An Investigation of the Quality of Air Data in Beijing

Qiuzi Fu, Zhongnan Fang, Sofia B. Villas-Boas, and George Judge[1]

July 16, 2014

Abstract

With an eye to providing a methodology for identifying the quality of reported air quality data, we use Benford second digit reference distribution to track data of the Air Quality Index (AQI) for 35 monitors over the period 2013-2014 in Beijing. This reference, known as Benford's law, is presented in many naturally occurring numerical data sets and portrays quite closely the air quality monitor data reported by the US Embassy in Beijing over the same time period. In contrast, we find that for a substantial number of Beijing monitors, especially those with large traffic nearby, the reported daily air quality data depart significantly from the Benford reference distribution, while the hourly real-time data fit well. This raises concerns relative to the potential errors in the AQI data and signals coming from air quality monitors. We find that the data deviation level is significantly correlated with the average pollution level and monitors' geographical features like housing prices nearby and traffic flows. If indeed air quality reports contain measurement errors, this could have serious health implications in that individuals do not engage in the right avoidance behaviors for bad air quality days.

Keywords: air monitor data, AQI, reporting, environmental quality, Benford's law, second digit distributions, **JEL classification**: C10, C24.

[1] We thank Junjie Zhang for very helpful comments, Xiaolei Wang, Junjie Zhang, the Beijing Municipal Environmental Protection Bureau (Beijing EPB) and the Embassy of the United States Beijing for granting us access to the data. Fu: Renmin University of China; Fang: Xiamen University; Villas-Boas and Judge: Department of Agricultural and Resource Economics, University of California, Berkeley. Corresponding author: sberto@berkeley.edu.

"The furthest distance in the world, is not the distance between life and death, but when you can't see me when I stand next to you on the street in Beijing"[2]

- *A Chinese Twitter user.*

1. Introduction

Beijing, the capital of China with splendid history, often hits the world headlines for its bad air quality. The air condition in China has long been a source of concern and was brought to the forefront during the Beijing Olympic Games in 2008, when the Chinese government has implemented strict policies to reduce local emissions [1].

Before 2013, the Ministry of Environment Protection (MEP) measured air quality by an index called "API" (Air Pollution Index). This index is criticized for its low standard and accuracy which covers limited pollutants and practitioners. Moreover, China defined a day with an API at or below 100 as a "Blue Sky Day". If the number of "Blue Sky Day" accounts for no less than 80% in a calendar year, the city would be qualified for the "National Environmental Protection Model City Award". It results in a significant discontinuity at the threshold of 100.[2-4]

At the beginning of 2013, the MEP released the official (trial) revisions (GB3095-2012) to the Ambient Air Quality Index (AQI). This is a notable shift from its previous form API) [11]. The indicator of PM2.5 (particular matter smaller or equals 2.5 micrometers in grain diameter, fine particle) was officially added to the new AQI system [12]. Another major change in policy is that the Beijing government stopped using "Blue Sky Day" to evaluate annual air quality performance. Instead, they now use the AQI level, since location-specific data are more precise than one uniform index for the whole city

[2] Originally from https://www.chinadialogue.net/blog/5610-Beijing-residents-find-humour-despite-air-pollution-/en.

[13]. Although this is a significant step forward in measuring air quality, there are still questions about the air quality situation provided by the government.

Starting from the beginning of 2013, heavy haze and smog often appeared in Beijing. At the same time, some Chinese tweeter users had found that the air quality data published by Chinese government deviated greatly from independent data measured by Embassy of the United States in Beijing [5]. Though it turned out to be a confusion resulting from different calculation methods, it has led to public panic over the air quality situation, since inaccuracy of air quality data can lead to huge losses. For example, releases of air quality measures to the public can lead to avoidance behavior on bad-air days and, as a consequence, reduce the incidence of morbidity and mortality outcomes[3]. In particular, those at risk of being negatively affected by pollution may have relatively strong incentives to adopt compensatory/avoidance behaviors. For example, some researchers have found that people responded to information about air quality, with smog alerts leading to significantly reduced attendance at major outdoor facilities in Los Angeles[8].

If the information on which choice is based is flawed, serious health and economic implications follow[9]. Other damages of such occurrences are hard to measure when the population face degraded air quality. Last year alone, almost 700 flights were forced to be cancelled at Beijing airports due to the smog. [10]. Consequently, there is a

[3] As more data on air quality become available, there has been increasing evidence linking measures of pollution positively and significantly to mortality (See 6. Chay, K.Y. and M. Greenstone, *The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession.* The quarterly journal of economics, 2003. **118**(3): p. 1121-1167. and, specifically for China, see 7. Tanaka, S., *Environmental Regulations in China and their impact on air pollution and infant mortality.* Job Market Paper, Boston University, http://bellarmine2. lmu. edu/economics/papers/TANAKA_JMP. pdf, 2010. morbidity, and to reductions in life expectancy 3. Chen, Y., et al., *The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality.* Journal of Environmental Economics and Management, 2013. **66**(3): p. 424-443.).

pressing need for some objective ways to track and screen the performance and integrity of important air quality indicators.

The situation above has inspired us to look into data quality. As the "Blue Sky Day" criteria don't exist anymore, we cannot use the original discontinuity test. Instead, with an eye to providing a methodology for tracking the quality of reported AQI, we use Benford's second digit reference distribution to track the AQI for 35 monitors over the period 2013-2014 in Beijing. This reference, known as Benford's law, is present in many naturally occurring numerical data sets and portrays quite closely with the air quality monitor data reported by the US Embassy in Beijing over the same time period. The method we propose contrasts a reference distribution known as Benford's law, separately to US Embassy contemporaneous monitor data and the distribution of the empirical second digits (SD) of the AQI. According to this law, in many naturally occurring numerical data sets, the digits follow a logarithmic weakly monotonic distribution. Given this empirical phenomenon, Benford's first significant digit (FSD) law has been used in many settings to check data integrity[4]. In these data checks, not only the FSD but also higher digits have been used to detect potential irregularities. In the case of the AQI, the first digits do not often vary much over limited periods of time. Therefore, Benford's law does not span the nine FSD space. In this case, it is possible to use

[4] Among many who have used Benford's law to check the validity of purported scientific data in the social sciences see 14. Varian, H.R., *BENFORDS LAW*, 1972, AMER STATISTICAL ASSOC 1429 DUKE ST, ALEXANDRIA, VA 22314. p. 65-&., 15. Giles, D.E., *Benford's law and naturally occurring prices in certain ebaY auctions.* Applied Economics Letters, 2007. **14**(3): p. 157-161., 16. Tam Cho, W.K. and B.J. Gaines, *Breaking the (Benford) law: Statistical fraud detection in campaign finance.* The American Statistician, 2007. **61**(3): p. 218-223., and 17. Judge, G. and L. Schechter, *Detecting problems in survey data using Benford's law.* Journal of Human Resources, 2009. **44**(1): p. 1-24.. See 18. Abrantes-Metz, R. and P. Bajari, *Screen for Conspiracies and their Multiple Applications.* Antitrust, 2009. **24**: p. 66. how statistical methods have started to be used in antitrust and finance to detect a variety of conspiracies and manipulations.

Benford second digit (SD) law as a reference distribution to track the AQI empirical second digit (0 to 9) frequencies.

We are not the first to report on the reliability of Beijing air quality data. In particular Andrews [4] and Chen et al[3], using different approaches, look into China air quality data. Before 2013, all the published data are in API form and, with the "Blue Sky Day" criteria, most methods fall into a discontinuity test around that threshold (e.g. see Ghanem and Zhang[2]). In related work, Wang et al [19] emphasize the importance of controlling for meteorological conditions; Wang et al. [20], find that the black carbon concentration was significantly better during the periods of traffic control than without the traffic control; The United Nations [21], while controlling for meteorological factors, examine Beijing's before, during, and immediately after the Games, indicating that officially reported API data shows a nationwide trend toward a better air; Finally, Viard and Fu [22] find that traffic restrictions led to improvements in the API.

Our research is also related to a broader literature on environmental policies. Several studies in the US have documented the health effects of air pollution [6, 23, 24], the effect of environmental policies on polluting industries [25-27], and the social costs of environmental policies [28]. Most of these studies suggest that air quality improvement is a long-term process and largely depends on the dynamic interplay of government policies and private compliance. In contrast, the actions that China undertook were largely government-driven, much more intensive, and implemented in a relatively short period. These features help us to separate the effects of Chinese efforts from other confounding factor. They also help us understand how much air quality improvement can be achieved if a government is willing and able to implement intensive measures in a short time.

To summarize, most previous papers that investigate Beijing air data quality, employ a discontinuity test based on environmental policy contexts. However, as the key “Blue Sky Day” policy has been abandoned, this method cannot be used to detect data quality. “Forensic economics” relies on several techniques to detect data quality [29]: one is to compare the reported data with other data sources, which we accomplish by using two kinds of data published by government, and the other is to search for unusual patterns embedded in the data, which we approach it by using Benford’s Second Digit Law as a reference to compare the air quality data to.

We find that in a substantial number of Beijing monitors, the reported daily air quality data depart significantly from the expected Benford second-digit reference distribution. This raises potential concerns relative to the unbiased quality of the signals coming from air quality monitors.

The paper proceeds as follows. In the next section we define the background and measurement of air quality in Beijing. In section 3 we demonstrate how air quality data sets confirm Benford’s Law using US Embassy data. Then we use this reference distribution to assess the quality of two kinds of Chinese AQI data and, in auxiliary analysis, we correlate departures from the reference with monitor characteristics. The last section 4 concludes and discusses implications of the findings.

2. Background

2.1. What is the new AQI and its public health function?

The Beijing AQI has been used for reporting daily air quality in the city of Beijing since Jan 1st 2013 [30]. It represents, in a numeric form, the clean or polluted nature of the air. Based on this number, the public can infer what associated health effects might be a concern for them. These are mainly short-term health effects that one may experience within a few hours or days after breathing polluted air. The Beijing Municipal Environmental Protection Bureau (Beijing EPB)5 is the official organization in charge of monitoring, calculating, and publishing the AQI in Beijing.

According to the new Ambient Air Quality Standards[11], Beijing EPB calculates the AQI for five major air pollutants: ground-level ozone, particle pollution (also known as particulate matter), carbon monoxide, sulfur dioxide, and nitrogen dioxide. For each of these pollutants, the MEP has established national air quality standards to protect public health.

The AQI ranges from 0 to 500 indicating increasing level of health concerns. For example, an AQI value of 50 represents good air quality with little potential to affect public health, while an AQI value over 300 represents hazardous air quality. The purpose of the AQI is to help the public understand what local air quality means to people's health. To make the index easier to understand, the AQI is divided into six categories.

Figure 1. Categories corresponding to Levels of Health Concern

[5] Introduction of Beijing EPB: http://www.bjepb.gov.cn/. They publish 24-hour air quality data of 35 monitors in Beijing in this website: http://zx.bjmemc.com.cn/

Air Quality Index (AQI) Values	Levels of Health Concern	Colors	AQI level
When the AQI is in this range:	..air quality conditions are:	...as symbolized by this color:	
0---50	Good	Green	First Level
51---100	Moderate	Yellow	Second Level
101---150	Unhealthy for Sensitive Groups	Orange	Third Level
151 to 200	Unhealthy	Red	Fourth Level
201 to 300	Very Unhealthy	Purple	Fifth Level
301 to 500	Hazardous	Maroon	Sixth Level

Source: MEP[11] and EPA[31]

Each category corresponds to a different level of health concern: (1) a "Good" AQI is from 0 to 50, where air quality is considered satisfactory, and air pollution poses little or no risk; (2) a "Moderate" AQI is in the range 51 – 100, where air quality is acceptable; however, for some pollutants there may be a moderate health concern for a very small number of people. For example, people who are unusually ozone sensitive may experience respiratory symptoms; (3) "Unhealthy for Sensitive Groups" The AQI falls within 101 - 150. Although the general public is not likely to be affected at this AQI range, people with lung diseases, older adults and children are at a greater risk from exposure to ozone, whereas people suffering from heart and lung diseases, older adults and children are at greater risk from the presence of particles in the air; (4) An "Unhealthy" AQI is in the range of 151 - 200. Everyone may begin to experience some adverse health effects, and members of the sensitive groups may experience more serious effects; (5) A “Very Unhealthy" AQI is 201 - 300. This would trigger a health alert signifying that everyone may experience more serious health effects; Finally (6) a "Hazardous" AQI is greater than 300. This triggers a health warning of emergency conditions and the entire population is more likely to be affected.

2.2. How Beijing's AQI is calculated and published

Scale. The concentrations of five main pollutants are measured at 35 monitoring stations throughout Beijing, over a 24-hour period (noon to noon). The average daily concentration of each pollutant is then converted into a normalized index, which means that each pollutant (ground-level ozone, particle pollution, nitrogen dioxide, carbon monoxide, and sulfur dioxide) is given its own AQI score (called the individual air quality index, and denoted by IAQI), using standard formulas developed by the MEP[32, 33].

According to the MEP, the IAQI is determined by

$$IAQI_p = \frac{IAQI_{Hi} - IAQI_{Lo}}{BP_{Hi} - BP_{Lo}}(C_p - BP_{Lo}) + IAQI_{Lo}$$

For each of the six pollutants (SO_2, NO_2, PM10, CO, O3 (ozone)-1hr and O3(ozone)-8 hour averages) the $IAQI_p$ measures the individual AQI for the pollutant p. The concentration of the pollutant is measured by C_p. If a city has more than one monitoring station, then the average of the pollutant concentrations is used. In the denominator, BP_{Hi} is the pollutant with the highest concentration and BP_{Lo} is the pollutant with the lowest concentration. The numerator, $IAQI_{Hi}$ is the index score of BP_{Hi} (on the IAQI scale of 0-500), and $IAQI_{Lo}$ is the index score of BP_{Lo} (on the IAQI scale of 0-500), where the conversion for each pollutant is listed in Table 1.

Table 1. Concentration normalization table for pollutants in the AQI

IAQI	SO2_24hr average (μg/mg^3)	SO2_1hr average (μg/mg^3)	NO2_24hr average (μg/mg^3)	NO2_1hr average (μg/mg^3)	PM10_24 hr average (μg/mg^3)	CO_24hr average (μg/mg^3)	CO_1hr average (μg/mg^3)	O3_1hr average (μg/mg^3)	O3_8hr average (μg/mg^3)	PM2.5_ 24hr average (μg/mg^ 3)
0	0	0	0	0	0	0	0	0	0	0
50	50	150	40	100	50	2	5	160	100	35
100	150	500	80	200	150	4	10	200	160	75
150	475	650	180	700	250	14	35	300	215	115
200	800	800	280	1200	350	24	60	400	265	150
300	1600	null	565	2340	420	36	90	800	800	250
400	2100	null	750	3090	500	48	120	1000	null	350
500	2620	null	940	3840	600	60	150	1200	null	500

Source: MEP[11]

In the First column of Table 1, the ranges of the index are listed. In the other columns the 24 average values corresponding to each pollutant that fall into the first column ranges in each row are reported. Sulfur dioxide, nitrogen dioxide, and carbon monoxide are denoted as SO2, NO2, CO, respectively. An entry of "null" means that the value for that particular cell exceeds the scale. "Particle pollution" is measured by PM10_24hr (Particles between 2.5 and 10 micrometers in diameter are referred to as "coarse particles") and PM2.5_24hr (The smallest particles, those 2.5 micrometers or less in diameter, also called "fine" particles). Finally, Ground-level ozone" is measured by O3_1hr and O3_8hr.

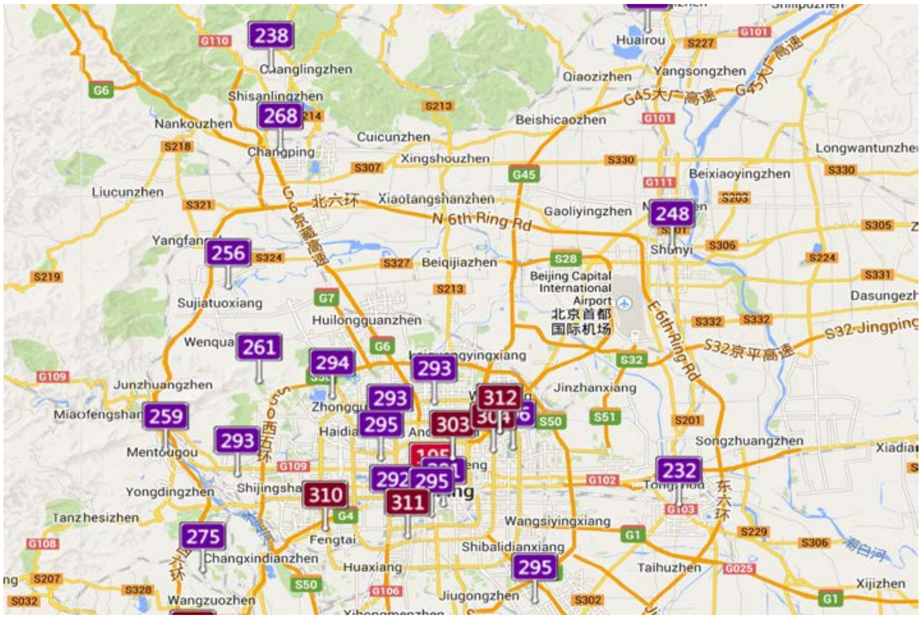
The highest of these *IAQI* values (that is, the primary pollutant) is reported as the AQI value for that day (MEP, 2012).

$$AQI = \max(IAQI_1, IAQI_2, IAQI_3, ..., IAQI_n)$$

Cross-section Monitors. There are 35 monitoring stations in Beijing and their locations are given in Figure 2, which depicts a cross-section of the AQI numbers reported for each monitor in a particular day. These monitors are classified into three types with different functions: (i) an Urban Evaluation Spot; (ii) a Suburb Control Spot; and (iii) a Traffic

Pollution Control Spot. Every day the Beijing EPB publishes AQI in two ways, for each of the 35 monitoring stations and reports those values to the public. We will describe these two kinds of data in detail later.

Figure 2. Cross Section of AQI in each Monitoring Station in Beijing



Source: Beijing EPB. A cross-sectional sample of AQI on a random day. Maroon points show hazardous air quality, purple ones very unhealthy.

3. Benford's Law and Air Quality Data

In order to identify the quality of air quality data reported, we use a method that is based on the empirical phenomenon known as Benford's law (1938).

3.1 Benford's Law

According to this law, in many naturally occurring numerical data sets, the leading digits are not uniformly distributed but instead follow a logarithmic weakly monotonic distribution. Under Benford's law the digit distributions for the first significant digit (FSD) and second digit (SD) are given by

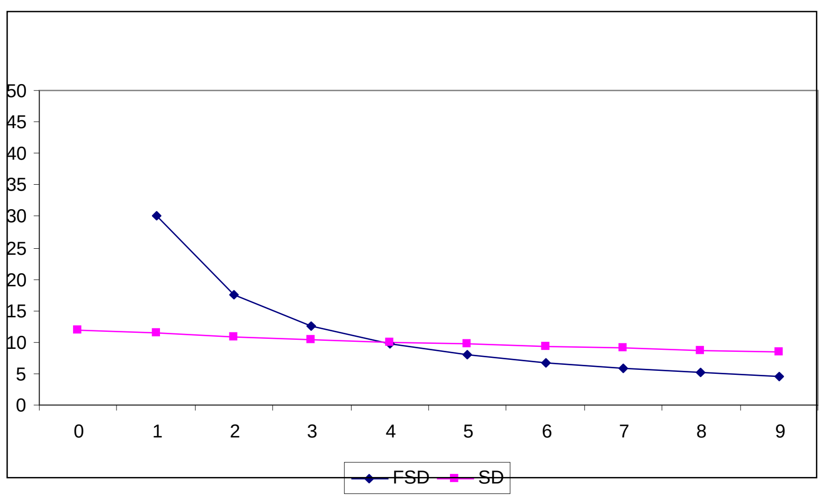
Prob(FSD= d) $= \log_{10}(1 + d^{-1})$, d = 1, 2, …, 9 (2.1.)

And by

Prob(SD= d) $= \sum_{k=1}^{9} \log_{10}\left(1 + (10k + d)^{-1}\right)$, d = 0,..,9, (2.2)

respectively. A graph of the first two digits' distributions is presented in Figure 3.

Figure 3. Benford's Law First and Second Digits Frequencies



Since Benford (1938), others have published studies showing that Benford's FSD law applies to a surprisingly large number of data sets, including populations of cities, electricity usage, and the daily returns to the Dow Jones. Although the FSD's of the AQI for the time period under study do not span the nine-digit space, due to the scale of the index, the SD (see equation 2.2) and following digits data may be expected to naturally do so.

Data set that obeys Benford's law generally follows four main rules. All data must be recorded in the same physical unit; the data set may not contain any inherent minima or maxima; the data set may not include any assigned numbers, which do not arise from any ‘‘natural’’ process; a data set should tend to have more small than large numbers, which also accords with the ‘‘natural’’ development process[34]. Our AQI data only violates the second rule (with maxima 500) but since we only look at second digit, so it has no influence here.

In the next section we demonstrate goodness of fit with Benford's law of two kinds of government data. It both shows how correctly measured hourly air quality data sets confirm to Benford's Law and how badly daily data fits. Then we use this law to assess the quality of the daily AQI data and explore more implications.

For each series of the 35 monitored spots originating data sets, we consider the empirical SD distribution of the Air Quality data. To evaluate the divergence between Benford and the data empirical second digit (SD) distributions in the cross section across the different monitor source locations and the embassy, we use the traditional Chi-Square goodness-of-fit test specified in equation 3. This corresponds to estimating for every monitored spot, a statistic that is obtained as the squared differences of the Benford and the empirical frequencies in certain data bins. e_i is the observed frequency in each bin in the empirical data and b_i is the frequency expected by Benford. Then the chi-square statistic obtained for each data originating spot is given by

$$\chi^2 = \sum_{i=0}^{9} (e_i - b_i)^2 / b_i \, , \qquad (3)$$

This statistic for our data has *9* degrees of freedom with the *10*, *5*, and *1* percent critical significance values of 14.98, 16.92 and 21.97, respectively.

In what follows next, we will use the Benford's second digit (SD) reference distribution to test US embassy monitor data, and to two kinds of Chinese Air Quality SD empirical distributions for the year 2013.

3.2. The Data

The data we use are monitor-specific and originate from monitors spread across the city including one in the US embassy.

Beijing EPB reports two kinds of data, the real-time one and daily one.

Hourly real-time data. Following the same national standard of AQI by MEP[11], Beijing EPB strictly report real-time AQI for all 35 spots every hour [6]. Generally speaking, it releases concentration level of all the above pollutants, and then gives the real-time AQI generated from the equation. So there are 24 hourly real-time AQIs for each monitor each day (24-hour moving average). As it just appear temporarily without available historical archives, we get data from December 5h, 2013 till now from a data scraping website[7].

Daily data. Beijing EPB releases the official daily AQI for all 35 spots one day after in Beijing from January 1st 2013, also according to the new standard. So there is one AQI value for each spot each day. Most people use this kind of data for air quality reference.

[6] See the real-time data here: http://zx.bjmemc.com.cn/
[7] The website is made by Xiaolei Wang: http://beijingair.sinaapp.com/

Relationship between these two kinds of data has been ambiguously described in the new standard too. “Real-time data should be ‘modified’ and ‘audited’ before being published as daily data”[11]. So now we don’t know which kind is more authentic. We will use Benford’s law to find it out later.

U.S. Embassy data. The U.S. Embassy has an air quality monitor to measure PM 2.5 particulates as an indication of the air quality on the Embassy compound located in the Chaoyang district. As the embassy only measures one pollutant, then the PM 2.5’s IAQI is just AQI in that hour. Each hour, the U.S. Embassy publishes a PM2.5 value and thus there are 24 AQI numbers each day, measuring that single air pollutant that we use in our analysis.

3.3. US Embassy data and Benford’s SD distribution

In Table 2, we calculated the SD frequencies, using data from US embassy[8]. Comparing its distribution with the theoretical reference Benford’s SD distribution we note, that the empirical SD of US embassy data and the Benford’s SD distribution are closely related. In terms of a goodness-of-fit, the chi-square of US Embassy is 9.44 (1% significant) we cannot reject the null hypothesis of the distribution equality for the Embassy.

Table 2. Empirical SD frequency for AQI data from US Embassy to Benford’s SD Reference Distribution.

[8] Data from the US Embassy can be found here: http://www.stateair.net/web/historical/1/1.html

digit	US Embassy SD Frequency	Benford SD Frequency
0	0.16	0.12
1	0.10	0.11
2	0.09	0.11
3	0.12	0.10
4	0.08	0.10
5	0.10	0.10
6	0.11	0.09
7	0.09	0.09
8	0.08	0.09
9	0.07	0.09
Chi-square	**9.44**	
1% significant	14.98	
5% significant	16.92	
10% significant	21.97	

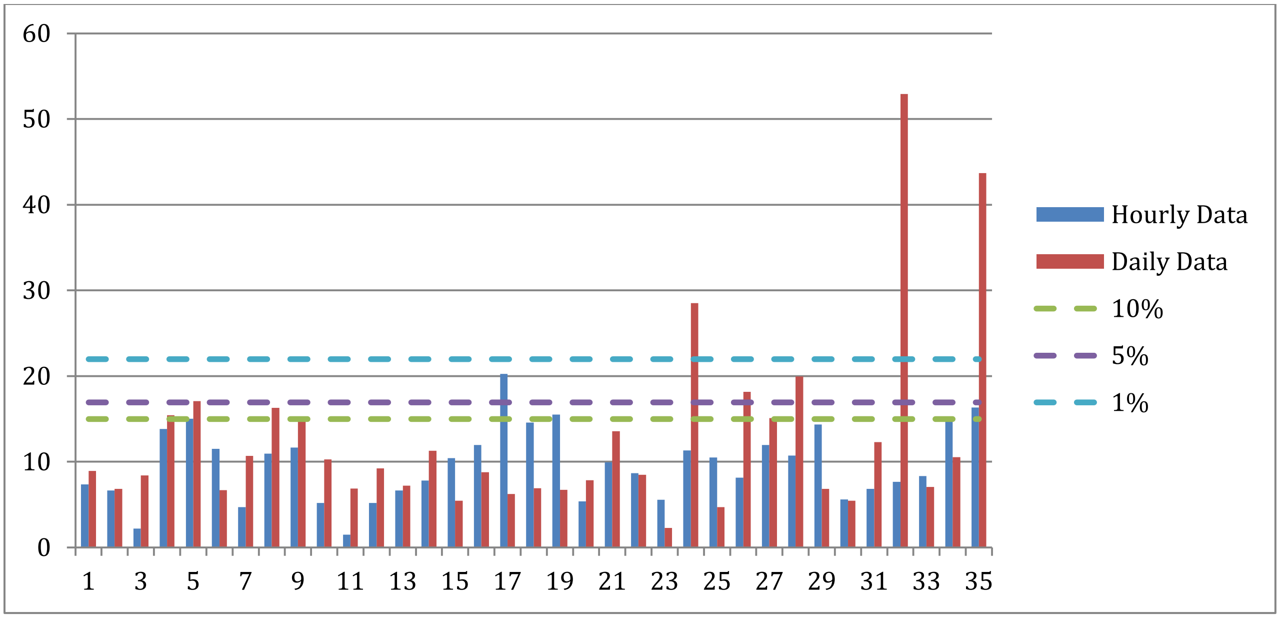
Source: Author's calculations.

The take away from this table is that US Embassy is providing air quality data during this period that is similar to Benford. Enlightened by this, we will use this tool to test two kinds of Chinese air data and explore more implications.

3.4. Compare two kinds of Chinese data

As the government publishes two kinds of data, one question naturally comes out. Which of these two kinds is more reliable? To deal with it, we made the following comparison according to Benford's law.

Figure 5. Chi-Square Statistics of two kinds of data, measuring Differences in Distribution of Each monitor spot empirical SD distribution and Benford SD reference



Source: Author's calculations.

This graph compares chi-squares between two kinds of data. For each monitor, we measured the departure of its digit distribution to Benford and tested significance. To be precise, we have more hourly real-time observations than daily data and significance test of chi-squares depends on data amount. So we randomly choose same amount of observations from two data sets. The above graph is one of the test results (they are all stable and similar). From the chat, it's easy to find that hourly data, though after transformation, generally follow Benford's law within the 10% significance level but daily data from several certain monitors are suspicious.

Before this confirmation, we cannot directly use Benford's law here, as we don't know the original distribution of this data set. AQI data are transformed from criteria pollutant concentrations. We are not sure whether this transformation affects distribution neither.

Since Benford's article appeared, there have been numerous attempts by mathematicians, physicists and amateurs to "prove" Benford's law, but there have been a main stumbling block, that is, some data sets satisfy the law and some do not, and there never was a very clear definition of an general statistical experiment that would predict which would and which would not. So if current data follow Benford's law closely, or if hypotheses of unbiased random samples from random distributions seem reasonable, then the predicted data should also follow Benford's law closely. Such a "Benford-in, Benford-out" test is at best only a double-check on reasonableness, since the law says nothing about the raw data themselves[35]. That is to say, every time before we use Benford's Law to check for data falsification, we are required to prove the assumption that true data follow Benford distribution[36].

Now the hourly real-time data have already fit Benford's law. So this comparison has laid a solid foundation for our following analysis. Now we can confidently use Benford's law to test unreliable daily air data.

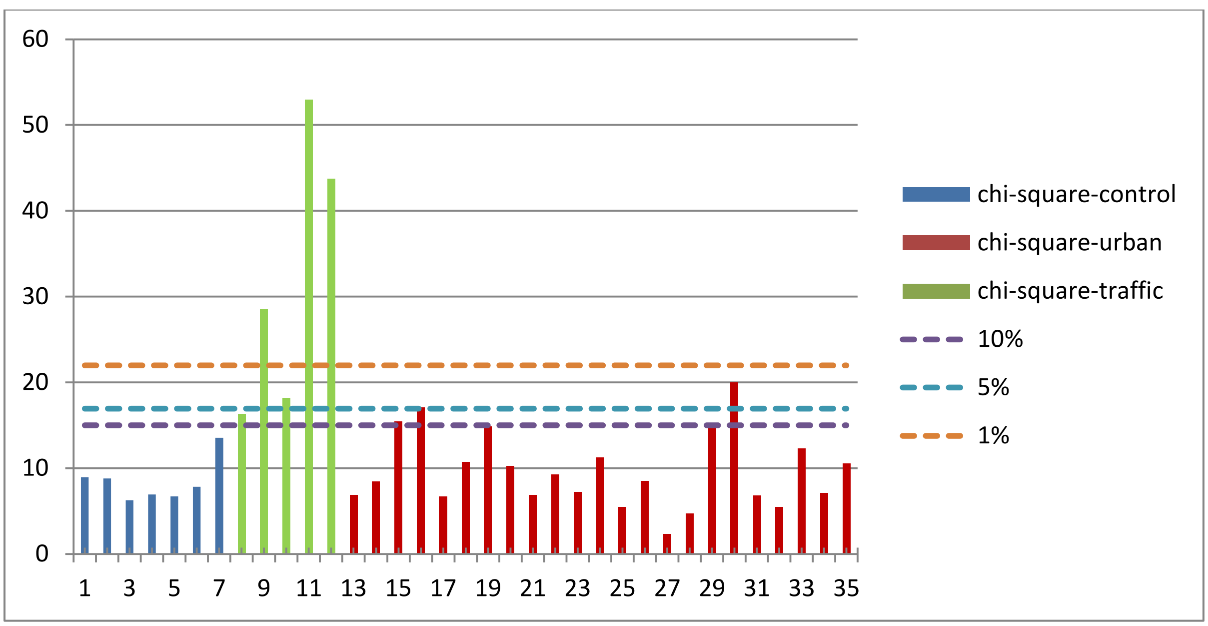
3.5 Cross-Sectional Patterns of Departure from the Reference SD distribution

As the next step, to assess the possibility of manipulation in daily data, for each of the 35 spot monitors separately, we compare the SD distribution to Benford and estimate the 35 chi-square statistics of departure from this reference distribution. As a second step, having investigated the possibility of manipulation, the question is why some monitors but not others have evidence consistent with this outcome. To answer this question we relate the chi-square statistics (which, being large and above the critical values, are suggestive of manipulation) to the observable features of each of the 35 monitor types.

We do this by using a linear regression of the chi-squares on observable cross sectional spot specific variables: a spot indicator (35 different spots), the spot type, the AQI level, the type of pollutant that underlies each spot monitor's AQI, and other spot specific characteristics (their distance to city center, housing price nearby[9], average and peak traffic flowand private car/taxi percentage nearby[37]).

For each of the 35 monitors, we use the chi-square statistics to measure departure from the Benford reference distribution and assess the *possibility of manipulation*. These estimates are reported in Figure 6 below as well as the 1 percent, 5 percent and 10 percent significance critical values.

Figure 6. Chi-Square Statistics of the Difference in Distribution of Each monitor spot empirical SD distribution and Benford SD reference, by spot type.



Source: Author's calculations.

[9] Source: http://bj.ganji.com/fang5/

These chi-square statistics are reported for each of the 35 monitor spots, as well as the 1 percent, 5 percent and 10 percent significance critical values, in blue, light green, and dark green respectively. In Figure 6 the chi-square results reflect that while most spots have data quite similar to the Benford reference, monitors 5, 24, 32 and 35 differ and are statistically significantly different from the Benford reference. We also highlighted chi-square of different monitor type into different color. Arranging by spot type, it's easy to find that five of the eight highest chi-squares originate from Traffic Pollution Control spots (that's all Traffic Pollution Control spots too), while Urban Evaluation Spots and Suburb Control Spots appear to agree with Benford's Law better.

The next question is why some monitors agree, but not others and why these data are different from the data reference? Motivated by the suggestive descriptive fact that Traffic Pollution Control spots appear to suffer from potential manipulation, we use a linear statistical model to investigate this issue. We take advantage of the fact that we observe different features of these spots (their distance to city center, housing price nearby, average and peak traffic flow and private car/taxi percentage nearby, etc.). Table B1 in Appendix B contains the summary statistics of the dependent variable (the chi-squares) and of all the available variables we correlate to the estimated chi-squares (see Table B2 in Appendix B). Using a principal components analysis (see Appendix B for details), we find further that (i) air quality data from monitors that have high housing prices nearby, large traffic flows around, and are closer to the city center are more correlated with higher manipulation levels, and (ii) data from monitors that report worse average air quality level and more PM2.5 demonstrate possible manipulation.

4. Conclusion

To investigate the manipulation of air quality data, we have analyzed the Chinese AQI, using as a methodology, the second digit distribution variant of Benford's law. We find that the US embassy data and hourly real-time data follow Benford's second distribution closely while some of the daily AQI data do not follow such distribution. Particularly, Traffic Pollution Control monitors have the highest divergence from the reference distributions. We further find suggestive evidence that air quality data from monitors that have high housing prices nearby, large traffic flows around, and are closer to the city center are more correlated with higher manipulation levels. Interestingly, monitors that report worse average air quality level and more PM2.5 have larger divergence from the reference.

The fact that data reported is not revealing actual air quality can have very serious health effects. Therefore, devising ways to monitor the data quality is very important. In fact, the Standing Committee of China's National People's Congress (NPC), the country's top legislature, has voted to adopt revisions to the Environmental Protection Law. The revised Environmental Protection Law is the first change to the legislation in 25 years[38]. In this most powerful legislation, the environmental category entry 68 point 6 says, "Officials of bureaus who manipulate or counterfeit environmental data will be held responsible and face punishment"[39]. Hopefully, this kind of strict legislation could improve data quality in the future.

Following this paper, we analyze the quality of air quality data over time, given that its quality is important to then evaluate policies. We follow a methodology similar to this present paper, but now use roll out periods over time to assess the conformity of the air data to the Benford reference distributions. This is a useful extension of this work given that is it important to have good data to assess the success of recent strict policies to improve air quality implemented by the Chinese government.

References

1. Chen, Y., et al., *Gaming in Air Pollution Data?: Lessons from China*, 2013, National Bureau of Economic Research.
2. Ghanem, D. and J. Zhang, *Effortless Perfection: Do Chinese Cities Manipulate.* Available at SSRN, 2013.
3. Chen, Y., et al., *The promise of Beijing: Evaluating the impact of the 2008 Olympic Games on air quality.* Journal of Environmental Economics and Management, 2013. **66**(3): p. 424-443.
4. Andrews, S.Q., *Inconsistencies in air quality metrics:'Blue Sky'days and PM10 concentrations in Beijing.* Environmental Research Letters, 2008. **3**(3): p. 034009.
5. C., Z. *潘石屹的环保经.* 2013; Available from: http://finance.sina.com.cn/roll/20130409/022115079732.shtml.
6. Chay, K.Y. and M. Greenstone, *The impact of air pollution on infant mortality: evidence from geographic variation in pollution shocks induced by a recession.* The quarterly journal of economics, 2003. **118**(3): p. 1121-1167.
7. Tanaka, S., *Environmental Regulations in China and their impact on air pollution and infant mortality.* Job Market Paper, Boston University, http://bellarmine2. lmu. edu/economics/papers/TANAKA_JMP. pdf, 2010.
8. Neidell, M., *Information, Avoidance behavior, and health the effect of ozone on asthma hospitalizations.* Journal of Human Resources, 2009. **44**(2): p. 450-478.
9. He, G., *The Effect of Air Pollution on Cardiovascular Mortality: Evidence from the 2008 Beijing Olympic Games.* Working paper U.C.Berkeley, 2014.
10. Karimi, S.J.F., *Beijing adopts emergency measures for 'hazardous' pollution*, in *CNN*2013.
11. (MEP), T.M.o.E.P., *Technical Regulation on Ambient Air Quality Index Daily Report GB3095-2012*, 10.2012.
12. Gao, F., *Evaluation of the Chinese New Air Quality Index (GB3095-2012): Based on Comparison with the US AQI System and the WHO AQGs.* 2013.
13. Wang, S. *"蓝天"数下岗"浓度"值登场.* 2012; Available from: http://epaper.jinghua.cn/html/2012-06/06/content_797393.htm.
14. Varian, H.R., *BENFORDS LAW*, 1972, AMER STATISTICAL ASSOC 1429 DUKE ST, ALEXANDRIA, VA 22314. p. 65-&.
15. Giles, D.E., *Benford's law and naturally occurring prices in certain ebaY auctions.* Applied Economics Letters, 2007. **14**(3): p. 157-161.

16. Tam Cho, W.K. and B.J. Gaines, *Breaking the (Benford) law: Statistical fraud detection in campaign finance.* The American Statistician, 2007. **61**(3): p. 218-223.
17. Judge, G. and L. Schechter, *Detecting problems in survey data using Benford's law.* Journal of Human Resources, 2009. **44**(1): p. 1-24.
18. Abrantes-Metz, R. and P. Bajari, *Screen for Conspiracies and their Multiple Applications.* Antitrust, 2009. **24**: p. 66.
19. Wang, W., et al., *Atmospheric Particulate Matter Pollution during the 2008 Beijing Olympics.* Environmental Science & Technology, 2009. **43**(14): p. 5314-5320.
20. Wang, X., et al., *Evaluating the air quality impacts of the 2008 Beijing Olympic Games: on-road emission factors and black carbon profiles.* Atmospheric Environment, 2009. **43**(30): p. 4535-4543.
21. UNEP, *Independent Environmental Assessment: Beijing 2008 Olympic Games*, 2009.
22. Viard, V.B. and S. Fu, *The effect of Beijing's driving restrictions on pollution and economic activity.* 2013.
23. Almond, D., et al., *Winter Heating or Clean Air? Unintended Impacts of China's Huai River Policy.* American Economic Review, 2009. **99**(2): p. 184-90.
24. Currie, J. and M. Neidell, *Air pollution and infant health: what can we learn from California's recent experience?* The Quarterly Journal of Economics, 2005. **120**(3): p. 1003-1030.
25. Becker, R. and V. Henderson, *Effects of air quality regulations on polluting industries.* Journal of Political Economy, 2000. **108**(2): p. 379-421.
26. Henderson, V., *Effects of air quality regulation*, 1995, National Bureau of Economic Research.
27. List, J.A., et al., *Effects of environmental regulations on manufacturing plant births: evidence from a propensity score matching estimator.* Review of Economics and Statistics, 2003. **85**(4): p. 944-952.
28. Hazilla, M. and R.J. Kopp, *Social cost of environmental quality regulations: A general equilibrium analysis.* Journal of Political Economy, 1990. **98**(4): p. 853-73.
29. Zitzewitz, E., *Forensic economics.* Journal of Economic Literature, 2012. **50**(3): p. 731-769.
30. Bureau, B.M.E.P. *1 月 1 日起我市将按新标准发布空气质量信息.* 2012.12; Available from: http://www.bjepb.gov.cn/bjepb/323474/331443/331810/331826/451742/index.html.
31. EPA. *Air Quality Index (AQI) - A Guide to Air Quality and Your Health*. 2014; Available from: http://airnow.gov/index.cfm?action=aqibasics.aqi.
32. Hsu, A. *China's new Air Quality Index: How does it measure up?* 2012.
33. Hsu, A. *China amends air quality measures but misses key pollutant – PM 2.5.* 2011; Available from: http://hsu.me/2011/03/china-amends-air-quality-measures-but-misses-key-pollutant-pm-2-5/.
34. Nigrini, M.J., *Digital Analysis Using Benford's Law: Tests & Statistics for Auditors*. 2000: Global Audit Publications.

35. Hill, T.P., *The First Digit Phenomenon: A century-old observation about an unexpected pattern in many numerical tables applies to the stock market, census statistics and accounting data.* American Scientist, 1998: p. 358-363.
36. Holz, C., *The Quality of China's GDP Statistics.* 2013.
37. 北京交通发展研究中心, *2013 年北京交通发展年报*. 2014.
38. *China's NPC adopts revised Environmental Protection Law*. 2014; Available from: http://news.xinhuanet.com/english/video/2014-04/25/c_133289139.htm.
39. *中华人民共和国环境保护法（自 2015 年 1 月 1 日起施行）*, MEP, Editor 2014.