Hybrid Contracts, Multitasking, and Incentives:

1 Introduction

Multitasking agency problems have been a well-studied topic in economics (Garicano and Rayo, 2016). Pioneering works, exemplified by Holmstrom and Milgrom (1991), Baker (1992), and Hart et al. (1997), have shown how to design linear incentive contracts in the presence of multitasking. However, linear incentive contracts are not the only way to address multitasking problems. In reality, we often observe hybrid contracts which require a minimum level of performance for one task while incentivizing the pursuit of high performance for the other.[1] For instance, in research universities, where tenure evaluation hinges on academic research performance, candidates are also required to meet minimum teaching standards (e.g., a teaching load of no less than some credits with a "tolerable" teaching assessment). Similarly, subcontracts in both business and public sectors usually set explicit prices for supplied goods or services (thereby creating strong incentives for cost reduction), yet a minimum quality standard is also enforced. In these settings, agents face a high-powered incentive for one task (e.g., academic research or cost reduction) while they also need to fulfill a minimum standard for another (e.g., teaching or product/service quality). Despite the widespread use of such hybrid incentive contracts in the real world, there remains a gap in our understanding of agents' strategic responses to them, particularly when the minimum performance standard is costly to meet and the status of the standard fulfillment could be revealed over time.[2]

In this paper, we aim to investigate how agents respond to hybrid incentive contracts with minimum standards in a multitasking context, using China's environmental regulation as a case study. China provides an excellent context for our analysis for two reasons. First, China has long been implementing a kind of hybrid incentive scheme for its local officials who are held accountable for certain minimum targets, such as pollution controls, and in the meantime seek to achieve high performances of GDP growth and fiscal revenues which help enhance their chance of promotion. Second, in the past decade, China has managed to achieve a significant reduction in air pollution while maintaining a reasonably high economic growth rate. According to a recent study (Greenstone et al., 2021), China reduced its air pollution by 40% during 2013-2020, which contributed to over three-quarters of the global decline in pollution. In the meantime, it achieved an economic growth rate of about 6.5.

The presence of hybrid contracts in China's environmental regulation poses interesting research questions. The minimum targets for air pollution controls require Chinese local officials to achieve a minimum number of days with good or moderate air quality within a year. Failures to meet them would result in severe punishment although over-fulfilling them yields no extra rewards. To what extent do the minimum targets influence local officials' efforts to control air pollution? What happens if they achieve the minimum targets ahead of time in a

[1] Existing literature has examined target-based non-linear contracts as well, but primarily focused on a single-task context. As an example, see Finan et al. (2017) for a review of studies on public sectors.

[2] Most previous studies have mainly focused on evaluating the overall effects of installing performance-based contracts (sometimes mentioning hybrid contracts in background information) without closely scrutinizing the interactions between minimum standards and high-powered incentives. See Heinrich (2002) for a review of studies on public sectors and Selviaridis and Wynstra (2015) for a review of more recent work.

given year or if they realize there is no way to fulfill the minimum targets for the rest of the year? How does the pressure to pursue economic growth (e.g., the task with high-powered incentives) affect officials' strategic responses to different completion statuses of the minimum targets?

However, it is challenging to empirically address these questions since the top-down assignment of the minimum air quality targets to specific regions could be endogenous. It is also rarely feasible to track the completion statuses of targets in real-time, not to mention agents' dynamic re-optimization over multitasking efforts. Our paper exploits a unique feature of China's air quality targets and employs a regression discontinuity design (RDD) to tackle potential identification challenges. As mentioned above, the minimum targets typically specify a minimum number of days (e.g., 250 days) in a year with good or moderate air quality (measured by the *Air Quality Index*, *AQI* for short). This cut-off feature of targets enables us to define two specific statuses of target completion, *doomed failure* and *early fulfillment*. The first status refers to a situation where no matter what to do during the rest time of the year, the target cannot be fulfilled, while the second refers to a situation where the target is fulfilled before the end of the year. Since daily AQI performance data are publicly released, it is fairly straightforward to identify the exact date when a prefectural city hits either of the two completion statuses of minimum targets. Then we can easily observe the immediate and potentially discontinuous responses of local officials when they find out the exact completion status.

Another challenge in investigating local officials' responses to minimum air quality targets is data manipulation by local officials. Before 2012, Chinese local officials were notoriously known for manipulating air quality data and cheating their superiors (Chen et al., 2012). Yet beginning in 2012, the Chinese central government mandated the nationwide installation of a real-time automatic monitoring system and issued a full battery of policies to crack down on data-manipulating behaviors (Greenstone et al., 2022). We also conduct a formal test of the credibility of the AQI data during our sample period and find no evidence of data manipulation. This test is critical for our RDD identification: As long as local officials cannot perfectly manipulate air pollution data, there should exist randomness in threshold dates associated with target completion statuses; thus, a discontinuous change in air quality performance or firm emissions surrounding those threshold dates would render us a causal interpretation of the observed pattern.

Our study exploits a novel dataset containing prefecture-day level air quality from January 1, 2015 to December 31, 2019, air quality targets assigned for prefectural cities as well as detailed information on incumbent prefectural leaders. To further provide micro-foundation evidence, we also obtain firm-day level SO_2 emissions from the Continuous Emission Monitoring System. We define the running variable of our RDD estimation as the time relative to the threshold date of doomed failure or early fulfillment. The baseline results show that a prefecture's air quality performance tends to improve significantly after a doomed failure occurs, which indicates that after finding out it is impossible to meet the minimum target, the local bureaucrat does not give up completely and continues to exert efforts to protect the environment. This result is consistent with our finding that the likelihood that an official will be punished increases with the severity of the failure. However, when it is clear that the minimum target has been met early, the air quality deteriorates immediately afterward. If we further take the pressures of pursuing economic growth into consideration, it is shown that when local officials are under less pressure to promote local economic growth, the improvement in air quality performance after a doomed failure is strengthened, and the deterioration in air quality performance after early fulfillment is weakened. We also provide evidence that firm-day SO_2 emissions exhibit similar patterns around the threshold dates, which indicates a potential channel that local officials could pass on their pressure of fulfilling AQI targets to local firms through a set of temporary regulatory measures (see more details in Section 2.3).

To deepen our understanding of the above empirical findings, we construct an illustrative model that formally specifies the mechanisms under which hybrid contracts function in a multitasking setting. We introduce intrinsic motivations for local environmental quality and investigate the interactions between hybrid contracts (which impose the extrinsic motivations) and intrinsic motivations.[3] Our model rationalizes the empirical patterns documented above and yields new predictions regarding the roles played by intrinsic motivations, political incentives, and central monitoring. First, intrinsic motivations could mitigate the incentives of strategic responses in either target completion status, and the asymmetric responses to the doomed failure and early fulfillment of minimum targets reported above are mainly driven by those with low intrinsic motivations. Second, for officials who have low intrinsic motivations to preserve local air quality, stronger career concerns make them more sensitive to hybrid contracts than those with weaker career concerns. Third, stricter central government supervision over environmental protection could enhance the improvement in air quality performance after doomed failure and restrain the deterioration in air quality performance after early fulfillment. We put these predictions to the data and find consistent evidence.

Our research directly speaks to three strands of literature. First, by embedding hybrid contracts into a multitasking agency model, this paper closely relates to the literature on both multitasking agency problems (Holmstrom and Milgrom, 1991; Baker, 1992; Hart et al., 1997) and target-based incentive contracts, especially in public sectors (Heckman et al., 1997; Courty and Marschke, 2004; Prendergast, 2007; Heckman et al., 2011; Newman and Azevedo, 2013). The multitasking literature focuses on linear incentive contracts and highlights the distortions caused by the combination of strong incentives for some easily measurable tasks and weak or no incentives for other poorly measurable ones. In order to correct distortions, weakening the incentives for *all* tasks is often suggested. Our study contributes to this line of literature by showing how hybrid incentive contracts with minimum targets serve as a specific incentive scheme to motivate Chinese local officials to strike a balance between air pollution controls and economic development.

The literature on target-based incentive contracts mainly focuses on a single-task environment (Finan et al., 2017). The existing scholarship has documented a lot of evidence on

[3]In this paper, we use the term "intrinsic motivations" in contrast with extrinsic motivations (more precisely local officials' political or career incentives). We are agnostic about the sources of intrinsic motivations for environmental protection, which might derive from inherent preferences or affinities driven by local connections (Persson and Zhuravskaya, 2016).

agents' gaming incentives when performance targets are narrowly measured or tied to a specific timetable. In a recent study on China's environmental regulation of water pollution, He et al. (2020) presents evidence that local officials enforced tighter regulation on polluters immediately upstream of water monitoring stations than on those immediately downstream since the stations only measure emissions from upstream. Several other studies have detected evidence that local officials manipulate data to "game" air pollution targets imposed by the central government (Chen et al., 2012; Greenstone et al., 2022). We also find evidence for local officials' shirking behavior in air pollution controls right after the early fulfillment of targets, but from the viewpoint of hybrid contracts, this strategic response is a natural response to the minimum target which is designed and intended by the principal, rather than a "gaming" behavior or an unintended consequence of the incentive design. Furthermore, our analysis suggests that hybrid contracts will guide local officials to make trade-offs between economic growth and environmental protection in accordance with local economic and environmental conditions.

Second, our study also adds significantly to a large and growing literature on the political economy of environmental regulation, especially in the context of China (List and Sturm, 2006; Burgess et al., 2012; Chen et al., 2012; Zheng et al., 2014; Kahn et al., 2015; Jia, 2017; Karplus and Wu, 2020; He et al., 2020; Greenstone et al., 2022).[4] The study closest to ours is Chen et al. (2018a). They use a difference-in-differences identification strategy and find evidence that Chinese prefectural cities located in the Two-Control Zone, an area subject to stricter environmental regulation, reduced their pollution emissions at the cost of local economic growth. While we share a common theme about multitasking local officials' responses to environmental regulation in China, our study differs from theirs in three important ways. First, in contrast to their focus on the overall effect of environmental regulation on pollution emissions and economic growth, our paper explicitly analyzes strategic responses of local officials to completion statuses of minimum targets and interactions between minimum targets and high-powered incentives in hybrid contracts. Second, we focus on the effects of the AQI-based minimum targets, which have recently replaced the Two-Control Zone policy as currently the major regulatory scheme of air quality controls over the last decade. Third, our analysis introduces intrinsic motivations into the context of environmental regulation and highlights their important role in shaping Chinese local officials' incentives.

Finally, our paper is closely related to the economics literature on interactions between intrinsic and extrinsic motivations. Although a lot of existing theoretical studies have discussed the effect of intrinsic motivations (Murdock, 2002; Bénabou and Tirole, 2003, 2006, 2016b; Besley and Ghatak, 2005, 2006; Prendergast, 2007), empirical evidence is rare and mostly comes from experimental studies (Gneezy and Rustichini, 2000a,b; Ashraf et al., 2014). If we interpret local affinities as intrinsic motivations and political incentives as extrinsic motivations, our paper provides real-context evidence of how interactions between intrinsic and extrinsic motivations affect local officials' strategic reactions to hybrid incentive contracts with minimum targets. Consistent with theoretical predictions, we show supportive evidence that

[4] See Zheng and Kahn (2013) for an extensive review of the literature on the political economy of China's urban environmental regulation.

non-outsiders' intrinsic motivations help remedy the lack of incentives to exceed the minimum targets if they are fulfilled early.

The remainder of the paper is organized as follows. Section 2 delves into the institutional background on China's hybrid incentive contracts, particularly the minimum targets in air pollution controls. It also discusses the contingency measures available for local governments to control air pollution, along with the central government's recent efforts to crack down the manipulation of air pollution data. Section 3 introduces our data and sample construction and presents some supportive evidence of punishment for failing to achieve minimum targets. Section 4 describes our empirical identification strategy, RDD, and presents the baseline empirical results. Section 5 builds a simple theoretical model to pin down the mechanisms under which hybrid contracts interact with intrinsic motivations. Section 6 carries out additional empirical tests derived from our model. Section 7 briefly concludes.

2 Institutional Background

2.1 Hybrid Incentive Contracts in the Chinese Government

Chinese local governments have played important roles in various aspects of economic growth and social development (Xu, 2011; Zheng and Kahn, 2013; Karplus et al., 2021). Since economic growth has traditionally dominated Chinese local officials' performance evaluations (Li and Zhou, 2005; Xu, 2011; Yao and Zhang, 2015), some other government tasks, such as environmental protection, received little attention from local officials. In 2006, as a response to deteriorating ecological conditions and increasing urban-rural disparities, hybrid contracts were introduced in the five-year plan for the first time.[5]

Specifically, the central government of China simultaneously set *anticipatory* and *obligatory* targets to local governments, which correspond to different incentive schemes. Anticipatory targets (e.g., growth targets of GDP and fiscal revenues) serve as guiding signals with no binding power but represent the key concerns of local officials, whose promotion is closely linked to the performance of these indicators (Li et al., 2019). By contrast, obligatory targets, such as emission reduction and poverty alleviation, set bottom lines and binding constraints. If local officials fail to fulfill their obligatory targets, they lose their chance of promotion or are even removed from their posts during the evaluation year. But over-fulfilling the obligatory indicators receives no extra reward. This dual-target scheme in China's context fits perfectly into the hybrid incentive contracts we previously described. Anticipatory targets correspond to *high-powered incentives* and could be summarized as "the higher, the better," while obligatory targets serve as *minimum standards* and could be described as "hitting the threshold but no more."

[5]China's multiple levels of government have regularly set top-down targets since the planned economy period. During the reform era, mandatory targets were replaced by (non-binding) guiding targets, and those related to economic and social development were regularly released in *Five-Year Plans* and annual *Reports on the Work of Government* by multiple levels of government. See Kennedy and Johnson (2016) and Li et al. (2019) for more detailed descriptions of China's target setting scheme.

While environmental protection has been introduced into the obligatory indicators for local officials since 2005, the specific content of the targets has changed over time. From 2005 to 2010, China's regulatory agencies focused on indicators such as energy consumption per GDP or CO_2 emissions. As time went on, air pollution controls gradually gained attention from China's government, domestic citizens, and international communities. The sudden burst of haze in 2013, the most severe air pollution in a half-century, became a trigger to push the Chinese government to shift the focus of environmental regulation from emission intensity to the ambient concentration of pollutants.[6]

2.2 Minimum Standards in Air Pollution Controls

When it comes to the specific design of air pollution control targets, AQI has been adopted as a contractible indicator since 2016. As stated in the *Ambient Air Quality Standards (GB 3095-2012)*, AQI is a composite index constructed from the concentration of six major pollutants: sulfur dioxide (SO_2), nitrogen dioxide (NO_2), carbon monoxide (CO), ozone (O_3), fine particulate matter ($PM_{2.5}$), and inhalable particulate matter (PM_{10}). AQI is first computed hourly at the monitoring station level according to a deterministic function and then averaged on the prefecture-day basis to depict a city's overall daily air quality.[7] It takes values in the range of [0,500] and air quality grades are defined as shown in the first two columns below:

AQI Value Range	Air Quality Grade	NAQI Value Range
0∼50	Good	450∼500
51∼100	Moderate	400∼449
101∼150	Lightly Polluted	350∼399
151∼200	Moderately Polluted	300∼349
201∼300	Heavily Polluted	200∼299
300∼500	Severely Polluted	0∼200

The AQI target is typically specified as *a minimum number of days with good or moderate air quality in a year*, or equivalently, *a minimum number of days with* $AQI \leq 100$. Monitoring stations publicly release hourly readings of AQI and pollutant concentrations. Since the total number of days in a calendar year is fixed (either 365 or 366), it is straightforward to monitor the target completion status on a daily basis. Each day, officials can observe whether their city is likely to demonstrate one of two distinctive states of target fulfillment: (1) doomed failure – the number of days remaining in the year is insufficient to meet the target even if air quality remains good or moderate for the rest of the year or (2) early fulfillment – the AQI target is met before the end of the year regardless of air quality readings in the remaining days of the year.

As will be shown in our subsequent analysis, the continuous and punctual detection of air quality and target completion status enables us to employ an RDD to resolve identification issues. Previous studies usually assume that the agent cannot observe whether the performance target is met or not until all of their efforts have been exerted. In our case, local officials can

[6] See Karplus et al. (2021) for a detailed review.

[7] See *Technical Regulation for Ambient Air Quality Index (on trial) (HJ 633-2012)* for a detailed description of the functional relationship between AQI and concentration of the six pollutants.

continuously observe the target fulfillment status and decide whether to continue to exert efforts (and how much), which implies that there could be a discontinuous change in incentives at the time an official learns that the completion status is early fulfillment or doomed failure.

It should be noted that punishments are not imposed indiscriminately on any degree of failure to meet obligatory targets. A barely fulfilled case may be tolerated, but a severe failure may trigger a serious warning or punishment. In other words, in the case of failure, local officials' punishment is based on the gap between the target and actual performance. We traced cases of warnings or criticism by the Ministry of Ecology and Environment (MEE)[8] or superior governments in official documentation and online reports, and found that for serious warnings or criticism, such as publicizing the names of local officials who failed to fulfill the targets on the government websites or circulated documents, the stated reasons for criticism or punishment usually involved expressions such as "severely lagging behind," "failing by a large margin," or "ranked last among all cities."[9] In Section 3.3, we will provide more supportive evidence of the correlation between the likelihood of punishment and the severity of the failure to meet the obligatory target. The dependence of punishment on the severity of the failure in air pollution controls implies that local officials cannot simply give up completely in the case of doomed failure; instead, they have to demonstrate further efforts in order to avoid a serious failure.

2.3 Contingency Measures Available for Local Governments

On top of the centrally-imposed air pollution control measures (e.g., tightening vehicle emission standards nationally), Chinese local governments have certain discretion to affect air pollution levels in their jurisdictions, even in a short time window.[10] The policy measures taken by local governments can be roughly divided into two categories. The first one relates to reinforced cleaning, such as rainfall enhancement or increased cleaning frequency by sanitation workers. The second and more important batch of policies points to regulations on air pollution sources. Specifically, on the household side, local governments could tighten their restrictions on polluting behaviors (e.g., straw burning) and encourage eco-friendly lifestyles (e.g., driving less). As for firms, the contingency plans authorize relevant government departments to directly impose production restrictions on high-emission plants. In extreme circumstances, local governments might even directly close factories or shut down construction sites to meet air pollution control targets.[11]

---

[8]The MEE was established after the reconstruction of the ministries under the State Council in 2018. The previous Ministry of Environment Protection was merged into this new ministry. We refer to both ministries as the MEE.

[9]For example, there was a news report that the MEE summoned prefectural leaders of three prefectures for "ranking the last three among cities under special monitoring." See https://www.sohu.com/a/230340946_120802 (accessed on June 7, 2022).

[10]See, for example, the emergency response plans released by Heilongjiang Province on https://www.hlj.gov.cn/n200/2021/0315/c668-11015508.html (accessed on June 7, 2022).

[11]For example, many local governments suspended production at power plants in order to alleviate air pollution for the upcoming 2008 Olympics Game (Chen et al., 2013). A recent example is the Linyi City of Shandong Province which, after being publicly criticized by the MEE in 2015, directly cut off the electricity supply to some manufacturing enterprises, which resulted in huge losses and irreversible damage to production equipment. See the detailed public report on https://www.thepaper.cn/newsDetail_forward_1347676 as an example (accessed

The regulation measures described above provide potential channels through which local governments and bureaucrats could immediately respond to the completion statuses of AQI targets. Motivated by this fact, we will conduct analysis in Section 4.3.3 to look at the firm emissions surrounding the threshold dates of different target completion statuses, which should indirectly reflect the changes of temporary regulatory measures taken by local officials to fulfill their AQI targets.

While local governments enjoy some discretionary power in taking emergency regulatory measures, they could affect local air quality only in a limited way. The air quality is determined by a set of factors some of which are outside the reach of local governments, such as photochemical reactions, weather conditions, and air pollution from neighboring regions. The impossibility of precise control of AQI guarantees the validity of the RDD model, which requires that agents could not precisely manipulate the running variable.[12]

2.4 Crackdown on Data Manipulation

It has been well documented that China's air pollution data have been subject to widespread fraud (Chen et al., 2012; Ghanem and Zhang, 2014; Greenstone et al., 2022). This fraud was enabled by the fact that local governments managed their own air-quality monitoring devices and self-reported air-quality data to the central government. The situation has changed dramatically since the administrative and legislative branches have devoted considerable efforts to crack down on data manipulation over the last several years in three main ways.

First, the MEE upgraded its monitoring technology and strengthened its policy enforcement to improve data quality. It introduced an automatic monitoring system from 2013 to 2015, which facilitated the real-time sharing of data with the central government and follow-up quality assurance through statistical tools. Greenstone et al. (2022) observed a sudden jump in pollutant concentration after the adoption of the automatic monitoring technology, and interpreted it as a signal of improvement in the quality of air pollution data. Besides the introduction of new data-collection technology, the MEE also managed to take over the 1,436 air quality monitoring stations from local governments by the end of 2016. Second, the MEE supplemented these technological and managerial upgrades with stricter surveillance. For instance, it regularly dispatched the Central Environmental Inspection Team to provinces to supervise local law enforcement and detect any signs of data manipulation. Third, the National People's Congress revised the *Environmental Protection Law* in 2014 by adding Article 17, which states that "monitoring institutions and the persons in charge thereof shall be responsible for the truth and accuracy of monitoring data."

Given the above measures taken by the central government, data manipulation in our sample period of 2015–2019 should not be a key concern. However, as a double check, we would provide direct empirical evidence in Section 3.3 to corroborate the reliability of the air pollution data in our sample.

on June 7, 2022).

[12] In the context of this article, it translates into the requirement that local governments should not be able to determine the specific threshold dates, which will be defined in Section 4.

3 Data and Summary Statistics

3.1 Data Collection

The data used in this paper were mainly collected from the following sources.

Prefectural Targets for Air Pollution Controls. We first collected prefecture-level, AQI-based targets from the publicly available documents issued by provincial governments. If related documents were not accessible on the Internet, we submitted applications for information disclosure to the relevant government departments such as the general office of the provincial government or the provincial Department of Ecology and Environment.[13] These targets are most commonly expressed as a minimum number of days with AQI $\leq$ 100 in a year; alternative expressions include the relative increase in the number of days with AQI $\leq$ 100 or scheduled improvement of air quality in multi-year contracts. To facilitate comparison, we standardized these targets into a common expression based on the rules described in Table B1.

AQI and Normalized Version. The AQI indices employed in this paper were compiled by the China Air Quality Online Monitoring and Analysis Platform, which utilizes a web crawler to fetch historical air quality records published by the MEE and releases them for free.[14] We fill in missing values resulting from temporary failures of the web crawler with data from another independent source that copies MEE data using similar technology.[15] We further define normalized AQI (NAQI) as 500$-$AQI. This normalization helps facilitate the interpretation of empirical results later, because a larger NAQI represents a higher level of air quality.[16]

Meteorological Conditions. Weather data were acquired from a public website that posts prefecture-day meteorological conditions,[17] including maximum and minimum temperatures, a category variable for wind speed, and two dummy variables for rainfall and snow, respectively.

Information of Prefectural Party Secretaries. We digitized the resumes of prefectural party secretaries from media reports and other related sources such as Baidu for cross-validation. We first extracted the dates they assumed office to confirm which party secretary was in charge on a particular day. We also constructed a series of variables to capture differences in bureaucrats' political incentives and personal closeness to the jurisdictions they governed, including the age at the time of their appointment, a proxy for the intensity of career concerns (Wang et al., 2020), a dummy for whether an incumbent has worked in the same prefecture before, and a dummy for whether the incumbent was locally promoted.

[13] We are required to properly preserve the documents and data acquired through application procedures and cannot circulate them without the permission from relevant government departments. Therefore, we cannot disclose the target numbers. However, local governments are obligated to reply to applications for information from Chinese citizens, so readers who are interested in the original data could directly submit applications for data release to provincial governments or departments of ecology and environment.

[14] The website is `https://www.aqistudy.cn/historydata/` (accessed on June 7, 2022).

[15] The website is `https://quotsoft.net/air/` (accessed on June 7, 2022).

[16] The original expression of AQI targets, i.e., *a minimum number of days with AQI* $\leq$ *100* will be changed to *a minimum number of days with NAQI* $\geq$ *400*.

[17] The website is `http://www.tianqihoubao.com/lishi/` (accessed on June 7, 2022).

Firm Information and SO_2 Emissions. To provide a micro-foundation for the empirical results derived from aggregate indices at the prefecture level, we obtained data on firm-day SO_2 emissions from the Continuous Emission Monitoring System (CEMS).[18] The MEE has made public a firm list and required that those listed firms should install the system and upload the real-time emissions of pollutants on websites operated by provincial governments. We first collected the firm list from publicly available documents and then fetched the emission data from provincial websites. To facilitate heterogeneity analysis among firms, we further obtained information on ownership types, employment size, and industries of the listed firms from Tianyancha.

AQI Targets and Completion. We first document stylized facts about AQI targets at the prefecture-year level. Figure 1 plots the distribution of AQI targets and completion rates along with estimated kernel density. Completion rates are computed at the prefecture-year level and equal to the actual number of days with AQI $\leq$ 100 divided by the targeted number. Rates equal to or larger than 100 indicate target fulfillment. Figure 2 plots the geographic distribution of completion status by year on a map, indicating prefectures that fail (streak) or succeed (color) to meet their targets. Figure 1 Panel B and Figure 2 illustrate that target achievement varies considerably in both cross-sectional and time dimensions. This implies that the AQI targets assigned by most provinces are both challenging and attainable.[20]

Credibility of AQI Data. Figure 1 Panel B also shows that prefectural completion rates ranged from 47–133%, with considerable dispersion but bunching to the right of 100%, which denotes exact fulfillment. An intuitive explanation for this bunching is that local bureaucrats falsified the air quality data when they failed to meet their targets. To rule out the possibility of data manipulation, we therefore plot the McCrary (2008) test of our AQI data in Panel A of Figure 3. Panel B extracts the same test from Chen et al. (2012), which examines Chinese local officials' widespread manipulation of air quality data during 2000–2009. Chen et al. (2012) interpreted the discontinuity around 100 in Panel B as manipulation of the API.[21] However, Panel A demonstrates that our AQI data exhibit no significant jump around 100. This is not surprising because, as introduced in Section 2.4, the central government has implemented a battery of policies to crack down on data manipulation and imposed severe punishments for this misbehavior. If data manipulation is not a driving force behind the bunching around 100%, then it may be caused by a lack of incentives for over-fulfillment if the target is met early. We explore this explanation further in Section 4.3.

Punishment on Failing AQI Targets. For AQI targets to serve as an effective minimum performance standard, failing to meet them must incur punishment, and the probability or severity of punishment should increase with the severity of the failure. That is, the minimum standard should specify the following probability of being punished:

$$P(y) = \begin{cases} 0 & , \quad \text{if } y \geq y_0 \\ p(y_0 - y), & \text{if } y < y_0 \end{cases}$$

where the increasing function, $p(\cdot)$, is the probability of being punished when the threshold is not met. Figure 4 and Table 2 provide supportive evidence for this scheme. We focus on two main forms of punishment implemented by MEE. The first is a face-to-face admonition (*yuetan*), in which the MEE summoned prefectural leaders for a face-to-face criticism of their failures in environmental protection. Second, since 2016, the MEE began regularly dispatching Central Environmental Inspection Team to provinces (*ducha*) and to "name and shame" specific

[20] We have 1,154 prefecture-year observations for which targets and actual performance are both available; 669 are fulfillment year, which means that the average rate of completion is 59%.

[21] API was calculated from the concentration of SO_2, NO_2 and PM_{10} according to *Ambient Air Quality Standards (GB 3095-1996)*, which is the former version of AQI and has been replaced by AQI since 2012.

prefectures for failures (*dianming*).[22] We collected public information about those two kinds of supervision, and construct a prefecture-year level dummy that denotes whether prefectural leaders were summoned for an administrative talk or name-picked during inspection in the next year. Figure 4 displays the relationship between performance in AQI targets and the probability of being punished. The prefecture-year differences between the actual percentage of days with AQI $\leq$ 100 and the targeted percentage are displayed on the horizontal axis; values equal to or larger than 0 indicate fulfillment. It is noteworthy that the fitted line in the fulfillment region is quite flat and remains close to zero,[23] while the fitted line in the failure region is much steeper and the probability of being punished increases with the failure margin.

To further account for potential confounding factors, Table 2 presents regression-based evidence. Based on the severity of failures compared with the AQI targets, we split the failure cases into two categories at the median value of failure margin: moderate or severe failures. Using fulfillment cases as the control group, the empirical results show that if local officials fail by a margin above the median, the probability of being punished nearly doubles (i.e., 0.116 versus 0.067). In summary, we find some supportive evidence of the association between a failure to achieve the AQI targets and the likelihood of punishment.

4 Econometric Strategy and Baseline Results

The two components of hybrid contracts function in distinct ways. The high-powered incentive component (e.g., the pursuit of economic growth) remains effective until the end of the year, while the cut-off feature of minimum air quality targets results in local officials' discontinuous responses. To facilitate our identification, we will implement the RDD estimation and first focus on how local officials respond to minimum air quality targets. Subsequently, we will consider the pressures these officials face in pursuing economic growth, which enables us to investigate the role of hybrid incentive contracts in shaping officials' behaviors in a multitasking context.

4.1 Definition of Threshold Date in RDD

The RDD we employ is a modified version of the classical RDD that has been increasingly popular in environmental economics due to the availability of high-frequency data (Hausman and Rapson, 2018). The running variable is time relative to the threshold date. More specifically, for prefecture c and year y, date d would be defined as the threshold in RDD if either of the following two conditions have held since date d:

$$(\text{\# of days with AQI } \leq 100 \text{ until } d-1) + (\text{\# of days remaining since } d) < \textit{Target},$$
$$\text{or, } (\text{\# of days with AQI } \leq 100 \text{ until } d-1) \geq \textit{Target}.$$

[22] Previous studies have closely examined those policies. See Shi et al. (2017) for a review of the first and Karplus and Wu (2020) for the second.

[23] The probability is not exactly equal to zero because the AQI target is only one of the environmental protection targets; those who succeed in this target could still fail in other dimensions of environmental protection (e.g., curbing water pollution) and be punished.

"*Target*" denotes "the target number of days with AQI $\leq$ 100" and "# of days remaining" is computed by subtracting the number of days that have passed (i.e., $d-1$) from the total number of days in the year (365 or 366). The first inequality indicates the status of *doomed failure*: prefecture c performs so poorly from January 1^{st} to date $d-1$ that it is impossible to meet the target in year y even if AQI $\leq$ 100 in all remaining days in the year. The second inequality indicates *early fulfillment*: prefecture c meets the target ahead of schedule on date $d-1$ regardless of what happens to the air quality in all remaining days in the year.[24]

Based on the rule introduced above, for prefecture c and year y, local bureaucrats would exert efforts to improve the air quality from January 1^{st} until a sudden change in incentives on the threshold date d, after which the target completion status shifts to either *doomed failure* or *early fulfillment*. To examine the different responses to different completion statuses of minimum targets of air pollution controls, we categorize the data into two subsamples at the prefecture-year level according to whether the target assigned to prefecture c in year y has been met; we call these the *failure year* and *fulfillment year* subsamples, respectively.

4.2 Estimation Procedure

Following the empirical design suggested by Hausman and Rapson (2018) and Li et al. (2020), we implement RDD analyses in two steps. First, we estimate the following equation using the full sample:

$$NAQI_{cyd} = a + m_c + l_d + g' X_{cyd} + z_{cyd}. \tag{1}$$

We then use the residual $NAQI\_Res_{cyd}$ obtained from equation (1) as the outcome variable for a subsample RDD estimation as follows:

$$NAQI\_Res_{cyd} = p + bD_{cyd} + f(t_{cyd}) + D_{cyd} \times f(t_{cyd}) + e_{cyd}. \tag{2}$$

The main idea behind the two-step procedure is that we first regress $NAQI$ against prefecture and calendar date fixed effects (denoted by m_c and l_d) along with a series of controls (X_{cyd}) to purify the outcome variable. Since the relationship between $NAQI$ and regressors holds for all observations, we could utilize the whole sample in this step and make full use of the dataset in hand, regardless of whether the prefecture receives a target from the provincial government in that year. We then use the less noisy outcome variable (i.e., the residual) for subsequent RDD analyses, which can be free of complex specifications and focused on estimating the discontinuous changes in air quality around the threshold date.

Subscripts c, y and d denote prefecture c, date d and year y, respectively. $NAQI\_Res_{cyd}$ represents residuals obtained from equation (1). X_{cyd} consists of a series of control variables and

[24]For example, a prefecture with a target of 250 days with AQI $\leq$ 100 met that standard on 199 of the first 314 days of the year. If the air quality on the 315^{th} day was still bad, the prefecture would be faced with a situation of doomed failure starting from the 316^{th} day until the end of the year. That is, even if air quality remained good or moderate in all the remaining 50 days (365 – 315 = 50), the total number of days up to standard would be 199 + 50 = 249, which would fail to meet the target. Alternatively, assume the prefecture already had 249 days up to standard during the first 314 days of the year. If air quality on the 315^{th} day was also good or moderate (thus 250 in total), the prefectural government would face early fulfillment starting from the 316^{th} day, because the target has already been achieved; it does not matter if all remaining 50 days are heavily polluted.

fixed effects. First, since weather conditions have profound effects on air pollution, we add the following variables to control for *daily weather conditions*: maximum and minimum temperatures, rainfall indicator, snow indicator, and a categorical variable for wind speed coded 1–4. Second, since cities may exhibit distinctive yearly and seasonal patterns of air quality, we interact prefecture fixed effects with the following dummies to address *prefecture-specific seasonality*: $m_c \times Day\ of\ Week_d$, where $Day\ of\ Week_d$ = {Monday, Tuesday, ..., Sunday}; $m_c \times Week\ of\ Month_d$, where $Week\ of\ Month_d$ = {Week1, ..., Week6}; $m_c \times Month\ of\ Year_d$, where $Month\ of\ Year_d$ = {January, ..., December} and $m_c \times Year_d$, where $Year_d$ = {2015, 2016, 2017, 2018, 2019}. The third component of X_{cyd} is leader fixed effects of incumbent prefectural party secretaries. The personal characteristics of Chinese local leaders, such as abilities, motivations, and expertise, affect regional development (Yao and Zhang, 2015) as well as air pollution controls. Since city party secretaries often move laterally between prefectures, we are able to assign a unique ID to a party secretary who may have worked for multiple prefectures. Including person fixed effects helps rule out the unobserved heterogeneity for prefectural leaders.

t_{cyd} is the running variable, and a negative (positive) value denotes the number of relative days before (after) the threshold date. D_{cyd} is a dummy variable for completion status that takes a value of 1 if $t_{cyd} \geq 0$ and 0 otherwise. $f(t_{cyd})$ is polynomial functions of the running variable; its interaction with D_{cyd} further allows for distinctive trends of the outcome variable on different sides of the threshold. In the baseline setting, bandwidths are 35 and 15 for failure year and fulfillment year subsamples, and the functional form of $f(t_{cyd})$ is specified as locally linear with a uniform kernel. Standard errors are clustered at the prefecture level. We would also demonstrate the robustness of our baseline estimates to alternative bandwidths, kernel functional forms, and heteroscedasticity-robust standard errors.

Two additional issues arise from the definition of threshold dates. First, recall that the completion status of prefecture c on date d is derived from air quality between January 1^{st} and date $d - 1$. Therefore the AQI on the day before the threshold date is mechanically determined by the following definition: prefecture c will always change to the status of *doomed failure* (*early fulfillment*) after a polluted (unpolluted) day. To solve this problem and ensure the validity of the RDD, we follow the “donut RDD" method proposed by Barreca et al. (2011) and Hausman and Rapson (2018) and exclude observations in the immediate vicinity of the cutoff date. That is, we drop observations with t = 0 or -1, which correspond to the threshold date and the previous day. The second additional issue is that the number of remaining days after the threshold day until the end of the year is systematically different in failure vs. fulfillment years. As displayed in Figure 1 Panel A, except for some special cases with extremely low targets, most targets are larger than a half-year (i.e., 183 days); some even approach 360 or more. Generally speaking, the number of remaining days is usually greater in the failure year subsample, which allows for wider ranges of bandwidth selection. As a result, we set bandwidths as 35 and 15, respectively, in a baseline estimation and we would also show that our results are robust to using alternative bandwidths.

4.3 Baseline Empirical Results

4.3.1 Graphic Evidence and Regression Results from RDD

In this subsection, we investigate local bureaucrats' responses after the threshold date. Panels A and B of Figure 5 show the graphic evidence from the baseline RDD setting for the failure year and fulfillment year subsamples, respectively.

We begin by probing the situation in which prefectures failed to achieve their targets. Panel A shows that doomed failure did not cause the official to completely give up. On the contrary, the air quality performance sharply improved after the threshold date. This is not surprising, given that the extent of the failure affects the likelihood or severity of punishment. Local leaders were thus incentivized to narrow the gap between outcomes and minimum targets to reduce the likelihood of warnings or criticism from their superiors.

Panel B clearly shows that if the target is met early, the air quality deteriorates afterward. This finding is perfectly consistent with the obligatory nature of AQI targets in a multitasking environment: the over-fulfillment of targets gains little, but costs much by crowding out the efforts of promoting economic development, which contributes significantly to local officials' career advancement. We interpret such responses as the result of the salient feature of obligatory targets in hybrid contracts – the goal of "hitting the target only but no more."

Table 3 reports the estimates from regression analyses. The results from the failure year and fulfillment year subsamples are displayed in the first three and last three columns, respectively. Empirical specifications in Columns (1) and (4) are the same as the baseline setting introduced in Section 4.2 and correspond to the graphic evidence discussed above. In Columns (2) and (5), kernel functions are changed into triangular ones that weight observations based on their distance from the threshold date. Columns (3) and (6) allow for quadratic polynomials instead of local linear ones to account for more flexible trends on both sides of the cutoff (Gelman and Imbens, 2019), and the corresponding RDD plots are displayed in Panels C and D of Figure 5. Standard errors (in parentheses) are clustered at the prefecture level; we also report heteroscedasticity-robust standard errors (in brackets), as suggested by Lee and Lemieux (2010). The RDD estimates in Table 3 remain highly robust under different specifications and deliver findings similar to what we observed in Figure 5.[25]

Since there is no consensus on the criterion of bandwidth selection in the econometric literature, we experiment on a range of bandwidths and show the robustness of our findings. The range is centered on bandwidths in the baseline setting with a length of 20, i.e., 25–45 for the failure year sample and 5–25 for the fulfillment year sample. Figure 6 plots the point estimates and 95% confidence intervals with bandwidths on the horizontal axes. Both panels show that our key estimates remain stable in both magnitude and significance level, though the magnitude of point estimates in the right panel escalates as bandwidths narrow, which may partially result from the decrease in the size of observations and statistical power.

[25]Since standard errors clustered at the prefecture level are generally larger, for the sake of prudence, we would implement statistical inference with them by default in the following analyses.

4.3.2 The Pressures on Local Officials to Pursue Economic Performance

As mentioned at the beginning of this section, we aim to introduce some proxies for local officials' concerns about economic growth performance, and we will investigate how these pressures to achieve high economic performance affect their responses to minimum air quality targets. This analysis will shed light on how the two components of hybrid contracts – high-powered incentives and minimum standards – shape different responses from local officials.

Prefectural officials' economic performance is usually benchmarked against that of their peers in the same province (Chen et al., 2005; Landry et al., 2018). For instance, prefectural governments' quarterly work reports often contain statements such as "our quarterly growth rate ranked highly among prefectures within the same province" or "our growth rate was above the provincial level." Based on these observations, we constructed two measures to capture the pressure of economic growth faced by prefectural leaders. First, prefecture c in quarter q would be classified as "having a low burden of GDP growth" if its quarterly growth rate in quarter $q-1$ was above the median growth rate of all prefectures in the province. The second measure is whether the prefecture's quarterly growth rate in quarter $q-1$ was above the provincial growth rate.[26] We manually collected prefecture-quarter GDP growth rates from the official websites of prefectural governments or bureaus of statistics and then supplemented this information with provincial quarterly growth rates from the National Bureau of Statistics.

Table 4 reports the estimation results when we bifurcate the sample according to the relative burden of economic growth. The first two columns in each panel provide supportive evidence of the conflict between economic development and environmental protection. That is, if prefectural leaders performed poorly in economic growth in the last quarter and are thus under greater pressure to demonstrate growth, the motivation to improve air quality after doomed failure would be weaker. This result is reflected in a smaller magnitude of the estimated coefficient in Column (2). Correspondingly, Column (4) implies that prefectural officials under more pressure to meet economic growth targets would shift their efforts to economic development immediately after fulfilling their AQI targets, and air quality would deteriorate by large margins. The estimate in Column (3) is smaller in magnitude and statistically insignificant, suggesting that officials from prefectures with good GDP performance are less motivated to shift their efforts away from air pollution controls when their AQI targets are met early.[27]

4.3.3 Firm-Level Evidence

In the previous analysis, we uncovered two major facts. First, there are discontinuous changes in air quality around the threshold date of target completion. Second, when local officials are under less pressure to promote local economic growth, the improvement in air quality performance after a doomed failure is strengthened, and the deterioration in air quality performance

[26] The provincial GDP growth rate is approximately a weighted average of prefectural growth rates, with the weights being the respective GDP levels of each prefecture within the province. In essence, we identify prefectures with a low burden to demonstrate economic growth by comparing their performance with the median or weighted average of prefectures in the same province, respectively.

[27] We use the bootstrap method to show that the differences between estimates from subsample regressions (i.e., Column (1) vs. Column (2), and Column (3) vs. Column (4)) are statistically significant. See Appendix Figure C1 for more details.

after early fulfillment is alleviated. We interpret these findings as evidence for local officials' strategic responses to hybrid contracts in a multitasking environment. However, the extent to which this interpretation can hold depends on the linkage between air quality changes observed from the data and the regulatory responses of local government officials surrounding threshold dates. In what follows, we aim to provide evidence to substantiate this connection.

As introduced in Section 2.3, Chinese local governments have the discretion to affect air pollution levels in their jurisdictions by taking certain emergency response plans, such as closing factories, prohibiting fuel-inefficient vehicles, or shutting down construction sites. Existing literature has documented evidence about how the pollution emissions of Chinese firms are affected by regulatory or emergency measures taken by local governments (Karplus et al., 2018; Buntaine et al., 2022; Agarwal et al., 2023). Inspired by this literature, we conduct an empirical analysis to examine how the status of target completion translates into the pollution emission behavior of firms in response to local government emergency policies. The firm-day emission data are collected from the CEMS with a focus on SO_2, which is the main pollutant produced by manufacturing firms. The RDD model is specified and estimated in the same way as introduced in Section 4.2, except that all prefecture fixed effects are replaced by firm fixed effects.

As shown in Table 5, firms' responses in SO_2 emissions are highly consistent with the findings documented in Table 3. First, SO_2 emissions are significantly reduced after the arrival of a doomed failure while increasing after the threshold date of early fulfillment.[28] We intend to interpret these empirical findings as a result of local government emergency policies because firms had no incentive to respond to the threshold dates in such a way if there were no pressures from local governments.

To further corroborate our argument, we examine heterogeneous responses exhibited by different firms in three dimensions. First of all, in Panel A of Table 6, we examine how state-owned enterprises (SOEs) and non-SOEs differ in responses to different target completion statuses. It is shown that when the target is doomed to fail, the SO_2 emission of SOEs is significantly reduced while non-SOEs' emission has no significant changes. However, if the target is fulfilled in advance and there is no pressure to improve air quality performance, SOEs' emission change is statistically insignificant while non-SOEs' emission experiences a significant increase. The results are quite intuitive in the sense that SOEs are directly controlled by local governments, and thus more likely than non-SOEs to yield to the pressure of air pollution controls imposed by local governments. Second, we define those firms which supply water, electricity, and heating for local people and manage public facilities as firms that specialize in public service. We then divide the sample into public service firms and other firms which are mostly manufacturing firms (e.g., metallurgy plants). It is expected that public service firms whose key role is to ensure common people's livelihood should be less responsive to the temporary pressure of air pollution controls than non-public service firms. Consistent with this expectation, Panel B of Table 6 shows that the emission of public service firms seems to be insensitive to target completion statuses, while the asymmetric responses to target comple-

[28] Notice that the better air qualities correspond to higher values of NAQI and thus the lower levels of SO_2 emissions. So the estimates in Table 3 and 5 have opposite signs.

tion statuses, which are similar to those documented in the baseline regressions, are mainly driven by non-public service firms. Finally, we divide the sample into high-pollution and low-pollution firms. All firms covered in the sample are already listed for close monitoring by the government due to their relatively high pollution level, but they still differ in terms of SO_2 emission levels. We calculate the mean SO_2 emission for each firm in the year just before the implementation of AQI targets and then divide them into two groups according to whether the initial emission level is above the provincial median or not. If there is an urgent need to improve air quality, high-pollution firms would face higher pressure from the government than those low-pollution ones. Panel C of Table 6 shows clearly that the emission of high-pollution firms rather than low-pollution firms exhibits significant response patterns under different target completion statuses which are consistent with what we observed in the baseline regressions.

Our analysis in Section 4.3.2 has shown that the pressure of seeking high economic performance significantly affects changes in air quality upon learning the completion status of minimum targets. We interpret this finding as evidence of local officials' responses to hybrid incentive contracts in air quality regulation. Given this, one may wonder whether there is firm-level evidence to support our interpretation, as we have presented in Table 6. For this purpose, we narrow our sample to productive firms (as defined in Panel B of Table 6), which are expected to contribute significantly to regional economic growth. We hypothesize that local officials will tend to relax regulatory constraints on those larger productive firms in their jurisdictions if they face heavy pressure for economic growth, all else being equal. Due to data limitations, we define larger productive firms as those with employment size above the median employment scale of CEMS firms within the same prefecture.[29] Table 7 reports regression results for this subsample. No matter how we define the relative pressure to achieve high economic growth, a consistent data pattern emerges: if a prefecture faces a doomed failure of completing minimum targets, the subsequent reductions in firm-level SO_2 emission are significantly weakened when local officials confront with a heavier pressure of seeking high economic growth; if a prefecture achieves its target ahead of time, the deterioration in SO_2 emission is aggravated for those under a heavier economic growth pressure.[30] This firm-level evidence, alongside the results reported in Tables 5 and 6, supports our interpretation that results in Tables 3 and 4 are driven by local officials' responses to hybrid incentive contracts in environmental regulation.

There is a caveat to our firm-level analysis: closing down factories is just one of the policy tools at local governments' disposal to enforce environmental regulation, and the SO_2 emis-

[29] As explained in footnote 19, our source for firm-level data, Tianyancha, only provides the size of each firm's workforce in terms of the number of employees contributing to social security. This definition of employment size may not include employees who do not pay for social security benefits, introducing potential bias of "missing" employees. However, this measure can serve as a reasonable proxy for firm size for two reasons. First, the actual size of a firm's workforce is typically closely related to the number of employees on their social security payroll. Second, larger firms are less likely to be significantly affected by the bias introduced by "missing" employees.

[30] If we rerun regressions in Table 7 for the subsample of either public service firms or productive firms with employment size lower than the median, we obtain insignificant results for either status of target completion. The results are available upon request.

sion of industrial production only account for a portion of air pollution. As such, we regard the evidence in Tables 5, 6, and 7 as only preliminary in nature. However, this evidence does suggest that the discontinuous changes of aggregate AQI performance, observed in Section 4.3.1, are closely correlated with the discontinuous polluting behaviors of certain local firms around threshold dates. These firm-level reactions reflect the pressures from local governments to enact emergency action plans in air pollution controls.

To summarize, regression results from the RDD analyses indicate that after crossing the doomed failure threshold, local bureaucrats continue to make efforts to improve air quality rather than giving up completely, while slacking off after early fulfillment is achieved. It is also shown that local officials' strategic responses to air quality targets are amplified by their concerns for pursuing high economic performance: greater pressure for economic growth weakens the improvement in air quality after a doomed failure and aggravates the deterioration of air quality after an early fulfillment. These empirical findings highlight the trade-offs between promoting economic growth and controlling air pollution when Chinese local officials faced hybrid incentive contracts designed by the central government.

In the upcoming section, we will construct a coherent theoretical model to rationalize the empirical findings documented so far. We will also extend the model to understand the role of political incentives in hybrid incentive contracts, as well as interactions between hybrid incentive contracts and intrinsic motivations in a multitasking context. Previous studies have documented evidence for the important roles of political incentives and intrinsic motivations in the context of China's political economy (e.g., Persson and Zhuravskaya, 2016; Li et al., 2019). In light of the existing literature, it would be intriguing to explore whether and how political incentives and intrinsic motivations affect the strategic responses to hybrid incentive contracts.

5 Theoretical Model

We develop a stylized model to rationalize the empirical findings presented in the previous section. Our objective is to utilize the most straightforward model possible to explain how Chinese local officials respond strategically to the completion status of minimum targets in hybrid contracts. To illustrate, our benchmark model takes local officials' incentive schemes as given and considers their optimal efforts in a multi-period and multitasking setup.

5.1 Model Setting

There is a representative agent who can exert effort in two periods $t = 1, 2$. The effort choice in each period t is two-dimensional: (x_t, z_t), where x_t denotes the effort exerted to promote economic growth and z_t denotes the effort exerted to protect the environment. We use this two-period model as a shortcut to investigate the change in incentives from the initial uncertain stage to the early fulfillment / doomed failure stage within the year. Hence, we do not impose any discounting between these two periods.

We make two assumptions about how air quality is determined. First, we denote y_t as the air quality in period t (i.e., the number of days with good or moderate air quality in period t).

We assume that $y_1 = z_0 + z_1 - x_1$ and $y_2 = y_1 + z_2 - x_2$. Basically, the exogenous parameter z_0 corresponds to the pre-determined air quality at the beginning of the model, which reflects the stock of pollutants in the air. If the agent does nothing in either period (i.e., $z_1 = x_1 = z_2 = x_2 = 0$), the air quality will stay at z_0. The agent's efforts in period 1 will change the stock of pollutants in such a way that high efforts to protect the environment will reduce the stock while high efforts to promote economic growth will increase the stock. The net effect is measured as $z_1 - x_1$; hence the air quality in period 1 becomes $y_1 = z_0 + z_1 - x_1$. Similarly, the air quality in period 2 is given by $y_2 = y_1 + z_2 - x_2$.

Second, there are two states of the world $s = g, b$ regarding the value of z_0. When $s = g$ (good), $z_0 = z_0^g$ is so high that the agent can fulfill the environmental target y_0 without exerting any effort; when $s = b$ (bad), $z_0 = z_0^b$ is so low that the agent fails to meet the environmental target y_0 for sure. The prior belief is that $s = g$ and $s = b$ occur with probability a and $1 - a$, respectively. We introduce these two states to capture the uncertainty faced by the agent about how easy to fulfill the environmental target. To simplify our analysis, we assume that the states are not known to the agent in period 1 but are perfectly revealed at the beginning of period 2. In other words, the beginning of period 2 is the threshold date, at which the agent knows for sure whether the status is early fulfillment or doomed failure.

Next, we make the following assumptions about how the agent's final payoff is determined to capture important features of local officials' incentive schemes. First, if the agent fulfills environmental target y_0 (i.e., the total number of days with good or moderate air quality exceeds y_0: $y_1 + y_2 \geq y_0$), then his or her expected reward is $f(x_1 + x_2)v$, where $x_1 + x_2$ is interpreted as the performance in economic growth. This assumption is consistent with the empirical literature on the promotion of China's local officials: Those in prefectures with higher GDP growth rates are more likely to be promoted (see, e.g., Li and Zhou (2005)). So we can interpret $f(x_1 + x_2)$ as the probability of promotion, satisfying $f' > 0$ and $f'' \leq 0$. Moreover, the pre-determined constant v measures the value of promotion.

Second, if the agent fails to meet the environmental target y_0 (i.e., $y_1 + y_2 < y_0$), then he or she has zero probability of promotion. Moreover, there is a probability p of being punished (i.e., retirement or being investigated). If the agent is punished, his loss of utility is a pre-determined constant u. Therefore, in this case, the agent's expected loss is $-pu$. As discussed in Section 2.2, the probability p depends on the distance between the realized air quality and the target: $\Delta_y = y_0 - y_1 - y_2$. We further assume that p is convex and strictly increasing in Δ_y.

Finally, the agent's total effort cost is a function of total efforts $(x, z) = (x_1 + x_2, z_1 + z_2)$, written as $C(x_1 + x_2, z_1 + z_2)$. As in Holmstrom and Milgrom (1991) and Bénabou and Tirole (2016a), we assume that $C_{xz} \geq 0$, meaning that the two efforts are substitutes. For simplicity, we focus on the following convenient specification of the cost function throughout this paper: $C(x, z) = \frac{1}{2}c$

c (z)]86 F11 2]TJ/F76 10.9091 Tf 4.543 0 Td [(1)]TJ/F112 11.3673 Tf 7.658 0 Td [(()]TJ/F76 10.9091 Tf 4.866 0 Td [(x)]TJ/F72 7.9701 Tf 6.331 -1.811 Td [(1)]TJ/112 11.3673 T

in the bad state, it is:

$$-p$$

small, we can always get both $z_1 > 0$ and $x_2^g > 0$.

PROPOSITION 2. *Assume $a \geq \frac{1}{2}$. If k is sufficiently close to zero, we must have $z_1 > 0$ and $x_2^g > 0$.*[32]

In the literature, it is common to assume that k is small. For example, Holmstrom and Milgrom (1991) assume that $k < \sqrt{c_x c_z}$. In this case, the interior solutions of x_2^g and z_2^b explain our observations after early fulfillment and doomed failure.

Notice that the change of air quality after the state is revealed, proxied by $y_2 - y_1$, depends on the net effect of efforts, $z_2 - x_2$. After early fulfillment, the net effect of efforts is $z_2^g - x_2^g < 0$, and hence there is a deterioration in air quality. Yet after doomed failure, the agent knows that the state is bad and has to exert greater efforts to protect the environment. The net effect of efforts is $z_2^b - x_2^b > 0$, and hence there is an improvement in air quality.

The parameter value of v is highly correlated with political incentives. For example, a younger agent will expect to receive more rent from promotions and hence have a higher v. Based on the first-order conditions, we can immediately derive how the change in air quality depends on political incentives.

PROPOSITION 3. *Suppose $a \geq \frac{1}{2}$ and $k < \sqrt{c_x c_z}$ is small enough such that the optimal solutions z_1, x_2^g and z_2^b are interior. Then the optimal solutions satisfy that $\frac{\P x_2^g}{\P v} > 0$ and $\frac{\P z_2^b}{\P v} > 0$.*

The intuition of the above comparative statice result is as follows. As v increases, an agent would like to exert more effort to promote economic growth when the state is good. This naturally leads to a decrease in z_1 because a lower z_1 decreases the marginal cost of exerting x_2^g. Then in the bad state, the agent will exert greater effort z_2^b because (1) a lower z_1 implies that the agent must exert greater effort in period 2 to decrease the probability of punishment; and (2) a lower z_1 also decreases the marginal cost of exerting z_2^b. Therefore, as political incentives increase, both the improvement in air quality after doomed failure and the deterioration in air quality after early success will increase.

5.3 Model Extensions

In this section, we extend the benchmark model to include the impact of agents' intrinsic motivations and supervision from above. These extensions yield more testable implications.

5.3.1 The Effects of Agents' Intrinsic Motivations

In reality, the agent may also receive intrinsic value from improving air quality. In period 1, the improvement is given by $y_1 - z_0$ while in period 2, it is denoted as $y_2 - z_0$. Therefore, the total utility is modeled as $q[j(y_1 - z_0 + y_2 - z_0)]$, where q measures the relative importance of environmental protection in the agent's utility function, and j is a strictly increasing and concave function. Given the specification of air quality $y_1 = z_0 + z_1 - x_1$ and $y_2 = y_1 + z_2 - x_2$, we can rewrite $y_1 - z_0 + y_2 - z_0$ as:

$$y_1 - z_0 + y_2 - z_0 = 2(z_1 - x_1) + z_2 - x_2,$$

[32] In the poof of Proposition 5, we show the exact requirements on the parameter value of k in a simple numerical example.

which is independent of the initial condition z_0.

Due to the intrinsic motivation, in the good state, the agent's total utility becomes:

$$f(x_1 + x_2)v + q[j(2(z_1 - x_1) + z_2 - x_2)] - C(x_1 + x_2, z_1 + z_2);$$

in the bad state, it becomes:

$$-p(\Delta_y)u + q[j(2(z_1 - x_1) + z_2 - x_2)] - C(x_1 + x_2, z_1 + z_2).$$

Notice that our benchmark model corresponds to the special case of $q = 0$. And we analyze the general case of $q > 0$ in this section. We also solve the problem through backward induction and find that similar to the $q = 0$ case, our first result establishes the optimality of the corner solution.

PROPOSITION 4. *Assume that $a \geq \frac{1}{2}$. The optimal efforts satisfy $x_1 = z_2^g = x_2^b = 0$.*

At first glance, Proposition 4 derives the same results as Proposition 1. However, the existence of intrinsic motivations significantly changes the equilibrium values of z_1, x_2^g, and z_2^b. From the expression of intrinsic motivations, $q[j(2(z_1 - x_1) + z_2 - x_2)]$, we can see that z_1 has a greater impact on intrinsic value due to the long-lasting effect of environmental protection effort exerted in period 1. Therefore, as q increases, it is natural to increase z_1 accordingly. Compared with the $q = 0$ case, the first derivative with respect to x_2^g implies a sharp decrease in x_2^g for two reasons: First, the intrinsic value of environmental protection makes the agent less willing to exert effort to promote economic growth. Second, a higher z_1 increases the marginal cost of exerting effort on economic growth when $k > 0$. Therefore, when q becomes large enough, it is possible that the optimal x_2^g becomes zero.

As q increases, there are two opposing effects on z_2^b. On the one hand, the higher intrinsic value of environmental protection forces the agent to exert greater effort; on the other hand, a higher z_1 makes the effort more costly. The long-lasting effect of environmental protection effort exerted in period 1 implies a crowd-out of z_2^b by z_1 since the effort z_1 is relatively cheaper. As a result, we also expect to see $z_2^b = 0$ when q becomes sufficiently large.

We use a simple numerical example to formally illustrate the above intuitions. In this example, we assume that all of f, p, and j are linear, and denote $I_f = f'$, $I_p = p'$, and $I_j = j'$. Then the first derivatives with respect to z_1, x_2^g, and z_2^b are given by:

$$a(2qI_j - c_z z_1 - kx_2^g) + (1 - a)(2uI_p + 2qI_j - c_z(z_1 + z_2^b));$$

$$vI_f - qI_j - c_x x_2^g - kz_1;$$

and

$$uI_p + qI_j - c_z(z_1 + z_2^b).$$

Notice that each of the above first derivatives equals zero when the corresponding optimal solution is interior, and is less than zero otherwise. Our next proposition shows that for any fixed parameters $a \geq \frac{1}{2}$ and $k < \sqrt{c_x c_z}$ such that the optimal solutions x_2^g and z_2^b are interior when $q = 0$, there always exists $\bar{q} > 0$ such that the optimal solutions x_2^g and z_2^b are zero when $q \geq \bar{q}$.

PROPOSITION 5. *Fix any $a \geq \frac{1}{2}$ and $k < \sqrt{c_x c_z}$. Suppose that $k < \frac{a}{1-a} \frac{I_f v}{\underline{I}_p u} cz$. Then there exists $\bar{q} > 0$ such that $x_2^g > 0$ and $z_2^b > 0$ when $q = 0$; and $x_2^g = z_2^b = 0$ when $q \geq \bar{q}$.*

Proposition 5 justifies our intuitive discussions stated below Proposition 4. In particular, starting from optimal solutions z

>a-f-z

2 2 if(solutwe g b ... z cel nc)] ... f -1.533 -2.645 Td [(22)]T8 11.3673 Tf 8.48 0 Td [( )]T] 0 g8 11.3673 fd [(

when the state is bad, it is:

$$-p(\Delta_y)u + p_b(z_2 - x_2) - \frac{1}{2}c_x(x_1 + x_2)^2 - \frac{1}{2}c_z(z_1 + z_2)^2.$$

The functions p_g and p_b can be interpreted as rewards / punishments based on the agent's performance in period 2. For example, in the good state, if an agent is caught slacking on environmental protection, he or she can be punished even if they meet their environmental target; in the bad state, if the agent is found to work hard on environmental protection, they may be punished less severely for failing to meet the target. Our next result shows that under certain assumptions, introducing p_g and p_b improves the agent's performance in environmental protection in the second period regardless of the state.

PROPOSITION 6. *Suppose both p_g and p_b are linear satisfying $p_g' = l_g$ and $p_b' = l_b$. If l_g and l_b satisfy*

$$-a\,l_g + (1-a)(l_p u - l_b) < 0.$$

Then compared with the benchmark model, the introduction of p_g and p_b leads to a larger $z_2^g - x_2^g$ and a larger $z_2^b - x_2^b$.

This proposition implies that if the incentive scheme can be designed to incorporate continuous supervision and surveillance, then the agent will strategically respond to this oversight by delivering a greater (lesser) improvement in air quality after doomed failure (early fulfillment).

6 Additional Empirical Tests Derived from the Model

Our theoretical model rationalizes the empirical findings reported in Tables 3 and 4 and generates new insights about the effects of intrinsic motivations, political incentives, and central monitoring on local officials' responses to different target fulfillment statuses. In this section, we empirically test these new theoretical predictions in turn.

6.1 Heterogeneous Responses Exhibited by Officials

Propositions 4 and 5 highlight the potential mitigating role played by intrinsic motivations. That is, if the intrinsic value q becomes sufficiently large, agents would not be sensitive to either doomed failure or early success, and thus air quality should not exhibit discontinuous changes. We follow Persson and Zhuravskaya (2016) and introduce a proxy for (a lack of) intrinsic motivations for environmental quality – whether the local official is an outsider. Since Persson and Zhuravskaya (2016) only present evidence for the positive effect of local experience on local public goods such as basic education, it is unclear whether this effect holds for local environmental protection. To support our proxy for local officials' intrinsic motivations to improve environmental quality, Appendix D provides more evidence for the positive effect of bureaucrats' local experience on local environmental protection.

We define the party secretary of prefecture c as an outsider if he or she is appointed from other prefectures or departments of provincial governments, and as a non-outsider if he or she is promoted within prefecture c.[33] Furthermore, considering that some local bureaucrats may be repeatedly moved among jurisdictions repeatedly during their careers, even if the party secretary of prefecture c was laterally moved from prefecture b, he may still have worked in prefecture c earlier in his or her career. To account for this possibility, we compute the total length of time the incumbent party secretary has served in prefecture c throughout his/her career before being appointed the party secretary, and further define those with no previous work experience in the prefecture as outsiders, which is a stricter version of outsiders.

We further split the subsamples of failure and fulfillment years according to whether the incumbent party secretary is an outsider, which generates four subsamples: failure year & outsider, failure year & non-outsider, fulfillment year & outsider, fulfillment year & non-outsider. Table 8 reports the results from subsample regressions where outsiders in Panels A and B are defined differently as described above. Consistent with predictions derived in Section 5.3.1, the results show that local bureaucrats defined as outsiders exhibit responses similar to our baseline results. The estimated coefficients in Columns (1) and (3) in both panels are significant at the same level with greater magnitude compared with Columns (1) and (4) in Table 3. However, those non-outsiders change little in the intensity of air pollution control, which is reflected by the insignificant estimates in Columns (2) and (4). Taken together, we find strong evidence that outsiders respond more sensitively or strategically to target fulfillment statuses than non-outsiders. This stark difference can be understood as supportive evidence of the mitigating effect of intrinsic motivations à la Persson and Zhuravskaya (2016) on the "gaming" incentive of local officials to regulate air quality.

6.2 Effects of Political Incentives in the Absence of Intrinsic Motivations

The theoretical model highlights the role of local officials' political incentives. Proposition 3 predicts that if we focus on local officials with no intrinsic motivations, both air quality improvement after doomed failure and deterioration after early fulfillment increase with the intensity of bureaucrats' political motivation. To empirically test this prediction, we construct two measures to capture the intensity of bureaucrats' political incentives. First, we follow the method proposed by Yu et al. (2016) and He et al. (2020) and define high-incentive officials as those younger than 57, which utilizes the informal rule in China that the probability of promotion drops dramatically after this age. Second, we compute career-concern intensity for every party secretary in the same way as Wang et al. (2020). For each province in each year, we pool all incumbent party secretaries and obtain the median career-concern intensity; we consider those above the median to be high-incentive officials. This measure is based on the locality of political tournaments in China, in which prefectural party secretaries mainly compete with their peers in other prefectures within the same province. It has been the case since the personnel control reforms launched in 1984, which empowered provincial governments to appoint prefecture-level officials in their provinces.

[33] The definition of an outsider is relatively clean-cut in the sense that the incumbent party secretary has never worked in the prefecture before and thus has little local connection. However, we could not simply regard the other ones as insiders or locals. Some of them may have worked in the specific prefecture for some period of time and move to another prefecture and then come back after a while. In this case it is hard to measure the extent by which they are locally connected or oriented. Therefore we refer them broadly as "non-outsiders."

In the theoretical model, we discuss the role of political incentives in the case of $q = 0$ for simplicity. To keep the empirical design consonant with it, we first bifurcate local bureaucrats into high- / low-incentive subsamples and then restrict regression analyses to outsiders who have never worked in the prefectures where they are currently taking office (i.e., the stricter definition of outsiders introduced in Section 6.1). Again, we create four subsamples for heterogeneity analysis: failure year & low-incentive, failure year & high-incentive, fulfillment year & low-incentive, fulfillment year & high-incentive.

Table 9 reports estimates from the subsample regressions. High-incentive bureaucrats are defined as those younger than 57 in Panel A and those with above-median career-concern intensity in Panel B. If we compare estimates in Columns (1) and (3) with those in Columns (2) and (4), it is obvious that high-incentive officials react much more intensely than their low-incentive counterparts when faced with either doomed failure or early fulfillment, as implied by the estimates in Columns (2) and (4), which dominate in both magnitude and significance level, regardless of how "high incentive" is defined.

6.3 Effects of Central Supervision

Proposition 6 predicts that stricter central government surveillance will motivate prefectural leaders to exert more effort in case of doomed failure and to alleviate the deterioration in the case of early fulfillment. The central government surveillance campaigns allow us to test this prediction. In 2013, the MEE issued a list of 79 cities that have been closely watched by the central government since then. As stated in a publicly accessible document, the MEE evaluates those cities on a monthly basis and discloses the bottom 10 cities every month, which imposes a heavy burden on cities on the watchlist. The watchlist was expanded to 164 cities in the second half year of 2018.[34] We classify prefectures on the watchlist as "strictly supervised" and those not on the list as "loosely supervised."

Table 10 shows that the regular surveillance from the central government has effectively constrained bureaucrats' strategic shirking on air pollution controls. Comparing the estimates in the first and last two columns clearly indicates that the improvement after doomed failure is strengthened (and the deterioration after early fulfillment is weakened) in prefectures under stricter supervision. Furthermore, consistent with the assumption made in Section 5.3.2 ($q = k = 0$), the results from subsample regressions in Panels B and C suggest that the empirical findings are mainly driven by outsiders, who derive little intrinsic value from air quality improvement and are thus more sensitive to central supervision.

7 Conclusion

This paper examines how Chinese local officials respond strategically to minimum air quality standards, particularly when regional economic development – a factor closely tied to their career advancement – is a primary concern. We leverage several institutional features of China's

[34] See Shi et al. (2019) for a comprehensive introduction of institutional details and an evaluation of its effectiveness.

environmental regulation, such as top-down target setting, the obligatory and quantitative nature of air pollution control targets, and the government's recent crackdown on data manipulation to overcome empirical challenges associated with measuring the impact of such hybrid contracts. Using a modified version of RDD, we find that local bureaucrats continue to make efforts in environmental protection even when the minimum target is doomed to fail, but tend to ease off these efforts if they meet their targets before the year's end. Our empirical results also reveal that the pressure for seeking high economic growth changes the responses of local officials to target fulfillment statuses: a greater pressure for economic growth weakens air quality improvements after doomed failure and aggravates its deterioration after early fulfillment.

We further develop a theoretical model to rationalize the empirical findings documented above and extend it to understand the interplay between hybrid incentive contracts and intrinsic motivations. Both model predictions and empirical findings show that those local officials with low affinities for local environmental protection or stronger political incentives exhibit a more strategic response to the completion status of obligatory targets. Moreover, stricter supervision from the central government could enhance air quality improvement after doomed failure and mitigate the deterioration after early fulfillment. Using China's environmental regulation as a case study, our analysis offers fresh insights into how hybrid contracts worked in a multitasking context.

References

Agarwal, S., Y. Han, Y. Qin, and H. Zhu (2023). Disguised Pollution: Industrial Activities in the Dark. *Journal of Public Economics 223*, 104904.

Ashraf, N., O. Bandiera, and B. K. Jack (2014). No Margin, No Mission? A Field Experiment on Incentives for Public Service Delivery. *Journal of Public Economics 120*, 1–17.

Baker, G. P. (1992). Incentive Contracts and Performance Measurement. *Journal of Political Economy 100*(3), 598–614.

Barreca, A. I., M. Guldi, J. M. Lindo, and G. R. Waddell (2011). Saving Babies? Revisiting the Effect of Very Low Birth Weight Classification. *Quarterly Journal of Economics 126*(4), 2117–2123.

Bénabou, R. and J. Tirole (2003). Intrinsic and Extrinsic Motivation. *Review of Economic Studies 70*(3), 489–520.

Bénabou, R. and J. Tirole (2006). Incentives and Prosocial Behavior. *American Economic Review 96*(5), 1652–1678.

Bénabou, R. and J. Tirole (2016a). Bonus Culture: Competitive Pay, Screening, and Multitasking. *Journal of Political Economy 124*(2), 305–370.

Bénabou, R. and J. Tirole (2016b). Mindful Economics: the Production, Consumption, and Value of Beliefs. *Journal of Economic Perspectives 30*(3), 141–164.

Besley, T. and M. Ghatak (2005). Competition and Incentives with Motivated Agents. *American Economic Review 95*(3), 616–636.

Besley, T. and M. Ghatak (2006). Sorting with Motivated Agents: Implications for School Competition and Teacher Incentives. *Journal of the European Economic Association 4*(2-3), 404–414.

Buntaine, M., M. Greenstone, G. He, S. Wang, M. Liu, and B. Zhang (2022). Does the Squeaky Wheel Get More Grease? The Direct and Indirect Effects of Citizen Participation on Environmental Governance in China. Working Paper.

Burgess, R., M. Hansen, B. A. Olken, P. Potapov, and S. Sieber (2012). The Political Economy of Deforestation in the Tropics. *Quarterly Journal of Economics 127*(4), 1707–1754.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica 82*(6), 2295–2326.

Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2012). Gaming in Air Pollution Data? Lessons from China. *B.E. Journal of Economic Analysis & Policy 13*(3).

Chen, Y., G. Z. Jin, N. Kumar, and G. Shi (2013). The Promise of Beijing: Evaluating the Impact of the 2008 Olympic Games on Air Quality. *Journal of Environmental Economics and Management 66*(3), 424–443.

Chen, Y., H. Li, and L.-A. Zhou (2005). Relative Performance Evaluation and the Turnover of Provincial Leaders in China. *Economics Letters 88*(3), 421–425.

Chen, Y. J., P. Li, and Y. Lu (2018). Career Concerns and Multitasking Local Bureaucrats: Evidence of a Target-based Performance Evaluation System in China. *Journal of Development Economics 133*, 84–101.

Chen, Z., M. E. Kahn, Y. Liu, and Z. Wang (2018). The Consequences of Spatially Differentiated Water Pollution Regulation in China. *Journal of Environmental Economics and Management 88*, 468–485.

Courty, P. and G. Marschke (2004). An Empirical Investigation of Gaming Responses to Explicit Performance Incentives. *Journal of Labor Economics 22*(1), 23–56.

Finan, F., B. A. Olken, and R. Pande (2017). The Personnel Economics of the Developing State. In *Handbook of Economic Field Experiments*, Volume 2, pp. 467–514. Elsevier.

Garicano, L. and L. Rayo (2016). Why Organizations Fail: Models and Cases. *Journal of Economic Literature 54*(1), 137–192.

Gelman, A. and G. Imbens (2019). Why High-order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics 37*(3), 447–456.

Ghanem, D. and J. Zhang (2014). Effortless Perfection: Do Chinese Cities Manipulate Air Pollution Data? *Journal of Environmental Economics and Management 68*(2), 203–225.

Gneezy, U. and A. Rustichini (2000a). A Fine is a Price. *Journal of Legal Studies 29*(1), 1–17.

Gneezy, U. and A. Rustichini (2000b). Pay Enough or Don't Pay at All. *Quarterly Journal of Economics 115*(3), 791–810.

Greenstone, M., G. He, R. Jia, and T. Liu (2022). Can Technology Solve the Principal-Agent Problem? Evidence from China's War on Air Pollution. *American Economic Review: Insights 4*(1), 54–70.

Greenstone, M., G. He, S. Li, and E. Y. Zou (2021). China's war on pollution: Evidence from the first 5 years. *Review of Environmental Economics and Policy 15*(2), 281–299.

Hart, O., A. Shleifer, and R. W. Vishny (1997). The Proper Scope of Government: Theory and an Application to Prisons. *Quarterly Journal of Economics 112*(4), 1127–1161.

Hausman, C. and D. S. Rapson (2018). Regression Discontinuity in Time: Considerations for Empirical Applications. *Annual Review of Resource Economics 10*, 533–552.

He, G., S. Wang, and B. Zhang (2020). Watering Down Environmental Regulation in China. *Quarterly Journal of Economics 135*(4), 2135–2185.

Heckman, J., C. Heinrich, and J. Smith (1997). Assessing The Performance of Performance Standards in Public Bureaucracies. *American Economic Review 87*(2), 389–395.

Heckman, J., C. Heinrich, and J. Smith (2011). Performance Standards and the Potential to Improve Government Performance. Working Paper.

Heinrich, C. J. (2002). Outcomes-Based Performance Management in the Public Sector: Implications for Government Accountability and Effectiveness. *Public Administration Review 62*(6), 712–725.

Holmstrom, B. and P. Milgrom (1991). Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics and Organization 7*, 24–58.

Jia, R. (2017). Pollution for Promotion. Working Paper.

Kahn, M. E., P. Li, and D. Zhao (2015). Water Pollution Progress at Borders: The Role of Changes in China's Political Promotion Incentives. *American Economic Journal: Economic Policy 7*(4), 223–42.

Karplus, V. J. and M. Wu (2020). Crackdowns in Hierarchies: Evidence from China's Environmental Inspections. Working Paper.

Karplus, V. J., J. Zhang, and J. Zhao (2021). Navigating and Evaluating the Labyrinth of Environmental Regulation in China. *Review of Environmental Economics and Policy 15*(2), 300–322.

Karplus, V. J., S. Zhang, and D. Almond (2018). Quantifying Coal Power Plant Responses to Tighter SO_2 Emissions Standards in China. *Proceedings of the National Academy of Sciences 115*(27), 7004–7009.

Kennedy, S. and C. K. Johnson (2016). *Perfecting China, Inc.: China's 13th Five-Year Plan*. Rowman & Littlefield.

Landry, P. F., X. Lü, and H. Duan (2018). Does Performance Matter? Evaluating Political Selection along the Chinese Administrative Ladder. *Comparative Political Studies 51*(8), 1074–1105.

Lee, D. S. and T. Lemieux (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature 48*(2), 281–355.

Li, H. and L.-A. Zhou (2005). Political Turnover and Economic Performance: The Incentive Role of Personnel Control in China. *Journal of Public Economics 89*(9-10), 1743–1762.

Li, P., Y. Lu, and J. Wang (2020). The Effects of Fuel Standards on Air Pollution: Evidence from China. *Journal of Development Economics 146*, 102488.

Li, X., C. Liu, X. Weng, and L.-A. Zhou (2019). Target Setting in Tournaments: Theory and Evidence from China. *Economic Journal 129*(623), 2888–2915.

List, J. A. and D. M. Sturm (2006). How Elections Matter: Theory and Evidence from Environmental Policy. *Quarterly Journal of Economics 121*(4), 1249–1281.

McCrary, J. (2008). Manipulation of the Running Variable in the Regression Discontinuity Design: A Density Test. *Journal of Econometrics 142*(2), 698–714.

Murdock, K. (2002). Intrinsic Motivation and Optimal Incentive Contracts. *RAND Journal of Economics 33*(4), 650–671.

Newman, J. L. and J. P. Azevedo (2013). Setting Reasonable Performance Targets for Public Service Delivery. Working Paper.

Persson, P. and E. Zhuravskaya (2016). The Limits of Career Concerns in Federalism: Evidence from China. *Journal of the European Economic Association 14*(2), 338–374.

Prendergast, C. (2007). The Motivation and Bias of Bureaucrats. *American Economic Review 97*(1), 180–196.

Selviaridis, K. and F. Wynstra (2015). Performance-Based Contracting: A Literature Review and Future Research Directions. *International Journal of Production Research 53*(12), 3505–3540.

Shi, C., F. Guo, and Q. Shi (2019). Ranking Effect in Air Pollution Governance: Evidence from Chinese Cities. *Journal of Environmental Management 251*, 109600.

Shi, Q., S. Chen, and F. Guo (2017). Administrative Interviews of China MEP and Environmental Governance: An Example of Air Pollution. *Statistical Research 10*, 88–97. (in Chinese).

Wang, Z., Q. Zhang, and L.-A. Zhou (2020). Career Incentives of City Leaders and Urban Spatial Expansion in China. *Review of Economics and Statistics 102*(5), 897–911.

Xu, C. (2011). The Fundamental Institutions of China's Reforms and Development. *Journal of Economic Literature 49*(4), 1076–1151.

Yao, Y. and M. Zhang (2015). Subnational Leaders and Economic Growth: Evidence from Chinese Cities. *Journal of Economic Growth 20*(4), 405–436.

Yu, J., L.-A. Zhou, and G. Zhu (2016). Strategic Interaction in Political Competition: Evidence from Spatial Effects Across Chinese cities. *Regional Science and Urban Economics 57*, 23–37.

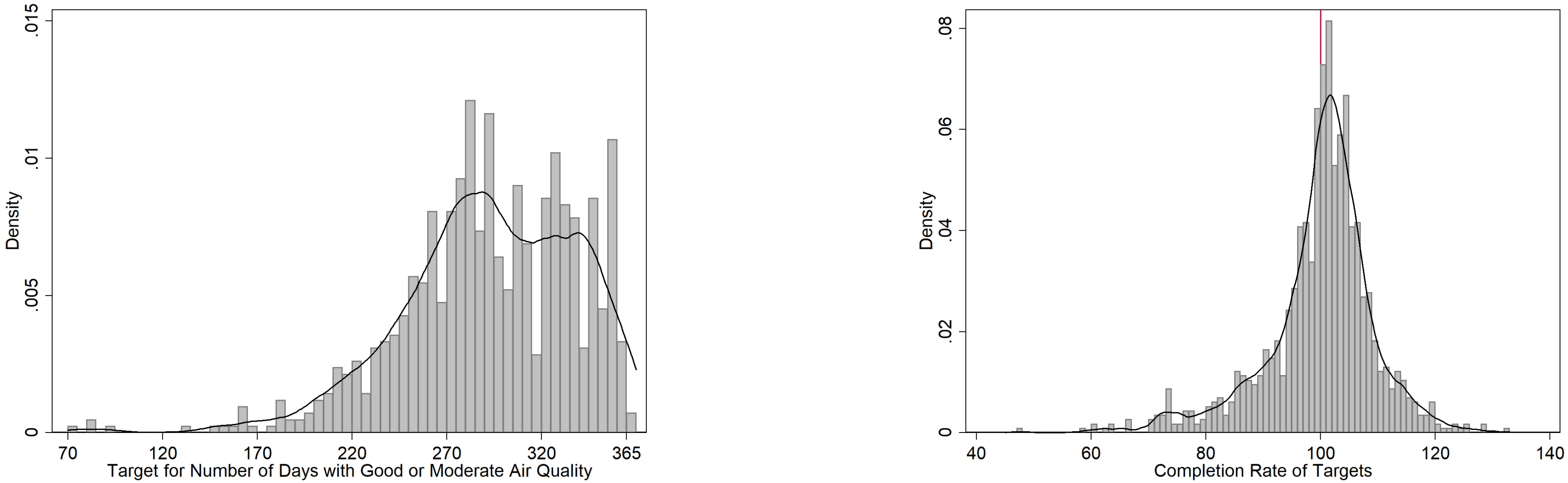
Zheng, S. and M. E. Kahn (2013). Understanding China's Urban Pollution Dynamics. *Journal of Economic Literature 51*(3), 731–772.

Zheng, S., M. E. Kahn, W. Sun, and D. Luo (2014). 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.

Figures and Tables

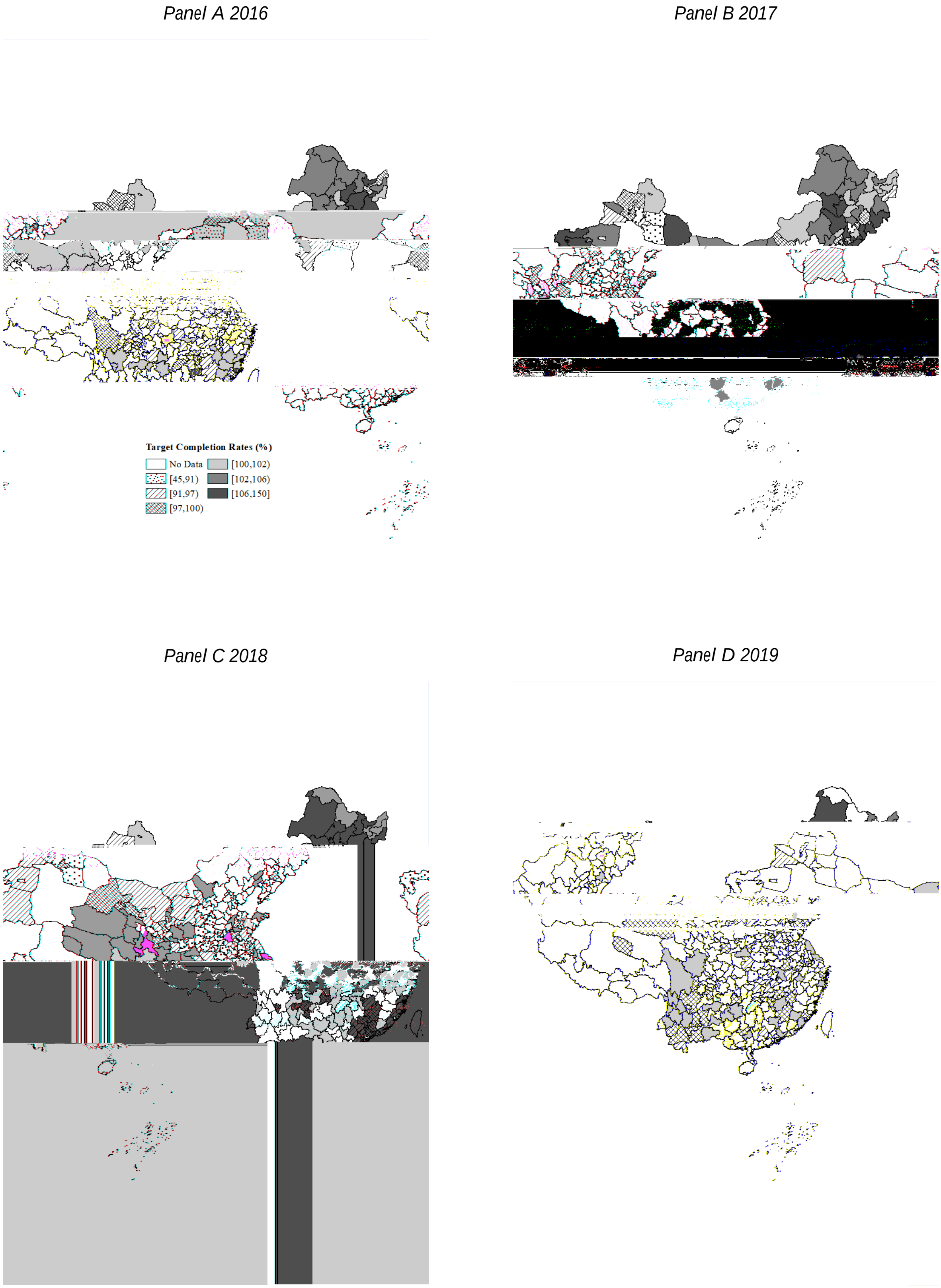
Figure 1. Distribution of Targets and Completion Rates

Panel A Distribution of Targets *Panel B Distribution of Completion Rates*



Notes: This graph plots the histogram of prefecture-year targets and completion rates along with the estimated kernel density. The completion rate equals to (actual percentage of days with AQI $\leq$ 100 / target percentage of days with AQI $\leq$ 100) $\times$ 100. That is, completion rates no less than 100 imply that the targets have been fulfilled.

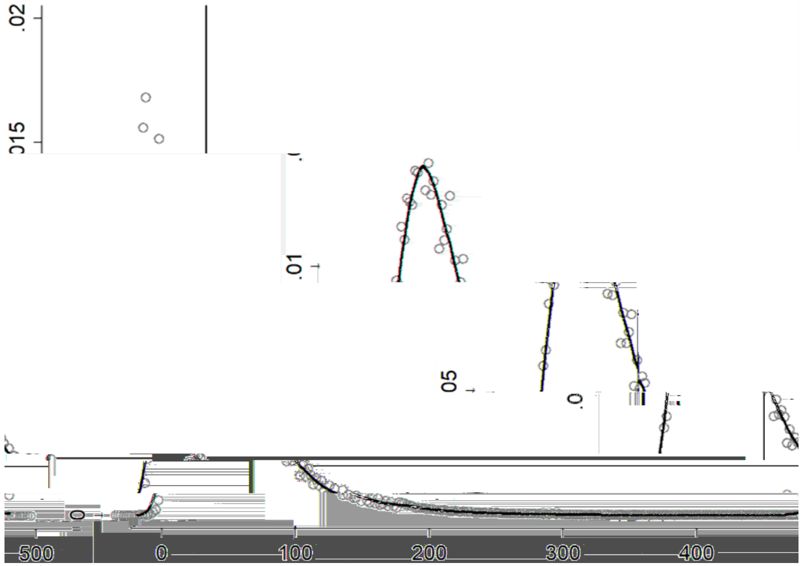
Figure 2. Geographical Distributions of Target Completion Rates



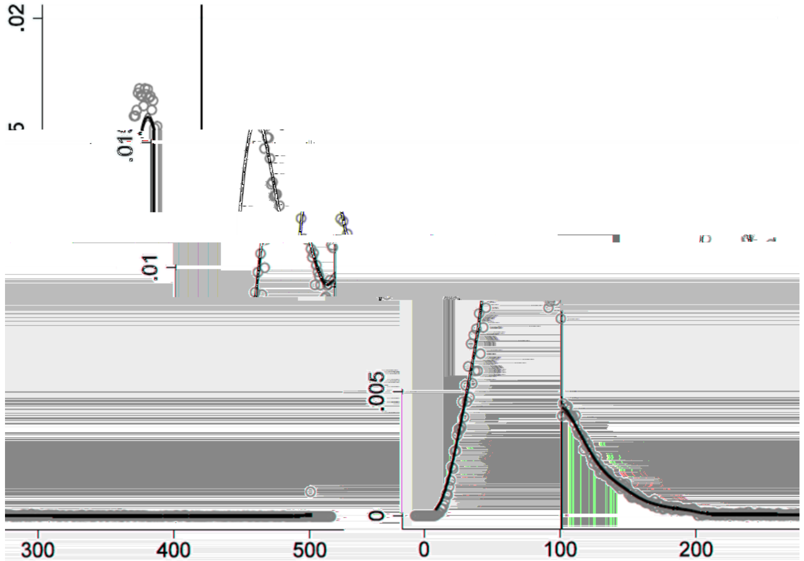
Notes: This graph plots the prefectural completion rates of targets on China's map by year. It shows that completion statuses vary considerably both cross-sectionally and over time.

Figure 3. McCrary Test of AQI and Discontinuity around 100

Panel A McCrary Test of AQI in Our Sample

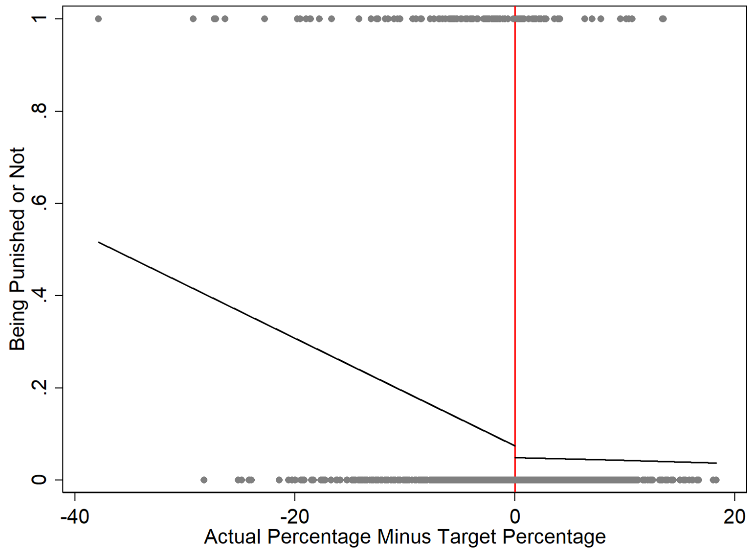


Panel B McCrary Test of API in Chen et al. (2012)



Notes: Panel A plots the graph of the McCrary test of the AQI data used in this paper and Panel B is extracted from Chen et al. (2012). The obvious discontinuity around 100 in Panel B was interpreted as the clue for manipulation of API in Chen et al. (2012) while it is not observed in our AQI data.

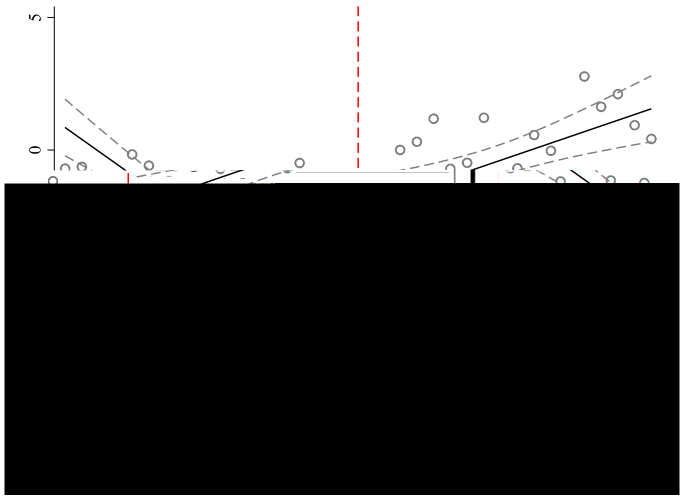
Figure 4. Supportive Evidence for Punishment on Failing AQI Targets



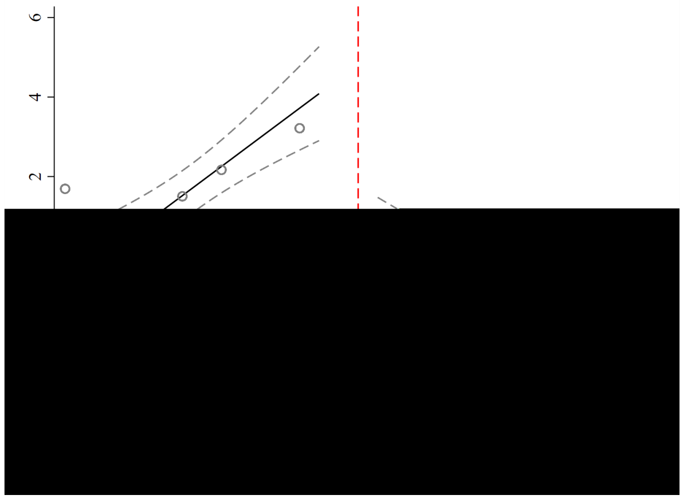
Notes: This graph shows the relationship between performance in AQI targets and the probability of being punished. The prefecture-year differences between the actual percentage of days with AQI ≤ 100 and the target percentage are displayed on the horizontal axis; values equal to or larger than 0 indicate fulfillment.

Figure 5. Graphic Evidence from RDD

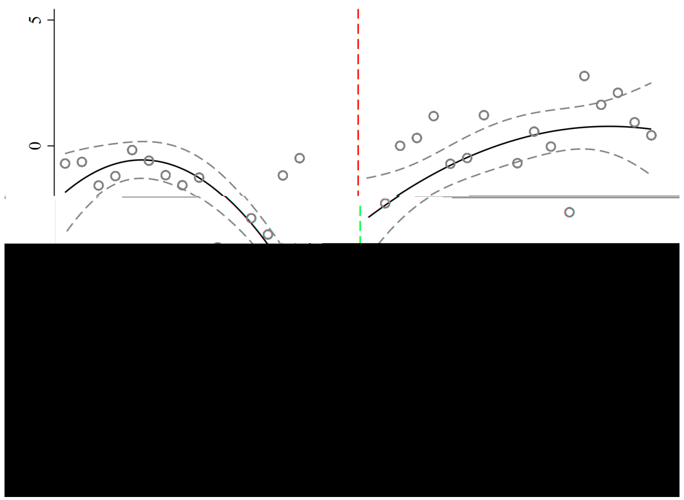
Panel A Failure Year (Local Linear)



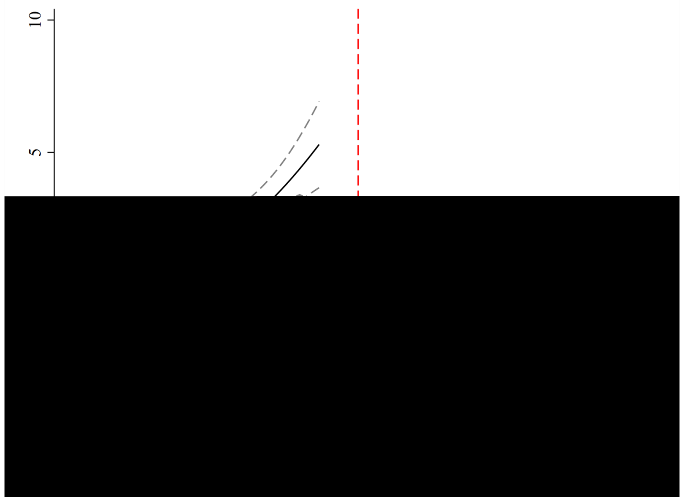
Panel B Fulfillment Year (Local Linear)



Panel C Failure Year (Quadratic Polynomials)



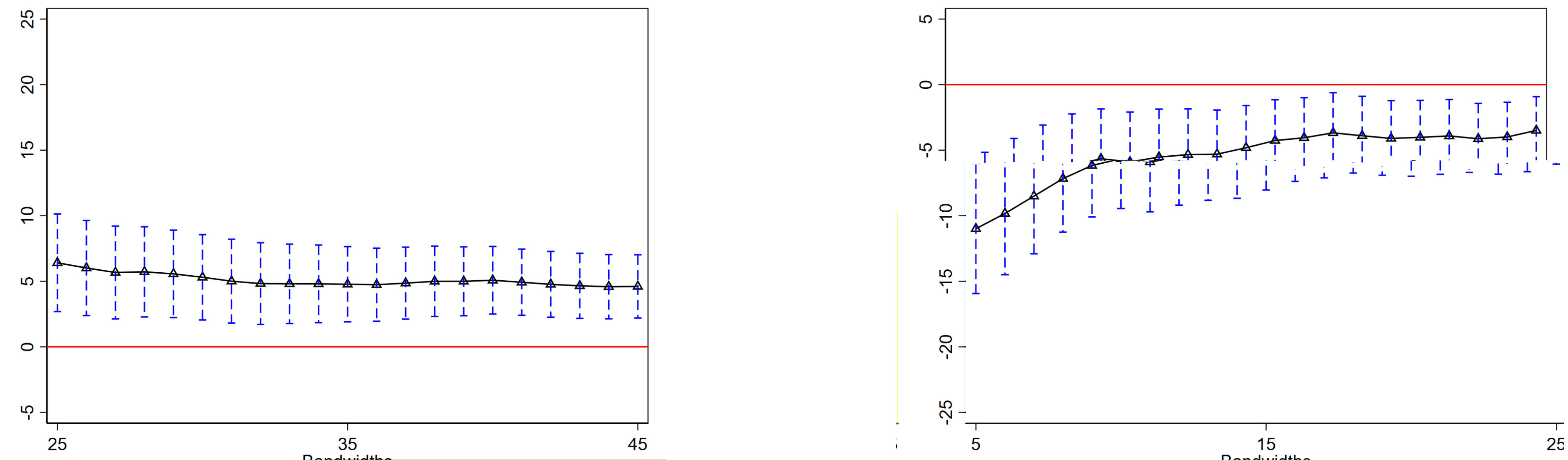
Panel D Fulfillment Year (Quadratic Polynomials)



Notes: This figure presents the graphic evidence from the baseline RDD setting. The RDD polynomials are specified as local linear in Panels A and B, and as quadratic in Panels C and D. The hollow circles represent bin means, while the solid and dashed lines indicate fitted lines accompanied by 95% confidence intervals.

Figure 6. Robust Checks of Bandwidth Selection

Panel A Failure Year *Panel B Fulfillment Year*



[redacted]

Notes: This graph shows point estimates along with 95% confidence intervals from RDD under different bandwidths. We experiment on bandwidths within a range with a length of 10 which is centered around the bandwidth in baseline setting (i.e., 35 and 15 for failure and fulfillment subsamples). The optimal bandwidths computed from the method proposed by Calonico et al. (2014) are also covered, which are 28 and 13 for each subsample, respectively. Standard errors are clustered at the prefecture level.

Table 1. Summary Statistics

	N	Mean	Std. Dev.	Min	Max
Panel A Prefecture-Year Variables					
Actual Percentage of Days with AQI $\leq$ 100	1554	79.93	15.94	19.73	100
Target Percentage of Days with AQI $\leq$ 100	1154	79.99	13.09	20	100
Completion Rates of Targets	1154	99.84	9.63	47.68	132.51
Panel B Prefecture-Day Variables					
Normalized AQI (NAQI)	538889	421.91	46.32	0	497
Daily Minimum Temperature	538889	10.6	11.49	-42	35
Daily Maximum Temperature	538889	20.07	10.92	-34	47
Dummy for Precipitation	538889	0.35	0.48	0	1
Dummy for Snow	538889	0.03	0.16	0	1
Category of Wind Speed	538889	1.43	0.7	1	4
Panel C Characteristics of Prefecture-Party Secretaries					
Dummy for Outside Promoted Bureaucrats	788	0.652	0.477	0	1
Dummy for Bureaucrats without Local Work Experience	788	0.59	0.492	0	1
Age of Incumbent Party Secretaries	788	52.25	2.99	41.67	58.58
Career-Concern Intensity	788	0.285	0.127	0	0.816
Panel D Firm-Day Variables					
SO_2 Emissions	542594	41.452	55.486	0	472.47
Panel E Characteristics of Firms					
Dummy for SOEs	428	0.117	0.322	0	1
Dummy for Public Service	427	0.482	0.500	0	1

Table 2. Regression Evidence for Punishment on Failing AQI Targets

Dependent Variable	*Being Named or Summoned for Administrative Talk in the Next Year*
Moderate Failure	0.067**
	(0.030)
Severe Failure	0.116**
	(0.049)
Prefecture Fixed Effects	YES
Year Fixed Effects	YES
Number of Observations	877

Notes: This table reports estimates from a prefecture-year level regression to provide supportive evidence for the incentive scheme of obligatory targets. The outcome variable is a dummy variable denoting whether the prefecture is summoned for an administrative talk or name-picked during the inspection in the next year. Explanatory variables are two dummies indicating whether the prefecture has failed by a moderate or large margin this year, with fulfillment cases as a control group. Standard errors are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.

Table 3. RDD Estimates and Robustness Checks

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent Variable	*Residual of Normalized AQI*					
Target Completion Status	*Failure*			*Fulfillment*		
Post	4.774	5.820	7.660	-4.814	-5.942	-8.015
	(1.456)***	(1.736)***	(2.447)***	(1.637)***	(1.746)***	(2.511)***
	[0.898]***	[1.045]***	[1.512]***	[1.126]***	[1.245]***	[1.928]***
Dep. Var. Mean	-1.779	-1.779	-1.779	0.576	0.576	0.576
Bandwidth	35	35	35	15	15	15
Kernel Function	Uniform	Triangular	Uniform	Uniform	Triangular	Uniform
RD Specification	LLR	LLR	Quadratic Polynomials	LLR	LLR	Quadratic Polynomials
Number of Observations	28701	28701	28701	15701	15701	15701

Notes: This table reports estimates from RDD estimation described by equations (1) and (2). "LLR" stands for local linear regression. Standard errors clustered at the prefecture level and heteroscedasticity-robust standard errors are reported in parentheses and brackets respectively with ***, **, * denoting significance at the 1%, 5%, 10% level.

Table 4. Effects of Pressures on Local Officials to Pursue Economic Performance

	(1)	(2)	(3)	(4)
Dependent Variables	*Residual of Normalized AQI*			
Target Completion Status	*Failure*		*Fulfillment*	
Subsample	*Low Burden of Economic Growth*	*High Burden of Economic Growth*	*Low Burden of Economic Growth*	*High Burden of Economic Growth*
Panel A Subsample Criterion: Above / Below the Median Growth Rate within the Province				
Post	6.164***	4.661**	-1.455	-7.822***
	(2.228)	(1.971)	(2.437)	(2.584)
Number of Observations	14878	12756	7123	6826
Panel B Subsample Criterion: Above / Below the Provincial Growth Rate				
Post	5.558**	4.166**	-2.062	-7.992***
	(2.152)	(2.076)	(2.382)	(2.623)
Number of Observations	16750	11335	7801	6988
Bandwidth	35	35	15	15
Kernel Function	Uniform	Uniform	Uniform	Uniform
RD Specification	LLR	LLR	LLR	LLR

Notes: This table reports results from subsample regressions to incorporate economic performance pressures into the analysis. In Panel A, a heavier economic performance pressure refers to the situation that the GDP growth rate of the city in the last quarter is below the median growth rate of all prefectures within the province. In Panel B, heavier economic pressure is defined as a situation where a prefecture's quarterly growth rate is below the provincial one, which is a weighted average of prefectural growth rates. Standard errors clustered at the prefecture level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level. Bootstrap tests are reported in C1

Table 5. Micro Evidence: Firms' Emission

	(1)	(2)
Dependent Variable	*Residual of SO_2 Emission*	
Target Completion Status	*Failure*	*Fulfillment*
Post	-1.808*	2.600**
	(0.971)	(1.007)
Dep. Var. Mean	-0.572	0.109
Bandwidth	35	15
Kernel Function	Uniform	Uniform
RD Specification	LLR	LLR
Number of Observations	43628	19089

Notes: This table reports estimates from RDD estimation following the specification employed in baseline analysis except that the data are granular at the firm-day level. Standard errors clustered at the firm level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.

Table 6. Micro Evidence: Heterogeneity Analysis

	(1)	(2)	(3)	(4)
Dependent Variables	*Residual of SO_2 Emission*			
Target Completion Status	*Failure*		*Fulfillment*	
Panel A Firms Ownership Types				
Subsample	*Non-SOEs*	*SOEs*	*Non-SOEs*	*SOEs*
Post	-1.434 (1.113)	-8.863* (4.614)	2.354** (1.093)	-2.173 (4.301)
Number of Observations	30372	4830	13420	1476
Panel B Firm Industries				
Subsample	*Public Service*	*Productive Firms*	*Public Service*	*Productive Firms*
Post	-1.030 (0.866)	-3.690** (0.780)	-0.336 (0.182)	4.299** (0.165)
Number of Observations	16268	18838	7591	7305
Panel C Initial Level of SO_2 Emissions				
Subsample	*Low Level*	*High Level*	*Low Level*	*High Level*
Post	-0.337 (1.059)	-3.843* (2.181)	0.728 (1.206)	4.208** (1.901)
Number of Observations	19704	16370	8064	8200
Bandwidth	35	35	15	15
Kernel Function	Uniform	Uniform	Uniform	Uniform
RD Specification	LLR	LLR	LLR	LLR

Notes: This table reports the results of heterogeneous analysis at the firm level. The firms specialized in public service including firms supplying water, electricity, and heating for local people and other firms managing public facilities. Standard errors clustered at the firm level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.

Table 7. Micro Evidence: Effects of Economic Performance Pressures

	(1)	(2)	(3)	(4)

Table 8. Effects of Intrinsic Motivations: Outsiders vs. Non-Outsiders

	(1)	(2)	(3)	(4)
Dependent Variable	*Residual of Normalized AQI*			
Target Completion Status	*Failure*		*Fulfillment*	
Subsample	*Outsiders*	*Non-Outsiders*	*Outsiders*	*Non-Outsiders*
Panel A Outsiders Defined According to Previous Working Location				
Post	6.714***	1.329	-6.517***	-1.465
	(1.871)	(2.263)	(2.095)	(2.538)
Number of Observations	18420	10281	10431	5270
Panel B Outsiders Defined According to Work Experience throughout the Career				
Post	5.681***	3.602	-6.509***	-2.398
	(1.940)	(2.176)	(2.219)	(2.404)
Number of Observations	16520	12181	9224	6477
Bandwidth	35	35	15	15
Kernel Function	Uniform	Uniform	Uniform	Uniform
RD Specification	LLR	LLR	LLR	LLR

Table 9. Effects of Political Incentive Intensities

	(1)	(2)	(3)	(4)
Dependent Variables	*Residual of Normalized AQI*			
Target Completion Status	*Failure*		*Fulfillment*	
Subsample	*Low-Incentive*	*High-Incentive*	*Low-Incentive*	*High-Incentive*
Panel A Political Incentive Measured by Age				
Post	0.998	7.002***	-4.540	-6.933***
	(3.184)	(2.241)	(5.604)	(2.424)
Number of Observations	3688	12832	1585	7639
Panel B Political Incentive Measured by Career-Concern Intensity				
Post	5.229**	6.222**	-4.029	-9.126***
	(2.410)	(3.007)	(3.122)	(3.125)
Number of Observations	8939	7581	4599	4625
Bandwidth	35	35	15	15
Kernel Function	Uniform	Uniform	Uniform	Uniform
RD Specification	LLR	LLR	LLR	LLR

Notes: This table reports results from subsample regressions to shed light on the role of political incentives. In Panel A, we follow the method proposed by Yu et al. (2016) and He et al. (2020) and define the high-incentive group as those younger than 57, the cutoff age for a discontinuous drop in promotion probability. In Panel B, the criterion for splitting the subsample is whether the career-concern intensity is above the median of all incumbent party secretaries within the same province in that year. The career concern intensity is computed according to Wang et al. (2020). Standard errors clustered at the prefecture level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.

Table 10. Effects of Central Supervision

	(1)	(2)	(3)	(4)
Dependent Variables	*Residual of Normalized AQI*			
Target Completion Status	*Failure*		*Fulfillment*	
Subsample	*Loosely Supervised*	*Strictly Supervised*	*Loosely Supervised*	*Strictly Supervised*
Panel A Full Sample Regression				
Post	3.352	6.265***	-5.355**	-3.428
	(2.214)	(1.841)	(2.070)	(2.243)
Number of Observations	15347	13354	11284	4417
Panel B Subsample Regression Using Outsiders				
Post	3.214	8.209***	-7.372**	-4.874*
	(2.981)	(2.566)	(3.076)	(2.644)
Number of Observations	8765	7755	6093	3131
Panel C Subsample Regression Using Non-Outsiders				
Post	3.508	3.860	-3.027	0.068
	(3.211)	(2.779)	(2.762)	(4.172)
Number of Observations	6582	5599	5191	1286
Bandwidth	35	35	15	15
Kernel Function	Uniform	Uniform	Uniform	Uniform
RD Specification	LLR	LLR	LLR	LLR

Notes: This table reports results that examine the effects of supervision. The strictly supervised subsample refers to those cities closely watched by the central government, which were selected by MEE in 2013. The number of cities in this group was expanded from 79 to 164 in the second half year of 2018. Standard errors clustered at the prefecture level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.

Appendix A Mathematical Appendix

Proof of Proposition 1

Proof. In period 2, if the state is revealed to be good, then the agent solves the following maximization problem:

$$h_g(x_1, z_1) = \max_{x_2, z_2}\{f(x_1, x_2)v - C(x_1 + x_2, z_1 + z_2)\}.$$

Obviously, the optimal solution satisfies $z_2^g = 0$. The first derivative with respect to x_2 is

$$f'(x_1 + x_2)v - c_x(x_1 + x_2) - kz_1.$$

If there is an interior optimal solution x_2^g, then the first order condition implies:

$$f'(x_1 + x_2^g)v - c_x(x_1 + x_2^g) - kz_1 = 0.$$

Otherwise, the optimal x_2^g is the corner solution $x_2^g = 0$ and hence we should have

$$f'(x_1)v - c_x x_1 - kz_1 \leq 0.$$

Moreover, from the envelope theorem, we have:

$$\frac{\P h_g}{\P x_1}(x_1, z_1) = f'(x_1 + x_2^g)v - c_x(x_1 + x_2^g) - kz_1 \leq 0$$

and

$$\frac{\P h_g}{\P z_1}(x_1, z_1) = -c_z z_1 - k(x_1 + x_2^g) \leq 0.$$

In period 2, if the state is revealed to be bad, then the agent solves the following maximization problem:

$$h_b(x_1, z_1) = \max_{x_2, z_2}\{-p(\Delta_y)u - C(x_1 + x_2, z_1 + z_2)\}.$$

Obviously, the optimal solution $x_2^b = 0$ and the first derivative with respect to z_2^b is:

$$p'u - c_z(z_1 + z_2^b) - kx_1.$$

We can also apply the envelope theorem, and obtain:

$$\frac{\P h_b(x_1, z_1)}{\P x_1} = -2p'u - c_x x_1 - k(z_1 + z_2^b) < 0;$$

$$\frac{\P h_b(x_1, z_1)}{\P z_1} = 2p'u - c_z(z_1 + z_2^b) - kx_1.$$

Finally, in period 1, the agent solves the following problem:

$$\max_{x_1, z_1}\{ah_g(x_1, z_1) + (1 - a)h_b(x_1, z_1)\}.$$

We have already shown that $x_2^b = z_2^g = 0$. To prove $x_1 = 0$, denote $\Omega(x_1, z_1) = ah_g(x_1, z_1) + (1-a)h_b(x_1, z_1)$. From the envelope theorem, we obtain the first derivative of Ω with respect to x_1:

$$\frac{¶\Omega(x_1, z_1)}{¶x_1} = a\left[f'(x_1 + x_2^g)v - c_x(x_1 + x_2^g) - kz_1\right] - (1-a)\left[2p'u + c_x x_1 + k(z_1 + z_2^b)\right] < 0, \quad \forall x_1 \geq 0.$$

Therefore, it is always optimal to set $x_1 = 0$. Moreover, $x_1 = 0$ implies the first derivative of Ω with respect to z_1:

$$\frac{¶\Omega(x_1, z_1)}{¶z_1} = -a(c_z z_1 + kx_2^g) + (1-a)\left[2p'u - c_z(z_1 + z_2^b)\right].$$

Suppose by contradiction that the optimal $z_2^b = 0$. This implies that $p'u - c_z z_1 \leq 0$ or $2p'u - c_z z_1 \leq c_z z_1$. As a result, $a \geq \frac{1}{2}$ implies that

$$\frac{¶\Omega(x_1, z_1)}{¶z_1} \leq -a(c_z z_1 + kx_2^g) + (1-a)c_z z_1 \leq 0, \quad \forall z_1 \geq 0.$$

So, it is always optimal to set $z_1 = 0$ as well. But then the first derivative with respect to z_2^b becomes $p'u > 0$, a contradiction! Therefore, we conclude that $z_2^b > 0$. □

Proof of Proposition 2

Proof. If $k = 0$, the first derivative with respect to x_2^g becomes $f'(x_2^g)v - c_x x_2^g > 0$ at $x_2^g = 0$. Since the inequality is strict, it still holds for k sufficiently small. Hence, we have $x_2^g > 0$ for k sufficiently close to zero.

Similarly, if $k = 0$,

$$\frac{¶\Omega(x_1, z_1)}{¶z_1} = -ac_z z_1 + (1-a)c_z(z_1 + z_2^b) > 0$$

at $z_1 = 0$. So, we conclude that $z_1 > 0$ for k sufficiently close to zero. □

Proof of Proposition 3

Proof. If the optimal solutions z_1, x_2^g and z_2^b are interior, the first order conditions imply that:

$$f'(x_2^g)v - c_x x_2^g - kz_1 = 0;$$

$$p'u - c_z(z_1 + z_2^b) = 0;$$

and

$$-a(c_z z_1 + kx_2^g) + (1-a)c_z(z_1 + z_2^b) = 0.$$

When $k = 0$, x_2^g is purely determined by the equation $f'(x_2^g)v = c_x x_2^g$ while (z_1, z_2^b) are determined by equations $p'u = c_z(z_1 + z_2^b)$ and $(2a-1)z_1 = (1-a)z_2^b$. Based on the assumptions of f, it is straightforward to see that $\frac{¶x_2^g}{¶v} > 0$, and $\frac{¶z_2^b}{¶v} = 0$.

When $k > 0$, total differentiation with respect to the first order conditions yields that

$$-(c_x - f''v)\frac{c_z}{ak}\frac{(ac_z + p''u)dz_1}{c_z + p''u} + kdz_1 = f'dv.$$

The assumptions of $f'' \leq 0$ and $p'' \geq 0$ together with the assumption $k < \sqrt{c_x c_z}$ imply that

$$k - (c_x - f''v)\frac{c_z}{ak}\frac{ac_z + p''u}{c_z + p''u} < 0,$$

and hence $\frac{\P x_1}{\P v} < 0$. Consequently, we obtain $\frac{\P x_2^g}{\P v} > 0$, and $\frac{\P z_2^b}{\P v} > 0$. □

Proof of Proposition 4

Proof. With a little abuse of notations, we define

$$h_g(x_1, z_1) = \max_{x_2, z_2}\{f(x_1 + x_2)v + q[j(2(z_1 - x_1) + z_2 - x_2)] - C(x_1 + x_2, z_1 + z_2)\};$$

$$h_b(x_1, z_1) = \max_{x_2, z_2}\{-p(\Delta_y)u + q[j(2(z_1 - x_1) + z_2 - x_2)] - C(x_1 + x_2, z_1 + z_2)\};$$

and

$$\Omega(x_1, z_1) = ah_g(x_1, z_1) + (1 - a)h_b(x_1, z_1).$$

Obviously, the optimal solution x_2^b should be zero. From the envelope theorem, it is straightforward to see that $\frac{\P h_b(x_1,z_1)}{\P x_1} < 0$ and $\frac{\P h_g(x_1,z_1)}{\P x_1} < 0$. Therefore, $\frac{\P \Omega(x_1,z_1)}{\P x_1} < 0$ and it is optimal to set $x_1 = 0$.

Now suppose by contradiction that z_2^g is interior. There are two cases to consider. In the first case, both z_2^g and z_2^b are interior. Then the first order conditions imply:

$$q\,j' = c_z(z_1 + z_2^g) + kx_2^g \quad \text{and} \quad p'u + q\,j' = c_z(z_1 + z_2^b). \tag{3}$$

Apply the envelope theorem and we can get:

$$\frac{\P h_g(x_1, z_1)}{\P z_1} = 2q\,j' - c_z(z_1 + z_2^g) - kx_2^g > 0; \tag{4}$$

and

$$\frac{\P h_b(x_1, z_1)}{\P z_1} = 2(p'u + q\,j') - c_z(z_1 + z_2^b) > 0. \tag{5}$$

Therefore, we must have

$$\frac{\P \Omega(x_1, z_1)}{\P z_1} > 0, \quad \forall z_1. \tag{6}$$

Hence, the optimal z_1 goes to infinity. But this cannot be optimal as it implies a payoff of $-\infty$.

In the other case, $z_2^b = 0$, then we can similarly apply the envelope theorem and get

$$\frac{¶\Omega(x_1, z_1)}{¶z_1} = a\left[c_z(z_1 + z_2^g) + kx_2^g\right] + (1-a)\left[2(p'u + q\,j') - c_z\right.$$

and

$$z_2^b = \frac{I_p u + I_j q}{c_z} - z_1.$$

It is straightforward to see that z_1 is linearly increasing in q while both x_2^g and z_2^b are linearly decreasing in q. Therefore, when q becomes sufficiently large, either x_2^g or z_2^b has to be zero.

If the optimal solutions satisfy $x_2^g = 0$ and $z_2^b > 0$, then the first derivative with respect to z_1 is strictly positive at $z_1 = 0$. Hence, we get

$$a(2q I_j - c_z z_1) + (1-a)c_z(z_1 + z_2^b) = 0;$$

and

$$u I_p + q I_j - c_z(z_1 + z_2^b) = 0.$$

The solutions are

$$z_1 = \frac{(1+a) I_j q + (1-a) I_p u}{c_z}$$

and

$$z_2^b = \frac{I_p u + I_j q}{c_z} - z_1.$$

It is also the case that z_1 is linearly increasing in q while z_2^b is linearly decreasing in q. Therefore, when q becomes sufficiently large, z_2^b has to be zero as well.

If the optimal solutions satisfy $x_2^g > 0$ and $z_2^b = 0$, we also get $z_1 > 0$ because otherwise the first derivative with respect to z_2^b is strictly positive at $z_2^b = 0$. Hence, we solve the following system of equations:

$$a(2q I_j - c_z z_1 - k x_2^g) + (1-a)(2u I_p + 2q I_j - c_z z_1) = 0;$$

and

$$v I_f - q I_j - c_x x_2^g - k z_1 = 0.$$

It is straightforward to see that z_1 is linearly increasing in q while x_2^g is linearly decreasing in q. As a result, we conclude that when q becomes sufficiently large, it must be the case that $z_1 > 0$ and $x_2^g = z_2^b = 0$.

In the second case, $k \geq \frac{1-a}{a} \frac{I_p u}{I_f v} cx$ and hence $z_1 = 0$ when $q = 0$. So we solve the following system of equations when q is sufficiently close to zero:

$$v I_f - q I_j - c_x x_2^g = 0;$$

and

$$u I_p + q I_j - c_z z_2^b = 0.$$

Notice that the optimal solution

$$x_2^g = \frac{v I_f - q I_j}{c_x}$$

is linearly decreasing in q and hence $x_2^g = 0$ when q is sufficiently large. Once $x_2^g = 0$, the first derivative with respect to z_1 is strictly positive at $z_1 = 0$ and hence the optimal solution z_1 becomes interior. Then we are back to the first case discussed above, under which $z_1 > 0$ and $x_2^g = z_2^b = 0$ as we keep increasing q. $\square$

Proof of Proposition 6

Proof. In the new problem, the first derivative with respect to z_2^g is $I_g - c_z(z_1 + z_2^g) \leq 0$; the first derivative with respect to z_2^b is $I_p u + I_b - c_z(z_1 + z_2^b) \leq 0$; and hence the first derivative with respect to z_1 satisfies

$$-ac_z(z_1 + z_2^g) + (1-a)(2I_p u - c_z(z_1 + z_2^b)) \leq -aI_g + (1-a)(I_p u - I_b) < 0.$$

Thus, it is optimal to set $z_1 = 0$. Then, we have the optimal $z_2^g = \frac{I_g}{c_z} > 0$, and the optimal $z_2^b = \frac{I_p u + I_b}{c_z}$, which is larger than the one in the benchmark model. Meanwhile, the optimal x_2^g decreases to $\frac{I_f v - I_g}{c_x}$ and the optimal x_2^b stays at zero. Hence, both $z_2^g - x_2^g$ and $z_2^b - x_2^b$ become larger. $\square$

Appendix B Rules of Unifying Target Presentations

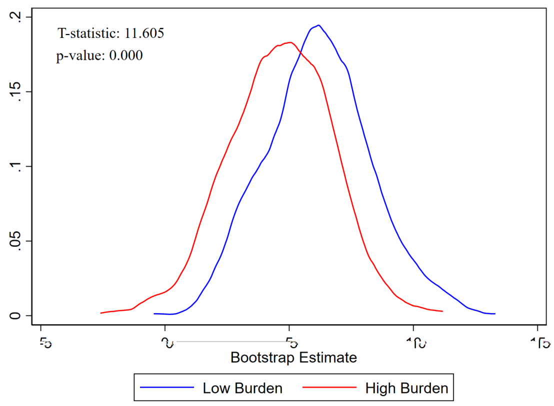
Table B1. Rules of Unifying Target Presentations

Expression of Targets	Rules of Presentation Transfer	Notes
The number of days with good or moderate air quality in prefecture c and year y should be no less than X.	X	"# of days with good or moderate air quality" is the default expression and does not need adjustment.
The percentage of days with good or moderate air quality in prefecture c and year y should be no less than W%.	$W\% \times 365$	365 would be replaced by 366 in 2016.
The number of days with good or moderate air quality in prefecture c and year y should further increase by V days on the basis of year y-1 (assuming the actual number in year y-1 to be U).	U + V	The actual number of days with good or moderate air quality in prefecture c and year y-1 is computed from air quality data and cross-validated with officially published local environmental bulletin.
The number of days with good or moderate air quality in prefecture c should increase to Q in y+5, on the basis of year y, with yearly progress no slower than 20%, 40%, 60%, 80% and 100% (assuming the actual number in year y to be P).	P + (Q-P)/5$\times$k (k = 1,2,3,4,5)	The actual number of days with good or moderate air quality in city c and year y is computed from air quality data and cross-validated with officially published local environmental bulletin. Targets for year y+1, y+2,..., y+5 are implied by the left equation with k = 1,2,..., 5 respectively.

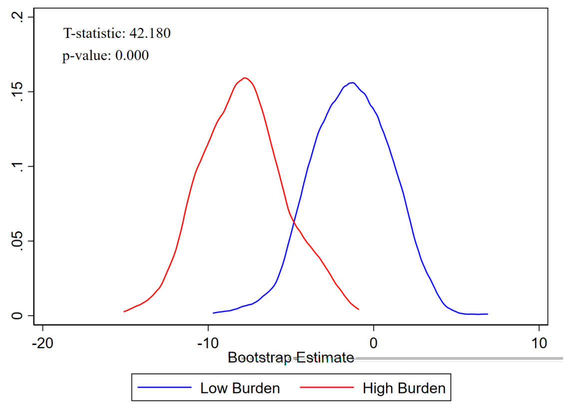
Appendix C Results from Bootstrap Tests

Figure C1. Bootstrapped Estimates for Different Economic Burden

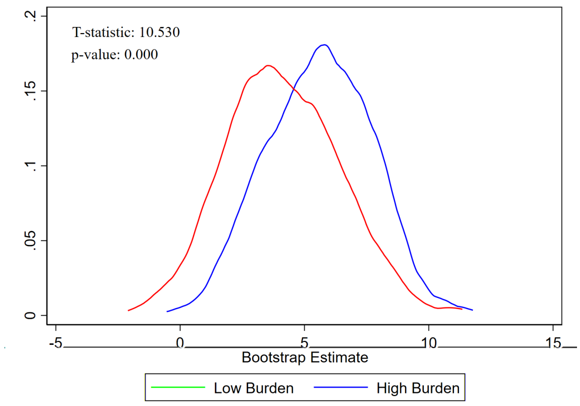
Panel A Failure Year (Criterion 1)



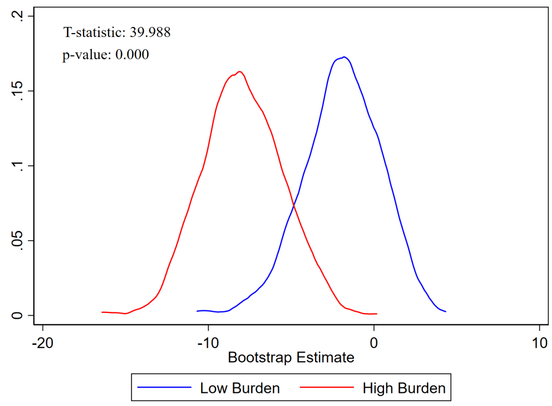
Panel B Fulfillment Year (Criterion 1)



Panel C Failure Year (Criterion 2)



Panel D Fulfillment Year (Criterion 2)



Appendix D Evidence for the Effects of Bureaucrats' Local Experience on Environmental Protection

In this appendix, we aim to provide more evidence to corroborate the linkage between the extent of prefecture officials' local experience (as a proxy for bureaucrats' intrinsic motivations for local environmental quality à la Persson and Zhuravskaya (2016)) and their efforts devoted to environmental protection. Specifically, we collected two sets of outcome variables that may reflect officials' differential preferences over the local environment and examined their correlation with the duration of officials' local experience. The first one is the intensity of policies on environmental regulation, which is constructed from manually collected annual reports on the work of prefectural governments. Following the same method as Chen et al. (2018a) and Chen et al. (2018b), we search for a list of keywords related to environmental protection in the transcripts of government work reports and identify sentences containing these words as environmental-regulation-related content.[35] Then we compute the total number of words in these sentences as a proxy for the intensity of attention to environmental protection. The second set includes the green coverage rate and completed investment in landscaping at the prefecture-year level, which are extracted from *China Urban Construction Statistical Yearbook* and depict the effectiveness of and expenditure in environmental policies, respectively.

All estimates in Table D1 collectively deliver similar findings. That is, prefectural leaders with local work experience (i.e., non-outsiders) exhibit a stronger preference for environmental regulation, regardless of the specific outcome variables. Furthermore, if we decompose the dummy for non-outsiders into two separate ones according to the length of local experience, those with a longer duration of local experience seem to pay more attention to environmental protection than those with shorter local experience. We have controlled for prefecture fixed effects, province-year fixed effects, personal characteristics of party secretaries, and time-variant city attributes in those regressions. Those results provide supportive evidence that the distinction between outsiders and non-outsiders at least partially captures some heterogeneity of local officials in intrinsic motivations for environmental protection.

[35] We experiment with two different lists of keywords. First, we follow the practice in the existing literature and utilize a relatively narrow list: *low carbon (Ditan)*, *energy consumption (Nenghao)*, *protecting the environment (Baohu Huanjing)* and *environmental protection (Huanjing Baohu)*. Considering that we use data as late as 2019 and there have been some newly invented political expressions, we further expand the list and add *environmental law enforcement (Huanjing Zhifa)*, *lucid waters and lush mountains (Lvshui Qingshan)* and *beautiful China (Meili Zhongguo)*.

Table D1. Bureaucrats' Heterogeneity in Intrinsic Motivations for Environmental Protection

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Dependent Variable	*Log (Size of Sentences Containing Key Words)*		*Log (Size of Sentences Containing Key Words)*		*Green Coverage Rate of Built District*		*Completed Investment in Landscaping*	
Non Outsider	0.278*		0.275**		1.181***		0.825	
	(0.153)		(0.118)		(0.387)		(0.529)	
Non Outsider-Weak		0.206		0.257*		1.123**		-0.094
		(0.197)		(0.148)		(0.526)		(0.565)
Non Outsider-Strong		0.377**		0.300**		1.246***		1.861**
		(0.154)		(0.132)		(0.447)		(0.800)
Dep. Var. Mean	4.772	4.772	5.081	5.081	37.723	37.723	4.125	4.125
Control Variables	YES	YES	YES	YES	YES	YES	YES	YES
Prefecture Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Province-Year Fixed Effects	YES	YES	YES	YES	YES	YES	YES	YES
Number of Clusters	296	296	296	296	273	273	273	273
Number of Observations	1142	1142	1142	1142	819	819	819	819

Notes: (1) Outcome variables in the first four columns are logarithmic total words of sentences that contain the keywords listed in footnote 35, with Columns (1) and (2) for the narrowly defined version and (3) and (4) for the extended list. (2) *Non-Outsider* is a dummy variable indicating that the incumbent party secretary has worked in the prefecture before. We further categorize *Non-Outsider* into two separate dummies indicating whether the work experience is longer than the median (strong) or not (weak). (3) We control for the following personal characteristics in all columns: a dummy for a native, age, a dummy for older than 57, educational level, and dummies for work experience in departments of provincial or central governments. We further include the logarithmic total number of words of government reports in Columns (1) - (4) and population density in Columns (5) - (8) to account for potential scale effects on outcome variables. Standard errors clustered at the prefecture level are reported in parentheses with ***, **, * denoting significance at the 1%, 5%, 10% level.