Online Appendix of "Starving and Deceiving? How Disasters Reshape Politicians' Incentives to Lie"

Shuo Chen, Xinyu Fan, and Xuanyi Wang[1]

Online Appendix A. Construction of GDP Manipulation

This section elaborates our method to calculate GDP manipulation. In short, we aim to recover the true GDP data from indicators that are closely related with economic growth (thus with GDP), and are less manipulable. In particular, we follow the estimation methods of Pinkovskiy and Sala-i-Martin (2016) and Clark et al. (2020). We start by collecting data of satellite light and air pollution – frequently used proxies for economic activities that are less manipulable by local officials. We include the sources of all data in appendix Table A1. We proceed to aggregate the data to county-level in annual bases, as the input variables of our GDP prediction. The input variables also include year and county dummies.

[1] Chen, Department of Economics, Fudan University, Shanghai, China, e-mail: cs@fudan.edu.cn; Fan: Cheung Kong Graduate School of Business, Beijing, China, e-mail: xyfan@ckgsb.edu.cn; Wang: Department of Economics, University of Zurich, Zurich, Switzerland, e-mail: xuanyi.wang@uzh.ch.

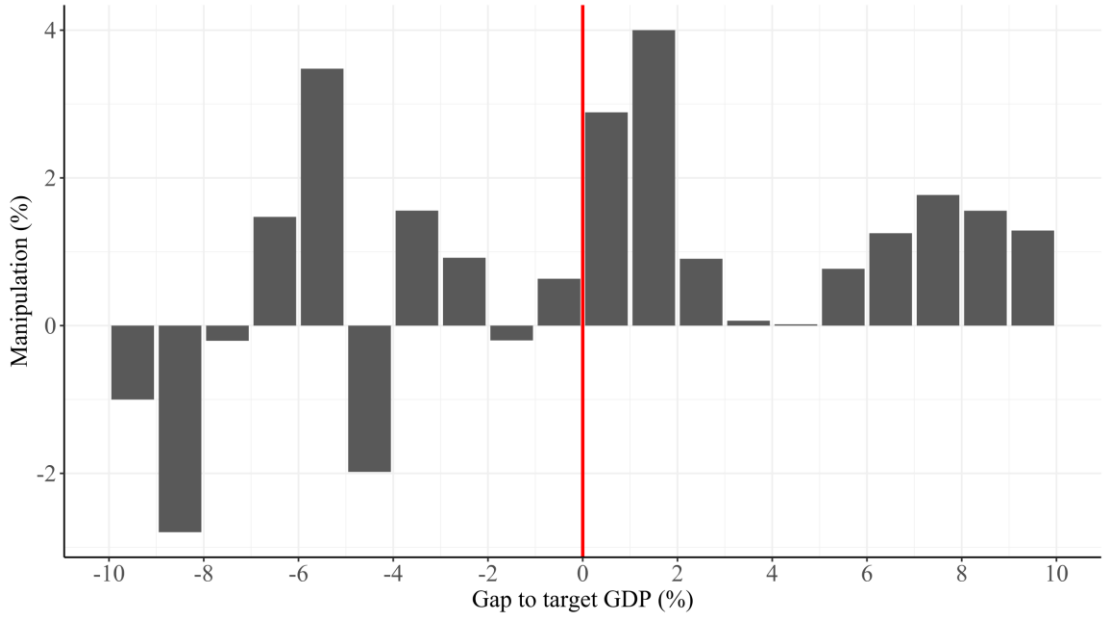
Table A1. Sources of Variables

Variables	Sources
NO2 (3-year moving average)	European Space Agency (ESA)
NO2 (1-year moving average)	National Aeronautics and Space Administration (NASA)
SO2	European Commission (EC)
PM10	EC
PM2.5	NASA, EC
CO2	EC
OC	EC
BC	EC
CO	EC
NH3	EC
NMVOC	EC
NOx	EC
Population	Oak Ridge National Laboratory
Nighttime light	NASA

Next, we select the machine learning models. We divide the sample into 70% training set and 30% validation set, and use popular models such as random forest, support vector machine, neural network, and XGBoost for training. We then test the predictive power of each model, among which XGBoost outperforms the rest. Therefore, we construct the outcome variables using XGBoost.

We justify the credibility of our estimates through three avenues. First, appendix Figure A1 depicts the distribution of GDP manipulation, where the horizontal axis marks GDP growth targets, and the vertical axis marks the frequency. As shown in the figure, the predicted manipulation is positive when the reported GDP is slightly above the target, and the gap shrinks when the reported GDP is far greater than the target. This is consistent with the intuition of manipulation, that individuals are strongly motivated to pass the threshold when their performance is slightly below the target, while the incentives of manipulation are less salient when the performance is well-above the target already. On the other hand, when the reported GDP is less than the target, the gap oscillates around zero – consistent with the presence of truth-telling officials.

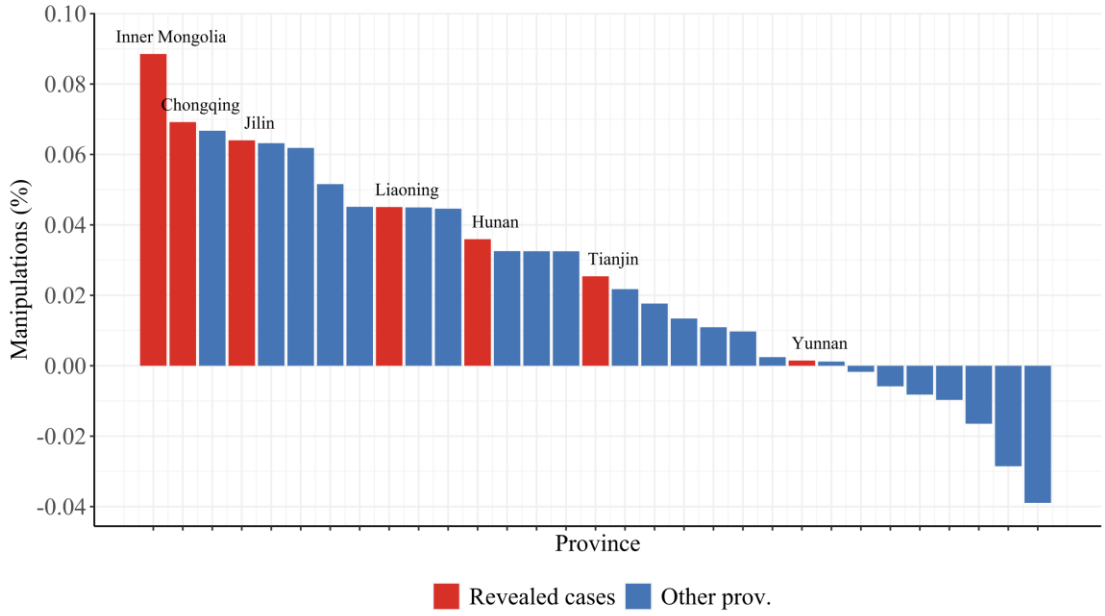
Figure A1. Distribution of GDP manipulations



Note: We only present the gap to target GDP ratio between -10% and 10% for conciseness.

Next, we cross-compare our estimated manipulation with exposed GDP manipulation case in reality. As shown in appendix Figure A2, the red bars indicate the exposed provinces in reality – they are also predicted to have manipulated local GDP in our estimation, which adds to the credibility of our estimates.

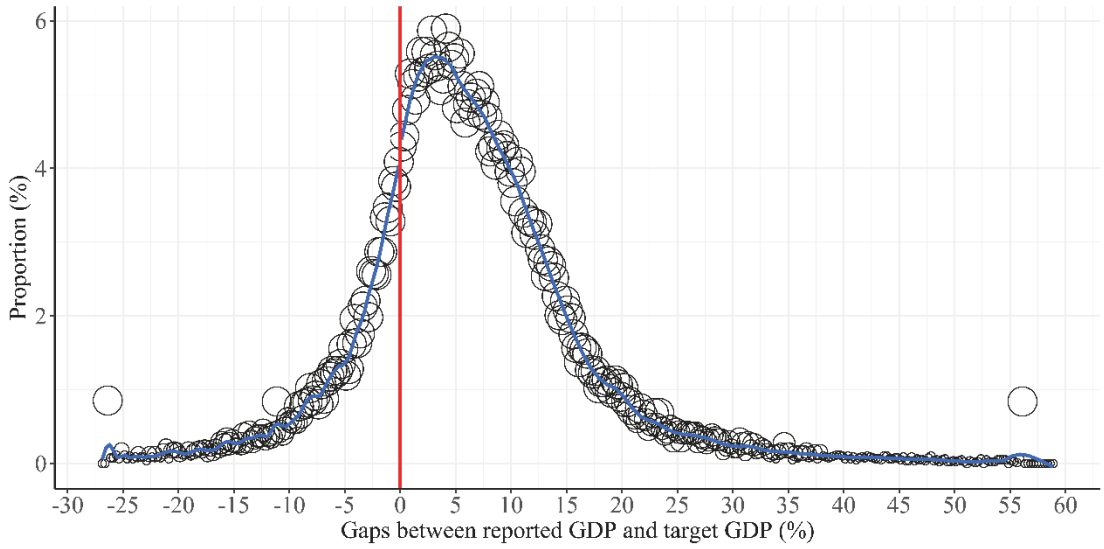
Figure A2. GDP manipulations across provinces



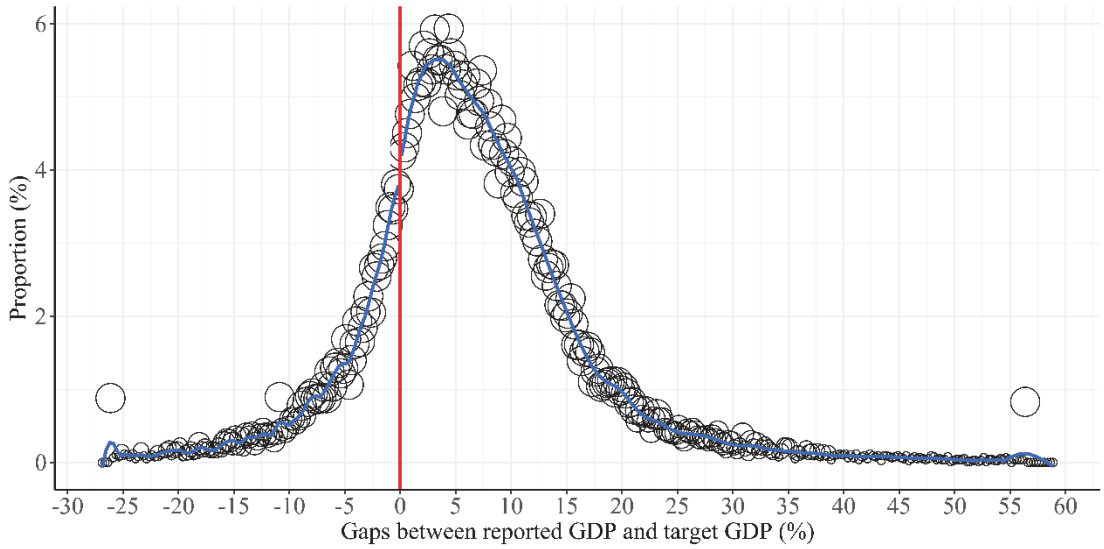
Lastly, we compare our calculated manipulations with the estimates by Henderson et al. (2012) based on satellite light data. The two estimates are highly correlated at 1% significance level.

Online Appendix B: Figures and Tables

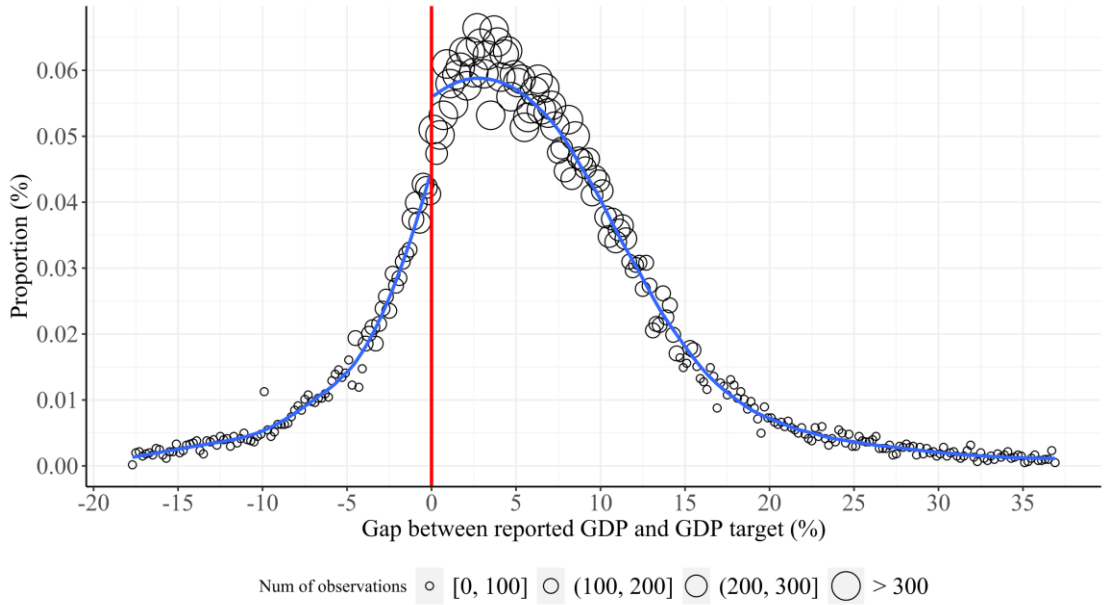
Figure B1. GDP Distribution after Manual Revision of Targets



(A) Target GDP + 0.1%



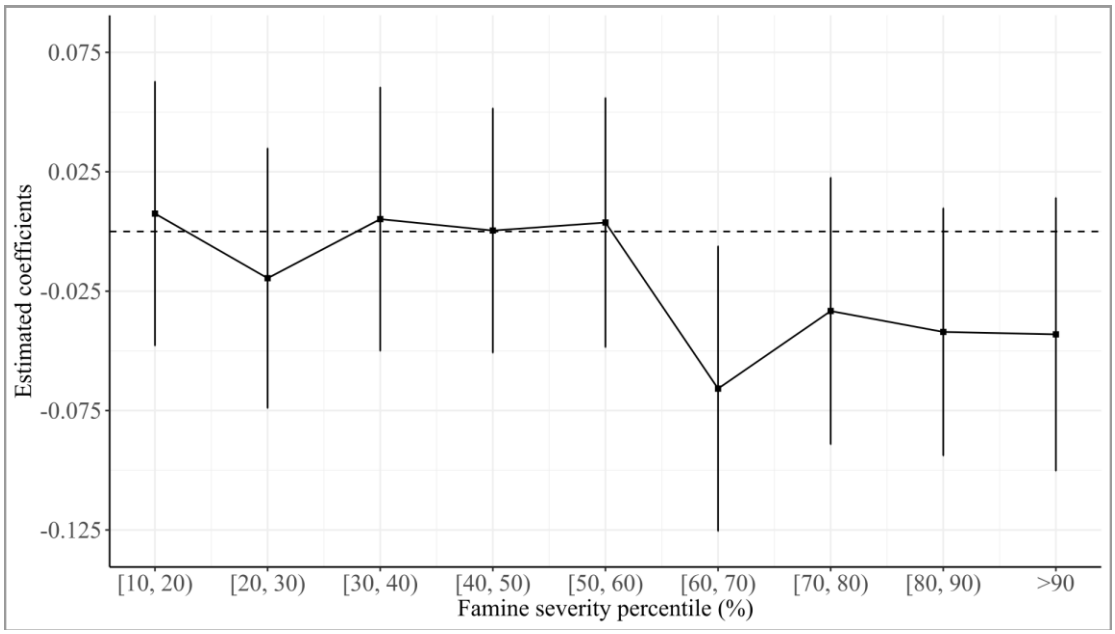
(B) Target GDP – 0.1%



(C) Real target GDP

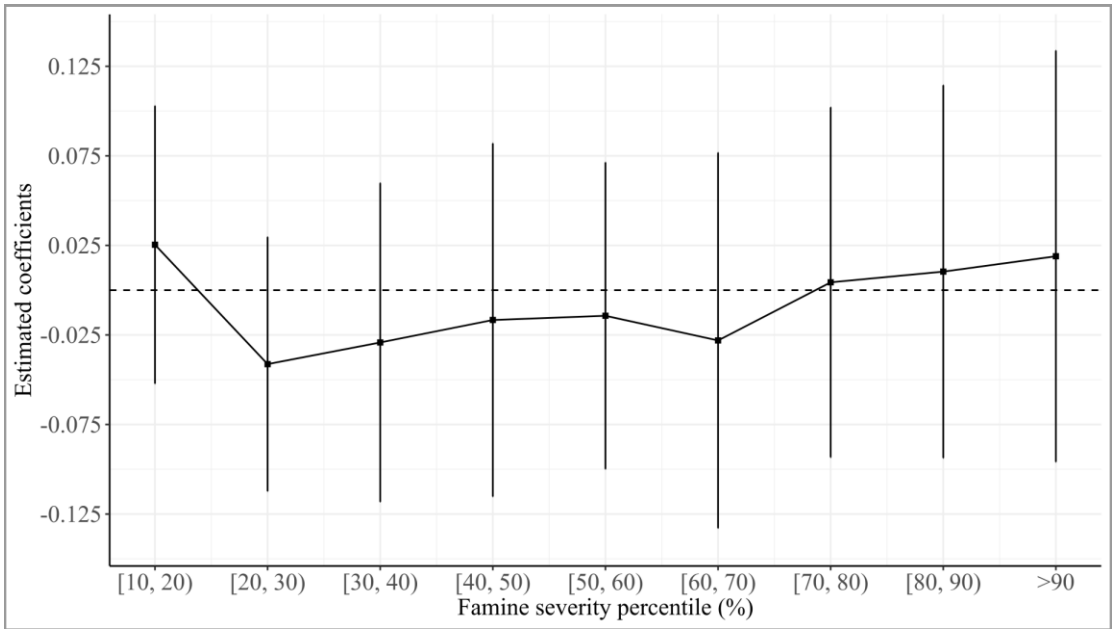
Notes: Panel (A) and (B) depict the distribution of the gap between reported GDP and the "placebo targets": target GDP + 0.1% in Panel (A), and target GDP – 0.1% in Panel (B), respectively. Both distributions are smooth. However, Panel (C) depicts the distribution of the gap between reported GDP and the true targets: a jump exists around 0, which points to the possibility of manipulation.

Figure B2. Severity and Rectification



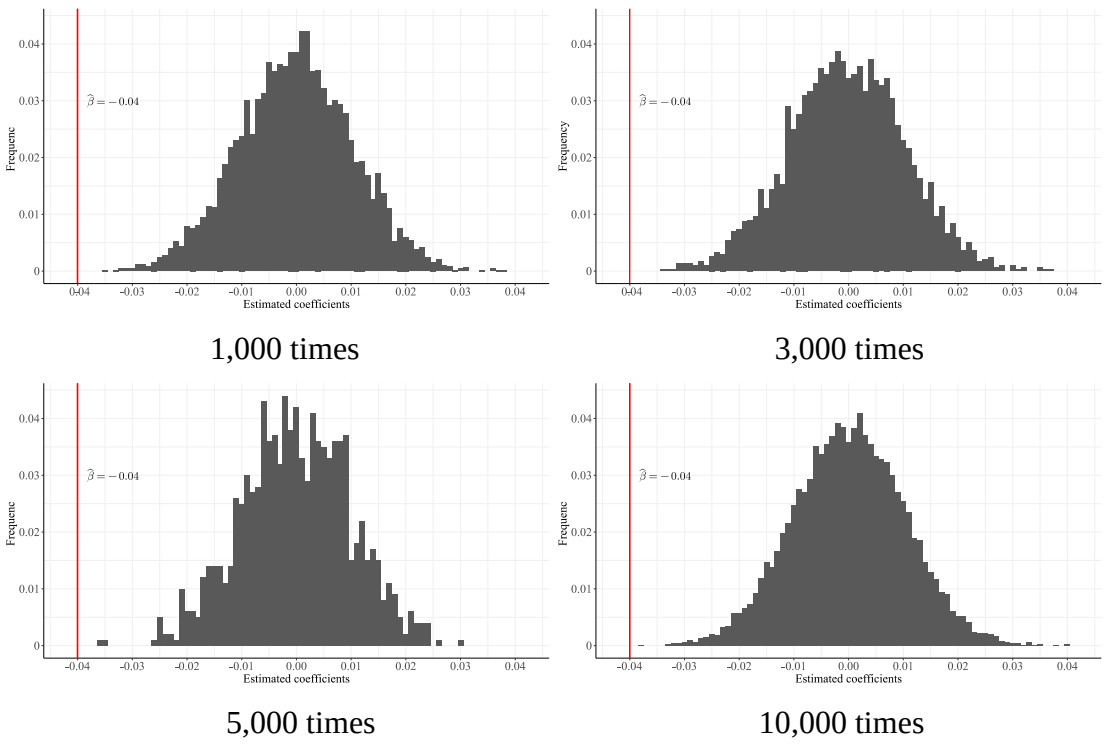
Notes: This figure replaces the binary famine exposure in the baseline to a generalized interaction term of the cohort dummies and deciles of famine severity, with the bottom 10% of the regions least hit by the famine as base group. As shown, when famine was not as severe, officials continued to lie in public offices. However, when the local famine was more severe, the officials became more reluctant to lie when they served in public offices later on.

Figure B3. Selection Effects of the Famine



Notes: This figure re-estimate Figure B2 for samples in which individuals were more than six years old when the famine hit – a group with more stable risk preferences, thus the test captures the selection effects of the famine. As shown, the impacts are insignificant regardless of famine severity. Therefore, our baseline results are less likely driven by selection or survivor's bias.

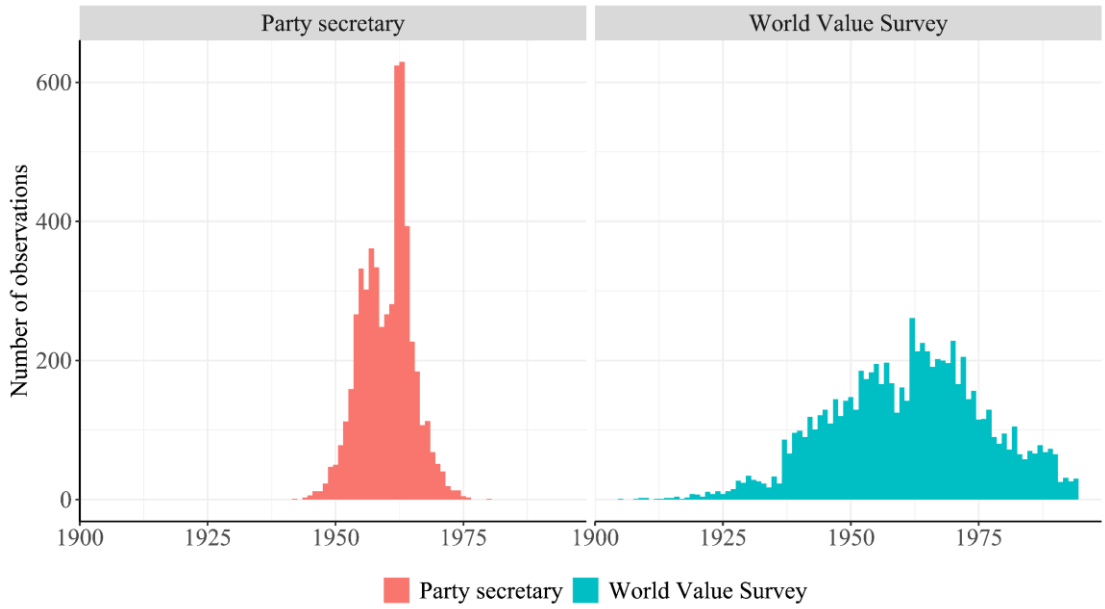
Figure B4. Random Sample Trial



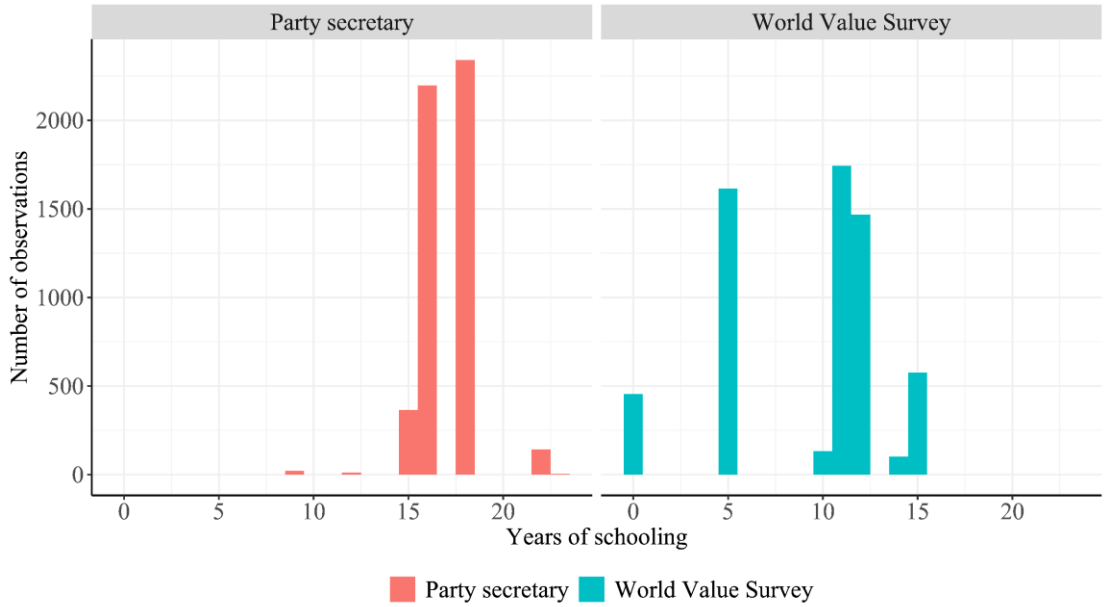
Notes: The figure depicts the distribution of coefficients in random sample trials. Specifically, we

randomly assign a famine severity to each county, and repeat our baseline according to the randomized famine. We replicate the random trials by 1,000, 3,000, 5,000, and 10,000 times, where Panel (A) to (D) present the distributions, respectively. Most of the coefficients are normally distributed, centring around zero, which suggests that our results are unlikely driven by spurious regression.

Figure B5. Sample differences: Local officials vs. WVS



(A) Year of birth



(B) Years of schooling

Notes: Panel (A) and (B) compares the systematic differences of the officials' cohort and the WVS respondents, in year of birth and educational levels. Such systematic differences urge us to use the propensity weighting method by Hirano and Imbens (2001) and Hirano et al. (2003) to address the issue in Table 3.

Table B1. Direction of Manipulation

	(1) Absolute Value	(2) Over-reports	(3) Under-reports
Childhood × Severity	-0.020	-0.039	-0.008
	(0.009)	(0.018)	(0.015)
Observations	5,076	2,614	2,462
R-squared	0.348	0.498	0.525
Personal features	YES	YES	YES
Birth place Fixed Effects	YES	YES	YES
Birth year Fixed Effects	YES	YES	YES
Provincial Time Trend	YES	YES	YES

Notes: The econometric specification is similar with the baseline. In Column 1, the dependent variable is the absolute value of the manipulation, while Column 2 and 3 reports the results on the sub-sample of positive and negative manipulation, respectively. As shown, famine exposure mainly restrains manipulation driven by over-reports, instead of under-reports.

Table B2. Concentrated Sample

	(1) 200 samples around cut-off	(2) 250 samples around cut-off
Childhood famine exposure	-0.050	-0.045
	(0.027)	(0.024)
Gender	-0.016	-0.015
	(0.011)	(0.016)
Education	-0.006	-0.002
	(0.001)	(0.004)
Constant	0.004	0.007
	(0.011)	(0.051)
Observation	200	250
R-squared	0.601	0.578
Birth place Fixed Effects	YES	YES
Birth year Fixed Effects	YES	YES
Provincial Time Trend	YES	YES

Notes: This table re-estimates the baseline using two concentrated sub-samples, to rule out the impact of outliers: Column 1 with 200 samples around the cut-off where reported GDP equals target GDP; and Column 2 the other one with 250 samples. As shown, famine exposure restrains manipulation in both cases.

Table B3. Additional controls

	(1)	(2)	(3)	(4)	(5)	(6)
Famine exposure	-0.025	-0.028	-0.023	-0.023	-0.020	-0.027
	(0.010)	(0.010)	(0.012)	(0.013)	(0.010)	(0.010)
First year				0.005	0.003	0.003
				(0.003)	(0.002)	(0.002)
Years in office					-0.001	-0.002
					(0.001)	(0.001)
Retired						-0.031
						(0.019)
Observations	14,458	13,776	9,773	9,773	9,773	9,773
R-squared	0.400	0.405	0.491	0.491	0.491	0.491
FE's	YES	YES	YES	YES	YES	YES
Personal characteristics	NO	YES	YES	YES	YES	YES
Socioeconomic indicators	NO	NO	YES	YES	YES	YES

Notes: Column 1 presents the results without additional controls. Column 2-3 gradually add the individual characteristics (gender and educational levels) and socioeconomic indicators (logged local population and logged GDP per capita) as additional controls, while Column 4-6 add the features of officials' career concerns. Column 4 adds a dummy taking value one if the official is in the first year of the tenure, and zero otherwise. This captures the extra incentives upon taking offices. Column 5 adds the number of years under which the official serves as the county party secretary, and Column 6 adds a dummy taking value one if the official is above 50 years old – and thus in the later stage of the career – and zero otherwise. The cut-off age of 50 years follows that in Yao and Zhang (2015). As shown, the effect of famine exposure remains robust in all columns.