Accountability, Career Incentives, and Local Bureaucrats: Evidence of A Target-Based Performance Evaluation System in China

Yvonne Jie Chen* National University of Singapore

Pei Li† Xiamen University

Yi Lu‡ National University of Singapore

March 15, 2016

Abstract

This paper studies the effect of a target-based performance evaluation system for high-level bureaucrats in China. In 2006, the Chinese central government made sulphur dioxide (SO_2) emissions cuts a main performance evaluation criterion for prefectural city mayors and party secretaries. Using a difference-in-difference framework, we compare emissions reductions in the so-called Two Control Zone cities, where more stringent emissions quotas were imposed, with reductions in non-control cities. Our results suggest that the target-based evaluation system is effective in achieving the targeted policy output. Adding the emission cuts into local bureaucrats' performance evaluations significantly reduces actual SO_2 emissions. By exploring a rich set of official characteristics, we further analyze the channels through which the policy operates. We show that the target-based evaluation system changes bureaucrats' behavior and induces them to exert more effort on controlling environmental pollution.

*sppcj@nus.edu.sg
†bengcome@gmail.com
‡justinly6@gmail.com

1 Introduction

Over the past decade, experience with the Millennium Development Goals shows that good governance supports sustained progress towards economic growth, poverty reduction and equality. Good governance has also become the center of discussion for the UN post-2015 development agenda (UNESCO [2012]). The quality of state institutions, including the performance of central and local government bureaucracies, holds the key to building effectively functioning governance in many developing countries (Evans and Rauch [1999], Rauch and Evans [2000]). How bureaucrats can be motivated to achieve efficient public management as well as public goods and services delivery is, therefore, an important question for policymakers and scholars.

In this paper, we study the effect of a target-based performance evaluation system on the behavior of local bureaucrats. Career concern theories suggest that top bureaucrats are largely driven by the outcomes of their mandated tasks (Holmstrom [1982], 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 non-public sectors. In addition to career motives, studies show that bureaucrats are also motivated by peer pressure, social norms, and bureaucrats' own "glow of warmth" (Wilson [1989], Maskin and Tirole [2004]).

Various incentive schemes for bureaucrats have been studied in the empirical literature. Previous studies have generated mixed evidence on the effectiveness of these schemes. The public administration literature has questioned incentivizing public sector workers with pay-for-performance. It is only shown to be effective when bundled with other incentives and proper management practice(Besley [2004], Perry et al. [2009]). Khan et al. [2016] show that an evaluation system based purely on performance-pay for tax collectors successfully increased tax revenue by fourteen percent, with no significant political cost. Recently randomized studies have been used to estimate the effectiveness of monetary versus implicit incentive schemes such as career advancement(Duflo et al. [2012], Ashraf et al. [2015]). These

studies, however, focus heavily on front-line public sector workers rather than bureaucrats at the management level.

In this paper, we examine the effect of a target-based performance evaluation system implemented in China in 2006. China has been the fastest-growing major economy over the past three decades, but growth comes at the cost of heavy environmental pollution[1]. The central government has realized the pressing environmental issues since the 1990s and has gradually shifted its focus from GDP growth to a more sustainable approach that places a much heavier weight on environmental protection. One of the major policy initiatives proposed by the central government in the 1990s was the 1998 establishment of the Two Control Zones policy (for control of sulphur dioxide and acid rain), hereafter TCZ. The main objective of the TCZ policy was to reduce SO_2 emissions, which have long been a major contributor to China's ambient air pollution[2],[3]. Based on the TCZ policy guideline, 162 out of 286 prefectural cities were assigned to be TCZ cities and had to follow more strict environmental regulations than the non-TCZ regions (Figure 1). However, the effect of the TCZ policy on SO_2 emissions reductions was temporary, and emissions continued to increase after a short dip in 1999-2000 (Figure 2). In order to reach the total SO_2 emissions cut targets (also known as the "green goal"), the central government imposed nationwide emissions quotas in 2000, with more stringent measures applied to TCZ cities. In 2006, further administrative regulations were imposed and the emissions quota was brought into the evaluation system, making local bureaucrats subject to administrative demerit or removal from office if the quota was not met.

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

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

[3]Existing studies have shown that this policy of TCZ affects investment (Lu et al. [2015]), exports (Hering and Poncet [2014]) and health (Tanaka [2015]).

Using the implementation of the target-based performance evaluation system imposed by the central government on TCZ cities and non-TCZ cities, we find that local bureaucrats exert more effort on SO_2 reduction when the emissions quota is built into their performance evaluation. Compared to the control group (non-TCZ cities), TCZ prefectural cities exhibited a similar pattern of reduction in SO_2 before 2006; after SO_2 reduction became part of the performance measures in 2006, TCZ prefectural cities experienced much higher SO_2 reductions.

China provides a good setting to study the effectiveness of target-based performance incentives for bureaucrats for several reasons. First, previous studies suggest that top local bureaucrats in China, including provincial and prefectural city government leaders, are largely motivated by their career concerns (Chen et al. [2005], Li and Zhou [2005], Jia [2014]). In China, the most important career consideration for local bureaucrats is promotion along the government hierarchy. Specifically, a city's China Communist Party (CCP) secretary aspires to be promoted to the provincial government, while the city mayor seeks promotion to city CCP secretary or to the provincial government. This career incentive is inline with bureaucrats in many other countries around world. Findings in China, therefore, help build general understanding of the impact of a target-based performance evaluation system on bureaucratic behaviors. Second, at the prefectural city level, city mayors and party secretaries oversee the delivery of public goods and services subject to goals and regulations set by the central government. They also have a certain degree of autonomy to implement policies at their own discretion. Hence the delivery of public goods and services varies widely across China depending on the characteristics and competence of local officials. The regional variation allows us to identify the policy impact by exploring the rich data on output and government official characteristics at the prefectural city level.

Our paper adds new empirical evidence on incentivizing bureaucrats with concrete performance targets to the economics literature. Lockwood and Porcelli [2013] show that the comprehensive performance assessment system in England has increased public service qual-

ity but has no significant effect on efficiency overall. Recent studies in Nigeria (Rasul and Rogger [2015a,b]) suggest that concrete targets have a detrimental effect that leads to dysfunctional responses among bureaucrats. The failure of performance-based incentives is usually attributed to the multitask nature of bureaucratic jobs and their intrinsic motives (Dixit [2002], Alesina and Tabellini [2008]). As bureaucrats usually juggle multiple tasks or multiple dimensions of a task, concert performance targets potentially cause social distortion by requiring bureaucrats to exert effort on less productive outcomes (Rasul and Rogger [2015b]). Emphasizing output as a performance measure would also likely crowd out the intrinsic motivations for bureaucrats (Benabou and Tirole [2003]). Our results, however, suggest that a target-based evaluation system can be effective in motivating bureaucrats to shift effort towards the task that is more heavily weighted by the evaluators and achieving the targetted policy output.

Concerns on gaming the target-based performance evaluation system were also brought up in previous studies (Bevan and Hood [2006], Chen et al. [2012], Gao [2015]. We address the issue with detailed firm level data from SO_2 intensive sectors[4]. We find that it is much less likely to establish a new coal fire plant in TCZ cities after the implementation of the target-based evaluation system. SO_2 intensive firms are also more likely to adopt desulfurization technologies. These evidence suggests that our results are probability not driven by gaming or data manipulation.

Our paper also adds to the literature on career incentives and occupation selection in the public sector. Studies on organization theories have shown that employees tend to sort themselves according to the job attributes and incentives provided. Job incentives would hence affect the performance outcomes through selection on employees (Roy [1951], Oyer and Schaefer [2010], Lazear and Oyer [2012],Ashraf et al. [2015]). In analyzing the mechanisms for the new evaluation system, we collected a large amount of information on individual characteristics, political tenure and job turnover for city mayors and secretaries from various

[4]See section 3.1 for definitions of SO_2 intensive sectors and firms

Chinese sources. We also extracted words of mentions for environmental related topics and other government objectives from city governments' Annual Work of the Government Report (*zhengfu gongzuo baogao*) between 2001 and 2013. Our analyses show that SO_2 reductions were not a result of office turnover or sorting of bureaucrats into local offices. This strengthened our hypothesis that the observed emissions reduction was the result of changed behavior and focus of local bureaucrats.

Our finding contributes to a 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 is usually associated with political connections, promotion probability and other political concerns (Zheng et al. [2014], Jia [2014], Kahn et al. [2015]). Our paper shows that the readiness of the center to reward and punish local officials on the basis of their environmental performance could also motivate local government officials to reach the green goal. [5]

The remainder of the paper is organised as follows. Section 2 provides a review of the policy background on the TCZ policy and performance evaluation system. We then present data description, empirical strategy and results in sections 3 and 4.

2 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 (FYP) in 1995, China's central government began to set limits on total SO_2 emissions. In 1998, the State Council instituted an SO_2

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

reduction program, known as the TCZ policy, to limit ambient SO_2 pollution and to curb the growing incidence of acid rain. The TCZ policy encompasses 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. Aggregately, the TCZ regions cover 11 percent of China's territory (see Figure 1) and are responsible for over 60 percent of China's total SO_2 emissions.[6] A total emissions control policy was incorporated into China's 10th FYP in 2000 to work in concert with the TCZ policy[7]. The goal set by the 10th FYP was to reduce SO_2 emissions by 10 percent nationwide and by 20 percent for the TCZ regions, as compared to 2000 levels[8].

To force localities to adhere to the national policy, two administrative enforcement mechanisms were put into place beginning in 2005. First, starting from Dec 2005, local 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[9]. The top-down target-based responsibility scheme ensures that local officials are tied to satisfying higher-level mandates for career advancement and legitimacy[10]. Second, the central government announced in 2007 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. SO_2 emissions reductions hence became one of the most important factor in the assessment of local cadres. Local government leaders who fail to meet the annual SO_2 emissions reduction targets would be penalized with an

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

[7]Relevant Chinese document for the 10th FYP guideline for TCZ policy is available at following government website: `http://www.ynepb.gov.cn/xxgk/read.aspx?newsid=11012`

[8]A series of policies were implemented in TCZ regions in order to meet the emissions reduction targets, including the closure of high SO_2 concentration coal mines, more control of SO_2 emissions from coal-fired power plants and industrial boilers, and the introduction of more energy-saving and low pollution technologies (Cao et al. [2009]). More generous fiscal policies were also implemented in TCZ regions, including subsidies for energy saving projects and the establishment of an SO_2 quota trading system. Higher pollution fees were also imposed in some regions.

[9]Relevant Chinese document is available at following government website: `http://www.gov.cn/zwgk/2005-12/13/content_125680.htm`

[10]In the meantime, the central government banned new construction projects in regions with unsatisfactory SO_2 emissions reduction, in an effort to align incentives for economic growth with local governments' achievement of the green goal.

administrative demerit or removal from office[11]. 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.

Before the administrative enforcement policies were imposed, the green goal set in China's 10th FYP (in 2000) was not strictly implemented. By the end of 2005, SO_2 emissions not only had not reduced, but had in fact increased by approximately 28 percent from 2000 levels (see Figure 2). After 2005, the consistent emphasis by the central government on the need to reduce SO_2 emissions has made it incrementally more difficult for both 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 was finally met and an overall 14.29 percent reduction achieved nationwide (see Figure 2).

3 Estimation Strategy

3.1 Data and Variables

Data for the empirical analyses presented in this section are 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 meteorological conditions of each city spanning the years from 2001 to 2013. Summary statistics are presented in Table 1.

Pollution Data

The industrial SO_2 emissions and wastewater discharge data are collected from the annual statistical yearbooks of the Chinese provinces, supplemented by *China City Statistical*

[11]Relevant Chinese document is available at following government website: `http://www.gov.cn/jrzg/2007-06/03/content_634545.htm`

Yearbooks 2004-2014 and *China Environmental Yearbooks* 2002-2014[12]. We use emission data between 2001 and 2013 for analyses presented in this paper. Data before 2001 were subject to significant missing values hence were not included in the study. The city-level annual SO_2 emissions are calculated by summing up emissions from all reporting firms in a given year. For each reporting firm, the State Environmental Protection Agency (SEPA) records its SO_2 emissions through direct monitoring. If direct monitoring is not feasible, SEPA would use reverse engineering measures computed from the industrial waste reported by the firm. The list of reporting firms is determined by SEPA and is subject to minor revisions each year. The city-level coal-fired generation capacity data are aggregated from the plant-level data documented in the *Compilation of Statistical Materials of Electric Power Industry* 2001-2010,[13] and the SEPA's annual communiqués on the shutdown of small thermal power plants.[14] The PM2.5 data are extracted from the grid data of *Global Annual PM2.5 Grids* from MODIS and MISR Aerosol Optical Depth, which provides a continous surface of concentrations of particulate matter of 2.5 micrometers or smaller[15].

Other City-Level Variables

The city slope and elevation are extracted from the *Shuttle Radar Topographic Mission 90m Digital Elevation Model* data. We collect city wind speed, temperature and precipitation data from *the China Meteorological Data Sharing Service System* (http://cdc.cma.gov.cn/). Soil pH data are 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. A dirty sector is defined by matching the 2-digit industry classification code to that of major SO_2 pollution sectors identified by the *Report of China Pollution Source Census 2007*. As the only pollution source census in China up to this date, the 2007 census examined 5.92 million

[12]According to the China National Pollution Census 2008, industrial activity is responsible for 91.4 percent of total SO_2 emissions. The remaining part is contributed by domestic activity (8.6 percent).

[13]This dataset covers all thermal power plants with installed capacity of 6,000 KW and above.

[14]SEPA has stopped publicly publishing the list of small thermal power plants shutdown annually in 2011.

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

pollution sources across all provinces in China. The final report identified six industries (as shown in Table 2) 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. The fatality data of workplace accidents are collected from the official website of the China State Administrative of Work Safety (http://www.chinasafety.gov.cn). We aggregate the daily data by year and city.

Characteristics of City Government Officials

We collect detailed demographic data including age, education level and length of service at current position. The detailed biographical data are compiled manually by the authors from various publicly accessible sources, including the online biographical database China Vitae (`http://www.chinavitae.com/`).

Mentions of Environmental Protection

We conduct a thorough content analysis of the city governments' Annual Work of the Government Report between 2001 and 2013. The Annual Work of the Government Report is usually presented by the city mayor at the 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 the general population. Each report outlines the target GDP growth rate for the year as well as other policy objectives for the city government. The reports are not available in any existing database or statistical year books. We collect information by manually searching more than 3,700 reports via city governments' websites and local newspapers. Our paper is, to our knowledge, the first attempt to use information extracted from the Annual Work of the Government Report for empirical analysis. Specifically, we search for the words "environmental protection" (*huan bao* or *huanjing baohu* in Chinese) and "emissions reductions" (*jian pai* in Chinese), and calculate the ratio of these mentions to the total number of words in the report. We use the ratio as a measure of effort that a city government leader places on environmental issues.

3.2 Estimation Framework

To investigate whether the performance-based evaluation system affects government officials' behavior in implementing and enforcing environmental regulations, we compare the emissions of SO_2 in TCZ cities (those cities with strict environmental regulations) before and after 2006 (the year that the performance-based evaluation system was adopted) with that of non-TCZ cities during the same period, or a difference-in-differences (DD) estimation. Specifically, our regression equation is

$$LnSO2_{ct} = \beta TCZ_c * Post05_t + \delta_c + \gamma_t + \mathbf{Z_c} * \mathbf{f(t)} + \varepsilon_{ct}, \tag{1}$$

where $LnSO2_{ct}$ is the logarithm of SO_2 emission in city c at year$t$; TCZ_c equals 1 if city c was designated as a TCZ city in 1998 and 0 otherwise; $Post05_t$ takes a value of 1 if $t > 2005$ and 0 otherwise; δ_c is 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, following the suggestion by Bertrand et al. [2004].

We expect β to be negative if the performance-based evaluation system is effective. This is because TCZ cities were imposed larger targeted reductions in SO_2 in 2000 and the new evaluation system was introduced in 2006. However, the unbiased estimation of β requires that SO_2 emissions in TCZ cities would have followed the same time trend as that in non-TCZ cities after 2005 if a performance-based evaluation system had not been imposed by the central government in 2006. 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 these unobserved differences changed their values in 2006, 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 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 between 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 on the TCZ selection. 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 collect two sets of variables as our TCZ selection variables. The first set of variables reflect conditions of the natural environment based on the 1995 APPCL amendment: *Roughness* (the standard deviation of slope), *Elevation* (average elevation in km), *Wind Speed* (annual average wind speed in 1990-1995), *Precipitation* (annual average precipitation in 1990-1995), *Soil pH* (average pH level of topsoil), *Coldness* (percentage of days with temperature of 5^oor below in 1990-1995). The second set of variables reflect pre-policy SO_2 pollution levels, which include *Coldness* (percentage of days with temperature of 5^oor below in 1990-1995) and the total employment in "dirty sectors" in 1996[16]. To flexibly control for trends in SO_2 emissions generated by these selection variables over time, we interact them with a

[16]Ideally we should include pre-policy level SO_2 level as a determinant for TCZ designation for each city but unfortunately the data are not available. We therefore 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. When the daily outdoor temperature drops to 5°C or below for a few days, the north part 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.

third-order polynomial function of time $\mathbf{Z_c} * \mathbf{f(t)}$.

As further checks on our identifying assumption in equation 1, we conduct several sensitivity exercises, including a graphic verification of whether TCZ and non-TCZ cities had similar trends in SO_2 emissions before the treatment, a placebo test with random assignment of TCZ status and timing of the adoption of the performance-based evaluation system, and falsification tests with other pollutants that the policies did not address. For details on these robustness checks, see Section 4.2.

4 Empirical Findings

4.1 Main Results

Table 3 reports main results from equation 1. Column 1 presents results from a simple DD estimation with only 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 2006, this result implies that the new evaluation system has disciplined or incentivized bureaucrats to enforce the environmental regulations.

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 $TCZ_c * Post05_t$, despite a drop of around 40 percent in the magnitude. In column 3, we further include, as a control variable, an interaction between TCZ_c and $Year05_t$ to investigate whether bureaucrats anticipated the new performance-

based evaluation system in 2006 and hence changed their behavior before the adoption (known as the Ashenfelter dip). We find a statistically insignificant estimate of $TCZ_c * Year05_t$ with barely any change in magnitude, suggesting no Ashenfelter dip in our research setting.

As the new performance-based evaluation system consists of two measures implemented respectively in 2005 and 2007, we divide our regressor of interest into two, $TCZ_c * Year06\_07_t$ and $TCZ_c * Post2007$, to incorporate the effects from the two post-treatment periods. As shown in column 4, both of the two regressors are negative and statistically significant, with the latter having about 78 percent larger magnitude, suggesting that both of the two measures in the new performance-based evaluation system matter.

4.2 Robustness Checks

Our aforementioned estimates hinge on the identifying assumption that conditional on the determinants for TCZ selection, the central government did not select TCZ cities on the remaining factors with the expectation that eight years later it would pass a new performance-based evaluation system and these remaining factors would change the trends of SO_2 emissions between TCZ and non-TCZ cities differently after the new evaluation system. To further verify that our identifying assumption sustains, we report in this section a battery of robustness checks.

Similar 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 SO_2 emissions before the treatment. Specifically, if similar time trends of SO_2 emissions existed for a long period before the treatment and suddenly diverged right after the treatment, it may indicate that the differential trends may be caused by the treatment. Figure 3 plots the time trend of differences in SO_2 emissions between TCZ and non-TCZ cities over 2002-2013 relative

to the difference in 2001. Each dot in Figure 3 estimates the additional gap in emissions between TCZ and non-TCZ cities using the gap in 2001 as the baseline[17]. Indeed, we find that the differences are close to zero from 2001 to 2005, suggesting a common trend between TCZ and non-TCZ cities for the five years before the treatment. Meanwhile, the difference dropped to negative immediately after the adoption of the treatment in 2006 and continued to be negative with increasing magnitude. These results confirm that the treatment in 2005 generated the differential trends in SO_2 emissions between TCZ and non-TCZ cities.

A placebo test with randomization of the treatment. As a second robustness check, we randomly assign the TCZ status to cities TCZ_c^{false} and the timing of the adoption of the new performance-based evaluation system $Post_t^{false}$, and then construct a new regressor of interest $TCZ_c^{false} * Post_t^{false}$. Given this random data generating process, a necessary condition for satisfying our identifying assumption is that $TCZ_c^{false} * Post_t^{false}$ shall produce zero effect on SO_2emissions; 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 distribution of 1,000 estimated coefficients is plotted in Figure 4. It is centered around zero and only 1.1 percent of the estimates are more negative than our true estimate in column 2 of Table 3. These results suggest that there likely were not substantial omitted variables in our specification[18].

Alternative pollutant measures. We use PM2.5 as an alternative measure of SO_2 emissions. PM2.5 data have been used extensively in the environment literature and are shown to be a reliable measure of air quality. PM2.5 particles can be directly emitted or produced from emitted precursors. Pollutants such as SO_2, nitriate (NOx) 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. The estimation result for the policy effect on PM2.5 particles is reported in column 5 of Table (3). The estimated coefficient is negative and statistically significant. A

[17]All other control variables are the same as column 2 in Table 3.

[18]See Chetty et al. [2009] and La Ferrara et al. [2012] for review and application of the method

decrease in PM2.5 as a result of SO_2 control policy in absence of regulations for the other reactors supports our hypothesis that the policy has effectively reduced SO_2 emissions.

Analyses with untargeted pollutants. Our estimates would be biased if, in 2006, along with the change in the evaluation system, there were other differential changes between TCZ and non-TCZ cities such as industrial structure and technology upgrading that affected environmental conditions. However, the TCZ policy exclusively targeted SO_2 emissions and did not address other pollution measures, such as wasterwater. Hence, if our aforementioned estimates are mainly due to an evaluation system that made bureaucrats more concerned about the enforcement of the TCZ policy, we shall not spot sizable differentials among alternative pollution measures between TCZ and non-TCZ cities after 2006. The estimation result for the policy effect on wastewater is reported in column 6 of Table 3. The estimated coefficient is small in magnitude and insignificant. The result lends further support to the satisfaction of our identifying assumption.

4.3 Effect of the Administrative Enforcement on GDP Growth and Industrial Sectors

First, we estimate the effect of the performance evaluation system on GDP growth rate and target GDP rate at the city level. Historically, GDP growth has been the indicator of competency of local government officials and has been used as a basis for promotion (Jia [2014], Xu [2011]). As outlined in Appendix A, when there is a trade-off between multiple tasks, bureaucrats will exert more effort on the task that's more heavily weighted by their supervisors. Hence, the shift of focus to environmental protection may lead to a reduced emphasis on economic growth by the city government. To this end, we examine the target GDP growth rate mentioned in the city governments' Annual Work of the Government Report, in both absolute terms and relative terms (that is, relative to the target GDP growth rate by the provincial government). Estimation results are reported in columns 1 and 2 of Table 4. Consistently, we find that after 2005, TCZ city had lower GDP growth rates, as

shown in columns 3 in Table 4, indicating a shift of focus away from economic growth in response to the new evaluation system.

In order to examine the policy effect on industrial sectors, we first examine the probability of coal fire plants entry and exit at the city level. To control SO_2 emissions, the TCZ policy restricted the opening of coal fire plants and demanded the closing of outdated coal fire plants in the TCZ cities. Hence, if our aforementioned estimates are caused by a change in enforcement of the TCZ policies due to the new evaluation system in 2006, we shall find consistent patterns of opening and closing of coal fire plants, specifically, less openings and more closings in TCZ cities after 2005. Indeed, the regression results in columns 1 and 2 of Table 5 confirm these results: $TCZ_c * Post05_t$ is negative and statistically significant in the regression of plant openings, and positive and statistically significant in the regression of plant closings. We also check the probability of installing desulfurization equipment at firm level. We expect that the performance evaluation would lead to a higher probability of installing desulfurization equipment among TCZ cities as an enforcement measure taken by local bureaucrats to reduce SO_2 emissions. Results in column 3 of Table 5 confirm our hypothesis. The coefficient for $TCZ_c * Post05_t$ is positive and statistically significant, meaning that firms in TCZ cities are indeed more likely to install these environmental protection technologies. These results provide micro-evidence that the target-based performance evaluation system has led to sectoral shifts and technology changes at the plant level.

4.4 Mechanisms and Alternative Explanations

We have documented that SO_2 emissions dropped in TCZ cities relative to that of non-TCZ cities after 2005. The change in outputs can be explained by theories of career concerns as outlined in Appendix A. Motivated by the theoretical works of Holmstrom [1989] and Alesina and Tabellini [2007], we hypothesize that bureaucrats exert more effort on controlling SO_2 emissions when the cut quota is built into the evaluation system.

We first provide a content analysis of the city governments' Annual Work of the Government Report. We use this to examine the emphasis a city government places on environmental issues in the coming year. Specifically, we search for the words "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 column 4 of Table 4. 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.

We also examine the effect of random onset of fatal accidents in a city on SO_2 emissions levels. The new performance-based evaluation system reduces the probability of promotion if the bureaucrats did not meet the annual SO_2 emissions cut target. However, other policies also affect a bureaucrat's likelihood of promotion. In particular, a provision issued by the central government in 2001 declared that "when an extremely fatal accident happens, the local government officials would be penalized based on the severity of the accident, either through demotion or dismissal, or through criminal procedures if dereliction of duty is involved"[19]. Hence, in the case of a fatal accident, officials' promotion probability would be negatively affected[20]. In the event of a fatal accident, the new performance evaluation system would incentivize bureaucrats to exert even more effort towards SO_2 emissions cuts so that they can compensate for the fatal accident[21]. Following this reasoning, we examine whether the occurrence and severity of an accident further increase the gap in the reduction

[19]The Chinese document can be found on the website of China government at: `http://www.sxaj.gov.cn/flfg/fg/gjfl/200905/881148.html`. A workplace accident is regarded as very serious if there are more than 10 fatalities.

[20]Fisman and Wang [2015] show that firms with political connections do not receive favorable treatment when fatal accidents happen and connections-mortality relationship is reduced in provinces where local officials are subject to certain safety targets. Their evidence suggests that it is difficult for local officials to avoid regulatory responses when fatal accidents happen.

[21]Only 3 out of 2499 city secretaries/mayors had been demoted because of accidents during our targeted period. As a result, the incidence of fatal accidents merely reduces the likelihood of promotion; it does not completely destroy the possibility of promotion.

of SO_2 emissions between TCZ and non-TCZ cities after 2005 in Table 6. The triple interaction term, $TCZ_c * Post05_t * Fatal\_Accident_{ct}$, is negative and statistically significant, regardless of $Fatal\_Accident_{ct}$ is measured by a dummy variable indicating with or without fatal accidents or a continuous variable capturing the death toll. Our results indicate that local bureaucrats treat SO_2 reductions as a coping mechanism when an unforseenable event happens and exert more effort on environmental issues in order to make up for what has happened.

An alternative explanation for the observed reduction in SO_2 is occupation selection. Many previous studies suggest that incentive packages and attributes of a job sort individuals into the "right" type of jobs (Ashraf et al. [2014, 2015], Roy [1951]). Hence it is possible that more environmentally conscious bureaucrats were put into office in TCZ cities after 2005 as compared to non-TCZ cities.

In China, the most important bureaucrat at the city level is the secretary of the CCP, who is in charge of the city's major decisions; the second most important bureaucrat is the city mayor, who is responsible for the daily operation of the government. To understand whether our findings are driven by occupation selection of bureaucrats with different environmental orientations across TCZ and non-TCZ cities, we conduct two sets of exercises.

First, we check whether the pattern of length of appointment and frequency of office turnover diverge between TCZ and non-TCZ cities after 2005. Specifically, we examine the probability of having a new CCP secretary and city mayor, the term length of the incumbent CCP secretary and city mayor, and the promotion of CCP secretary and city mayor. Regression results are reported in Table 7, with columns 1-3 for city mayors and columns 4-6 for CCP secretaries. Consistently, we find all the estimates are statistically insignificant and small in magnitude. These results indicate that TCZ and non-TCZ cities had similar patterns of CCP secretary/city mayor turnover, term length of the incumbent CCP secretary/city mayor, and promotion probability of CCP secretary/city mayor after 2005.

Second, we examine various individual characteristics of city CCP secretaries and city mayors to check whether TCZ and non-TCZ cities attracted different types of bureaucrats after 2005. Specifically, we look at age, education level, and whether the bureaucrats are working in their hometown in Table 8. We also look at various indicators of bureaucrats' working experience, including service length in the concerned city, previous experience in a firm, previous experience in upper-level government, and promotion from lower-level county government in Table 9. Across all the regressions, we consistently find that the coefficients of $TCZ_c * Post05_t$ are statistically insignificant and small in magnitude. These results indicate that TCZ and non-TCZ cities have similar governing bureaucrats after 2005.

In summary, while we are unable to directly check whether TCZ and non-TCZ cities have governing bureaucrats with different preferences or competence levels for implementing environment regulations after 2005, our analyses show that these two groups of cities have similar types of bureaucrats and no difference in office turnover and length of service. This helps dispel the argument that our results are driven by sorting of more environmentally concious bureaucrats into positions in TCZ cities than non-TCZ cities in the post-2005 period. Our tests also do not suggest that the observed results were due to divergent data manipulation patterns between TCZ and non-TCZ cities.

5 Conclusion

In this paper we examine the effect of a target-based performance evaluation system on SO_2 emissions reduction. We find that when SO_2 emissions quota were built into local bureaucrats' evaluation system, SO_2 emissions were significantly reduced. Our analysis suggests that the new evaluation system induced local bureaucrats to exert more effort on environmental issues. Using a comprehensive dataset that we constructed, we observe a clear shift of effort and focus from GDP output to environmental protection among these local bu-

reaucrats. Our finding does not suggest that the emissions reductions were a result of office turn-over or sorting.

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's more heavily weighted by their supervisors. It might also crowd out bureaucrats' intrinsic motivation. Our analysis shows a clear pattern of local bureaucrats making a conscious trade-off between GDP 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.

Appendix A

A Conceptual Framework

In this section, we present a conceptual framework that illustrates the behavior of bureaucrats in response to a target-based evalaution system. Following the seminal work by Dewatripont et al. [1999] and Alesina and Tabellini [2007], bureaucrats are partly motivated by their "career concerns". In China, career concerns for bureaucrats include moving up the hierarchical leadership ladder or moving to a state-owned enterprise.

Following Alesina and Tabellini [2007], suppose that a bureaucrat focuses on two policy outcomes, y_1 and y_2. To achieve each outcome, the bureaucrat has to exert effort level a_1 and a_2, respectively. Let the cost of exerting effort be $C(a_1, a_2)$ and the cost function be strictly increasing and convex. This cost is assumed to be public information. The bureaucrat also cares about the evaluation by her supervisor as it matters for her career concerns. The policy output for each period is therefore $y_1 = \eta_i + a_1 + \epsilon_1$ and $y_2 = \eta_i + a_2 + \epsilon_2$, respectively, where η_i is the ability for a typtical bureaucrat i drawn from a distribution with mean $\bar{\eta_i}$ and variance σ_{η_i}. ϵ_1, ϵ_2 are random draws from a i.i.d. distribution. Suppose the aggregate performance of the bureaucrat is assessed by $x = \phi \times y_1 + (1 - \phi) \times y_2, 0 < \phi < 1$, where ϕ is the weight assigned by the supervisor to each implementation output. x therefore is the performance measure of the bureaucrat according to some stated goals.

Let δ be the market value of the signaled ability, then return for the bureaucrat can be written as $R^B(a_{1,}a_2) = \delta E(\eta_i|x)$[22]. The goal for the bureaucrat is to choose her effort level so that $R^B(a_1, a_2) - C(a_1, a_2)$ is maximized. First order conditions of the maximization problem yield

$$\begin{aligned} \delta\phi &= C_1(a_{1,}a_2) \\ \delta(1-\phi) &= C_2(a_{1,}a_2) \end{aligned}$$

[22]Note that $x = \phi y_1 + (1-\phi)y_2 = \phi(\eta_i + a_1 + \epsilon_1) + (1-\phi)(\eta_i + a_2 + \epsilon_2) = \eta_i + \phi a_1 + \phi\epsilon_1 + a_2(1-\phi) + \epsilon_2(1-\phi)$. Therefore $E(\eta_i|x) = x - \phi a_1 - (1-\phi)a_2 - E[\phi\epsilon_1 + (1-\phi)\epsilon_2|x]$ and can be written as a function of $a_1$and a_2.

where C_1, C_2 is the partial derivative of the cost function with respect to effort a_1, a_2. It is easy to show that $\frac{\phi}{(1-\phi)} = \frac{C_1}{C_2}$. This means that a necessary condition for optimization is that the ratio of the marginal cost of effort level between the two tasks is equal to the relative weight assigned by the evaluator. Given the convex cost structure, a higher weight leads to increased effort in a particular task.

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.

N. Ashraf, O. Bandiera, and B. K. Jack. No margin, no mission? a field experiment on incentives for public service delivery. *Journal of Public Economics*, 120:1–17, 2014.

N. Ashraf, O. Bandiera, and S. S. Lee. Do-gooders and go-getters: Career incentives, selection, and performance in public service delivery. Working Paper, 2015.

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.

T. Besley. Paying politicians: Theory and evidence. *Journal of the European Economic Association*, 2:193–215, 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.

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*, 3:231–250, 2009.

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. 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.

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

E. Duflo, R. Hanna, and S. P. Ryan. Incentives work: Getting teachers to come to school. *American Economic Review*, 102(4):1247–1278, 2012.

P. B. Evans and J. E. Rauch. Bureaucracy and growth: A cross-national analysis of the effects of "weberian" state structures on economic growth. *American Sociological Review*, 64(5):748–765, 1999.

R. Fisman and Y. X. Wang. The mortality cost of political connections. *Review of Economic Studies*, Forthcoming, 2015.

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

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

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

B. Holmstrom. Design of incentive schemes and the new soviet incentive model. *European Economic Review*, 17(2):127–148, 1982.

B. Holmstrom. Agency costs and innovation. *Journal of Economic Behavior & Organization*, 12(3):305–327, 1989.

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.

R. X. Jia. Pollution for promotion. Working Paper, 2014.

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.

A. Q. Khan, A. I. Khwaja, and B. A. Olken. Tax farming redux: Experimental evidence on performance pay for tax collectors. *Quarterly Journal of Economics*, Forthcoming, 2016.

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.

E. P. Lazear and P. Oyer. Personnel economics. In R. Gibbons and J. Roberts, editors, *Handbook of Organizational Economics*. Princeton University Press, 2012. NBER Working Paper No. 13480.

H. B. Li and L. A. Zhou. Political turnover and economic performance: the incentive role of personnel control in china. *Journal of Public Economics*, 89:1743–1762, 2005.

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.

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. Mimeo, 2015.

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

P. Oyer and S. Schaefer. Personnel economics: Hiring and incentives. In *Handbook of Labor Economics*, volume 4. Elsevier, 2010. NBER Working Paper No. 15977.

J. L. Perry, T. Engbers, and S. Y. Jun. Back to the future? performance-related pay, empirical research, and the perils of persistence. *Public Administration Review*, 69(1): 39–51, 2009.

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, 2015a.

I. Rasul and D. Rogger. Management of bureaucrats and public service delivery: Evidence from the nigerian civil service. Working Paper, 2015b.

J. E. Rauch and P. B. Evans. Bureaucratic structure and bureaucratic performance in less developed countries. *Journal of Public Economics*, 75:49–71, 2000.

A. D. Roy. Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2): 135–146, 1951.

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

UNESCO. Thematic think piece: Governance and development. Technical report, UNDESA, UNDP, UNESCO, 2012.

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. Cost of pollution in china. Technical report, Washington, DC: World Bank, Washington, DC, 1997.

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

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

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.

Table 1: Summary Statistics

Key Variables	**Definition**	**Mean**	**S.D.**
TCZ	=1 if a city is a TCZ city; =0 otherwise	0.566	0.496
SO_2	Total industrial SO_2 emissions (10^4 tons)	6.064	6.078
PM2.5	Concentrations of PM2.5 (micrograms/m3)	31.661	14.112
Wastewater	Total industrial waste water discharges (10^7 tons)	7.734	10.44
CoalFirePlant_Capacity			
Entry	New fossil-fuel power plant electricity generation capacity (10^4KW)	45.050	176.998
Exit	Closed fossil-fuel power plant electricity generation capacity (10^4 KW)	0.204	0.760
Fatal Accident			
Happened	=1 if an extremely fatal accident happened; =0 otherwise	0.169	0.370
Death Toll	Total fatality of the extremely fatal accidents (100 persons)	0.046	0.161
Gov_EcoIndex	Count of "green" keywords in the government work report		
	divided by total word count of the report in this year ($\times 10^3$)	0.399	0.287
GDP Real	Real GDP Growth (%)	12.820	3.430
GDPTarget	GDP growth target (%)	12.180	2.850
GDPTarget_Gap	City GDP growth target minus provincial GDP growth target	2.251	2.269
City mayer characteristics			
Age	Age	50.460	4.153
Education	=1 if the with a master or higher degree; =0 otherwise	0.687	0.467
Hometown	=1 if born in this city; =0 otherwise	0.090	0.289
New	=1 if in the first yea of his tenure; =0 otherwise	0.296	0.456
Tenure	Length of tenure in the current position	2.615	1.589
Service Length	Total service years in this city	6.912	9.542
Promotion	=1 if promoted this year; =0 otherwise	0.188	0.391
As a leader in a			
Firm	=1 if worked as the CEO in a big enterprise; =0 otherwise	0.197	0.398
Up-tier government	=1 if worked in a provincial/central government; =0 otherwise	0.518	0.500
County	=1 if worked as a county governor/secretary; =0 otherwise	0.428	0.495
City secretary characteristics			
Age	Age	52.600	4.056
Education	=1 if the with a master or higher degree; =0 otherwise	0.666	0.472
Hometown	=1 if born in this city; =0 otherwise	0.053	0.225
New	=1 if in the first yea of his tenure; =0 otherwise	0.266	0.442
Tenure	Length of tenure in the current position	2.877	1.757
Promotion	=1 if promoted this year; =0 otherwise	0.139	0.346
Service Length	Total service years in this city	6.740	8.004
As a leader in a			
Firm	=1 if worked as the CEO in a big enterprise; =0 otherwise	0.163	0.370
Up-tier government	=1 if worked in a provincial/central government; =0 otherwise	0.589	0.492
County	=1 if worked as a county governor/secretary; =0 otherwise	0.396	0.490
Firm variables			
Desulfurization_Rate	A fossil-fuel power plant's desulfurization rate (%)	22.620	40.03
Firm_Size	A fossil-fuel power plant's electricity generation capacity (10^4 KW)	19.108	41.17
Firm_Age	A fossil-fuel power plant's age (year)	21.540	19.03
Other Controls			
Roughness	Standard deviation of land gradient	19.989	12.044
Elevation	Average elevation (km)	0.521	0.593
Wind_Speed	Annual average wind speed 1990-1995 (0.1m/s)	22.831	9.494
Coldness	Percent of days with temperatures of 5° or below in one year 1990-1995(%)	0.189	0.155
Precipitation	Annual average precipitation 1990-1995 (mm)	985.578	497.783
Soil_PH	Average soil PH level of the topsoil (0-30 cm)	3.291	1.036
Dirty_Labor	Total no. of employees in dirty sector in 1996 (10,000 persons)	10.496	9.776

Note: Extremely fatal accidents are officially defined as accidents with fatality over 10. "Green" keywords include "environmental protection" (*huan bao* or *huanjing baohu*) and "pollutant emissions reduction" (*jian pai*). Data sources are described in full in section 4.1.

Table 2: SO_2 Intensive Industries

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%

Note: Data source: China Pollution Source Census 2007.

Table 3: Effect of Performance Evaluation on Industrial SO_2 Emissions, PM2.5 and Wastewater

Dependent variable	Log(SO_2)				Log(PM2.5)	Log (Wastewater)
	(1)	(2)	(3)	(4)	(5)	(6)
TCZ × Post2005	-0.234***	-0.145**	-0.154**		-0.013*	-0.030
	(0.061)	(0.062)	(0.068)		(0.007)	(0.056)
TCZ × Year2005			-0.039			
			(0.050)			
TCZ × Post2007				-0.166**		
				(0.072)		
TCZ × Year06_07				-0.093*		
				(0.048)		
City dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Control × T		Yes	Yes	Yes	Yes	Yes
Control × T^2		Yes	Yes	Yes	Yes	Yes
Control × T^3		Yes	Yes	Yes	Yes	Yes
No. of observations	3,714	3,714	3,714	3,714	2,574	3,715
R-squared	0.852	0.861	0.861	0.861	0.260	0.868
Year coverage	2001-2013	2001-2013	2001-2013	2001-2013	2001-2010	2001-2013
No. of clusters	286	286	286	286	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. All observations are at the city-level. The variable "control" denotes seven key criteria the central government used in selecting the TCZ cities. In columns 1-4, the dependent variable is the natural log of industrial SO2 emissions. Interactions of the seven key selection variables with a third-order polynomial function of time are included in columns 2-6. The PM2.5 data obtained from the MODIS and MISR Aerosol Optical Depth are substituted for industrial SO2 emissions in column 5. In column 6, industrial waste water is used as the dependent variable for placebo test. All regressions control for year fixed effects and city fixed effects. The standard errors are reported in parentheses, clustered by cities.

Table 4: Effect of Performance Evaluation on GDP and GDP Target (2001-2013)

Dependent variable	GDPTarget	GDPTarget_Gap	GDPReal	Gov_EcoIndex
	(1)	(2)	(3)	(4)
TCZ × Post2005	-0.747***	-0.554***	-0.881***	0.047**
	(0.225)	(0.208)	(0.326)	(0.021)
City dummy	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes
Control × T	Yes	Yes	Yes	Yes
Control × T^2	Yes	Yes	Yes	Yes
Control × T^3	Yes	Yes	Yes	Yes
No. of observations	3,715	3,715	3,718	3,718
Adjusted R-squared	0.613	0.475	0.467	0.523
No. of clusters	286	286	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. All observations are at the city-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. The standard errors are reported in parentheses, clustered by cities.

Table 5: Effect of Performance Evaluation on SO_2 Intensive Sector

Dependent variable	Log (1+CoalFirePlant_Capacity)		Desulfurization_Rate
	Entry	Exit	
	(1)	(2)	(3)
TCZ × Post2005	-2.012***	0.726***	3.873*
	(0.478)	(0.273)	(2.246)
TCZ			1.931
			(1.483)
Log(Firm_Size)			7.719***
			(1.280)
Log(Firm_Age)			1.931
			(1.483)
Firm dummy			Yes
City dummy	Yes	Yes	
Year dummy	Yes	Yes	Yes
Control × T	Yes	Yes	Yes
Control × T^2	Yes	Yes	Yes
Control × T^3	Yes	Yes	Yes
Year coverage	2001-2010	2001-2008,2010	2001-2010
No. of observations	2,860	2,574	22,100
Adjusted R-squared	0.210	0.260	0.695
No. of clusters	286	286	277

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. In columns 1-2, observations are at the city-level. In column 3, observations are at firm-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. In column 3, natural log of firm age and firm size are further added as control variables. The standard errors are reported in parentheses, clustered by cities.

Table 6: The Impact of Fatal Accidents on SO_2 Emissions (2001-2013)

Dependent variable	Log (SO_2)	
	Happened	Death toll
	(1)	(2)
TCZ × Post2005× Fatal_Accident	-0.154*	-0.027*
	(0.085)	(0.016)
TCZ × Post2005	-0.120*	-0.184***
	(0.063)	(0.067)
Fatal_Accident	-0.076	-0.014
	(0.055)	(0.010)
TCZ × Fatal_Accident	0.108*	0.182
	(0.064)	(0.012)
Fatal_Accident × Post2005	0.126*	0.023*
	(0.070)	(0.013)
City dummy	Yes	Yes
Year dummy	Yes	Yes
Control × T	Yes	Yes
Control × T^2	Yes	Yes
Control × T^3	Yes	Yes
No. of observations	3,714	3,714
R-squared	0.861	0.861
No. of clusters	286	286

Note: ***, **, * denote significance at 1%, ** at 5%, and * at 10%. All observations are at the city-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. The standard errors are reported in parentheses, clustered by cities.

Table 7: Political Tenure and Promotion of City Mayors and CCP Secretaries (2001-2013)

Dependent Variable	City Mayor			CCP Secretary		
	New	Log(Tenure)	Promotion	New	Log(Tenure)	Promotion
	(1)	(2)	(3)	(4)	(5)	(6)
TCZ × Post2005	0.026	-0.005	0.053	0.091	-0.065	-0.109
	(0.073)	(0.044)	(0.102)	(0.077)	(0.049)	(0.108)
City dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Control × T	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^2	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^3	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	3,718	3,718	3,718	3,718	3,718	3,718
Adjusted R-squared		0.067			0.059	
Pseudo R-squared	0.072		0.113	0.071		0.155
No. of clusters	286	286	286	286	286	286

Note: All observations are at the city-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. The standard errors are reported in parentheses, clustered by cities.

Table 8: Characteristics of Mayors and Secretaries (2001-2013)

Dependent Variable	Mayor			Secretary		
	Log(Age)	Education	Hometown	Log(Age)	Education	Hometown
	(1)	(2)	(3)	(4)	(5)	(6)
TCZ × Post2005	0.014	0.105	0.445	-0.011	0.225	-0.004
	(0.009)	(0.227)	(0.442)	(0.008)	(0.214)	(0.422)
City dummy	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes
Control × T	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^2	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^3	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	3,718	3,718	3,718	3,718	3,718	3,718
Adjusted R-squared	0.291			0.364		
Pseudo R-squared		0.365	0.563		0.356	0.540
No. of clusters	286	286	286	286	286	286

Note: All observations are at the city-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. The standard errors are reported in parentheses, clustered by cities.

Table 9: Political and Work Experience of Mayors and Secretaries (2001-2013)

Dependent variable	Mayor				Secretary			
	Log(Service Length)	As a leader in a			Log(Service Length)	As a leader in a		
		Firm	Up-tier government	County		Firm	Up-tier government	County
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
TCZ × Post2005	0.084	-0.186	-0.180	-0.069	-0.059	0.247	0.091	-0.288
	(0.115)	(0.270)	(0.191)	(0.197)	(0.097)	(0.288)	(0.203)	(0.224)
City dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year dummy	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control × T	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^2	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Control × T^3	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
No. of observations	3,718	3,718	3,718	3,718	3,718	3,718	3,718	3,718
Adjusted R-squared	0.237				0.186			
Pseudo R-squared		0.389	0.246	0.265		0.412	0.327	0.338
No. of clusters	286	286	286	286	286	286	286	286

Note: All observations are at the city-level. All regressions control for year fixed effects, city fixed effects, interactions of the seven key selection variables with a third-order polynomial function of time. The standard errors are reported in parentheses, clustered by cities.

Figure 1: Two Control Zone

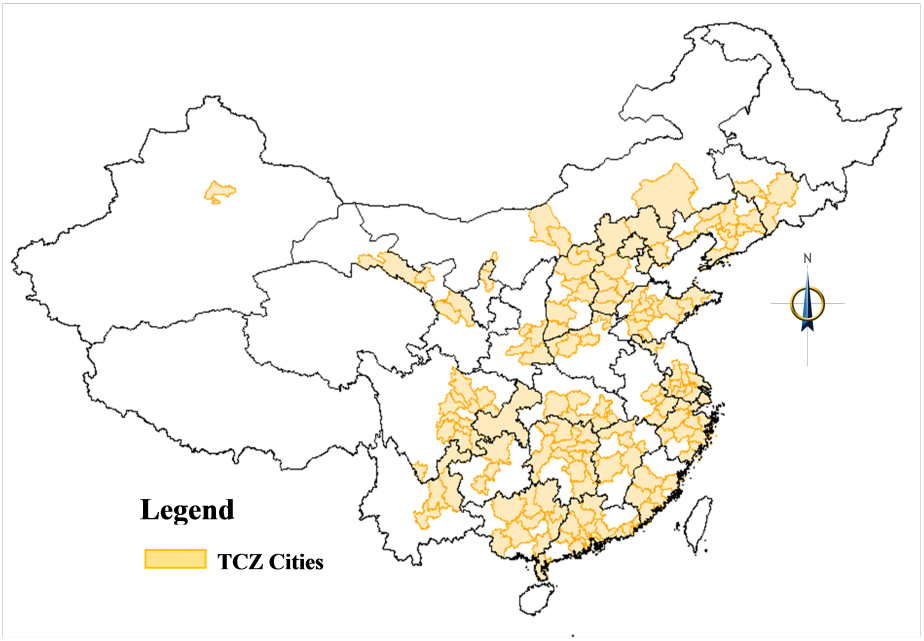
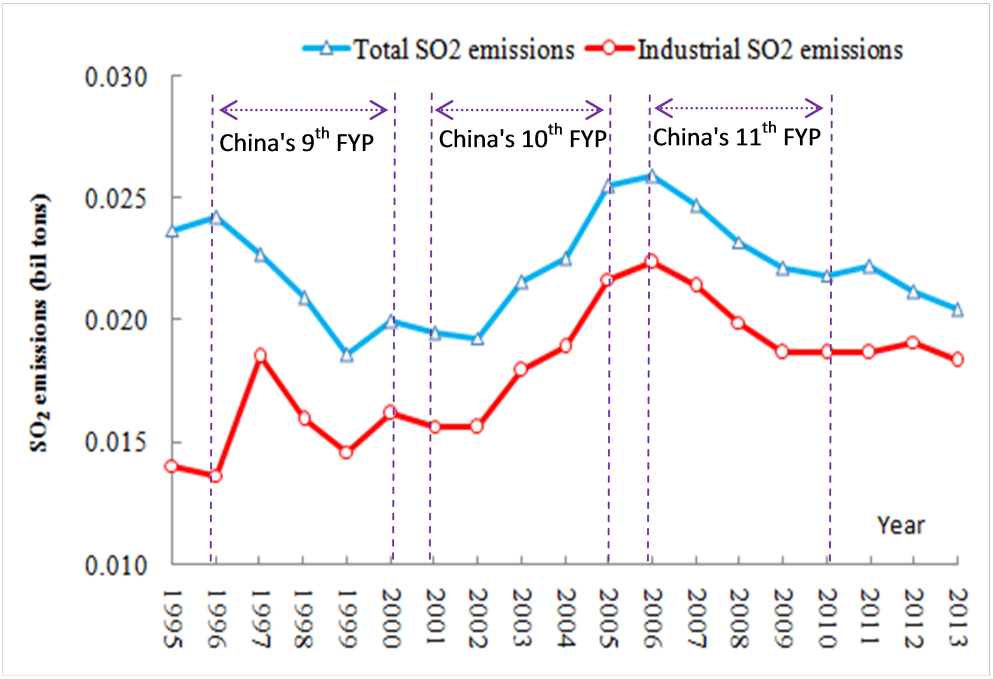
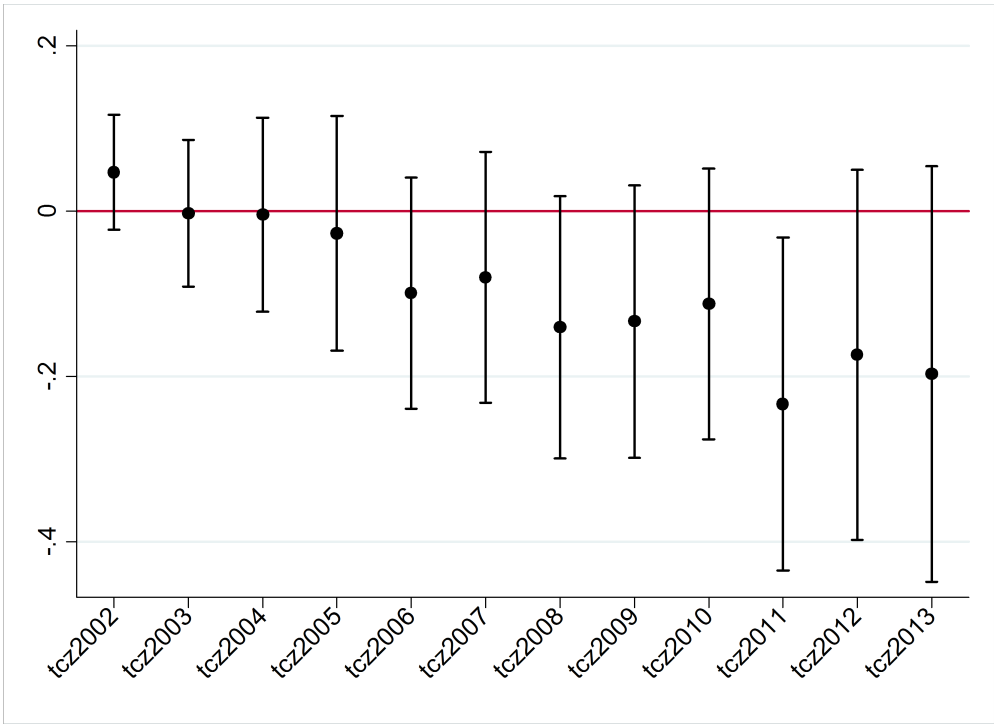


Figure 2: SO_2 Emission Trend 1995-2013



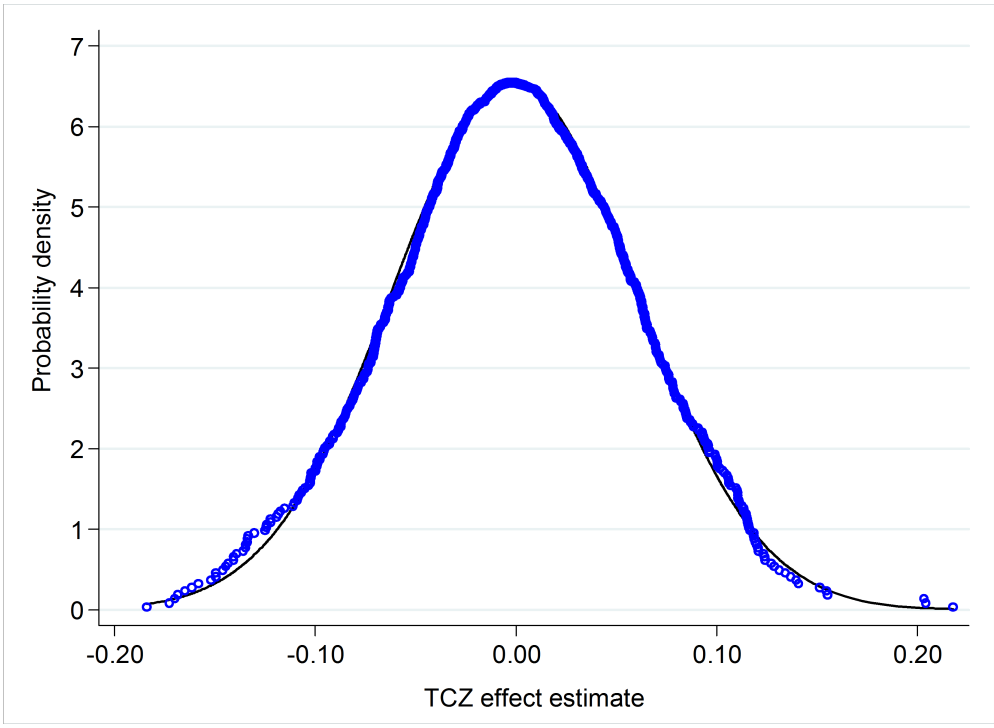
Data Source: China Statistical Yearbook (various years)

Figure 3: Estimated Coefficients of SO_2 Emissions Gap between TCZ and non-TCZ cities



Note: Each dot estimates the additional gap in emissions between TCZ and non-TCZ cities using the gap in 2001 as the baseline. All other control variables follow the same specification as column 2 in Table 3

Figure 4: Placebo Test



Note: The figure shows the cumulative distribution of the estimated coefficients from 1,000 simulations. The distribution is generated according to the placebo test discussed in section 4.2