Testing the emission reduction claims of CDM projects using the Benford's Law

Cole, Matt; Maddison, David; Zhang, Liyun

DOI:
10.1007/s10584-019-02593-5

Document Version
Publisher's PDF, also known as Version of record

Citation for published version (Harvard):
Cole, M, Maddison, D & Zhang, L 2019, 'Testing the emission reduction claims of CDM projects using the Benford's Law', *Climatic Change*. https://doi.org/10.1007/s10584-019-02593-5

Climatic Change
https://doi.org/10.1007/s10584-019-02593-5

Testing the emission reduction claims of CDM projects using the Benford's Law

Matthew A. Cole[1] · David J. Maddison[1] · Liyun Zhang[1]

Received: 8 March 2019 / Accepted: 30 October 2019/Published online: 22 November 2019

Abstract
Benford's Law suggests that the first digits of numerical data are heavily skewed towards low numbers. Data that fail to conform to Benford's Law when conformity is to be expected may have been manipulated. Using Benford's Law, we conduct digital frequency analysis on the emission reduction claims of Clean Development Mechanism projects. Digital frequency analysis indicates that although emission reduction claims made in project design documents do not conform to Benford's Law, we cannot reject the null hypothesis that data on certified emission reductions do. Benford's Law offers a rapid, low-cost means of identifying possible instances of data manipulation.

Keywords Clean Development Mechanism · Emission reductions · Benford's Law · Data manipulation

1 Introduction

It has long been understood that if data underpinning environmental policy are unreliable then the policies based upon these data may be inefficient. One cause of data unreliability is deliberate manipulation by those reporting it. The threat of deliberate manipulation of GHG emission data is, however, inadequately addressed in the climate change literature. This is not because of any consensus that the scope for misreporting GHG emissions is limited and certainly not because the consequences of misreporting are benign. Indeed, it is widely agreed that formulating an appropriate response to climate change is vital and the integrity of GHG emission data is the cornerstone upon which national and international climate policies are built.

One technique that has been used to examine the integrity of data rests upon a statistical phenomenon known as Benford's Law (BL). A large number of processes have been shown to result in data conforming to BL and when these processes occur the distribution of the first (and subsequent) digits does not follow a uniform distribution, but is instead heavily skewed towards low

✉ Matthew A. Cole
m.a.cole@bham.ac.uk

[1] Department of Economics, Birmingham Business School, University of Birmingham, Birmingham B15 2TT, UK

numbers. Nonconformity with BL in instances where conformity was to be expected is viewed as indicative of possible data mishandling or even manipulation, and BL is for this reason widely used in forensic auditing e.g. Durtschi et al. (2004), Dlugosz and Müller-Funk (2009) and Nigrini (2012). There are also a few applications of BL examining the integrity of environmental data although surprisingly, none of these involves climate policy or greenhouse gases (GHGs).

This paper uses BL to examine the integrity of emission reduction data from Clean Development Mechanism (CDM) projects. The CDM has played a key role in international climate policy but many have expressed concerns about the additionality of such projects and the claimed extent of any emission reductions. Some commentators have even concluded that future research into the CDM should focus on the 'development of fraud detection systems for the CDM and…quantitative and qualitative screening of the claims of additionality as part of a broader system of fraud identification.' (Drew and Drew 2010, page 250).

Apart from being a pillar of global climate policy, the CDM provides an interesting application of BL for other reasons. First, there is a clear incentive for manipulating emission reductions since projects with greater emission reductions will be more attractive. Second, it is possible to observe emission reductions both prior to and after significant auditing activity. Specifically, we take data from the United Nations Framework Convention on Climate Change (UNFCCC) CDM website and examine the expected emissions reduction (EER) claims contained in project design documents (PDDs) and the eventual issuance of certified emission reductions (CERs). Third, it is possible to stratify the data enabling us to locate the source of any nonconformity. Fourth, there are good reasons to suppose that, in the absence of data manipulation, CDM GHG emission reduction projects should conform to BL in view of the processes involved. Our reasons for believing CDM GHG emission reductions ought to conform to BL are discussed below.

In sum, this paper provides the first statistical analysis of the integrity of emission reductions from CDM projects; the first application of BL to any baseline-and-credit system; and the first application to GHGs. More generally, and in contrast to the largely US focus of the existing BL literature, this is also the first study to use BL to examine environmental data across a range of countries.

To anticipate our main findings, there is evidence that EERs do not always conform to BL. Such a finding is not inconsistent with the suggestion that there is data manipulation at this point in the CDM project cycle. It is notable, however, how well BL often describes CDM GHG emission reductions. This finding throws into sharp relief those instances where it does not and supports in our view, the application of BL to other GHG emission datasets. Most importantly, however, we cannot reject null hypothesis that the distribution of CERs conforms to BL.

The remainder of the paper is organised as follows. Section 2 explains BL and the statistical processes that give rise to it. Section 3 reviews previous attempts to apply BL to emissions and pollution concentrations. Section 4 describes the CDM project cycle and section 5 presents data on CDM GHG emission reductions. Section 6 discusses the various tests used by us to assess conformity with BL. Section 7 presents the results which are then further explored in section 8. Finally, section 9 considers the role BL might play in testing for manipulation of GHG emission data, whether by firms or governments.

2 Benford's Law

BL is a counterintuitive property of data whereby the leading digit of a number ($d1$) is more likely to be small than large. According to BL, the probability of $d1 = d$ is given by $log_{10}\left(\frac{d+1}{d}\right)$

so the probability that $d1 = 1$ is 30.1% decreasing to 4.6% for $d1 = 9$. A similar rule applies to the second ($d2$) and subsequent digits, although after the fourth digit, each digit appears with a frequency of approximately 10.0%.

The highly nonuniform distribution of leading digits was discovered by Newcomb (1881) but popularised by Benford (1938). The idea was prompted by the observation that the first few pages in logarithmic tables always wear out more quickly. Benford (op cit) collected 20,229 observations on data including data on river lengths, populations, scientific constants and numbers contained in newspapers and confirmed that in these data, the distribution of digits followed BL.

Data conforming to BL is widely encountered throughout the social, physical and biosciences (see e.g. Pietronero et al. 2001; Nigrini and Miller 2007; Fu et al. 2007; Tam Cho and Gaines 2007; Judge and Schechter 2009; Moret et al. 2009; Joannes-Boyau et al. 2015).

The ubiquity of BL is due to many common statistical processes having BL as a limiting distribution. For example, geometric sequences will generally follow BL (Raimi 1976). Lemons (1986) considers the possible distributions of pieces of some conserved quantity. Assuming each distribution is equally likely, the expected distribution turns out to conform to BL. When random variables are repeatedly multiplied, divided or raised to an integer power, the data that ensue eventually conform to BL (Boyle 1994). If distributions are randomly selected and samples are randomly drawn from these distributions, and furthermore supposing that the average distribution is scale or base-invariant (see below), the distribution of significant digits will conform to BL (Hill 1998). Lastly, any distribution will approximate BL if it spans several orders of magnitude and is smooth (Fewster 2009).

A Benfords set also possesses the property of scale invariance (Pinkham 1961). This means that units of measurement do not influence whether data conform to BL. For example, irrespective of whether monetary amounts are expressed in Dollars, Sterling or Euros, this should not affect whether data conform to BL. The digit frequencies corresponding to BL are the only set of frequencies possessing this property (Hill 1995).

Some explanations appear more useful than others in understanding why particular data conform to BL. In the case of stock markets, an explanation for conformity with BL based on geometric sequences seems appropriate. If, however, the challenge is to explain why the size of freshwater lakes conforms to BL, the most appropriate explanation is that they form part of a conserved quantity. The reason why accounting data conforms to BL might be because such data are the product of price × quantity.

We expect CDM GHG emission reduction project data to conform to BL. First, the data comprises random samples from random distributions, each one relating to a different sort of project. As argued by Hill (op cit), this is sufficient to generate a distribution that conforms to BL. Second, the distribution of emission reductions spans many orders of magnitude (see below) and there is no reason why the distribution should not be smooth. Finally, emission reduction claims are the consequence of mathematical operations e.g. emissions avoided multiplied by the number of wind turbines (although multiple mathematical operations are generally required before a distribution conforms closely to BL).

Because we are relying in part on the argument that multiple orders of magnitude combined with smoothness are a reason for expecting conformity with BL, it is perhaps worth providing a more in depth explanation of the point.

Conformity with BL amounts to the claim that the areas under the probability density function of emission reductions between the values of 1.0–1.9, 10.0–19.9, 100.0–199.9 and

1000.0–1999.9 tons of carbon etc. sum to 0.301. Now take the Log_{10} of these quantities to obtain $Log_{10}(1.0)$-$Log_{10}(1.9)$, $Log_{10}(10.0)$-$Log_{10}(19.9)$, $Log_{10}(100.0)$-$Log_{10}(199.9)$, $Log_{10}(1000.0)$-$Log_{10}(1999.9)$ etc. The distance on the horizontal axis corresponding to each of these 'stripes' is also 0.301.

Now consider the areas under the probability density function between the values of 2.0–2.9, 20.0–29.9, 200.0–299.9 and 2000.0–2999.9 tons of carbon etc. Here, conformity with BL amounts to the claim that the areas under the probability density function amounts to 0.176. Once more taking the Log_{10} of the quantities, the distance on the horizontal axis corresponding to each of these stripes is 0.176. The same process is performed where the leading digit is 3–9. The distances on the horizontal axis corresponding to each of these stripes range from 0.124 to 0.045.

We now make the simple observation that the combined lengths of the distances on the *x*-axis provide an approximation of the areas under the entire probability density function. This approximation, moreover, becomes better and better as the number of orders of magnitude increases. The approximation could, however, still be poor if the probability density function is not smooth.

Having explained when BL holds and why we expect it to do so, here it is important to mention situations where BL does not apply. Assigned numbers such as telephone numbers do not follow BL. Numbers influenced by human psychology, including manipulated numbers, likewise tend not to conform to BL. Indeed, the manner in which a distribution containing manipulated numbers departs from BL reveals the emphasis placed on particular numbers. When individuals manipulate numbers they likely use heuristic techniques. Collins (2017) contends that someone using a computer keyboard is likely to use their more dominant index and middle fingers and hence type 4, 5, 6 or 7 more often than other digits.

Other situations where BL does not apply are those where there is a minimum (other than zero), or a maximum. Critically for our purposes, however, there is no minimum or maximum value for the CDM GHG emission reductions.

There are some forms of data manipulation that BL cannot detect e.g. multiplying all figures in a dataset by the same amount (a consequence of the property of scale invariance). BL cannot distinguish between accidental mishandling and deliberate manipulation. Even the innocent practice of rounding-off terminal digits results in nonconformity with BL. Most fundamentally there is the problem of type-I and type-II errors when testing the null hypothesis that data conform to BL. Put differently, statistical tests may point to possible data manipulation when it is absent and the absence of data manipulation when it is present. BL is accordingly best viewed only as a first step in examining data for possible data manipulation and not as a substitute for auditing.

3 Applications of Benford's Law to analysing emissions and pollution concentrations

The first attempt to apply BL to emission data is Dumas and Devine (2002). Using data from North Carolina 1996–1998, they analyse firms' self-reported emissions of volatile organic compounds. They argue that these data should conform to BL in the absence of manipulation or other forms of data mishandling because the data exhibits multiple orders of magnitude and can be considered as a random sample from random distributions. Tests reveal that not only is there nonconformity with BL, there is, as expected, evidence of statistically significant

underreporting of between 9.5 and 10%. They suggest that when evidence of nonconformity is more apparent in some sectors than others, it might pay the regulator to focus attention on the nonconforming sectors.

According to de Marchi and Hamilton (2006), BL offers a simple way to conduct triage on self-reported emission data, although critically, the test does not allow the regulator to determine precisely which plants are misreporting data (if any). They use data from the US Toxic Release Inventory (TRI). These data refer to both emissions and off-site transfers of 12 chemicals. Two substances, nitric acid and lead, fail to conform to BL with the distribution of $d1$ containing too many 2s and 5s.

Fugitive and stack emissions, as well as off-site transfers, have all been subject to analysis using BL. Critically, however, some of these are easier to conceal than others. Using data on lead from the US TRI, Zahran et al. (2014) demonstrate that data for off-site transfers conforms most closely to BL. Changes in the extent of nonconformity with BL are also shown to coincide with changes in emission monitoring regulations.

Apart from emissions and off-site transfers, BL has also been applied to pollution concentrations. Brown (2005) analyses the $d1$ frequencies of 8 selected UK datasets comprising pollution concentrations. Results indicate that the annual average and weekly concentrations of 12 measured heavy metals at 17 monitoring sites conform very closely to BL. The key determinant of whether data conform to BL appears to be how many orders of magnitude are present in the data, with more than four orders of magnitude resulting in a close correspondence. Brown (op cit) also examines the effects of data mishandling where, for a percentage of the data, $d1$ is dropped such that $d2$ is then misinterpreted as $d1$, $d3$ as $d2$ and so on. Conformity with BL is very sensitive to data being mistreated in this way. Another form of accidental data mishandling discussed is where $d1$ and $d2$ become accidentally transposed.

Fu et al. (2014) analyse data from the Chinese Air Quality Index (AQI) taken from 35 different sites in Beijing. They use BL to test the frequency of $d2$. The data the authors use relates to 2013–2014, the period immediately after the Blue Sky Days initiative had ended. Blue Sky Days are days when the Chinese AQI is below a threshold of 100 and this was once used as a performance indicator for evaluating local officials (hence the incentive to misreport). They find that although hourly data conform to BL, daily data do not.

One way to validate the use of BL is by comparing data collected by different parties, one of which has no incentive to manipulate the data. Data manipulation is suspected if data collected by the disinterested party conforms to BL whereas the other data does not. Using this approach, Stoerk (2016) compares Beijing Municipal Environmental Protection Bureau (BMEPB) air pollution data with data recorded by an air pollution monitor run by the US embassy. Whereas data from the US embassy conforms to BL throughout the entire period of investigation, conformity is poor in early years for the BMEPB data. This discrepancy persists when aerosol optical density data from satellites is used instead of the US embassy data.

Recently, Beiglou et al. (2017) use BL to examine wastewater discharges in Ohio. Data are taken from 223 facilities and cover a period of 3 years. Measurements relate to a variety of wastewater parameters: microbial, nutrients, metals and solids. The authors screen the data prior to analysis. Parameters not spread over at least one order of magnitude were deemed unsuitable for analysis using BL. The authors find that conformity with BL differs greatly across different wastewater parameters.

A notable feature of the literature is the desire to demonstrate that the data are such that BL ought to hold absent manipulation or other sorts of data mishandling. An alternative strategy is to validate BL by demonstrating that BL holds only when either the opportunity/incentive to

manipulate the data is absent. Interestingly, there are no published attempts to substantiate the use of BL by demonstrating that numbers with excess frequencies are associated with use of heuristic techniques. There are likewise no attempts to analyse environmental data known with certainty to have been manipulated. Finally, there are no prior attempts to use BL to analyse GHG emissions; something we now seek to correct.

4 The Clean Development Mechanism

The CDM is an arrangement under the Kyoto Protocol whereby emission reduction projects in developing countries earn CERs which can then be sold. The two objectives of the CDM are to help Annex I Parties cost-effectively meet part of their emission reduction targets under the Kyoto Protocol and to assist non-Annex I Parties in achieving sustainable development. CDM projects can be bilateral whereby an Annex I country develops a project or unilateral whereby a non-Annex I country develops projects and sells the CERs.

A significant body of research examines the accomplishments of the CDM (Olsen 2007; Sutter and Parreno 2007; Lecocq and Ambrosi 2007; Paulsson 2009); opinion is, however, divided over the success of the scheme. Some point to its achievements in reducing the costs of mitigating GHGs and in promoting a degree of north-south technology transfer (Aslam 2001; Haites et al. 2006; Pearson 2007; Dechezleprêtre et al. 2008, 2009; Seres et al. 2009). Others allege that CDM projects do not help countries to achieve sustainable development (Anagnostopoulos et al. 2004; Gundimeda 2004; Karakosta et al. 2009; Nussbaumer 2009; Bumpus and Cole 2010; Crowe 2013). Many researchers express scepticism about the additionality of CDM projects (Zhang and Wang 2011). The transactions costs of CDM projects (which include the costs of auditing) moreover appear to be high (Krey 2005; Bellassen et al. 2015).

The CDM project cycle involves seven steps. First, the project participants prepare a PDD with a detailed description of the proposed CDM project, including estimated emission reductions, a methodology supporting their estimates and importantly, evidence of the additionality of the project. The PDD is then submitted to and reviewed by an accredited designated operational entity (DOE) contracted by the project participants. These DOEs are approved third-party auditors. After the review, the DOE proceeds with the validation of the CDM project by preparing a validation report which confirms that the proposed project is a valid project. The PDD is then made publicly available by the DOE on the UNFCCC CDM website for comments. Second, the project developer secures a letter of approval from the designated national authority (DNA) of the host country. The letter of approval confirms that the project meets the host country's sustainable development criteria, complies with the country's laws and regulations and fulfils any other requirement specified by the DNA. Third, following the host country's approval, the DOE validates the PDD. Fourth, after determining that the proposed project meets all relevant requirements of the CDM, the DOE submits the project to CDM Executive Board (EB) with request for registration. The project is registered if there are no objections from member countries or at least three EB members. Fifth, if the project is registered and operating, the project participant monitors actual emission reductions made by the project according to approved methodology and submits a monitoring report to the DOE. Sixth, the DOE verifies the actual emission reductions and if satisfied prepares verification and certification reports. The DOE who verifies the emission reductions cannot be the same that validates the project except for small-scale projects. The DOE submits these

reports to EB with request for issuance of CERs. Finally, EB issues CERs to the project participants through the CDM registry. A significant number of projects fail to complete the project cycle (Cormier and Bellassen 2013).

In view of the linking of the two schemes, it is interesting to compare the auditing requirements of the CDM with those of the EU emissions trading scheme (EUETS) where the auditing requirements seem less stringent (Warnecke 2014). Some researchers have also drawn attention to the critical role of the DOEs and argued that DOEs are affected by misaligned incentives: DOEs might not want to be too strict if the effect is to scare away new business (Drew and Drew, op cit).

China is the largest host country for CDM projects in terms of the number of projects and the largest supplier of CERs on the CDM market. *Measures for the Operation and Management of Clean Development Mechanism Projects* is the regulatory framework for CDM implementation in China. This includes detailed guidance on the eligibility, application and approval procedures for CDM projects (Zhang and Wang 2011; Fay et al. 2011). All hydro and wind projects as well as all new combined cycle natural gas power plants are required to be submitted through the CDM (Karakosta et al. 2009).

5 Testing emission reduction data from CDM projects using Benford's Law

Data are sourced from the UNFCCC CDM website. The dataset includes detailed information on CDM projects, including project title, project type, project classification, host country, methodology, project status, type of crediting period and CERs issued. There are in the dataset we use 12,880 CDM activities in total including 12,382 project activities and 498 programmes of activities. The dataset that we use was last updated on 10 October 2016. It includes projects starting the validation process from 1 December 2003 to 5 October 2016.

Table 1 describes the data in terms of project status. Project status is also presented for the two main host countries: China and India. Of the 12,880 projects contained in the data, 8038 (62.40%) are registered. The next largest category refers to 2856 projects (22.17%) whose validation was terminated.

Turning to type of project, Table 2 reveals that the most common project involved wind (3065 projects corresponding to 23.79%) closely followed by hydro (3020 projects corresponding to 23.44%). For wind, these percentages are even higher for China and India (32.04% and 32.43% respectively). Whereas in China, hydro constitutes 34.50% of all projects, in India, the figure is only 9.78%. India has numerous biomass energy projects (659) constituting 19.60% of Indian projects. By contrast, China has only 220 such projects constituting only 4.37% of all its projects.

Table 3 presents information on EERs and CERs broken down for China and India. Note that in total, there are 12,675 observations on EERs rather than 12,880. The reason for this discrepancy is that some projects do not include information on emission reductions. Likewise, the number of observations on EERs for registered projects is 8035 rather than 8038 since for three projects, this information is missing. In no case has the same project inadvertently been included multiple times (the unique project identifiers contained in the data are all different). Evident from comparing the minimum and maximum values in Table 3 is the fact that both EERs and CERs the data span 4–6 orders of magnitude. This means that if the probability density function of emission reductions is smooth, these data ought to conform to BL.

Table 1 Projects by status

	All		China		India	
Status	No.	%	No.	%	No.	%
Bundled Registered: Ref 0784	1	0.00	0	0	0	0
Deregistered	2	0.01	0	0	1	0.02
Pending publication	187	1.45	34	0.67	67	1.99
Registered	8038	62.40	3806	75.60	1660	49.39
Rejected	276	2.14	96	1.90	54	1.60
Requesting registration	6	0.04	0	0	3	0.08
Review requested	2	0.01	0	0	1	0.02
Validation public	632	4.90	85	1.68	324	9.63
Validation replaced	812	6.30	236	4.68	216	6.42
Validation terminated	2856	22.17	760	15.09	1016	30.22
Withdrawn	67	0.52	17	0.33	19	0.56
Withdrawn before publication	1	0.00	0	0	0	0

Listed are the numbers and shares of CDM projects of different status. Projects hosted in China and India are separately identified

Table 2 Project type by country

	All		China		India	
Project type	No.	%	No.	%	No.	%
Afforestation	20	0.15	3	0.05	1	0.02
Agriculture	6	0.04	1	0.01	4	0.11
Biogas	1	0.00	0	0	1	0.02
Biomass energy	1506	11.69	220	4.37	659	19.60
Cement	98	0.76	46	0.91	27	0.80
CO_2 usage	6	0.04	0	0	0	0
Coal bed/mine methane	183	1.42	167	3.31	4	0.11
EE households	240	1.86	29	0.57	85	2.52
EE industry	339	2.63	24	0.47	229	6.81
EE own generation	837	6.49	496	9.85	242	7.20
EE service	104	0.80	3	0.05	47	1.39
EE supply side	211	1.63	43	0.85	76	2.26
Energy distribution	40	0.31	20	0.39	6	0.17
Fossil fuel switch	255	1.97	53	1.05	86	2.55
Fugitive	99	0.76	4	0.07	19	0.56
Geothermal	44	0.34	3	0.05	0	0
HFCs	27	0.20	13	0.25	9	0.26
Hybrid renewables	27	0.20	0	0	4	0.11
Hydro	3020	23.44	1737	34.50	329	9.78
Landfill gas	637	4.94	156	3.09	59	1.75
Methane avoidance	1195	9.27	149	2.95	101	3.00
Mixed renewables	20	0.15	4	0.07	7	0.20
N_2O	117	0.90	51	1.01	9	0.26
PFCs and SF6	24	0.18	2	0.03	2	0.05
Reforestation	90	0.69	8	0.15	26	0.77
Solar	553	4.29	175	3.47	187	5.56
Tidal	2	0.01	0	0	0	0
Transport	72	0.55	10	0.19	29	0.86
Wind	3065	23.79	1613	32.04	1090	32.43
Not specified	42	0.32	4	0.07	23	0.68

Listed are the numbers and shares of CDM projects of different types. Projects hosted in China and India are separately identified

Table 3 Expected and certified emission reductions by country

	Country	Number of projects	Percentage of projects	Mean	Std. Dev.	Min.	Max.
Expected emission reductions	All	12675	100	131941.8	443409.9	13	21931348
	China	5010	39.5	158514.5	420950.4	465	10437249
	India	3338	26.4	96640.7	520092.4	13	21931348
Of which registered projects	All	8035	100	128839.2	387223.5	13	10437249
	China	3806	47.4	157327.5	441039.2	465	10437249
	India	1660	20.7	72401.3	271649.2	13	4017202
Certified emission reductions	All	2854	100	497722.7	3205171	3.91	74928479
	China	1451	50.8	603326.3	3682392	715.2	63011172
	India	597	20.9	241138.4	1487168	16.2	29688497

Listed are summary statistics of expected and certified emission reductions from all CDM projects, as well as those hosted in China and India

6 Methodology

We start by visually comparing the frequencies of the significant digits of the observed data with the expected (Benford) frequencies. Then, we use the χ^2, Kolmogorov-Smirnov (K-S) and Kuiper statistical tests described below to test for conformity with BL the distributions of $d1$, $d2$ and $d1d2$. The latter combination is preferred because it captures more information; there are 90 possible digit combinations (10–99 inclusive). We delete all positive numbers that are less than 10 as they do not have a second digit.

The χ^2 test provides good insight into the general fit over the entire range of the distribution. Other things being equal a higher χ^2 value indicates a larger deviation of the observed frequencies from the expected Benford frequencies. The χ^2 statistic is for $d1$ calculated as follows:

$$\chi 2 = N\sum_{d=1}^{d=9} \frac{\left(p(d)_{\text{Observed}} - p(d)_{\text{Benford}}\right)^2}{p(d)_{\text{Benford}}}$$

where N is the sample size and $p(d)_{\text{Observed}}$ and $p(d)_{\text{Benford}}$ are respectively, the observed and Benford frequencies for digit d. For $d1$, the χ^2 is calculated with $9 - 1 = 8$ degrees of freedom. This test may be extended to examine the distribution of $d2$ and $d1d2$. For $d2$, the degrees of freedom are however, $10 - 1 = 9$ and for $d1d2$, the degrees of freedom are $90 - 1 = 89$. The null hypothesis of this test is that the distribution of digital frequencies observed corresponds to BL. An important shortcoming of the χ^2 test is that it is dependent on sample size. As the sample size increases, the probability of rejecting the null hypothesis grows. What this means is that the χ^2 test statistic cannot be used to compare two distributions in terms of how well they conform to BL when sample sizes differ.

The K-S test may also be used to test whether two underlying probability distributions differ. The K-S statistic D is calculated as follows:

$$D = sup\left|F(d)_{\text{Benford}} - F(d)_{\text{Observed}}\right|$$

where *sup* is the supremum function and $F(d)$ is the cumulative distribution of d. The K-S test relies upon the maximum absolute difference between the theoretical and observed cumulative

distributions. The null hypothesis of this test is also that the distribution of digital frequencies observed corresponds to BL. Although widely used to compare discontinuous distributions, an important weakness is that the critical values assume a continuous distribution. The critical values are too high when the distribution is discontinuous.

The Kuiper test provides another means of determining whether two probability distributions differ. Giles (2007) recommends using this test rather than the K-S test to investigate whether data conform to BL. It suffers from the same problem as the K-S test in that the critical values assume the distribution is continuous. The test involves calculating the maximal extent of the deviation of $F(d)_{\text{Observed}}$ over $F(d)_{\text{Benford}}$ (D^+) as well as the maximal extent of the deviation of the observed cumulative distribution below the theoretical one (D^-). Where N is the sample size, the Kuiper test statistic V_N^* is as follows:

$$V_N^* = V_N\left(N^{\frac{1}{2}} + 0.155 + 0.24N^{-\frac{1}{2}}\right)$$

Where:

$$V_N = D^+ + D^-$$

Finally, we include the Kullback-Leibler (K-L) statistic. This provides a measure of the information lost when one distribution is used to approximate another. We use the K-L statistic as a means of ranking distributions in terms of their conformity with BL. Where $p(d)_{\text{Observed}}$ represents the observed distribution and $p(d)_{\text{Benford}}$ represents the Benford distribution, the K-L statistic D_{KL} is for $d1$:

$$D_{\text{KL}}\left(p(d)_{\text{Observed}} \| p(d)_{\text{Benford}}\right) = \sum_{d=1}^{d=9} p(d)_{\text{Observed}} ln \frac{p(d)_{\text{Observed}}}{p(d)_{\text{Benford}}}$$

The K-L measure can be extended to $d2$ and $d1d2$.

7 Results

Commencing with the full sample of EERs contained in the PDD data, the correspondence with BL is poor. Table 4 contains the χ^2, K-S and Kuiper test results for $d1$, $d2$ and $d1d2$. All the tests reject at the 1% level of significance the null hypothesis that the distribution conforms to BL.

With a dataset that ought and yet fails to conform to BL, it can be helpful to stratify the data in an attempt to identify the source of nonconformity. Accordingly, we consider separately the EER claims contained in the PDDs for China and India. The χ^2, K-S and Kuiper tests for $d1$, $d2$ and $d1d2$ are again reported in Table 4.

For the EERs contained in the PDDs of CDM GHG emission reduction projects in China, both the graphs and statistical tests once more point to nonconformity with BL. Figure 1 reveals excessive frequencies for $d1$ = 1, 8 and 9. For $d2$, the digits 0, 1 and 2 appear more frequently than expected whilst 3–9 appear less frequently. For $d1d2$, excess frequencies are observed for 10, 11, 12, 35, 41, 46, 78 and 81–99.

By contrast, the graphs of $d1$, $d2$ and $d1d2$ for India (see Fig. 2) show a remarkably good fit. The χ^2, K-S and Kuiper tests also fail to reject the null hypothesis that the data conform to BL (apart from the χ^2 test for $d1$ which is ambiguous: significant at the 5% but not at the 1% level of confidence). These results are not a consequence of the fact that there are more

Table 4 Statistical analysis of expected and certified emission reductions

	Country	*N*	Digit(s)	χ^2 test	K-S test	Kuiper test	K-L statistic
Expected emission reductions	All countries	12675	*d*1	123.833**	0.025**	2.89**	0.00463
			*d*2	42.679**	0.021**	2.46**	0.00165
			*d*1*d*2	354.234**	0.026**	5.63**	0.01356
	China	5010	*d*1	482.492**	0.082**	7.53**	0.04133
			*d*2	103.336**	0.061**	4.32**	0.00988
			*d*1*d*2	993.413**	0.083**	11.27**	0.08571
	India	3338	*d*1	18.455*	0.016	1.56	0.00282
			*d*2	4.132	0.008	0.55	0.00061
			*d*1*d*2	93.871	0.017	1.76*	0.01413
Of which registered projects	All countries	8035	*d*1	228.099**	0.046**	4.53**	0.01283
			*d*2	58.686**	0.033**	2.97**	0.00357
			*d*1*d*2	513.353**	0.046**	7.63**	0.02963
	China	3806	*d*1	604.124**	0.105**	8.52**	0.06613
			*d*2	134.971**	0.074**	4.61**	0.01672
			*d*1*d*2	1182.408**	0.107**	12.32**	0.12945
	India	1660	*d*1	13.895	0.012	0.87	0.00410
			*d*2	9.827	0.012	0.93	0.00295
			*d*1*d*2	82.084	0.013	0.92	0.02527
Certified emission reductions	All countries	2854	*d*1	9.523	0.020	1.10	0.00167
			*d*2	4.208	0.009	0.52	0.00074
			*d*1*d*2	72.568	0.023	1.26	0.01344
	China	1451	*d*1	17.834*	0.027	1.07	0.00623
			*d*2	3.573	0.010	0.50	0.00123
			*d*1*d*2	103.381	0.028	1.44	0.03631

Reported are the statistics of the χ^2, K-S and Kuiper tests and K-L statistic for *d*1, *d*2 and *d*1*d*2 of EERs for CDM projects and those registered, including projects hosted in all countries, as well as China and India. Statistics for CERs of all countries and those hosted in China are also reported. '**' and '*' denote a significant difference from the Benford distribution at 1% and 5% level of significance respectively

observations for China than India because the K-L test confirms that much more information is lost when the Benford distribution is used to approximate the observed distribution for China than for India. Similar results are obtained even if the sample is restricted to those projects whose website status indicates that they are registered.

We now analyse the CERs issued to each project. These have been subject to additional auditing. Beginning once more with the full sample, the visual conformity of *d*1, *d*2 and *d*1*d*2 with BL is now excellent (see Fig. 3). The χ^2, K-S and Kuiper tests contained in Table 2 confirm the distributions of *d*1, *d*2 and *d*1*d*2 possess the expected frequencies.

Given that the digital frequency analysis of EERs found in the PDDs of projects in China does not correspond to BL, we repeat the analysis for CERs for projects in China. Now the results are quite different: although the χ^2 test for *d*1 is ambiguous, the χ^2 tests for *d*2 and *d*1*d*2 are not significant, and neither the K-S nor the Kuiper tests are significant either.

8 Discussion

Before investigating the source of nonconformity with BL for the case of Chinese EERs, we explore whether our inability to reject the null of conformity with BL in the case of India

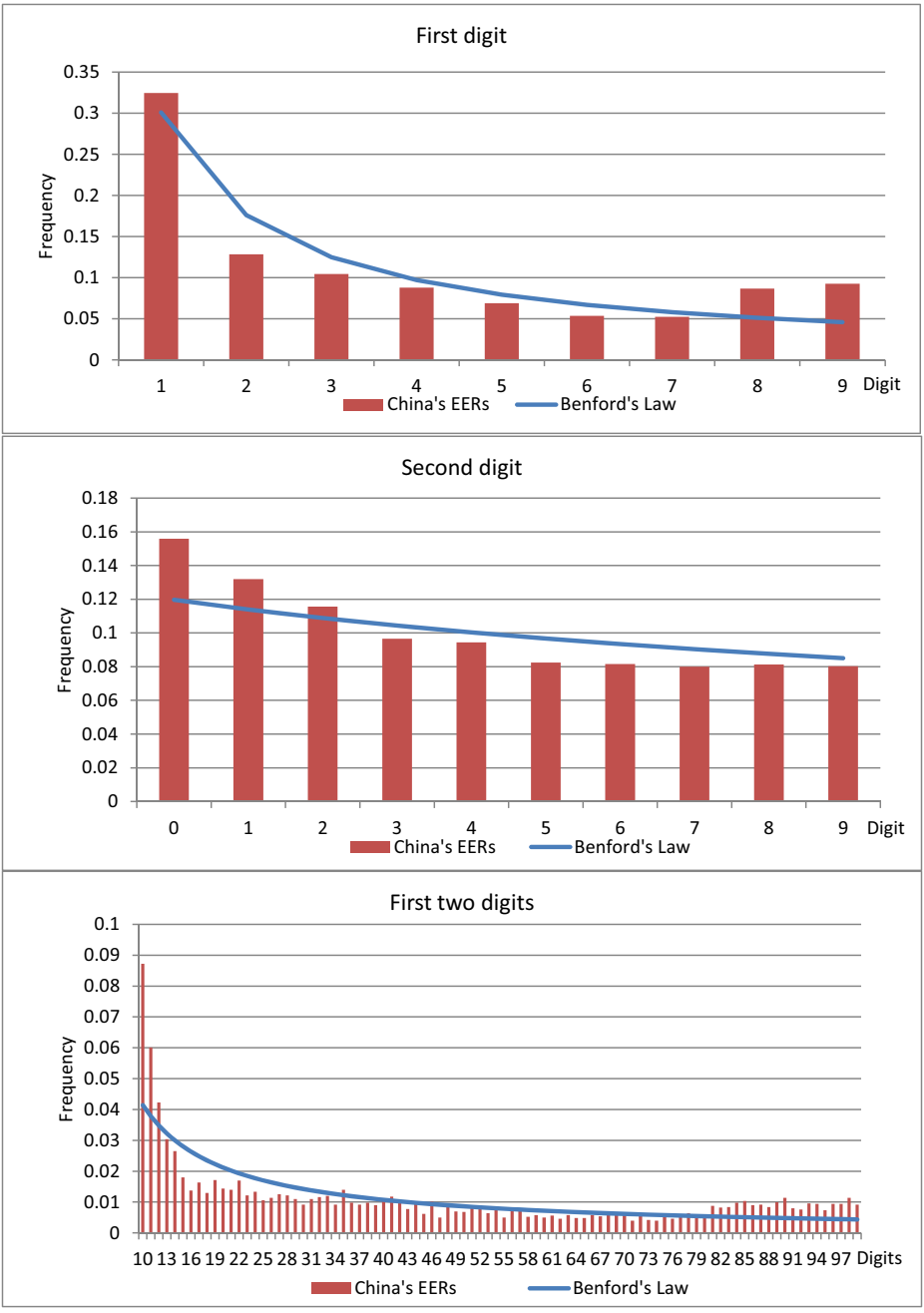

Fig. 1 Digital frequency analysis of expected emission reductions for projects in China. Plotted are the frequencies of each digit for $d1$, $d2$ and $d1d2$ for EERs of CDM projects hosted in China as well as the corresponding expected frequencies from Benford's Law

happens for other countries too. We use the K-L statistic to compare the loss of information from using BL to approximate the distribution of digits observed for these countries.

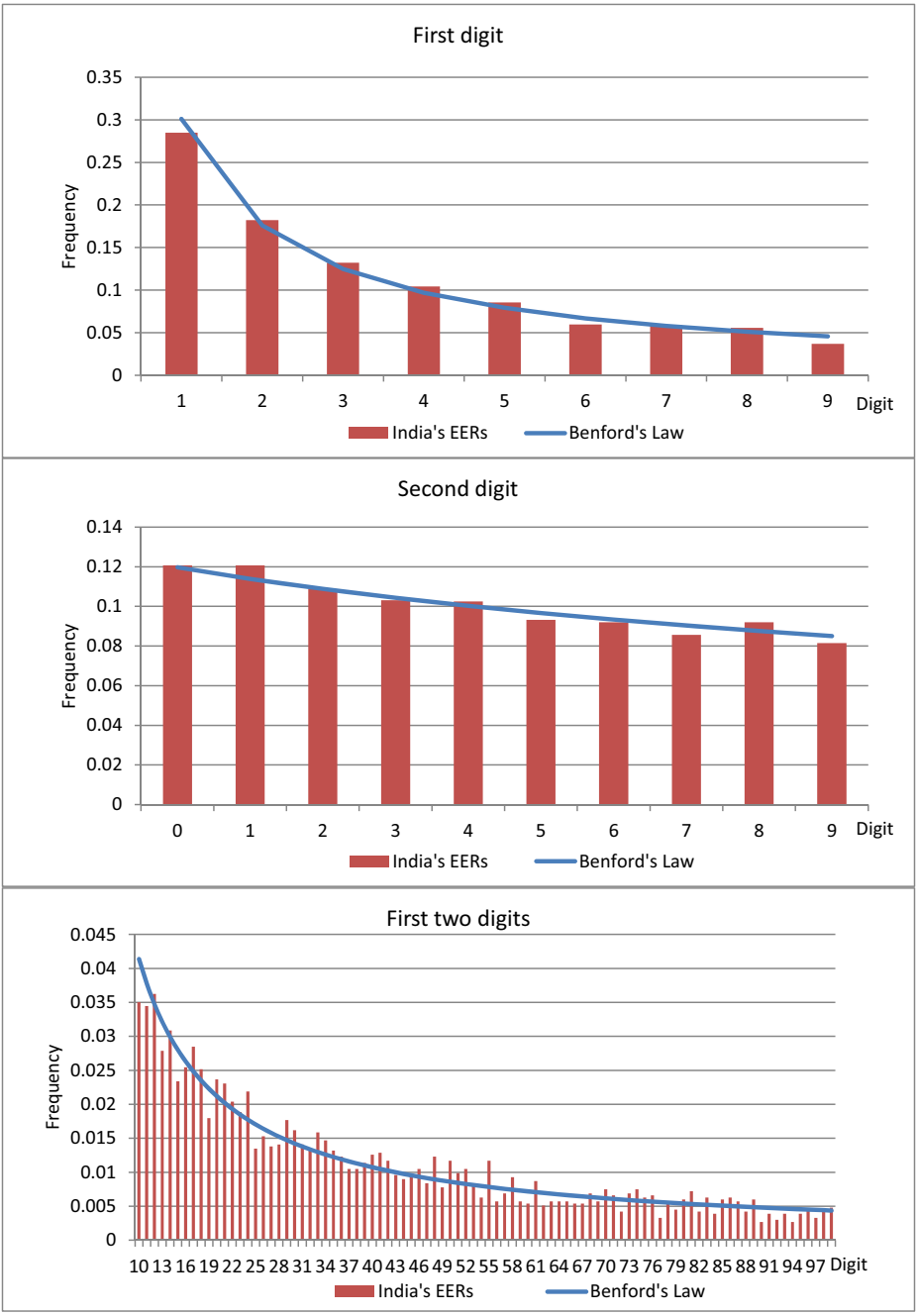

Fig. 2 Digital frequency analysis of emission reduction claims for projects in India. Plotted are the frequencies of each digit for $d1$, $d2$ and $d1d2$ for EERs of CDM projects hosted in India as well as the corresponding expected frequencies from Benford's Law

Examining EERs for CDM projects in Brazil and Mexico, we are unable to reject the null hypothesis that the distributions of $d1$, $d2$ and $d1d2$ conform to BL although in two instances

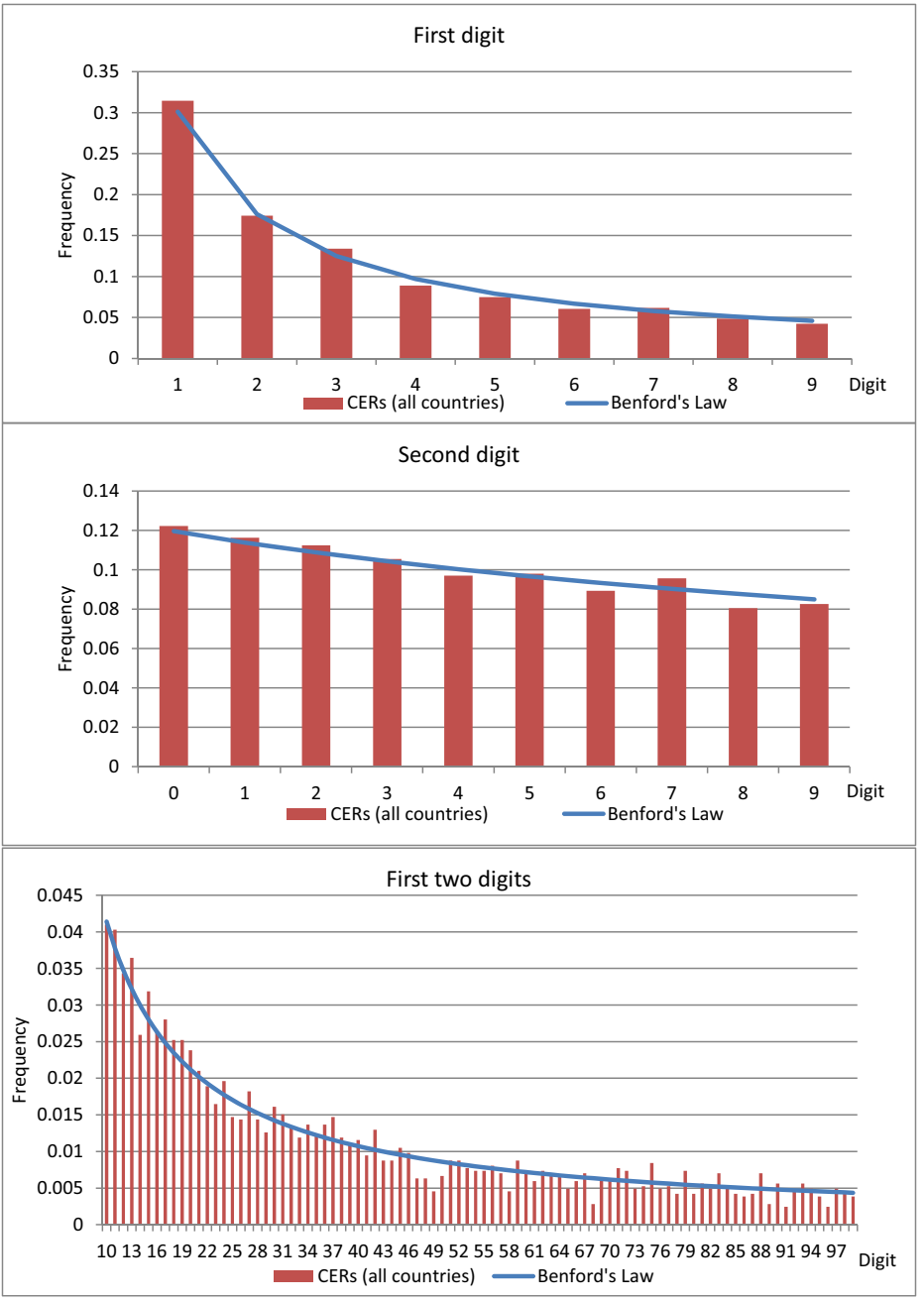

Fig. 3 Digital frequency analysis of certified emission reductions (all countries). Plotted are the frequencies of each digit for $d1$, $d2$ and $d1d2$ for CERs of CDM projects as well as the corresponding expected frequencies from Benford's Law

the test is ambiguous (see Table 5). The K-L statistics are lower than for China in the case of $d1$ but in the case of $d2$, the statistic for Mexico is higher. Given the small number of projects,

some $d1d2$ combinations are missing e.g. there are no projects in Brazil beginning with the digit combination 74.

To determine whether nonconformity in Chinese EERs is caused by projects of a particular type, we drop major project categories, one by one. Dropping hydro projects, however, fails to resolve the problem of nonconformity with BL with all tests continuing to reject the null hypothesis that the distribution of digits conforms to BL at the 1% level of confidence.

Dropping wind projects by contrast changes things. Now, whether the null hypothesis is rejected depends on the specific test and the digit(s) under consideration. For $d1$, although the χ^2 and Kuiper tests are significant at the 1% level of confidence, the K-S test is significant only at the 5% level of confidence. None of the tests is significant even at the 5% level of confidence for $d2$. For $d1d2$, the χ^2 and Kuiper tests are significant at 1% level of confidence, but the K-S test is significant only at the 5% level of confidence. More importantly, the K-L statistics are lower: 0.005 for $d1$ for wind compared with 0.041 for all projects.

Based on these results, we investigate whether the distribution of digits for EERs for wind projects in China conforms to BL. Here all the tests reject the null of conformity at the 1% level of significance. By contrast, for EERs for wind projects in India, none of the tests rejects the null of conformity to BL, even at the 5% level of confidence. The K-L statistics too are much higher for China.

Table 5 Further results

	Country	N	Digit(s)	χ^2 test	K-S test	Kuiper test	K-L statistic
EERs	Brazil	748	$d1$	15.541*	0.035	1.619	0.010
			$d2$	12.375	0.027	0.752	0.008
			$d1d2$	95.236	0.042	1.866*	NA
	Mexico	324	$d1$	10.815	0.038	1.039	0.017
			$d2$	10.858	0.045	0.977	0.017
			$d1d2$	94.206	0.048	1.318	NA
	China excluding hydro	3275	$d1$	998.164**	0.134**	12.681**	0.131
			$d2$	182.448**	0.101**	5.811**	0.026
			$d1d2$	1820.464**	0.144**	16.148**	0.226
	China excluding wind	3397	$d1$	40.180**	0.025*	2.687**	0.005
			$d2$	6.059	0.007	0.914	0.000
			$d1d2$	156.82**	0.026*	3.010**	0.023
	China wind	1613	$d1$	1928.362**	0.284**	16.213**	0.481
			$d2$	352.843**	0.192**	7.742**	0.096
			$d1d2$	3437.143**	0.296**	22.541**	NA
	India wind	1089	$d1$	12.263	0.033	1.108	0.005
			$d2$	8.575	0.013	0.597	0.003
			$d1d2$	90.121	0.035	1.621	0.040
CERs	China wind	547	$d1$	10.833	0.050	1.203	0.010
			$d2$	9.217	0.020	0.902	0.008
			$d1d2$	89.381	0.056	1.687	0.085
EERs	China wind (if CERs issued)	547	$d1$	356.551**	0.260**	8.927**	0.321
			$d2$	190.133**	0.271**	6.403**	0.167
			$d1d2$	1023.707**	0.384**	11.843**	NA

Reported are the statistics of the χ^2, K-S and Kuiper tests and K-L statistic for $d1$, $d2$ and $d1d2$ of EERs of CDM projects hosted in Brazil and Mexico, as well as projects hosted in China excluding hydro and wind, and only wind projects in China and India. The same statistics are reported for CERs of China's wind projects and the EERs of wind projects that went on to earn CERs. '**' and '*' denote a significant difference from the Benford distribution at 1% and 5% level of significance respectively

Analysing the distribution of digits for CERs for 547 wind projects in China, we are unable to reject the null of conformity with BL using any test and moreover, the K-L statistics point to only a minor loss of information from using BL to approximate the distribution. By contrast, the distribution of digits for EERs for those exact same 547 wind projects that went on to earn CERs continues to display nonconformity with BL. All the tests are significant at the 1% level of confidence and the K-L tests indicate a significant loss of information from using BL as an approximation. Even if EERs are measured in terms of annual emission reductions and CERs are in terms of cumulative emission reductions, it is hard to reconcile these findings because of the scale invariance property of Benford sets.

With data manipulation, there is usually a plausible direction of bias. Along with evidence that the distribution of digits does not conform to BL, this provides another opportunity to assess whether data might have been manipulated. The Distortion Factor (DF) model was developed by Nigrini (1996) to measure the bias in US income tax data. This is achieved by comparing the mean of the reported numbers after they have been 'collapsed' (so that they all fall in the range 10–99) with the mean of the numbers in a Benford set. The DF model rests on a number of assumptions. First, those who manipulate data do so in a way which does not alter the order of magnitude (this is deemed too suspicious). Second, the percentage change caused by data manipulation is on average the same across all orders of magnitude. The expected mean (EM) of the collapsed dataset is as follows:

$$EM = \frac{90}{N\left(10^{1/N}-1\right)}$$

And the DF is given by:

$$DF = \frac{100(AM-EM)}{EM}$$

However, whilst there is fairly obviously underreporting of taxable income, with a baseline-and-credit system such as the CDM, the incentive is to exaggerate (although in the case of CDM projects, the dependency of registration fees on EERs might point in the opposite direction). For the full sample of EER claims in PDDs, the DF test displayed in Table 7 shows that reported numbers are indeed 2.92% above those normally found in a Benford set whereas for China, they are on average 8.58% higher. Both distortion factors are, moreover, significant at the 1% level of significance (Table 6).

Table 6 Distortion factor analysis

	Country	*N*	Observed mean	Expected mean	Distortion factor	Std. Dev.	*Z*-statistic
Expected emission reductions	All	12675	40.223	39.082	0.029	0.005	5.146**
	China	5010	42.431	39.077	0.085	0.009	9.517**
Of which registered projects	All	8035	41.023	39.080	0.049	0.007	6.979**
	China	3806	43.406	39.074	0.110	0.010	10.714**

Reported are the results of the distortion factor analysis for EERs of all and registered CDM projects. Projects hosted in China are separately identified. The null hypothesis is that the observed mean of the collapsed data is equal to the expected mean for that of a Benford's set. '**' and '*' denote statistical significance at 1% and 5% respectively

Table 7 The effect of deliberate data manipulation on tests for nonconformity: The case of Indian expected emission reductions

Percentage of data manipulated	Digit(s)	χ^2 test	K-S test	Kuiper test	K-L statistic
1	*d*1	19.854*	0.019	1.724	0.00301
	*d*2	4.409	0.008	0.572	0.00065
	*d*1*d*2	103.252	0.019	1.863	0.01552
5	*d*1	26.076**	0.032**	2.262**	0.00394
	*d*2	4.066	0.005	0.546	0.00060
	*d*1*d*2	117.564*	0.032**	2.401**	0.01758
10	*d*1	48.555**	0.049**	3.026**	0.00741
	*d*2	7.831	0.0121	0.702	0.00118
	*d*1*d*2	130.721**	0.050**	3.195**	0.01968
25	*d*1	147.510**	0.092**	5.342**	0.02336
	*d*2	34.871**	0.026*	1.535	0.00534
	*d*1*d*2	264.180**	0.095**	5.546**	0.03986

Reported are the statistics of the χ^2, K-S and Kuiper tests and K-L statistic for *d*1, *d*2 and *d*1*d2* of EERs for CDM projects hosted in India. A percentage of these data have been contaminated by adding 1 to every digit apart from when the digit is 9. '**' and '*' denote a significant difference from the Benford distribution at 1% and 5% level of significance respectively

We cannot be certain why EER claims contained in the PDDs of projects in China do not conform to BL or why, according to the DF test, the EER claims appear to have been inflated. Our findings are, however, not inconsistent with the possibility that data for some Chinese CDM projects, particularly those involving wind, might have been manipulated. For concerns about the integrity of Chinese data, see Hsu et al. (2012), Ghanem and Zhang (2014), Zheng et al. (2014), Liang et al. (2016), Morris and Zhang (2017), Stoerk (2016) and Brombal (2017). Nevertheless, it is important to stress that, even if the reason EERs do not conform to BL is data manipulation, attempts at data manipulation have not survived the full auditing process: data on CERs conforms to BL. In addition, we cannot rule out the possibility that, rather than manipulation, data from the PDDs of Chinese projects might at some point have been mishandled. Also possible is that statistical processes resulting in conformity with BL do not describe the processes generating EER claims contained in Chinese PDDs (although the ones from India and two other host countries do).

Finally, we investigate the sensitivity of tests used to detect data manipulation by randomly replacing a percentage of observations with data that we have deliberately manipulated. More specifically, data for EERs for India are altered by adding 1 to each digit unless the digit is already 9 in which case it is left unaltered. This obviously inflates expected emission reductions. For example, with this heuristic, 4,486,341 becomes 5,597,452 and 64,996 becomes 75,997. We then examine how the tests for nonconformity with BL respond as the percentage of manipulated observations increases.

The results of this admittedly simple experiment are presented in the Table 7. They indicate that as the percentage of data that is manipulated increases, the tests against the null of conformity to BL quickly begin to show statistical significance. For example, when 5% of the data are manipulated, the χ^2 test for *d*1 becomes statistically significant at the 1% level of confidence along with the K-S and Kuiper tests.

9 Conclusion

This paper examines the integrity of the emission reduction claims of CDM projects by subjecting EERs and CERs to digital frequency analysis. We find that EERs do not always

conform to BL, specifically those from China which distortion factor analysis suggests might have been inflated. Our findings are therefore not inconsistent with the possibility that the EERs in Chinese CDM projects, and particularly those involving wind energy, have been manipulated. Interestingly, however, we cannot reject the null hypothesis that the distribution of CERs does conform to BL, implying that the full CDM auditing process is effective.

Given the prevalence of self-reporting emissions, the growing use of regulatory systems that incentivise the manipulation of emission data and the high resource costs of environmental auditing, we believe that digital frequency analysis, which is both rapid and low-cost, has an important role to play in the analysis of self-reported GHG emissions and emission reductions. BL can improve the chance of detecting data manipulation compared with random auditing. We also perceive a role for employing the same techniques to analyse countries' self-reported GHG emissions. We suggest that emission (and emission reductions) data for GHGs whether reported by firms or governments be routinely screened for nonconformity with BL whilst recognising that this is only a first step and will never supplant the need for environmental auditing.

One interesting case study might be to use BL to analyse data from carbon trading schemes. The number of facilities involved in the EUETS for example, is extremely large and includes many different countries across which regulation is perhaps, unevenly applied. The pressure on those entities liable to the EUETS to cheat is intensified by competition from unregulated entities outside the EU. Furthermore, although with more than 14,000 installations, the EUETS carbon trading scheme is the largest, there are numerous other carbon trading schemes in operation. BL might also be a tool with which the relevant authorities could scrutinise data from the Chinese emission trading scheme, which is expected to launch soon.

Finally, we need further methodological advances in the use of BL. We need to find ways to demonstrate that departures from BL really do indeed point to data manipulation or mishandling rather than something else. One way to do this is through comparing conformity with BL before and after some policy change influencing firms' or even governments' decision whether to misreport data. Another important task will be to establish criteria for the use of BL e.g. how many orders of magnitude should be present in the data in order to expect it to conform to BL?

Acknowledgements We would like to thank three anonymous referees and Richard Tol, Heinz Welsch and Tong Zhu for comments on an earlier draft.

References

Anagnostopoulos K, Flamos A, Kagiannas AG, Psarras J (2004) The impact of Clean Development Mechanism in achieving sustainable development. Int J Environ Pollut 21(1):1–23

Aslam MA (2001) Technology transfer under the CDM-materializing the myth in the Japanese context? Clim Pol 1:451–464

Beiglou PHB, Gibbs C, Rivers L, Adhikari U, Mitchell J (2017) Applicability of Benford's Law to compliance assessment of self- reported wastewater treatment plant discharge data. Journal of Environmental Assessment Policy and Management 19(04):1750017

Bellassen V, Stephan N, Afriat M, Alberola E, Barker A, Chang J-P, Chiquet C, Cochran I, Deheza M, Dimopoulos C, Foucherot C, Jacquier G, Morel R, Robinson R, Shishlov I (2015) Monitoring, reporting and verifying emissions in the climate economy. Nat Clim Chang 5(4):319–328
Benford F (1938) The law of anomalous numbers. Proc Am Philos Soc 78(4):551–572
Boyle J (1994) An application of fourier series to the most significant digit problem. Am Math Mon 101(9):879–886
Brombal D (2017) Accuracy of environmental monitoring in China: exploring the influence of institutional, political and ideological factors. Sustainability 9(3):324
Brown RJ (2005) Benford's Law and the screening of analytical data: the case of pollutant concentrations in ambient air. Analyst 130(9):1280–1285
Bumpus AG, Cole JC (2010) How can the current CDM deliver sustainable development? Wiley Interdiscip Rev Clim Chang 1(4):541–547
Collins JC (2017) Using Excel and Benford's Law to detect fraud: learn the formula's functions and techniques that enable efficient Benford analysis of data sets. J Account 223(4):44–50
Cormier A, Bellassen V (2013) The risks of CDM projects: how did only 30% of the expected credits come through? Energy Policy 54:173–183
Crowe TL (2013) The potential of the CDM to deliver pro-poor benefits. Clim Pol 13(1):58–79
de Marchi S, Hamilton JT (2006) Assessing the accuracy of self-reported data: an evaluation of the toxics release inventory. J Risk Uncertain 32(1):57–76
Dechezleprêtre A, Glachant M, Ménière Y (2008) The Clean Development Mechanism and the international diffusion of technologies: an empirical study. Energy Policy 36(4):1273–1283
Dechezleprêtre A, Glachant M, Ménière Y (2009) Technology transfer by CDM projects: a comparison of Brazil, China, India and Mexico. Energy Policy 37(2):703–711
Dlugosz S, Müller-Funk U (2009) The value of the last digit: statistical fraud detection with digit analysis. Adv Data Anal Class 3(3):281–290
Drew JM, Drew ME (2010) Establishing additionality: fraud vulnerabilities in the Clean Development Mechanism. Account Res J 23(3):243–253
Dumas CF, Devine JH (2002) Detecting evidence of non compliance in self-reported pollution emissions data: an application of Benford's Law. In American Agricultural Economics Association Annual Meeting, Tampa, FL. Available at http://agecon.lib.umn.edu
Durtschi C, Hillison W, Pacini C (2004) The effective use of Benford's law to assist in detecting fraud in accounting data. J Forensic Account 5(1):17–34
Fay J, Kapfudzaruwa F, Matheson S (2011) A comparative policy analysis of the Clean Development Mechanism in South Africa and China. Climate and Development 4(1):40–53
Fewster RM (2009) A simple explanation of Benford's Law. Am Stat 63(1):26–32
Fu D, Shi YQ, Su W (2007) A generalized Benford's law for JPEG coefficients and its applications in image forensics. In: Electronic Imaging 2007. International Society for Optics and Photonics, p 65051L
Fu Q, Fang Z, Villas-Boas SB, Judge G (2014) An investigation of the quality of air data in Beijing. Available at http://are.berkeley.edu/~sberto/BeijingJuly16.pdf
Ghanem D, Zhang J (2014) 'Effortless perfection:' do Chinese cities manipulate air pollution data? J Environ Econ Manag 68(2):203–225
Giles DE (2007) Benford's law and naturally occurring prices in certain ebaY auctions. Appl Econ Lett 14(3):157–161
Gundimeda H (2004) How 'sustainable' is the 'sustainable development objective' of CDM in developing countries like India? Forest Policy Econ 6(3):329–343
Haites E, Duan MS, Seres S (2006) Technology transfer by CDM projects. Clim Pol 6(3):327–344
Hill TP (1995) Base-invariance implies Benford's Law. Proc Am Math Soc 123(3):887–895
Hill TP (1998) The first digit phenomenon. Am Sci 86:358–363
Hsu A, de Sherbinin A, Shi H (2012) Seeking truth from facts: the challenge of environmental indicator development in China. Environ Dev 3:39–51
Joannes-Boyau R, Bodin T, Scheffers A, Sambridge M, May SM (2015) Using Benford's law to investigate natural hazard dataset homogeneity. Sci Rep 5
Judge G, Schechter L (2009) Detecting problems in survey data using Benford's Law. J Hum Resour 44(1):1–24
Karakosta C, Doukas H, Psarras J (2009) Directing Clean Development Mechanism towards developing countries' sustainable development priorities. Energy Sustain Dev 13(2):77–84
Krey M (2005) Transactions costs of unilateral CDM projects in India – results from an empirical survey. Energy Policy 33:2385–2397
Lecocq F, Ambrosi P (2007) The Clean Development Mechanism: history, status, and prospects. Rev Environ Econ Policy 1(1):134–151
Lemons DS (1986) On the numbers of things and the distribution of first digits. Am J Phys 54:816–817

Liang X, Li S, Zhang S, Huang H, Chen SX (2016) PM2.5 data reliability, consistency, and air quality assessment in five Chinese cities. J Geophys Res-Atmos 121(17):10220–10236
Moret MA, de Senna V, Santana MC, Zebende GF (2009) Geometric structural aspects of proteins and Newcomb-Benford Law. Int J Mod Phys C 20:1981–1988
Morris SD and Zhang J (2017) Validating China's output data using satellite observations. Available at SSRN: https://ssrn.com/abstract=2920650
Newcomb S (1881) Note on the frequency of use of the different digits in natural numbers. Am J Math 4(1):39–40
Nigrini MJ (1996) A taxpayer compliance application of Benford's law. J Am Taxat Assoc 18(1):72–91
Nigrini MJ (2012) Benford's Law: applications for forensic accounting, auditing, and fraud detection. Series: Wiley corporate F and A
Nigrini MJ, Miller SJ (2007) Benford's law applied to hydrology data–results and relevance to other geophysical data. Math Geol 39:469–490
Nussbaumer P (2009) On the contribution of labelled Certified Emission Reductions to sustainable development: a multi-criteria evaluation of CDM projects. Energy Policy 37(1):91–101
Olsen KH (2007) The Clean Development Mechanism's contribution to sustainable development: a review of the literature. Clim Chang 84(1):59–73
Paulsson E (2009) A review of the CDM literature: from fine-tuning to critical scrutiny? Int Environ Agreements: Polit Law Econ 9(1):63–80
Pearson B (2007) Market failure: why the Clean Development Mechanism won't promote clean development. J Clean Prod 15(2):247–252
Pietronero L, Tosatti E, Tosatti V, Vespignani A (2001) Explaining the uneven distribution of numbers in nature: the laws of Benford and Zipf. Phys A: Stat Mech Appl 293(1):297–304
Pinkham R (1961) On the distribution of first significant digits. Ann Math Stat 32:1223–1230
Raimi RA (1976) The first digit problem. Am Math Mon 8(7):521–538
Seres S, Haites E, Murphy K (2009) Analysis of technology transfer in CDM projects: An update. Energy Policy 37(11):4919–4926
Stoerk T (2016) Statistical corruption in Beijing's air quality data has likely ended in 2012. Atmos Environ 127:365–371
Sutter C, Parreno JC (2007) Does the current Clean Development Mechanism (CDM) deliver its sustainable development claim? An analysis of officially registered CDM projects. Clim Chang 84(1):74–90
Tam Cho WK, Gaines BJ (2007) Breaking the (Benford) law: statistical fraud detection in campaign finance. Am Stat 61(3):218–223
Warnecke C (2014) Can CDM monitoring requirements be reduced while maintaining environmental integrity. Clim Pol 14(4):443–466
Zahran S, Iverson T, Weiler S, Underwood A (2014) Evidence that the accuracy of self-reported lead emissions data improved: a puzzle and discussion. J Risk Uncertain 49(3):235–257
Zhang J, Wang C (2011) Co-benefits and additionality of the Clean Development Mechanism: an empirical analysis. J Environ Econ Manag 62(2):140–154
Zheng S, Kahn ME, Sun W, Luo D (2014) Incentives for China's urban mayors to mitigate pollution externalities: the role of the central government and public environmentalism. Reg Sci Urban Econ 47:61–71