Benford’s Law and COVID-19 Reporting*

Christoffer Koch†  Ken Okamura‡
Research Department  Saïd Business School
Federal Reserve Bank of Dallas  University of Oxford

April 28, 2020

Abstract

Trust in the reported data of contagious diseases in real time is important for policy makers. Media and politicians have cast doubt on Chinese reported data on COVID-19 cases. We find Chinese confirmed infections match the distribution expected in Benford’s Law and are similar to that seen in the U.S. and Italy and thus find no evidence of manipulation. Policy makers in the rest of the world should trust the Chinese data and formulate policy accordingly.

Keywords: Corona COVID-19, statistical reporting, government accountability

JEL Codes: E61, E65, H41, I18,

Highlights:

- We find no evidence of manipulation of Chinese COVID-19 data using Benford’s Law.
- Models on the trade-off between growth and deaths can be calibrated with Chinese data.
- Future Chinese data post-quarantine should guide policy in other countries.

---

*The views expressed in this paper are those of the author and are not necessarily reflective of views at the Federal Reserve Bank of Dallas or the Federal Reserve System. Any errors or omissions are the sole responsibility of the authors.

†2200 North Pearl Street, Dallas, TX 75201, United States. christoffer.koch@dal.frb.org
‡Park End Street, Oxford, OX1 1HP, United Kingdom. ken.okamura@sbs.ox.ac.uk

This preprint research paper has not been peer reviewed. Electronic copy available at: https://ssrn.com/abstract=3586413
1 Chinese Reporting on the Coronavirus

Contrary to popular speculation, we find no evidence that the Chinese massaged their COVID-19 statistics. We use a statistical fraud detection technique, Benford’s (1938) Law, to assess the veracity of the statistics. This empirical finding is important because China was affected first. Policies to combat the global pandemic are informed by its response. Skepticism about the Chinese data may result – and may already have resulted – in poor policy choices or result in the public not accepting policy decisions to the detriment of society.

The media frequently claim the Chinese government has understated the numbers of those affected. Politicians echo these claims with President Trump declaring the reported death toll and infections seemed “a little bit on the light side”. The on-going doubts over the credibility of its published data are problematic as it impacts subsequent policy choices by countries that saw epidemics later. Recent papers that rely on Chinese data for calibration and analysis include: Models of economic activity and the trade-off with deaths such as Atkeson (2020), Jones, Philippon, and Venkateswaran (2020) and Alvarez, Argente, and Lippi (2020); Fang, Wang, and Yang (2020) predict the effect of movement restrictions on the spread of the disease; the infection model produced by Imperial College London Ferguson et al. (2020) that informed U.K. government policy. Since the policy choice of many countries has been to undertake social distancing, travel bans, and lockdowns patterned after the successful choices made by China, it is important that policy makers know the data is reliable.

Lack of confidence in the Chinese data may have led to a slower response in Europe to the emergent pandemic. Chinese provinces neighboring Hubei province, the Chinese epicentre, imposed movement controls, quarantines and checks on January 23rd at a time when the number of confirmed cases in Hubei was 444 and the number of deaths was 17. Most major cities in Hubei, including the provincial capital Wuhan, were also locked down at this time. In comparison Italy, Europe’s initial pandemic hotspot, reached 445 cases on February 26th and 17 deaths the following day. Italy imposed a regional quarantine on 50,000 people on February

---

1“What to make of China’s Coronavirus figures” Foreign Policy, April 1st 2020. James Palmer. "Can China’s COVID-19 statistics be trusted?” The Diplomat, March 26th 2020. Scott Romaniuk and Tobias Burgers.
2One interesting note is that an early and influential paper on the macroeconomic effects of pandemics Eichenbaum, Rebelo, and Trabandt (2020) does not rely on the Chinese data for calibration in the main.
3See Qiu, Chen, and Shi (2020).
4Hubei province has a population of 58.5 million, the city of Wuhan has a population of 11 million, making these entities roughly similar to Italy and the Lombardy region.
Italy did not, however, impose a full regional lockdown until March 7\textsuperscript{th}. It even took until March 9\textsuperscript{th} for the lockdown to be extended to the entire nation. Similarly, restrictions on international travel were too mild. France made travel advisories against travelling to part of Northern Italy and merely requiring self-quarantine for those returning from these areas on February 23\textsuperscript{nd}. Austria and Switzerland waited until March 10\textsuperscript{th} and March 8\textsuperscript{th} to impose controls on cross-border movement. By February 26\textsuperscript{th}, Hubei had seen cases rise to 65,187 and deaths to 2,615.

Chinese data from February had already shown the effectiveness of the quarantine measures in slowing the progress of the disease, both within Hubei and in other Chinese provinces see Figure 1. The daily percentage change in the number of cases fell from over 100\% to 30\% within a week. Within two weeks it had fallen to 10\%. By the end of February, the daily percentage change in the number of confirmed cases in China had fallen to 1\%. Yet, governments in Europe chose to delay their responses or to engage in far less aggressive measures. The popular media was skeptical about the effectiveness of quarantines\textsuperscript{5} and of the decline in Chinese cases.\textsuperscript{6,7}

This skepticism about politically motivated manipulation of Chinese state statistics is deeply rooted. Anecdotal and academic evidence point to lower level officials manipulating data to meet targets. In 2007, Chinese Premier Li Kejiang called all GDP measures “man-made and therefore not reliable” when discussing data on Liaoning province.\textsuperscript{8} Lyu, Wang, Zhang, and Zhang (2018) find evidence that regional growth rates are manipulated to meet growth targets. Specifically, in the case of the SARS outbreak in 2002-3, criticism of the Chinese response surfaced. The World Health Organization (WHO) suspected that China underreported the number of cases (see Parry, 2003; Ashraf, 2003).\textsuperscript{9}

\textsuperscript{5}“Why we should be skeptical of China’s Coronavirus quarantine” Washington Post, January 24\textsuperscript{th} 2020. Howard Markel. “As Coronavirus fears intensify, effectiveness of quarantines is questioned”. New York Times, January 26\textsuperscript{th} 2020. Chris Buckley, Raymond Zhong, Denise Grady, and Roni Caryn Rabin.

\textsuperscript{6}“Are China’s Coronavirus figures reliable?” Foreign Policy, February 19\textsuperscript{th} 2020. James Palmer.

\textsuperscript{7}One reason for the skepticism was the jump in cases on February 12\textsuperscript{th} when doctors in Hubei were instructed to change the definition of cases from confirmed by laboratory tests to those with evidence of congestion on a chest x-ray. This single data point has no significance for the analysis of conformity with Benford’s Law.

\textsuperscript{8}US diplomatic cable March 15\textsuperscript{th}, 2007. Leaked to Wikileaks.

\textsuperscript{9}Since the SARS outbreak, international reporting requirements saw major changes in the form of the International Health Regulations 2005 that bound countries to “greater notification of public health emergencies of international concern” and improvements to Chinese domestic disease control efforts. In the 2013 H7N9 outbreak the Chinese public health system worked well (Wang, 2013).
Some data manipulation clearly took place early in the epidemic. The number of cases reported by the Wuhan authorities was “frozen” at 41 during the Hubei provincial Chinese People’s Political Consultative Conference and the Wuhan People’s Congress (Lianghui) between January 12th and 17th, 2020. John Mackenzie, a member of the WHO emergency committee told the Financial Times on February 5th that China must have been withholding information on new cases from the WHO. By January 20th, Chinese President Xi Jinping issued instructions to all levels of government to put people’s safety first and take effective measures to combat the epidemic. We will test the hypothesis that local governments ignored this admonition and chose to manipulate their COVID-19 case numbers.

2 Benford’s Law

Benford’s Law is used to detect fraud or flaws in data collection based on the distribution of the first digits of observed data. A Benford distribution of first digits arises naturally for processes that are exponential with multiple changes of magnitude, Michalski and Stoltz (2013). The spread of COVID-19 demonstrates exponential growth and changes of magnitude.

The frequency with which the first digit is “1” is 30.1%, the first digit is “2” is 17.6% etc, declining to the first digit being “9” only 4.6% of the time. Since it takes a 100% increase to go from “1” to “2” and a mere 11.1% increase to go from “9” to “1”, this logarithmic distribution makes sense, as long as the data series is sufficiently large and has a number of changes of magnitude. See Table 1

\[ P(d) = \log_{10} \left( \frac{1 + d}{d} \right) \quad \text{for} \quad d \in (1, \ldots, 9) \]  
(Benford’s Law)

| First Digit | 1   | 2   | 3   | 4   | 5   | 6   | 7   | 8   | 9   |
|-------------|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| Benford Distribution Probability | 0.301 | 0.176 | 0.125 | 0.097 | 0.079 | 0.067 | 0.058 | 0.051 | 0.046 |

Table 1: Benford’s Law Distribution of First Digit

10The most obvious example being the treatment of Dr Li Wenliang, the doctor pressured by police for spreading rumors who died after contracting the coronavirus Green (2020).

11“WHO expert says China too slow to report coronavirus cases” Financial Times, February 5th 2020. Primrose Riordan and Sue-Lin Wong.
The use of Benford’s Law to detect fraud has been widely demonstrated in economics and accounting (Varian, 1972). In economics, Benford’s Law has been used to determine if economic statistics have been manipulated by governments: Nye and Moul (2007), Gonzales-Garcia and Pastor (2009), Rauch, Göttche, Brähler, and Engel (2011), Michalski and Stoltz (2013) and Holz (2014). In accountancy, Nigrini (1996) examined taxpayer records and used it to detect tax evasion. It has also been used in epidemiology, Idrovo, Fernández-Niño, Bojórquez-Chapela, and Moreno-Montoya (2011) examined the reported weekly number of confirmed cases from 35 countries during the A(H1N1) pandemic in 2009.

3 Fraud Detection Suggests No Data Manipulation

We compile data from the Johns Hopkins University Corona Virus Research Center, for Chinese provinces and U.S. states. and for Italian provinces we use data from the daily Dipartimento della Protezione Civile bulletins. We use the daily number of cases/deaths/recoveries announced for the analysis. We focus on periods when the number of national cases grows by at least 10% as after this the data series no longer follow an exponential path and the distribution of first digits is not likely to follow Benford’s Law (Table 2). The spread and scale of cases and deaths are shown in Figure 2 - 7 for China, U.S. and Italy and in Figure 8 for recoveries in China. For confirmed cases the data series have multiple changes in magnitude for most Chinese provinces, U.S. states and Italian regions. Deaths do not show these multiple changes in magnitude for Chinese provinces and only to a limited extent in the U.S. and Italy. Chinese provincial recoveries do show multiple changes in magnitude. More than two changes of magnitude are desirable for a series to follow Benford’s Law, which is why we focus on the number of confirmed cases.

The number of confirmed cases understates the true number of infections as China was unable to test those who did not present at hospitals. However, the lack of tests is a problem in both Italy and the U.S. The number of cases detected by health authorities in countries lacking testing capacity is still valuable, see Harris (2020).

Figure 9 shows that for Chinese provinces, U.S. states and Italian regions the number of confirmed cases the distribution of the first digits shows a decline from 1 to 9 in line with the

---

12Data last retrieved on April 12, 2020.
expected distribution of Benford’s Law. Figure 10 show the distribution of deaths is out of line with Benford for Chinese provinces, but is reasonably close for the U.S. and Italy. Figure 11 shows for Chinese provinces the number of recoveries appears close to Benford’s Law.

Tests of significance for Benford’s Law require that the "true" distribution should follow the Benford distribution. Our null hypothesis is that the observed distribution follows the theoretical (Benford) distribution. The most common test is the Chi-Square test of Goodness of Fit:

$$D^2 = n \sum_{d=1}^{9} \frac{(h_d - p_d)^2}{p_d}$$

(Chi-Square test)

Where $n$ denotes the number of observations, $h$ is the observed frequencies of the digits and $p$ is the Benford’s Law distribution. We also use a Kuiper test (a modified Kolmogorov–Smirnov test).

$$T_K = (D_n^+ + D_n^-) \left[ \sqrt{n} + 0.155 + \frac{0.24}{\sqrt{n}} \right]$$

(Kuiper test)

Where $D_n^+ = \sup(H_d - P_d)$ and $D_n^- = \sup(P_d - H_d)$ and $H_d$ and $P_d$ represent the cumulative frequencies of the first digit $d$ in the observed data and the Benford distribution. We also calculate the $m$ (max) statistic proposed by Leemis, Schmeiser, and Evans (2000) where $m = \max_{d=1,...,9} |h_d - p_d|$ and the $d$ (distance) statistic proposed by Tam Cho and Gaines (2007) $d = \sqrt{\sum_{d=1}^{9}(h_d - p_d)^2}$. We use the Benford-specific tests of significance provided by Morrow (2014) for the Kuiper, $m$ and $d$ tests.

Table 3 displays our results. The Chi-Square test does not support a Benford distribution...
| Country   | Category         | Leading Digit | Kuiper | χ²-Stat | dₚ* | mₚ* | V˚ |
|------------|------------------|---------------|--------|---------|-----|-----|----|
| China      | Confirmed Cases  | 166 99 72 44 50 37 34 22 12 | 536    | 11.08*** | 0.85 | 0.54 | 0.60 |
| US         | Confirmed Cases  | 394 213 146 92 101 71 66 55 55 | 1,193  | 10.49*** | 1.28* | 1.00** | 1.13 |
| Italy      | Confirmed Cases  | 145 81 57 43 42 28 23 26 18 | 463    | 2.65*    | 0.47 | 0.26 | 0.34 |
| China      | Deaths           | 73 9 6 3 2 2 2 1 1 99 | 90.54*** | 4.63*** | 4.34*** | 5.02*** |
| US         | Deaths           | 225 149 74 63 38 34 43 25 19 | 670    | 26.31*** | 1.81*** | 1.20** | 1.72** |
| Italy      | Deaths           | 151 78 50 35 33 26 19 15 14 | 421    | 11.14*** | 1.33** | 1.15** | 1.19 |
| China      | Recoveries       | 210 127 79 47 37 34 34 21 29 | 618    | 16.84*** | 1.50** | 0.96* | 1.88*** |

Notes: ***, **, and * denotes statistical significance at the 1%, 5%, and 10% level.

Table 3: Table of First Digit Distribution and Test of Significance

except for Italian regions’ number of confirmed cases, where the null hypothesis is rejected at the 10% level. However, as noted in other papers on Benford’s Law, the Chi-Square test is extremely sensitive for large sample sizes and tends to reject statistical significance even for small differences. The Kuiper test does not reject the null hypothesis that the distribution is Benford for China, U.S. and Italy for confirmed cases. For China and Italy, the d and m tests also do not reject the null. For the U.S. the m test rejects the null at the 5% level and the d test rejects the null at the 10% level. These tests of significance are supportive of the view that for confirmed cases, the distribution of the first digits follows Benford’s Law for all three countries.

In general, for deaths most of the tests reject the null hypothesis at the 1% level although for Italian deaths the Kuiper test does not reject the null and the d and m tests reject the null at the 5% level. For US deaths the m test rejects the null at the 5% level. The US and Italy are closer to Benford than China probably because they have a higher number of deaths and a wider geographical spread of deaths. Finally, Chinese recoveries show the Kuiper test rejects the null at the 1% level, the d rejects the null at the 10% level and m rejects the null at the 5% level.
4 Conclusion

China’s distribution of first digits for confirmed cases is in line with Benford’s Law. Thus we reject the hypothesis that the Chinese data has been manipulated. It also matches the distribution found in the United States and Italy. An advantage of Benford’s Law is the inherent difficulty of coordinating a distortion of figures in real time on a panel basis. It is possible to create data series that fit Benford’s Law (Diekmann, 2007). To manipulate the Chinese data in this fashion would require someone to coordinate daily announcements across all provinces while accurately forecasting future infection rates. This is improbable.

As China is at least a month ahead of Europe and six weeks ahead of the United States, its data should be used not only for calibration of models to inform policy measures to slow infection, but also for guidance in the lifting of stay-at-home orders.
Figures

Figure 1: Confirmed Cases and Daily Rates of Change in Hubei and China excluding Hubei

Figure 2: Maximum Daily Confirmed Cases by Chinese Province

Source: Johns Hopkins University.
Figure 3: Maximum Daily Confirmed Cases by U.S. State

Figure 4: Maximum Daily Confirmed Cases by Italian Region
Figure 5: Maximum Daily Deaths by Chinese Province

Figure 6: Maximum Daily Deaths by U.S. State
Figure 7: Maximum Daily Deaths by Italian Region

Figure 8: Maximum Daily Recoveries by Chinese Province
Figure 9: Confirmed Cases in Chinese Provinces, U.S. States and Italian Regions
Figure 10: Deaths in Chinese Provinces, U.S. States and Italian Regions

Figure 11: Recoveries in Chinese Provinces
References

Alvarez, F. E., D. Argente, and F. Lippi (2020). *A Simple Planning Problem for COVID-19 Lockdown*. Working Paper 26981. National Bureau of Economic Research.  
doi: [10.3386/w26981](https://doi.org/10.3386/w26981).

Ashraf, H. (2003). “China finally throws full weight behind efforts to contain SARS”. *The Lancet* 361.9367, p. 1439.  
doi: [10.1016/S0140-6736(03)13150-9](https://doi.org/10.1016/S0140-6736(03)13150-9).

Atkeson, A. (2020). “What Will Be the Economic Impact of COVID-19 in the US? Rough Estimates of Disease Scenarios”. Working Paper Series 26867.  
doi: [10.3386/w26867](https://doi.org/10.3386/w26867).

Benford, F. (1938). “The Law of Anomalous Numbers”. *Proceedings of the American Philosophical Society* 78.4, pp. 551–572.  
doi: [http://www.jstor.org/stable/984802](http://www.jstor.org/stable/984802).

Diekmann, A. (2007). “Not the First Digit! Using Benford’s Law to Detect Fraudulent Scientific Data”. *Journal of Applied Statistics* 34.3, pp. 321–329.  
doi: [10.1080/02664760601004940](https://doi.org/10.1080/02664760601004940).

Eichenbaum, M. S., S. Rebelo, and M. Trabandt (2020). *The Macroeconomics of Epidemics*. Working Paper 26882. National Bureau of Economic Research.  
doi: [10.3386/w26882](https://doi.org/10.3386/w26882).

Fang, H., L. Wang, and Y. Yang (2020). *Human Mobility Restrictions and the Spread of the Novel Coronavirus (2019-nCoV) in China*. Working Paper 26906. National Bureau of Economic Research.  
doi: [10.3386/w26906](https://doi.org/10.3386/w26906).

Ferguson, N., D. Laydon, G. Nedjati Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhatia, A. Boonyasiri, Z. Cucunuba Perez, and G. Cuomo-Dannenburg (2020). “Report 9: Impact of non-pharmaceutical interventions (NPIs) to reduce COVID19 mortality and healthcare demand”.  

Gonzales-Garcia, J and G Pastor (2009). “Benford’s Law and Macroeconomic Data Quality”. *International Monetary Fund Working Paper 2009-10*.  
doi: [10.5089/9781451871579.001](https://doi.org/10.5089/9781451871579.001).
Green, A. (2020). “Li Wenliang”. *The Lancet* 395.10225, p. 682.  
DOI: 10.1016/S0140-6736(20)30382-2.

Harris, J. E. (2020). *The Coronavirus Epidemic Curve is Already Flattening in New York City*. Working Paper 26917. National Bureau of Economic Research. 
DOI: 10.3386/w26917.

Holz, C. A. (2014). “The Quality of China’s GDP Statistics”. *China Economic Review* 30, pp. 309–338. 
DOI: 10.1016/j.chieco.2014.06.009.

Idrovo, A., J. Fernández-Niño, I Bojórquez-Chapela, and J Moreno-Montoya (2011). “Performance of public health surveillance systems during the influenza A (H1N1) pandemic in the Americas: testing a new method based on Benford’s Law“. *Epidemiology & Infection* 139.12, pp. 1827–1834. 
DOI: 10.1017/S095026881100015X.

Jones, C. J., T. Philippon, and V. Venkateswaran (2020). *Optimal Mitigation Policies in a Pandemic: Social Distancing and Working from Home*. Working Paper 26984. National Bureau of Economic Research. 
DOI: 10.3386/w26984.

Leemis, L. M., B. W. Schmeiser, and D. L. Evans (2000). “Survival Distributions Satisfying Benford’s Law”. *The American Statistician* 54.4, pp. 236–241. 
DOI: 10.1080/00031305.2000.10474554.

Lyu, C., K. Wang, F. Zhang, and X. Zhang (2018). “GDP Management to Meet or Beat Growth Targets”. *Journal of Accounting and Economics* 66.1, pp. 318–338. 
DOI: 10.1016/j.jacceco.2018.07.001.

Michalski, T. and G. Stoltz (2013). “Do Countries Falsify Economic Data Strategically? Some Evidence That They Might”. *The Review of Economics and Statistics* 95.2, pp. 591–616. 
DOI: 10.1162/REST_a_00274.

Morrow, J. (Aug. 2014). *Benford’s Law, Families of Distributions and a Test Basis*. CEP Discussion Papers dp1291. Centre for Economic Performance, LSE. 
DOI: https://ideas.repec.org/p/cep/cepdp/dp1291.html.

Nigrini, M. J. (1996). “A Taxpayer Compliance Application of Benford’s Law”. *The Journal of the American Taxation Association* 18.1, p. 72.
Nye, J. and C. Moul (2007). “The political economy of numbers: on the application of Benford’s law to international macroeconomic statistics”. *The BE Journal of Macroeconomics* 7.1.  
*doi*: 10.2202/1935-1690.1449.

Parry, J. (2003). “WHO is worried that China is under-reporting SARS”. *BMJ* 326.7399, p. 1110.  
*doi*: 10.1136/bmj.326.7399.1110-a.

Qiu, Y., X. Chen, and W. Shi (2020). “Impacts of Social and Economic Factors on the Transmission of Coronavirus Disease 2019 (COVID-19) in China”. *Journal of Population Economics* 33.4, Forthcoming.  
*doi*: 10.1101/2020.03.13.20035238.

Rauch, B., M. Göttscbe, G. Brähler, and S. Engel (2011). “Fact and Fiction in EU-Governmental Economic Data”. *German Economic Review* 12.3, pp. 243–255.  
*doi*: 10.1111/j.1468-0475.2011.00542.x.

Tam Cho, W. K. and B. J. Gaines (2007). “Breaking the (Benford) law: Statistical fraud detection in campaign finance”. *The American Statistician* 61.3, pp. 218–223.

Varian, H. R. (1972). “Benford’s Law (Letters to the Editor)”. *American Statistician* 26.3, p. 65.  
*doi*: 10.1080/00031305.1972.10478934.

Wang, Y. (2013). “The H7N9 influenza virus in China–changes since SARS”. *N Engl J Med* 368.25, pp. 2348–2349.  
*doi*: 10.1056/NEJMp1305311.