Online Appendix

Growth at the Margin: Political Incentives and Firm Behavior in China

Kefan Chen

Department of Economics, Lingnan University

Appendix A Supplementary Data Details

Appendix A.1 Growth Targets and Actual Growth

Data on China's economic growth targets are compiled from publicly available government documents and statistical records. Central growth targets are outlined in key documents such as the *Five-Year Plan for National Economic and Social Development*[1] and the annual *Report on the Work of the Government*[2], which serve as official declarations of national macroeconomic priorities.[3] These sources, available on government websites, offer detailed information on annual GDP growth targets, fiscal revenue goals, social development objectives, and other related metrics. At the provincial and prefectural levels, growth targets are

[1]First introduced in 1953, these plans are developed by the National Development and Reform Commission (NDRC) and approved by the National People's Congress (NPC). Each plan sets specific goals for economic growth, industrial restructuring, technological innovation, environmental protection, and social welfare. While originally a hallmark of China's centrally planned economy, the Five-Year Plans have evolved into a hybrid framework, blending market-driven strategies with state-led guidance. They provide a policy roadmap for all levels of government, businesses, and investors, ensuring that national objectives are aligned with local implementation efforts. The plans also reflect China's long-term vision, balancing short-term economic goals with sustainable development and societal progress.

[2]Typically delivered by local city governors during the annual session of the local People's Congress, these reports provide a comprehensive review of the government's achievements over the past year and outline the policy priorities, economic targets, and social development goals for the year ahead. Key areas covered include GDP growth, employment, fiscal policies, environmental protection, and technological innovation. Serving as a guiding framework, the report aligns the activities of government agencies, regional authorities, and state-owned enterprises with national objectives. Often regarded as a barometer of China's economic and political direction, it reflects the broader priorities of the central government while addressing regional challenges.

[3]These documents are typically published early in the year, usually between January and February. Higher-level government work reports are released before those of lower-level governments, ensuring that growth targets are set in advance and reflect the priorities of higher authorities. This sequencing allows local targets to be established after observing the objectives outlined by upper-level leadership.

extracted from local *five-year and annual plans*, which adapt national objectives to region-specific contexts. These documents are publicly released as finalized reports, mirroring the publication process at the central level. The data collection aims to compile a nationwide dataset spanning 20 years. This process involves manually extracting information from official documents and websites of 339 prefecture-level cities across all 32 provincial-level administrative units in China, covering the past two decades.

Economic performance data supplement these growth targets, providing realized growth rates, per capita GDP, and other key indicators. These are sourced from the *China City Statistical Yearbook*[4] and regional statistical publications, enabling comparisons between targeted and achieved economic outcomes across national, provincial, and prefectural levels.

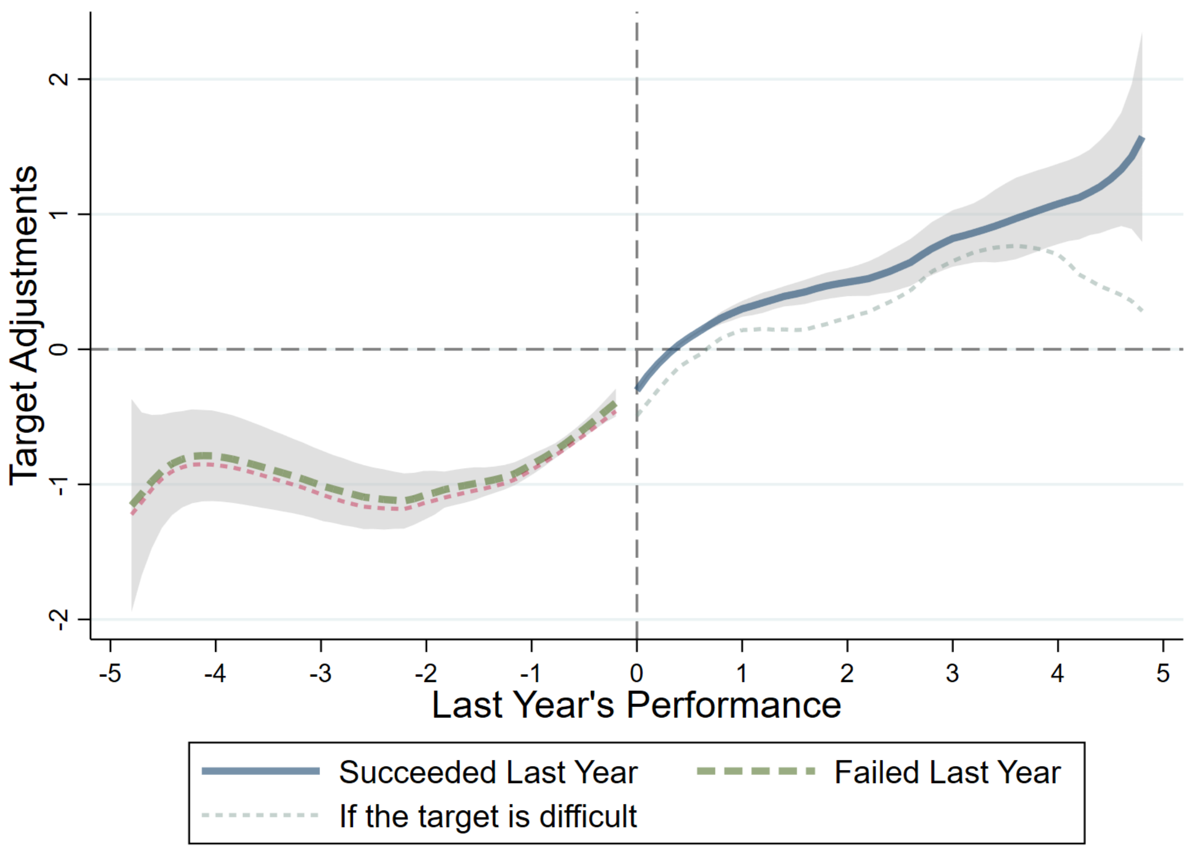
Appendix A.1.1 City Party Secretaries

The dataset on Chinese city party secretaries[5], including their demographics and promotion outcomes, was meticulously compiled from official government websites and news platforms such as *Xinhua Net*, The dataset covers two decades of leadership transitions across the previously mentioned 339 cities, including observations on over 1,800 party secretaries. This dataset captures key demographic and professional attributes, such as age, gender, place of

[4] An authoritative publication that provides comprehensive statistical data on the social and economic development of cities across China. Published annually by the National Bureau of Statistics of China. It covers a wide range of indicators, including population demographics, GDP, industrial output, infrastructure development, education, public health, environmental protection, and fiscal revenue. The yearbook organizes data by city and year, enabling detailed comparisons across different regions and time periods.

[5] The decision to use data on party secretaries rather than mayors stems from the fundamental structure of China's political system, where the Chinese Communist Party (CCP) wields authority over all levels of government. As the top-ranking CCP official in a city, the party secretary holds ultimate decision-making power, particularly in critical areas such as political appointments, policy priorities, and resource allocation. In contrast, the mayor, while responsible for the city's day-to-day administration and economic management, operates under the direct oversight of the party secretary. This makes the party secretary the most influential figure in shaping city-level outcomes.

Figure A1: Target Ratcheting



Notes: This figure plots current-year growth targets against prior-year performance, specifically whether the previous target was met or missed. This shows that cities with stronger past performance tend to set higher targets for the following year. When a city meets or exceeds its target, the subsequent year's target is typically adjusted upward—albeit conservatively, usually by less than half of the previous year's excess. In contrast, downward adjustments following a missed target are smaller and more constrained. Notably, there is no discontinuity in target adjustment at the point where the prior-year target was just met, suggesting that the observed bunching at the target threshold is more likely driven by discontinuities in promotion probabilities rather than in the target-setting process itself.

origin, education level[6], tenure length[7] prior positions, and career experiences. For simplicity, party secretaries will be referred to as governors throughout the remainder of this paper.

The centerpiece of the dataset is information on officials' career trajectories and promotion outcomes. In this context, a promotion is defined as an increase in an official's rank within the Communist Party hierarchy,[8] for example, advancing from rank r to $r + 1$. This paper leverages the *Regulations on the Management of Civil Servants' Positions and Ranks (2006)* to determine the administrative rank associated with each position held at different career points. Promotions are specifically examined in relation to the official's next position after serving as a city governor, with a focus on whether their advancement is linked to the city's economic performance during their tenure. At the end of a governor's tenure, their career outcomes fall into one of three categories[9]:

[6] Chinese officials typically acquire their education through three primary pathways. The first is traditional academic study, culminating in a degree upon graduation. The second involves earning a degree while concurrently holding an official position; these selective part-time programs are not open to the general public. The third pathway is through party schools, which exclusively admit Communist Party members and primarily focus on party-related knowledge rather than standard academic curricula. An official's educational background may include a combination of these pathways, but degrees obtained through traditional academic study are generally considered the most reflective of an individual's capabilities. Nevertheless, this paper does not delve into these complexities and instead uses the highest degree obtained by the official, regardless of whether it was earned through part-time study or from a party school.

[7] In China, the official term for party secretaries is five years, but in practice, the average tenure in the sample is less than four years due to frequent political rotations and reshuffles. The timing of promotion or reassignment is uncertain, as no fixed legal framework dictates these transitions. Instead, officials are often reassigned or promoted under the party's meritocratic system, which emphasizes career advancement based on performance, typically measured by economic and political achievements.

[8] Some studies have adopted alternative definitions of promotion, including transitions to more influential roles within the same rank (Zeng and Zhou, 2024). These roles are considered important because they significantly increase the likelihood of subsequent promotion. However, this paper retains the most direct definition of promotion—advancement to a higher rank—due to the subjective and potentially contentious nature of defining "more important" positions.

[9] There is indeed a fourth category includes officials who are removed from their positions due to criminal convictions or expulsion from the Communist Party. While the frequency of such cases has increased in recent years, these instances were rare during the sample period analyzed in this paper and are therefore excluded from consideration.

$$
\text{Promotion} = \begin{cases} 1 & \text{if promoted to a higher administrative rank,} \\ 0 & \text{if transferred to an equivalent-level position,} \\ -1 & \text{if retired (typically due to reaching the retirement age).} \end{cases}
$$

Thus, Promotion = 1 indicates a rise by at least one administrative rank after the end of their term, Promotion = 0 represents lateral transfers[10], and Promotion = −1 reflects retirement [11] The appendix includes an example of an official's resume, illustrating the typical format and details encountered during data collection.

Appendix A.1.2 One Example CV

[10]In China, it is rare for officials to exit the government job market after completing a term in office. Government officials typically have limited or no "outside options" and, are usually reassigned to positions of equivalent rank if not promoted. Similarly, demotions are uncommon. The most likely scenario for sidelined officials is reassignment to less influential departments at the same rank, effectively marginalizing them. This reassignment can be interpreted as a form of demotion or a step toward retirement.

[11]Officials reassigned to positions in the National People's Congress (NPC) or the Chinese People's Political Consultative Conference (CPPCC) after reaching a certain age are often considered effectively retired, despite retaining their rank and benefits (Qiao, 2013). Accordingly, this paper categorizes such officials as retired rather than assigning them to other classifications.

Figure A2: Example CV

Personal Résumé

Education

- **1985.07–1989.08**: Student of Chinese Language and Literature at Changchun Normal College
- **1989.08–1990.11**: Teacher at Tonghua Normal School
- **1990.11–1991.10**: Deputy Secretary to the Youth League at Changchun Normal College
- **1991.10–1994.09**: Secretary to the Youth League at Changchun Normal College

Rank 1
- **1996.04:** Joined the Chinese Communist Party

Rank r
- **2001.08-2008.03:** The Standing Committee Member of the Liaoyuan Municipal Party Committee, Jilin Province

Rank r
- **2008.03-2010.04:** The Standing Committee Member of the Songyuan Municipal Party Committee, Jilin Province

Rank r+1
- **2010.04-2011.03:** The Deputy Secretary of the Songyuan Municipal Party Committee, Jilin Province

Rank r+1
- **2012.09-2015.09:** The Director of the Department of Human Resources and Social Security of Jilin Province

Rank r+1
- **2015.09-2019.09:** The Secretary of the Baishan Municipal Party Committee, Jilin Province

Rank r+2
- **2019.09-2020.06:** The Vice Governor of the People's Government of Jilin Province

Appendix A.1.3 CIED Financial Statements

This paper's micro-level firm data primarily comes from the China Industrial Enterprise Database (CIED), a comprehensive dataset maintained by the National Bureau of Statistics of China. It includes all state-owned industrial enterprises and non-state-owned enterprises with annual revenue above specific thresholds—5 million RMB prior to 2011 and 20 million RMB afterward. The dataset primarily focuses on three sectors: mining, manufacturing, and the production and supply of electricity, gas, and water, with manufacturing enterprises accounting for over 90% of the sample. Covering the period from 1998 to 2014, it comprises millions of observations, forming a substantial but unbalanced panel dataset[12]. The database contains two broad categories of information: (1) basic firm characteristics, such as enterprise codes, names, ownership types, addresses, and workforce details, and (2) financial and operational variables, including total assets, fixed assets, sales revenue, value-added, R&D expenditure, and profits. With approximately 130 indicators, the CIED provides a rich resource for analyzing firm performance, productivity, and behavior across various dimensions.

Despite its inherent limitations,[13] the CIED database remains one of the few widely used datasets for studying micro-level manufacturing firms in China. To address its shortcomings, this paper complements the analysis with data from the China Stock Market and Accounting Research (CSMAR) database.[14] Although the CSMAR database includes a smaller number

[12]Note that the databases on economic targets and official information cover the period from 2000 to 2021, spanning a longer timeframe compared to the firm-level dataset.

[13]While widely regarded as a valuable resource for microeconomic and firm-level research, the database has several notable limitations that researchers must address (Brandt et al., 2014). First, it suffers from inconsistencies in identifiers, such as enterprise codes and names, complicating the construction of accurate longitudinal panel data. Frequent changes due to restructuring or renaming can result in overestimating the number of unique firms or failing to match records for the same firm over time. Second, the database contains significant data gaps and variable inconsistencies. Furthermore, discrepancies in the definitions and measurements of variables like "capital" can introduce measurement errors and affect research outcomes. Third, the dataset includes outliers and anomalous values in variables such as profits, assets, and production, requiring rigorous cleaning and filtering to ensure reliability. Lastly, a selection bias is inherent in its design, as it covers only state-owned enterprises and non-state enterprises exceeding the revenue threshold, excluding smaller firms. This exclusion limits its representation of China's industrial sector, particularly for studies on market dynamics or firm heterogeneity.

[14]The CSMAR database is a comprehensive dataset that focuses on publicly listed firms in China, offering detailed information on financial performance, corporate governance, and stock market indicators. Compared to the CIED, CSMAR provides higher data quality, with fewer missing values and greater consistency, owing

of firms, it offers significantly greater precision and a broader range of variables, making it a valuable complement to the analysis. This paper follows the data-matching approach outlined by Brandt et al. (2012)[15] and adopts the cleaning methods proposed by Lam et al. (2017) to construct a coherent panel dataset.[16] The final sample comprises over 200,000 distinct firms and more than 600,000 firm-year observations.[17]

Appendix A.1.4 Energy Consumption and Pollution Records

To measure firms' genuine economic activities, this paper relies on data on industrial pollution and energy consumption, primarily sourced from the China Industrial Enterprise Pollution Database (CIEPD). This database records information on 27 types of industrial pollutants, energy consumption, and pollution treatment indicators. Compiled with approval from the National Bureau of Statistics and designed by the Ministry of Environmental Protection, it focuses on state-owned enterprises and large non-state-owned industrial firms. The dataset offers comprehensive firm-year indicators, covering raw material usage, energy consumption, solid waste, gaseous emissions, and water pollutants. For each pollutant type, it provides detailed information on emission quantities and treatment status, offering valuable insights into firms' environmental impact and mitigation efforts. Notably, the China

to its emphasis on publicly traded companies subject to stricter regulatory oversight. However, its primary limitation lies in its narrower scope, as it excludes private and smaller firms, making it less representative of China's broader industrial landscape.

[15]Firms are matched using a combination of identifiers, including enterprise codes, names, legal representatives, addresses, postal codes, industry codes, main products, administrative districts, and establishment years. This methodology ensures consistency by reconnecting firms that underwent restructuring or reorganization, resulting in changes to enterprise codes or names, thereby preserving continuity across transitions. Industry codes are also standardized to ensure compatibility across different classification systems. Specifically, the 1994 GB/T4754-1994 classification is aligned with the 2002 GB/T4754-2002 standard, and the 2013 classification is similarly mapped to the 2002 standard. This harmonization enables consistency across different time periods and enhances the comparability of industry-level data.

[16]Samples with missing or fewer than 10 employees were excluded from the analysis. Observations inconsistent with generally accepted accounting principles, such as those with profit margins exceeding 1 or negative net fixed asset values, were also removed. Additionally, samples missing key financial indicators, such as industrial output value or industrial sales value, were excluded. Following Lam et al. (2017), the 2010 data, which has been widely questioned for its reliability, was omitted from the study. To maintain continuity, 2009 and 2011 were treated as consecutive years in the analysis. Continuous explanatory variables were winsorized at the 1st and 99th percentiles to mitigate the influence of extreme outliers.

[17]However, 30% of the firms have only one year of observation, while the longest observed firms have up to 16 years of data.

Industrial Enterprise Pollution Database (CIEPD) can be matched with the China Industrial Enterprise Database (CIED), enabling the construction of firm-year observations that integrate financial indicators with detailed data on pollutants and energy consumption.[18]

Distinguishing between pollutant emission and pollutant generation is crucial due to the heterogeneity in firms' environmental management capacities. A firm's actual pollutant generation may not equal its emissions, as pollutant generation consists of two components: emissions and the amount treated through pollution control measures.[19] Given the variation in energy types and pollutant categories across industries, firms typically consume different types of energy and produce distinct pollutants. To enable meaningful inter-firm comparisons and to avoid inter-correlation among energy sources or pollutants, it is essential to aggregate all energy types and pollutants into unified metrics. For pollutants, this study converts various pollutants generated into standardized pollution equivalent units[20] based on the *Pollutant Discharge Fee Collection Standards (2003)*. This regulation assigns specific pollution equivalents to different pollutants[21] and sets discharge fees at 0.7 RMB per pollution equivalent for water pollutants and 0.6 RMB for air emissions. Following Li and Chen (2019), this study applies a weighting adjustment to aggregate pollution equivalents from industrial water pollutants and air emissions, based on the relative discharge fee ratio between the two categories. For energy consumption, China employs "standard coal" as the benchmark for measuring and converting various forms of energy usage. Following the *General Principles for Calculation of Comprehensive Energy Consumption (2008)*, this study

[18]Following the approach of Chen and Chen (2019), the initial matching is based on firm names and legal representative codes. This is followed by matching using organization codes and years. The resulting matched dataset is then consolidated, with duplicate entries removed to ensure accuracy and consistency.

[19]Formally, pollutant generation equals pollutant emissions plus pollutant treatment.

[20]The concept of pollution equivalents refers to a standardized measure that compares the harmful effects, toxicity to organisms, and treatment costs of various pollutants to a baseline pollutant. Pollution equivalents serve as a comprehensive metric to assess the environmental impact of different pollutants or emission activities, considering both their detrimental effects on the environment and the technical and economic feasibility of treatment. For pollutants within the same medium, equal pollution equivalent values indicate a roughly equivalent level of environmental harm.

[21]The pollutants included in the analysis primarily consist of: chemical oxygen demand (COD) and ammonia nitrogen emissions from industrial wastewater; sulfur dioxide (SO_2), nitrogen oxides (NO_x), smoke, and industrial dust emissions from industrial air pollutants.

converts different types of energy inputs[22] into their equivalent standard coal consumption, incorporating it as an energy input alongside capital and labor. The detailed procedures for calculating pollution equivalents and standard coal equivalents are provided in the next subsections.

Appendix A.1.5 Pollution Equivalents

To standardize all pollutants into comparable units, the emission quantities are first converted into kilograms. Then, pollutant equivalents are calculated using the following formula:

$$\text{Pollutant Equivalent Quantity} = \frac{\text{Emission Quantity of the Pollutant (kg)}}{\text{Equivalent Quantity Value of the Pollutant (kg)}}$$

The Equivalent Quantity Value of the Pollutant is determined according to the specifications outlined in Table 1-5 of the *Pollutant Discharge Fee Collection Standards (2003)*.

Appendix A.1.6 Aggregate Energy Consumption

Each type of energy consumption is first converted into its respective unit: water consumption is measured in 10,000 tons, electricity consumption in 10,000 kWh, coal usage in 10,000 tons, natural gas usage in 10,000 cubic meters, gasoline usage in 10,000 tons, diesel usage in 10,000 tons, and district heating in 10,000 GJ. The standardized energy consumption is then calculated by converting all energy variables into equivalent standard coal consumption using the following formula:

$$\text{Standard Coal Equivalent} = (\text{Water Consumption} \times 0.0002429)$$
$$+(\text{Electricity Consumption} \times 1.229) + (\text{Coal Usage} \times 0.7143)$$
$$+(\text{Natural Gas Usage} \times 13.3) + (\text{Gasoline Usage} \times 1.4714)$$

[22] The energy inputs calculated in this study include total coal consumption, total fuel oil consumption, total clean gas consumption, and total industrial water usage.

$$+(\text{Diesel Usage} \times 1.4571) + (\text{District Heating} \times 0.03412)$$

This calculation consolidates all energy types into a unified measure of standard coal equivalent, ensuring consistency and comparability across observations *(General Principles for Calculation of Comprehensive Energy Consumption (2008))*.

Appendix B Further Details on Survival Analysis

Appendix B.1 Logit Model

The discrete time logit model is a widely used framework in survival analysis, particularly suitable for datasets where events are recorded at discrete intervals, such as annually, monthly or events observed at the end of each governor's term. The fundamental equation of the discrete time logit model expresses the log-odds of the hazard probability as a linear function of covariates and a baseline hazard term:

$$\log\left(\frac{h(t)}{1-h(t)}\right) = c(t) + \beta' X$$

where $h(t)$ represents the conditional probability that a governor is promoted during interval t, given that they have not been promoted before t. The baseline hazard function, $c(t)$, captures the temporal dependency of the promotion hazard, and $\beta' X$ is a linear combination of covariates:

$$\beta = [\beta_1, \beta_2, \beta_3, \ldots, \beta_k]^\top, \quad X = [X_1, X_2, X_3, \ldots, X_k]^\top,$$

$X_1, X_2, \ldots, X_k$ represent the covariates influencing promotion probabilities. These include factors such as the GDP growth gap (actual minus target), the governor's age, experience in higher-level government positions, and education level. Rearranging this equation, the hazard probability can be expressed as:

$$h(t) = \frac{\exp(c(t) + \beta' X)}{1 + \exp(c(t) + \beta' X)}.$$

This framework allows the inclusion of covariates that vary across both individuals and time, reflecting the dynamics of political performance and its influence on career outcomes. In this study, the baseline hazard function $c(t)$ is specified to capture the duration dependence inherent in governor promotions. A common specification is a logarithmic form (Jenkins, 2005):

$$c(t) = (q-1)\ln(t),$$

where q determines whether the hazard rate increases, decreases, or remains constant over time. This flexibility is particularly relevant given the structured evaluation cycles in China's political system, where promotion opportunities are tied to term length. In addition to the parametric specification, this study also considers a non-parametric approach to account for the baseline hazard. The non-parametric method involves creating dummy variables for each discrete time interval to represent the baseline hazard $c(t)$ without imposing a specific functional form. For example, if J is the maximum observed time interval, the baseline hazard is expressed as:

$$c(t) = \sum_{j=1}^{J} \gamma_j \mathbb{I}(t=j),$$

The estimation of the discrete time logit model is conducted using maximum likelihood estimation (MLE). For each individual i, the contribution to the likelihood depends on whether the event of interest (promotion) is observed ($d_i = 1$) or censored ($d_i = 0$) during a given interval t. The log-likelihood function is then written as:

$$\ln L = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[d_{it} \ln h(t) + (1-d_{it}) \ln\left(1-h(t)\right) \right].$$

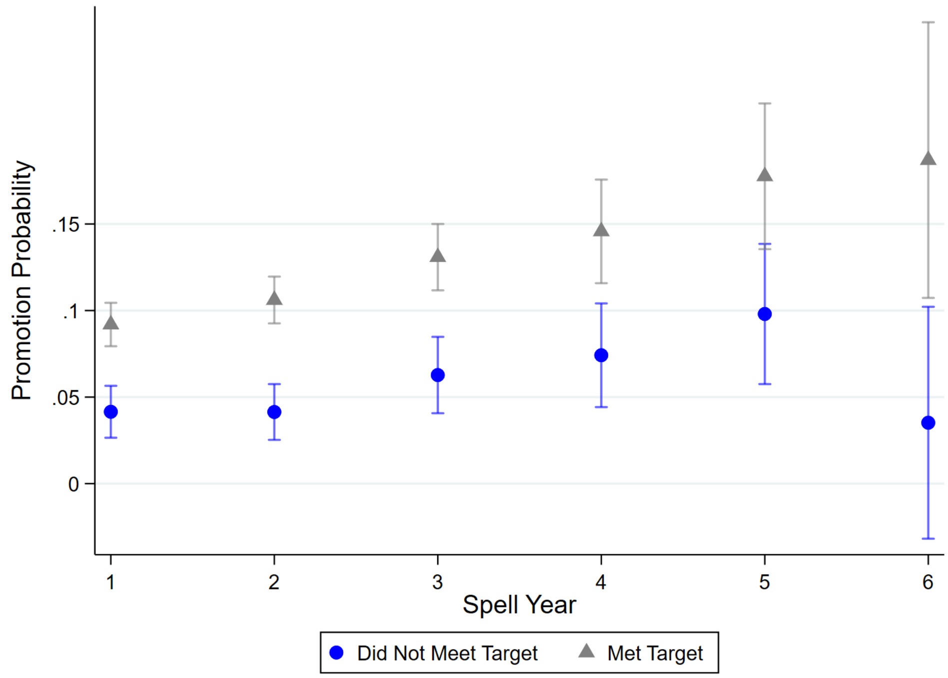
Substituting the hazard function $h(t)$ for the discrete time logit model the log-likelihood becomes:

$$\ln L = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[d_{it} \left(c(t) + \beta' X \right) - \ln\left(1 + \exp(c(t) + \beta' X)\right) \right].$$

Once the parameters $\hat{\beta}$ and $\hat{c}(t)$ are estimated, the hazard rate for individual i at time t is calculated using:

Figure B1: Promotion Probability by Time in Office

Two Types of Performance



Notes: Figure B1 separates officials into two groups—those who met their GDP growth targets and those who did not—and plots their promotion probabilities as a function of tenure length. Each group is constructed conditionally: for example, the second group includes only those who failed to meet the target in the first year, with subsequent groups defined analogously

$$\hat{h}_i(t) = \frac{\exp\left(\hat{c}(t) + \hat{\beta}' X_i\right)}{1 + \exp\left(\hat{c}(t) + \hat{\beta}' X_i\right)}.$$

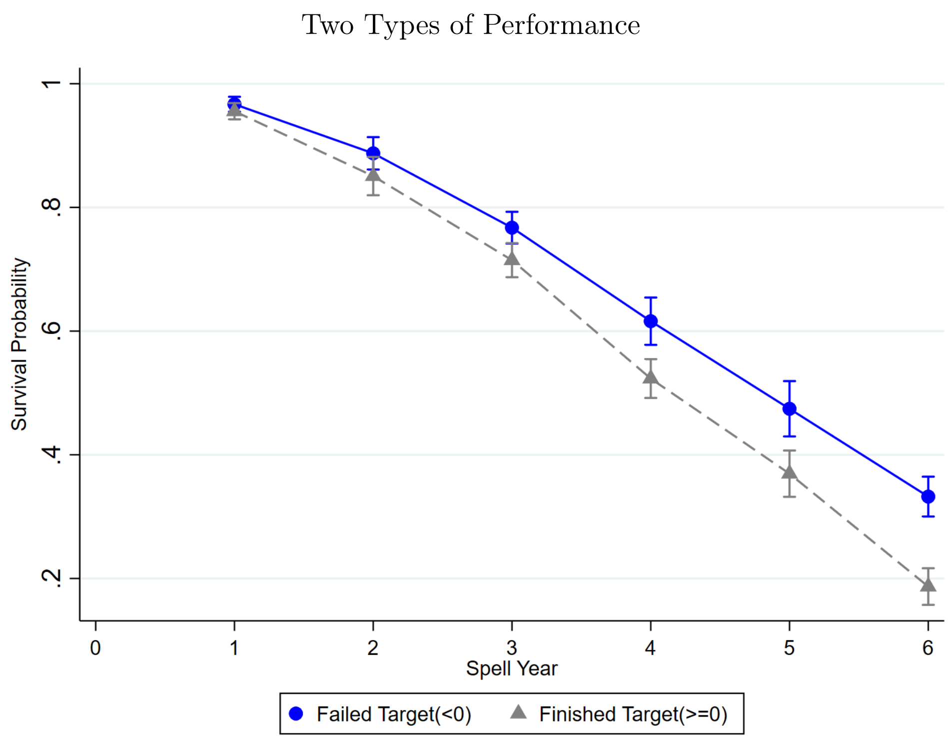
The corresponding survival probability, which represents the likelihood of not being promoted up to time t, is given by:

$$\hat{S}_i(t) = \prod_{j=1}^{t} \left[1 - \hat{h}_i(j)\right].$$

Here, the product accumulates the probabilities of not experiencing the event in each interval from the start until t.

Figure B2 shows the survival probabilities of the two groups of officials as a function of

Figure B2: Survival Rate by Time in Office



Notes: Figure B2 divides officials into two groups: those who achieved their GDP growth targets and those who did not. The figure illustrates the survival probabilities of two groups of officials as a function of tenure length. The second group includes only observations where the target was not achieved in the first year, and the same logic applies to subsequent groups. In other words, this graph tracks the survival rate over time conditional on initial performance, failed the first year target.

survival length. Survival probability reflects the likelihood that an official remains in their current position without being promoted. For both groups, survival probability declines as length increases, consistent with the cumulative risk of promotion (or exit) over time. Officials who meet their GDP growth targets (Gap ≥ 0) consistently have lower survival probabilities than those who fail to meet targets (Gap < 0). This inverse relationship highlights the faster promotion rates of target-achieving officials, who are more likely to leave their current positions for higher roles. The steep decline in survival probabilities for target-achieving officials reflects their accelerated career trajectories. In contrast, officials who fail to meet targets experience a slower decline in survival probabilities, indicating longer tenures in their current roles due to delayed promotion or stagnation. This divergence underscores the meritocratic nature of the promotion process, officials who fail to meet growth targets are promoted more slowly, resulting in higher survival probabilities in their current positions.

Appendix B.2 Non-parametric Baseline Hazard

The first column of Table B1 presents the results of the logit model using a non-parametric specification for the baseline hazard. The odds ratio for the Actual-Target Gap indicates that a one-unit increase in the gap raises the odds of promotion by approximately 9.5%. Similarly, the odds of promotion increase by 3.1% for each additional year of the governor's age, 16.2% for higher education levels, and 15% for prior governmental experience. All results are highly significant and align closely with those from the parametric model, underscoring their robustness. The primary focus of this non-parametric analysis is the dummy variables $j = 1$ to $j = 9$, which represent the log-hazard for each discrete time interval, capturing the non-parametric baseline hazard. These coefficients offer insights into the evolution of promotion probabilities over time. All $j = 1$ to $j = 9$ coefficients are negative, indicating a generally low baseline hazard across these intervals. The magnitude of the coefficients decreases over time (i.e., becomes less negative), suggesting a rising hazard rate as tenure lengthens.[23] However, the relatively modest magnitude of the $j = 9$ coefficient suggests that promotion likelihood may stabilize or even decline slightly at extended durations. This non-monotonicity could stem from institutional or political constraints, where prolonged tenure might indicate stagnation or reduced opportunities for advancement (Li and Zhou, 2005). In summary, the non-parametric baseline hazard reveals an increasing promotion probability over time, particularly during early intervals.

Appendix B.3 Complementary Log-Log Model

The second model employed in this study is the complementary log-log (cloglog) model[24]. Both the cloglog and logit models are widely used to estimate discrete-time event probabilities; however, they differ in their underlying assumptions and functional forms. Specifically, the cloglog model is based on a proportional hazards framework, modeling the hazard rate as:

$$h(t) = 1 - \exp\left(-\exp\left(c(t) + \beta' X\right)\right),$$

In contrast, the logit model assumes proportional odds and uses the logistic function $\frac{\exp(\cdot)}{1+\exp(\cdot)}$. The cloglog model approximates continuous-time proportional hazards models, such as the Cox model (Cox, 1972), when time is divided into discrete intervals. This characteristic makes it especially appropriate for this study, as promotions are recorded on an annual basis, even though the political processes driving them likely unfold in a more continuous manner. On the other hand, the logit model is more flexible, as it does not assume a connection to continuous-time processes. Substituting the cloglog hazard function $h(t)$, the log-likelihood becomes:

[23] For example, the coefficient for $j = 1$ is -4.752, while for $j = 9$, it is -1.675. This pattern reflects an increasing probability of "failure" with longer survival, although the trend is not strictly monotonic.

[24] also known as the *Prentice-Gloeckler (1978) model*

Table B1: Survival Analysis with Non-parametric Baseline Hazard

	Promotion Dummy		
	(1)	(2)	(3)
	Logit	Complementary log-log	Generalized Gamma
Actual-Target Gap	0.091***	0.081***	0.111***
	(0.020)	(0.018)	(0.024)
Age of Governor	0.031***	0.027***	0.058***
	(0.012)	(0.010)	(0.018)
Education Level	0.150**	0.132**	0.153
	(0.068)	(0.061)	(0.100)
Prior Experience	0.140***	0.125***	0.230***
	(0.046)	(0.042)	(0.072)
$j = 1$	-4.752***	-4.540***	-6.712***
	(0.640)	(0.565)	(0.983)
$j = 2$	-3.800***	-3.635***	-5.532***
	(0.644)	(0.567)	(0.973)
$j = 3$	-3.865***	-3.690***	-5.429***
	(0.655)	(0.576)	(0.966)
$j = 4$	-3.116***	-3.026***	-4.387***
	(0.660)	(0.581)	(0.946)
$j = 5$	-2.991***	-2.912***	-3.887***
	(0.672)	(0.591)	(0.947)
$j = 6$	-2.888***	-2.816***	-3.464***
	(0.692)	(0.605)	(0.975)
$j = 7$	-3.053***	-2.983***	-3.284***
	(0.726)	(0.638)	(1.046)
$j = 8$	-2.981***	-2.915***	-2.651**
	(0.818)	(0.714)	(1.192)
$j = 9$	-1.675*	-1.927***	-1.012
	(0.904)	(0.728)	(1.445)
Observations	4508	4508	4504

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: Similar to Table 1, this table presents the results of survival analysis using a non-parametric baseline hazard. The three columns correspond to the results of three different models, all utilizing the same covariates. Dummy variables $j = 1$ to $j = 9$ represent the log-hazard for each discrete time interval (duration interval) and define the non-parametric baseline hazard. These dummy variables are associated with specific time periods, and their coefficients provide insight into how the hazard rate evolves across different durations. The coefficients for j_1 to j_9 are all negative, indicating that the baseline hazard remains generally low during these intervals. However, the coefficients tend to become less negative over time, suggesting a rising hazard rate across intervals, albeit not in a strictly monotonic fashion. This pattern is consistent with the idea that promotion probabilities may increase as time elapses, a common characteristic observed in many survival processes.

$$\ln L = \sum_{i=1}^{N} \sum_{t=1}^{T_i} \left[d_{it} \cdot \ln\left(1 - \exp\left(-\exp\left(c(t) + \beta' X\right)\right)\right) - (1 - d_{it}) \cdot \exp\left(-\exp\left(c(t) + \beta' X\right)\right) \right].$$

The hazard rate and survival probability are calculated in a manner similar to that described in previous section.

The second column of Table 1 presents the results from the complementary log-log (cloglog) model, which estimates the hazard rate of promotion under the proportional hazards assumption. The coefficients are slightly smaller than those obtained from the logit model.[25] The odds ratios derived from the cloglog model demonstrate consistent and highly significant relationships between the covariates and the likelihood of promotion. Specifically, a one-unit increase in the Actual-Target Gap corresponds to an approximate 8.3% increase in the promotion probability. Similarly, the effects of other covariates are broadly consistent with those observed in the logit model. These results further confirm the robustness of the covariates across different survival modeling approaches. Column 2 of Table B1 provides the results of the cloglog model with a non-parametric baseline hazard. The observed differences between this specification and the cloglog model with a logarithmic baseline hazard are comparable to those noted between the two corresponding logit specifications.

Appendix B.4 Survival Model with Frailty

In survival analysis, accounting for unobserved heterogeneity, often termed frailty, is essential (Heckman and Singer, 1984). Frailty represents individual-specific factors that influence the hazard rate but are not directly observed in the data. Neglecting frailty can result in biased parameter estimates, particularly when systematic differences between individuals are omit-

[25] The coefficients are slightly smaller than those in the logit model, reflecting differences in the underlying assumptions of the two approaches. The logit model assumes a symmetric cumulative distribution function (logistic distribution), which assigns more weight to outcomes in the middle of the probability range. This often results in larger coefficients when covariates have a strong effect. In contrast, the cloglog model assumes an asymmetric distribution, placing greater emphasis on extreme values. This asymmetry can dampen coefficient magnitudes, particularly if the effects of covariates are less pronounced in these tails.

ted. In this study, frailty captures unobserved factors affecting city governors' promotion probabilities, such as personal political networks or managerial skills, which are challenging to quantify. Incorporating frailty addresses several potential biases. First, it prevents misestimation of duration dependence, avoiding exaggerated declines or understatements of the hazard rate over time. Second, it ensures that the effects of observed covariates—such as GDP growth performance, prior government experience, and education—remain accurate and are not diminished by omitted unobserved variables. One way to incorporate frailty is by introducing a random effect into the hazard function, which scales the hazard rate for each individual (Meyer et al., 1991). The frailty-adjusted hazard function is expressed as:

$$h(t|X, v) = 1 - \exp\left(-\exp\left(c(t) + \beta' X + \ln(v)\right)\right),$$

where v introduces individual-level variability in the hazard. The corresponding survivor function is:

$$S(t|X, \sigma^2) = \int_0^\infty S(t|X, v) f(v; \sigma^2) dv,$$

where $S(t|X, v) = \exp\left(-\sum_{j=1}^{t} \exp(c(j) + \beta' X + \ln(v))\right)$, and $f(v; \sigma^2)$ is the probability density function of v, typically assumed to follow a Gamma distribution with a mean of 1 and variance σ^2. The log-likelihood function is given by:

$$\ln L = \sum_{i=1}^{N} \ln\left[(1 - d_i) A_i + d_i B_i\right],$$

where d_i is an indicator variable that equals 1 if the event (promotion) occurs for individual i and 0 otherwise. The terms A_i and B_i are components of the likelihood that account for the integration over the frailty distribution. The component A_i captures the likelihood of observing no promotion up to the last interval T_i, and is defined as:

$$A_i = \left[1 + \sigma^2 \sum_{j=1}^{T_i} \exp\left(c(j) + \beta' X_{ij}\right) \right]^{-\frac{1}{\sigma^2}},$$

where $c(j)$ represents the baseline hazard at time j, and $\beta' X_{ij}$ denotes the linear combination of covariates for individual i at time j. The term B_i accounts for the likelihood contribution from the event occurring in the last observed interval. It is specified as:

$$B_i = \begin{cases} 1 - A_i, & \text{if } T_i = 1, \\ \left[1 + \sigma^2 \sum_{j=1}^{T_i - 1} \exp\left(c(j) + \beta' X_{ij}\right)\right]^{-\frac{1}{\sigma^2}} - A_i, & \text{if } T_i > 1. \end{cases}$$

As $\sigma^2 \to 0$, the frailty model simplifies to the standard complementary log-log model without frailty, where unobserved heterogeneity is ignored.

The third column of Table 1 presents the results of the survival model with frailty. Compared to the logit and cloglog models, the coefficients in the frailty model are larger in magnitude. This reflects the model's ability to capture unobserved individual differences that might otherwise dilute the effects of observable covariates. While the direction and significance of the results remain consistent with previous models, the larger coefficients emphasize the amplified roles of all the covariates when individual heterogeneity is accounted for. A one-unit increase in the GDP growth gap raises the promotion probability by 10.3% ($p < 0.01$), while each additional year of the governor's age increases it by 4.6% ($p < 0.01$). Prior governmental experience boosts promotion probability by 17% ($p < 0.01$), and higher education levels are associated with an 11% increase ($p < 0.05$). These findings underscore the importance of both observable and unobservable factors in shaping promotion outcomes. The non-parametric results also remain largely consistent.

Appendix C Additional in Bunching Estimation

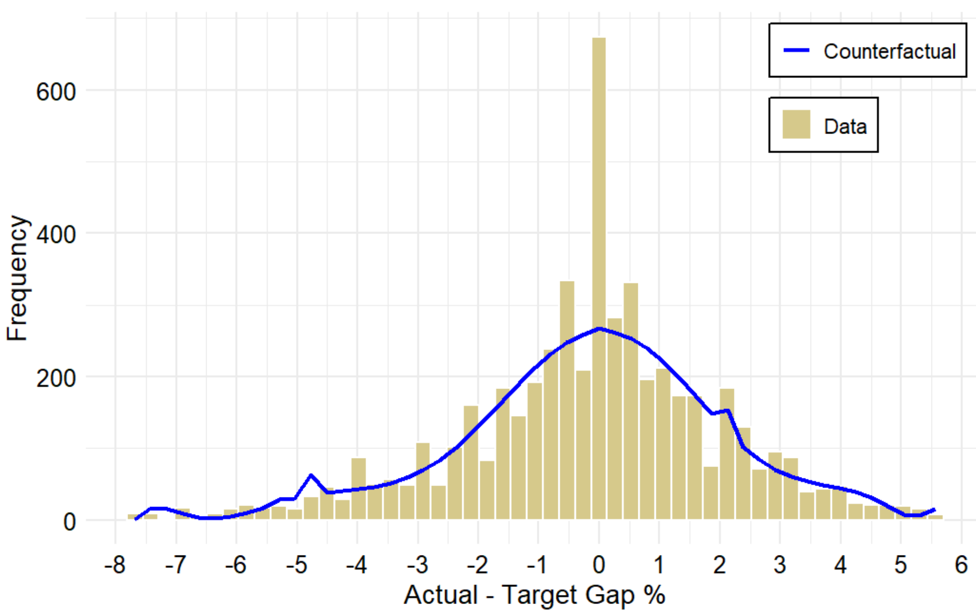
Appendix C.1 Round Number Bunching And Fixed Effects

Targets are often set as round numbers, such as 6 or 7.5 percent in a given year, prompting self-reported actual values to align similarly in order to achieve a gap of zero. However, the reference point effect (Kahneman and Tversky, 1979) can cause round numbers to attract more bunching due to their simplicity, neatness, and convenience. This means that observed bunching may be influenced by factors unrelated to incentive changes. Therefore, I control for round number bunching to prevent overestimating responses at round-numbered kinks. Politicians may bunch at these points for reasons beyond simply meeting the target. This accounts for the perception that gaps, such as 1 percent below or 2.5 percent above, are considered "rounder" and thus may attract more bunching. Failure to control for round number bunching can significantly bias the bunching estimate upwards if z^* is also a round number (Kleven and Waseem, 2013; Dube et al., 2020). Additionally, since politicians might prefer consistently exceeding the target to distinguish themselves from others, I include fixed effects for bins in the set of $K = \{z_1^*, z_2^*\}$ that are outside the excluded region but included in the counterfactual estimation. This helps to net out any potential influence of nearby kinks from the counterfactual. Not controlling for other bunching masses can exert a downward bias on the bunching estimate at z^* by inflating the counterfactual estimate (Mavrokonstantis and Seibold, 2022). Formally, I re-estimate the observed frequency in bin j by incorporating round-number dummies (the third term) and nearby-kink dummies (the fourth term):

$$C_j = \sum_{i=0}^{p} \beta_i (z_j)^i + \sum_{i=z_L}^{z_U} \gamma_i \mathbb{1}\Big[z_j = i\Big] + \sum_{r \in R} \rho_r \mathbb{1}\left[\frac{z_j}{r} \in \mathbb{N}\right] + \sum_{k \in K} \theta_k \mathbb{1}\Big[z_j \in K \wedge z_j \notin [z_l, z_u]\Big] + v_j$$

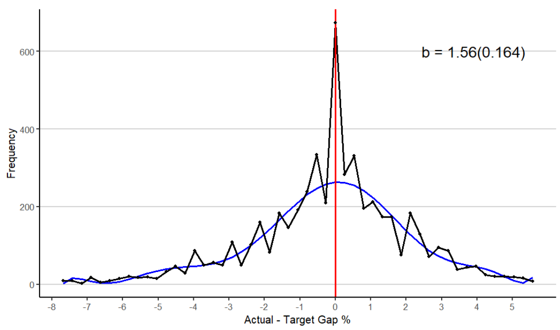
Panel C and Panel D of Figure C2 and second lines of Table 2 show the results after adding these controls. The estimates remain significant.

Figure C1: Counterfactuals with Controls

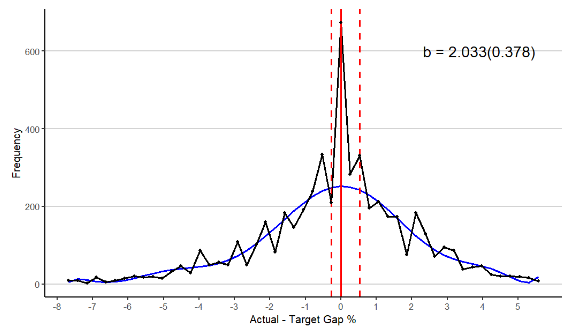


Notes: "Round number" effects are captured by indicators for whether the gap is a multiple of 0.5%, addressing the possibility that officials disproportionately report figures rounded to psychologically salient values. "Nearby kink" controls target other regions in the distribution where abnormal clustering is observed outside the main bunching region. Two such intervals are identified: [2%, 3%] and [–1%, –0.2%]. After incorporating these controls, the estimated excess mass declines, indicating that part of the observed bunching is attributable not only to promotion incentives but also to reporting tendencies favoring round numbers and specific intervals.

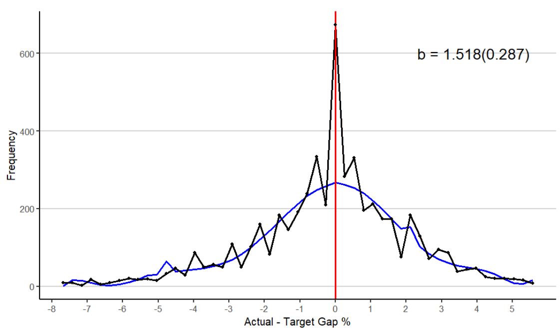
Figure C2: Gap Distribution And Bunching Estimator



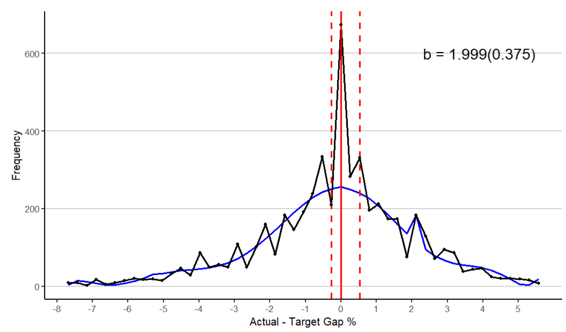
(a) Panel A: Sharp Bunching



(b) Panel B: Diffuse Bunching



(c) Panel C: Sharp Bunching: Round #



(d) Panel D: Diffuse Bunching: Round #

Notes: Figure C2 illustrates the empirical distribution of the actual minus target gap (black line) alongside the estimated counterfactual distribution (blue line). The estimated excess mass is shown in the upper right corner of each panel, with bootstrapped standard errors in parentheses. Panels A and B present the baseline estimates: sharp bunching focuses on a gap of zero as the bunching threshold, while diffuse bunching considers a range (in this case, [-0.2, 0.5]) as the bunching region. The selection of the bunching window is based on visual inspection. Panels C and D show sharp and diffuse bunching, respectively, with added controls for round numbers and nearby kinks. "Round number" refer to the inclusion of dummies for multiples of 0.5% in the actual minus target gap, examining whether these values induce additional bunching in the gap distribution. Specifically, I controlled for whether the gaps are multiples of 0.5%. "Nearby kinks" dummies address other regions of extra mass in the distribution outside the main bunching region. In my analysis, I identified two distinct intervals of extra mass: between 2% and 3%, and between -1% and -0.2%. The estimates of excess mass decrease after adding more controls, suggesting that beyond the incentive to meet targets, officials tend to report GDP figures that gravitate towards round numbers.

Appendix C.2 Integration Constraint Correction

The initial estimate $\hat{b}_0$ may be biased because it overlooks how politicians internally respond. For example, if halfway through the year they discover that economic growth is below expectations, they might intensify efforts in the latter half to ensure they meet the growth target. Consequently, the empirical distribution below z^* naturally appears lower than it would be without incentives. Thus, the straightforward calculation mentioned earlier tends to overestimate the excess mass by disregarding the additional mass at the kink originating from points to the left of z^*. As a result, the observed distribution below the bunching region inaccurately represents the true counterfactual due to the shift it has undergone.

Following the approach outlined by Chetty et al. (2011) for adjusting the counterfactual density to the left of the kink, known as the integration constraint correction, involves shifting the counterfactual distribution upwards on the left side z^* until the number of observations in the empirical distribution aligns with the count in the counterfactual distribution. Despite some skepticism[26], subsequent literature has affirmed the plausibility and validity of this approach (Mortenson and Whitten, 2020; Bergolo et al., 2021). This approach specifically involves defining the counterfactual distribution $\hat{C}_j$ as the fitted values obtained from the regression:

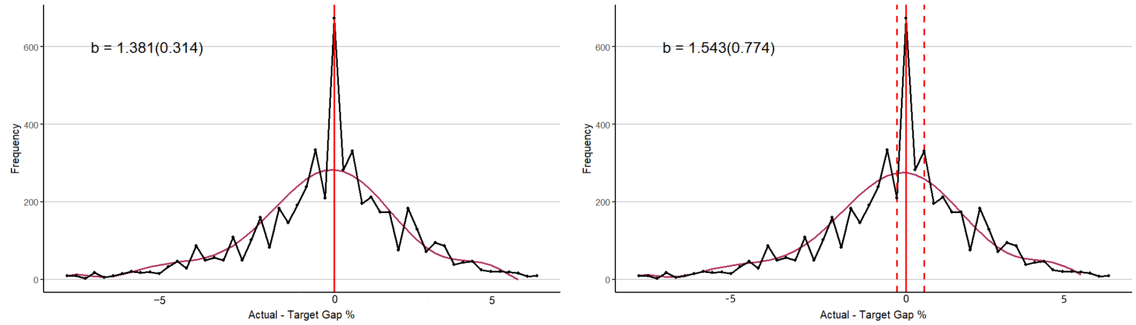
$$C_j\left(1+\mathbb{1}[j<z_L]\frac{\hat{B}_0}{\sum_{j=z^*-1}^{\infty} C_j}\right)=\sum_{i=0}^{p}\beta_i(z_j)^i+\sum_{i=z_L}^{z_U}\gamma_i\mathbb{1}\Big[z_j=i\Big]+\sum_{r\in R}\rho_r\mathbb{1}\Big[\frac{z_j}{r}\in\mathbb{N}\Big]$$

$$+\sum_{k\in K}\theta_k\mathbb{1}\Big[z_j\in K\wedge z_j\notin[z_l,z_u]\Big]+v_j$$

The calculation proceeds iteratively until it reaches a fixed point. The empirical estimate of $\hat{B}$ is then derived from this corrected counterfactual representing the excess mass around the kink relative to the average density of the counterfactual distribution between $[z_L, z_U]$:

[26] Conflicting perspectives in the literature regarding the adjustment of the counterfactual density in the presence of significant extensive responses (Kleven, 2016)

Figure C3: Bunching Estimator After Correction



(a) Panel A: Sharp Bunching (b) Panel B: Diffuse Bunching

Notes: Figure C3 presents the adjusted sharp and diffuse bunching estimators following the integration constraint correction. This correction involves shifting the counterfactual distribution upward to the right of the bunching region.

$$\hat{b} = \frac{\hat{B}_0}{[(z_U - z_L)/\delta]^{-1} \sum_{j=z_L}^{z_U} \hat{C}_j}$$

Note that the denominator of $\hat{b}$ is the average density of counterfactual gap distribution within the bunching window. Panel A and Panel B of Figure C3 illustrate the sharp bunching and diffuse bunching after the correction, respectively. The corrected estimate of excess mass is presented in the third lines of Table 2. As anticipated, a marginal decline in significance is observed following the correction. This is attributed to the upward adjustment of the counterfactual below z^*.

Appendix D Connected Sample

Appendix D.1 Connected Cities

Table D1 presents the "connectedness" of cities within the sample. Among the 339 cities analyzed, 59 are classified as "isolated," as no governors in these cities have served in equivalent positions in other cities during the past 20 years. These cities cannot be linked by movers. The remaining 280 cities have at least one mover—an official who, during their career, held the same position in another city.[27] The isolated cities in the sample include data on 317 officials, while the non-isolated cities account for observations from over 1,500 officials. The spatial distribution of non-isolated cities is relatively uniform, covering cities across 31 provinces, which represents 97% of China's total provincial-level administrative regions. The lower part of Table D1 provides details on non-isolated cities, which can be further grouped based on their networks. Most connected groups are small, typically linked by only one or two officials, and these connections often occur within the same province. However, Group 1 stands out as a significantly larger subsample, encompassing three-quarters of all non-isolated cities through a single interconnected network. The remaining one-quarter of non-isolated cities form separate, independent networks. The subsequent analysis focuses primarily on all non-isolated cities, with additional emphasis on Group 1 cities.

Appendix D.2 Link to Promotion Outcomes

This subsection serves as a robustness check for the previous sections on political incentives by linking leaders' individual contributions to closing the growth gap with their promotion outcomes. The aim is to mitigate the influence of external factors, such as city-level economic conditions and other sources of endogeneity. It is worth noting that to address the inherent indeterminacy of the fixed-effects model, a constraint ($\sum_i \theta_i = 0$) was imposed on leader fixed effects. This constraint ensures that the estimated fixed effects represent each leader's

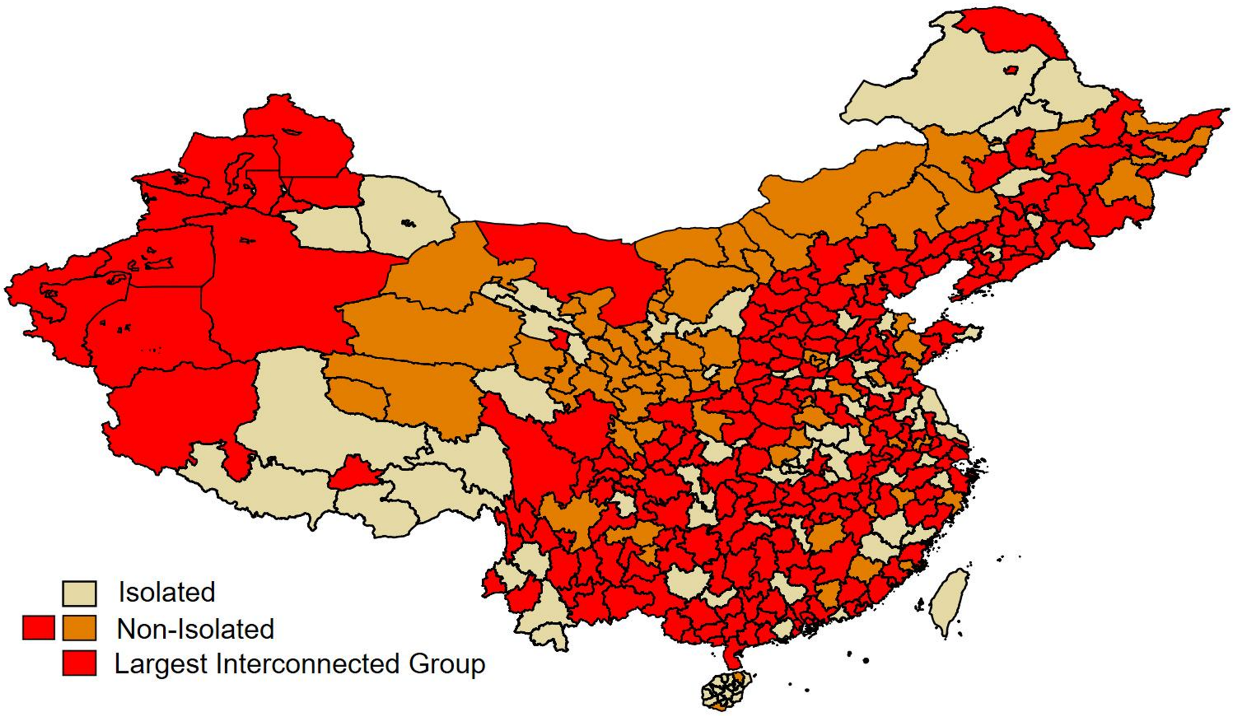
[27] Most movers were transferred within the same province or to neighboring provinces.

Table D1: Connected Groups

Group	Observations	Officials	Movers	Cities	Prov. Invol.
Isolated	1353	317	0	59	25
Non-isolated	6274	1532	274	280	31
Interconnected Groups within Non-isolated Cities					
Group 1	4695	1139	220	208	27
Group 2	552	136	24	26	4
Group 3	136	36	7	6	2
Group 4	92	20	4	4	3
Group 5	67	18	2	3	1
Group 6	60	14	2	3	1
Group 7	69	18	2	3	1
Group 8	70	15	1	3	2
Group 9	46	11	1	2	1
Group 10	44	11	1	2	1
Group 11	46	13	1	2	2
Group 12	44	14	1	2	1
Group 13	46	13	1	2	1
Group 14	46	12	1	2	1
Group 15	44	10	1	2	1
Group 16	40	10	1	2	1
Group 17	46	9	1	2	1
Group 18	44	13	1	2	1
Group 19	43	7	1	2	1
Group 20	44	13	1	2	1
Total	6274	1532	274	280	

Notes: This table summarizes the distribution of observations across isolated and non-isolated groups, including a breakdown of interconnected groups within non-isolated cities. The columns in the table provide information on the number of observations, officials, movers, cities, and provinces involved in each group. Provinces involved refers to the number of distinct provinces represented by the cities within each group. Movers denotes officials who held the same position (governor) in (at least) two different cities. Groups 1–20 represent mutually-exclusive, single interconnected networks of cities.

Figure D1: Connected Cities

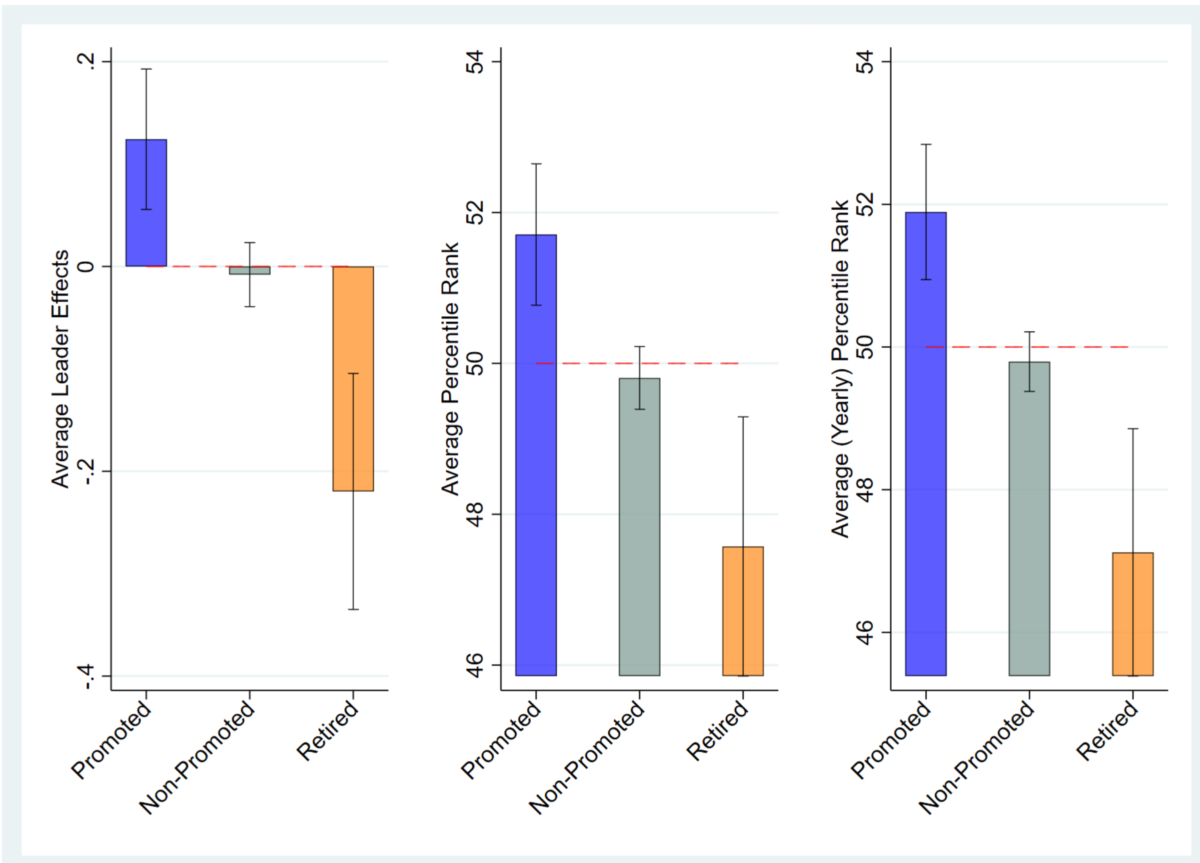


deviation relative to the average of all leaders, rather than their absolute contribution[28]. Without an external reference point or absolute benchmark, the model cannot attribute economic growth solely to the effect of a particular leader but can only identify their relative ranking among all leaders. This approach is akin to estimating relationships within a network of interconnected nodes: one can determine how much stronger one node is compared to another, but not the absolute strength of each node. This section, in fact, examines the relationship between leaders' relative rankings in meeting growth targets and their promotion outcomes. If the findings from the previous sections are robust, leaders with higher relative contributions to closing the growth gap should demonstrate a greater likelihood of promotion.

Figure D2 illustrates the average leader fixed effects across different promotion outcomes: promoted, non-promoted, and retired. The middle panel presents the average percentile rankings of leader fixed effects after converting them into their relative ranks within the

[28]For example, $\theta_i > 0$ indicates that the leader's contribution to meet target exceeds the average, while $\theta_i < 0$ suggests it falls below the average.

Figure D2: Average Leader Effects by Promotion Status



Notes: This figure presents the average leader effects across different promotion outcomes, along with the corresponding percentile ranks derived from their distribution. The rightmost panel illustrates leaders' yearly percentile ranks, calculated annually to provide a year-specific relative ranking.

distribution. The rightmost panel further refines this analysis by calculating yearly percentile rankings to account for potential variations in promotion competition across different years. The results show that, on average, promoted officials exhibit above-average (rank $> 50\%$) performance in meeting growth targets compared to their peers. Non-promoted officials, including those laterally transferred, demonstrate performance close to but slightly below the average. Retired officials, on the other hand, perform the worst on average, likely reflecting their diminished incentives for achieving growth targets due to limited promotion prospects, often related to age or retirement eligibility (Zeng and Zhou, 2024). The average percentile rank results align with these findings and remain robust, whether calculated across the full sample or on a year-by-year basis. Notably, the average rank of officials corresponding to city-year observations where the growth gap equals zero (i.e., targets are precisely met) is approximately 51%. This is very similar to the average rank of officials who were promoted, suggesting that promoted officials are, on average, likely to have successfully achieved their targets.

Figure D3 illustrates the relationship between an official's promotion probability and their rank in the distribution of leader effects. While some variation exists, the overall trend indicates that officials with higher relative contributions to meeting growth targets are more likely to be promoted.

The regression results presented in Table D2 further reinforce the relationship depicted in Figure D3. The dependent variable is a promotion dummy, and the demographic controls used at the leader level are nearly identical to those in Section 4.1. However, the main independent variable here is the leader's percentile rank in contributing to closing the growth gap, rather than the actual-minus-target gap. The left three columns report results for all non-isolated cities, while the right three columns focus on the Group 1 cities subsample.[29] Columns (3) and (6) use yearly-calculated ranks as the main independent variable instead of ranks derived from the entire sample. The results, however, remain highly consistent regard-

[29]Group 1 cities represent the largest subset of non-isolated cities connected by a single, cohesive network, mutually exclusive from other groups.

Figure D3: Promotion Probability by Percentile Rank of Leader Effects

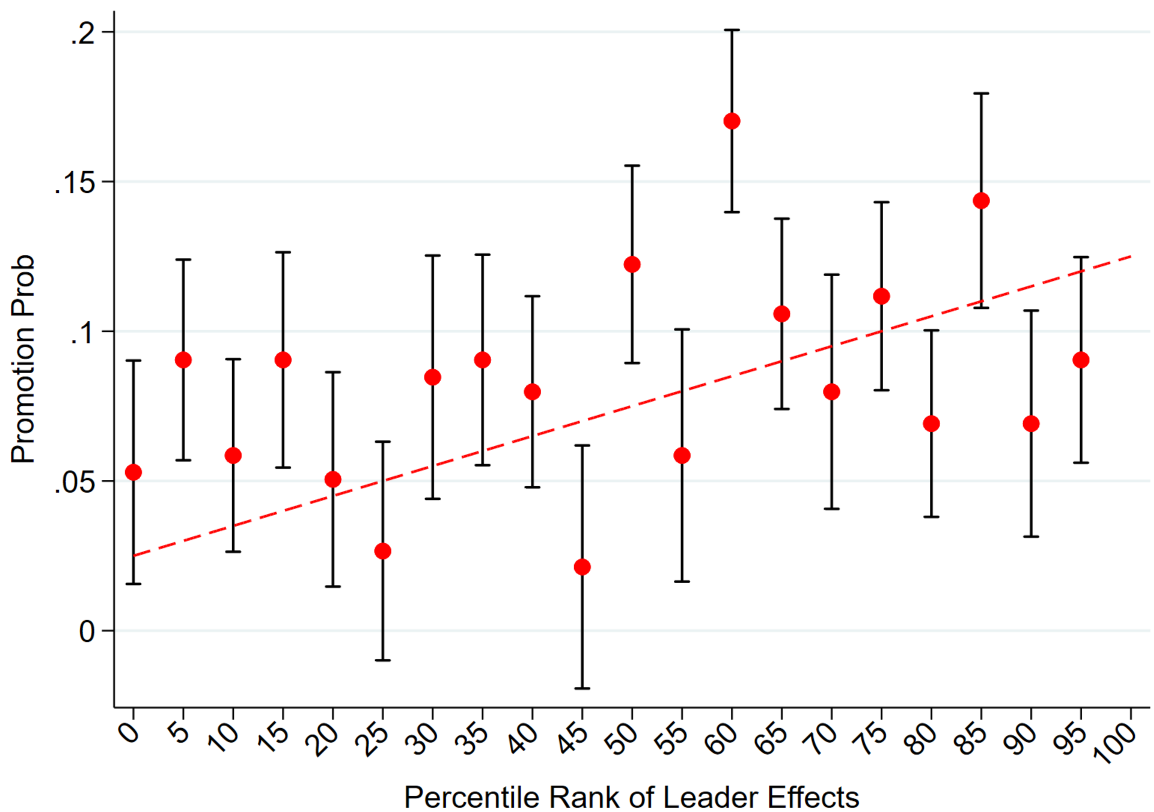


Table D2: Leader Effects and Promotion Outcomes

	Promotion Dummy					
	All Non-Isolated Cities			*Group 1 Cities*		
	(1)	(2)	(3)	(4)	(5)	(6)
Percentile Rank	0.0015**	0.0018**	0.0019**	0.0016**	0.0017**	0.0017**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Tenure		-0.0931***	-0.0932***		-0.0956***	-0.0957***
		(0.013)	(0.013)		(0.016)	(0.016)
Prior Experience		0.0814***	0.0813***		0.0846***	0.0847***
		(0.024)	(0.024)		(0.029)	(0.029)
Education Level		0.1776***	0.1776***		0.1757***	0.1758***
		(0.037)	(0.037)		(0.047)	(0.047)
Age of Governor		0.0003	0.0003		0.0011	0.0011
		(0.006)	(0.006)		(0.008)	(0.008)
Observations	3857	3775	3775	2726	2656	2656
Year FE	✓	✓	✓	✓	✓	✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents regression results examining the relationship between a leader's fixed-effects percentile rank and their promotion outcomes. The left panel includes all non-isolated cities, while the right panel focuses on Group 1 cities—defined as the largest mutually exclusive group of cities connected by a single, interconnected network within the non-isolated sample. Columns (1) and (4) exclude demographic characteristic controls, while Columns (2) and (5) incorporate them. Columns (3) and (6) use year-by-year percentile ranks instead of full-sample rankings.

less of which rank measure is used. Specifically, the findings suggest that a 1% improvement in a leader's rank relative to others increases their promotion probability by 0.17–0.18%. For example, a 20% improvement in rank corresponds to a 3.4–3.6% higher likelihood of promotion. Given the average promotion probability is only about 8–10%, this effect is substantial. Additionally, longer tenure is associated with a lower likelihood of promotion, consistent with the idea that high-performing officials are more likely to be promoted quickly—aligning with the findings in Section 4.1. The influence of other demographic characteristics on promotion remains largely consistent with earlier conclusions. Results from the full sample and the Group 1 subsample are also remarkably similar, further validating the robustness.

Appendix E Robustness Tests

Appendix E.1 Validation with Listed Firms Subsample

The first robustness test narrows the sample to focus exclusively on listed firms within the entire sample, using an alternative data source. Previous studies have raised concerns about the overall quality of the CIED dataset, citing missing key variables or low-quality observations for some firms in certain years (Dai et al., 2021; Brandt et al., 2014). In contrast, listed firms typically provide higher-quality, standardized data due to regulatory requirements, minimizing the risk of measurement errors in critical variables like inventory changes (Chen et al., 2015). Moreover, listed firms often represent larger, more economically significant enterprises whose behavior is likely to be more responsive to political pressures, making them a valuable focus for this analysis (Piotroski et al., 2015). Although listed firms are not fully representative of all firms in the economy, focusing on them reduces the sample size to approximately one-tenth of the original dataset.

Table E1 presents results specifically for listed firms, with data sourced from the CSMAR database rather than the previously used CIED, serving as a robustness check for Table ??. The magnitude of changes in inventory is even larger, which aligns with expectations, as listed firms are typically larger, more responsive, and possibly more susceptible to any influence. The significance of sales decreases slightly, potentially due to the alternative dataset lacking direct sales data for listed firms, requiring the use of proxies. Nonetheless, the overall direction of the results remains consistent with the previous findings.[30]

Appendix E.2 Placebo Analysis

In this subsection, a Placebo Analysis, or Pseudo Gap Test, is performed, conceptually resembling a falsification test. Specifically, alternative artificial thresholds for the running

[30]The sales data from the CIED is intentionally not used in this analysis, not only due to the limited sample size but also because one of the primary objectives of this robustness test is to address potential unreliability in the original dataset.

Table E1: Robustness Test (1): Listed Firms Subsample

Panel A: Inventory Changes

	$\ln \text{Inventory}_t - \ln \text{Inventory}_{t-1}$				
	(1)	(2)	(3)	(4)	(5)
β	0.068***	0.070***	0.070***	0.067***	0.068***
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Firm FE		✓	✓	✓	✓
Industry FE			✓	✓	✓
City Trends				✓	✓
Controls					✓
Observations	18026	18026	18026	18026	17985
R^2	0.0050	0.0063	0.0063	0.0084	0.0128

Panel B: Sales (a proxy)

	$\ln \text{Sales}_t$				
	(1)	(2)	(3)	(4)	(5)
β	0.018	0.020**	0.020**	0.015*	0.014*
	(0.012)	(0.009)	(0.009)	(0.009)	(0.009)
Firm FE		✓	✓	✓	✓
Industry FE			✓	✓	✓
City Trends				✓	✓
Controls					✓
Observations	19162	19162	19162	19162	19119
R^2	0.0058	0.0096	0.0096	0.0228	0.0381

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.
Notes: This table replicates the specification in Table **??**, but limits the sample to publicly listed firms. Direct sales data are not available for listed firms; instead, sales are approximated by multiplying net accounts receivable by the accounts receivable turnover ratio. The results remain consistent, especially for inventory change regressions.

variable are created, such as ± 2 or ± 3. This involves redefining the dummy variable D in the main specification by shifting the critical point for target completion—for instance, setting gap ≥ -3 or gap ≥ 2. The regression discontinuity design is then re-estimated around each pseudo gap to evaluate whether significant discontinuities (jumps) appear at these artificial thresholds. The primary goal of this analysis is to rule out spurious patterns (Athey and Imbens, 2017). Ideally, pseudo gaps should be placed as far as possible from the actual target completion threshold (gap $= 0$) to minimize the influence of diffuse bunching, as discussed in Section 4. Nevertheless, the analysis considers all pseudo gaps at ± 1, ± 2, and ± 3.

Table E2 presents the results of the pseudo gap analysis for inventory changes and sales. Column (1) corresponds to the last column of Table 4, representing the main results. Subsequent columns report the outcomes under alternative definitions of the target completion threshold. The findings reveal that, apart from the actual threshold, most results are statistically insignificant. This suggests that the observed anomalies in firms' behavior are strongly tied to whether the target was genuinely met, rather than to noise, data peculiarities (e.g., irregularities in the data distribution), model misspecification, or spurious correlations. Notably, under the gap ≥ -1 scenario—where the threshold shifts to consider gap ≥ -1 as meeting the target—both inventory changes and sales show some positive significance. This may be attributed to diffuse bunching, where observations near the true target threshold reflect increased efforts to meet the goal. The closer a city comes to achieving its target, the greater the likelihood of intensified efforts, a pattern further corroborated by the heterogeneity analysis in the next section.[31] Importantly, no other points, particularly those to the right of the threshold, display unusual jumps, further reinforcing the robustness of the results.

Appendix E.3 Alternative Dependent Variables

[31] It is worth noting that the "treatment" in this study is not as "clean" as in other RDD applications, where agents below the threshold are entirely unaffected by the treatment. However, this does not substantially undermine the overall robustness of the results.

Table E2: Robustness Test (2): Placebo Analysis

Panel A: Placebo – Inventory Changes

	ln Inventory$_t$ − ln Inventory$_{t-1}$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$D_{jt} = 1$ if	gap ≥ 0	gap ≥ 1	gap ≥ 2	gap ≥ 3	gap ≥ -1	gap ≥ -2	gap ≥ -3
β	0.061***	-0.000	-0.000	0.003	0.064***	-0.000	-0.003
	(0.008)	(0.002)	(0.002)	(0.005)	(0.003)	(0.002)	(0.002)
Observations	189981	189981	189981	189981	189981	189981	189981
R^2	0.0017	0.0013	0.0013	0.0013	0.0017	0.0015	0.0017

Panel B: Placebo – Sales

	ln Sales$_t$						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
$D_{jt} = 1$ if	gap ≥ 0	gap ≥ 1	gap ≥ 2	gap ≥ 3	gap ≥ -1	gap ≥ -2	gap ≥ -3
β	0.044***	-0.002	0.015	0.015	0.004*	-0.000	-0.005
	(0.004)	(0.005)	(0.014)	(0.031)	(0.002)	(0.003)	(0.003)
Observations	279519	279519	279519	279519	279519	279519	279519
R^2	0.0098	0.0015	0.0089	0.0088	0.0014	0.0015	0.0017

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents robustness test results for placebos in inventory change and sales. Panel A reports results for inventory change, while Panel B reports similar results for sales. Each column corresponds to a specific gap value (Gap=1, Gap=2, Gap=3, etc). This means that the gap definition used in each column corresponds to different threshold specifications. Column (1), where gap = 0, aligns with the last column of Table 4, representing the main results. Subsequent columns redefine the breakpoint location, effectively redefining the values of the dummy variable D. For instance, gap = 2 indicates that $D = 1$ if gap ≥ 2. Overall, the results remain robust under these alternative definitions.

Figure E1: Change in Inventory by Gap

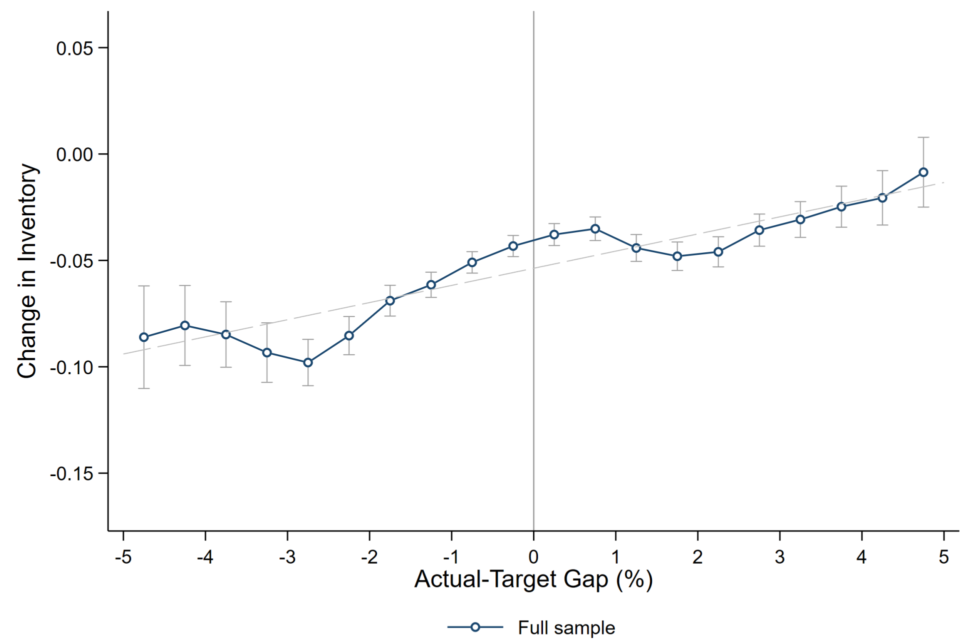


Figure E2: Ln of Sales by Gap

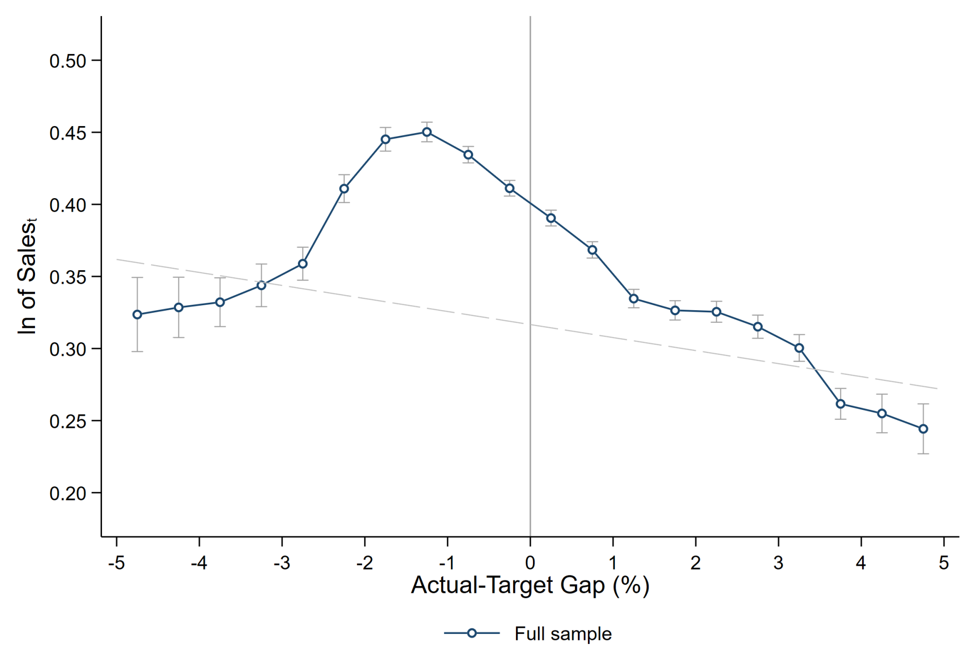


Figure E3: Supplemental: Change in Inventory by Gap

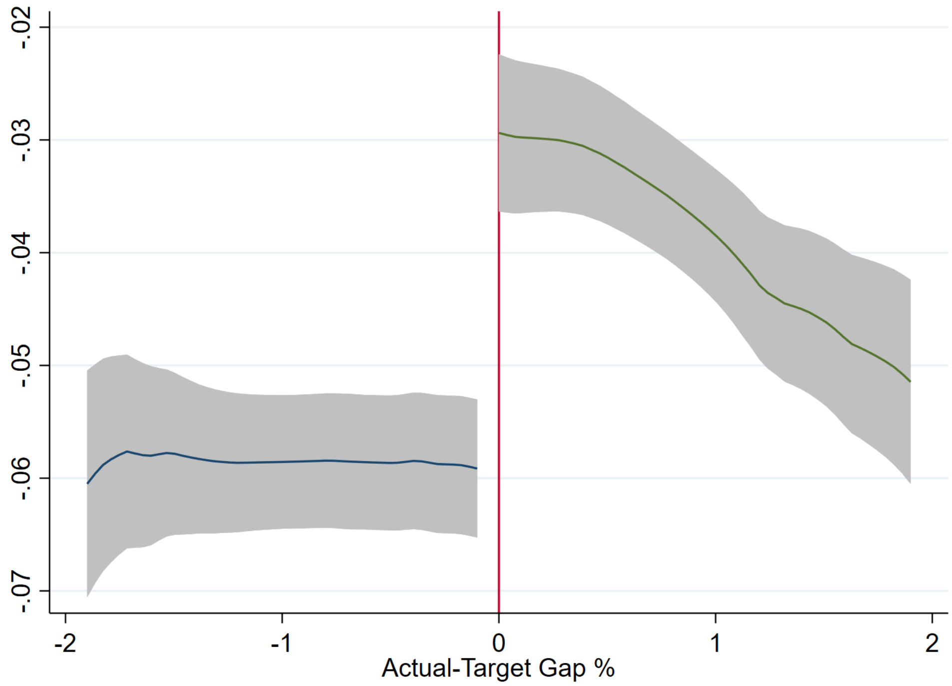


Figure E4: Supplemental: Ln of Sales by Gap

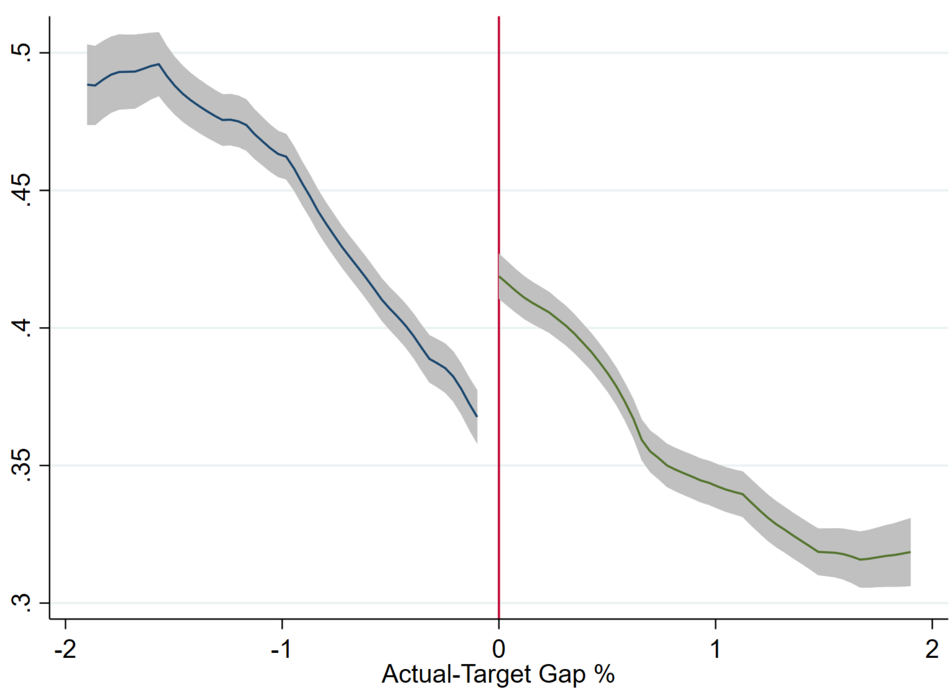


Table E3: Robustness Test (3): Alternative Dependent Variables

	Output		*Revenue*		*Overproduction*		*Intermediate Inputs*	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
β	0.072***	0.067***	0.044***	0.043***	0.063***	0.053***	-0.064***	-0.065***
	(0.005)	(0.005)	(0.004)	(0.004)	(0.008)	(0.008)	(0.005)	(0.005)
Firm FE	✓	✓	✓	✓	✓	✓	✓	✓
Industry FE	✓	✓	✓	✓	✓	✓	✓	✓
City Trends	✓	✓	✓	✓	✓	✓	✓	✓
Controls		✓		✓		✓		✓
Observations	295435	293358	335378	330269	188225	185473	297170	296970

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents the results of the same specification when the dependent variable is replaced with alternative measures: output, revenue, intermediate input, and an indicator of whether the firm engaged in overproduction during the period. Overproduction is an accounting-based measure, and the detailed calculation steps are provided in Appendix E.

In this subsection, the main dependent variable is replaced with alternative firm-level indicators closely tied to local GDP calculations to test the robustness of the results. The first alternative is output, which directly contributes to the GDP formula. The second is revenue, a close proxy for sales. The third is overproduction, a composite measure incorporating both sales and inventory.[32] In addition, the last one is intermediate input, which contributes negatively to GDP calculations.

Table E3 reports the results for four alternative dependent variables. Variables contributing positively to GDP calculations show a strong positive correlation with target completion, while those contributing negatively are negatively correlated with target completion. The proxy for output, gross industrial output value, may reflect price effects and is thus used primarily as a robustness check. Revenue results closely mirror those for sales in both magnitude and significance. Overproduction, a constructed measure combining inventory and sales, also exhibits a positive and significant relationship with target completion. However, it is derived from indirect calculation rather than direct data. Intermediate inputs, on the other hand, are negatively correlated with target completion. Due to substantial data gaps in many years, this variable serves as a reference rather than a key focus. Overall, the results remain robust when alternative dependent variables are used.

Appendix E.4 Calculation of Overproduction

Overproduction is measured by analyzing the abnormal levels of production relative to expected levels, given sales and other production inputs. The steps to calculate overproduction are detailed below.

To estimate normal production levels, the following regression model is used:

[32] Overproduction, an accounting concept, is typically defined as abnormal production levels exceeding expected levels (Roychowdhury, 2006). It also serves as a substitute for inventory, as overproduction increases inventory levels, directly impacting GDP under the production approach.

$$\frac{\text{PROD}_{it}}{\text{Assets}_{it-1}} = \alpha_1 \frac{1}{\text{Assets}_{it-1}} + \alpha_2 \frac{\text{Sales}_{it}}{\text{Assets}_{it-1}} + \alpha_3 \frac{\Delta\text{Sales}_{it}}{\text{Assets}_{it-1}} + \alpha_4 \frac{\Delta\text{Sales}_{it-1}}{\text{Assets}_{it-1}} + \epsilon_{it},$$

where:

- PROD_{it}: Total production for firm i in year t, calculated as the cost of goods sold plus the change in inventory.
- Assets_{it-1}: Total assets in the previous period.
- Sales_{it}: Sales in year t.
- ΔSales_{it}: Change in sales from year $t-1$ to year t.
- $\Delta\text{Sales}_{it-1}$: Change in sales from year $t-2$ to year $t-1$.
- ϵ_{it}: Residual, which captures the deviation from expected production levels.

The residual term ϵ_{it} from the regression represents the abnormal production, which is used as a proxy for overproduction.

Appendix E.5 Alternative Measure of Target Completion

The running variable can also be replaced with an alternative measure. Previously, target completion was determined by whether a city's actual GDP growth rate exceeded the economic growth target set by the municipal government (gap ≥ 0). An alternative measure, however, uses provincial-level data: the difference between the province's annual GDP growth rate and its growth target. While a province meeting its target does not guarantee that all cities within it achieved their respective goals, provincial governments serve as the direct superiors of city governors and have authority over personnel decisions. Thus, provincial-level performance provides a meaningful proxy for the pressure local officials face to meet growth targets.

Table E4 presents the results using the provincial actual-target gap as the running variable. The dependent variables, firm-level ln(inventory changes) and ln(sales), remain consistent with the main analysis. The results for inventory changes are highly robust, with coefficients showing similar size, sign, and significance compared to those obtained using the city-level gap. For sales, the signs and significance are consistent, but the coefficients are larger in magnitude. This indicates that the sales results are also relatively robust. The increased magnitude for sales may result from the provincial-level GDP gap exhibiting less variation than the city-level gap. Moreover, the higher volatility in sales data compared to inventory changes likely amplifies this effect, as reduced variation in the provincial gap makes its influence more pronounced, resulting in larger coefficients for sales.

Table E4: Robustness Test (4): Alternative Measure of Target Completion

Panel A: Inventory Changes

	ln of Inventory$_t$ − ln of Inventory$_{t-1}$				
	(1)	(2)	(3)	(4)	(5)
β'	0.039***	0.060***	0.093***	0.093***	0.094***
	(0.005)	(0.006)	(0.009)	(0.009)	(0.009)
Firm FE		✓	✓	✓	✓
Industry FE			✓	✓	✓
City Trends				✓	✓
Controls					✓
Observations	233686	233686	209951	209951	208685
R^2	0.0008	0.0012	0.0023	0.0023	0.0026

Panel B: Sales

	ln of Sales$_t$				
	(1)	(2)	(3)	(4)	(5)
β'	-0.070***	0.078***	0.083***	0.083***	0.086***
	(0.006)	(0.003)	(0.004)	(0.004)	(0.004)
Observations	339496	339496	311053	311053	309183
R^2	0.0009	0.0066	0.0064	0.0066	0.0142
Firm FE		✓	✓	✓	✓
Industry FE			✓	✓	✓
City Trends				✓	✓
Controls					✓

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents the results when the method for determining target completion is changed from "city actual growth - city target growth ≥ 0" to "provincial actual growth - provincial target growth ≥ 0." The dependent variables remain firm-level ln(inventory changes) and ln(sales).

Appendix F Heterogeneity Analysis

Appendix F.1 Provincial Pressure and Quarterly Inventory

If growth targets effectively influence firm behavior, firms in regions under greater growth pressure are likely to respond more actively to meet these targets. However, quantifying the level of growth pressure faced by governments is challenging and subject to debate. One approach, explored in prior research, considers pressure exerted by higher-level governments (Li et al., 2018; Chen et al., 2020). Since 2005, provincial governments in China have published cumulative GDP growth rates at the end of each quarter as part of quarterly assessments, offering a valuable measure of local growth pressure. For instance, if mid-year GDP growth (cumulative growth at the end of second quarter) falls short of the annual target, firms are expected to respond more intensely in the second half of the year to make up for the shortfall.[33]

Table F1: Yearly Target Completion Rate Based on Success or Failure of Previous Quarters

	If Succeeded Q1	If Succeeded Q2	If Succeeded Q3
Target Completion Rate	81.92%	85.72%	93.67%
	If Failed Q1	If Failed Q2	If Failed Q3
Target Completion Rate	19.28%	10.48%	4.49%

Table F1 reports the annual target completion rate[34] conditional on a city's GDP performance in the first three quarters. Results show a strong positive association between quarterly performance and year-end target fulfillment, particularly as the year progresses. Cities with GDP growth already above target by Q3 are more likely to achieve the annual

[33] Ideally, having quarterly GDP data at the city level would be optimal, as it aligns with the dimensionality of the firm-level data observed. However, city-level quarterly GDP data is not available. Nevertheless, using provincial-level data can still serve as a reasonable proxy for city-level data (as demonstrated section E.5) and also to some extent reflect the pressure exerted by higher-level governments.

[34] defined as the probability of meeting or exceeding the GDP growth target by the end of the year (Q4)

goal. In contrast, if a city falls short of the target in Q1, the probability of recovery by year-end drops to below 20%. These findings suggest that early- and mid-year economic performance serves as a strong predictor of final outcomes, both for external observers and for local officials.

To test this hypothesis and further supplement the earlier findings, quarterly firm-level data is required. Fortunately, all listed firms in China release quarterly financial reports, including inventory data. By combining these two, the following specification is estimated to explore this relationship:

$$y_{ijtq} = \beta_1 \text{Gap}_{tqp} + \beta_2 Q_c + \sum_{k>c} \beta_k Q_c \times q_k + \sum_{r=1}^{3} \gamma_r q_r + \Gamma X_{it} + \sum_j \phi_j t + \mu_i + \lambda_s + \epsilon_{itq}$$

Here, y_{ijtq} represents the change in inventory for firm i in city j, year t, and quarter q. Gap_{tqp} denotes the difference between actual GDP growth in quarter q and the annual GDP target for province p in year t. Q_c is a dummy variable for $c = [1, 2, 3]$, indicating whether GDP growth by the end of the first c quarters failed to meet the annual target.[35] The term $\beta_k Q_c \times q_k$ captures the cumulative effects of missing the target during the first c quarters on subsequent quarters. For example, if $c = 1$, this interaction term captures the impact of the first quarter's GDP growth falling below the annual target on the outcomes in the second, third, and fourth quarters, with the fourth quarter representing annual GDP growth. $\sum_{r=1}^{3} \gamma_r q_r$ represents quarter fixed effects, with the fourth quarter serving as the reference group. The terms ΓX_{it}, $\sum_j \phi_j t$, μ_i, and λ_s correspond to firm-specific covariates, city-specific time trends, firm fixed effects, and industry fixed effects, respectively, same as in the previous sections.

Distinguishing between cases where the annual GDP target was ultimately achieved, despite quarterly shortfalls, and those where the target was missed is essential. Quarterly GDP underperformance is strongly correlated with failing to meet the annual target. For

[35] For instance, $Q_2 = 1$ implies that by the end of the second quarter, actual GDP growth was below the annual target, while $Q_3 = 1$ indicates that GDP growth remained below target by the end of September.

instance, among observations where first-quarter GDP fell short of the target, about 20% managed to achieve the annual goal. This figure drops to 11% for cases where GDP for the first two quarters remained below the target and plummets to just 3% when the shortfall persisted through the first three quarters. When the pressure to meet targets becomes overwhelming or the probability of success too low, the marginal returns to effort diminish. Significant effort may not sufficiently raise the likelihood of promotion, or the rewards for promotion may appear inadequate. In such cases, an official's participation constraint may no longer hold, leading them to withdraw from the "promotion contest" by exerting little or no effort (Li et al., 2019). This distinction aims to identify officials who remain active participants—likely putting in effort to meet the targets—and separate them from those who effectively abandon the contest or reduce their efforts.

Table F2 examines the effect of "quarterly GDP falling short of the annual target" on firms' quarterly inventory changes. Columns (1)–(3) present results for provinces where officials actively participated and ultimately met their annual targets. Columns (4)–(6) mirror these results but use the full sample, including cases where officials reduced efforts or abandoned the goal. The coefficients for Gap_{tqp} are positive and significant across all specifications, indicating that larger gaps between quarterly GDP and the annual target are associated with greater increases in quarterly inventory. This positive correlation is intuitive. Columns (1) and (4) assess the impact of missing the GDP target in the first quarter on inventory changes in the subsequent three quarters. Firms begin increasing inventory from the second quarter, with the effect intensifying as the year-end approaches. Similar trends are observed in the full sample, though with slightly lower magnitudes and significance. Columns (2) and (5) analyze cases where GDP fell short of the target during the first two quarters. Firms start increasing inventory in the third quarter, with a continued and growing effect through the fourth quarter. The full sample shows comparable patterns. Finally, Columns (3) and (6) focus on cases where GDP was below target for the first three quarters. In Column (3), where officials ultimately met the annual target, inventory increases peaked in

Table F2: Growth Pressure Effects on Quarterly Inventory

	Quarterly Change in Inventory					
	Subsample: Finished Provincial Annual Target			*All Sample*		
	(1)	(2)	(3)	(4)	(5)	(6)
Gap_{tqp}	0.001**	0.001***	0.001***	0.002***	0.002***	0.002***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Quarter 2 effect	0.028***			0.013*		
	(0.007)			(0.007)		
Quarter 3 effect	0.030***	0.023***		0.017**	0.013**	
	(0.007)	(0.008)		(0.007)	(0.005)	
Quarter 4 effect	0.044***	0.040***	0.059**	0.021**	0.017**	0.009
	(0.009)	(0.010)	(0.023)	(0.008)	(0.007)	(0.007)
Q_1	-0.042***			-0.028***		
(if failed quarter 1)	(0.006)			(0.006)		
Q_2		-0.035***			-0.024***	
(if failed quarters 1-2)		(0.005)			(0.004)	
Q_3			-0.012			-0.015***
(if failed quarters 1-3)			(0.009)			(0.003)
q_1 (Quarter 1 FE)	0.046***	0.042***	0.039***	0.086***	0.083***	0.080***
	(0.004)	(0.004)	(0.004)	(0.007)	(0.007)	(0.007)
q_2 (Quarter 2 FE)	-0.009***	-0.009**	-0.011***	0.031***	0.034***	0.031***
	(0.003)	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)
q_3 (Quarter 3 FE)	-0.007*	-0.007**	-0.007**	0.041***	0.041***	0.043***
	(0.004)	(0.003)	(0.003)	(0.006)	(0.006)	(0.005)
Observations	33086	33086	33086	70379	70379	70379

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents the results of the above specification. Columns (1)–(3) are based on a subsample of provinces that met their annual growth targets by year-end, even if they failed to meet the target in certain quarters. Columns (4)–(6) mirror Columns (1)–(3) but use the full sample. Column (1) and Column (4) examine the effects of a province failing to meet its GDP growth target in the first quarter on subsequent quarters. Columns (2) and (5) consider the impact of consecutive GDP shortfalls in the first and second quarters on the remaining quarters. Finally, Columns (3) and (6) analyze the effects of GDP shortfalls across the first three quarters on the final quarter. The terms q_1, q_2, and q_3 represent quarter fixed effects.

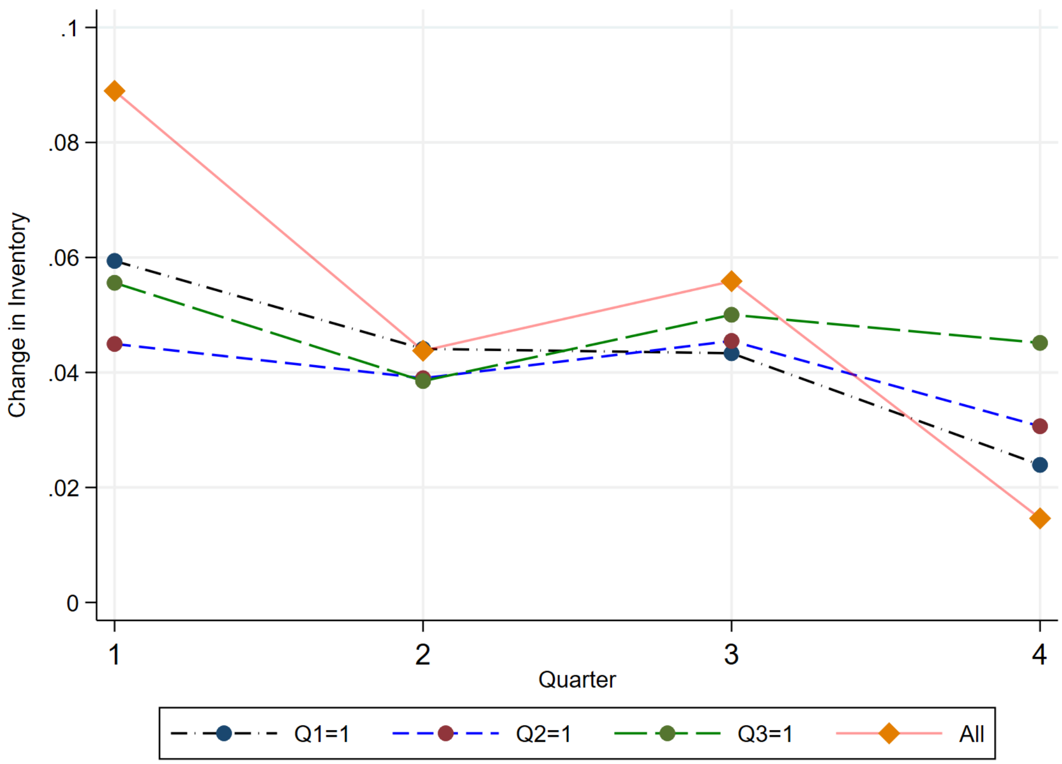
the fourth quarter, reflecting a final effort to close the gap. By contrast, Column (6) shows no significant inventory increase in the fourth quarter, suggesting officials deemed the target unattainable, ceased efforts, or faced insurmountable economic barriers. This aligns with the fact that 97% of cases with GDP shortfalls in the first three quarters failed to recover by year-end. The coefficients for Q_1, Q_2, and Q_3 are negative, reflecting weaker economic performance in provinces that failed to meet their targets. However, the presence of growth pressure offsets these negative values.

Figure F1 visualizes the results from Table F2, depicting firms' quarterly inventory changes and their trends under varying quarterly performance levels. The figure shows that when GDP falls short in the first quarter (black line), firms compensate with a steeper recovery trajectory in subsequent quarters relative to the benchmark.[36] Similar catch-up patterns are observed for second-quarter (blue line) and third-quarter (green line) shortfalls. Notably, when GDP underperforms across all three quarters, the fourth quarter experiences the highest inventory growth as firms strive to close the gap. A stepwise pattern emerges in the fourth-quarter growth rates for $Q3$, $Q2$, and $Q1$, with diminishing effects as earlier-quarter shortfalls are considered. This sequential trend is particularly noteworthy.

A similar analysis can be conducted using the CIED manufacturing firm dataset. Although the CIED data is only available annually and therefore unsuitable for direct comparison, provincial-level quarterly GDP data provides a useful proxy for measuring political pressure from higher-level governments. In China, higher-level authorities exert significant control over lower-level governments, including the power to appoint and dismiss officials. This hierarchical structure compels local governments to align with top-down directives and macroeconomic policies (Wu and Chen, 2016). When a province's quarterly GDP falls short of its target, it reflects growth pressure from higher authorities, prompting provincial leaders

[36] The change in inventory is generally highest in the first quarter, as firms often negotiate new supplier contracts at the beginning of the year, leading to bulk purchases during Q1. This results in a significant accumulation of raw material inventories and intermediate goods. Conversely, the change in inventory is usually lowest at year-end, as firms strategically deplete existing stocks in Q4 to minimize excess inventory carrying costs. This practice is especially prevalent in industries dealing with perishable goods or rapidly depreciating inventory.

Figure F1: Change in Inventory by Quarter
with Distinct Quarterly Performance



Notes: This figure illustrates firms' quarterly changes in inventory across quarters 1 through 4, along with trends in inventory changes under varying levels of quarterly performance, conditioned on meeting the year-end target. $Q1 = 1$ indicates that GDP in the first quarter fell short of the target, $Q2 = 1$ signifies that GDP in both the first and second quarters consecutively fell short, and $Q3 = 1$ reflects underperformance in all three quarters. The trend line for the full sample serves as a benchmark for comparison.

to push cities within their jurisdiction to meet annual targets. In contrast, cities that have already achieved their goals face less pressure, as local governors have weaker incentives to stimulate further growth. Additionally, if the provincial government fails to meet its target, underperforming city governors are often the first to be held accountable. Thus, this analysis focuses on cities that have not yet achieved their annual targets.

Table F3 examines the impact of provincial-level growth pressure on firms' inventory changes. The table closely mirrors the structure and elements of Table F2.[37] Here, D' indicates that a city failed to meet its annual growth target, which is negatively correlated with inventory changes. The interaction term $Q_c \times D'$ captures the influence of provincial quarterly underperformance on city-level outcomes. The results reveal that while failing to meet annual targets (D') is generally associated with lower inventory changes, weak provincial quarterly GDP performance prompts cities that failed to meet their targets to make greater efforts to bridge the gap. Regardless of whether the shortfall occurs in the first quarter or extends across multiple quarters, cities react by ramping up inventory adjustments. In other words, while failing to meet city targets is generally linked to reduced inventory changes, this effect is mitigated under stronger provincial pressure. In the subsample, weaker provincial quarterly GDP performance leads to increased efforts at the city level, revealing a stepwise pattern of inventory growth based on the duration of the province's shortfall within a year, consistent with earlier findings. Although the results remain significant, the stepwise pattern is absent in the full sample, likely due to the inclusion of disengaged officials who put forth limited effort.

Although listed firms represent a smaller and less representative sample, and quarterly GDP data is available only at the provincial level, the combined findings from Tables F2 and F3 strengthen the overall validity of the analysis. Together, the results suggest that when officials recognize that their current efforts may fall short of achieving targets, they are likely to intensify their efforts to close the gap.

[37] Since quarterly GDP data is only available starting from 2005, while the CIED database spans 2000–2015, the total sample size for this analysis is reduced to two-thirds of the original.

Table F3: Higher-Level Pressure and Unmet Targets

	Yearly Change in Inventory					
	Subsample: Finished Provincial Annual Target			*All Sample*		
	(1)	(2)	(3)	(4)	(5)	(6)
D'	-0.017***	-0.018***	-0.013***	-0.018***	-0.018***	-0.014***
(if $\text{gap}_{jt} < 0$)	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.002)
$Q_c \times D'$	0.021***	0.024***	0.043**	0.014***	0.016***	0.010***
	(0.005)	(0.005)	(0.022)	(0.002)	(0.002)	(0.003)
Q_1	-0.034***			-0.032***		
(if failed quarter 1)	(0.002)			(0.002)		
Q_2		-0.038***			-0.035***	
(if failed quarters 1–2)		(0.003)			(0.002)	
Q_3			-0.035***			-0.028***
(if failed quarters 1–3)			(0.008)			(0.002)
Observations	153,413	153,413	153,413	231,606	231,606	231,606

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents firms' yearly changes in inventory, accounting for growth pressure from higher-level governments, using the same database as the main results. The structure of the table mimics that of Table F2, with results reported for both a subsample and the full sample to exclude discouraged officials. D' is a dummy variable equal to 1 if the city fails to meet its target. The interaction term $Q_c \times D'$ captures the impact of poor provincial quarterly economic performance on cities' target completion. Here, weak provincial quarterly economic performance is interpreted as growth pressure exerted by higher-level governments.

Appendix F.2 Sample Restriction Around the Threshold

For cities with a small negative GDP gap, the additional effort required to close the gap is relatively minimal, increasing the likelihood of local governments exerting extra effort (e.g., incentivizing firms) to meet the target. In contrast, cities with a large negative gap face low probabilities of success, leading local officials to perceive such efforts as unproductive and abandon attempts to catch up. This analysis examines how the relationship between city-level GDP targets and firm-level outcomes evolves when the sample is restricted to cities with GDP gaps closer to the threshold. The aim is to determine whether the discontinuity in firm outcomes becomes more pronounced when focusing on cities narrowly missing or just meeting their targets. Narrowing the sample to cities near the threshold also minimizes the influence of extreme values and outliers, as cities far from the threshold are less representative of marginal dynamics and may introduce unnecessary variation into the estimates.

Table F4 presents the results of narrowing the sample to smaller intervals around the target threshold. In each panel, the first column corresponds to the last column of Table 4. For cities within 1 percentage point of the target, shown in Column (2), meeting the GDP target increases inventory changes by approximately 3.1%. This effect is twice as large as that observed in the full sample, indicating that cities closer to the target exert greater pressure on firms to adjust inventories. This reflects heightened incentives or coercion as cities approach the threshold.[38] When the sample is further narrowed to cities within 0.5 percentage points of the target, as shown in Column (3), meeting the GDP target increases inventory changes by 3.7%. The effect continues to grow, suggesting that firms in cities very close to the threshold are more responsive to growth incentives. This is likely due to the intensified urgency for local governments to meet their targets, where the marginal benefit of additional effort is greater. These results align with earlier findings from the

[38] It is important to note that cities ending up within 0.5% or 1% of their growth targets is the result of significant effort. Without such deliberate efforts, cities might fall further behind their targets, potentially by a larger margin, such as 2%. This aligns with the purpose of this analysis: outcomes closer to the target—whether slightly exceeding or falling just short—are more likely the product of intensified efforts.

Table F4: Sensitivity with Narrower Bandwidths

Panel A: Inventory Changes			
	All (1)	$Gap \in [-1, 1]$ (2)	$Gap \in [-0.5, 0.5]$ (3)
β	0.061***	0.182***	0.209***
	(0.008)	(0.020)	(0.039)
Firm FE	✓	✓	✓
Industry FE	✓	✓	✓
City Trends	✓	✓	✓
Controls	✓	✓	✓
Observations	189,981	85,652	53,770
R^2	0.0017	0.0039	0.0088

Panel B: Sales			
	All (1)	$Gap \in [-1, 1]$ (2)	$Gap \in [-0.5, 0.5]$ (3)
β	0.044***	0.091***	0.102***
	(0.004)	(0.007)	(0.018)
Firm FE	✓	✓	✓
Industry FE	✓	✓	✓
City Trends	✓	✓	✓
Controls	✓	✓	✓
Observations	279,519	137,580	80,833
R^2	0.0098	0.0121	0.0190

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Notes: This table presents the results when the sample is restricted to smaller intervals around the target threshold. Both inventory changes and sales are analyzed. In each panel, the first column corresponds to the last column of Table 4. The second column shows the results when the sample is restricted to cases where the gap falls within the range of -1 to 1. Similarly, the third column further narrows the sample to cases where the gap lies within the range of -0.5 to 0.5.

bunching analysis, where effects were shown to increase as observations concentrated near the threshold, reinforcing the validity of both sets of results.

Appendix G True Output Proportion

Using the assumptions outlined in Section 6, the relationship between pollution and target completion can be expressed using the chain rule:

$$\frac{\partial p_{it}}{\partial D_{it}} = \frac{\partial p_{it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}}$$

The relationship between total output Y_{it} and D_{it} is expressed as:

$$\frac{\partial Y_{it}}{\partial D_{it}} = \frac{\partial Y_{T,it}}{\partial D_{it}} + \frac{\partial Y_{M,it}}{\partial D_{it}}.$$

True output is only affected indirectly through energy:

$$\frac{\partial Y_{T,it}}{\partial D_{it}} = \frac{\partial Y_{T,it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}}$$

Substituting:

$$\frac{\partial Y_{it}}{\partial D_{it}} = \frac{\partial Y_{T,it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}} + \frac{\partial Y_{M,it}}{\partial D_{it}}$$

The proportion of true output's contribution to the total output jump at the target threshold can be expressed as:

$$\text{True Output Proportion} = \frac{\frac{\partial Y_{T,it}}{\partial D_{it}}}{\frac{\partial Y_{it}}{\partial D_{it}}} = \frac{\frac{\partial Y_{T,it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}}}{\frac{\partial Y_{it}}{\partial D_{it}}} = \gamma_1$$

Using the pollution-energy relationship, the proportion of true output, normalized by pollution, can be derived. From the assumption:

$$\frac{\partial p_{it}}{\partial e_{it}} = \alpha \cdot \frac{\partial Y_{T,it}}{\partial e_{it}}$$

the energy elasticity of true output[39] can be expressed as:

$$\frac{\partial Y_{T,it}}{\partial e_{it}} = \frac{\frac{\partial p_{it}}{\partial e_{it}}}{\alpha}$$

Using the relationship:

$$\frac{\partial p_{it}}{\partial D_{it}} = \frac{\partial p_{it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}}$$

we can rearrange to solve for $\frac{\partial e_{it}}{\partial D_{it}}$:

$$\frac{\partial e_{it}}{\partial D_{it}} = \frac{\frac{\partial p_{it}}{\partial D_{it}}}{\frac{\partial p_{it}}{\partial e_{it}}}$$

Substituting these back into the true output proportion gives:

$$\text{True Output Proportion} = \frac{\frac{\partial p_{it}}{\partial e_{it}} \cdot \frac{\partial e_{it}}{\partial D_{it}}}{\alpha \cdot \frac{\partial Y_{it}}{\partial D_{it}}} = \frac{\frac{\partial p_{it}}{\partial e_{it}} \cdot \frac{\frac{\partial p_{it}}{\partial D_{it}}}{\frac{\partial p_{it}}{\partial e_{it}}}}{\alpha \cdot \frac{\partial Y_{it}}{\partial D_{it}}} = \frac{\frac{\partial p_{it}}{\partial D_{it}}}{\alpha \cdot \frac{\partial Y_{it}}{\partial D_{it}}} = \gamma_2$$

[39] γ_1 represents the lower bound of the true output proportion since it is estimated using only energy, labor, and capital to predict true output. However, true output is clearly influenced by additional factors beyond these three.These include technological efficiency (Syverson, 2011), natural conditions (Ploeg, 2011), and firm-specific characteristics such as economies of scale (Allcott et al., 2016), capital quality (Fleisher et al., 2010), and accumulated experience. Conversely, γ_2, when setting $\alpha = 1$, serves as an upper bound for the true output proportion. In reality, the proportionality constant α in China is likely greater than 1 for several reasons. First, many industries, such as steel, cement, and manufacturing, are highly energy-intensive but often operate inefficiently, resulting in a low marginal contribution of energy to true output (Wang et al., 2019). Second, China's reliance on coal and other high-pollution energy sources leads to a disproportionately large increase in pollution for each additional unit of energy consumed. Third, regions with weak enforcement of environmental regulations often lack effective pollution control technologies, further amplifying the pollution impact of energy use. By assuming $\alpha = 1$, the analysis simplifies the relationship, treating the efficiency of turning energy into pollutants as equivalent to the efficiency of turning energy into true output. This simplification provides an upper bound without introducing excessive additional assumptions about α.

Figure G1: Energy Consumption

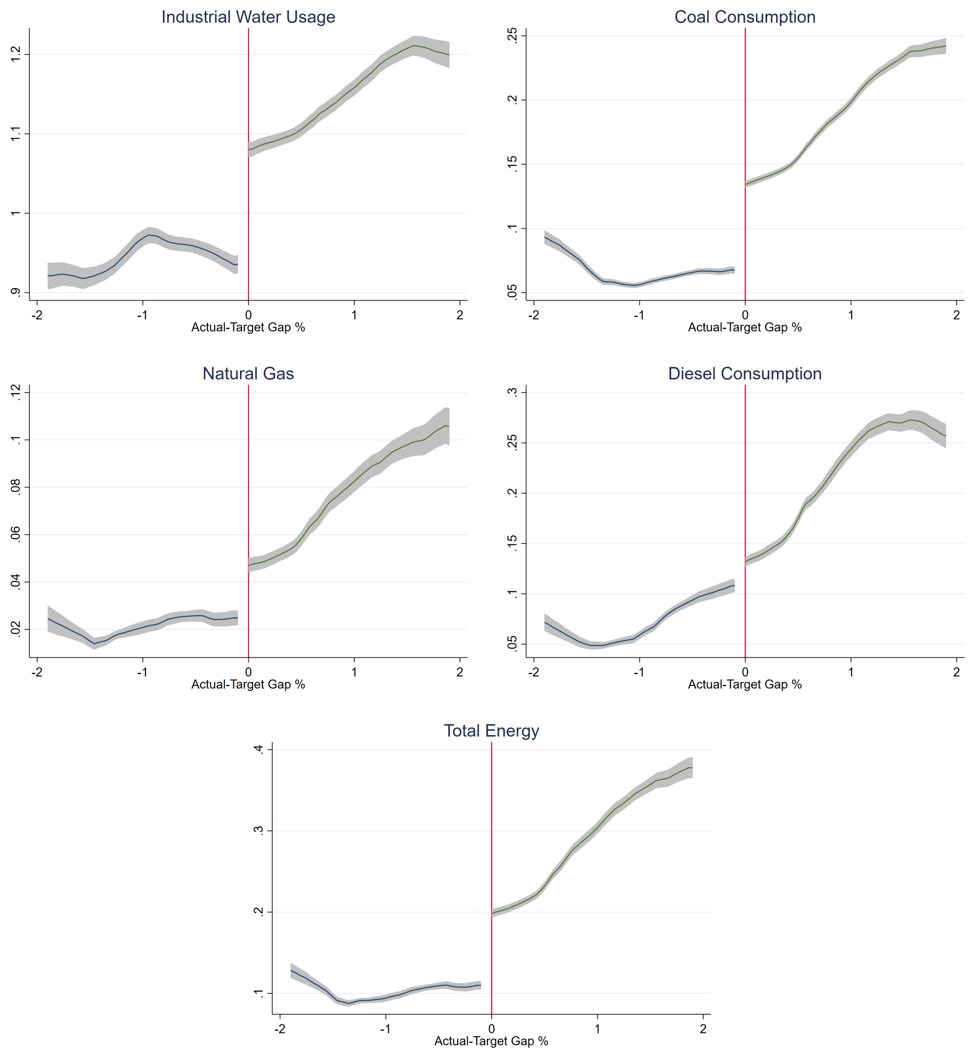
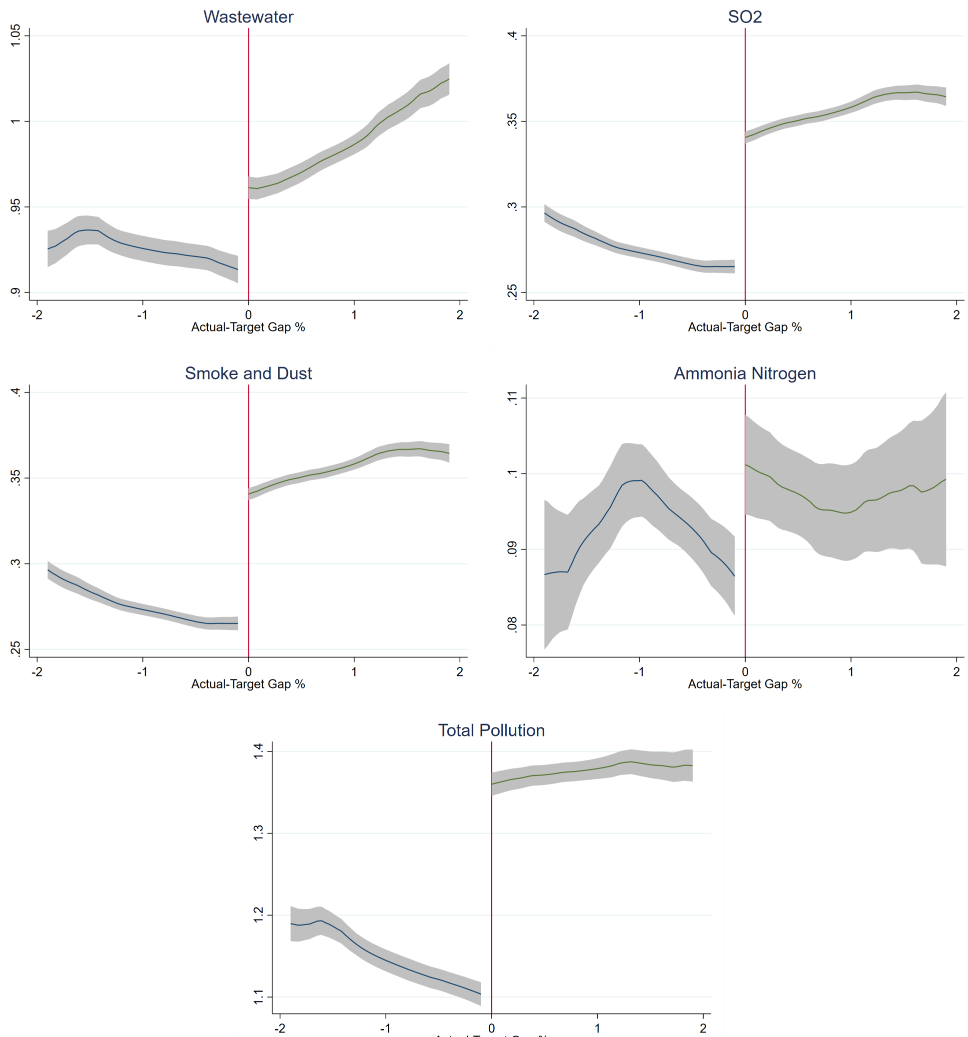


Figure G2: Pollutants



References

Allcott, H., Collard-Wexler, A., and O'Connell, S. D. (2016). How do electricity shortages affect industry? evidence from india. *American Economic Review*, 106(3):587–624.

Athey, S. and Imbens, G. W. (2017). The state of applied econometrics: Causality and policy evaluation. *Journal of Economic Perspectives*, 31(2):3–32.

Bergolo, M., Burdin, G., De Rosa, M., Giaccobasso, M., and Leites, M. (2021). Digging into the channels of bunching: Evidence from the uruguayan income tax. *The Economic Journal*, 131(639):2726–2762.

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

Brandt, L., Van Biesebroeck, J., and Zhang, Y. (2014). Challenges of working with the chinese nbs firm-level data. *China Economic Review*, 30:339–352.

Chen, S., Miao, B., and Shevlin, T. (2015). A new measure of disclosure quality: The level of disaggregation of accounting data in annual reports. *Journal of Accounting Research*, 53(5):1017–1054.

Chen, X., Cheng, Q., and Hao, Y. e. a. (2020). Gdp growth incentives and earnings management: evidence from china. *Review of Accounting Studies*, 25:1002–1039.

Chen, Z. and Chen, Q. (2019). Energy utilization efficiency of chinese enterprises: Heterogeneity, influencing factors, and policy implications. *China Industrial Economics*, (12):78–95.

Chetty, R., Friedman, J. N., Olsen, T., and Pistaferri, L. (2011). Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from danish tax records. *The Quarterly Journal of Economics*, 126(2):749–804.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society: Series B (Methodological)*, 34(2):187–202.

Dai, Y., Li, X., and Liu, D. e. a. (2021). Throwing good money after bad: Zombie lending and the supply chain contagion of firm exit. *Journal of Economic Behavior & Organization*, 189:379–402.

Dube, A., Jacobs, J., Naidu, S., and Suri, S. (2020). Monopsony in online labor markets. *American Economic Review: Insights*, 2(1):33–46.

Fleisher, B., Li, H., and Zhao, M. Q. (2010). Human capital, economic growth, and regional inequality in china. *Journal of Development Economics*, 92(2):215–231.

Heckman, J. and Singer, B. (1984). A method for minimizing the impact of distributional assumptions in econometric models for duration data. *Econometrica: Journal of the Econometric Society*, pages 271–320.

Jenkins, S. P. (2005). Survival analysis. Unpublished manuscript, Institute for Social and Economic Research, University of Essex, Colchester, UK.

Kahneman, D. and Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47(2):363–391.

Kleven, H. J. (2016). Bunching. *Annual Review of Economics*, 8(1):435–464.

Kleven, H. J. and Waseem, M. (2013). Using notches to uncover optimization frictions and structural elasticities: Theory and evidence from pakistan. *The Quarterly Journal of Economics*, 128(2):669–723.

Lam, W. R., Schipke, M. A., and Tan, Y. e. a. (2017). *Resolving China's zombies: Tackling debt and raising productivity*. International Monetary Fund.

Li, H. and Zhou, L. A. (2005). Political turnover and economic performance: the incentive role of personnel control in china. *Journal of Public Economics*, 89(9-10):1743–1762.

Li, J., Xia, J., and Zajac, E. J. (2018). On the duality of political and economic stakeholder influence on firm innovation performance: Theory and evidence from chinese firms. *Strategic Management Journal*, 39(1):193–216.

Li, P. and Chen, Y. (2019). Environmental regulation, firm bargaining power, and green total factor productivity. *Finance & Trade Economics*, 40(11):144–160.

Li, X., Liu, C., and Weng, X. e. a. (2019). Target setting in tournaments: theory and evidence from china. *The Economic Journal*, 129(623):2888–2915.

Mavrokonstantis, P. and Seibold, A. (2022). Bunching and adjustment costs: Evidence from cypriot tax reforms. *Journal of Public Economics*, 214:104727.

Meyer, J. M., Eaves, L. J., Heath, A. C., and Martin, N. G. (1991). Estimating genetic influences on the age-at-menarche: A survival analysis approach. *American Journal of Medical Genetics*, 39(2):148–154.

Mortenson, J. A. and Whitten, A. (2020). Bunching to maximize tax credits: Evidence from kinks in the us tax schedule. *American Economic Journal: Economic Policy*, 12(3):402–432.

Piotroski, J. D., Wong, T., and Zhang, T. (2015). Political incentives to suppress negative information: Evidence from chinese listed firms. *Journal of Accounting Research*, 53(2):405–459.

Ploeg, F. (2011). Natural resources: curse or blessing? *Journal of Economic Literature*, 49(2):366–420.

Qiao, K.-y. (2013). A reexamination of the mechanism of official promotion tournament in china: Evidence from provincial and municipal governments. *Financial Research*, 39(4):4.

Roychowdhury, S. (2006). Earnings management through real activities manipulation. *Sloan School of Management, Massachusetts Institute of Technology*.

Syverson, C. (2011). What determines productivity? *Journal of Economic Literature*, 49(2):326–365.

Wang, J., Hu, M., and Tukker, A. e. a. (2019). The impact of regional convergence in energy-intensive industries on china's co2 emissions and emission goals. *Energy Economics*, 80:512–523.

Wu, M. and Chen, B. (2016). Assignment of provincial officials based on economic performance: Evidence from china. *China Economic Review*, 38:60–75.

Zeng, J. and Zhou, Q. (2024). Mayors' promotion incentives and subnational-level gdp manipulation. *Journal of Urban Economics*, 143:103679.