
Dependent variable mean	9.783	5.950	8.385
City dummy	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes
TCZ Controls × f(t)	Yes	Yes	Yes
Other Controls × Post2005	Yes	Yes	Yes
No. of observations	4,285	1,716	4,285
Year coverage	2001-2015	2003-2008	2001-2015
Adjusted R-squared	0.752	0.779	0.860
No. of clusters	286	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level. Panel B reports the differences between regression coefficients for TCZ × Post2005 on Log(SO_2) emissions (see column 3 of Table 3) and log emissions of untargeted pollutants presented in this table. In Panel A, the standard errors are reported in parentheses and clustered at city level. In Panel B, the standard errors are computed using the delta method.

Table 6

Policy Measures

Dependent variable	Desulfurization_ Rate	Thermal Power Pant Log (1+Plant_Capacity)		All Dirty Sector Log (0.1+Number_Firms)		SO_2_Remove_ Rate	Clean_Energy_ Share
		Entry	Exit	Entry	Exit		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
TCZ × *Post2005*	3.701*	-1.960***	0.793***	-0.035	0.154*	6.651***	1.458
	(2.116)	(0.475)	(0.285)	(0.142)	(0.087)	(2.405)	(1.877)
Dependent variable mean	22.621	3.422	1.512	1.066	2.684	31.256	31.681
City dummy		Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes
TCZ Controls × f(t)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Other Controls × Post2005	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Firm_Size	Yes						
Firm_Age	Yes						
Firm dummy	Yes						
No. of observations	22,100	2,860	2,574	2,288	2,288	2,288	2,860
Year coverage	2001-2010	2001-2010	2001-2008,2010	2001-2008	2001-2008	2003-2010	2001-2010
Adjusted R-squared	0.696	0.211	0.262	0.525	0.593	0.695	0.848
No. of clusters	277	286	286	286	286	286	286

Notes: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. In column 1, observations are at firm-level. In column 1, natural log of firm age and firm size are further added as control variables. In columns 2-7, all observations are at the city-level. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level.

Table 7

Revealed Preference by Local Bureaucrats

Dependent variable	Government_EcoIndex	Target_GDP_Growth
	(1)	(2)
TCZ × Post2005	0.048**	-0.735***
	(0.020)	(0.220)
Dependent variable mean	0.420	11.822
City dummy	Yes	Yes
Year dummy	Yes	Yes
TCZ Controls × f(t)	Yes	Yes
Other Controls × Post2005	Yes	Yes
No. of observations	4,288	4,284
Year coverage	2001-2015	2001-2015
Adjusted R-squared	0.495	0.608
No. of clusters	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level.

Table A.1

Dirty Sectors (SO_2 Emission Intensive Sectors)

Sector (2-digit)	SO_2 emissions share in total industrial SO_2 emissions
Production and Supply of Electric Power & Heat Power	50.4%
Non-metallic Mineral Products	12.7%
Pressing of Ferrous Metals	10.4%
Raw Chemical Materials and Chemical Products	6.1%
Smelting and Pressing of Non-ferrous Metals	5.8%
Processing of Petroleum, Coking, Pressing of Nuclear Fuel	3.1%
Sum	88.5%

Data source : *China Pollution Source Census* 2007.

Table A.2
Robustness Tests

Dependent variable	Log(SO_2)	SO_2	Log(SO_2)
	(1)	(2)	(3)
TCZ × Post2005	-0.142**	-0.215	-0.244***
	(0.067)	(0.281)	(0.064)
Dependent variable mean	1.338	5.948	1.338
City dummy	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes
TCZ Controls × f(t)	Yes	Yes	
Other Controls × Post2005	Yes	Yes	
Log(Investment_PerCapita)	Yes		
Log(GDP_PerCapita)			Yes
Manu_GDPShare			Yes
Coal_Abundance			Yes
Trade_Intensity			Yes
No. of observations	4,283	4,283	4,283
Year coverage	2001-2015	2001-2015	2001-2015
Adjusted R-squared	0.847	0.889	0.839
No. of clusters	286	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. "TCZ Controls × f(t)" includes interactions between a third-order polynomial function of time and seven key criteria that the central government used to select TCZ cities. These variables are: roughness, elevation, wind speed, precipitation, soil pH level, coldness and dirty sector employment (see appendix table A1 for definitions of dirty sectors). "Other Controls × Post2005" includes interactions between a dummy variable Post2005 and a set of variables that are known in the literature to be affected by the TCZ policy, which includes FDI, mortality and dirty sectors' export sales share. The standard errors are reported in parentheses and clustered at city level.

Table A.3
Balance Tests

Variables	TCZ cities	Non-TCZ cities	Unconditional diff.	Conditional diff.
	(1)	(2)	(3)	(4)
Panel A: Selection criteria				
Average elevation	0.493	0.558	-0.064	
(km)	[0.544]	[0.654]	(0.073)	
Land surface roughness	19.858	20.161	-0.303	
	[10.818]	[13.563]	(1.485)	
Percent of days with temperatures of 5° or below 1990-1995	15.833	22.887	-7.054***	
(%)	[14.303]	[16.279]	(1.844)	
Annual wind speed 1990-1995	21.925	24.013	-2.088*	
(0.1m/s)	[9.157]	[9.863]	(1.141)	
Average soil PH level of the topsoil	3.296	3.285	0.011	
	[1.037]	[1.044]	(0.124)	
No. of employees in dirty sector 1996	13.372	6.738	6.634***	
(10,000 persons)	[11.535]	[4.782]	(1.003)	
Annual average precipitation 1990-1995	10.602	8.881	1.721***	
(m)	[4.921]	[4.920]	(0.587)	
Panel B: Other characteristics				
Total lights at night	5.880	3.867	2.012***	0.378
($\times 10^{-4}$)	[5.272]	[3.224]	(0.505)	(0.352)
Population density	4.510	3.220	1.291***	0.196
(100 persons per sq. km)	[3.131]	[2.556]	(0.336)	(0.281)
Captial-to-labor ratio	3.342	2.485	0.857**	0.375
(1,000 Yuan per worker)	[2.352]	[3.262]	(0.346)	(0.265)
SOE employment share	48.358	52.087	-3.729**	-2.420
(%)	[13.989]	[12.757]	(1.587)	(1.703)
Foreign firm industrial output share	8.072	4.786	3.286***	1.100
(%)	[9.609]	[7.221]	(0.995)	(0.957)
Government education expenditure per capita	1.635	1.230	0.405***	0.183
(10 Yuan per capita)	[1.826]	[0.706]	(0.157)	(0.161)
Road density	7.897	6.325	1.572***	0.811
(km of road per 100 sq. km of land area)	[5.712]	[4.082]	(0.580)	(0.501)
No. of international sister cities	3.531	1.306	2.224***	0.108
	[5.137]	[2.088]	(0.445)	(0.422)
Observations	162	124		

Note: This table reports the summary statistics of our treatment and control samples. Panel A shows the comparison of selection criteria between the treatment and control groups. Panel B compares the treatment and control groups on various economic and social development variables in 2001, both before and after controlling for the selection criteria. Columns 1 and 2 show means and standard deviations in square brackets. Column 3 reports the unconditional difference between the treatment and control group. Column 4 reports the conditional difference of these characteristics of a regression on the treatment dummy controlling for the selection criteria. The standard errors are reported in parentheses.