A Censored Maximum Likelihood Approach to Quantifying Manipulation in China's Air Pollution Data

Dalia Ghanem, Shu Shen, Junjie Zhang

Abstract: Data manipulation around cutoff points is observed in economics broadly and in environmental and resource economics in particular. This paper develops a simple and tractable censored maximum likelihood approach to quantify the degree of manipulation in China's air pollution data around the "blue-sky day" cutoff. We construct annual measures of manipulation for 111 Chinese cities. For Beijing, we estimate 4%–16.8% of manipulation among reported blue-sky days annually, which translate to an estimated total of 208.1 manipulated blue-sky days between 2001 and 2010. For the remaining cities reporting pollution data over the 10-year period, we estimate a 93.9 average for the total number of manipulated blue-sky days with a 395.9 maximum. Using LASSO shrinkage, we examine the relationship between manipulation and local official characteristics, and find a positive correlation between manipulation and having an elite-educated party secretary, robust to numerous checks. Further empirical analysis suggests that promotion considerations may help explain this finding.

JEL Codes: C24, Q52, Q53, Q54

Keywords: maximum likelihood, censoring, generalized beta distribution, environmental policy, pollution control, principal-agent problem

DATA MANIPULATION AROUND CERTAIN CUTOFF POINTS is observed widely in economics and other social sciences when agents, such as employees or government officials, are incentivized to meet specific cutoff values of performance evaluation

Dalia Ghanem is at the University of California, Davis (dghanem@ucdavis.edu). Shu Shen is at the University of California, Davis (shushen@ucdavis.edu). Junjie Zhang is at Duke Kunshan University and Duke University (junjie.zhang@duke.edu). We would like to thank Tim Beatty, Graham Elliott, Mark Jacobsen, and Kevin Novan for helpful discussions. Junjie Zhang thanks the support from the National Natural Science Foundation of China (project 71773043, 71773062).

Dataverse data: https://doi.org/10.7910/DVN/HJWORA

Received January 21, 2019; Accepted April 16, 2020; Published online August 4, 2020.

JAERE, volume 7, number 5. 2333-5955/2020/0705-00XX$10.00 https://doi.org/10.1086/709649

metrics that are self-reported. For example, Fisman and Wang (2017) find that Chinese government officials decrease accidental deaths to meet certain ceiling numbers used in their performance evaluations; Foremny et al. (2017) document that local governments in Spain inflate total local population data above cutoff values to obtain higher per capita grant allocations. Numerous studies have documented and measured manipulation in high-stakes test scores (Figlio and Getzler 2002; Jacob and Levitt 2003; Jacob 2005; Figlio 2006; Reback and Cullen 2006; Dee et al. 2011; Diamond and Persson 2016; Dee et al. 2019). In environmental economics, manipulation in principal-agent settings constitutes a major challenge to the effective implementation and evaluation of climate and pollution control policies (e.g., Blonz 2019; Cole et al. 2019). The extent of manipulation in this context is a key parameter from a policy evaluation perspective.

China's environmental regulatory system is an important setting where these principal-agent issues arise. In order to incentivize its local officials to improve air quality, the Chinese central government included the number of "blue-sky" days in a year as one of the metrics used in its evaluation of local officials. A blue-sky day is a day with Air Pollution Index (API) below 101. Between 2001 and 2010, the API and its criteria pollutants were reported by data collection agencies overseen by the city governments. Numerous studies (Andrews 2008a, 2008b; Chen et al. 2012; Ghanem and Zhang 2014; Fu et al. 2014) document evidence of manipulation of China's air pollution data, especially PM_{10} concentrations, around the blue-sky day cutoff using a number of statistical tests (e.g., McCrary 2008).[1] However, these studies do not quantify the magnitude of manipulation. Furthermore, highly statistically significant evidence of its existence does not imply a larger magnitude of manipulation. As a result, the extent to which China's policy truly increased the number of blue-sky days remains an open question. The goal of this paper is to shed light on this question by constructing a panel of annual manipulation measures of air pollution data for 111 reporting Chinese cities between 2001 and 2010. We further illustrate that accounting for manipulation is crucial to identify and estimate the effect of any policy aimed at curbing air pollution during this time period.

In our analysis of city-level air quality data in China, we develop a censored maximum likelihood estimation (MLE) procedure to quantify the magnitude of data manipulation around the policy cutoff. Specifically, we formalize the problem of identifying manipulation measures with the potential outcomes framework (Rubin 1974) and show that in

1. Fu et al. (2014) and Stoerk (2016) use Benford's law to test for the presence of manipulation in China's air quality data. However, this is a general test for data irregularity and does not specifically test for threshold manipulation. To test for the presence of a discontinuity, Chen et al. (2012) use the McCrary (2008) test as well as the Burgstahler and Dichev (1997) test. The latter test exploits a property of the cdf of an i.i.d. random variable which holds if the distribution is continuous. Specifically, when considering equally spaced bins, the probability that the random variable lies within a specific one equals the average of the neighboring bins by the continuity of the distribution as shown in Takeuchi (2004).

the presence of manipulation around a cutoff point we can observe an interval-censored version of the true pollutant concentration. Hence, censored MLE is a convenient approach to estimation. The key assumptions we impose are that manipulation only occurs within a window around the cutoff, which we refer to as the manipulation window, and that the direction of manipulation is known. Both assumptions are motivated by our empirical setting and have been imposed explicitly or implicitly in previous work quantifying threshold manipulation in other settings (e.g., Diamond and Persson 2016; Foremny et al. 2017; Dee et al. 2019). In our implementation of the censored MLE procedure, we use a flexible class of distributions well suited for positive, continuous random variables, the generalized beta distribution of the second kind (GB2) (McDonald 1984; McDonald and Mantrala 1995). It nests the lognormal, Gamma and Weibull distributions, which were found to fit the distribution of pollutant concentrations well (e.g., Holland and Fitz-Simons 1982).

Using our proposed method, we estimate two manipulation measures for PM_{10} concentrations for all city-year combinations available in our data set: (i) the proportion of manipulation among blue-sky days and (ii) the number of manipulated blue-sky days. In a detailed analysis of Beijing's data, we find that between 2001 and 2010, 4.0%–16.8% of Beijing's blue-sky days in a year were manipulated to meet the blue-sky day cutoff. The highest proportion of manipulation we estimate is in 2006 and 2007, whereas the lowest is in 2008, the year of the Beijing Olympic Games. These proportions allow us to estimate the number of manipulated blue-sky days in Beijing, which range from 10.7 to 41.3 per year and sum to a total of 208.1 during the 10-year period. When we zoom in to days that are manipulable (i.e., days with air quality readings close to yet exceeding the blue-sky day cutoff), we find very high rates of data manipulation among those days.

When we estimate the two manipulation measures for all reporting cities in our data set, we document an average 3.1% of manipulation among reported blue-sky days per year, whereas the average number of manipulated blue-sky days in a year is 8.3. Although these averages may seem modest, the distribution of both measures has a long right tail. The maximum annual proportion of manipulation among blue-sky days we estimate is 30.5%, and the maximum number of manipulated blue-sky days is 83.5 in a year. We also find substantial temporal and geographical heterogeneity in the magnitude of manipulation around the blue-sky day cutoff.

Finally, since the number of blue-sky days is one of the performance criteria for local officials in China, we investigate the possibility of correlation between local official characteristics and the heterogeneous manipulation behavior we document. To do so, we use LASSO to select the key predictors of manipulation among a large set of demographic, education, and experience variables for both party secretaries and mayors. It is important to emphasize that our implementation of the LASSO procedure here is purely as a predictive method to examine the correlation between local official characteristics and manipulation behavior. The most pronounced result we find in our LASSO analysis

is that having an elite-educated party secretary in power is positively correlated with city-level manipulation. We specifically find that having such a secretary is associated with a statistically significant 1% increase in the proportion of manipulation among blue-sky days, which is 30% of the mean of this proportion in our sample. This result is robust to numerous checks of our LASSO strategy.

In order to interpret this seemingly surprising correlation, we examine the relationship between manipulation and future promotion for elite-educated as well as other party secretaries. We find that the correlation between manipulation and future promotion to certain province-level positions is positive and larger for the former relative to the latter. We find similar patterns in the correlation between GDP and manipulation for elite-educated and other party secretaries. While this analysis suggests that promotion incentives and a prioritization of economic growth may be possible explanations for the correlation between manipulation and party secretary's elite education status, there are other factors in this complex administrative environment that we cannot separate in our data. A thorough examination of these issues is beyond the scope of this paper and constitutes a priority for future work.

The first contribution of this paper is methodological. While motivated by environmental data manipulation in China, our censored MLE approach is general and can be used to measure the extent of threshold manipulation in other principal-agent settings as well as in regression discontinuity designs (see, e.g., Imbens and Lemieux 2008). The proposed procedure has several advantages over the polynomial fitting approach adopted from the excess bunching literature (cf. Chetty et al. 2011). Papers employing this procedure estimate the counterfactual data distribution using polynomials, which can lead to noisy fits sensitive to the degree of polynomials.[2] In addition, the resulting densities are not guaranteed to be nonnegative and may not even integrate to one. We provide a clear empirical illustration of those issues in section 1.4. As a result, practitioners implementing the method often need to attempt polynomials of different orders to ensure that general properties of distribution functions will not be violated. This specification searching leads to the well-known post-selection inference problem and hence invalidates standard inference procedures, including the parametric bootstrap method adopted in the empirical literature. Our procedure is better suited to estimating counterfactual data distributions and can obviate the need for specification searching by employing a flexible class of distributions in the MLE, such as the GB2 class employed in our air pollution application.

This paper also contributes to the literature on data manipulation in principal-agent settings in environmental and resource economics (Blonz 2019; Cole et al. 2019) broadly as well as to the literature on air quality data manipulation in China (Andrews 2008a, 2008b; Chen et al. 2012; Fu et al. 2014; Ghanem and Zhang 2014; Liang et al. 2016;

2. Similar arguments about the flaws of polynomial fitting are also made by Gelman and Imbens (2018) in the context of regression discontinuity designs

Stoerk 2016) more specifically. While the aforementioned literature found statistical evidence consistent with manipulation for many cities in China prior to 2013, the McCrary (2008) statistic used in this literature is not monotonic in the degree of manipulation.[3] Hence, it cannot be used to quantify manipulation. This paper fills this gap in the literature by providing annual measures of manipulation for reporting cities during that period. It thereby sheds light on the degree of effectiveness of China's policy to increase the number of blue-sky days in the first decade of the twenty-first century.

Our analysis also highlights the importance of accounting for manipulation in evaluating environmental policies. To provide an example, according to the reported data, the proportion of blue-sky days in Beijing sees a 7% increase in 2008, when a bundle of permanent and temporary pollution control policies were put in place for the Olympic Games. Using our censored MLE approach, we estimate the counterfactual distribution of Beijing's air quality data and document a 15% increase in the proportion of blue-sky days between 2007 and 2008. Our estimated manipulation measures suggest that the difference between our estimated 15% increase compared to the 7% increase obtained from the reported data is driven by substantial overreporting of blue-sky days in 2007.

Finally, our empirical analysis relates to the literature on political economy considerations in environmental regulation as well as the literature on meritocratic promotion and its unintended consequences. The Chinese central government often uses career advancement incentives to induce desirable economic, social, and environmental outcomes from local officials (Li and Zhou 2005; Xu 2011). A growing body of literature questions whether meritocratic promotions can effectively lead to these desired outcomes and points to their potential unintended consequences (Kahn et al. 2015; Fisman and Wang 2017; Jia 2017; Shi et al. 2020). Our paper provides further evidence that some officials' characteristics that matter for promotion are also correlated with manipulation. In particular, education level was documented as an important determinant of promotion (Shih et al. 2012).

The rest of the paper is organized as follows. Section 1 presents the identification and estimation of manipulation measures using our proposed censored MLE method and compares it with the polynomial fitting approach. Section 2 provides annual measures of manipulation for all cities in our data set over the 10-year period we examine. Section 3 examines the predictors of manipulation among local official characteristics. Section 4 concludes.

1. ESTIMATING MANIPULATION MEASURES VIA CENSORED MLE

In this section, we first describe our air pollutant concentration data and background. Then, we outline our strategy to identify and estimate measures of manipulation in China's air pollution data around the blue-sky day cutoff.

3. For a graphical illustration of this point, see fig. A1 (figs. A1–A12 are available online).

1.1. Background and Data Summary

China has utilized a unique approach of regular performance evaluation and promotion incentives to induce its local officials to comply with centrally mandated economic, social, and environmental targets. During 2001–10, the Chinese central government used the number of blue-sky days as one of the performance metrics to evaluate local officials. In particular, the tenth and eleventh Five-Year Plans (2001–5 and 2006–10) set specific targets for the annual proportion of blue-sky days.[4] This naturally leads to an incentive to manipulate API to be less than 101 as long as the manipulation is hard to detect by the public.

We use city-level daily PM_{10} concentrations (density of fine particles of diameter less than 10 micrometers, reported in mg/m^3) for 111 cities from 2001 to 2010, which are collected by the China National Environmental Monitoring Center (CNEMC), an affiliate of the Ministry of Ecology and Environment of China (formerly Ministry of Environmental Protection). One of the advantages of using 2001–10 as our sample period is that the air pollution standard (GB3905-1996 rev) does not vary during these years. The cutoff for blue-sky days for PM_{10} concentrations is 0.15 mg/m^3. The API is a dimensionless index composed of three criteria pollutants: PM_{10}, SO_2, and NO_2 (see Ghanem and Zhang [2014] for details on the construction of the API). We focus on PM_{10} since it was the dominant pollutant between 2001 and 2010. Specifically, 74% of the reported non-blue-sky days are classified as such due to the reported PM_{10} concentrations.

Our data set is a mere compilation of the daily PM_{10} concentration data reported by the city governments. In 2001, 47 cities including Beijing started to report their air pollutant concentrations. The number of reporting cities increases to 107 in 2003 and stabilizes at 111 starting in 2006. After a city has started to report air pollutant concentrations, there are still occasional missing values since a city is required to use a minimum of 12 hours of effective monitoring to report a daily concentration for PM_{10}. However, once a city starts reporting, the number of missing days is negligible, as illustrated in table 1. For instance, in 2010, we only have 6 missing days for all 111 cities in total.

In addition to the number of city-day observations, table 1 also reports key summary statistics for daily PM_{10} concentrations for all cities as they roll into our data set as well as for the 47 cities that started to report daily PM_{10} concentrations in 2001. We find that the mean, standard deviation, and maximum of daily PM_{10} concentrations have declined modestly over the sample period.

Next, to motivate our analysis of data manipulation, we provide several interesting summary statistics of Beijing's PM_{10} concentrations in figure 1: (1) the annual proportion of days with excellent air quality (i.e., days with PM_{10} concentrations below 0.05 mg/m^3 corresponding to an API of 50 or less), (2) the annual proportion of blue-sky days (i.e., days with PM_{10} concentrations below 0.15 mg/m^3 corresponding to an API of 100 or

4. Source: http://www.gov.cn/zhengce/content/2008-03/28/content_4877.htm (in Chinese; retrieved on March 8, 2018).

Table 1. Summary Statistics for Daily PM_{10} Concentrations

	Full Data Sample					Cities Reporting since 2001				
Year	Observations	Mean	SD	Min	Max	Observations	Mean	SD	Min	Max
2001	15,127	.112	.084	.003	2.561					
2002	26,048	.120	.093	.004	2.010	17,151	.109	.086	.004	1.667
2003	36,052	.117	.078	.001	2.720	17,138	.101	.066	.004	.972
2004	39,827	.112	.070	.006	2.042	17,202	.098	.064	.006	1.274
2005	39,782	.099	.061	.006	2.410	17,154	.092	.063	.006	2.410
2006	40,514	.100	.064	.003	1.860	17,155	.095	.069	.003	1.860
2007	40,404	.094	.054	.007	1.348	17,108	.091	.059	.007	1.348
2008	40,622	.088	.052	.006	1.511	17,201	.088	.057	.006	1.195
2009	40,504	.087	.052	.003	1.764	17,154	.086	.058	.003	1.764
2010	40,509	.088	.056	.004	1.230	17,154	.087	.062	.004	1.230

Note. The table reports the number of observations as well as summary statistics for daily PM_{10} concentrations in our sample. We report the annual summary statistics for our full data sample as well as for the subsample of 47 cities that started reporting in 2001. Note that there were 366 days in 2004 and 2008.

less), and (3) the annual proportion of days with unhealthy air quality (i.e., days with PM_{10} concentrations exceeding 0.35 mg/m^3 corresponding to an API of 200 or higher). We first note that both the reported blue-sky days and days with excellent air quality experience a substantial increase in 2008, when a bundle of permanent and temporary policies were put in place to curb air pollution during the Olympic Games. These policies included driving restrictions as well as relocation of heavy-polluting industries.

As discussed earlier, the proportion of blue-sky days is an important environmental target linked to performance evaluation of local officials during the time period we examine. Since local officials in China are expected to show continuous improvement in performance evaluation metrics (Xu 2011), it is not surprising to see that it is the only series that has a clear upward trend over our entire sample period. The proportions of days with excellent and unhealthy air quality, on the other hand, oscillate from year to year. This motivates the following sections where we design an identification and estimation strategy to quantify city-specific annual measures of manipulation around the blue-sky day cutoff.

1.2. Identification of Manipulation Measures

Let X be the observed air pollutant concentration that is subject to potential manipulation, and c be the cutoff that incentivizes data manipulation, that is, the blue-sky cutoff. The observed pollutant concentration X is a combination of true and manipulated data. Formally, $X = (1 - Z)X(0) + ZX(1)$, where $X(0)$ is true data, $X(1)$ is manipulated data, and Z is an unobserved binary indicator for data manipulation. We are interested in quantifying the magnitude of data manipulation among blue-sky days

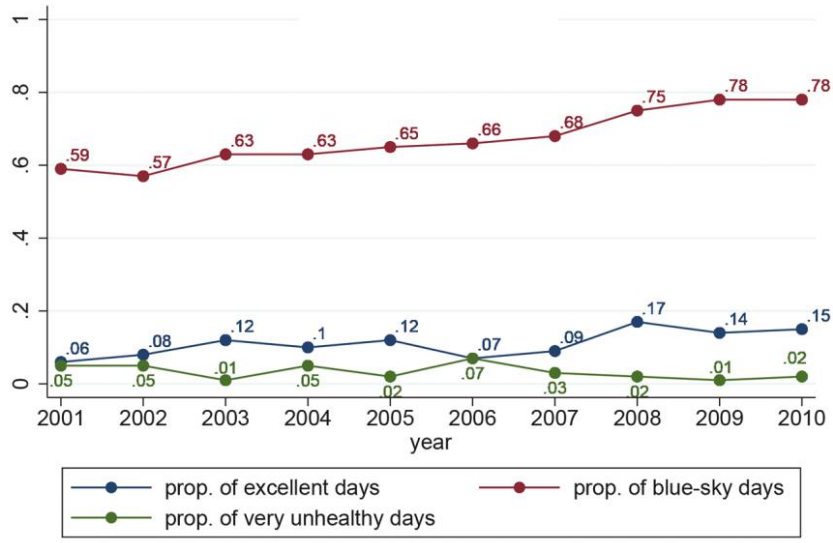


Figure 1. Proportions of days with excellent, blue-sky, and unhealthy air quality. The figure presents the time series plots for the annual proportion of days with excellent air quality (excellent: PM_{10} concentrations less than 0.05 mg/m^3), the annual proportion of blue-sky days (blue-sky: PM_{10} concentrations less than 0.15 mg/m^3), and the annual proportion of unhealthy days (unhealthy: PM_{10} concentrations larger than 0.35 mg/m^3).

and several other measures of the degree of manipulation. To establish identification, we make the following assumptions motivated by our empirical setting.

Assumption 1: $Z = 0$ if $X(0) \leq c$; $X(1) \leq c$.

Assumption 2: $P(Z = 1|X(0) = x) = 0$ for all $x \notin [\underline{x}, \bar{x}]$, where $c \in [\underline{x}, \bar{x}]$ and $0 < P(\underline{x} \leq X(0) \leq \bar{x}) < 1$.

Assumption 3: The cumulative distribution function (cdf) of $X(0)$ is $G(.; \theta)$, where $G(.; \theta)$ is a known function with density $g(.; \theta)$ and θ is an unknown finite-dimensional parameter.

Figure 2 provides a graphical illustration of the above assumptions. The first assumption implies that data manipulation is unidirectional with known direction and that manipulation moves data across the cutoff value c. This assumption is plausible since the number of blue-sky days is used as part of the performance evaluation of city officials, and therefore it is reasonable to expect pollutant concentrations to be manipulated downward and across the blue-sky threshold. There is no marginal benefit for the local

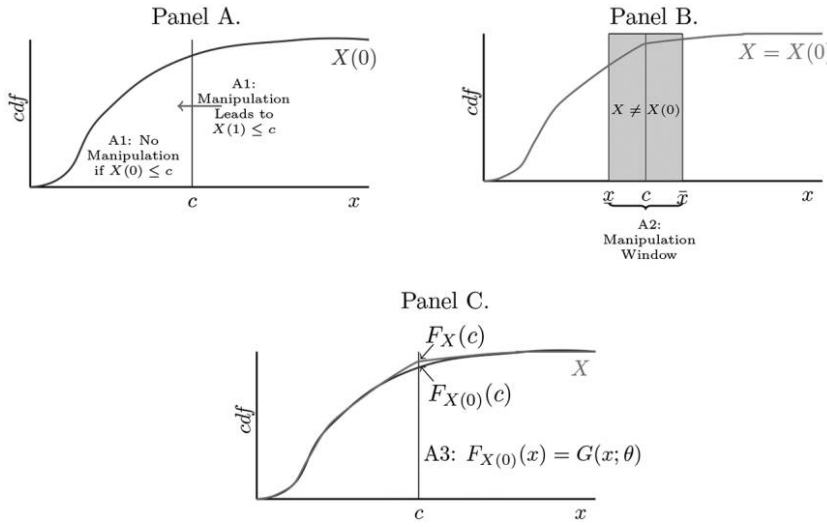


Figure 2. Graphical illustration of assumptions. Assumptions 1–3 are referred to as A1–A3 in this figure. The figure illustrates how the three assumptions together allow us to identify the proportion of manipulation, that is, $\lambda \equiv P(Z = 1) = F_X(c) - F_{X(0)}(c) = F_X(c) - G(c;\theta)$. Color version available as an online enhancement.

official to manipulate an observation that already meets the target or to manipulate an observation without making it meet the prescribed cutoff.

The identification of the unmanipulated data distribution follows from assumptions 2 and 3. Assumption 2 restricts manipulation to occur only within a manipulation window and implies that an interval-censored version of the true pollutant concentration is observed by the researcher, as illustrated in panel B of figure 2. Formally, $X(0) = X$ if $X \notin [\underline{x}, \bar{x}]$, and $X(0) \in [\underline{x}, \bar{x}]$ otherwise. This assumption is plausible in our setup for two reasons. First, to ensure that the public cannot detect air pollution manipulation, we do not expect manipulation to occur if the pollutant concentration exceeds the cutoff for blue-sky days by a large amount. For the same reason, we expect manipulated data to be close to the blue-sky day cutoff.

Finally, assumption 3 imposes that the class of parametric distributions, $G(.;\theta)$, for the true pollutant concentration is known, which is referred to as the counterfactual data distribution in the literature. Since θ is finite-dimensional, $G(.;\theta)$ can be identified from the interval-censored version of $X(0)$ and we can estimate it by censored MLE. In this paper, we use the generalized beta distribution of the second kind (GB2), a large class of distributions nesting the lognormal, Gamma and Weibull, which were found to fit the distribution of pollutant concentrations well (e.g., Holland and Fitz-Simons 1982). This class of distributions was developed to model income distributions and hence is well suited for positive, continuous random variables (McDonald 1984; McDonald

and Mantrala 1995). Furthermore, the distributional assumption is testable in our setup by comparing the fit provided by the parametric distribution with the empirical cdf outside the manipulation window using appropriate distance measures, such as Kolmogorov-Smirnov.

Given the above assumptions, it is straightforward to show that the overall proportion of manipulation λ is equal to the difference between the observed and counterfactual distributions of the pollutant concentration evaluated at the cutoff c. Specifically,

$$\begin{aligned}\lambda &\equiv P(Z = 1) = P(X(1) \leq c, Z = 1)\\ &= P(X(0) \leq c, Z = 0) + P(X(1) \leq c, Z = 1) - P(X(0) \leq c, Z = 0)\\ &= P(X \leq c) - P(X(0) \leq c) = F_X(c) - G(c; \theta).\end{aligned}$$

The first equality follows from the definition of X. The third equality follows from assumption 1 since there is no manipulation if the true data are already below the cutoff c and all manipulation moves data below c. The last equality holds following the definition of $G(.; \theta)$ and λ.

Once λ is identified, we can identify the proportion of manipulation among blue-sky days, that is, the proportion of manipulation among days reportedly meeting the performance cutoff,

$$\mu \equiv P(Z = 1|X \leq c) = \frac{\lambda}{F_X(c)}.$$

Since manipulation only occurs within the manipulation window, another measure of manipulation that could be of interest is the proportion of manipulation among all manipulable data (of the true air pollutant concentration),

$$\nu \equiv P(Z = 1|c < X(0) \leq \bar{x}) = \frac{\lambda}{G(\bar{x}; \theta) - G(c; \theta)},$$

which is also called the proportion of in-range manipulation in the literature (see, e.g., Dee et al. 2019).

1.3. Estimation

Let X_{itd} denote the reported PM_{10} concentration in city i on day d in year t, which is possibly manipulated within the window $[\underline{x}, \bar{x}]$. Let $G(.; \theta_{it})$ be the counterfactual cdf of the unmanipulated PM_{10}, which belongs to the GB2 class by assumption, and $g(.; \theta_{it})$ be its corresponding pdf. The distributions are allowed to vary across cities and years. Let T_{it} be the total number of days observed in year t for city i. The parameter θ_{it} can be consistently estimated for each city i and year t by $\hat{\theta}_{it}$, the maximizer to the MLE objective function,

$$\sum_{d=1}^{T_{it}} \{1\{x_{itd} \notin [\underline{x}, \bar{x}]\} \log g(x_{itd}; \theta_{it}) + 1\{x_{itd} \in [\underline{x}, \bar{x}]\} \log(G(\bar{x}; \theta_{it}) - G(\underline{x}; \theta_{it}))\}.$$

We implement the estimation in Stata. Our codes for the censored GB2 objective function are modified from the GB2LFIT Stata module developed by Jenkins (2014).

The manipulation window, $[\underline{x}, \bar{x}]$, is specified by the user and should be informed by the empirical context. In our setup, we expect manipulation to move pollutant concentrations to values just below the cutoff of 0.15, because there is no incentive for officials to further underreport the data once the blue-sky day cutoff is met. We hence choose our baseline manipulation window to be asymmetric, with the length of the interval above the cutoff to be double its length below the cutoff. In general, using a wide manipulation window can ensure its inclusion of all manipulated data. However, it can negatively impact estimation efficiency through censoring out proportionately more observations. Given the heterogeneity of the PM_{10} distributions across the different cities and years in our data set, we set our baseline manipulation window to be [0.135, 0.18]. This window choice is large enough to capture possibly manipulated data but yet ensures that there is a sufficient number of observations in the continuous part of the different distributions we examine. We also report results using alternative window choices to illustrate the robustness of our estimation results (see app. C; apps. A–C are available online).

Given the estimator of the counterfactual PM_{10} distribution, $G(.; \hat{\theta}_{it})$, the estimation of the different measures of manipulation we examine is straightforward. The annual proportion of manipulation, or the unconditional probability of manipulation among all days in a year, is estimated by

$$\hat{\lambda}_{it} = \hat{F}_{X_{it}}(c) - G(c; \hat{\theta}_{it}),$$

where $\hat{F}_{X_{it}}(c)$ denotes the empirical cdf of daily PM_{10} concentrations in city i in year t evaluated at the cutoff c. The proportion of manipulation among all reported blue-sky days in a year is estimated by

$$\hat{\mu}_{it} = \hat{\lambda}_{it} / \hat{F}_{X_{it}}(c).$$

We can also use $\hat{\lambda}_{it}$ to estimate the number of manipulated blue-sky days, or $\hat{m}_{it} = \hat{\lambda}_{it} T_{it}$. Finally, the proportion of in-range manipulation, or the proportion of manipulation among manipulable non-blue-sky days, is estimated by

$$\hat{\nu}_{it} = \hat{\lambda}_{it} / \left(G(\bar{x}; \hat{\theta}_{it}) - G(c; \hat{\theta}_{it}) \right).$$

1.4. Comparison with Polynomial Fitting

As discussed in the introduction, the predominant approach to quantify data manipulation in the applied economics literature is polynomial fitting. Here, we provide an empirical illustration of the issues we raised in the introduction regarding this method.

Figure 3 reports the estimated counterfactual distribution of PM_{10} concentrations for Beijing and Shanghai in 2007 using both the proposed censored MLE approach and the polynomial fitting approach adopted in Dee et al. (2019) and Foremny et al.

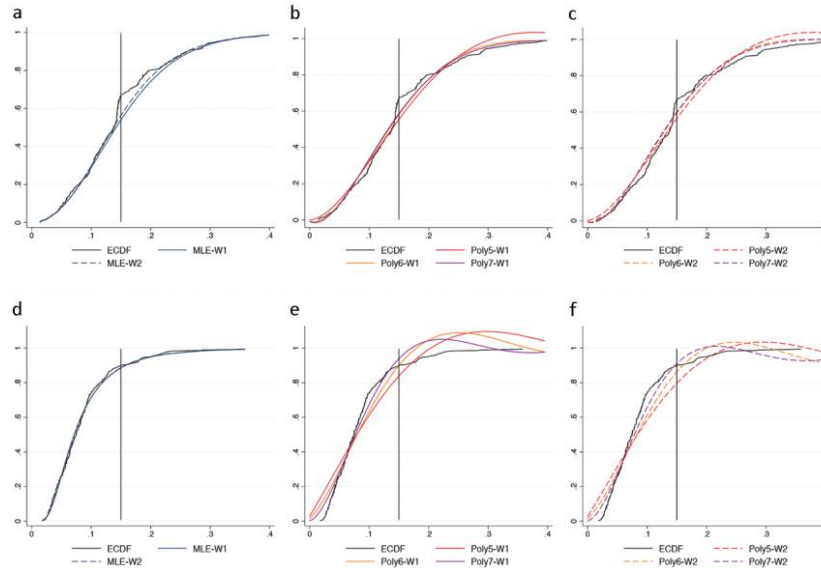


Figure 3. Examples of counterfactual distribution estimates: PM_{10} concentration of Beijing and Shanghai 2007. ECDF denotes the empirical cdf of the observed data, MLE-W1 and MLE-W2 denote the censored MLE estimator of the counterfactual cdf using two different manipulation windows [0.135, 0.22] (W1) and [0.135, 0.18] (W2), respectively. For $q = 5, 6, 7$, Poly$q$-W1 and Poly$q$-W2 denote the qth order polynomial estimator of the counterfactual cdf using manipulation windows W1 and W2, respectively. The vertical line marks the blue-sky day cutoff of $c = 0.15$.

(2017). We include similar figures for all other years in our sample period in the appendix.

The polynomial fitting approach in Dee et al. (2019) and Foremny et al. (2017) first bins the observed data to obtain discretized frequencies ($p_{s,it}$). The discretized frequencies are then regressed on high-order polynomials of the center points of the bins (C_s) using data outside the manipulation window. Specifically,

$$p_{s,it} = \sum_{k=1}^{q} \pi_{k,it}(C_s)^k + \sum_{k \in \mathcal{M}_{[\underline{x},\bar{x}]}} \gamma_{k,it} 1\{C_s = k\} + \epsilon_{s,it},$$

where $\mathcal{M}_{[\underline{x},\bar{x}]}$ is the set of discretized concentrations that fall inside the manipulation window. In our application, the reported PM_{10} concentrations are discretized into 0.005 mg/m^3 bins and $\mathcal{M}_{[\underline{x},\bar{x}]} = \{\underline{x}, \underline{x} + 0.005, \underline{x} + 0.01, \ldots, \bar{x}\}$. Given the regression estimates $\{\hat{\pi}_{k,it}\}_{k=1}^{q}$, the counterfactual PM_{10} distribution is estimated by $\hat{F}_{X_{it}(0)}(C_s; \hat{\pi}_{1,it}, \ldots, \hat{\pi}_{q,it}) = \Sigma_{j=1}^{s}\Sigma_{k=1}^{q}\hat{\pi}_{k,it}(C_j)^k$, which can be used to construct different measures of data manipulation discussed in the previous section.

The plots in figure 3 illustrate that the MLE approach using the GB2 distribution fits the data outside the manipulation window very well, which is not surprising given the flexibility of this distribution class and its suitability for pollutant concentrations. Furthermore, the fit provided by the GB2 distribution is not very sensitive to the choice of manipulation window. Recall that according to our assumptions, the counterfactual and the observed distribution should agree outside the manipulation window. Figure 3 illustrates that the estimated counterfactual distribution provides a similarly good fit to the empirical cdf outside the manipulation window for both window choices we consider. As a result, the estimate of λ, which is the vertical distance between the empirical cdf and the fitted distribution using censored MLE evaluated at the threshold, that is, $\hat{F}_{X_{it}}(c) - G(c; \hat{\theta}_{it})$, is very similar regardless of the choice of manipulation window.[5]

The polynomial approach, on the other hand, provides a somewhat noisy fit for Beijing's distribution and a very poor fit for Shanghai's. In general, our empirical illustration shows that there is no guarantee that the estimator of the unmanipulated data distribution obtained from the polynomial fitting approach will satisfy general properties of distribution functions. As a result, specification searching is often necessary to provide a reasonable fit for the data distribution outside the manipulation window. Figure 3 also illustrates that the fit of the polynomial estimates can be very sensitive to the choice of manipulation window relative to the censored MLE approach. The resulting estimated proportion of manipulation would hence suffer from these same issues.

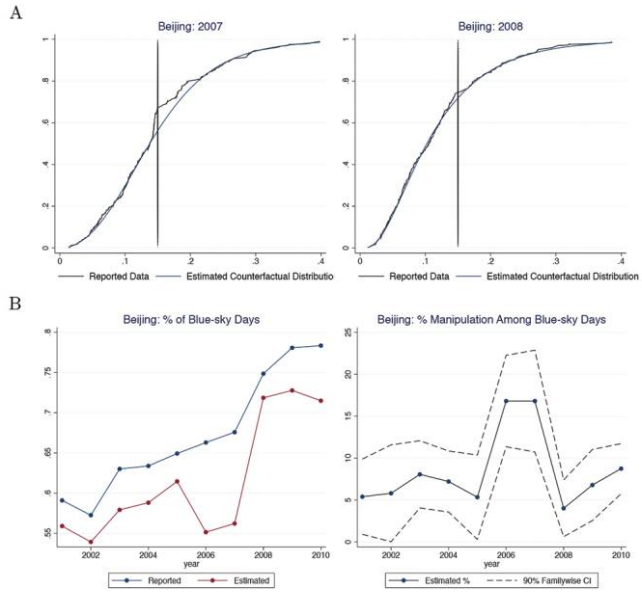
2. QUANTIFYING MANIPULATION IN CHINA'S AIR POLLUTION DATA

In this section, we present our estimation results for different measures of manipulation for all cities and years in our sample period, subject to data availability. We apply our censored MLE procedure to estimate $G(.; \theta_{it})$ for cities i and years t using the manipulation window [0.135, 0.18]. Given the importance of Beijing and its air quality issues, we first provide a detailed analysis of manipulation of Beijing's PM_{10} concentrations and then summarize the results for all cities in our data set.

2.1. An Analysis of Beijing's Air Pollution Data

Panel A of figure 4 presents our estimates of the counterfactual PM_{10} distribution for Beijing in 2007 and 2008. The two graphs put the manipulation behavior in these adjacent years into sharp contrast. Using the empirical distribution of PM_{10}, Beijing reported 68% blue-sky days in 2007 and 75% blue-sky days in 2008. However, if we take data manipulation into account by using the estimated counterfactual distribution, the proportion of blue-sky days we estimate increased from 57% to 72% in 2008, the year of

5. For censored MLE estimation results with several alternative manipulation windows, please refer to fig. A10.



C

	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	Total
# of Manipulated Blue-sky Days	10.7	12.1	18.6	16.7	12.6	40.7	41.3	11.0	19.3	25.0	208.1

Figure 4. Manipulation measures for Beijing's PM_{10} concentrations. *A*, Estimation of counterfactual data distribution; panel *A* plots the empirical cdf and the censored MLE counterfactual cdf of PM_{10} in 2007 and 2008. *B*, Proportion of reported and estimated blue-sky days; the left figure of panel *B* plots the time series of the proportion of blue-sky days in a year using the reported PM_{10} concentration data and the estimated true PM_{10} concentration distribution. The right figure of panel *B* plots the time series of the estimated proportion of manipulation among blue-sky days in a year. The 90% family-wise confidence intervals are calculated using a nonoverlapping block-bootstrap procedure with 7-day block length and 200 bootstrap replications. Panel *C* reports the number of manipulated blue-sky days in Beijing for each year in our sample period, which is calculated by $\hat{m}_{it} = (\hat{F}_{X_{it}}(0.15) - G(0.15; \hat{\theta}_{it})) \cdot T_{it}$ as defined in section 1.3, where i denotes Beijing.

the Olympic Games. Our estimation results suggest that Beijing drastically manipulated air quality data around the blue-sky day cutoff in 2007. Interestingly, as a by-product, our results also indicate that the measures taken by Beijing to improve air quality during the Olympic Games were much more effective than the reported data imply, at least when measured by the number of blue-sky days.

Panel B of figure 4 presents the key manipulation measures we estimate for Beijing. The left graph plots the time series of Beijing's proportion of reported blue-sky

days, $\hat{F}_{X_{it}}(0.15)$, against the estimated proportion of unmanipulated blue-sky days, $G(0.15; \hat{\theta}_{it})$. According to the estimated counterfactual distribution, the proportion of blue-sky days in Beijing does not show clear signs of improvement before 2008, unlike the reported data, which imply a steadily increasing proportion of blue-sky days over the 10 years. The right graph reports the proportions of manipulation among blue-sky days over the 10-year span as well as the 90% family-wise confidence band. All estimated proportions are statistically significant.[6] The smallest proportion of manipulation we estimate for Beijing is 4% in 2008, whereas the largest estimated proportion is 16.8% in 2006 and 2007.[7] Using the estimated proportions of manipulation, we estimate the number of manipulated blue-sky days, which are reported in panel C. Our estimates range between 10.7 and 41.3 manipulated blue-sky days per year implying a total of 208.1 manipulated blue-sky days over the entire period.

As discussed in section 1, not all days that miss the blue-sky day cutoff are manipulable in our empirical setting. Our baseline manipulation window [0.135, 0.18] defines PM_{10} readings between the cutoff 0.15 and the upper bound 0.18 to be data that are manipulable. This allows us to compute the degree of manipulation among manipulable days. Figure 5 plots the estimated annual proportions of manipulable versus manipulated data as well as the proportion of in-range manipulation in Beijing. The figure illustrates that the proportion of manipulable data is relatively stable over our sample period, whereas the proportion of manipulation as well as in-range manipulation varies from year to year. We specifically find that in 2006–7 most manipulable days are manipulated to meet the blue-sky day cutoff. The proportions of in-range manipulation in those years exceed a striking proportion of 80%.

The manipulation measures in figures 4 and 5 shed light on the discrepancy between the trends of reported blue-sky days and other summary statistics in figure 1 pointed out above. For instance, if we look at the proportion of days with excellent and unhealthy air quality reported in figure 1, we find that their values in 2006–7 are similar to those reported in 2001–2, which suggest that all four years had relatively similar air quality. However, the proportion of blue-sky days is about 9%–11% higher in 2006–7 than its 2001–2 levels. A potential explanation for this discrepancy is provided by our estimated proportions of manipulation among blue-sky days, which are three times as high in 2006–7 as in 2001–2. Last but not least, our estimates further suggest that a large proportion of the reported increase in blue-sky days after 2008 is due to data manipulation,

6. Figure A10*A* compares these estimates with those obtained using alternative choices of manipulation windows and shows that all the estimates obtained from the different manipulation windows lie within the 90% confidence band of the benchmark estimates.

7. The greater extent of manipulation we estimate in those two years could also be due to intertemporal shifting of production from 2008 to those years as well as infrastructural investments in preparation for the Games.

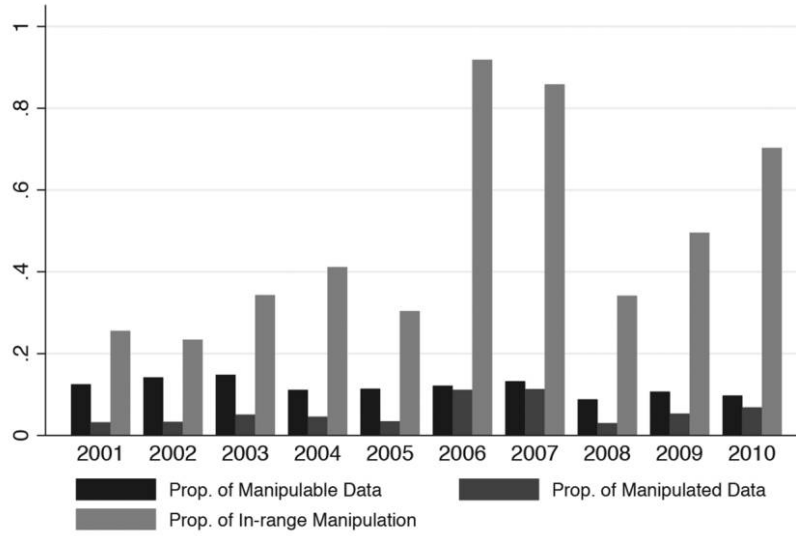


Figure 5. Manipulable data, manipulated data, and in-range manipulation. The proportions are calculated using the censored MLE estimate of the counterfactual distribution with manipulation window [0.135, 0.18]; "prop. of manipulable data" is the proportion of data within the manipulation window, but above the blue-sky day cutoff according to the counterfactual distribution, that is, $G(0.18; \hat{\theta}_{it}) - G(0.15; \hat{\theta}_{it})$; "prop. of manipulated data" is the overall proportion of manipulation, that is $\hat{\lambda}_{it}$; "prop. of in-range manipulation" is the proportion of manipulation among the manipulated data, that is, $\hat{\nu}_{it} = \hat{\lambda}_{it}/(G(0.18; \hat{\theta}_{it}) - G(0.15; \hat{\theta}_{it}))$, where i denotes Beijing.

which is intuitive because many of the policies implemented in 2008 to curb air pollution were temporary.

2.2. Overview of Manipulation Behavior in Chinese Cities

In this section, we summarize the estimated manipulation measures for all cities and years in our data set. Panel A of table 2 presents the summary statistics for the proportion of manipulation among blue-sky days ($\hat{\mu}_{it}$) and the number of manipulated blue-sky days ($\hat{m}_{it} = \hat{\lambda}_{it} \cdot T_{it}$). The average proportion of manipulation among blue-sky days is estimated at 3.1% with a standard deviation of 4.4%. The average number of manipulated blue-sky days is about 8.3 per year with a 11.1 standard deviation. Despite these modest averages, the distribution of both manipulation measures has a long right tail. The highest proportion of manipulation among blue-sky days we estimate is 30.5% (Shijiazhuang in 2001) and the largest number of manipulated blue sky days we find is 83.5 days in a year (Zaozhuang in 2009).

To quantify cumulative exposure to manipulated blue-sky days, panel B of table 2 provides estimates of the total number of manipulated blue-sky days over our sample

Table 2. Summary Statistics of Manipulation Measures of PM_{10} Concentrations: All Cities

	A. Summary Statistics for All Cities and Years		
	Mean	SD	Maximum
Proportion of manipulation among blue-sky days $\hat{\mu}_{it}$	3.1%	4.4%	30.5%
Estimated number of manipulated blue-sky days ($\hat{m}_{it}$)	8.3	11.1	83.5
No. of city × year observations			1,012

	B. Summary Statistics of the Estimated Total of Manipulated Blue-Sky Days							
	Cities Reporting since 2001 Total: 2001–10				All Reporting Cities Total: 2006–10			
	No. Cities	Mean	SD	Max	No. Cities	Mean	SD	Max
Estimated total of manipulated blue-sky days	46	93.9	88.5	395.9	110	39.9	40.2	153.0

Note. Panel A reports the summary statistics for all cities and years in our data set for the following variables: (1) $\hat{\mu}_{it} = \hat{\lambda}_{it}/\hat{F}_{X_{it}}(0.15)$, (2) $\hat{m}_{it} = \hat{\lambda}_{it} \cdot T_{it}$, where $\hat{\lambda}_{it} = \hat{F}_{X_{it}}(0.15) - G(0.15; \hat{\theta}_{it})$ as defined in sec. 1.3. Manipulation measures are estimated for city-year combinations with at least 100 daily readings. As pointed out above, most of the missing days are a result of a substantial proportion of cities starting to report air quality data after 2001, some of which enter our data set in the middle of their first year. Panel B reports (1) summary statistics for the estimated total number of manipulated blue-sky days between 2001 and 2010 for cities that have continuously reported PM_{10} data over the entire sample period with a minimum of 100 daily observations per year, and (2) summary statistics for the total of manipulated blue-sky days between 2006 and 2010 for all cities in our data set reporting a minimum of 100 daily observations per year.

period. We first provide summary statistics of the total number of manipulated blue-sky days between 2001 and 2010 for cities that report PM_{10} concentrations during the entire period. For these cities, we estimate an average of 93.9 manipulated blue-sky days over the 10 years with a maximum of 395.9 (Shenyang). We also estimate the total number of manipulated blue-sky days between 2006 and 2010 for all cities in our data set. For these last 5 years of the decade we examine, we find the total number of manipulated blue-sky days to be 39.9 on average with a maximum of 153 (Shenyang).

Next, we examine the geographical distribution of the number of manipulated blue-sky days. Figure 6 displays maps with the number of manipulated blue-sky days for selected years in our sample period, specifically 2006 and 2010. The maps for all other years in our data set are included in the appendix. Figure 6 illustrates substantial geographical heterogeneity in manipulation behavior. Furthermore, the maps (including those reported in the appendix for other years) also point to the variability in the geographical distribution from year to year, which suggests that there is potential heterogeneity within cities over time.

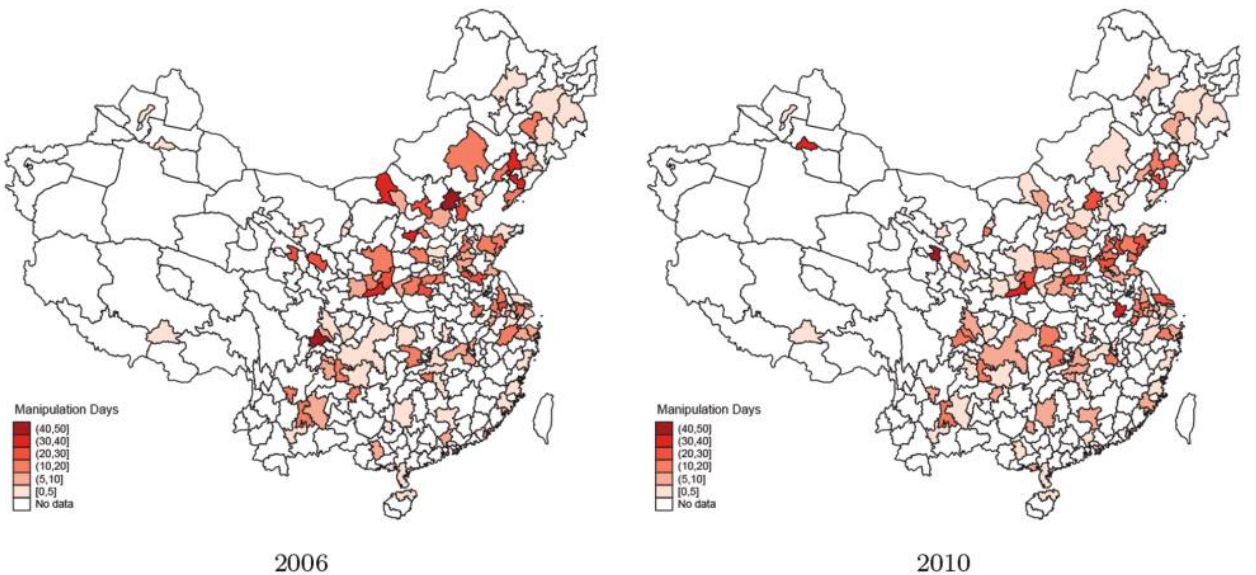


Figure 6. Geographical distribution of the estimated number of manipulated days. For each year, the map reports the number of manipulated blue-sky days for each city.

To further illustrate the within-city heterogeneity discussed above, figure 7 presents the counterparts of figure 4 for the remaining three province-level cities in China, specifically Shanghai, Tianjin, and Chongqing. Unlike Beijing, the time series of reported and estimated proportion of blue-sky days are very similar for Shanghai, and there is no statistical evidence of data manipulation around the blue-sky day cutoff. For both Tianjin (panel B) and Chongqing (panel C), we find statistically significant proportions of manipulation for the first part of the time period we examine; however, the trajectory of the proportion of manipulation among blue-sky days is quite different for each city. Our estimated proportions are more or less stable between 5% and 10% for Tianjin before 2005, whereas we find statistically significant estimates of 5% for Chongqing in 2001 followed by a dramatic increase and then decrease of the estimated proportions of manipulation. The highest proportion of manipulation for Chongqing is reached in 2003.

Overall, our results illustrate substantial heterogeneity across cities and years in the manipulation behavior. Our analysis suggests that the degree of effectiveness of China's policy to increase the proportion of blue-sky days varies considerably across cities and years.

3. PREDICTING MANIPULATION USING LOCAL OFFICIAL CHARACTERISTICS

Since the proportion of blue-sky days is incorporated in the performance evaluation of local officials, we next examine the key predictors of air quality data manipulation among resumé-type characteristics of city local officials using LASSO shrinkage. We first describe the institutional background and the data we collected on local officials. We then present our baseline LASSO specification, its results as well as multiple robustness checks.

3.1. Background and Data Summary

Each city in China has a city party secretary and a mayor. The city party secretary is the local representative of the Chinese Communist Party and hence the political leader of the city. The mayor, on the other hand, is the administrative manager of the city. Given that the party secretary is the political leader, he or she is arguably the more powerful local leader. The relative political power of the party secretary and mayor further depends on the administrative ranking of the city. For instance, in province-level cities, such as Beijing or Shanghai, the city party secretary is a member of the Politburo of the Communist Party of China. Hence, the party secretary in such a city has substantially greater political power and connections than the mayor. Further institutional details are provided in appendix A.

We assemble a unique city-level panel data set of 111 cities between 2001 and 2010, which includes party secretary and mayor characteristics as well as the annual measures of manipulation of blue-sky days that we constructed in the previous section.

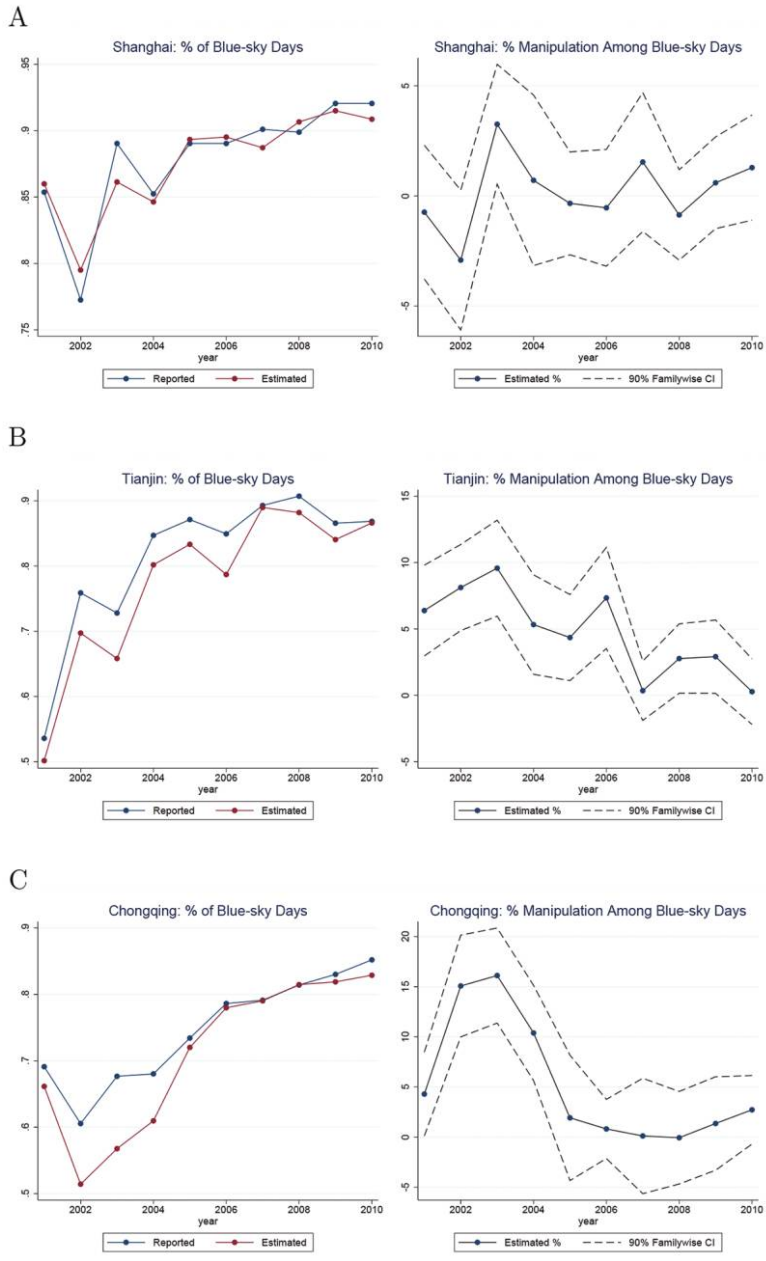


Figure 7. Reported versus estimated blue-sky days: Shanghai (*A*), Tianjin (*B*), and Chongqing (*C*). For each city, the left graph plots the time series of the proportion of blue-sky days in a year using the reported PM_{10} concentration data and the estimated true PM_{10} concentration distribution. The right graph plots the time series of the estimated proportion of manipulation among blue-sky days in a year. The 90% family-wise confidence intervals are calculated using a nonoverlapping block-bootstrap procedure with 7-day block length and 200 bootstrap replications.

The data set includes demographic, education, and work experience variables for all party secretaries and mayors who held office in the 111 cities between 2001 and 2010, subject to data availability, using their resumés available online.[8] There are broadly three categories of cities in our data set; specifically, province-level, subprovincial-level, and prefecture-level cities. To the best of our knowledge, this is the first data set to have such detailed information on local officials in China. Table 3 presents the summary statistics of the variables in our data set.

The demographic characteristics include gender (male or female) and ethnicity (Han or other). The overwhelming majority of party secretaries and mayors in our data set are male Han. There are only about 3% (2%) female Han and 9% (14%) male non-Han among party secretaries (mayors). There are no female non-Han party secretaries or mayors in our sample.

The education variables include a range of dummy variables for full-time and part-time degrees. For full-time educational degrees, we include dummy variables for college completion (Completed College), STEM majors (STEM Major), and attending an elite college (Elite College), which are highly selective universities in China.[9] Furthermore, we include a dummy variable indicating whether a local official entered college during the Cultural Revolution as a "Gong Nong Bing" college student (College Entrance during 1971–77), since the college admissions criteria were less academic and favored individuals with modest family backgrounds.[10] Similarly, we include a dummy variable that captures whether the local official was among the first two cohorts of college students

8. The data collection proceeded as follows. For each local official, whether party secretary or mayor, in power in a city and year in our data set, the variables in table A14 (tables A1–A14 are available online) are collected. Note that, in addition to the variables relevant to the local official, the city and year they were in power were also collected. Two data files are then produced, one for city party secretaries and another for mayors. Then, the two data sets are merged with the city-year panel of manipulation measures by city and year. Note that two cities, Zibo and Mudanjiang, exhibit reverse manipulation, so we exclude them from our LASSO analysis. Local officials are uniquely identified by their names in Chinese characters and their birth date. We have a total of 524 unique officials, for summary statistics on tenure and number of posts for officials in our data set see table A3.

9. We use the 1978 list of national key universities. The 88 listed universities include 16 comprehensive universities (e.g., Peking University), 51 science and technology institutes (e.g., Tsinghua University), 9 agricultural universities (e.g., Beijing Forestry University), 6 medical schools, 2 teachers' colleges, 2 foreign language schools, 1 law school, and 1 music conservatory.

10. During the Cultural Revolution, the college admission criteria put less emphasis on academic standards and favored students from peasant and working-class families (Chang 1974). Furthermore, much of the urban youth, who would otherwise enter college, were sent to rural areas to work. Hence, the first two college entrance exams after the Cultural Revolution were arguably the most competitive exams attracting many of those who were not allowed a university education during the Cultural Revolution.

Table 3. Local Official Characteristics: Summary Statistics

	Party Secretary				Mayor			
	Mean (SD) (1)	No. City × Year Obs. (2)	Mean (SD) (3)	No. City × Official Obs. (4)	Mean (SD) (1)	No. City × Year Obs. (2)	Mean (SD) (3)	City × Official Obs. (4)
Demographic characteristics:								
Male Han	.87	989	.89	316	.84	989	.84	332
Female Han	.03	989	.03	316	.02	989	.02	332
Male non-Han	.10	989	.09	316	.13	989	.14	332
Education:								
Full time:								
College	.62	989	.61	316	.53	989	.55	332
STEM major	.31	989	.32	316	.30	989	.31	332
Elite college	.28	989	.28	316	.25	989	.25	332
Entered college 1971–77	.18	989	.18	316	.15	989	.17	332
Entered college 1978	.26	989	.24	316	.25	989	.23	332
Part time:								
College	.39	989	.38	316	.45	989	.43	332
Graduate degree	.25	989	.25	316	.20	989	.19	332
Experience:								
Years in current post	2.33 (1.80)	989	2.12 (1.40)	316	2.18 (1.74)	989	1.97 (1.41)	332

Years to retirement	5.72	985	5.46	314	7.44	982	7.56	327
	(3.95)		(3.73)		(4.42)		(4.39)	
Current post in birth province	.58	989	.58	316	.61	989	.62	332
Previous experience:								
Enterprise	.39	973	.41	308	.47	944	.46	313
Research	.27	973	.25	308	.24	944	.25	313
Administrator in gov't or party organization	.28	973	.31	308	.36	945	.37	314
County mayor	.22	965	.22	307	.29	948	.28	315
County party secretary	.33	965	.33	307	.31	951	.32	316
City mayor	.57	966	.59	308	.14	949	.17	315
City party secretary	.28	965	.27	307	.10	946	.09	315
Central government	.16	989	.16	316	.13	989	.14	332
Location of previous posts:								
Current city	.86	989	.81	316	.91	989	.89	332
Current province	.98	989	.98	316	1.00	989	1.00	332
Other province	.25	989	.26	316	.22	989	.23	332

Note. This table reports two sample means and their respective number of observations for the local official characteristics we collect. Since party secretaries and mayors can hold office for multiple years in the same city, we report the averages across city-year observations as well as city-official observations. Columns 1 and 2 report the sample mean averaged across city-year observations and the number of observations, respectively. This corresponds to the panel that we use in our LASSO analysis. Columns 3 and 4 report the sample mean averaged across city-official observations and the number of observations, respectively. We also report the standard deviations in parentheses for continuous variables.

selected immediately after the Cultural Revolution (College Entrance in 1978),[11] indicating that the official had a strong academic background before entering college and received a higher quality college education.

The experience variables in our data set fall under three categories: current post, previous experience, and previous locations. For the current post, we include tenure in the current post (Years in Current Post) and years to retirement (Years to Retirement) as well as a dummy for whether the current post is in the official's birth province (Current Post in Birth Province). Furthermore, we have a host of dummy variables for previous government positions held, Administrator in Government or Party Organization, County Mayor, County Party Secretary, City Mayor, City Party Secretary, and Central Government.[12] Finally, we have indicator variables for whether the current official had any previous post in the current city (Current City), current province (Current Province), or another province (Other Province). Almost every local official in our sample had a previous post in the same province as their current position. Hence, we exclude this variable from our LASSO analysis.

The summary statistics presented in table 3 illustrate that on average party secretaries and mayors have similar education characteristics. In terms of experience, party secretaries on average have about 2 years less to retirement and are more likely to have served as city party secretaries or mayors prior to their current post. Otherwise, there seems to be little difference between party secretaries and mayors on average in terms of the other experience variables. In our discussion of the LASSO results, we use promotion data for party secretaries, which were collected from the local official resumés as the local official characteristics. The relevant summary statistics are presented in section 3.3.

3.2. LASSO Results

In this section, we examine the key predictors of air quality data manipulation among party secretary and mayor characteristics. Since we have a fairly large number of local official characteristics in our data set, specifically 46 (23 for each party secretary and mayor), we apply the LASSO shooting algorithm proposed by Belloni et al. (2014b) to select among these characteristics. The main advantage of this procedure is that it delivers valid post-selection inference that is robust to model selection errors. The key assumption of the LASSO method is the approximate sparsity assumption, which means that the relationship between the outcome variable and regressors can be well approximated by a linear function of a small number of regressors. Our results illustrate that the sparsity assumption is likely appropriate for our empirical setting.

Our baseline LASSO specification is presented below. Let $z_{it}^{m,j}$ denote the jth mayor characteristic, $z_{it}^{s,j}$ the jth party secretary characteristic in city i at year t, and M_{it} a

11. The first cohort of college students after the Cultural Revolution was called the seventy-seventh cohort but they in fact entered college in spring 1978. The second cohort entered college in fall 1978.

12. Note that counties have a lower administrative ranking than cities in the Chinese system.

manipulation measure. The LASSO model selection step is performed on the following regression equation

$$M_{it} = \sum_{j=1}^{K/2} \beta_j z_{it}^{m,j} + \sum_{j=K/2+1}^{K} \beta_j z_{it}^{s,j} + \gamma_p + \delta_r + \eta_t + u_{it}, \quad (1)$$

where γ_p, δ_r and η_t are province, city-rank, and year fixed effects, respectively.[13] These fixed effects are treated as control variables in the LASSO procedure, that is, the outcome and regressors are residualized using these fixed effects before selection is implemented. Hereafter, when we refer to a set of fixed effects as controls in a LASSO procedure, we mean that they are used to residualize the outcome and regressors prior to the selection step.

The following provides the penalized objective function of our procedure (Belloni et al. 2014a):

$$\hat{\beta} = \text{argmin}_{\beta} \sum_{i=1}^{n} \sum_{t=1}^{T} (M_{it} - \sum_{j=1}^{K/2} z_{it}^{p,j} \beta_j - \sum_{j=K/2+1}^{K} z_{it}^{m,j} \beta_j - \gamma_p - \delta_r - \eta_t)^2 + \lambda \sum_{j=1}^{2K} \left| \hat{l}_j \beta_j \right|, \quad (2)$$

where $\hat{l}_j$ is a regressor-specific data-driven penalty loading and λ is a penalty chosen according to Belloni et al. (2014b). The L^1 penalty in the LASSO shrinks some coefficients to zero and thereby performs variable selection. The post-LASSO regression is then performed on the selected variables. In our application of this procedure, we include all variables selected in the LASSO selection with either $\hat{\mu}_{it}$ or $\hat{m}_{it}$ as the dependent variable in our post-LASSO regression.[14] In addition to the post-LASSO regression, we also report the results of two-way fixed effects regressions, that is, regressions with city and year fixed effects.[15]

13. City-rank fixed effects include dummies for (1) province-level cities, (2) subprovincial cities, (3) vice subprovincial cities, and (4) prefecture-level cities defined based on the administrative ranking of the cities. The subprovincial category consists of 10 capital cities of important provinces as well as five other important noncapital cities, whereas the remaining capital cities are included in the vice subprovincial category.

14. The implementation here mimics the double LASSO procedure in Belloni et al. (2014b), where selection is performed twice on the outcome variable and the treatment to avoid any omitted variable bias.

15. When we include city and year fixed effects as controls in the LASSO procedure in lieu of the province, city-rank, and year fixed effects, none of the local official characteristics are selected. To rule out the possibility that this result is driven by low variability in the local official characteristics after controlling for city and year fixed effects, we compare the standard deviation of the fixed effects residualized variables with the original variables in tables A1, A2 and find substantial variation left in the variables even after removing city and year fixed effects. In addition, app. B presents LASSO and elastic net results with city and year fixed effects as controls using different information criteria, including AIC and BIC, to select the penalty level instead of the penalty level proposed in Belloni et al. (2014b).

Table 4. Predicting Manipulation Using Local Official Characteristics: Baseline LASSO Results

A. LASSO Selection Step

	Dependent Variable	
	Proportion of Manipulation	Manipulated Blue-Sky Days
Variables included for selection:		
Party secretary characteristics:		
Demographic	Yes	Yes
Education	Yes	Yes
Experience	Yes	Yes
Mayor characteristics:		
Demographic	Yes	Yes
Education	Yes	Yes
Experience	Yes	Yes
Variables selected by LASSO:		
Party secretary characteristics	Elite college	Elite college
Mayor characteristics	None	None
Variables included as controls:		
Province FE	Yes	Yes
City-rank FE	Yes	Yes
Year FE	Yes	Yes

B. Post-LASSO and Fixed Effects Regression

	Dependent Variable			
	Proportion of Manipulation		Manipulated Blue-Sky Days	
	(1)	(2)	(3)	(4)
Party secretary with elite college degree	.012	.010	2.84	2.30
	(2.76)	(2.27)	(2.59)	(2.11)
Post-LASSO with province, city-rank, and year FE	Yes	No	Yes	No
City and year FE	No	Yes	No	Yes
No. of city × year observations	989	989	989	989

C. Summary Statistics of Dependent Variables

	Mean	SD	Min	Max	No. City × Year Obs.
Proportion of manipulation ($\hat{\mu}_{it}$)	.03	.04	−.06	.31	989
Manipulated blue-sky days ($\hat{\mu}_{it}$)	8.58	10.85	−14.55	83.45	989

Table 4 (*Continued*)

C. Summary Statistics of Dependent Variables

	Mean	SD	Min	Max	No. City × Year Obs.
No. of daily observations by city-year (T_{it})	354.83	35.28	188	366	989

Note. Panel A: The LASSO selection step uses the automatic penalty level in Belloni et al. (2014b). Panel B: The *t*-statistics in parentheses are computed using city-level cluster-robust standard errors. Panel C: The summary statistics in parentheses are computed using the merged data set that includes manipulation measures as well as local official characteristics. In addition to the manipulation measures, we also report the summary statistics for the number of daily observations used to construct these measures for each city-year observation. FE = fixed effects.

The main finding from our baseline LASSO results is that the only variable selected as a predictor of manipulation is having a party secretary in power who obtained an undergraduate degree from an elite college, hereafter PSEC (Party Secretary Elite College). Table 4 reports the LASSO selection results in panel A and the post-LASSO and two-way fixed effects results in panel B. The LASSO selection procedure is performed using the data-dependent penalty level recommended in Belloni et al. (2014b). We find that having a party secretary with an elite college degree (PSEC = 1) is associated with a 1 percentage point increase in the annual proportion of manipulation (2.30 manipulated blue-sky days in a year), which is statistically significant at the 5% level. To put this finding into context, the average proportion of manipulation in our estimation sample is 3% with a standard deviation of 4% as presented in panel C of table 4. Hence, the average increase in the proportion of manipulation correlated with PSEC is 30% of its sample mean, whereas the average increase in manipulated blue-sky days is about 25% of its sample mean.

Next, we consider several alternative LASSO specifications. The first natural robustness check of our finding is to perform the LASSO procedure while excluding PSEC from the variables that are available for selection. In this case, no variables are selected, as presented in table 5. This finding supports the importance of the PSEC variable as the key predictor and that it is not masking the predictive power of other variables. The two other robustness checks we present below reduce the number of control variables we include in the LASSO selection step as well as the penalty level. Both checks allow the LASSO to select more variables, so that we can examine if they affect the magnitude and significance of the coefficient on the PSEC variable. Before we proceed to summarize the aforementioned robustnesss checks, we refer the reader to appendix B for additional robustness checks of the PSEC coefficient estimate, which consider using information criteria, such as the Akaike (AIC) and Bayesian (BIC) information criterion, to select the penalty level for both LASSO and elastic net.

Table 5. Alternative LASSO Specification I (PSEC Excluded from LASSO)

	Dependent Variable	
	Proportion of Manipulation	Manipulated Blue-Sky Days
Variables included for selection:		
Party secretary characteristics:		
Demographic	Yes	Yes
Education	All but PSEC	All but PSEC
Experience	Yes	Yes
Mayor characteristics:		
Demographic	Yes	Yes
Education	Yes	Yes
Experience	Yes	Yes
Variables selected by LASSO:		
Party secretary characteristics	None	None
Mayor characteristics	None	None
Variables included as controls:		
Province FE	Yes	Yes
City-rank FE	Yes	Yes
Year FE	Yes	Yes

Note. The LASSO selection step uses the automatic penalty level in Belloni et al. (2014b). PSEC = party secretary elite college; FE = fixed effects.

Table 6 reports the post-LASSO and fixed effects results for variables selected by LASSO using a smaller set of fixed effects than in the baseline specification. Columns 1–2 report the results with province and year fixed effects as controls only in the LASSO selection step, whereas columns 3–4 report the results with province fixed effects only. In all versions of the LASSO we consider here, the PSEC variable is selected and its magnitude is similar to the baseline specification. While other variables are selected, none of them are statistically significant, even at the 10% level, in the city and year fixed effects regressions.

All of the above LASSO results rely on the data-driven choice of the penalty level, λ, proposed in Belloni et al. (2014b), which is 279.53 for our baseline LASSO specification. Tables 7 and 8 present the post-LASSO and fixed effects results for variables selected by LASSO using lower values for this tuning parameter, specifically, 250, 200, 150, and 100. For $\lambda = 250$, only PSEC is selected. As a result, columns 1–2 of both tables are identical to our baseline post-LASSO and fixed effects results. For the remaining choices of λ, while additional variables are selected, the coefficient on PSEC is robust to the inclusion of these additional regressors in both post-LASSO and fixed effects regressions. Its statistical significance is maintained at even lower significance levels than in the baseline specification.

Table 6. Alternative LASSO Specification II: Relaxing the Control Variables

	A. Dependent Variable: Proportion of Manipulation			
	(1)	(2)	(3)	(4)
Party secretary with elite college degree (PSEC)	.016	.011	.013	.011
	(3.05)	(2.26)	(2.41)	(2.26)
Other party secretary characteristics:				
Previous experience as city party secretary	.006	−.001	.004	−.001
	(1.20)	(−.13)	(.96)	(−.13)
Mayor characteristics:				
Previous experience as city party secretary	.008	−.007	.005	−.007
	(1.17)	(−.82)	(.68)	(−.82)
Previous posts in other province	.005	.001	.009	.001
	(1.16)	(.25)	(1.83)	(.25)
Post-LASSO	Yes	No	Yes	No
City and year FE	No	Yes	No	Yes
No. city-year observations	925	925	925	925
	B. Dependent Variable: Manipulated Blue-Sky Days			
	(1)	(2)	(3)	(4)
Party secretary with elite college degree (PSEC)	3.59	2.41	3.20	2.41
	(2.88)	(2.01)	(2.42)	(2.01)
Other party secretary characteristics:				
Previous experience as city party secretary	1.49	−.28	1.29	−.28
	(1.27)	(−.24)	(1.17)	(−.24)
Mayor characteristics:				
Previous experience as city party secretary	1.93	−2.09	2.24	−2.09
	(1.38)	(−1.14)	(1.42)	(−1.14)
Previous posts in other province	1.03	−.25	1.61	−.25
	(.89)	(−.21)	(1.35)	(−.21)
Post-LASSO	Yes	No	Yes	No
City and year FE	No	Yes	No	Yes
No. of city × year observations	925	925	925	925

Note. The *t*-statistics in parentheses are computed using city-level cluster-robust standard errors. Columns 1 and 2 include regressors selected using LASSO with only province and time fixed effects as controls; cols. 3 and 4 include regressors selected using LASSO with only province fixed effects as controls. The post-LASSO regression in cols. 1 and 3 include province and year fixed effects (province fixed effects only). FE = fixed effects.

Table 7. Alternative LASSO Specification III: Relaxing the LASSO Penalty Level

	Dependent Variable: Proportion of Manipulation							
Variables Selected Using Penalty Level	250 (1)	250 (2)	200 (3)	200 (4)	150 (5)	150 (6)	100 (7)	100 (8)
PSEC	.012	.010	.012	.010	.014	.014	.013	.012
	(2.76)	(2.27)	(2.77)	(2.28)	(3.29)	(3.04)	(3.10)	(2.70)
Other party secretary characteristics:								
Male Han			.009	.006	.009	.007	.009	.006
			(2.02)	(1.32)	(1.92)	(1.36)	(1.79)	(1.22)
College entrance 1978					−.008	−.009	−.009	−.010
					(−2.30)	(−2.31)	(−2.45)	(−2.37)
Current post in birth province							−.007	−.007
							(−1.77)	(−1.73)
Previous experience:								
City party secretary							.001	−.002
							(.19)	(−.49)
City mayor							−.004	−.003
							(−1.24)	(−.80)
Previous posts in current city					.005	.005	.007	.006
					(1.52)	(1.48)	(2.22)	(1.96)

Mayor characteristics:								
Female Han							.017	.017
							(1.96)	(1.73)
STEM major					.009	.009	.008	.008
					(2.03)	(1.97)	(1.51)	(1.64)
Elite college degree							.002	.001
							(.37)	(.23)
Previous experience:								
City party secretary							−.000	−.008
							(−.04)	(−.94)
County party secretary					.005	.004	.006	.005
					(1.19)	(.91)	(1.50)	(1.14)
County mayor					.005	.005	.006	.006
					(1.11)	(.99)	(1.24)	(1.17)
Previous posts in other province					.006	.005	.005	.004
					(1.25)	(.86)	(1.16)	(.78)
Post-LASSO	Yes	No	Yes	No	Yes	No	Yes	No
City and year FE	No	Yes	No	Yes	No	Yes	No	Yes
No. of city × year observations	989	989	989	989	948	948	924	924

Note. The *t*-statistics in parentheses are computed using city-level cluster-robust standard errors. The LASSO selection step and post-LASSO regressions control for province, city-rank, and year fixed effects. PSEC = party secretary elite college; FE = fixed effects.

Table 8. Alternative LASSO Specification III: Relaxing the LASSO Penalty Level

	Dependent Variable: Manipulated Blue-Sky Days							
	Variables Selected Using Penalty Level							
	250	250	200	200	150	150	100	100
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
PSEC	2.84	2.30	2.81	2.29	3.14	2.87	2.98	2.51
	(2.59)	(2.11)	(2.60)	(2.11)	(2.91)	(2.60)	(2.72)	(2.27)
Other party secretary characteristics:								
Male Han			1.99	1.42	1.98	1.46	1.74	1.18
			(1.96)	(1.35)	(1.74)	(1.23)	(1.49)	(.99)
College entrance 1978					−1.75	−1.86	−1.87	−1.93
					(−1.81)	(−1.81)	(−2.00)	(−1.91)
Current post in birth province							−1.11	−1.12
							(−1.10)	(−1.00)
Previous experience:								
City party secretary							.24	−.50
							(.23)	(−.43)
City mayor							−1.26	−1.00
							(−1.36)	(−.99)
Previous posts in current city					1.84	1.72	2.36	2.02
					(2.08)	(1.94)	(2.69)	(2.31)
Female Han							3.36	3.16
							(2.08)	(1.54)
STEM major					2.33	2.25	1.92	1.93
					(2.26)	(2.06)	(1.64)	(1.69)
Elite college degree							.99	.70
							(.80)	(.58)
Previous experience								
City party secretary							−.55	−2.27
							(−.33)	(−1.15)
County secretary					1.38	1.19	1.76	1.55
					(1.20)	(.94)	(1.59)	(1.22)
County mayor					1.11	.95	1.23	1.09
					(.95)	(.76)	(1.04)	(.87)
Previous posts in other province					1.00	.52	.83	.35
					(.87)	(.40)	(.74)	(.26)

Table 8 (*Continued*)

	Dependent Variable: Manipulated Blue-Sky Days							
	Variables Selected Using Penalty Level							
	250 (1)	250 (2)	200 (3)	200 (4)	150 (5)	150 (6)	100 (7)	100 (8)
Post-LASSO	Yes	No	Yes	No	Yes	No	Yes	No
City and year FE	No	Yes	No	Yes	No	Yes	No	Yes
No. of city × year observations	989	989	989	989	948	948	924	924

Note. The *t*-statistics in parentheses are computed using city-level cluster-robust standard errors. The LASSO selection step and post-LASSO regressions control for province, city-rank, and year fixed effects (FE). PSEC = party secretary elite college; FE = fixed effects.

Finally, it is worth noting that among the additional variables selected by LASSO using lower penalty levels (tables 7, 8), there are only two that are statistically significant at the 5% level in the city and year fixed effects regression in column 8 of those tables. Notably, both variables are party secretary characteristics. The first is Previous Posts in Current City, which is positively correlated with manipulation, whereas the other variable is College Entrance in 1978, which is negatively correlated with manipulation. The former variable indicates that these party secretaries likely had the opportunity to build professional experience and connections in the city prior to their post as party secretary and hence may be viewed as a proxy for prior local professional experience as well as local connections. While the positive correlation of this variable with manipulation is more intuitively attributable to local connections, we cannot separate these different mechanisms in our data. The latter variable, College Entrance in 1978, equals one if the party secretary in power entered college in 1978, that is, immediately after the Cultural Revolution. Since academic criteria were not the primary determinants of college entrance during the Cultural Revolution, many individuals who were not able to enter college during those seven years took the college entrance exam in 1977 (to enter college in 1978). As a result, these individuals faced a 5% admissions rate, the lowest in China's history, which implies that College Entrance in 1978 is an indicator of strong academic ability. The negative correlation between manipulation and College Entrance in 1978 provides suggestive evidence that air pollution manipulation may be negatively correlated with academic ability.

3.3. Discussion

In addition to supporting the robustness of the PSEC-manipulation relationship, the previous results shed some light on the correlation between PSEC and manipulation. The PSEC variable captures multiple unobservable characteristics, including academic

ability and high-profile connections. Which of these unobservables is driving the relationship between manipulation and PSEC is an open question. The negative correlation between manipulation and College Entrance in 1978 provides suggestive evidence that manipulation and strong academic ability are negatively correlated. Hence, the interpretation of the PSEC-manipulation relationship that is most consistent with the patterns in our data is that connections is likely the unobservable that can explain the positive correlation between manipulation and PSEC.

We next consider the relationship between promotion and manipulation. Since the Chinese central government uses career advancement incentives to encourage local officials to meet its targets (Li and Zhou 2005; Xu 2011), promotion concerns are a likely factor behind this predictive relationship. Furthermore, previous work has found education level to be an important determinant of promotion (Shih et al. 2012). Since this previous literature suggests that promotion concerns may be a possible explanation for our findings, we examine the correlation between promotion and manipulation for elite-educated and other party secretaries. Table 9 presents the correlation between the proportion of manipulation and the promotion to higher administrative positions for elite-educated and other party secretaries. Since the probability of promotion is heterogeneous due to city-rank-specific or city-specific unobservables, we present the correlations controlling for heterogeneity in the mean at the city-rank level (I) and the city level (II). We find that manipulation and promotion, irrespective of position, are weakly positively correlated, specifically 0.02, for elite-educated party secretaries, whereas they are negatively correlated, specifically −0.05, for other party secretaries. When we look at the correlation between manipulation and promotion to specific positions, we find that this correlation differential is largest for those promoted to province-level National People's Congress/Chinese People's Political Consultative Conference (NPC/CPPCC) and the chief manager (party secretary) of a provincial department, specifically 0.11 and 0.16, respectively, for the elite-educated secretaries, whereas this correlation is −0.10 and zero, respectively, for the non-elite educated secretaries (col. 2 in table 9, panel B). The correlation differential is very small for those promoted to other province-level party or government positions as well as to positions at the central government. We also include the correlations controlling for city- and year-specific heterogeneity (III), which deliver qualitatively similar results. Hence, these patterns suggest that promotion concerns may help explain the correlation between manipulation and PSEC.

Since economic growth is an important factor in promotions of China's local officials, which has provided an explanation for increased pollution by connected leaders in Jia (2017), we compare the correlation between GDP and manipulation for party secretaries with elite-college degrees (PSEC = 1) with others (PSEC = 0). Table 10 presents the mean total GDP as well as by sector of cities conditional on PSEC status in panel A. Elite-educated party secretaries tend to serve in cities with higher GDP on average and the difference relative to other party secretaries is statistically significant at the 5% level, except for primary-sector GDP. Panel B presents the correlation between

Table 9. Correlation between Manipulation and Promotion of Party Secretaries

A. Mean of Promotion Variables Conditional on PSEC

	PSEC = 1	Difference (PSEC = 1) – (PSEC = 0)	*t*-Statistic
Promoted after current post	.32	–.09	–1.22
Promoted to:			
Province-level party/city government	.41	–.01	–.14
Province-level NPC/CPPCC	.21	–.02	–.37
Chief manager (party secretary) of provincial dept	.08	.00	–.02
Central government	.09	.05	1.09

B. Correlation of Manipulation and Promotion Conditional on PSEC

	I		II		III	
	PSEC = 1	PSEC = 0	PSEC = 1	PSEC = 0	PSEC = 1	PSEC = 0
Promoted after current post	.02	–.05	.02	–.05	–.02	–.06
Promoted to:						
Province-level party/city government	–.01	.04	–.02	.04	.00	.01
Province-level NPC/CPPCC	.09	–.09	.11	–.10	.09	–.07
Chief manager (party secretary) of provincial dept	.15	–.01	.16	.00	.14	.00
Central government	.09	.07	.09	.07	.05	.04
No. of city × year observations	266	670	266	670	266	670

Note. The *t*-statistics are obtained using city-level cluster-robust standard errors. To compute the above correlations I, we first remove city-rank-specific unobservables from both proportion of manipulation and the promotion variable in question. Specifically, for two variables $Z_{it}^1 = \mu_r^1 + \epsilon_{it}^1$ and $Z_{it}^2 = \mu_r^2 + \epsilon_{it}^2$, the above correlation estimates the correlation between ϵ_{it}^1 and ϵ_{it}^2. The correlations II are computed after removing city-specific unobservables from both the proportion of manipulation and the promotion variable in question. Specifically, for two variables $Z_{it}^1 = \alpha_i^1 + \epsilon_{it}^1$ and $Z_{it}^2 = \alpha_i^2 + \epsilon_{it}^2$, the above correlation estimates the correlation between ϵ_{it}^1 and ϵ_{it}^2. The correlations III are correlations of residuals after removing city-specific and year-specific unobservables. PSEC = party secretary elite college.

manipulation and GDP after accounting for city and year fixed effects conditional on PSEC. We find that GDP and manipulation are weakly positively correlated for elite-educated party secretaries (PSEC = 1) with a correlation coefficient of 0.04, whereas they are weakly negatively correlated for other party secretaries (PSEC = 0) with a –0.02 correlation coefficient. We find similar correlation patterns for the secondary and tertiary sector, whereas the correlation is negative for the primary sector

Table 10. PSEC, Economic Growth, and Manipulation

A. Mean of GDP Conditional on PSEC

	PSEC = 1	Difference (PSEC = 1) – (PSEC = 0)	*t*-Statistic
GDP	2095.9	801.8	2.34
GDP by sector:			
Primary	100.9	–2.1	–.19
Secondary	977.0	319.7	2.30
Tertiary	1,000.5	474.2	2.25

B. Within-Correlation of GDP and Manipulation Conditional on PSEC

	PSEC = 1	PSEC = 0
GDP	.04	–.02
GDP by sector:		
Primary	–.13	.00
Secondary	.02	–.02
Tertiary	.06	–.02
No. of city × year observations	274	708

Note. The *t*-statistics reported are obtained using city-level cluster-robust standard errors. To compute the above correlations, we first remove city- and year-specific unobservables from both the proportion of manipulation and the economic variable in question. Specifically, for two variables $Z_{it}^1 = \alpha_i^1 + \lambda_t^1 + \epsilon_{it}^1$ and $Z_{it}^2 = \alpha_i^2 + \lambda_t^2 + \epsilon_{it}^2$, the above correlations estimate the correlations between ϵ_{it}^1 and ϵ_{it}^2. PSEC = party secretary elite college.

conditional on PSEC = 1 and zero conditional on PSEC = 0. The results are intuitive since the primary sector is not a major contributor to PM_{10} pollution.

Overall, these patterns in the data provide suggestive evidence that the positive correlation between manipulation and having an elite-educated party secretary may be a result of career considerations. This is further supported by previous literature documenting that economic growth is the primary predictor of promotion for party secretaries (Zheng et al. 2013). However, given the complexity of the local leadership structure and the promotion criteria, further work is required to determine the mechanisms behind this relationship.

4. CONCLUSION

We present a simple and tractable strategy to estimate measures of manipulation in China's air pollution data around the blue-sky day cutoff. While the proposed method is motivated by the Chinese environmental regulatory system, it is general and can be applied in a variety of empirical settings in economics and other social sciences. We illustrate that it provides a convenient alternative to existing methods in terms of

practical implementation and the validity of standard inference. Given the connection between data manipulation and excess bunching (Saez 2010; Chetty et al. 2011; Bertanha et al. 2018), the procedure proposed in this paper may be extended to provide alternative methods to estimate the proportion of excess bunching. This is an important direction left for future work.

Using our proposed method, we quantify manipulation around the blue-sky day cutoff for air pollutant concentrations for reporting Chinese cities between 2001 and 2010. Our results suggest that the effectiveness of China's policy to induce its local leaders to comply with environmental targets varies substantially in the first decade of the twenty-first century. While the average manipulation measures we estimate are modest, the distribution of these measures has a long right tail. We also document substantial geographical and temporal heterogeneity.

Since days exceeding the blue-sky day cutoff would be labeled "unhealthy for sensitive groups," manipulation of such days to meet the blue-sky day cutoff implies that sensitive groups, which include young children and the elderly, will not be alerted on these days to take appropriate defensive measures to minimize the adverse effects of these unhealthy pollution levels (Neidell 2009; Zhang and Mu 2018). Our estimates of the total number of manipulated blue-sky days over the sample period suggest that the cumulative exposure of sensitive groups to manipulated blue-sky days over our sample period was sizable for many cities in our data set, with a maximum exceeding a full year of manipulated blue-sky days in a 10-year period.

Finally, our results point to the importance of political economy considerations in understanding the effectiveness of the environmental regulatory system in China. While our empirical analysis provides some suggestive evidence that promotion concerns may have adverse impacts on environmental compliance, further analysis that carefully considers the dual-head structure of local leadership and the career tournament system in China is required to provide conclusive evidence on this complex policy environment. This constitutes a priority for future work.

REFERENCES

Andrews, Steven Q. 2008a. Playing air quality games. *Far Eastern Economic Review*, 53–57.

———. 2008b. Inconsistencies in air quality metrics: "Blue sky" days and PM 10 concentrations in Beijing. *Environmental Research Letters* 3 (3): 034009.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014a. High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives* 28 (2): 29–50.

———. 2014b. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies* 81 (2): 608–50.

Bertanha, Marinho, Andrew H. McCallum, and Nathan Seegert. 2018. Better bunching, nicer notching. Unpublished manuscript.

Blonz, Josh. 2019. The welfare costs of misaligned incentives: Energy inefficiency and the principal-agent problem. Unpublished manuscript.

Burgstahler, David C., and Ilia D. Dichev. 1997. Earnings, adaptation and equity value. *Accounting Review* 72 (2): 187–215.

Chang, Parris H. 1974. The Cultural Revolution and Chinese higher education: Change and controversy. *Journal of General Education* 26 (3): 187–94.

Chen, Yuyu, Ginger Zhe Jin, Naresh Kumar, and Guang Shi. 2012. Gaming in air pollution data? Lessons from China. *B. E. Journal of Economic Analysis and Policy (Advances)* 13 (3): article 2.

Chetty, Raj, John N. Friedman, Tore Olsen, and Luigi Pistaferri. 2011. Adjustment costs, firm responses, and micro vs. macro labor supply elasticities: Evidence from Danish tax records. *Quarterly Journal of Economics* 126 (2): 749–804.

Cole, Matthew A., David J. Maddison, and Liyun Zhang. 2019. Testing the emission reductions claims of CDM projects using Benford's law. Unpublished manuscript.

Dee, Thomas S., Will Dobbie, Brian A. Jacob, and Jonah Rockoff. 2019. The causes and consequences of test score manipulation: Evidence from the New York Regents examinations. *American Economic Journal: Applied Economics* 11 (3): 382–423.

Dee, Thomas, Brian Jacob, Justin McCrary, and Jonah Rockoff. 2011. Rules and discretion in the evaluation of students and schools: The case of the New York Regents examinations. Center for Education Policy Analysis Working paper, Stanford University .

Diamond, Rebecca, and Petra Persson. 2016. The long-term consequences of teacher discretion in grading of high-stakes tests. NBER Working paper 22207, National Bureau of Economic Research, Cambridge, MA.

Figlio, David. 2006. Testing, crime and punishment. *Journal of Public Economics* 90:837–51.

Figlio, David, and Lawrence Getzler. 2002. Accountability, ability and disability: Gaming the system. NBER Working paper 9307, National Bureau of Economic Research, Cambridge, MA.

Fisman, Raymond, and Yongxiang Wang. 2017. The distortionary effects of incentives in government: Evidence from China's "death ceiling" program. *American Economic Journal: Applied Economics* 9 (2): 202–18.

Foremny, Dirk, Jordi Jofre-Monseny, and Albert Sollé Ollé. 2017. "Ghost citizens": Using notches to identify manipulation of population-based grants. *Journal of Public Economics* 154:49–66.

Fu, Qiuzi, Zhongnan Fang, Sofia B. Villas-Boas, and George Judge. 2014. An investigation of the quality of air data in Beijing. Unpublished manuscript.

Gelman, Andrew, and Guido Imbens. 2018. Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics* 37 (3): 1–10.

Ghanem, Dalia, and Junjie Zhang. 2014. "Effortless perfection": Do Chinese cities manipulate air pollution data? *Journal of Environmental Economics and Management* 68 (2): 203–25.

Holland, David M., and Terence Fitz-Simons. 1982. Fitting statistical distributions to air quality data by the maximum likelihood method. *Atmospheric Environment (1967)* 16 (5): 1071–76.

Imbens, Guido W., and Thomas Lemieux. 2008. Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142 (2): 615–35.

Jacob, Brian. 2005. Accountability, incentives and behavior: The impact of high-stakes testing in the Chicago public schools. *Journal of Public Economics* 89 (5–6): 761–96.

Jacob, Brian A., and Steven D. Levitt. 2003. Rotten apples: An investigation of the prevalence and predictors of teacher cheating. *Quarterly Journal of Economics* 118 (3): 843–77.

Jenkins, Stephen P. 2014. GB2LFIT: Stata module to fit generalized beta of the second kind distribution by maximum likelihood (log parameter metric). Statistical Software Components, Boston College, Department of Economics.

Jia, Ruixue. 2017. Pollution for promotion. Unpublished manuscript.

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

Li, Hongbin, and Li-An Zhou. 2005. Political turnover and economic performance: The incentive role of personnel control in China. *Journal of Public Economics* 89 (9–10): 1743–62.

Liang, Xuan, Shuo Li, Shuyi Zhang, Hui Huang, and Song Xi Chen. 2016. PM2.5 data reliability, consistency, and air quality assessment in five Chinese cities. *Journal of Geophysical Research: Atmospheres* 121 (17): 10220–36.

McCrary, Justin. 2008. Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics* 142 (2): 698–714.

McDonald, James B. 1984. Some generalized functions for the size distribution of income. *Econometrica* 52 (3): 647–63.

McDonald, James B., and Anand Mantrala. 1995. The distribution of personal income: Revisited. *Journal of Applied Econometrics* 10 (2): 201–4.

Neidell, Matthew. 2009. Information, avoidance behavior, and health: The effect of ozone on asthma hospitalizations. *Journal of Human Resources* 44:450–78.

Reback, Randall, and Julie Berry Cullen. 2006. Tinkering toward accolades: School gaming under a performance accountability system. NBER Working paper 12286, National Bureau of Economic Research, Cambridge, MA.

Rubin, Donald. 1974. Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology* 66 (5): 688–701.

Saez, Emmanuel. 2010. Do taxpayers bunch at kink points? *American Economic Journal: Economic Policy* 2 (3): 180–212.

Shi, Qingling, Chenchen Shi, and Feng Guo. 2020. National leaders' visits and temporary improvement of air quality: Evidence from Chinese cities. *Empirical Economics* 58:2105–27.

Shih, Victor, Christopher Adolph, and Mingxing Liu. 2012. Getting ahead in the Communist Party: Explaining the advancement of Central Committee members in China. *American Political Science Review* 106:166–87.

Stoerk, Thomas. 2016. Statistical corruption in Beijing's air quality data has likely ended in 2012. *Atmospheric Environment* 127:365–71.

Takeuchi, Yoshiyuki. 2004. On a statistical method to detect discontinuity in the distribution function of reported earnings. *Mathematics and Computers in Simulation* 64 (1): 103–11.

Xu, Chenggang. 2011. The fundamental institutions of China's reforms and development. *Journal of Economic Literature* 49 (4): 1076–1151.

Zhang, Junjie, and Quan Mu. 2018. Air pollution and defensive expenditures: Evidence from particulate-filtering facemasks. *Journal of Environmental Economics and Management* 92:517–36.

Zheng, Siqi, Mathew Kahn, Weizeng Sun, and Danglun Luo. 2013. Incentivizing China's urban mayors to mitigate pollution externalities: The role of the central government and public environmentalism. NBER Working paper 18872, National Bureau of Economic Research, Cambridge, MA.