BENFORD'S LAW AND FRAUD DETECTION

by
YANG LU

STEVEN J. MILLER, ADVISOR

A thesis submitted in partial fulfillment
of the requirements for the
Degree of Bachelor of Arts with Honors
in Mathematics

WILLIAMS COLLEGE
Williamstown, Massachusetts

May 14, 2014

ABSTRACT

Benford's Law describes the situation in which the frequency distribution of the first digits in a real-life data set does not follow a uniform distribution. Rather, the probability of a digit d occurring as the leading digit of a data point is the difference between $d+1$ and d on a logarithmic scale with base 10. This phenomenon has been regularly used by auditors as a tool to detect fraud. The bootstrap is a resampling method which can be used to estimate the sampling distribution of an estimator, first proposed by Efron in 1979 [Ef].

We apply the bootstrap method to find a way to reduce the number of data points required for effective fraud detection based on Benford's Law. Oftentimes an auditor may not have access to all the data; a method using a subset of the data for fraud detection makes the auditor's job possible and potentially saves both human and computational resources. We have found that assuming that a data set of size 5000 or more is either free of manipulation or a result of summing the original values and some values following either a normal distribution or a uniform distribution, we only require 5% of the data to detect potential data manipulation.

ACKNOWLEDGEMENTS

I am grateful to my thesis advisor Prof. Steven Miller for his invaluable guidance and inspiration throughout the year and at every stage of the thesis. I would also like to thank my second reader Prof. Klingenberg for his valuable comments.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. INTRODUCTION

Take a random data set consisting of numbers of any sort, the leading digit is defined by the first number which occurs in any data point. A question to ask is what is the probability of any of the digits from 1 through 9 occurring as the leading digit of a data point. The intuitive answer would be that it is equally likely for any of the nine non-zero digits to be the first digit of a number. However, this does not hold true in many cases. In fact, 1 occurring as the leading digit happens about 30% of the time; 2, 18% of the time. The probability of a digit d occurring as the leading digit is the difference between $d+1$ and d on a logarithmic scale with base 10. This phenomenon is known as the Benford's Law, first stated by Simon Newcomb in 1881 [Ne]. Scholars and practitioners have since studied various data sets and confirmed the phenomenon [CLTF, GM, NiMil]. They suggest that oftentimes the patterns of the first digits of authentic numbers should conform to Benford's Law. Therefore, accountants and auditors could compare the actual frequency distributions of the first digits with those predicted by Benford's Law and potentially determine the authenticity of a data set.

The bootstrap as a resampling method to estimate the sampling distribution of an estimator was first proposed by Efron in 1979 [Ef]. The method repeatedly resamples with replacement from one or several samples of a population and obtains estimates related to the population parameters.

In this paper, we study two related questions. The first is how different ways of manipulating data would have an effect on the frequency distribution of the leading digits and various deviation measures based on Benford's Law. The second is to explore whether there is a minimum number of data points required to detect fraudulent data when applying Benford's Law.

This is important in that fraudulent data, be it in accounting, tax reporting, or scientific experiments can misguide scholars and practitioners who rely on the data and results for further studying. Specifically, in the context of accounting and tax frauds, it would significantly harm investors in denying them the rights to be fully informed about the companies. It may also hurt the government's tax revenues. An efficient way to detect accounting frauds (either manipulating existing numbers or falsifying entries) will potentially help both the investors and the Internal Revenue Service (IRS). Oftentimes, an auditor may not have access to all the data; a method which utilizes a subset of the data for fraud detection makes the auditor's job possible and potentially saves both human and computational resources.

There are many ways for one to manipulate a data set. We study two major ways of manipulating a data set. One way is to sum the original data and a sequence of numbers which follows a certain distribution; the other is to replace a proportion of the original data with values following another distribution. We can compare the actual frequency distribution of the first digits with that predicted by Benford's Law, and then apply the bootstrap method to examine whether there is a minimum number of data points required to detect a non-conformity to Benford's Law.

In this paper we have found that assuming that a data set of size 5000 or more is either free of manipulation or a result of summing the original values and values following either a normal distribution or a uniform distribution, we only require 5% of the data to detect potential manipulation. However, it is more difficult to detect partial replacement of the original data with values following either a normal distribution or a uniform distribution. For a data set of size 5000 partially replaced with values following a normal distribution, only when the proportion of replacement reaches 0.9, we can use 9% of the data set to effectively detect non-conformity to Benford's Law. When the size of the data set becomes 10000, bootstrap allows us to use 10% of the data to detect any potential non-conformity to Benford's Law insofar as the proportion of replacement is 0.8 or larger. Assuming partial replacement of data with values following a uniform distribution, when the proportion of replacement is 0.4 or larger, we can apply the bootstrap method with 5% of the data to detect non-Benfordness.

Section 2 reviews some theoretical background on Benford's Law. In Section 3, we present theoretical calculations for both the convolution of a Benford's density and either a normal or a uniform density, and partial replacement of data by new data following either a normal or a uniform distribution. Section 4 empirically explores the relations among the deviation from Benford's Law, the parameterization of a normal or a uniform distribution, and the sizes of the data sets. We then investigate how bootstrap could potentially help identify deviations from Benford's Law. Section 5 provides background on bootstrap and reports the simulation results when we apply a bootstrap method for fraud detection. Section 6 performs an empirical analysis on financial filings from two companies and investigates the effectiveness of the bootstrap method on small data sets. Section 7 concludes.

2. BACKGROUND ON BENFORD'S LAW

This section provides some background on convolution, Benford's Law, and measures of conformity to Benford's Law.

2.1. **Sum of two random variables.**

One way a person can manipulate data is to add new values to the original data. The values added may follow a certain distribution such as a normal distribution or a uniform distribution. The rationale behind adding values which follow a normal distribution is that forgers may come up with a sequence of numbers with a bell-shaped distribution. Alternatively, the values added may follow a uniform distribution if a person is naively adding random numbers within a range of values, each of which has an equal probability of being chosen.

The operation of summing two or more densities is called convolution. In other words, convolution forms a linear combination of independent random variables. We can either work directly with probability density functions or use the inverse Fourier transform of the probability density functions to obtain analytical results of a convolution.

Let $Z = X + Y$ where X and Y are two random variables following certain distributions. The density function $h(x)$ of the random variable Z is the convolution of the two density functions $f(x)$ and $g(x)$ as shown by

$$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(x-\tau)g(\tau)d\tau, \tag{2.1}$$

where $f(x)$ and $g(x)$ are the density functions of the random variables X and Y.

Alternatively, we can apply the Fourier transform to calculate the convolution. Let F be the Fourier tranform operation. The Fourier transform of the density function $f(x)$ is given by $\hat{f}(t)$,

$$\hat{f}(t) = F(f(x)) = \int_{-\infty}^{\infty} f(x)e^{-2\pi itx}\mathrm{d}x, \tag{2.2}$$

and the inverse Fourier transform F^{-1} is

$$F^{-1}(f(x)) = \int_{-\infty}^{\infty} f(x)e^{2\pi itx}\mathrm{d}x. \tag{2.3}$$

The density function $h(x)$ can be calculated by the inverse Fourier transform as shown below

$$h(x) = (f * g)(x) = F^{-1}[F(f(x))F(g(x))] = F^{-1}[\hat{f}(t)\hat{g}(t)]. \tag{2.4}$$

2.2. **Benford's Law for the first digit.** Benford's Law for the first digit refers to the phenomenon in which the probability of a first digit d with base 10 occurring in a data set is $\log_{10}(1+\frac{1}{d})$. Yet in any real data set, it is impossible to observe that the exact frequency distribution of the first digits follows Benford's Law because d is an integer and $\log_{10}(1+\frac{1}{d})$ is irrational. The probability of a first digit d occurs in a finite data set cannot be irrational.

2.3. **Measures of deviation from the Benford's distribution.** There are various measures to gauge the extent to which the frequency distribution of the first digits of a data set deviates from the theoretical Benford's distribution.

One way to examine the amount of deviation is to use the Pearson's Chi-Square test since we are examining the goodness of fit of categorical data.

The Chi-Square test statistic can be calculated as follows:

$$\chi^2 = \sum_{d=1}^{9} \frac{(p(d) - f(d))^2}{f(d)}, \tag{2.5}$$

where $p(d)$ is the frequency distribution of the first digit d from the data and $f(d)$ is the frequency distribution of the first digit d based on Benford's Law.

We can obtain the p-value from the Chi-Square test with 8 degrees of freedom. The Chi-Square test statistic to obtain a 95 percentile at 8 degrees of freem is 15.5 and that to obtain a 99 percentile is 20.1. The null hypothesis is that the frequency distribution of the first digits based on the data set follows Benford's Law. Given any arbitrary level of significance α, if the p-value is smaller than the significance level α, we can reject the null hypothesis that the frequency distribution of the first digits based on the data set follows Benford's Law [Pi].

Another way to measure the extent of deviation is the Mean Absolute Deviation (MAD) which is the average absolute deviation of the frequency distribution of each non-zero digit as the first digit from the Benford's distribution [DN]. It follows the formula

$$\mathrm{MAD} = \frac{1}{9}\sum_{d=1}^{9} |p(d) - f(d)|, \tag{2.6}$$

where $p(d)$ is the frequency distribution of the first digit d from the data and $f(d)$ is the frequency distribution of the first digit d based on Benford's Law.

Drake and Nigrini [DN] provide the following guidelines for using MAD to test the goodness-of-fit of the first digits based on prior experience.

- **Close conformity** 0.000 - 0.004
- **Acceptable conformity** 0.004 - 0.008

- **Marginally acceptable conformity** 0.008 - 0.012
- **Non-conformity** Greater than 0.012

3. THEORY ON BENFORD'S LAW

We explore two different scenarios in which people could manipulate data and how the resulting frequency distributions of the first digits deviate from Benford's Law. We assume that the original data values follow a Benford's distribution with base 10 within the interval $[0.1, 1)$. In Section 3.1 and 3.2 we sum the original values and new values following a distribution other than a Benford's distribution; in Section 3.3 and 3.4, we replace a proportion of the original data values with values following another distribution. For each scenario, we include two cases (a normal distribution and a uniform distribution) to examine the extent to which each mixed distribution deviates from Benford's Law.

3.1. Sum of a Benford random variable and a normal random variable.

Let X and Y be two random variables following a normal distribution and a Benford's distribution, respectively. We have the density function $f(x)$ of the normal distribution with mean μ_x and standard deviation σ_x,

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{\frac{-(x-\mu_x)^2}{2\sigma_x^2}}, \tag{3.1}$$

and the density function $g(x)$ of the Benford's distribution with base 10 bounded by $[0.1, 1)$[1]

$$g(x) = \frac{1}{x\ln(10)}. \tag{3.2}$$

Let $Z = X + Y$, the density function $h(x)$ of the random variable Z is thus the convolution of the two density functions $f(x)$ and $g(x)$

$$h(x) = (f * g)(x) = \int_{-\infty}^{\infty} f(x-\tau)g(\tau)\mathrm{d}\tau. \tag{3.3}$$

Applying the Fourier transform to the convolution, we have the same equation as Equation (2.4)

$$h(x) = (f * g)(x) = F^{-1}[F(f(x))F(g(x))] = F^{-1}[\hat{f}(t)\hat{g}(t)], \tag{3.4}$$

and then solve the Fourier transforms of $f(x)$ and $g(x)$, respectively. For the Fourier transform of $f(x)$, we substitute $x' = x - \mu_x$ and use the Euler's formula $e^{-i\theta} =$

---

[1] We do not use the range from 1 to 10 for the Benford's distribution as the magnitude of possible values gets large and as a result, Mathematica is unable to numerically integrate the probability density functions required in later part of this section.

$\cos\theta - i\sin\theta$.

$$\begin{aligned}\hat{f}(t) &= F(f(x)) \\ &= \frac{1}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma_x^2}} e^{-2\pi i t(x'+\mu_x)}\mathrm{d}x' \\ &= \frac{e^{-2\pi i t\mu_x}}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma_x^2}} e^{-2\pi i t x'}\mathrm{d}x' \\ &= \frac{e^{-2\pi i t\mu_x}}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma_x^2}} [\cos(2\pi t x') - i\sin(2\pi t x')]\mathrm{d}x' \\ &= \frac{e^{-2\pi i t\mu_x}}{\sqrt{2\pi}\sigma_x}\int_{-\infty}^{\infty} e^{-\frac{x'^2}{2\sigma_x^2}} \cos(2\pi t x')\mathrm{d}x'. \end{aligned} \tag{3.5}$$

Since we know that

$$\int_0^{\infty} e^{-aT^2}\cos(2xT)dT = \frac{1}{2}\sqrt{\frac{\pi}{a}}e^{-\frac{x^2}{a}}, \tag{3.6}$$

let $x = \pi t$, $T = x'$, and $a = \frac{1}{2\sigma_x^2}$, we have

$$\hat{f}(t) = F(f(x)) = e^{-2\pi i t\mu_x} e^{-2\pi^2\sigma_x^2 t^2}. \tag{3.7}$$

For the Fourier transform of $g(x)$ within $[0.1, 1)$, we have

$$\begin{aligned}\hat{g}(t) &= F(g(x)) \\ &= \int_{0.1}^{1} g(x)e^{-2\pi i t x}\mathrm{d}x \\ &= \int_{0.1}^{1} \frac{1}{x\ln(10)}e^{-2\pi i t x}\mathrm{d}x \\ &= \int_{0.1}^{1} \frac{1}{x\ln(10)}[\cos(2\pi t x) - i\sin(2\pi t x)]\mathrm{d}x \\ &= \frac{-\mathrm{Ci}(0.2\pi t) + \mathrm{Ci}(2\pi t) + i(\mathrm{Si}(0.2\pi t) - \mathrm{Si}(2\pi t))}{\ln(10)}, \end{aligned} \tag{3.8}$$

where Ci is the cosine integral and Si is the sine integral.[2]

The inverse Fourier transform is thus:

$$\begin{aligned}F^{-1}[\hat{f}(t)\hat{g}(t)] &= \int_{-\infty}^{\infty} \frac{-\mathrm{Ci}(0.2\pi t) + \mathrm{Ci}(2\pi t) + i(\mathrm{Si}(0.2\pi t) - \mathrm{Si}(2\pi t))}{\ln(10)} * \\ &\quad e^{-2\pi i t\mu_f} e^{-2\pi^2\sigma_f^2 t^2} e^{2\pi i t x}\mathrm{d}t \\ &= \frac{1}{\ln(10)}\int_{-\infty}^{\infty} (-\mathrm{Ci}(0.2\pi t) + \mathrm{Ci}(2\pi t) + i(\mathrm{Si}(0.2\pi t) - \mathrm{Si}(2\pi t))) * \\ &\quad e^{-2\pi i t\mu_f + 2\pi i t x - 2\pi^2\sigma_f^2 t^2}\mathrm{d}t. \end{aligned} \tag{3.9}$$

---

[2] The last line of Equation (3.8) is obtained from Mathematica with the assumption that $t > 0$.

Since there is no closed-form solution to the inverse Fourier transform $F^{-1}[\hat{f}(t)\hat{g}(t)]$, we choose an appropriate bound for t to evaluate the integral numerically.

For illustrative purposes, we set the mean of the normal distribution to 0 and standard deviation to 0.1. We have the resulting probability density function as

$$\begin{aligned} F^{-1}[\hat{f}(t)\hat{g}(t)] &= \frac{1}{\ln(10)}\int_{-\infty}^{\infty} e^{2\pi itx-0.02\pi^2t^2} \\ &\quad *\Big[-\operatorname{Ci}(0.2\pi t)+\operatorname{Ci}(2\pi t)+ \\ &\quad i(\operatorname{Si}(0.2\pi t)-\operatorname{Si}(2\pi t))\Big]\mathrm{d}t. \end{aligned} \tag{3.10}$$

Since $f(t)$ is bounded by $(-\infty,\infty)$, $g(t)$ is bounded by $[0.1,1)$, and the standard deviation of the normal distribution is 0.1, we choose t to be bounded by $\pm$ 4 standard deviations of the normal distribution, that is $t\in(-0.3,1.4)$, to cover most of the values in the resulting distribution. The probability for a randomly generated number from a normal distribution to be 4 standard deviations away from the mean is around 0.01%.

The inverse Fourier transform is supposed to be the probability density function of the resulting distribution (the convolution of a Benford's density function bounded by $[0.1,1)$ and a normal density function). Hence, the integral is the probablity density function of a distribution and should return a value of 1 when evaluated between $(-\infty,\infty)$. As a sanity check, we let $x\in(-0.3,1.4)$ and numerically integrate the inverse Fourier transform; the integral does have a value very close to 1 from numerical integration in Mathematica.

Before calculating and presenting the frequency distribution of the first digits based on the convolution in which the normal distribution has mean 0 and standard deviation ranging from 0.1 to 1, we first explain our numerical integration procedure in detail:

(1) We set the range of t in Equation (3.9) as $(0.1-4\sigma,1+4\sigma)$ where σ is the standard deviation of the normal distribution in the convolution;

(2) We restrict x in Equation (3.9) to ranges which give the leading digits desired. For instance, to get the frequency distribution of 1 as the first digit, we evaluate the integral of the inverse Fourier transform over $[1,2)$, $(-2,-1]$, $[0.1,0.2)$, $(-0.2,-0.1]$, $[0.01,0.02)$, and $(-0.02,-0.01]$;

(3) We intentionally omit the ranges of values which are of the magnitude of 10^{-3} or less as the probability of the resulting values falling into these intervals is very small.

TABLE 1. Frequency distribution (in %) of the first digits of the convolution (Benford and normal)

Digit	Benford	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
1	30.10	19.61	19.79	21.58	23.90	26.66	29.54	32.23	34.52	36.29	37.52
2	17.61	18.29	15.42	14.17	13.14	12.26	11.72	11.67	12.12	12.99	14.15
3	12.49	14.64	14.26	13.21	12.31	11.41	10.53	9.72	9.04	8.54	8.26
4	9.69	11.44	12.48	11.99	11.33	10.63	9.92	9.23	8.58	7.99	7.46
5	7.92	9.44	10.58	10.59	9.85	9.74	9.23	8.69	8.16	7.65	7.17
6	6.69	8.15	8.91	9.13	8.93	8.78	8.46	8.08	7.68	7.27	6.87
7	5.80	7.22	7.46	7.69	7.80	7.78	7.65	7.43	7.16	6.85	6.53
8	5.12	6.29	6.13	6.35	6.62	6.79	6.83	6.76	6.61	6.41	6.17
9	4.58	4.82	4.83	5.14	5.32	5.84	6.02	6.09	6.05	5.95	5.79

Note: We report the frequency distribution of the first digits of the convolution and that of the theoretical Benford's distribution. The table shows different parameterizations of the normal distribution; the values of the standard deviation of the normal distribution vary from 0.1 to 1.

TABLE 2. MAD based on Table 1

Parameterization	MAD
N(0, 0.1)	0.023
N(0, 0.2)	0.028
N(0, 0.3)	0.026
N(0, 0.4)	0.023
N(0, 0.5)	0.022
N(0, 0.6)	0.019
N(0, 0.7)	0.020
N(0, 0.8)	0.022
N(0, 0.9)	0.023
N(0, 1.0)	0.024

Note: We report the MAD values for each convolution of the density of a Benford's distribution bounded by $[0.1, 1)$ and the density of a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1. An MAD value greater than 0.012 is considered as non-conformity [DN].

Table 1 reports the frequency distribution of the first digits of the convolution. The second column "Benford" shows the theoretical probabilities of the first digits based on Benford's Law. From the third column onwards, we report the frequency distributions of the first digits based on different standard deviations of the normal distribution in the convolution. Table 2 reports the MAD values based on Table 1. All MAD values in the table are greater than 0.012, indicating the data shows non-conformity to Benford's Law according to [DN].

3.2. **Sum of a Benford random variable and a uniform random variable.** Similarly, we explore the convolution of a Benford's density function and a uniform density function. Let $Z = X + Y$ where X and Y are random variables which follow a Benford's distribution and a uniform distribution, respectively. The density function of a Benford's distribution is $f(x)$ and the density function of a uniform distribution is $g(x)$:

$$f(x) = \frac{1}{x\ln(10)}, x \in [0.1, 1), \tag{3.11}$$

and

$$g(x) = \frac{1}{b-a}, x \in [a, b]. \tag{3.12}$$

The convolution of the two density functions $f(x)$ and $g(x)$ is

$$\begin{aligned} (f * g)(z) &= \int_{-\infty}^{\infty} f(z-y)g(y)\mathrm{d}y \\ &= \int_{-\infty}^{\infty} g(z-x)f(x)\mathrm{d}x. \end{aligned} \tag{3.13}$$

We set the lower bound of the uniform distribution to 0 and vary the upper bound from 0.1 to 1. When the uniform distribution is bounded by 0 and 1. For 1 to be the first digit of the convolution, there are two different cases:

(1) $X \in [0.1, 0.2)$ and $Y \in [0, 0.2 - X)$,
(2) $X \in [0.1, 1)$ and $Y \in [1 - X, 1]$,

The probability of 1 as the first digit (FD) is give by

$$\begin{aligned} \Pr(FD = 1) &= \int_{x=0.1}^{0.2}\int_{y=0}^{0.2-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\ &\quad \int_{x=0.1}^{1}\int_{y=1-x}^{1} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\ &= \int_{x=0.1}^{0.2}\int_{y=0}^{0.2-x} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\ &\quad \int_{x=0.1}^{1}\int_{y=1-x}^{1} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x \\ &= \left[\frac{0.2\ln(x)}{\ln(10)} - \frac{x}{\ln(10)}\right]_{x=0.1}^{0.2} + \left.\frac{x}{\ln(10)}\right|_{x=0.1}^{1} \\ &= \frac{0.2\ln(2) + 0.8}{\ln(10)} \\ &\approx 0.408. \end{aligned} \tag{3.14}$$

There are also two cases in which the sum of the two random variables X and Y has a first digit 2:

(1) $X \in [0.1, 0.2]$ and $Y \in [0.2 - X, 0.3 - X)$,
(2) $X \in [0.2, 0.3)$ and $Y \in [0, 0.3 - X)$.

Hence, the probability of 2 as the first digit is

$$\begin{aligned}
\Pr(FD = 2) &= \int_{x=0.1}^{0.2} \int_{y=0.2-x}^{0.3-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.2}^{0.3} \int_{y=0}^{0.3-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x=0.1}^{0.2} \int_{y=0.2-x}^{0.3-x} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.2}^{0.3} \int_{y=0}^{0.3-x} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x \\
&= \left.\frac{0.1\ln(x)}{\ln(10)}\right|_{x=0.1}^{0.2} + \left[\frac{0.3\ln(x)}{\ln(10)} - \frac{x}{\ln(10)}\right]_{x=0.2}^{0.3} \\
&\approx 0.0395. \qquad (3.15)
\end{aligned}$$

In general, there are two cases to obtain a leading digit i where $i > 1$:

(1) $X \in [0.1, 0.1 * i]$ and $Y \in [0.1 * i - X, 0.1 * (i + 1) - X)$,
(2) $X \in [0.1 * i, 0.1 * (i + 1))$ and $Y \in [0, 0.1 * (i + 1) - X)$.

The probability of the first digit as i where $i > 1$ is

$$\begin{aligned}
\Pr(FD = i) &= \int_{x=0.1}^{0.1*i} \int_{y=0.1*i-x}^{0.1*(i+1)-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.1*i}^{0.1*(i+1)} \int_{y=0}^{0.1*(i+1)-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x=0.1}^{0.1*i} \int_{y=0.1*i-x}^{0.1*(i+1)-x} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.1*i}^{0.1*(i+1)} \int_{y=0}^{0.1*(i+1)-x} 1 * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x \\
&= \left.\frac{0.1\ln(x)}{\ln(10)}\right|_{x=0.1}^{0.1*i} + \left[\frac{0.1*(i+1)\ln(x)}{\ln(10)} - \frac{x}{\ln(10)}\right]_{x=0.1*i}^{0.1*(i+1)} \\
&= \frac{(i+1)*\ln(i+1) - i*\ln(i) - 1}{10\ln(10)}. \qquad (3.16)
\end{aligned}$$

Equation (3.16) shows the closed-form solution for the frequency distribution of the first digits based on the convolution of a Benford's density function and a uniform density function bounded by $[0, 1]$. Table 3 reports the results.

TABLE 3. Frequency distribution (in %) of the first digits based on Equation (3.16)

First Digit	Benford	U[0,1]
1	30.10	40.80
2	17.61	3.95
3	12.49	5.43
4	9.69	6.52
5	7.92	7.40
6	6.69	8.12
7	5.80	8.75
8	5.12	9.29
9	4.58	9.78

Note: We report the frequency distribution of the first digits based on the Benford's distribution alone (Column 2) and the convolution of a Benford's density function and the density of a uniform distribution $U[0, 1]$ (Column 3) in percentages.

To fully examine the frequency distribution of the first digits of the resulting convolution, we explore the case of the upper bound of the uniform distribution being 0.1 later and later vary it from 0.1 to 1. For 1 to be the first digit based on the convolution of a Benford's density function and a uniform density function bounded by $[0, 0.1]$, there are two cases:

(1) $X \in [0.1, 0.2)$ and $Y \in [0, 0.2 - X)$,
(2) $X \in [0.9, 1)$ and $Y \in [1 - X, 0.1)$.

The probability of 1 as the first digit in this case is

$$\begin{aligned}
\Pr(FD=1) &= \int_{x=0.1}^{0.2}\int_{y=0}^{0.2-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.9}^{1}\int_{y=1-x}^{0.1} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x=0.1}^{0.2}\int_{y=0}^{0.2-x} \frac{1}{0.1-0} * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.9}^{1}\int_{y=1-x}^{0.1} \frac{1}{0.1-0} * \frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x \\
&= \frac{10}{\ln(10)}\left(\int_{x=0.1}^{0.2} \frac{y}{x}\bigg|_{y=0}^{0.2-x}\mathrm{d}x + \int_{x=0.9}^{1} \frac{y}{x}\bigg|_{y=1-x}^{0.1}\mathrm{d}x\right) \\
&\approx 0.190. \qquad (3.17)
\end{aligned}$$

We can generalize the formula to cases with i as the leading digit where $i > 1$:

(1) $X \in [0.1*(i-1), 0.1*i)$ and $Y \in [0.1*i - X, 0.1)$,
(2) $X \in [0.1*i, 0.1*(i+1))$ and $Y \in [0, 0.1*(i+1) - X)$.

The probability of the first digit being i when $i > 1$ is

$$\begin{aligned}
\Pr(FD=i) &= \int_{x=0.1*(i-1)}^{0.1*i}\int_{y=0.1*i-x}^{0.1} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.1*i}^{0.1*(i+1)}\int_{y=0}^{0.1*(i+1)-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\
&= \int_{x=0.1*(i-1)}^{0.1*i}\int_{y=0.1*i-x}^{0.1} \frac{10}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\
&\quad \int_{x=0.1*i}^{0.1*(i+1)}\int_{y=0}^{0.1*(i+1)-x} \frac{10}{x\ln(10)}\mathrm{d}y\mathrm{d}x \\
&= \frac{10}{\ln(10)}\left(\int_{x=0.1*(i-1)}^{0.1*i} \frac{1}{x}\bigg|_{y=0.1*i-x}^{0.1}\mathrm{d}x + \int_{x=0.1*i}^{0.1*(i+1)} \frac{1}{x}\bigg|_{y=0}^{0.1*(i+1)-x}\mathrm{d}x\right) \\
&= \frac{(i+1)\ln(i+1) - (2i)\ln(i) + (i-1)\ln(i-1)}{\ln(10)}. \qquad (3.18)
\end{aligned}$$

Table 4 reports the results of the convolution of a Benford's density function bounded by $[0.1, 1)$ and a uniform density function bounded by $[0, 0.1]$.

TABLE 4. Frequency distribution of the first digits based on Equation (3.17) and (3.18)

First Digit	Benford (%)	U[0,0.1] (%)
1	30.10	19.00
2	17.61	22.72
3	12.49	14.76
4	9.69	10.97
5	7.92	8.74
6	6.69	7.27
7	5.80	6.23
8	5.12	5.44
9	4.58	4.84

Note: We report the frequency distribution of the first digits based on the Benford's distribution alone (Column 2) and the convolution of a Benford's density function and a uniform density function of $U[0, 0.1]$ (Column 3) in percentage.

When the upper bound of the uniform distribution b is between 0.1 and 1, there are two cases such that the first digit is 1:

(1) $X \in [0.1, 0.2)$ and $Y \in [0, 0.2 - X)$,
(2) $X \in [\max\{0.1, 1 - b\}, 1)$ and $Y \in [1 - X, b)$.

The probability of 1 being the first digit is

$$\begin{aligned}\Pr(FD = 1) &= \int_{x=0.1}^{0.2} \int_{y=0}^{0.2-x} f(z-y)g(y)\mathrm{d}y\mathrm{d}x + \\ &\quad \int_{x=\max\{0.1,1-b\}}^{1} \int_{y=1-x}^{b} f(z-y)g(y)\mathrm{d}y\mathrm{d}x \\ &= \int_{x=0.1}^{0.2} \int_{y=0}^{0.2-x} \frac{1}{b-0}\frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x + \\ &\quad \int_{x=\max\{0.1,1-b\}}^{1} \int_{y=1-x}^{b} \frac{1}{b-0}\frac{1}{x\ln(10)}\mathrm{d}y\mathrm{d}x\end{aligned} \tag{3.19}$$

where $0.1 \leq b \leq 1$ and $x = \max\{0.1, 1 - b\}$.

For digits 2 to 9, there are two cases as well and the probability of i being the first digit where $i \in \{2, 3, \ldots, 9\}$:

(1) $X \in [\max\{0.1, 0.1 * i - b\}, 0.1 * i)$ and $Y \in [0.1 * i - X, A)$,
(2) $X \in [0.1 * i, 0.1 * (i + 1))$ and $Y \in [0, 0.1 * (i + 1) - X)$,

where

$$A = \begin{cases} b, & \min\{b, (i+1)*0.1 - \max\{0.1, 0.1*i - b\}\} = b \\ 0.1*(i+1) - X, & \text{otherwise.} \end{cases}$$

$$\begin{aligned} \Pr(FD = i) &= \int_{x=\max\{0.1,0.1*i-b\}}^{0.1*i} \int_{y=0.1*i-x}^{A} f(z-y)g(y)dydx + \\ &\quad \int_{x=0.1*i}^{0.1*(i+1)} \int_{y=0}^{0.1*(i+1)-x} f(z-y)g(y)dydx \\ &= \int_{x=\max\{0.1,0.1*i-b\}}^{0.1*i} \int_{y=0.1*i-x}^{A} \frac{1}{b\ln(10)}dydx + \\ &\quad \int_{x=0.1*i}^{0.1*(i+1)} \int_{y=0}^{0.1*(i+1)-x} \frac{1}{b\ln(10)}dydx, \end{aligned} \tag{3.20}$$

where $i \in \{2, 3, \ldots, 9\}$.

Table 5 reports the frequency distributions of the first digits based on the closed-form solution in Equation (3.19) and (3.20). Table 6 reports the MAD values based on different convolutions of two densities – a Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1. Among all the specifications of the upper bound, we observe that the MAD values are all greater than 0.012. All ten sets of the specifications have the MAD values indicating non-conformity to Benford's Law. The uniform distributions with upper bounds of 0.5 and 0.6 have the greatest MAD values.

Table 5. Frequency distribution (in %) of the first digits of convolution (Benford and uniform)

		Upper bound of the uniform distribution									
Digit	Benford	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
1	30.1	19.0	13.1	12.9	14.3	16.7	19.7	23.4	28.1	34.2	40.8
2	17.6	22.7	26.4	13.2	9.9	7.9	6.6	5.6	4.9	4.4	4.0
3	12.5	14.8	22.8	31.1	13.6	10.9	9.0	7.8	6.8	6.0	5.4
4	9.7	11.0	15.8	25.0	33.8	13.0	10.9	9.3	8.2	7.2	6.5
5	7.9	8.7	12.2	18.3	27.0	35.6	12.3	10.6	9.2	8.2	7.4
6	6.7	7.3	9.9	14.5	20.4	28.6	36.8	11.6	10.2	9.0	8.1
7	5.8	6.2	8.4	12.0	16.6	22.3	29.8	37.6	10.9	9.7	8.7
8	5.1	5.4	7.3	10.3	14.0	18.4	23.8	30.9	38.3	10.3	9.3
9	4.6	4.8	6.4	9.0	12.1	15.7	20.0	25.1	31.8	38.9	9.8

Note: We report the frequency distribution of the first digits of the convolution of a Benford's density function and a uniform density function with a lower bound of 0 and an upper bound ranging from 0.1 to 1. The values are calculated based on the closed-form solution in Equation (3.19) and (3.20).

Table 6. MAD based on Table 5

Upper Bound	MAD
0.1	0.025
0.2	0.063
0.3	0.099
0.4	0.121
0.5	0.132
0.6	0.132
0.7	0.122
0.8	0.102
0.9	0.080
1.0	0.054

Note: We report the MAD values based on the convolution of a Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1. The first column indicates the upper bounds of different uniform distributions; the second column reports the MAD values. An MAD value greater than 0.012 is considered as non-conformity to Benford's Law [DN].

3.3. **Replace with a normal random variable.** Another way to manipulate data is by replacement where a proportion of the original data are replaced with values which do not conform to Benford's Law. We assume that the original data follow a Benford's distribution. Similarly, we perform theoretical calculations to solve for the frequency distribution of each of the nine non-zero digits as the leading digits.

Suppose that there are two random variables X and Y where X follows the Benford's distribution bounded by the interval $[0.1, 1)$ and Y follows a normal distribution with mean μ and standard deviation σ. Assume that the original data conform to the Benford's distribution and that $m\%$ of the data are replaced by values following the aforementioned normal distribution. To calculate the frequency distribution of the leading digit d, we sum the frequency distribution of digit d in the Benford's distribution and that in the normal distribution. The frequency distribution of a leading digit d in the Benford's distribution is given by $\log_{10}(1 + \frac{1}{d})$ and the frequency distribution of the leading digit d in a normal distribution is the sum of the integrals within the intervals $[10 * d, 10 * (d + 1))$, $(-10 * (d + 1), -10 * d]$, $[d, d + 1)$, $(-(d + 1), -d]$, $[0.1 * d, 0.1 * (d + 1))$, $(-0.1 * (d + 1), -0.1 * d]$, $[0.01 * d, 0.01 * (d + 1))$, and $(-0.01 * (d + 1), -0.01 * d], \ldots$.

Let $F(x)$ be the cumulative density function of the normal distribution $N(\mu, \sigma)$. For simplicity, assume the mean of the normal distribution μ is 0. Thus, the cumulative density function of the normal distribution is symmetric about $\mu = 0$. The probabilty of a non-zero digit d being the first digit (FD) is

$$\begin{aligned}
\Pr(FD = d) &= (1-m)\log_{10}(1+\frac{1}{d}) + \\
&\quad m\sum_{n=-\infty}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] + \\
&\quad m\sum_{n=-\infty}^{\infty}[F(-d*10^{-n}) - F(-(d+1)*10^{-n})] \\
&= (1-m)\log_{10}(1+\frac{1}{d}) + \\
&\quad 2m\sum_{n=-\infty}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})], \qquad (3.21)
\end{aligned}$$

where $d \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$.

Since Equation (3.21) contains an infinite sum, we can approximate to some precision level and provide an error bound for the result.

Given the Mean Value Theorem, we know that for any n, there exists a value $c \in (d * 10^{-n}, (d+1) * 10^{-n})$ such that

$$\begin{aligned} \frac{F((d+1)*10^{-n}) - F(d*10^{-n})}{(d+1)*10^{-n} - d*10^{-n}} &= F'(c) \\ \frac{F((d+1)*10^{-n}) - F(d*10^{-n})}{10^{-n}} &= F'(c) \\ F((d+1)*10^{-n}) - F(d*10^{-n}) &= F'(c)*10^{-n}, \end{aligned} \tag{3.22}$$

where $f(x) = F'(x)$ is the probability density function of the normal distribution since the first derivative of a cumulative density function of a distribution is its probability density function. Thus, the probability density function $f(x)$ is

$$f(x) = F'(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}. \tag{3.23}$$

Since the probability density function for the half normal is monotonic, we know $F((d+1)*10^{-n}) - F(d*10^{-n}) \in (10^{-n} * f(d*10^{-n}), 10^{-n} * f((d+1)*10^{-n}))$. When we evaluate the frequency distribution of the leading digit d of the normal distribution to the precision interval of $[10^{-i} * d, 10^{-i} * (d+1))$, the probability of the digit d as the leading digit in the mixed distribution is:

$$\begin{aligned} \Pr(FD = d) &= (1-m)\log_{10}(1+\frac{1}{d}) + 2m\sum_{n=-\infty}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\ &= (1-m)\log_{10}(1+\frac{1}{d}) + \\ &\quad 2m\sum_{n=-\infty}^{i}[F((d+1)*10^{-n}) - F(d*10^{-n})] + \\ &\quad 2m\sum_{n=i+1}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\ &= (1-m)\log_{10}(1+\frac{1}{d}) + \\ &\quad 2m\sum_{n=-\infty}^{-1}[F((d+1)*10^{-n}) - F(d*10^{-n})] + \\ &\quad 2m\sum_{n=0}^{i}[F((d+1)*10^{-n}) - F(d*10^{-n})] + \\ &\quad 2m\sum_{n=i+1}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})], \end{aligned} \tag{3.24}$$

where $2m\sum_{n=-\infty}^{-1}[F((d+1)*10^{-n})-F(d*10^{-n})]$ is negligible in the estimation when we consider the probability of a two-digit integer occurs from a normal distribution with mean 0 and standard deviation less than 1, which is at least 10 standard deviations away from 0; $2m\sum_{n=0}^{i}[F((d+1)*10^{-n})-F(d*10^{-n})]$ can be calculated by a program; and $2m\sum_{n=i+1}^{\infty}[F((d+1)*10^{-n})-F(d*10^{-n})]$ is the error term which is omitted in the estimation

$$\begin{aligned}\epsilon &= \sum_{n=i+1}^{\infty}[F((d+1)*10^{-n})-F(d*10^{-n})]\\ &= \sum_{n=i+1}^{\infty}[10^{-n}*f(x)]\\ &= \sum_{n=i+1}^{\infty}[10^{-n}*\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}]\\ &= \frac{1}{\sigma\sqrt{2\pi}}\sum_{n=i+1}^{\infty}[10^{-n}*e^{-\frac{(x-\mu)^2}{2\sigma^2}}], \qquad (3.25)\end{aligned}$$

where $x \in (d*10^{-n},(d+1)*10^{-n})$.

Following Section 3.1, we calculate the frequency distribution of the leading digits based on the mixed data sets. When the standard deviation σ is 0.1, Equation (3.25) can be further simplified to:

$$\begin{aligned}\epsilon &= \frac{1}{\sigma\sqrt{2\pi}}\sum_{n=i+1}^{\infty}[10^{-n}*e^{-\frac{(x-\mu)^2}{2\sigma^2}}]\\ &= \frac{1}{0.1\sqrt{2\pi}}\sum_{n=i+1}^{\infty}[10^{-n}*e^{-\frac{x^2}{2*0.1^2}}]\\ &= \frac{10}{\sqrt{2\pi}}\sum_{n=i+1}^{\infty}[10^{-n}*e^{-50x^2}]. \qquad (3.26)\end{aligned}$$

We bound e^{-50x^2} with the interval $(0,1)$. Thus,

$$\begin{aligned}0 &< \epsilon < \frac{10}{\sqrt{2\pi}}\sum_{n=i+1}^{\infty}[10^{-n}*e^{-50x^2}]\\ 0 &< \epsilon < \frac{10}{\sqrt{2\pi}}\sum_{n=i+1}^{\infty}10^{-n}\\ 0 &< \epsilon < \frac{10}{9\sqrt{2\pi}}10^{-(i+1)}, \qquad (3.27)\end{aligned}$$

and when $i \geq 3$, the error term ϵ can be bounded by $(0,10^{-4})$.

Therefore, to obtain a precision level of 10^{-3} for the frequency distribution of a leading digit d, we choose i to be 3 and use Equation (3.28) to calculate the frequency distribution of the leading digit d,

$$\begin{aligned}\Pr(d) &= (1-m)\log_{10}(1+\frac{1}{d}) + 2m\sum_{n=-\infty}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\ &\approx (1-m)\log_{10}(1+\frac{1}{d}) + \\ &\quad 2m\sum_{n=0}^{3}[F((d+1)*10^{-n}) - F(d*10^{-n})]. \qquad (3.28)\end{aligned}$$

Table 7 reports the frequency distribution of the leading digits based on a final data set whose values are partially replaced by values following a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1. Each sub-header indicates the proportion of the data replaced. The frequency of the first digits does not vary much after controlling for the proportion of replacement. However, controlling for the standard deviation of the normal distribution, the frequency distribution of the first digits varies more as more data are replaced. The extent of deviation can be seen clearly in Table 8 which shows the MAD values for these mixed data sets. While holding the percentage of replacement constant, the MAD values for different standard deviations do not change much. However, in general the MAD value increases as the percentage of replacement increases. Figure 1 presents the same information graphically. Each subplot shows the MAD value based on data partially replaced by values following various normal distributions. Figure 2 displays the relationship among MAD, proportion of replacement, and standard deviation of the normal distribution. The MAD value is the lowest when the standard deviation of the normal distribution is around 0.8 across the proportions of data replaced.

TABLE 7. Frequency distribution (in %) of the partially replaced data set

Digit	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10% replaced										
1	30.68	30.53	29.73	29.29	29.25	29.47	29.82	30.17	30.47	30.69
2	17.13	18.12	18.02	17.71	17.43	17.22	17.08	17.03	17.05	17.14
3	12.10	12.56	12.88	12.82	12.67	12.51	12.38	12.26	12.17	12.11
4	9.52	9.48	9.88	10.00	9.96	9.87	9.77	9.68	9.60	9.53
5	7.89	7.65	7.92	8.12	8.17	8.15	8.09	8.02	7.96	7.90
6	6.75	6.47	6.57	6.78	6.89	6.91	6.89	6.85	6.81	6.76
7	5.90	5.63	5.62	5.78	5.91	5.97	5.99	5.97	5.94	5.91
8	5.24	5.01	4.94	5.03	5.15	5.24	5.27	5.28	5.27	5.25
9	4.71	4.51	4.42	4.45	4.55	4.64	4.70	4.72	4.72	4.71
20% replaced										
1	31.26	30.95	29.36	28.48	28.39	28.84	29.54	30.24	30.83	31.27
2	16.65	18.63	18.42	17.80	17.25	16.83	16.56	16.45	16.50	16.66
3	11.71	12.62	13.27	13.15	12.84	12.53	12.26	12.03	11.85	11.72
4	9.36	9.27	10.07	10.31	10.23	10.05	9.86	9.67	9.51	9.37
5	7.87	7.38	7.91	8.32	8.43	8.37	8.26	8.13	8.00	7.88
6	6.81	6.24	6.45	6.86	7.08	7.13	7.09	7.01	6.92	6.82
7	6.00	5.47	5.45	5.76	6.03	6.15	6.17	6.14	6.09	6.02
8	5.36	4.90	4.76	4.94	5.19	5.36	5.43	5.44	5.42	5.38
9	4.84	4.45	4.26	4.33	4.53	4.71	4.82	4.86	4.87	4.85
30% replaced										
1	31.83	31.38	28.99	27.67	27.53	28.21	29.25	30.31	31.19	31.86
2	16.17	19.14	18.83	17.90	17.07	16.44	16.04	15.87	15.94	16.19
3	11.31	12.69	13.66	13.48	13.01	12.55	12.14	11.80	11.53	11.34
4	9.19	9.06	10.26	10.62	10.50	10.23	9.94	9.67	9.42	9.21
5	7.84	7.11	7.91	8.53	8.68	8.60	8.43	8.23	8.04	7.86
6	6.86	6.01	6.32	6.94	7.27	7.34	7.28	7.16	7.03	6.89
7	6.10	5.31	5.27	5.75	6.14	6.32	6.36	6.32	6.23	6.13
8	5.49	4.80	4.58	4.86	5.23	5.48	5.59	5.61	5.57	5.51
9	4.97	4.39	4.10	4.21	4.51	4.78	4.94	5.01	5.02	4.99
40% replaced										
1	32.41	31.80	28.62	26.86	26.67	27.59	28.97	30.37	31.56	32.44
2	15.69	19.66	19.24	17.99	16.90	16.06	15.51	15.29	15.38	15.72
3	10.92	12.75	14.05	13.81	13.19	12.56	12.02	11.57	11.21	10.95
4	9.02	8.85	10.45	10.93	10.77	10.41	10.02	9.66	9.33	9.05
5	7.81	6.84	7.91	8.73	8.94	8.83	8.60	8.34	8.08	7.84
6	6.92	5.78	6.20	7.02	7.46	7.56	7.48	7.32	7.14	6.95
7	6.21	5.14	5.10	5.73	6.25	6.50	6.55	6.49	6.37	6.24
8	5.61	4.69	4.40	4.77	5.27	5.60	5.75	5.77	5.73	5.64
9	5.10	4.33	3.94	4.08	4.49	4.85	5.06	5.15	5.17	5.13
50% replaced										
1	32.98	32.23	28.25	26.04	25.82	26.96	28.69	30.44	31.92	33.02
2	15.21	20.17	19.65	18.09	16.72	15.67	14.99	14.71	14.83	15.25
3	10.53	12.82	14.45	14.14	13.36	12.58	11.90	11.33	10.88	10.57
4	8.85	8.64	10.64	11.24	11.04	10.59	10.10	9.65	9.24	8.89
5	7.79	6.56	7.90	8.93	9.19	9.05	8.77	8.44	8.12	7.83
6	6.97	5.55	6.07	7.11	7.65	7.77	7.67	7.48	7.25	7.01
7	6.31	4.98	4.92	5.71	6.36	6.67	6.73	6.66	6.52	6.35
8	5.73	4.58	4.22	4.68	5.31	5.72	5.91	5.94	5.88	5.77
9	5.23	4.27	3.78	3.96	4.47	4.91	5.18	5.30	5.31	5.27

Continued: Frequency distribution of the partially replaced distribution

Digit	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
60% replaced										
1	33.56	32.65	27.88	25.23	24.96	26.33	28.40	30.51	32.28	33.61
2	14.73	20.68	20.05	18.19	16.54	15.28	14.46	14.13	14.27	14.78
3	10.14	12.88	14.84	14.47	13.54	12.60	11.78	11.10	10.56	10.18
4	8.68	8.43	10.82	11.55	11.31	10.77	10.19	9.64	9.15	8.73
5	7.76	6.29	7.90	9.13	9.44	9.28	8.94	8.55	8.16	7.81
6	7.03	5.32	5.95	7.19	7.84	7.99	7.87	7.63	7.36	7.08
7	6.41	4.81	4.74	5.69	6.48	6.85	6.92	6.83	6.66	6.46
8	5.86	4.48	4.04	4.60	5.35	5.84	6.07	6.10	6.03	5.91
9	5.36	4.21	3.62	3.84	4.45	4.98	5.30	5.44	5.46	5.40
70% replaced										
1	34.14	33.08	27.51	24.42	24.10	25.70	28.12	30.58	32.65	34.19
2	14.25	21.19	20.46	18.28	16.36	14.89	13.94	13.55	13.71	14.30
3	9.74	12.95	15.23	14.79	13.71	12.62	11.66	10.87	10.24	9.80
4	8.52	8.22	11.01	11.86	11.58	10.95	10.27	9.63	9.06	8.57
5	7.73	6.02	7.90	9.34	9.70	9.51	9.11	8.65	8.20	7.79
6	7.09	5.09	5.82	7.27	8.03	8.20	8.07	7.79	7.47	7.14
7	6.51	4.65	4.57	5.67	6.59	7.02	7.11	7.00	6.80	6.57
8	5.98	4.37	3.86	4.51	5.39	5.97	6.23	6.27	6.19	6.04
9	5.49	4.15	3.46	3.71	4.43	5.05	5.42	5.59	5.61	5.54
80% replaced										
1	34.71	33.50	27.14	23.61	23.25	25.07	27.84	30.64	33.01	34.78
2	13.77	21.70	20.87	18.38	16.18	14.50	13.41	12.97	13.16	13.83
3	9.35	13.01	15.62	15.12	13.88	12.63	11.54	10.64	9.92	9.41
4	8.35	8.02	11.20	12.16	11.85	11.13	10.35	9.62	8.97	8.41
5	7.71	5.75	7.89	9.54	9.95	9.74	9.28	8.76	8.24	7.77
6	7.14	4.86	5.70	7.35	8.22	8.42	8.26	7.95	7.58	7.21
7	6.61	4.49	4.39	5.66	6.70	7.20	7.29	7.18	6.95	6.68
8	6.11	4.26	3.68	4.43	5.43	6.09	6.38	6.43	6.34	6.17
9	5.62	4.09	3.30	3.59	4.41	5.11	5.54	5.73	5.75	5.68
90% replaced										
1	35.29	33.93	26.77	22.80	22.39	24.44	27.55	30.71	33.38	35.36
2	13.29	22.21	21.28	18.48	16.00	14.12	12.89	12.39	12.60	13.36
3	8.96	13.07	16.01	15.45	14.06	12.65	11.43	10.40	9.60	9.03
4	8.18	7.81	11.39	12.47	12.12	11.31	10.43	9.61	8.89	8.25
5	7.68	5.48	7.89	9.74	10.21	9.96	9.45	8.86	8.29	7.75
6	7.20	4.63	5.58	7.44	8.41	8.64	8.46	8.11	7.69	7.27
7	6.72	4.32	4.22	5.64	6.82	7.37	7.48	7.35	7.09	6.79
8	6.23	4.16	3.50	4.34	5.47	6.21	6.54	6.60	6.49	6.30
9	5.75	4.02	3.14	3.46	4.38	5.18	5.66	5.87	5.90	5.82
100% replaced										
1	35.87	34.35	26.40	21.99	21.53	23.81	27.27	30.78	33.74	35.94
2	12.81	22.73	21.68	18.57	15.82	13.73	12.36	11.81	12.04	12.89
3	8.56	13.14	16.40	15.78	14.23	12.67	11.31	10.17	9.28	8.64
4	8.01	7.60	11.58	12.78	12.39	11.49	10.52	9.61	8.80	8.09
5	7.65	5.21	7.89	9.94	10.46	10.19	9.62	8.97	8.33	7.73
6	7.25	4.41	5.45	7.52	8.60	8.85	8.65	8.26	7.80	7.33
7	6.82	4.16	4.04	5.62	6.93	7.54	7.67	7.52	7.24	6.90
8	6.35	4.05	3.32	4.25	5.51	6.33	6.70	6.76	6.65	6.43
9	5.88	3.96	2.98	3.34	4.36	5.25	5.78	6.02	6.05	5.96

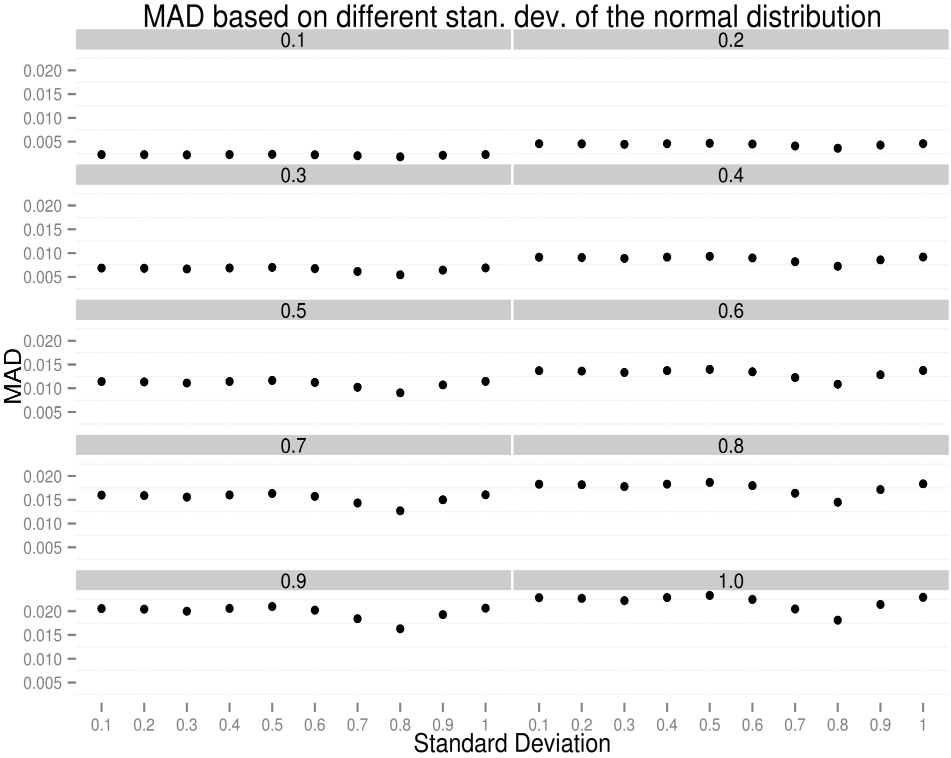
Note: We report the frequency distribution of the first digits of the mixed data. $(1-m)\%$ of the data follow a Benford's distribution; $m\%$ follow a normal distribution. The table reports different parameterizations of the normal distribution.

Table 8. MAD based on Table 7

	Percentage of Benford's distribution replaced (%)									
	10	20	30	40	50	60	70	80	90	100
N(0, 0.1)	0.002	0.005	0.007	0.009	0.011	0.014	0.016	0.018	0.021	0.023
N(0, 0.2)	0.002	0.005	0.007	0.009	0.011	0.014	0.016	0.018	0.020	0.023
N(0, 0.3)	0.002	0.004	0.007	0.009	0.011	0.013	0.016	0.018	0.020	0.022
N(0, 0.4)	0.002	0.005	0.007	0.009	0.011	0.014	0.016	0.018	0.021	0.023
N(0, 0.5)	0.002	0.005	0.007	0.009	0.012	0.014	0.016	0.019	0.021	0.023
N(0, 0.6)	0.002	0.004	0.007	0.009	0.011	0.013	0.016	0.018	0.020	0.022
N(0, 0.7)	0.002	0.004	0.006	0.008	0.010	0.012	0.014	0.016	0.018	0.020
N(0, 0.8)	0.002	0.004	0.005	0.007	0.009	0.011	0.013	0.014	0.016	0.018
N(0, 0.9)	0.002	0.004	0.006	0.009	0.011	0.013	0.015	0.017	0.019	0.021
N(0, 1.0)	0.002	0.005	0.007	0.009	0.011	0.014	0.016	0.018	0.021	0.023

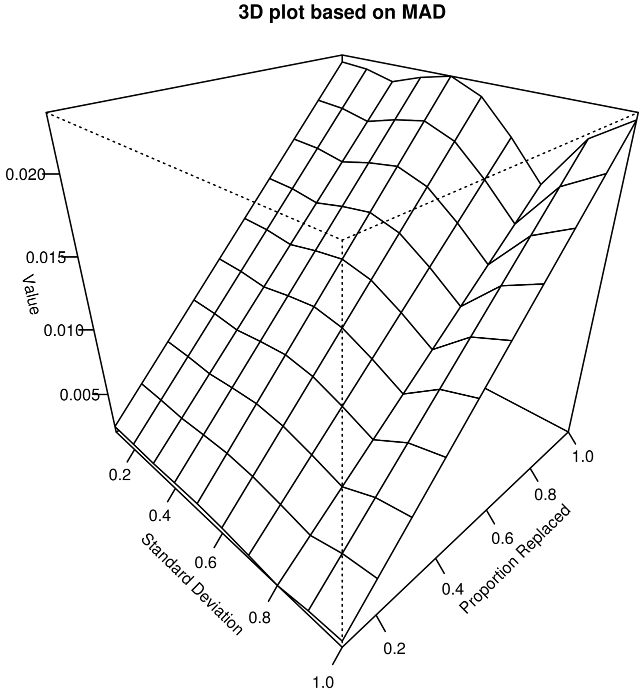
Note: We report the MAD values of mixed data, $m\%$ of which are replaced by values following a normal distribution. The table reports the results for different parameterizations of the normal distribution. An MAD value greater than 0.012 is considered as non-conformity [DN].

FIGURE 1. MAD based on Table 7



Note: We display the MAD values of mixed data, $m\%$ of which are replaced by values following a normal distribution. Each subplot shows a different proportion of the original data being replaced.

FIGURE 2. 3-D plot of based on Table 8



Note: We display the relationship among MAD, the proportion of replacement, and the standard deviation of the normal distribution. The MAD values are the lowest when the standard deviation of the normal distribution is around 0.8 across various proportions of data replaced.

3.4. **Replace with a uniform random variable.** We perform theoretical calculations to solve for the frequency distribution of each of the nine non-zero digits as the leading digits when we replace a proportion of data conforming to Benford's Law with data following a uniform distribution. Suppose there are two random variables X and Y followig a Benford's distribution bounded by $[0.1, 1)$ anda a uniform distribution between a and b, respectively. Assume that the original data conform to Benford's Law and that $m\%$ of the data are replaced by data generated from the aforementioned uniform distribution. To calculate the frequency distribution of the leading digit d, we sum the frequency distribution of digit d in the Benford's distribution and that in the uniform distribution. The frequency distribution of a digit d as the leading digit in a Benford's distribution is given by $\log_{10}(1 + \frac{1}{d})$ and the frequency distribution of a digit d as the leading digit in a uniform distribution is the sum of the integrals within intervals $[d, d+1)$, $(-(d+1), -d]$, $[0.1*d, 0.1*(d+1))$, $(-0.1*(d+1), -0.1*d]$, $[0.01*d, 0.01*(d+1))$, and $(-0.01*(d+1), -0.01*d]$, $\ldots$ where $d\ in[a, b]$.

Let $F(x)$ be the cumulative density function of the uniform distribution $U[a, b]$. Assume that the lower bound of the uniform distribution is 0 and the upper bound varies from 0.1 to 1. The probability of a digit d occurring as the first digit (FD) is

$$\begin{aligned}
\Pr(FD = d) &= (1-m)\log_{10}(1+\frac{1}{d}) + \\
&\quad m\sum_{n=-\infty}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\
&= (1-m)\log_{10}(1+\frac{1}{d}) + \\
&\quad m\sum_{n=-\infty}^{0}[F((d+1)*10^{-n}) - F(d*10^{-n})] + \\
&\quad m\sum_{n=1}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\
&= (1-m)\log_{10}(1+\frac{1}{d}) + \\
&\quad m\sum_{n=1}^{\infty}[F((d+1)*10^{-n}) - F(d*10^{-n})] \\
&= (1-m)\log_{10}(1+\frac{1}{d}) + \frac{m}{90b} + \\
&\quad m\left[\frac{(d+1)*10^{-1}}{b} - \frac{d*10^{-1}}{b}\right], \qquad (3.29)
\end{aligned}$$

where $d \in \{1, 2, 3, 4, 5, 6, 7, 8, 9\}$ and $m \sum_{n=-\infty}^{0} [F((d+1)*10^{-n}) - F(d*10^{-n})]$ is 0 becase the probability of a two-digit integer occurs from a uniform distribution between 0 and 1 is 0. Since the upper bound b ranges from 0.1 to 1, when the first digit d is greater than or equal to the upper bound b, the term $\left[\frac{(d+1)*10^{-1}}{b} - \frac{d*10^{-1}}{b}\right]$ is 0.

Table 9 reports the frequency distribution of the first digit based on data partially replaced by values following a uniform distribution with a lower bound 0 and an upper bound ranging from 0.1 to 1. The values are reported in percentages. Table 10 reports the MAD values for the data partially replaced by values following a uniform distribution. There is an monotonic increase in the values of MAD with the proportion of the data replaced, given the same upper bound.

TABLE 9. Frequency distribution (in %) of the partially replaced data set

	Upper bound of the uniform distribution									
Digit	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
10% replaced										
1	28.20	32.65	30.80	29.87	29.31	28.94	28.68	28.48	28.33	28.20
2	16.96	16.40	19.55	18.63	18.07	17.70	17.44	17.24	17.08	16.96
3	12.36	11.80	11.61	14.02	13.47	13.10	12.83	12.63	12.48	12.36
4	9.83	9.28	9.09	9.00	10.94	10.57	10.31	10.11	9.96	9.83
5	8.24	7.68	7.50	7.40	7.35	8.98	8.71	8.52	8.36	8.24
6	7.14	6.58	6.40	6.30	6.25	6.21	7.61	7.41	7.26	7.14
7	6.33	5.77	5.59	5.50	5.44	5.40	5.38	6.61	6.45	6.33
8	5.71	5.16	4.97	4.88	4.83	4.79	4.76	4.74	5.84	5.71
9	5.23	4.67	4.49	4.40	4.34	4.30	4.28	4.26	4.24	5.23
20% replaced										
1	26.30	35.19	31.49	29.64	28.53	27.79	27.26	26.86	26.55	26.30
2	16.31	15.20	21.49	19.64	18.53	17.79	17.26	16.87	16.56	16.31
3	12.22	11.11	10.74	15.55	14.44	13.70	13.17	12.77	12.46	12.22
4	9.98	8.86	8.49	8.31	12.20	11.46	10.93	10.53	10.22	9.98
5	8.56	7.45	7.08	6.89	6.78	10.04	9.51	9.11	8.80	8.56
6	7.58	6.47	6.10	5.91	5.80	5.73	8.53	8.13	7.82	7.58
7	6.86	5.75	5.38	5.19	5.08	5.01	4.96	7.42	7.11	6.86
8	6.31	5.20	4.83	4.65	4.54	4.46	4.41	4.37	6.56	6.31
9	5.88	4.77	4.40	4.22	4.11	4.03	3.98	3.94	3.91	5.88
30% replaced										
1	24.41	37.74	32.18	29.41	27.74	26.63	25.83	25.24	24.78	24.41
2	15.66	13.99	23.44	20.66	18.99	17.88	17.09	16.49	16.03	15.66
3	12.08	10.41	9.86	17.08	15.41	14.30	13.51	12.91	12.45	12.08
4	10.12	8.45	7.89	7.62	13.45	12.34	11.55	10.95	10.49	10.12
5	8.88	7.21	6.65	6.38	6.21	11.10	10.30	9.71	9.25	8.88
6	8.02	6.35	5.80	5.52	5.35	5.24	9.45	8.85	8.39	8.02
7	7.39	5.73	5.17	4.89	4.73	4.61	4.54	8.23	7.76	7.39
8	6.91	5.25	4.69	4.41	4.25	4.14	4.06	4.00	7.28	6.91
9	6.54	4.87	4.31	4.04	3.87	3.76	3.68	3.62	3.57	6.54
40% replaced										
1	22.51	40.28	32.88	29.17	26.95	25.47	24.41	23.62	23.00	22.51
2	15.01	12.79	25.38	21.68	19.45	17.97	16.91	16.12	15.50	15.01
3	11.94	9.72	8.98	18.61	16.39	14.90	13.85	13.05	12.43	11.94
4	10.26	8.04	7.30	6.93	14.70	13.22	12.16	11.37	10.75	10.26
5	9.20	6.97	6.23	5.86	5.64	12.16	11.10	10.31	9.69	9.20
6	8.46	6.24	5.50	5.13	4.91	4.76	10.37	9.57	8.96	8.46
7	7.92	5.70	4.96	4.59	4.37	4.22	4.11	9.04	8.42	7.92
8	7.51	5.29	4.55	4.18	3.96	3.81	3.70	3.62	8.01	7.51
9	7.19	4.97	4.23	3.86	3.63	3.49	3.38	3.30	3.24	7.19
50% replaced										
1	20.61	42.83	33.57	28.94	26.16	24.31	22.99	22.00	21.22	20.61
2	14.36	11.58	27.32	22.69	19.92	18.06	16.74	15.75	14.98	14.36
3	11.80	9.02	8.10	20.14	17.36	15.51	14.18	13.19	12.42	11.80
4	10.40	7.62	6.70	6.23	15.96	14.10	12.78	11.79	11.02	10.40
5	9.51	6.74	5.81	5.35	5.07	13.22	11.90	10.90	10.13	9.51
6	8.90	6.13	5.20	4.74	4.46	4.27	11.28	10.29	9.52	8.90
7	8.46	5.68	4.75	4.29	4.01	3.83	3.69	9.84	9.07	8.46
8	8.11	5.34	4.41	3.95	3.67	3.48	3.35	3.25	8.73	8.11
9	7.84	5.07	4.14	3.68	3.40	3.21	3.08	2.98	2.91	7.84

Continued: Frequency distribution (in %) of the partially replaced data set

	Upper bound of the uniform distribution									
Digit	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
60% replaced										
1	18.71	45.37	34.26	28.71	25.37	23.15	21.57	20.37	19.45	18.71
2	13.71	10.38	29.27	23.71	20.38	18.15	16.57	15.38	14.45	13.71
3	11.66	8.33	7.22	21.66	18.33	16.11	14.52	13.33	12.40	11.66
4	10.54	7.21	6.10	5.54	17.21	14.99	13.40	12.21	11.28	10.54
5	9.83	6.50	5.39	4.83	4.50	14.28	12.69	11.50	10.57	9.83
6	9.34	6.01	4.90	4.34	4.01	3.79	12.20	11.01	10.09	9.34
7	8.99	5.65	4.54	3.99	3.65	3.43	3.27	10.65	9.73	8.99
8	8.71	5.38	4.27	3.71	3.38	3.16	3.00	2.88	9.45	8.71
9	8.50	5.16	4.05	3.50	3.16	2.94	2.78	2.66	2.57	8.50
70% replaced										
1	16.81	47.92	34.96	28.48	24.59	21.99	20.14	18.75	17.67	16.81
2	13.06	9.17	31.21	24.73	20.84	18.25	16.39	15.00	13.92	13.06
3	11.53	7.64	6.34	23.19	19.30	16.71	14.86	13.47	12.39	11.53
4	10.69	6.80	5.50	4.85	18.46	15.87	14.02	12.63	11.55	10.69
5	10.15	6.26	4.97	4.32	3.93	15.34	13.49	12.10	11.02	10.15
6	9.79	5.90	4.60	3.95	3.56	3.30	13.12	11.73	10.65	9.79
7	9.52	5.63	4.33	3.68	3.30	3.04	2.85	11.46	10.38	9.52
8	9.31	5.42	4.13	3.48	3.09	2.83	2.65	2.51	10.18	9.31
9	9.15	5.26	3.97	3.32	2.93	2.67	2.48	2.34	2.24	9.15
80% replaced										
1	14.91	50.47	35.65	28.24	23.80	20.84	18.72	17.13	15.90	14.91
2	12.41	7.97	33.15	25.74	21.30	18.34	16.22	14.63	13.40	12.41
3	11.39	6.94	5.46	24.72	20.28	17.31	15.20	13.61	12.38	11.39
4	10.83	6.38	4.90	4.16	19.72	16.75	14.64	13.05	11.81	10.83
5	10.47	6.03	4.55	3.81	3.36	16.40	14.28	12.69	11.46	10.47
6	10.23	5.78	4.30	3.56	3.12	2.82	14.04	12.45	11.22	10.23
7	10.05	5.60	4.12	3.38	2.94	2.64	2.43	12.27	11.04	10.05
8	9.91	5.47	3.99	3.25	2.80	2.50	2.29	2.13	10.90	9.91
9	9.80	5.36	3.88	3.14	2.69	2.40	2.18	2.03	1.90	9.80
90% replaced										
1	13.01	53.01	36.34	28.01	23.01	19.68	17.30	15.51	14.12	13.01
2	11.76	6.76	35.09	26.76	21.76	18.43	16.05	14.26	12.87	11.76
3	11.25	6.25	4.58	26.25	21.25	17.92	15.54	13.75	12.36	11.25
4	10.97	5.97	4.30	3.47	20.97	17.64	15.25	13.47	12.08	10.97
5	10.79	5.79	4.13	3.29	2.79	17.46	15.08	13.29	11.90	10.79
6	10.67	5.67	4.00	3.17	2.67	2.34	14.96	13.17	11.78	10.67
7	10.58	5.58	3.91	3.08	2.58	2.25	2.01	13.08	11.69	10.58
8	10.51	5.51	3.84	3.01	2.51	2.18	1.94	1.76	11.62	10.51
9	10.46	5.46	3.79	2.96	2.46	2.12	1.89	1.71	1.57	10.46
100% replaced										
1	11.11	55.56	37.04	27.78	22.22	18.52	15.87	13.89	12.35	11.11
2	11.11	5.56	37.04	27.78	22.22	18.52	15.87	13.89	12.35	11.11
3	11.11	5.56	3.70	27.78	22.22	18.52	15.87	13.89	12.35	11.11
4	11.11	5.56	3.70	2.78	22.22	18.52	15.87	13.89	12.35	11.11
5	11.11	5.56	3.70	2.78	2.22	18.52	15.87	13.89	12.35	11.11
6	11.11	5.56	3.70	2.78	2.22	1.85	15.87	13.89	12.35	11.11
7	11.11	5.56	3.70	2.78	2.22	1.85	1.59	13.89	12.35	11.11
8	11.11	5.56	3.70	2.78	2.22	1.85	1.59	1.39	12.35	11.11
9	11.11	5.56	3.70	2.78	2.22	1.85	1.59	1.39	1.23	11.11

Note: We report the frequency distribution of the first digits of the mixed data sets including $(1-m)\%$ of the values following a Benford's distribution and $m\%$ of the values following a uniform distribution.

TABLE 10. MAD based on Table 9

Upper Bound	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0.1	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.048	0.054	0.060
0.2	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.048	0.054	0.060
0.3	0.006	0.012	0.018	0.023	0.029	0.035	0.041	0.047	0.053	0.059
0.4	0.006	0.011	0.017	0.023	0.028	0.034	0.040	0.045	0.051	0.057
0.5	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.048	0.054	0.060
0.6	0.006	0.012	0.018	0.023	0.029	0.035	0.041	0.047	0.053	0.059
0.7	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.047	0.053	0.059
0.8	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.048	0.054	0.060
0.9	0.006	0.012	0.018	0.024	0.029	0.035	0.041	0.047	0.053	0.059
1.0	0.006	0.012	0.018	0.024	0.030	0.036	0.042	0.048	0.054	0.060

Note: We report the MAD values based on mixed data which are partially repalced by values following a uniform distribution. The lower bounds of the uniform distributions are 0 and the upper bounds range from 0.1 to 1. An MAD value greater than 0.012 is considered as non-conformity to Benford's Law [DN].

4. SIMULATION STUDIES ON MIXED DISTRIBUTIONS

We conduct simulation studies to gain a better understanding of the frequency distribution of the first digits in the cases of convolution and partial replacement mentioned in Section 3. Section 4.1 reports the simulation results of a convolution of a Benford's density and a normal density. Section 4.2 presents the simulation results of mixed data sets where some of the original data following a Benford's distribution are replaced by values following a normal distribution. Since we can calculate the frequency distribution of the first digits based on the convolution of a Benford's density and a uniform density (Section 3.2) and those of partial replacement with data following a uniform distribution (Section 3.4) without numerical integration, we do not perform simulations for these two scenarios.

4.1. **Simulated convolution of Benford and normal densities.** We perform simulations to explore and potentially identify the relationship among different measures of deviation from Benford's Law, the parameterization of the normal distribution, and the number of data points used. The number of data points is important in that the more data there are, the more accurate the extent of deviation can be measured.

We start with the convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0 as shown in Section 3.1. Since we cannot draw from the Benford's distribution directly, we use the inverse cumulative density function method to randomly generate numbers from the Benford's distribution. We then randomly draw numbers from a normal distribution and sum them with those drawn from the Benford's distribution. We calculate various deviation measures including the Chi-Square test statistic, the p-value from the Chi-Square test, and the MAD values. We vary the standard deviation of the normal distribution and the sample size to observe the changes in various deviation measures.

The data generating procedure is as follows:

(1) The probability density function of the Benford's distribution bounded by $[0.1, 1)$ is $f(x) = \frac{1}{x\ln(10)}$; the cumulative density function is:

$$\begin{aligned} F(x) &= \int_{0.1}^{x} f(x)\mathrm{d}x \\ &= \frac{\ln(x) - \ln(0.1)}{\ln(10)}. \end{aligned} \tag{4.1}$$

(2) We draw a random number u from the uniform distribution bounded by the interval $[0, 1]$.

(3) Solve the value of x such that $F(x) = u$. It is clear that $x = e^{u\ln(10)+\ln(0.1)}$ and x is the random number drawn from the Benford's distribution.

(4) Randomly draw a number y from a normal distribution with mean μ and standard deviation σ.

(5) Obtain the sum z of x and y, $z = x + y$.

(6) Repeat the procedure to draw a different random number. The sample sizes are 10, 22, 46, 100, 215, 464, 1000, 2154, 4641, and 10000. These are the rounded numbers when the range from 10 to 10000 is divided into 10 equal intervals on the natural-logarithmic scale.

Table 11 and Table 12 report the test statistics based on the Chi-Square tests and the associated p-values based on the simulated data sets. Neither the test statistics or the p-values offer much evidence as to the conformity of the data sets to Benford's Law. Table 13 reports the MAD values based on the simulation results. All MAD values are greater than 0.012 which indicates non-conformity to Benford's Law according to [DN].

TABLE 11. Chi-Square test statistics based on simulations (convolution of Benford and normal)

	Sample Size									
St. Dev.	10	22	46	100	215	464	1000	2154	4642	10000
0.1	9.49	10.18	11.54	14.53	21.10	35.21	65.26	130.01	270.33	572.02
0.2	9.80	10.77	12.42	16.29	24.71	42.52	81.32	164.41	343.57	730.15
0.3	9.83	10.56	12.21	15.67	23.16	39.31	74.27	148.83	309.61	657.51
0.4	9.78	10.39	11.67	14.70	21.01	34.78	64.43	128.51	265.40	561.02
0.5	9.62	10.15	11.30	13.92	19.53	31.56	57.64	113.92	235.44	496.22
0.6	9.32	9.90	11.00	13.50	18.83	30.23	55.03	108.08	222.86	470.63
0.7	9.10	9.70	10.80	13.34	18.76	30.29	55.30	108.89	225.12	474.54
0.8	8.94	9.49	10.63	13.27	18.78	30.95	56.88	112.52	232.95	492.53
0.9	8.69	9.28	10.47	13.16	18.99	31.42	58.37	116.55	241.30	510.37
1.0	8.48	9.07	10.33	13.11	19.12	32.03	59.49	119.09	247.44	523.59

Note: We report the test statistics of the Chi-Square test based on simulated data sets. The data are a result of a convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0 and standard deviation ranging from 0.1 to 1. Results reported here are the average of 20000 times of simulations.

TABLE 12. P-values associated with the Chi-Square test based on simulations (convolution of Benford and normal)

	Sample Size									
St. Dev.	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.42	0.38	0.30	0.18	0.06	0.00	0.00	0.00	0.00	0.00
0.2	0.40	0.35	0.27	0.14	0.03	0.00	0.00	0.00	0.00	0.00
0.3	0.40	0.36	0.28	0.16	0.04	0.00	0.00	0.00	0.00	0.00
0.4	0.41	0.37	0.31	0.19	0.07	0.01	0.00	0.00	0.00	0.00
0.5	0.41	0.39	0.33	0.22	0.09	0.01	0.00	0.00	0.00	0.00
0.6	0.43	0.40	0.34	0.23	0.09	0.01	0.00	0.00	0.00	0.00
0.7	0.45	0.40	0.34	0.23	0.09	0.01	0.00	0.00	0.00	0.00
0.8	0.45	0.41	0.35	0.23	0.09	0.01	0.00	0.00	0.00	0.00
0.9	0.46	0.42	0.35	0.23	0.08	0.01	0.00	0.00	0.00	0.00
1.0	0.48	0.43	0.36	0.23	0.08	0.01	0.00	0.00	0.00	0.00

Note: We report the p-values associated with the Chi-Square test based on the simulations. This is based on the convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0 and standard deviation ranging from 0.1 to 1. The degrees of freedom to calculate the p-values are 8. Results are the average values of 20000 times of simulations.

TABLE 13. MAD based on simulations (convolution of Benford and normal)

	Sample Size									
St. Dev.	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.08	0.06	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.02
0.2	0.08	0.06	0.04	0.04	0.03	0.03	0.03	0.03	0.03	0.03
0.3	0.08	0.06	0.04	0.03	0.03	0.03	0.03	0.03	0.03	0.03
0.4	0.08	0.06	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.02
0.5	0.08	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02
0.6	0.08	0.05	0.04	0.03	0.02	0.02	0.02	0.02	0.02	0.02
0.7	0.08	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02
0.8	0.08	0.05	0.04	0.03	0.03	0.02	0.02	0.02	0.02	0.02
0.9	0.08	0.05	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.02
1.0	0.08	0.05	0.04	0.03	0.03	0.03	0.02	0.02	0.02	0.02

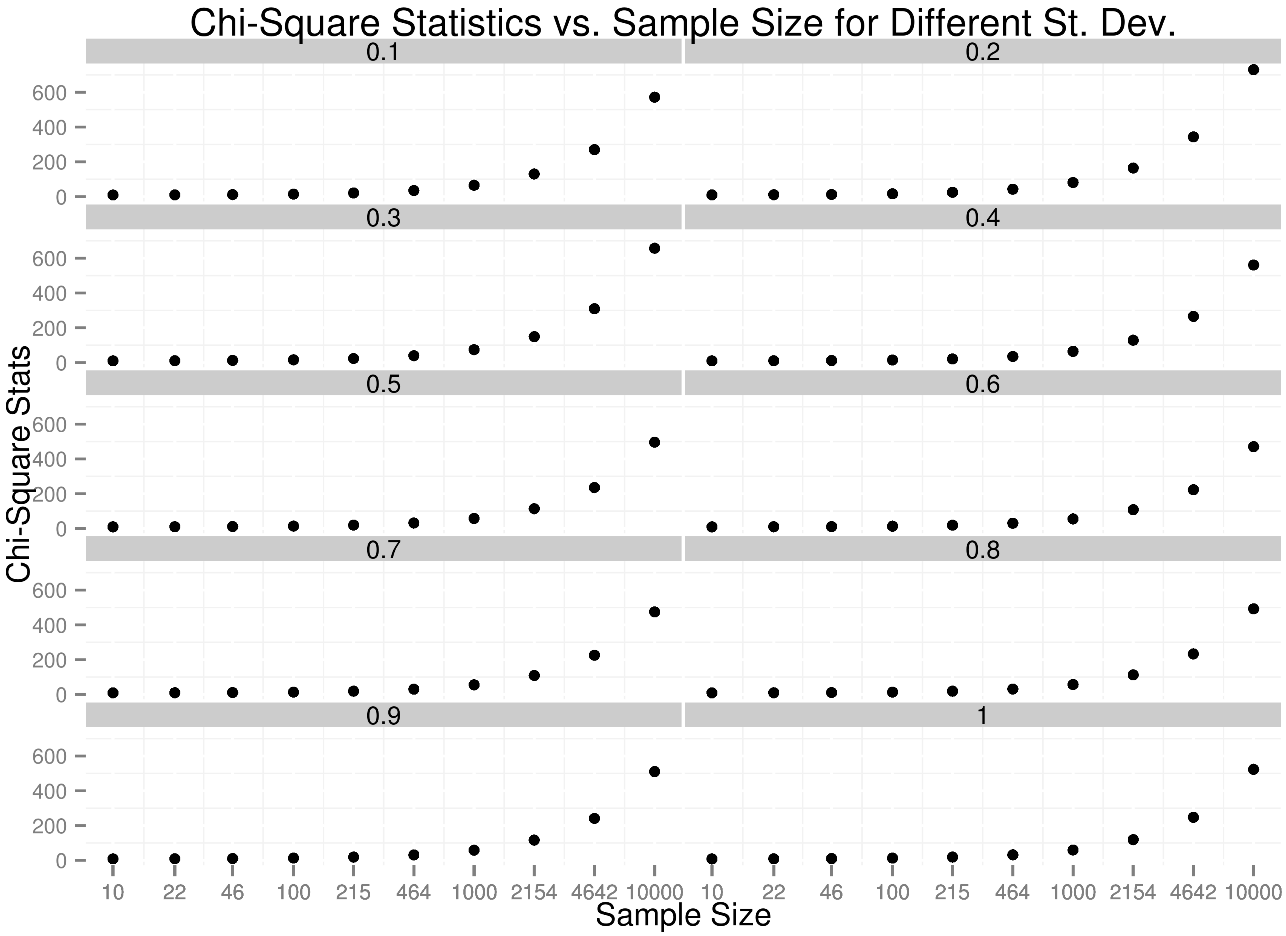
Note: We report the MAD values based on the simulations. This is based on the convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0 and standard deviation ranging from 0.1 to 1. Results are the average values of 20000 times.

Figure 3 shows the relationship between the Chi-Square test statistics and sample size given the standard deviation of the normal distribution. Each subplot represents a different standard deviation of the normal distribution included in the convolution. Controlling for standard deviation, the Chi-Square test statistics monotonically increase as the sample size increases. Once the sample size goes above 2000, the Chi-Square test statistics grow at a faster rate than those when the sample size is small.

Similarly, Figure 4 shows the relationship between MAD and sample size given the standard deviation of the normal distribution. Each subplot represents a different value of the standard deviation of the normal distribution included in the convolution. Figure 5 presents the relationship between MAD and standard deviation given the sample size. Each subplot represents a different sample size drawn in the simulation. There is a symmetric relation about the standard deviation 0.6 when the sample size is greater than 100.

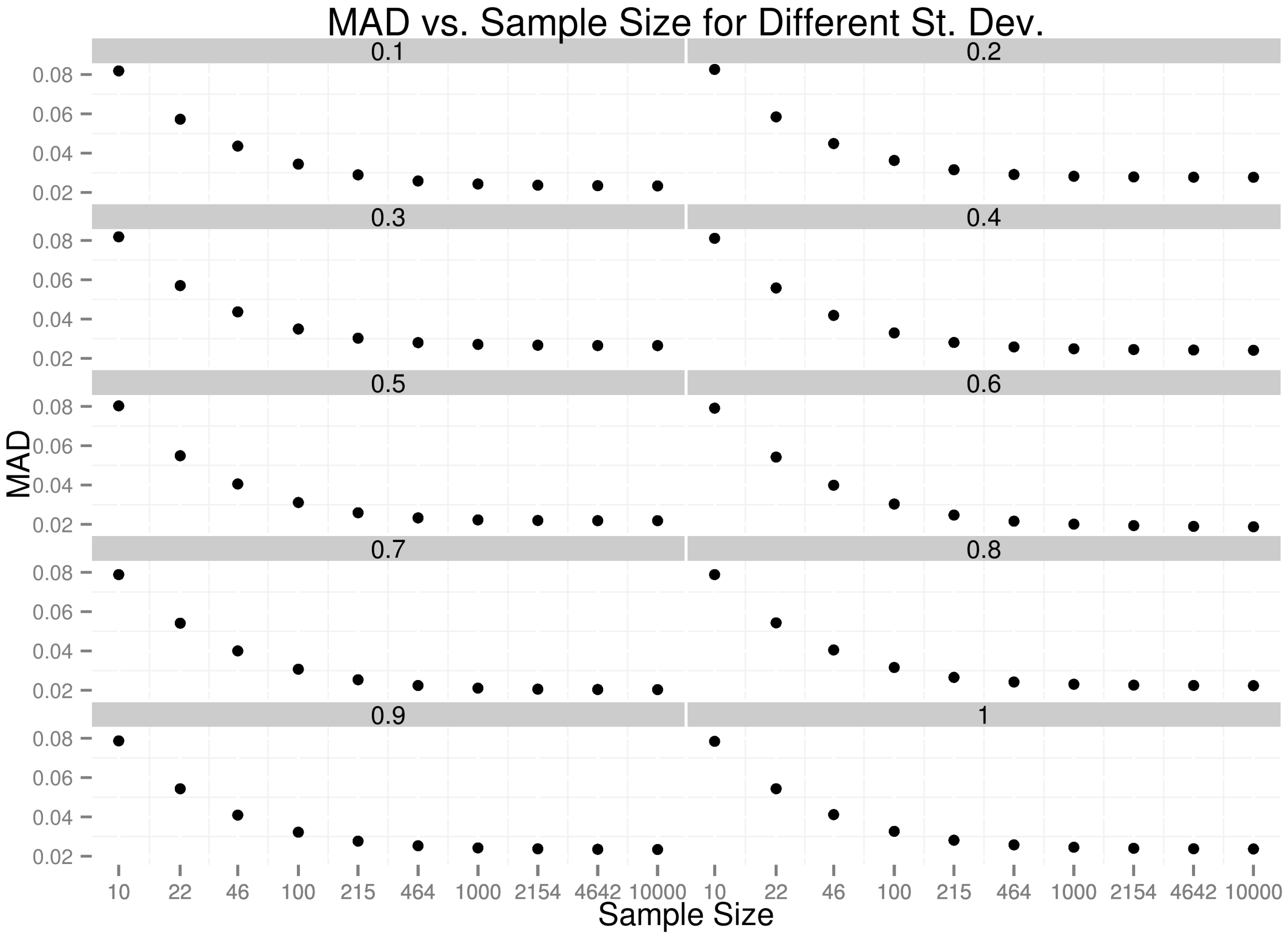
We also present the relationship among deviation measures, sample size, and the standard deviation of the normal distribution in a three-dimensional setting. Figure 6 shows the relationship for Chi-Square test statistics. Figure 7 is on the MAD values. The MAD values of various convolutions are all greater than 0.012 which means that the data values differ greatly from the theoretical Benford's distribution.

FIGURE 3. Chi-Square test statistics vs. sample size



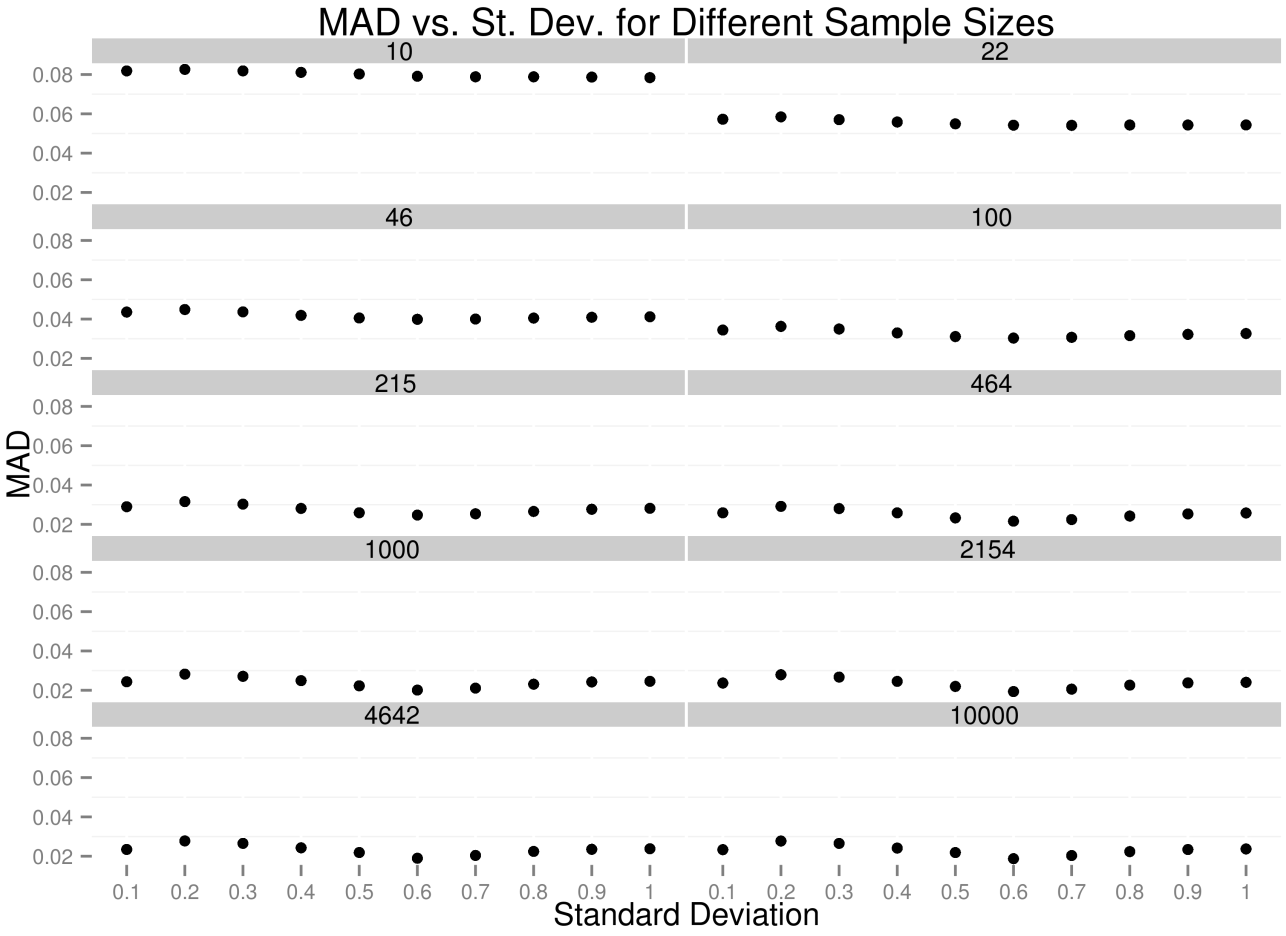
Note: We display the relationship between the Chi-Square test statistics and sample size given the standard deviation of the normal distribution. Each subplot represents a different standard deviation of the normal distribution included in the convolution. Sample sizes are displayed on the natural logarithmic scale.

FIGURE 4. MAD vs. sample size



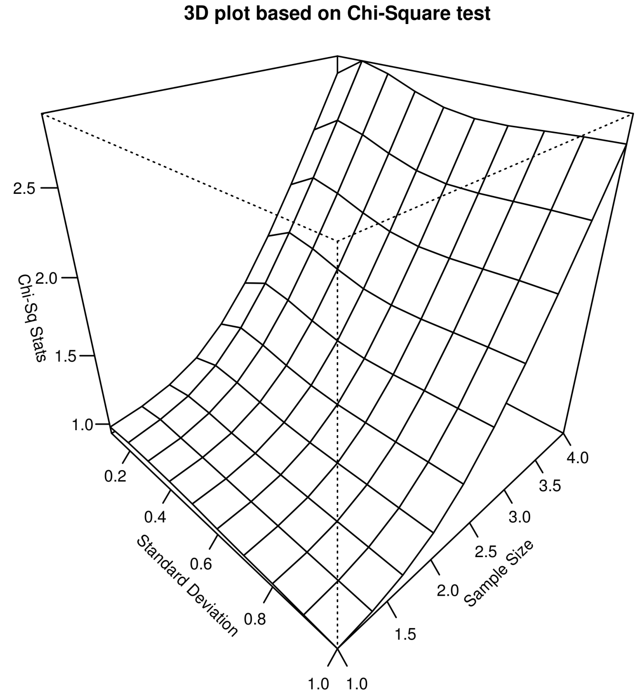
Note: We display the relationship between the MAD and sample size given the standard deviation of the normal distribution. Each subplot represents a different standard deviation of the normal distribution included in the convolution. The sample sizes are displayed on the natural logarithmic scale.

FIGURE 5. MAD vs. standard deviation



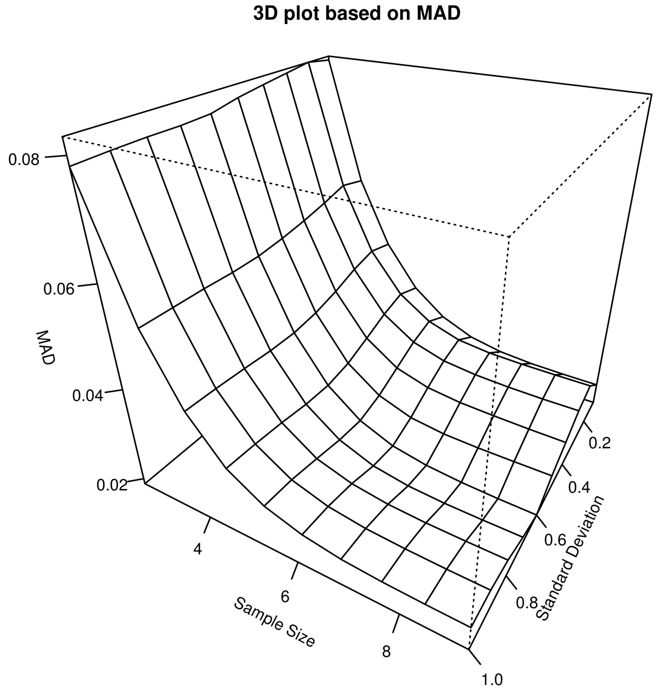
Note: We display the relationship between the MAD values and the standard deviation controlling for sample size. Each subplot represents a different sample size drawn in the simulation.

FIGURE 6. 3-D plot for Chi-Square test



Note: We display the relationship among the Chi-Square test statistic, sample size, and the standard deviation of the normal distribution in a 3-D plot. Both the sample size and the Chi-Square test statistic values are on the natural logarithmic scale.

FIGURE 7. 3-D plot for MAD



Note: We display the relationship among the Mean Absolute Deviation, sample size, and the standard deviation of the normal distribution in a 3-D plot. The sample size is on the natural logarithmic scale.

4.2. **Simulated replacement with a normal random variable.** Similar to Section 4.1, we perform simulations to explore and potentially identify the relationship among different deviation measures, the parameterization of the normal distribution, the number of data points used, and the proportion of the original data set replaced. We assume that the original data set conforms to Benford's Law.

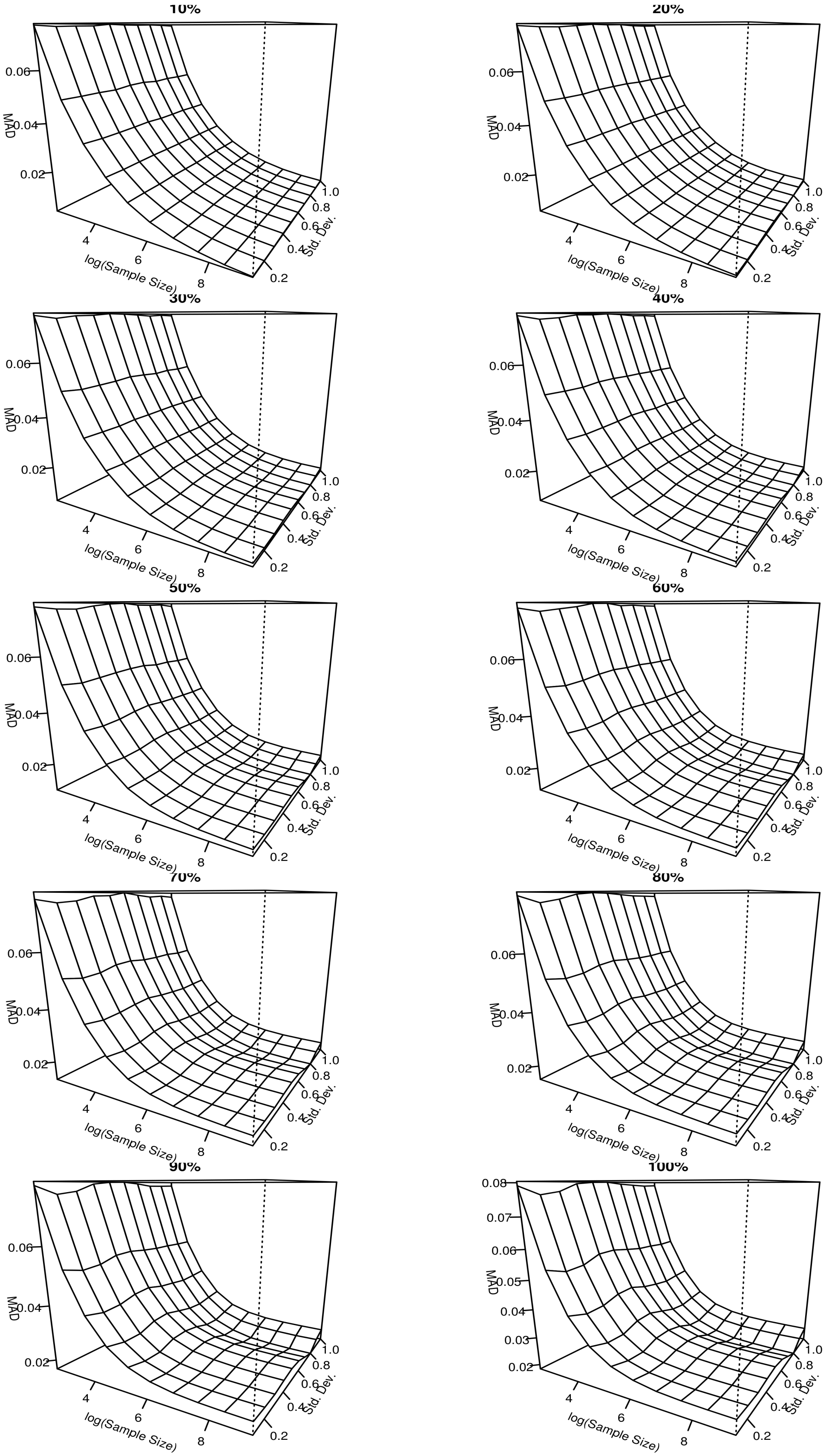
The data generating procedure is as follows:

(1) From Section 4.1, we know that the probability density function of the Benford's distribution is $f(x) = \frac{1}{x\ln(10)}$ and that the cumulative density function is $F(x) = \int_{0.1}^{x} f(x)\mathrm{d}x = \frac{\ln(x)-\ln(0.1)}{\ln(10)}$.
(2) We draw a random number u from the uniform distribution within the interval $[0, 1]$.
(3) Solve the value of x such that $F(x) = u$. It is clear that $x = e^{u\ln(10)+\ln(0.1)}$ and x is the random number drawn from the Benford's distribution.
(4) Repeat the three steps above for N numbers.
(5) For a fixed proportion m where $0 \leq m \leq 1$, draw $m * N$ values from the N numbers which follow the Benford's distribution.
(6) Randomly draw $(1-m) * N$ numbers from a normal distribution with mean μ and standard deviation σ.

Table 34 and 35 in the Appendices report the Chi-Square test-statistics and the associated p-values based on the simulated data, a proportion of which are replaced by data following a normal distribution.

Figure 8 shows the MAD values of the mixed data: $(1-m)\%$ of the data follow a Benford's distribution and $m\%$ follow a normal distribution. Each subplot shows a different proportion of the original data replaced. The standard deviation of the normal distribution range from 0.1 to 1. Across different standard deviations, as the sample size increases, the MAD values decrease. This could be because when the sample size is too small, the MAD values are not reflective of the conformity of the data to Benford's Law. As the sample size increases, the MAD values based on the larger sample provide a more accurate picture of whether a data set conforms to Benford's Law. Moreover, as expected, the higher the proportion of replacement is, the higher the MAD value becomes. When the proportion of replacement becomes 100%, the mixed data set becomes one whose values follow a normal distribution.

FIGURE 8. MAD based on simulations (partial replacement by values following a normal distribution



5. USE OF BOOTSTRAP FOR FRAUD DETECTION

5.1. **Background on bootstrap.** The bootstrap as a resampling method to estimate the sampling distribution of an estimator was first proposed by Efron in 1979 [Ef]. The basic idea is to repeatedly sample from one or several existing samples in order to make inference about the population. The bootstrap can be useful when the theoretical distribution of a statistic of interest is complicated. It is also attractive since inference on the population based on a sample is allowed when the closed-form of the underlying distribution is unknown.

There are various forms of the bootstrap method including non-parametric bootstrap, parametric bootstrap, balanced bootstrap, and smooth bootstrap. In this paper, we apply the basic form of the bootstrap - non-parametric bootstrap - with various sample sizes. Non-parametric bootstrap [Ef] treats the original data set as a whole and repeatedly draws a new observation with equal probability and with replacement.

Historically, the Internal Revenue Service (IRS) and other organizations have been interested in detecting fraud in tax and financial filings by large corporations. Methods based on Benford's Law have been used as screening tools for such fraud detection efforts. The IRS often aggregates huge amounts of data from tax forms and performs analyses in order to examine the extent to which the frequency distribution of the first digits deviates from Benford's Law.

We would like to find the minimum number of data points required for fraud detection based on Benford's Law. We do so by applying the bootstrap. Oftentimes the auditors may not be granted full access to the data on tax or other financial forms. The bootstrap method provides a unique way for auditors to circumvent the problem in that they no longer require the entire data set to make inference regarding the extent of non-conformity to Benford's Law. Also, knowing the number of data points required for fraud detection potentially reduces the costs of such detection efforts in terms of human resources and computational power.

5.2. **Simulation procedures.** In Section 3, we studied two distinct ways of manipulating a data set, namely, convolution and partial replacement of data points. For Benford's Law investigations, we are only interested in the frequency distribution of the leading digits of the resulting data values and not the final distribution of these data values. In this section, we artificially create data sets and apply the non-parametric bootstrap method on them. We compute several deviation measures including the p-values from the Chi-square tests and MAD based on the bootstrap and compare them with deviation measures based on the entire data sets. Specifically, we study how effectively we can detect data manipulation by bootstrapping with different sample sizes. The assumption is that the frequency distribution of the leading digits based on the original data conforms to Benford's Law.

The simulation steps are as follows:

1. To artificially create data for the following five scenarios – data following a Benford's distribution, data as a result of summing original values and new values following either a normal or a uniform distribution, partial replacement of the original data with values following either a normal or a uniform distribution.
2. To calculate various deviation measures based on the data including the Chi-Square test statistics, p-values based on the Chi-Square tests, and MADs.
3. To apply the bootstrap method with a subset of the data set and compute various deviation measures with a 95% confidence interval.
4. To compare the values of the deviation measures from bootstrapping with those from the original data set.

Parameters used in the bootstrap include the following.

- N, the size of the data set.
- n, the size of the bootstrap sample drawn from the data set, $n \leq N$.
- b, the number of bootstrapping iterations.
- μ and σ, mean and standard deviation of the normal distribution.
- u and v, the lower and upper bounds of a uniform distribution.
- m, the proportion of the data set replaced by data following another distribution. $0 \leq m \leq 1$.

In order to fully understand the effect of the bootstrap on the estimation of the deviation from Benford's Law, we vary the size of the bootstrap sample and that of the data set. Varying the size of the bootstrap sample allows us to observe whether there is a threshold number of data points which provides a reasonable estimate of the deviation from Benford's Law. Changing the size of the entire data set allows us to detect if the minimum number of data points required for the bootstrap varies with different sizes of the data sets.

We assume the size of the data set $N \in \{2500, 5000, 10000\}$ and the size of the bootstrap sample $n \in \{10, 22, 46, 100, 215, 464, 1000, 2154, 4641, 10000\}$ where applicable, and $n \leq N$ as those are the rounded numbers when splitting the number from 10 to 10000 into 10 equal intervals on the natural-logarithmic scale. We apply the same bootstrap sample sizes for all three different data sets. For simplicity, we set both the mean of the normal distribution and the lower bound of the uniform distribution to 0, namely, $\mu = 0$ and $u = 0$. We run 100 bootstrap iterations for each case.[3]

5.3. **Benford's distribution only.** We apply the non-parametric bootstrap method with different sample sizes on an artificially created data set. Data follow a Benford's distribution bounded by $[0.1, 1)$. This is the baseline case. Table 14 reports the mean MAD values and p-values of the Chi-Square tests when we apply the bootstrap with samples of various sizes. The MAD values are reported in percentages. As can be seen, the p-values do not indicate non-conformity to Benford's Law. This is as expected because the underlying data are drawn from a Benford's distribution. The mean MAD values go below the threshold of 1.2% based on [DN] only when the sample size becomes 1000 or larger. However, this is not indicative of the validity of the measure since it is only based on the mean MAD values.

Figure 9 shows the mean estimates of the MAD values and the 95% confidence intervals when we apply the bootstrap on data sets of various sizes. The black lines (middle) are the mean MAD estimates. The red lines (top and bottom) represent the 95% confidence intervals. The blue lines indicate the original values. Table 15 reports the smallest bootstrap sample sizes required such that the 95% confidence intervals include the actual MAD values based on the entire data sets.

These results imply that we need samples close to the entire data set for the bootstrap to produce a confidence interval which includes the true MAD values. However, these values may be overly theoretical in that in reality the MAD value of any data set may hardly approach a value that is very close to 0. On the other hand, the upper values of the 95% confidence intervals for all three data sets do not exceed the threshold MAD value of non-conformity recommended by [DN].

Table 16 reports the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values (in %) when we bootstrap with a different sample size in each row. Based on the MAD criteria by [DN], we require at least 464 out of 2500, 215 out of 5000, and 464 out of 10000 data points to have a lower bound of the 95% confidence interval smaller than 1.2% and thus not reject the claim that the data follows Benford Law. For a data set of size 5000 or larger, this bootstrap sample size represents about 5% of the entire data set. In other words, assuming that the data values conform to Benford's Law, the non-parametric bootstrap method

---

[3] We have checked the robustness of the results with 100 iterations as running the same bootstrap method with 200 iterations gives similar results.

only requires about 5% of the data not to reject the claim that the data conform to Benford's Law.

TABLE 14. MAD (in %) and p-values when applying bootstrap on data following Benford's distribution only

					MAD (in %)						
Size	Original	$N = 10$	$N = 22$	$N = 46$	$N = 100$	$N = 215$	$N = 464$	$N = 1000$	$N = 2154$	$N = 4642$	$N = 10000$
2500	0.3	10.6	7.7	5.4	3.5	2.5	1.5	1.0	0.6		
5000	0.3	9.6	6.1	4.7	2.7	2.1	1.6	1.1	0.7	0.4	
10000	0.2	9.6	6.3	5.6	4.3	2.6	1.6	1.0	0.7	0.4	0.0
					p-values						
Size	Original	$N = 10$	$N = 22$	$N = 46$	$N = 100$	$N = 215$	$N = 464$	$N = 1000$	$N = 2154$	$N = 4642$	$N = 10000$
2500	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
5000	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
10000	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: We report MAD (in %) and p-values from the Chi-Square test when we apply bootstrap on data sets of various sizes. Values in all three data sets follow a Benford's distribution bounded by $[0.1, 1)$. The "Size" column indicates the size of the entire data set. The "Original" column indicates the MAD values based on the entire data set. N refers to the size of a bootstrap sample. The number of bootstrapping iterations is 100.

TABLE 15. Minimum bootstrap sizes required to include the original MAD values (in %) in the 95% confidence intervals

Size	Bootstrap Size	Original MAD	Estimated MAD	Std. Dev.	Lower	Upper
2500	2154	0.28	0.55	0.14	0.27	0.84
5000	4642	0.25	0.41	0.11	0.19	0.63
10000	10000	0.21	0.32	0.07	0.17	0.47

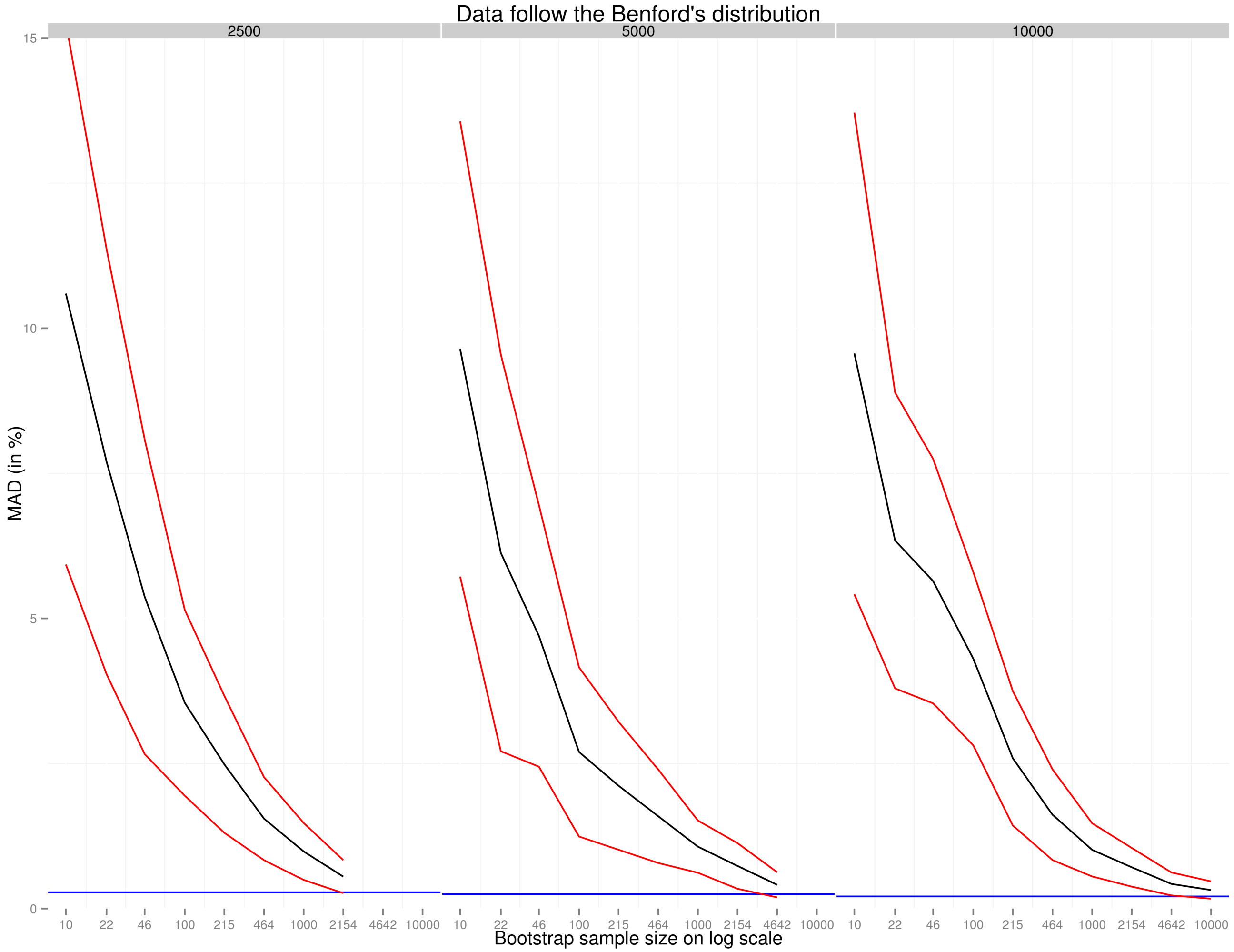
Note: We report the minimum bootstrap sample sizes such that the 95% confidence intervals of the estimated MAD values include the MAD values based on the entire data set. The MAD values and estimates are reported in percentages. These are reported in the "Bootstrap Size" column. The "Size" column indicates the overall size of the data set. The "Original MAD" column indicates the MAD value based on the entire data set. The "Estimated MAD" column reports the mean estimates of the MAD values based on 100 bootstrap iterations. The "Std. Dev." column reports the standard deviation of the bootstrap estimates. The "Lower" and "Upper" columns report the lower and upper bounds of the 95% confidence intervals based on the MAD estimates.

TABLE 16. Lower bound of the 95% confidence interval of MAD when bootstrapping on Benford distribution only

Data Set Size = 2500			
Sample Size	Mean	Std. Dev.	Lower
10	10.60	2.33	5.93
22	7.70	1.83	4.04
46	5.37	1.36	2.66
100	3.55	0.80	1.95
215	2.49	0.59	1.31
464	1.55	0.36	0.83
1000	0.99	0.25	0.50
2154	0.55	0.14	0.27
Data Set Size = 5000			
Sample Size	Mean	Std. Dev.	Lower
10	9.64	1.96	5.72
22	6.13	1.71	2.71
46	4.70	1.13	2.45
100	2.70	0.73	1.24
215	2.12	0.55	1.02
464	1.59	0.40	0.79
1000	1.07	0.22	0.62
2154	0.74	0.20	0.34
4642	0.41	0.11	0.19
Data Set Size = 10000			
Sample Size	Mean	Std. Dev.	Lower
10	9.57	2.07	5.42
22	6.34	1.27	3.79
46	5.64	1.05	3.54
100	4.31	0.75	2.81
215	2.59	0.58	1.43
464	1.62	0.39	0.84
1000	1.01	0.23	0.55
2154	0.71	0.17	0.38
4642	0.43	0.10	0.23
10000	0.32	0.07	0.17

Note: We report the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values when bootstrapping with a different sample size in each row. The MAD values and estimates are reported in percentages. The "Sample Size" column indicates the size of the bootstrap sample. The "Mean" column reports the mean estimates of the MAD values based on 100 times of bootstrapping. The "Std. Dev." column reports the standard deviation of the MAD values based on 100 times of bootstrapping. The "Lower" column reports the lower bound of the 95% confidence interval. [DN] states that data show non-conformity to the Benford's distribution when the MAD value is greater than 1.2%.

FIGURE 9. MAD (in %) and 95% confidence intervals based on bootstrap (Benford's only)



Note: We plot the MAD values in percentages (in black, top lines) and the 95% confidence intervals (in red, top and bottom lines) when we apply bootstrap on data sets whose values follow a Benford's distribution with the interval $[0.1, 1)$. The blue lines show the original MAD values for each data set.

5.4. **Sum of a Benford random variable and a normal random variable.** The second case we examine is that the distribution of the data points follows a convolution of a Benford's density and a normal density. The Benford's density is bounded by the interval $[0.1, 1)$ and the normal density is of mean 0 and standard deviation ranging from 0.1 to 1. Table 17 and Table 18 report the mean MAD values (in %) and the p-values based on the Chi-Square tests when we apply the bootstrap on data sets of different sizes, 2500, 5000, and 10000, respectively. Within each data set, we vary the bootstrap sample sizes in order to find the smallest size of bootstrap sample required to provide a 95% confidence interval which includes the MAD value of the entire data set.

According to the MAD criteria [DN], the MAD values can correctly alert us to any non-conformity to Benford's Law as both the original MAD values and all the MAD values from the bootstrap are greater than 1.2% in any of the specifications of the standard deviation across all three data sets.

Figure 10, Figure 11, and Figure 12 show the estimated MAD values in percentages (in black, middle lines) and the 95% confidence intervals (in red, top and bottom lines) based on the bootstrap method. The blue lines (horizontal) indicate the MAD values of the data sets as a whole. Across the three data sets which are of the size 2500, 5000, and 10000, the original MAD values as well as the mean estimates do not differ by much, implying that the bootstrap method produces consistent results across data sets of different sizes.

Figure 10 shows that in the worst case where the standard deviation of the normal distribution is 0.1, the bootstrap requires 464 data points in the sample such that the 95% confidence interval includes the original MAD value. In other specifications of the standard deviation, the minimum number of data points required in bootstrap is at most 215, which is roughly one tenth of the size of the original data set.

Figure 11 shows similar results where in the worst case, the bootstrap requires 464 data points, less than one tenth of the size of the original data to have the MAD values of the entire data set in the 95% confidence interval. In other cases, having a bootstrap sample of 215 is enough to produce a 95% confidence interval which includes the actual value of MAD. We observe similar results in Figure 12 where in the worst cases (the standard deviations are 0.3 and 0.9), it requires 1000 data points for the bootstrap sample, 10% of the data set, to have an MAD estimate within the 95% confidence interval.

Table 19 reports the mininum sizes of the bootstrap samples required such that the 95% confidence intervals include the MAD values based on the entire data sets.

In the Appendices, Table 32 reports the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values (in %) when we apply the bootstrap with different sample sizes on the data following a convolution of a Benford's density and a normal density. Except when we bootstrap 215 out of 5000 data points, the lower bounds of the 95% confidence interval of the estimated MAD values all exceed the threshold value 1.2%; the bootstrap method of all the sample sizes examined correctly identifies the non-conformity to Benford's Law.

TABLE 17. MAD (in %) when applying the bootstrap method (convolution of Benford and normal)

Data Set Size = 2500											
Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.3	13.5	9.4	5.9	4.7	3.9	3.1	2.5	2.4		
0.2	2.8	11.7	8.5	7.0	4.8	3.8	3.2	2.9	2.9		
0.3	2.7	13.9	10.1	7.3	5.4	3.6	3.3	2.6	2.7		
0.4	2.6	12.3	10.0	5.6	4.3	3.5	2.6	2.4	2.6		
0.5	2.1	9.4	7.7	5.5	3.8	3.0	2.7	2.4	2.4		
0.6	2.0	12.2	8.2	5.2	4.0	2.6	2.4	2.0	1.9		
0.7	2.0	9.6	7.7	4.7	4.0	2.9	2.7	2.3	2.3		
0.8	2.3	12.3	9.0	5.3	4.9	3.1	2.8	2.8	2.5		
0.9	2.5	11.3	7.9	5.9	4.4	3.1	2.6	2.7	2.5		
1.0	2.5	9.9	8.4	5.6	3.9	3.0	2.5	2.9	2.6		
Data Set Size = 5000											
Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.5	11.2	10.0	7.4	4.3	3.7	2.6	2.3	2.4	2.5	
0.2	2.8	11.5	9.2	7.4	4.7	4.4	3.5	2.8	2.9	3.1	
0.3	2.8	8.2	6.6	5.6	4.4	3.8	3.2	3.1	3.0	2.9	
0.4	2.3	10.8	7.1	6.2	3.4	2.8	2.6	2.6	2.6	2.5	
0.5	2.3	9.1	5.9	4.7	3.3	2.7	2.1	2.2	2.1	2.1	
0.6	1.9	9.5	7.4	5.8	4.6	3.2	2.3	2.1	1.9	1.9	
0.7	1.7	9.5	6.5	4.0	3.3	2.7	2.3	1.8	1.8	2.1	
0.8	2.5	8.3	5.4	5.1	4.6	3.7	2.8	2.4	2.3	2.5	
0.9	2.3	9.6	6.3	4.6	3.9	3.0	3.0	2.5	2.4	2.5	
1.0	2.3	11.5	7.6	6.1	3.9	2.4	2.5	2.6	2.4	2.5	
Data Set Size = 10000											
Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.4	10.5	6.9	6.1	4.8	3.5	3.1	2.8	2.6	2.4	2.4
0.2	2.6	12.6	8.2	5.6	3.3	3.0	3.0	2.9	2.7	2.5	2.8
0.3	2.6	14.5	9.2	7.7	5.4	4.4	4.0	3.2	3.0	2.8	2.8
0.4	2.5	12.2	7.8	6.5	4.3	3.6	3.5	2.5	2.5	2.5	2.4
0.5	2.0	11.2	7.5	6.1	4.0	3.2	3.3	2.4	2.4	2.4	2.2
0.6	2.1	10.5	7.6	5.7	3.5	3.4	3.1	2.5	2.1	2.1	2.0
0.7	1.9	12.0	7.2	6.2	3.5	3.0	2.7	2.5	2.3	2.2	2.2
0.8	2.4	12.1	7.6	7.2	3.9	3.7	3.0	2.8	2.4	2.3	2.3
0.9	2.3	11.6	6.4	5.2	4.0	3.8	3.2	2.8	2.6	2.4	2.3
1.0	2.3	9.7	6.7	6.0	4.4	4.2	3.2	2.9	2.6	2.5	2.5

Note: We report the MAD values (in %) based on applying the bootstrap method on data sets of various sizes. Within each data set, we vary the bootstrap sample sizes to obtain the MAD estimates. We report the mean estimates in the 3rd to 12th columns. The data in each of the three data sets follow a convolution of Benford's density with the interval $[0.1, 1)$ and a normal density with mean 0 and standard deviation varying from 0.1 to 1. The number of bootstrap iterations is 100.

TABLE 18. P-values from the Chi-Square tests when applying the bootstrap method (convolution of Benford and normal)

Data Set Size = 2500

Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.941	0.993	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	1.000	0.957	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	1.000	0.952	0.993	1.000	1.000	1.000	1.000	1.000	1.000		
0.4	1.000	0.890	0.962	0.999	1.000	1.000	1.000	1.000	1.000		
0.5	1.000	0.976	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
0.6	1.000	0.966	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
0.7	1.000	0.986	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
0.8	1.000	0.907	0.994	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	1.000	0.960	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	1.000	0.984	0.998	1.000	1.000	1.000	1.000	1.000	1.000		

Data Set Size = 5000

Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.972	0.972	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	1.000	0.980	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.4	1.000	0.986	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.5	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.6	1.000	0.992	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.7	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	1.000	0.997	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	1.000	0.976	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

Data Set Size = 10000

Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1.000	0.904	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.000	0.938	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	0.975	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	0.951	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.000	0.974	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.000	0.921	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	0.858	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	1.000	0.945	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	1.000	0.973	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

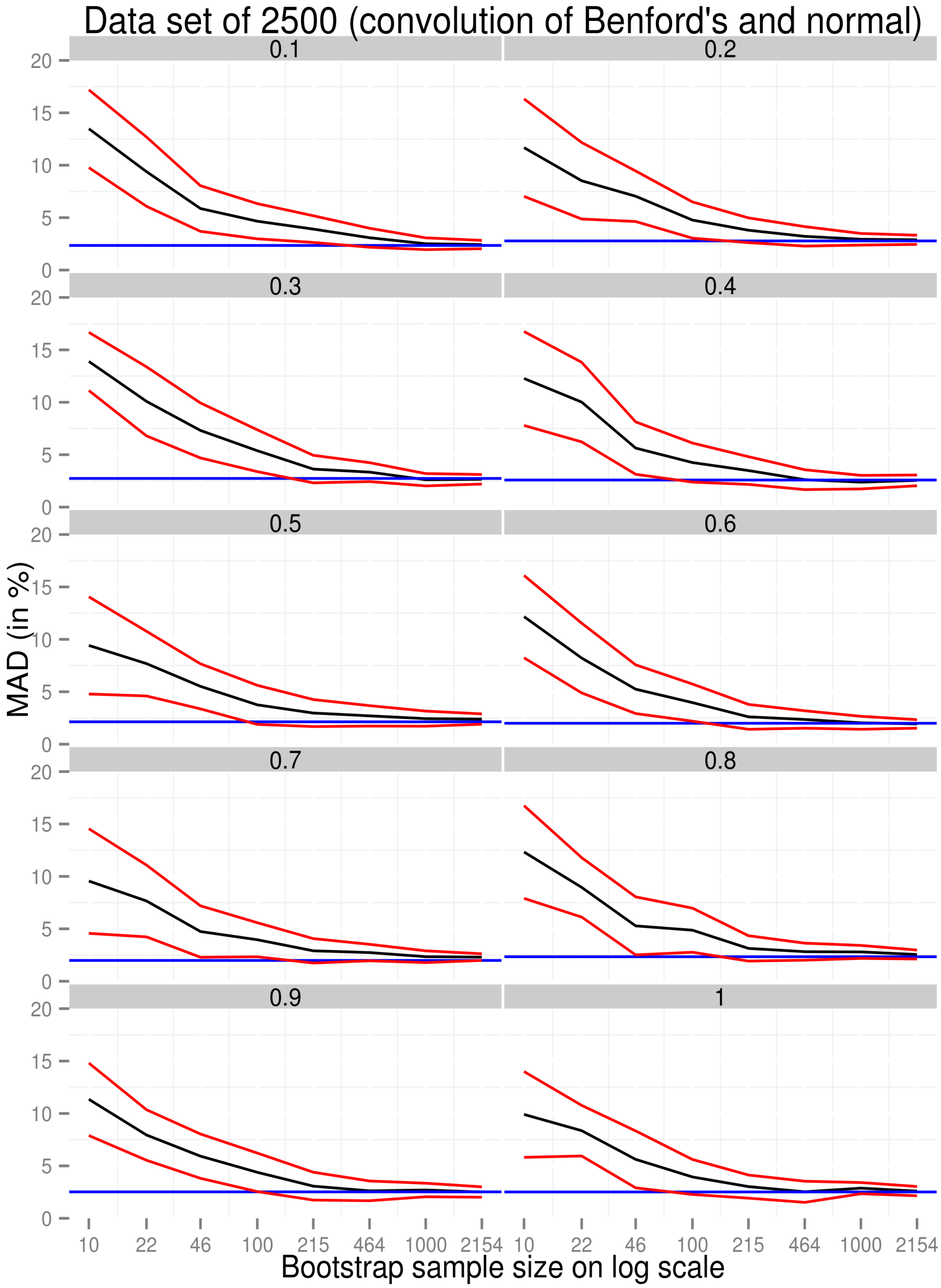
Note: We report the p-values (in %) from the Chi-Squared test based on bootstrapping on data of various sizes. Within each data set, we vary the bootstrap sample sizes to obtain the p-value estimates. We report the mean p-values in the 3rd to 12th columns. The data in each of the three data sets follow a convolution of Benford's density with the interval $[0.1, 1)$ and a normal density with mean 0 and standard deviation varying from 0.1 to 1. The number of bootstrap iterations is 100.

TABLE 19. Minimum sizes of bootstrap samples required to include the original MAD values in the 95% confidence intervals (convolution of Benford and normal)

Std. Dev.	2500	5000	10000
0.1	464	215	215
0.2	215	464	100
0.3	215	100	1000
0.4	100	100	215
0.5	100	100	215
0.6	215	464	100
0.7	215	100	100
0.8	215	464	100
0.9	215	46	1000
1.0	100	100	464

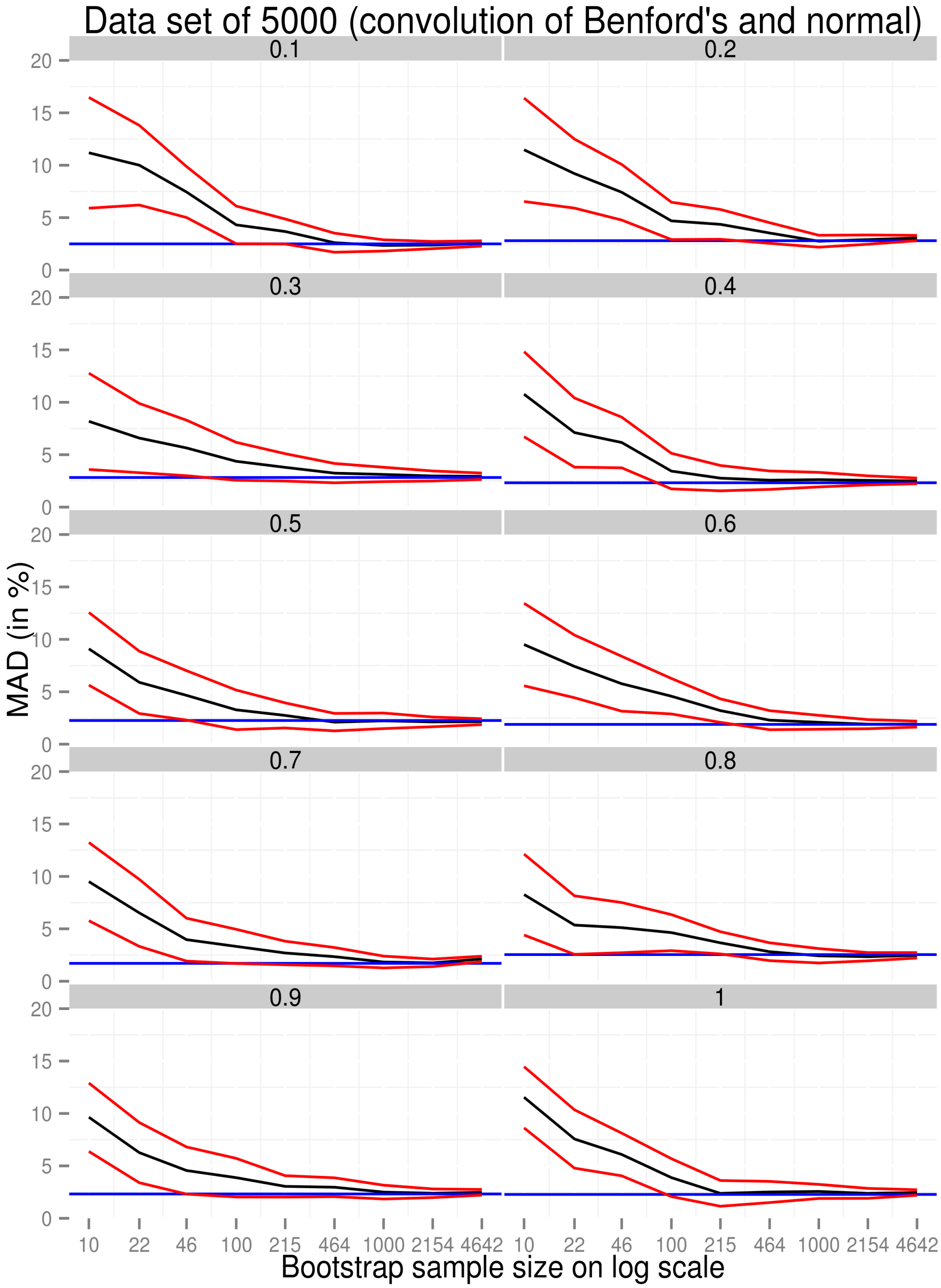
Note: We report the mininum sizes of the bootstrap samples required such that the 95% confidence intervals include the MAD values based on the entire data sets. The data in each of the three data sets follow a convolution of Benford's density with the interval $[0.1, 1)$ and a normal density with mean 0 and standard deviation varying from 0.1 to 1.

FIGURE 10. MAD (in %) and 95% confidence intervals of 2500 data points (convolution of Benford and normal)



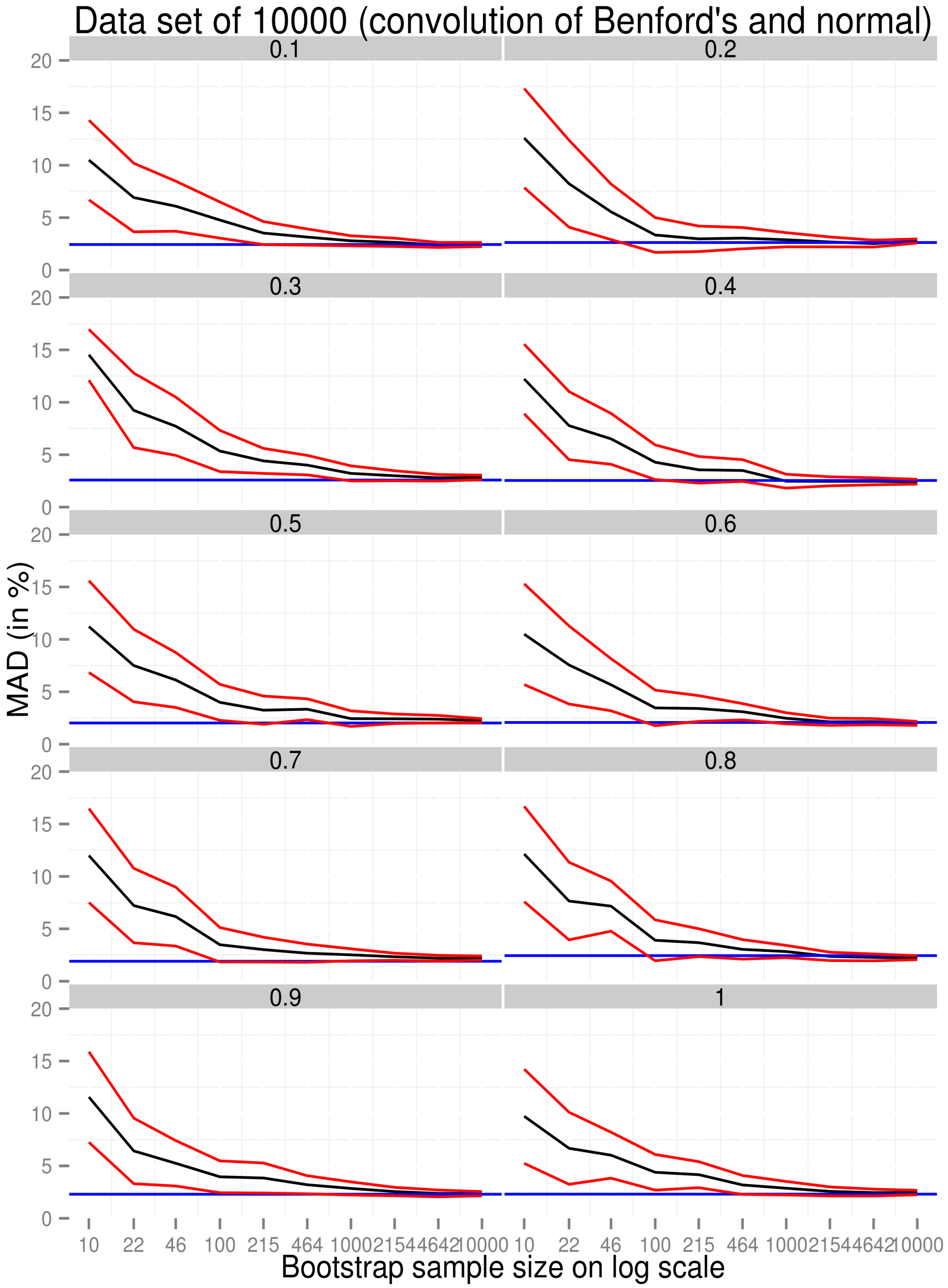
Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence interval (in red, top and bottom lines) when we apply the bootstrap method on data which follow a convolution of a Benford's density with the interval $[0.1, 1)$ and a normal density with mean 0. Each subplot shows a different standard deviation of the normal distribution. The blue lines (horizontal) show the original MAD values for each data set. The data set is of siez 2500.

FIGURE 11. MAD (in %) and 95% confidence intervals based of 5000 data points (convolution of Benford and normal)



Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence interval (in red, top and bottom lines) when we apply the bootstrap method on data which follow a convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0. Each subplot shows a different standard deviation of the normal distribution. The blue (horizontal) lines show the original MAD values for each data set. The data set is of size 5000.

FIGURE 12. MAD (in %) and 95% confidence intervals based of 10000 data points (convolution of Benford and normal)



Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence interval (in red, top and bottom lines) when we apply the bootstrap method on data which follow a convolution of a Benford's density bounded by $[0.1, 1)$ and a normal density with mean 0. Each subplot shows a different standard deviation of the normal distribution. The blue lines (horizontal) show the original MAD values for each data set. The data set is of size 10000.

5.5. **Sum of a Benford random variable and a uniform random variable.** Similarly, we apply the bootstrap with different sample sizes on the convolution of a Benford's density and a uniform density. The uniform density has a lower bound of 0 and an upper bound ranging from 0.1 to 1. Table 20 and Table 21 report the mean MAD values and the p-values based on bootstrapping of different sample sizes. We also observe that all the MAD values are above 0.012 and bring us to attention of possible non-conformity to Benford's Law.

Figure 13, Figure 14, and Figure 15 show the estimated MAD values (in black, middle lines) and the 95% confidence intervals (in red, top and bottom lines) from bootstrapping as well as whether the confidence intervals include the actual MAD values of the data sets (in blue, horizontal lines).

Similar to what is observed in Section 5.4, across all three data sets which are of the size 2500, 5000, and 10000, the original MAD values as well as the mean estimates do not differ much, implying that the bootstrap method produces consistent results across data sets of various sizes.

As to the number of data points required such that the 95% confidence interval includes the MAD value based on the entire data set, we examine the three data sets individually.

Figure 13 shows that in the worst case where the upper bound of the uniform distribution is 0.1, the bootstrap method requires 464 data points in the sample such that the 95% confidence interval includes the original MAD value. In other specifications of the upper bound, the number of data points required for applying bootstrapping is 215, which is roughly one tenth of the size of the original data set.

Figure 14 shows similar results. In the worst case in which the upper bound of the normal distribution is 0.6, the bootstrap method requires 464 data points, less than one tenth of the size of the original data. In other cases, having a bootstrap sample of 215 is enough to produce a 95% confidence interval which includes the actual value of MAD.

We observe similar results in Figure 15 where in the worst cases (the upper bounds are 0.3 and 1), it requires 1000 data points for the bootstrap sample, 10% of the data set. In other cases, 215 data points are required.

Table 33 in the Appendices reports the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values (in %) when we apply the bootstrap with different sample sizes on data following a convolution of a Benford density and a uniform density. The lower bounds of the 95% confidence interval of the estimated MAD values all exceed the threshold value 1.2% according to [DN]; the bootstrap method of all the sample sizes examined all correctly identifies the non-conformity to the Benfordness of the data set.

As described in Section 5.3, for a data set of size 5000 or more, we can apply the bootstrap method with only 5% of the data to correctly rule out the non-Benfordness given that the data set conforms to Benford's Law. Assuming that a data set, if manipulated, would be the result of a convolution of a Benford density and either a normal density or a uniform density, we only require 5% of the data set to determine whether the data set conforms to Benford's Law. This greatly reduces the requirement on the access to data when an auditor attempts to determine the Benfordness of a data set and makes it possible for such a test even if the auditor does not have access to the entire data set.

TABLE 20. MAD (in %) when applying the bootstrap method (convolution of Benford and uniform)

Data Set Size = 2500

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.7	8.9	8.1	4.7	4.7	3.6	3.0	3.0	2.7		
0.2	3.7	9.8	9.1	6.1	5.3	4.8	4.3	4.0	4.0		
0.3	4.8	12.0	10.4	6.7	7.0	5.9	5.3	5.1	5.0		
0.4	5.2	11.7	10.0	7.0	6.9	6.2	5.8	5.7	5.5		
0.5	5.5	13.6	10.2	7.2	6.6	5.9	6.1	5.9	5.6		
0.6	5.7	11.6	9.9	7.5	6.7	6.2	6.2	6.1	5.7		
0.7	5.7	12.1	11.5	7.5	6.5	5.9	6.0	6.0	5.7		
0.8	5.2	12.8	8.8	7.0	6.7	6.0	5.9	5.7	5.4		
0.9	4.7	12.3	8.8	6.5	6.0	5.6	5.6	5.3	5.2		
1.0	5.5	13.4	10.0	7.0	6.3	6.1	6.0	5.8	5.6		

Data Set Size = 5000

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.5	9.6	6.2	4.8	3.4	2.7	2.7	2.6	2.6	2.6	
0.2	3.7	11.6	7.3	5.8	5.0	4.5	3.8	3.7	3.7	3.8	
0.3	4.7	10.4	8.1	8.1	7.0	5.3	4.5	4.4	4.7	4.7	
0.4	5.2	10.4	8.9	7.6	6.8	5.7	4.8	4.8	5.1	5.2	
0.5	5.6	11.2	9.5	8.4	6.7	5.8	5.1	5.0	5.5	5.5	
0.6	5.5	11.8	10.7	8.5	6.8	5.9	5.2	5.1	5.5	5.6	
0.7	5.4	12.0	11.2	8.3	6.6	5.8	5.0	5.0	5.5	5.4	
0.8	5.3	11.0	11.2	8.5	7.0	5.7	5.0	4.5	5.0	5.1	
0.9	5.0	10.8	10.3	7.4	6.3	4.9	4.4	4.7	4.9	4.9	
1.0	5.3	10.6	8.5	6.9	5.8	4.8	4.9	5.1	5.4	5.4	

Data Set Size = 10000

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	2.5	9.3	6.3	6.7	5.0	3.2	3.0	2.7	2.5	2.4	2.5
0.2	3.8	9.3	7.7	6.4	5.4	4.4	4.4	4.2	3.9	3.8	3.9
0.3	5.0	8.7	7.7	6.6	6.0	5.5	4.7	4.6	4.5	4.5	4.8
0.4	5.1	10.2	9.5	7.4	6.0	5.3	5.0	5.1	5.1	5.1	5.3
0.5	5.4	9.3	8.3	8.0	6.2	5.5	5.6	5.5	5.4	5.3	5.4
0.6	5.7	10.5	8.9	8.3	6.4	5.7	5.7	5.3	5.3	5.3	5.4
0.7	5.3	10.8	8.7	8.2	6.4	5.6	5.6	5.2	5.1	5.1	5.2
0.8	4.7	10.3	9.1	8.2	6.2	5.5	5.3	4.7	4.6	4.7	4.7
0.9	5.0	10.8	8.8	7.8	6.4	5.4	4.9	4.6	4.8	4.8	4.9
1.0	5.4	11.5	9.3	6.7	5.7	5.8	5.2	5.2	5.3	5.3	5.4

Note: We report the MAD values (in %) based on bootstrap on data sets of various sizes. Within each data set, we vary the bootstrap sample sizes to obtain the MAD estimates. We report the mean estimates in the 3rd to 12th columns. The data in each of the three data sets follow a convolution of Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound 0 and an upper bound varying from 0.1 to 1. The number of bootstrap iterations is 100.

TABLE 21. P-values from the Chi-Square tests when applying the bootstrap method (convolution of Benford and uniform)

Data Set Size = 2500

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.998	0.997	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	1.000	0.993	0.995	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	1.000	0.964	0.993	1.000	1.000	1.000	1.000	1.000	1.000		
0.4	1.000	0.958	0.991	0.999	1.000	1.000	1.000	1.000	1.000		
0.5	1.000	0.816	0.975	0.997	1.000	1.000	1.000	1.000	1.000		
0.6	1.000	0.937	0.989	0.999	1.000	1.000	1.000	1.000	1.000		
0.7	1.000	0.960	0.961	0.999	1.000	1.000	1.000	1.000	1.000		
0.8	1.000	0.874	0.990	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	1.000	0.913	0.987	0.999	1.000	1.000	1.000	1.000	1.000		
1.0	1.000	0.924	0.954	0.998	1.000	1.000	1.000	1.000	1.000		

Data Set Size = 5000

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.979	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	1.000	0.979	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	1.000	0.972	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.4	1.000	0.982	0.996	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
0.5	1.000	0.956	0.978	0.994	1.000	1.000	1.000	1.000	1.000	1.000	
0.6	1.000	0.935	0.980	0.996	1.000	1.000	1.000	1.000	1.000	1.000	
0.7	1.000	0.919	0.968	0.995	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	1.000	0.943	0.969	0.995	0.999	1.000	1.000	1.000	1.000	1.000	
0.9	1.000	0.921	0.961	0.995	0.999	1.000	1.000	1.000	1.000	1.000	
1.0	1.000	0.989	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	

Data Set Size = 10000

Upper Bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.000	0.979	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.000	0.987	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.000	0.963	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.000	0.955	0.989	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.000	0.969	0.991	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.9	1.000	0.943	0.984	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000
1.0	1.000	0.918	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

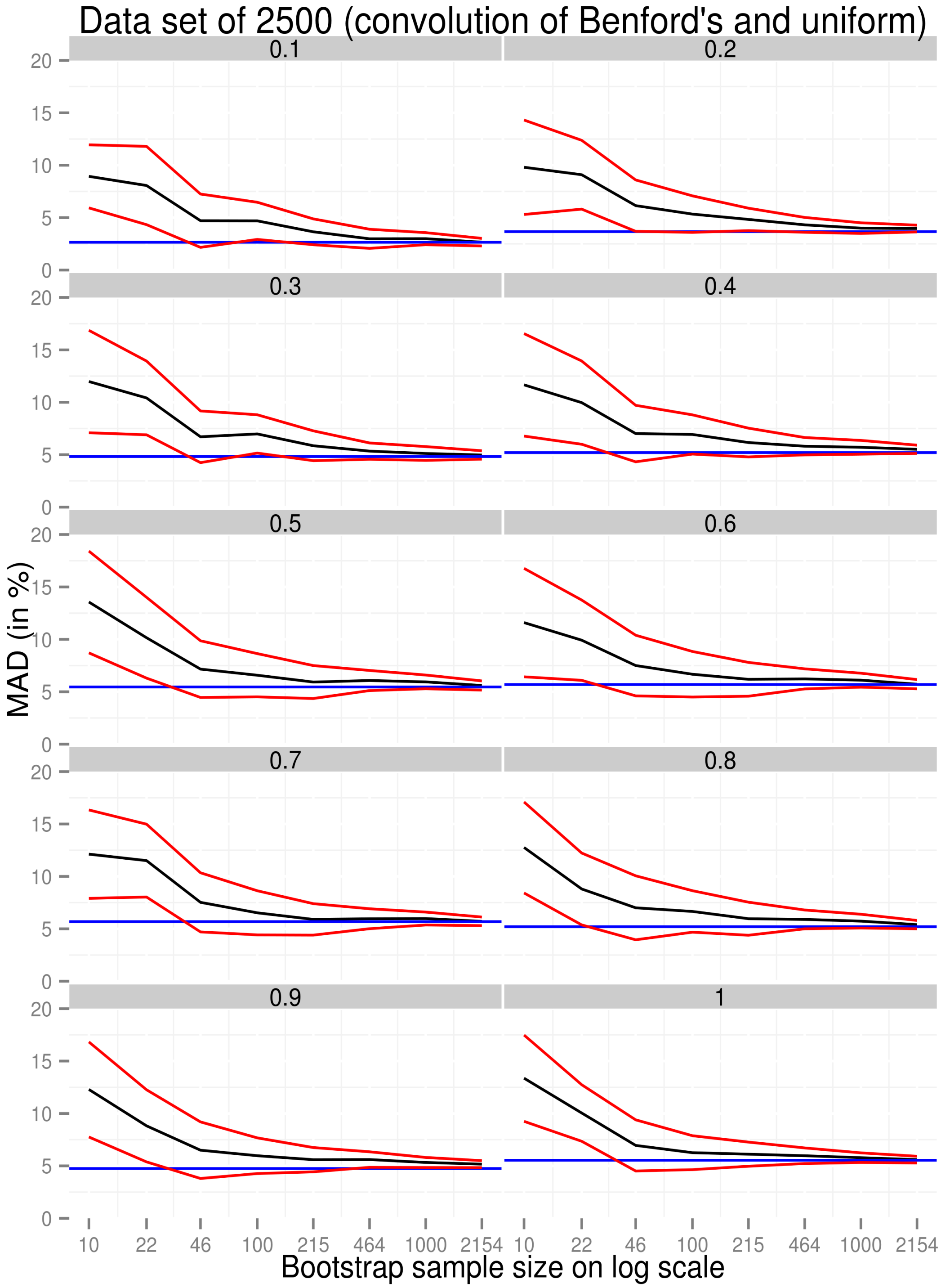
Note: We report the p-values (in %) from the Chi-Squared test based on bootstrap on data sets of various sizes. Within each data set, we vary the bootstrap sample sizes to obtain the p-value estimates. We report the mean p-values in the 3rd to 12th columns. The data in each of the three data sets follow a convolution of Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1. The number of bootstrap iterations is 100.

TABLE 22. Minimum sizes of bootstrap samples required to include the original MAD values in the 95% confidence intervals (convolution of Benford and uniform)

Upper Bound	2500	5000	10000
0.1	46	46	215
0.2	100	46	100
0.3	46	22	10
0.4	46	46	46
0.5	46	46	10
0.6	46	46	22
0.7	46	46	22
0.8	46	46	100
0.9	46	46	100
1.0	46	46	46

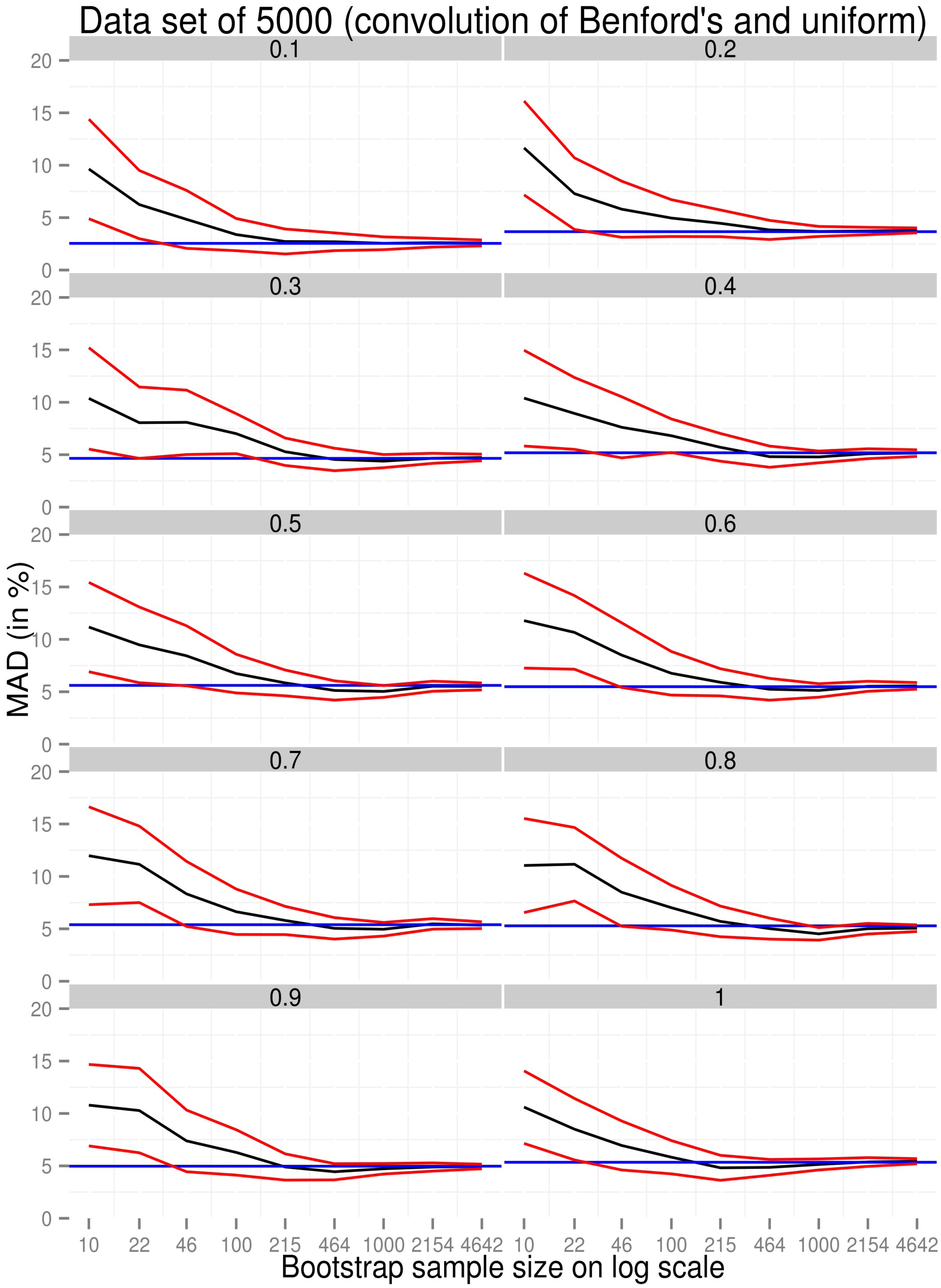
Note: We report the mininum sizes of the bootstrap samples required such that the 95% confidence intervals include the MAD values based on the entire data sets. The data in each of the three data sets follow a convolution of Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1.

FIGURE 13. MAD (in %) and 95% confidence intervals of 2500 data points (convolution of Benford and uniform)



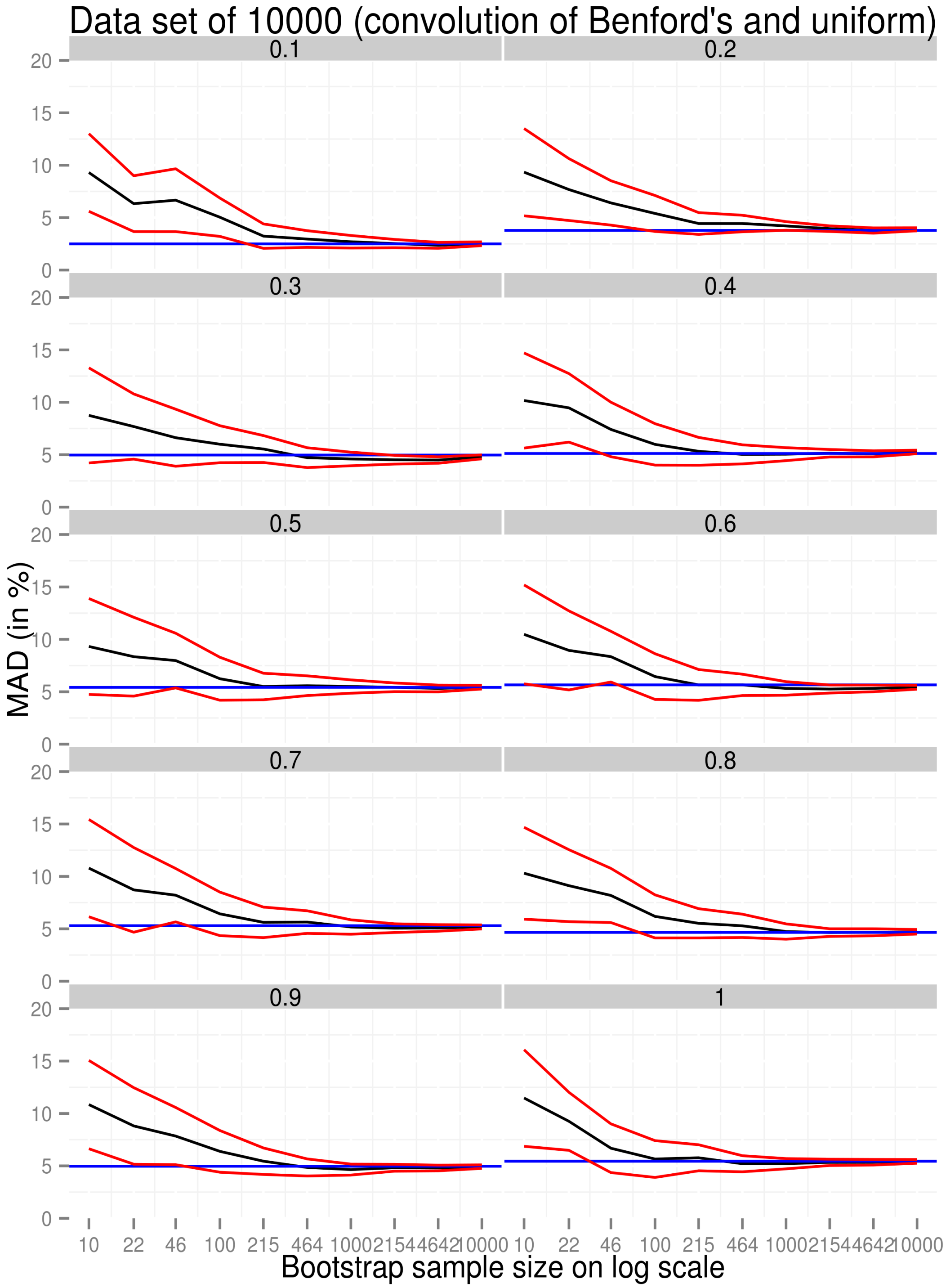
Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence intervals (in red, top and bottom lines) when we apply bootstrap on data which follow a convolution of a Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1 as shown by each subplot. The blue lines (horizontal) show the original MAD values for each data set. The data set is of size 2500.

FIGURE 14. MAD (in %) and 95% confidence intervals of 5000 data points (convolution of Benford and uniform)



Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence intervals (in red, top and bottom lines) when we apply bootstrap on data which follow a convolution of a Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1 as shown by each subplot. The blue lines (horizontal) show the original MAD values for each data set. The data set is of size 5000.

FIGURE 15. MAD (in %) and 95% confidence intervals of 10000 data points (convolution of Benford and uniform)



Note: We plot the MAD values in percentages (in black, middle lines) and the 95% confidence intervals (in red, top and bottom lines) when we apply bootstrap on data following a convolution of a Benford's density bounded by $[0.1, 1)$ and a uniform density with a lower bound of 0 and an upper bound ranging from 0.1 to 1 as shown by each subplot. The blue lines (horizontal) show the original MAD values for each data set. The data set is of size 10000.

5.6. **Replace with a normal random variable.** We also apply the bootstrap with different sample sizes on data sets in which originally all values follow a Benford's distribution bounded by $[0.1, 1)$ and are partially replaced by values following a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1. We apply the same procedure on these three data sets to make sure that the patterns are consistent across data sets of different sizes.

Table 23 reports the original MAD values for data sets of different sizes and standard deviations. Table 37 and Table 38 in the Appendices report the mean MAD values (in $\%$) and the p-values based on various bootstrap sample sizes. For a data set with 2500 values, based on the non-conformity MAD threshold of 0.012, when the proportion of replacement is no more than 30%, the MAD values computed based on data sets as a whole cannot identify non-conformity. When the proportion of replacement is 40%, the MAD values based on the entire data sets identify non-conformity in 3 out of 10 specifications of the normal distribution. When the proportion of replacement is 50% (60%), the MAD values alone can identify 6 (9) out of 10 specifications. Only when the proportion of replacement reaches 80% and beyond, the MAD values alone can identify non-conformity by having values greater than 0.012. As to a data set of size 5000, the MAD values alone do not identify non-conformity at all when the proportion of replacement is no more than 40%. As the proportion of replacement increases, the MAD values work better; however, MAD correctly identifies non-conformity when the proportion of replacement is 80% or more. Similarly, for a data set of size 10000, the MAD values correctly spot non-conformity 10 out of 10 cases when the proportion of replacement is 60% or more.

Table 24 reports the mininum sizes of the bootstrap samples required such that the 95% confidence intervals include the MAD values of the entire data sets. In general, across different proportions of replacement, the minimum size of the bootstrap samples increases with the proportion of replacement. This is reasonable in that as the proportion of replacement increases, the data deviate more from Benford's Law since the underlying distribution deviates more from the Benford's distribution.

Figure 16, Figure 17, and Figure 18 show the mean MAD values (in $\%$) based on the data sets partially replaced by values following a normal distribution. The original sizes of the data sets are 2500, 5000, and 10000, respectively. Each subplot shows a different proportion of the data replaced with all ten specifications of the standard deviation of the normal distribution. There is more dispersion of the MAD values when the data set is of size 2500. This is due to the fact that a data set of size

2500 is relatively small and there is more variation when sampling a subset of such a data set.

Unlike adding either a normal random variable or a uniform random variable to a Benford random variable, it is more difficult to detect partial replacement of data following a Benford's distribution with values following a normal distribution. Using the bootstrap method, for a data set of size 2500, we cannot effectively detect non-conformity to Benford's Law when the proportion of replacement is 0.6 or smaller. When the proportion of replacement is 0.7, we can use 100 data points (4% of the data set) to detect non-Benfordness except the case in which the standard deviation of the normal distribution is 0.3. When the proportion of replacement is 0.8 or higher, using 4% of the data allows us to detect non-conformity to Benford's Law at a significance level of 5%.

For a data set of size 5000, when the proportion of replacement is 0.7 or smaller, the bootstrap method does not allow us to detect non-Benfordness effectively. Oftentimes, the lower bounds of the 95% confidence intervals generated are smaller than 0.12%, the threshold value of non-conformity proposed by [DN]. When we use around 9.2% of the data set (464 out of 5000 data points), 80% of which are replaced by values following a normal distribution, the method detects 8 out of 10 different specifications of the normal distribution. It does not work effectively when the standard deviation of the normal distribution is 0.2 or 0.3. As the proportion of replacement reaches 0.9, we can use 9% of the data set to effectively detect non-Benfordness of the data.

As the size of the data set becomes 10000, the bootstrap method allows us to use 10% of the data to detect non-conformity to Benford's Law when the proportion of replacement is 0.8 or higher.

TABLE 23. MAD (in %) for data sets partially replaced by values following a normal distribution

Data Set Size = 2500

% replaced	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10	0.39	0.49	0.29	0.38	0.47	0.32	0.63	0.41	0.40	0.49
20	0.54	0.62	0.76	0.72	0.41	1.18	0.68	0.62	0.49	0.78
30	0.89	0.96	0.88	0.88	1.09	0.86	0.87	0.57	0.88	1.14
40	1.45	0.72	0.85	1.12	1.30	1.25	0.61	0.83	0.75	0.58
50	1.45	1.20	1.35	1.43	1.37	1.12	0.94	1.03	0.96	1.30
60	1.62	1.58	1.53	1.07	1.77	1.50	1.41	1.25	1.37	1.57
70	1.75	1.75	1.67	1.63	1.54	1.48	1.54	1.34	1.32	1.66
80	2.24	1.77	1.57	2.18	1.88	1.98	1.32	1.52	1.94	1.86
90	2.27	1.98	2.13	2.13	1.91	2.22	1.63	1.47	2.10	2.25
100	2.89	2.21	2.29	2.19	2.58	2.34	2.15	2.21	2.34	2.42

Data Set Size = 5000

% replaced	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10	0.23	0.36	0.41	0.30	0.44	0.58	0.51	0.28	0.34	0.28
20	0.69	0.59	0.68	0.65	0.44	0.52	0.74	0.24	0.53	0.83
30	0.63	0.65	0.71	0.99	0.75	0.89	0.77	0.63	0.82	0.68
40	0.96	0.99	0.87	0.99	1.18	0.82	1.04	1.01	0.75	0.59
50	1.16	1.14	1.32	1.51	1.20	1.48	1.02	0.83	1.26	1.02
60	1.49	1.41	1.78	1.34	1.24	1.55	1.11	1.19	1.06	1.36
70	1.72	1.64	1.71	1.57	1.50	1.50	1.53	1.12	1.52	1.62
80	2.13	2.04	1.53	1.76	1.64	1.91	1.63	1.44	1.53	1.78
90	1.90	2.29	1.86	2.25	2.24	2.01	1.60	1.43	1.69	2.19
100	2.34	2.09	2.19	2.17	2.22	1.76	2.01	1.71	2.16	2.32

Data Set Size = 10000

% replaced	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10	0.29	0.38	0.30	0.47	0.36	0.29	0.33	0.21	0.31	0.40
20	0.32	0.51	0.58	0.61	0.54	0.62	0.64	0.55	0.46	0.61
30	0.44	0.65	0.82	0.55	0.62	0.79	0.64	0.62	0.61	0.73
40	0.95	0.93	0.83	0.88	0.78	0.79	0.77	0.96	0.76	0.77
50	1.25	1.25	1.10	1.08	1.06	0.99	0.98	0.97	1.23	1.15
60	1.39	1.41	1.34	1.27	1.34	1.47	1.23	1.27	1.32	1.33
70	1.48	1.47	1.58	1.54	1.79	1.62	1.39	1.26	1.38	1.55
80	1.80	1.73	1.63	1.89	1.75	1.59	1.71	1.46	1.94	1.69
90	2.04	2.14	2.01	2.06	2.03	2.21	1.69	1.71	1.97	2.02
100	2.35	2.20	2.27	2.31	2.34	2.35	2.06	1.90	2.30	2.22

Note: We report the MAD values (in %) of the data sets whose values follow a Benford's distribution with the interval $[0.1, 1)$ partially replaced values following a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1.

TABLE 24. The minimum bootstrap sample sizes for MAD to be within the 95% confidence interval (replaced by uniform)

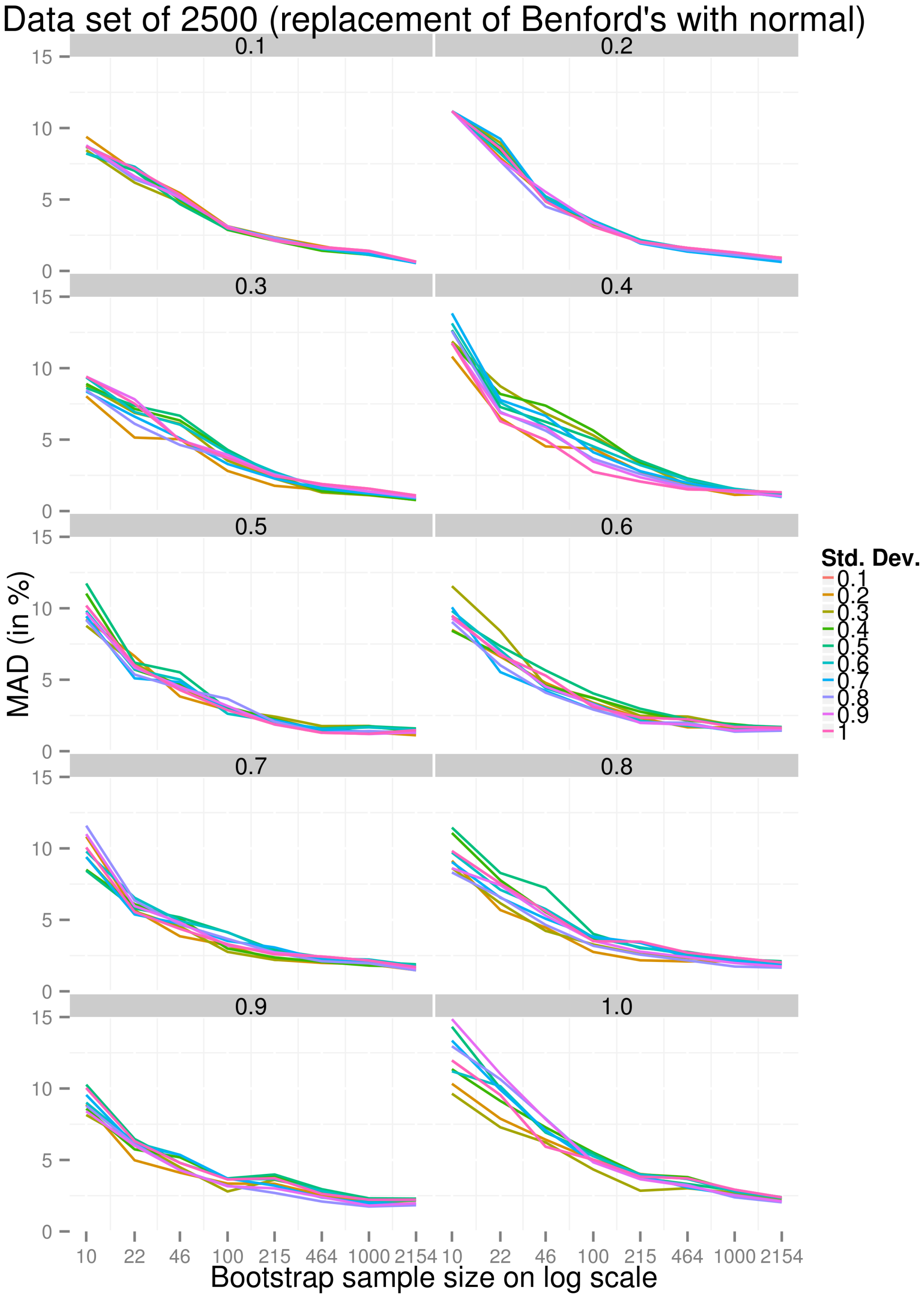
% replaced	Std. Dev.	2500	5000	10000
10	0.1	2154	4642	4642
10	0.2	2154	4642	4642
10	0.3		4642	4642
10	0.4	2154		4642
10	0.5	2154	1000	4642
10	0.6	2154	1000	4642
10	0.7	2154	1000	4642
10	0.8	2154	4642	4642
10	0.9	2154	2154	4642
10	1.0	2154	4642	2154
20	0.1	2154	1000	10000
20	0.2	1000	2154	10000
20	0.3	1000	1000	1000
20	0.4	1000	4642	464
20	0.5	2154		2154
20	0.6	215	4642	1000
20	0.7	464	1000	2154
20	0.8	2154		2154
20	0.9	2154	1000	4642
20	1.0	464	464	2154
30	0.1	2154	1000	10000
30	0.2	215	1000	2154
30	0.3	464	2154	1000
30	0.4	464	2154	1000
30	0.5	464	2154	2154
30	0.6	1000	2154	464
30	0.7	1000	2154	2154
30	0.8		2154	4642
30	0.9	1000	1000	2154
30	1.0	464	1000	1000
40	0.1	100	1000	1000
40	0.2	1000	464	2154
40	0.3	1000	1000	2154
40	0.4	1000	1000	464
40	0.5	1000	215	4642
40	0.6	1000		2154
40	0.7	2154	464	2154
40	0.8	1000	464	464
40	0.9	1000	1000	1000
40	1.0		4642	1000
50	0.1	100	1000	100
50	0.2	464	1000	464
50	0.3	100	464	215
50	0.4	100	215	1000
50	0.5	100	464	215
50	0.6	215	215	1000
50	0.7	464	215	464
50	0.8	464	2154	464
50	0.9	215	1000	215
50	1.0	215	1000	100

Continued: The minimum bootstrap sample sizes for MAD to be within the 95% confidence interval (replaced by uniform)

% replaced	Std. Dev.	2500	5000	10000
60	0.1	100	215	215
60	0.2	215	215	215
60	0.3	215	215	215
60	0.4		215	215
60	0.5	215	215	215
60	0.6	215	215	100
60	0.7	215	464	215
60	0.8	215	464	464
60	0.9	215	464	464
60	1.0	100	215	215
70	0.1	215	100	100
70	0.2	46	46	100
70	0.3	100	215	215
70	0.4	100	464	215
70	0.5	464	464	215
70	0.6	464	464	464
70	0.7	464	215	464
70	0.8	464	464	464
70	0.9	2154	464	215
70	1.0	215	100	100
80	0.1	100	215	215
80	0.2	100	215	464
80	0.3	215	215	100
80	0.4	100	215	100
80	0.5	215	215	100
80	0.6	100	100	100
80	0.7		215	215
80	0.8	215	215	215
80	0.9	100	215	100
80	1.0	100	215	215
90	0.1	100		100
90	0.2	464	464	100
90	0.3	46	4642	100
90	0.4	100	215	215
90	0.5	1000	1000	100
90	0.6	100	215	46
90	0.7	464		215
90	0.8	100		1000
90	0.9	46		215
90	1.0	100	4642	100
100	0.1	215	100	1000
100	0.2	2154	100	1000
100	0.3	215	100	1000
100	0.4	2154	100	2154
100	0.5	215	100	1000
100	0.6	1000	100	215
100	0.7	1000	100	4642
100	0.8	1000	100	1000
100	0.9	1000	100	1000
100	1.0	1000	100	2154

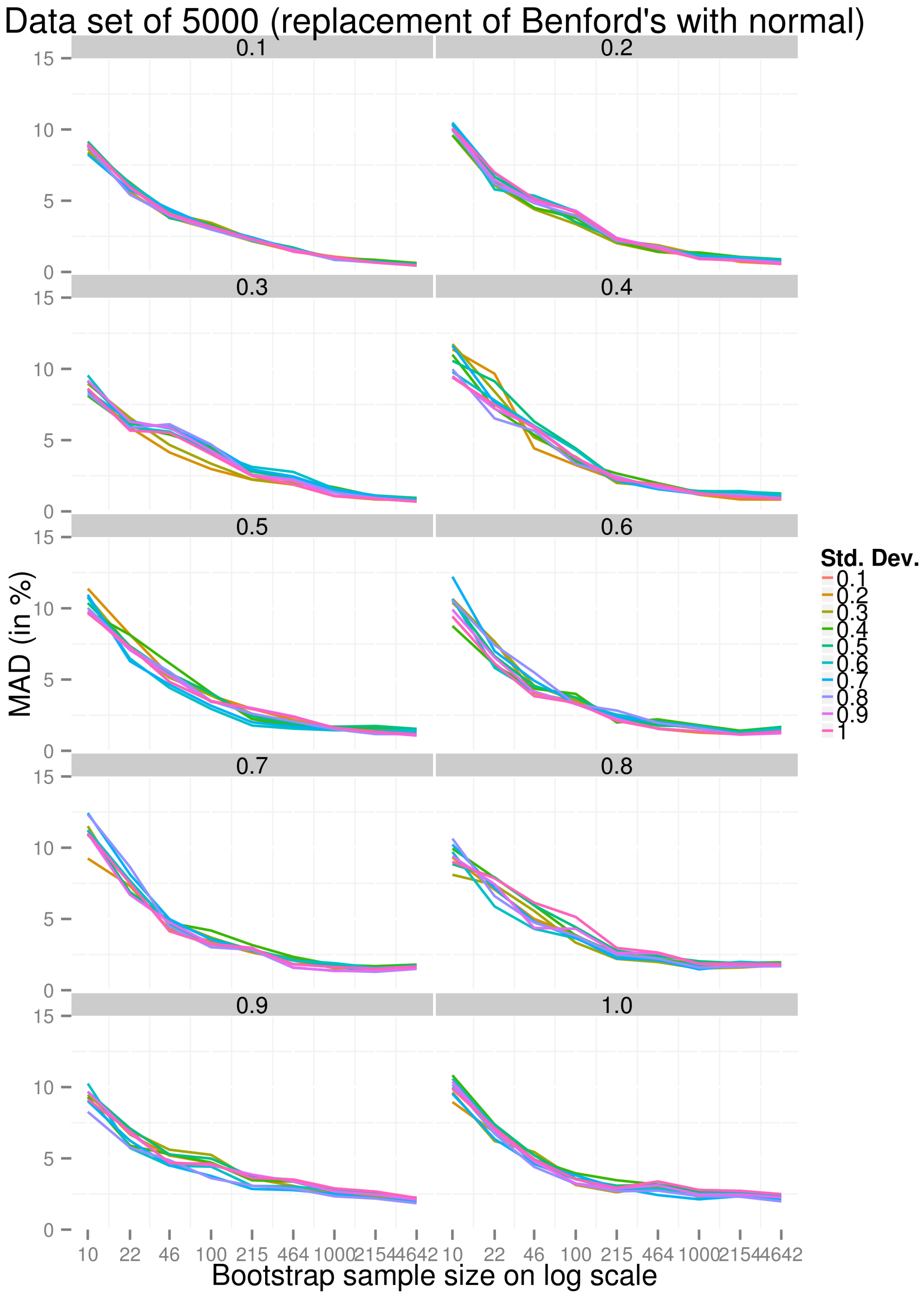
Note: We report the minimum sizes of bootstrap samples required such that the 95 % confidence intervals include the MAD value of the entire data set. The original data set conforms to Benford's Law and some of the values have been replaced by values following a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1. The number of bootstrap iterations is 100.

FIGURE 16. MAD (in %) from partial replacement with values following a normal distribution (size = 2500)



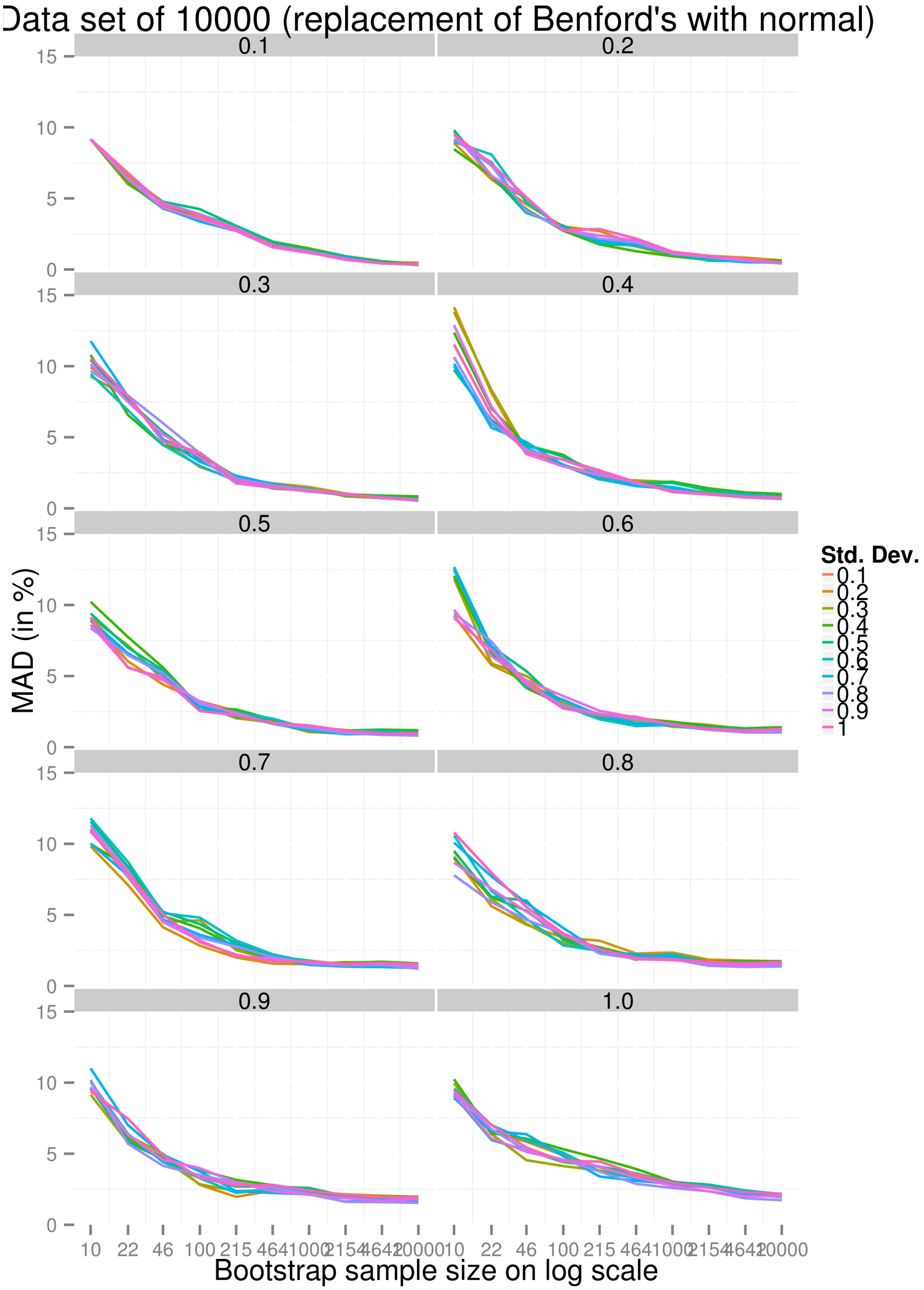
Note: We plot the MAD values (in %) based on bootstrapping on data sets whose values follow a Benford's distribution with the interval $[0.1, 1)$ partially replaced with values following a normal distribution with mean 0 and standard deviation ranging from 0.1 to 1. The data set is of size 2500. The proportion of replacement ranges from 0.1 to 1.

FIGURE 17. MAD (in %) from partial replacement with values following a normal distribution (size = 5000)



Note: We plot the MAD values of the data sets which are of partial replacement with values following a normal distribution. The data set is of size 5000. All the MAD values are calculated from simulations.

FIGURE 18. MAD (in %) from partial replacement with values following a normal distribution (size = 10000)



Note: We plot the MAD values of the data sets which are of partial replacement with values following a normal distribution. The data set is of size 10000. All the MAD values are calculated from simulations.

5.7. **Replace with a uniform random variable.** Table 25 reports the MAD values (in %) for data sets of different sizes and upper bounds. Table 39 and Table 40 in the Appendices report the mean MAD values (in %) and p-values based on various bootstrap sample sizes when we apply bootstrap on data sets of different sizes. The data are partially replaced with values following a uniform distribution between a lower bound 0 and an upper bound ranging from 0.1 to 1.

For a data set of size 2500, the MAD values alone do not identify non-conformity when the proportion of replacement is no more than 30%. The MAD values start to spot non-conformity as the proportion of replacement increases. When the proportion of replacement reaches 70% or beyond, the MAD values computed based on the entire data sets can correctly all 10 cases of non-conformity. As the size of a data set becomes 5000, the MAD values alone work 10 out of 10 specifications of the uniform distribution only when the proportion of replacement is 80% or more. However, the minimum proportion of replacement required for MAD alone to identify non-conformity decreases to 60% when the size of a data set is 10000.

Table 26 reports the mininum sizes of the bootstrap samples required such that the 95% confidence intervals include the MAD values based on the entire data sets. The bootstrap method does not work particularly well in reducing the number of data points required to produce a confidence interval which includes the MAD value of the entire data set. When the proportion of replacement is 0.5 or larger, the bootstrap method does not generate a confidence interval which includes the MAD values of the entire data sets.

Figure 19, Figure 20, and Figure 21 show the mean MAD values (in %) from data sets which are partially replaced by data following various uniform distributions. We observe a similar amount of dispersion in terms of the MAD values in all three plots as is in Figure 16.

Similar to Section 5.6, it is difficult to detect partial replacement of data with values following uniform distributions when the proportion of replacement is 0.3 or smaller. However, when the proportion of replacement is 0.4 or larger, we can apply the bootstrap method with 5% of the data to detect the non-Benfordness of a data set. The lower bounds of the 95% confidence intervals based on bootstrapping 5% of the partially replaced data are all above 1.2%, the threshold value recommended by [DN] for non-conformity to Benford's Law.

TABLE 25. MAD (in %) for data sets partially replaced by values following a uniform distribution

Data Set Size = 2500

% replaced	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
10	0.39	0.49	0.29	0.38	0.47	0.32	0.63	0.41	0.40	0.49
20	0.54	0.62	0.76	0.72	0.41	1.18	0.68	0.62	0.49	0.78
30	0.89	0.96	0.88	0.88	1.09	0.86	0.87	0.57	0.88	1.14
40	1.45	0.72	0.85	1.12	1.30	1.25	0.61	0.83	0.75	0.58
50	1.45	1.20	1.35	1.43	1.37	1.12	0.94	1.03	0.96	1.30
60	1.62	1.58	1.53	1.07	1.77	1.50	1.41	1.25	1.37	1.57
70	1.75	1.75	1.67	1.63	1.54	1.48	1.54	1.34	1.32	1.66
80	2.24	1.77	1.57	2.18	1.88	1.98	1.32	1.52	1.94	1.86
90	2.27	1.98	2.13	2.13	1.91	2.22	1.63	1.47	2.10	2.25
100	2.89	2.21	2.29	2.19	2.58	2.34	2.15	2.21	2.34	2.42

Data Set Size = 5000

% replaced	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
10	0.23	0.36	0.41	0.30	0.44	0.58	0.51	0.28	0.34	0.28
20	0.69	0.59	0.68	0.65	0.44	0.52	0.74	0.24	0.53	0.83
30	0.63	0.65	0.71	0.99	0.75	0.89	0.77	0.63	0.82	0.68
40	0.96	0.99	0.87	0.99	1.18	0.82	1.04	1.01	0.75	0.59
50	1.16	1.14	1.32	1.51	1.20	1.48	1.02	0.83	1.26	1.02
60	1.49	1.41	1.78	1.34	1.24	1.55	1.11	1.19	1.06	1.36
70	1.72	1.64	1.71	1.57	1.50	1.50	1.53	1.12	1.52	1.62
80	2.13	2.04	1.53	1.76	1.64	1.91	1.63	1.44	1.53	1.78
90	1.90	2.29	1.86	2.25	2.24	2.01	1.60	1.43	1.69	2.19
100	2.34	2.09	2.19	2.17	2.22	1.76	2.01	1.71	2.16	2.32

Data Set Size = 10000

% replaced	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
10	0.29	0.38	0.30	0.47	0.36	0.29	0.33	0.21	0.31	0.40
20	0.32	0.51	0.58	0.61	0.54	0.62	0.64	0.55	0.46	0.61
30	0.44	0.65	0.82	0.55	0.62	0.79	0.64	0.62	0.61	0.73
40	0.95	0.93	0.83	0.88	0.78	0.79	0.77	0.96	0.76	0.77
50	1.25	1.25	1.10	1.08	1.06	0.99	0.98	0.97	1.23	1.15
60	1.39	1.41	1.34	1.27	1.34	1.47	1.23	1.27	1.32	1.33
70	1.48	1.47	1.58	1.54	1.79	1.62	1.39	1.26	1.38	1.55
80	1.80	1.73	1.63	1.89	1.75	1.59	1.71	1.46	1.94	1.69
90	2.04	2.14	2.01	2.06	2.03	2.21	1.69	1.71	1.97	2.02
100	2.35	2.20	2.27	2.31	2.34	2.35	2.06	1.90	2.30	2.22

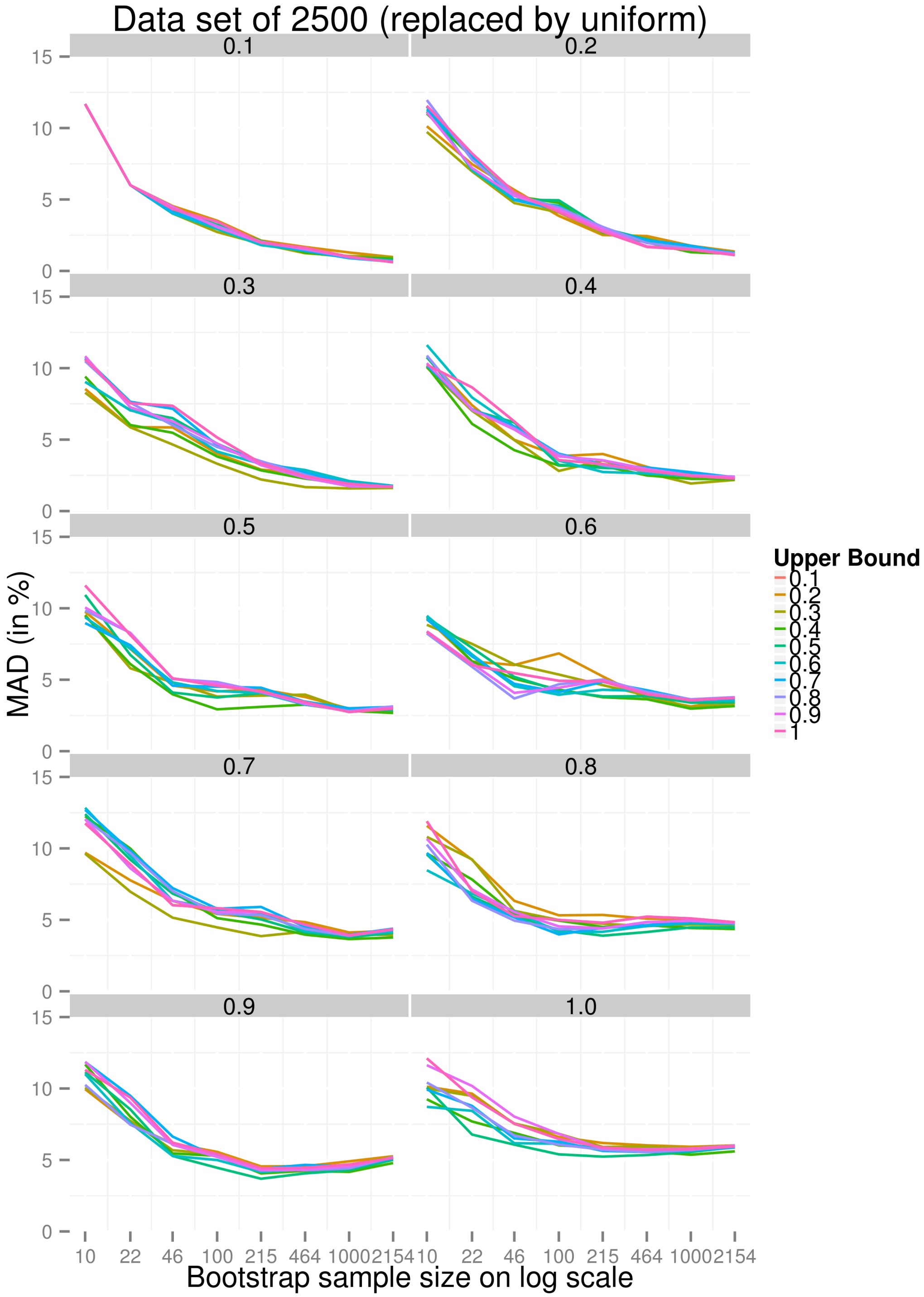
Note: We report the MAD values (in %) of the data sets whose values follow a Benford's distribution bounded by $[0.1, 1)$ partially replaced by data following uniform distributions with a lower bound of 0 and an upper bound ranging from 0.1 to 1.

TABLE 26. The minimum bootstrap sample sizes for MAD to be within the 95% confidence interval (replaced by uniform)

% Replaced	Upper Bound	2500	5000	10000
10	0.1	2154		
10	0.2		4642	
10	0.4			10000
10	0.5	1000	1000	
10	0.6		1000	
10	0.7	1000	1000	
10	0.8	2154		
10	0.9	2154		
10	1.0	2154		
20	0.3	2154	1000	
20	0.4	1000	1000	
20	0.6	1000		2154
20	0.7		1000	1000
20	1.0	2154		
30	0.3	464		
50	0.4	100		

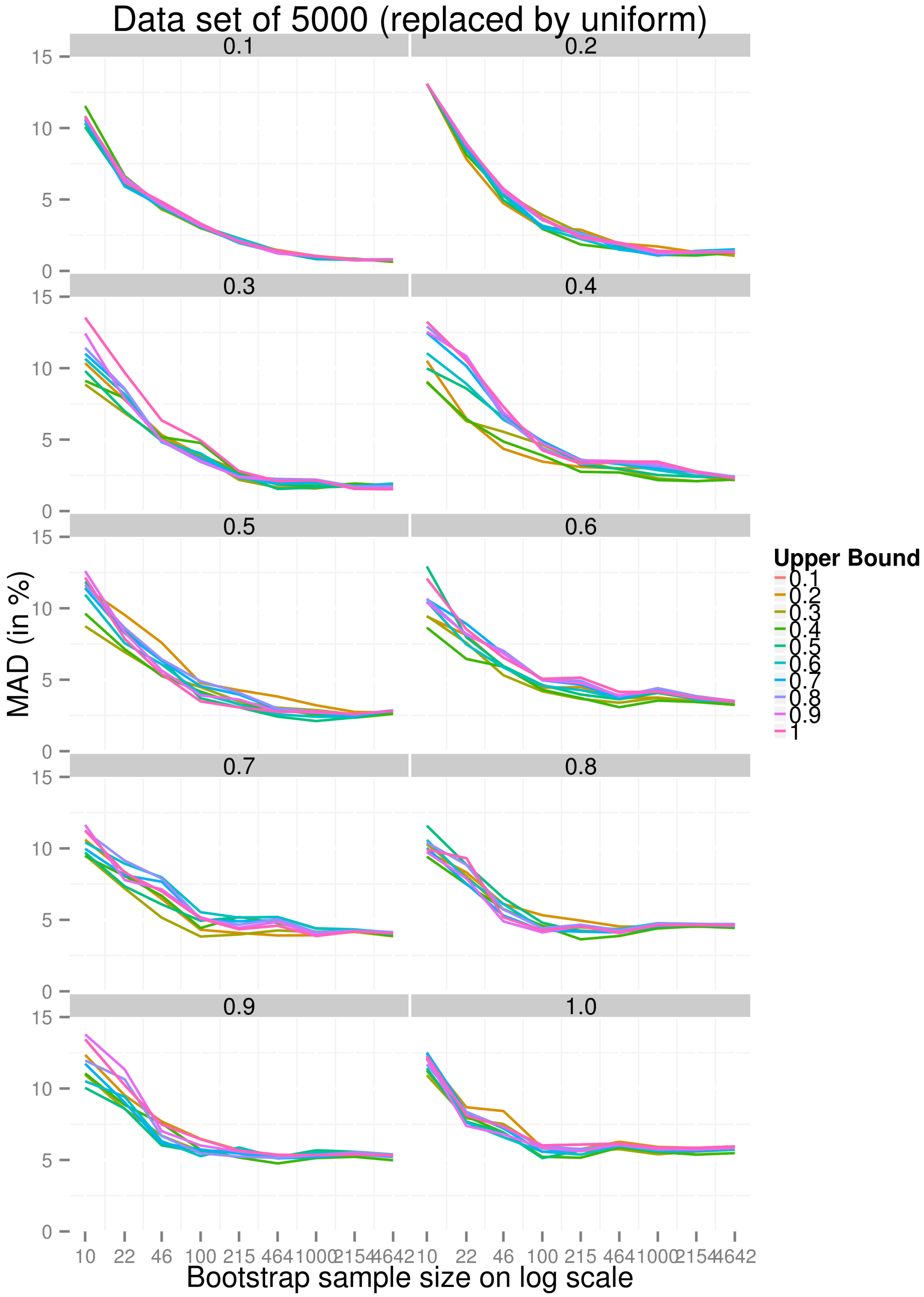
Note: We report the minimum sizes of bootstrap samples required such that the 95 % confidence intervals include the MAD value based on the entire data set. The original data conformed to Benford's Law and some of the data have been replaced by values generated from uniform distributions with a lower bound of 0 and upper bounds ranging from 0.1 to 1. The number of bootstrap iterations is 100.

FIGURE 19. MAD (in %) from partial replacement with data following uniform distributions (size = 2500)



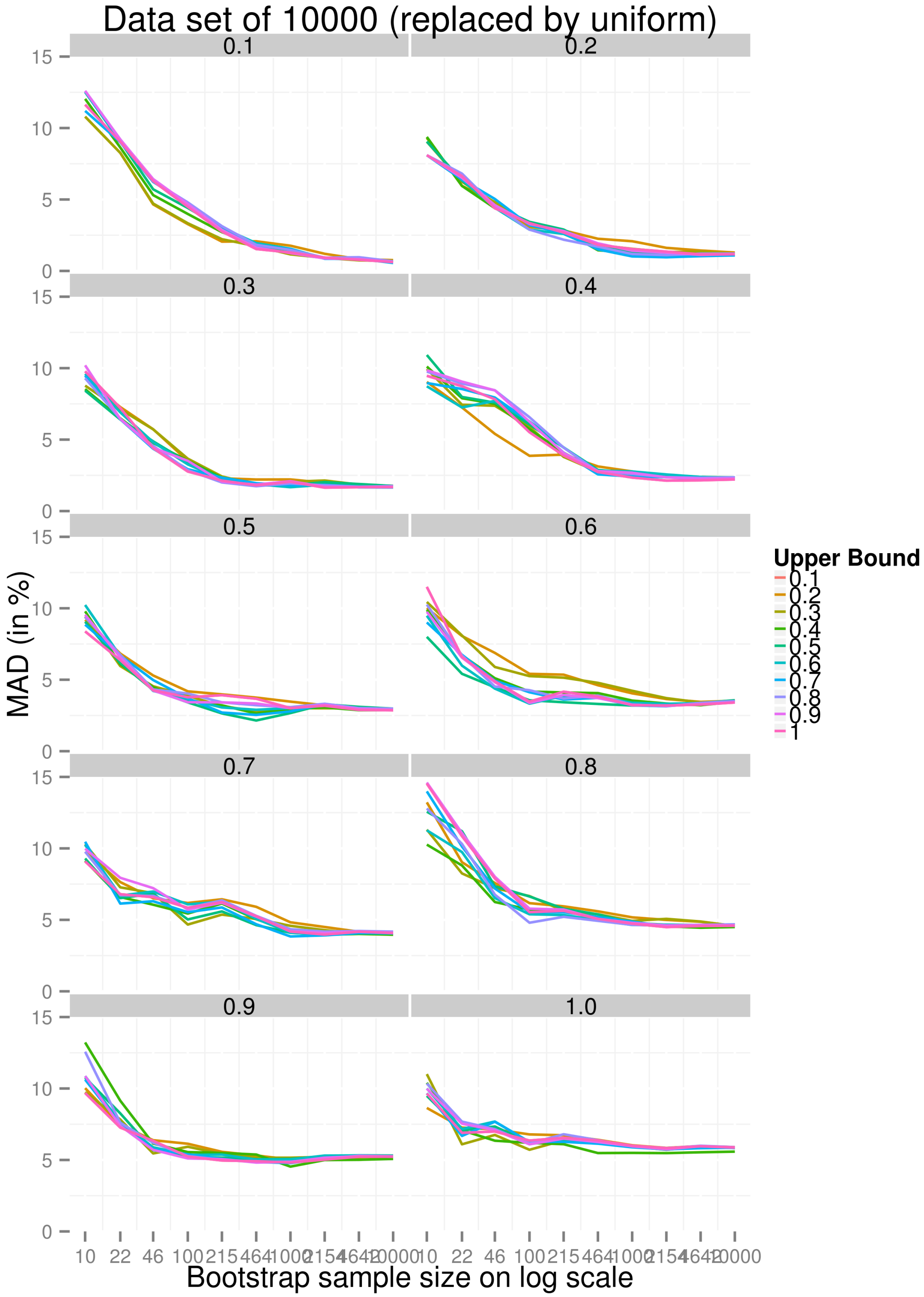
Note: We plot the MAD values (in %) based on bootstrapping on data sets which are partially replaced with data following uniform distributions with a lower bound 0 and upper bounds ranging from 0.1 to 1. The data set is of size 2500. The proportion of replacement of the data set ranges from 0.1 to 1.

FIGURE 20. MAD (in %) from partial replacement with data following uniform distributions (size = 5000)



Note: We plot the MAD values (in %) based on bootstrapping on data sets which are partially replaced with data following uniform distributions with a lower bound 0 and upper bounds ranging from 0.1 to 1. The data set is of size 5000. The proportion of replacement of the data set ranges from 0.1 to 1.

FIGURE 21. MAD (in %) from partial replacement with data following uniform distributions (size = 10000)



Note: We plot the MAD values (in %) based on bootstrapping on data sets which are partially replaced with data following uniform distributions with a lower bound 0 and upper bounds ranging from 0.1 to 1. The data set is of size 10000. The proportion of replacement of the data set ranges from 0.1 to 1.

6. EMPIRICAL DATA ANALYSIS ON SEC FILINGS

We apply Benford's Law to analyze the Form 10-K filings on the U.S. Securities and Exchange Commission's (SEC) Electronic Data-Gathering, Analysis, and Retrieval System (EDGAR)[4] for two different companies. According to the SEC, the annual report on Form 10-K "provides a comprehensive overview of the company's business and financial condition and includes audited financial statements".[5] This is not meant to be a comprehensive study on how closely the data values from the Form 10-K filings resemble the Benford's distribution but rather an exploratory one on whether Benford's law and bootstrapping can be used to detect fraudulent financial reporting.

We first obtain two data sets (one fraudulent and the other, not, at least not exposed) with accounting data from the SEC website directly. We then examine the frequency distribution of the first digits for both data sets and calculate the deviations from the theoretical Benford's distribution. We proceed to apply the bootstrap by repeatedly resampling each data set and checking the aforementioned distributions and deviations. We attempt to determine the threshold value of the bootstrap sample size such that the 95% confidence interval includes the actual deviation based on the entire data set.

The two companies we have examined are the Enron Corporation and the Intel Corporation. The Enron scandal broke out in the second half of year 2001 where Enron was discovered of committing creative accounting and fraudulent reporting. We obtained Enron's Form 10-K filings from EDGAR from year 1998 through 2001.[6] For comparison purposes, we also obtained the Form 10-K filings for the Intel Corporation during the same period. We extract numbers from the Form 10-K based on the following criteria:

(1) Obtain all the tables in the Form 10-K, excluding the content page;
(2) Exclude all figures from captions, footnotes, headers, and column and row names;
(3) Exclude all the derived numbers which are the results of arithmetic operations based on other numbers. For instance, total revenues and net incomes are derived numbers.

---

[4] See http://www.sec.gov/edgar.shtml for more details.

[5] http://www.sec.gov/answers/form10k.htm

[6] See `http://www.sec.gov/cgi-bin/browse-edgar?action=getcompany&CIK=0001024401&type=10-K&dateb=&owner=exclude&count=40`.

We first examine the frequency distribution of the first digits from the Enron 10-K filings annually. Table 27 reports the frequency distributions of the first digits by year. We compute the deviation from the theoretical Benford's distribution based on the Chi-Square tests and MAD and report the results in Table 28. The extent to which the frequency distribution deviates from the theoretical values is indistinguishable based on the p-values associated with the Chi-Square tests. However, the MAD values indicate that there is an increasing amount of deviation from year 1999 to year 2000 (from 0.007 to 0.013, a close to 50% increase). According to the criteria by [DN], only the numbers from 2000 show non-conformity to Benford's Law with an MAD value of 0.013. The MAD values from the other three years are of marginally acceptable conformity to Benford's Law.

TABLE 27. Freqency distribution of the first digits from Enron's 10-K filings

First Digit	Benford	1998	1999	2000	2001
1	30.1	29.3	29.3	31.7	28.6
2	17.6	20.1	18.8	14.4	15.6
3	12.5	12.6	13.2	12.1	12.6
4	9.7	9.4	8.4	8.3	9.0
5	7.9	8.0	8.0	7.9	7.5
6	6.7	7.1	7.3	8.2	8.2
7	5.8	5.1	6.0	7.9	8.4
8	5.1	4.6	5.6	5.7	5.2
9	4.6	3.6	3.4	3.8	4.9

Note: We have obtained Enron's Form 10-K filings from EDGAR and reported the frequency distribution of all the non-derived numbers reported from year 1998 to year 2001.

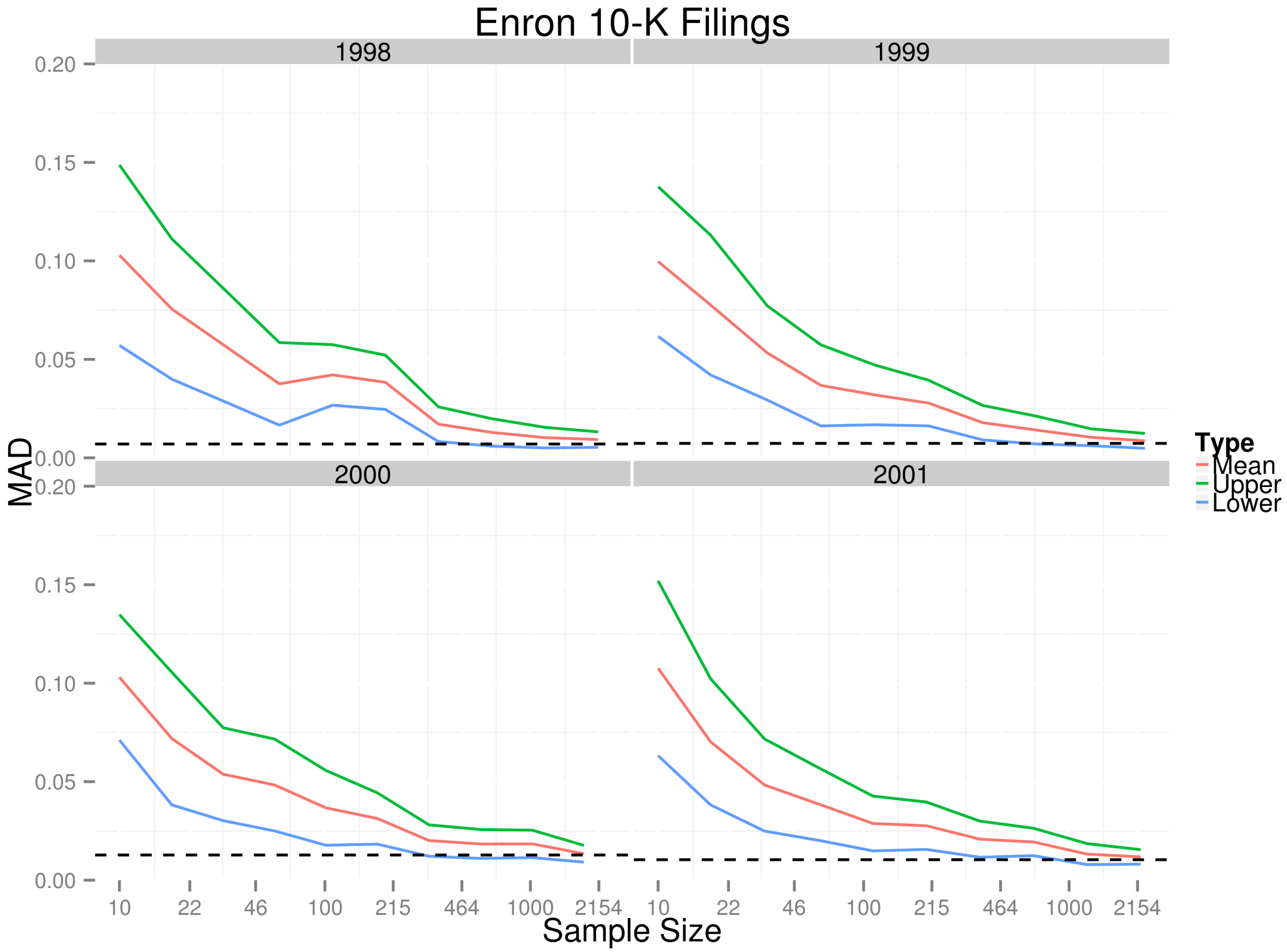
TABLE 28. P-values and the MAD values based on Enron's 10-K filings

	1998	1999	2000	2001
P Value	1.000	1.000	1.000	1.000
MAD	0.007	0.007	0.013	0.010

Note: We report the p-values associated with the Chi-Square tests and the MAD values for the frequency distribution of the first digits in Enron's Form 10-K filings. The first digits are the first numbers occurred in all the non-derived numbers reported from year 1998 to 2001.

We apply the bootstrap method on non-derived numbers extracted from the Form 10-K filed by Enron from 1998 to 2001. We do so with different bootstrap sample sizes in order to find out the minimum number of data points required to estimate the deviation. Figure 22 shows the mean estimates and the 95% confidence intervals for different bootstrap sample sizes. We apply 100 bootstrap iterations to obtain the mean estimates and the confidence intervals. The dotted black lines indicate the actual MAD values for each year. As the number of data points increases, the variation of the MAD estimates becomes smaller. For year 1998, 1999, and 2000, the 95% confidence intervals include the actual MAD values once the bootstrap sample size goes beyond 464.

FIGURE 22. MAD estimates from bootstrapping with different sample sizes based on Enron's 10K (1998-2001)



Note: We apply the bootstrap method with various sample sizes on data from the Enron's 10-K filings. We report the mean MAD values (red, middle lines) and the 95% confidence interval for the MAD estimates (green and blue, top and bottom lines). The black dotted line shows the MAD value based on the entire data set.

For comparison purposes, we obtain Form 10-K filings from the Intel Corporation during the same period (1889-2001) and conduct the same set of analyses. Table 29 reports the frequency distributions of the first digits from the non-derived numbers based on Intel's Form 10-K filings. Figure 23 shows the results of applying the bootstrap method with different sample sizes. The 95% confidence intervals include the mean estimate of the MAD values once the sample size is 215 or more.

Table 30 reports that in terms of MAD, the non-derived numbers from Intel's Form 10-K have more deviations from the theoretical Benford's distribution as compared to those from Enron. Further, all the MAD values based on the Intel Form 10-K filings are in the range of non-conformity. However, this alone does not reflect or imply that there is fraudulent accounting in either company. This brief empirical analysis of the two companies' financial filing shows that examining the frequency distribution of the first digits and the extent to which the data deviates from the Benford's distribution does not necessarily present us with evidence of forgery or data manipulation.

TABLE 29. Freqency distribution of first digits from Intel's filing (1998-2001)

First Digit	Benford	1998	1999	2000	2001
1	30.1	26.7	28.1	26.6	26.4
2	17.6	18.2	16.9	16.8	16.6
3	12.5	13.3	15.0	16.5	14.6
4	9.7	8.9	8.7	8.3	13.7
5	7.9	10.2	10.6	9.0	6.3
6	6.7	8.6	9.4	9.9	7.6
7	5.8	4.9	4.4	5.3	7.2
8	5.1	4.4	3.0	3.0	3.5
9	4.6	4.8	3.8	4.4	4.0

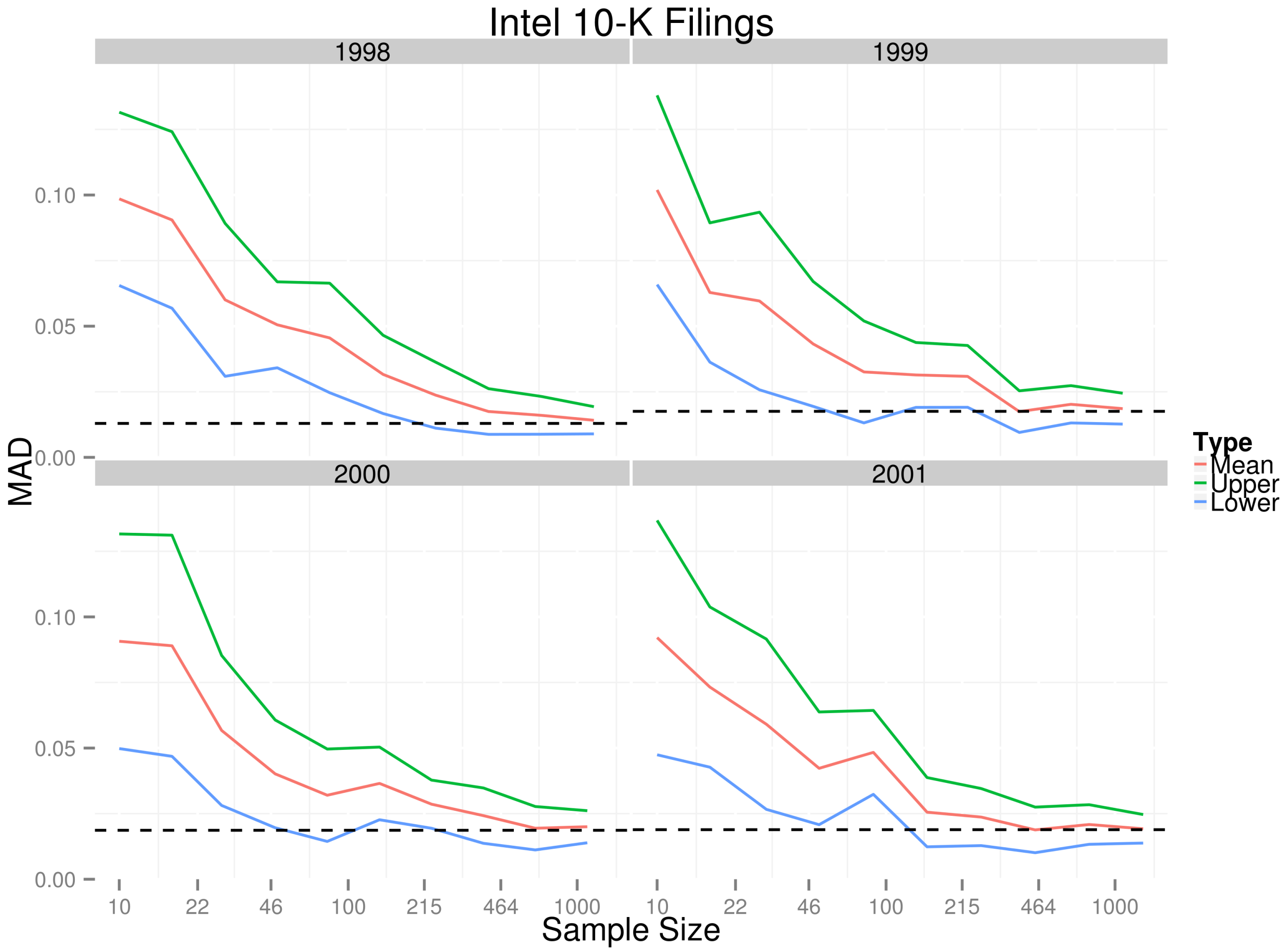
Note: We obtained Intel's 10-K filings from EDGAR and reported the frequency distribution of all the non-derived numbers reported from year 1998 to year 2001.

TABLE 30. P-values and the MAD values based on Intel's 10-K filing (1998-2001)

	1998	1999	2000	2001
P Value	1.000	1.000	1.000	1.000
MAD	0.013	0.018	0.019	0.019

Note: We report the p-values related to the Chi-Square tests and the MAD values for the frequency distributions of the first digits in Intel's Form 10-K filings. The first digits are the first numbers occurred in all the non-derived numbers reported from year 1998 to year 2001.

FIGURE 23. MAD estimates and intervals from bootstrapping with different sample sizes based on Intel's 10K (1998-2001)



Note: We apply the bootstrap method with various sample sizes on data from the Intel's 10-K filings. We show the mean MAD values (red, middle lines) and the 95% confidence interval for the MAD estimates (green and blue, top and bottom lines). The black dotted line shows the MAD value based on the entire data set.

7. CONCLUSION

Methods based on Benford's Law are often used as a screening tool to identify investigation targets for data fraud. Given the amount of time and computational resources required to process huge amounts of data, we started the paper with the goal of looking for a way to reduce the number of data points needed for Benford-based fraud detection.

We explored two major ways in which one could manipulate data, adding new values to existing ones and replacing existing data with new ones. We first reviewed the theory on Benford's Law and performed theoretical calculations concerning frequency distribution of the first digits of convolution and partial replacement of data, respectively. We also used simulations to gain a better understanding of convolution and partial replacement of the data.

In an attempt to look for any relationship among the amount of deviation from the Benford's distribution, sample size, attributes of the mixed distribution, and proportion of replacement where applicable, we have found various patterns but not any functional relations. Specifically, in the case of summing a Benford random variable and a normal random variable, across different standard deviations of the normal distribution, the MAD value decreases as the sample size increases. Yet for different sample sizes, the MAD values are symmetric about the standard deviation of 0.6. In the case of partial replacement with values following a normal distribution, across different proportions of replacement, we have observed similar shapes of the surface formed by MAD, logarithm of sample size, and standard deviation of the normal distribution.

Applying non-parametric bootstrap in simulations, we have explored the bootstrap as a possible way of reducing the number of data points needed for Benford-based fraud detection measures. We studied specifically the number of data points required for bootstrap to generate a 95% confidence interval which includes the MAD value of the whole data set as well as the minimum number of data points required to detect non-conformity to Benford's Law.

Regarding the inclusion of the MAD value of the entire data set, in the case of summing a Benford random variable and a normal random variable, we often require a tenth of the data for bootstrap for the estimated MAD to be in the 95% confidence interval. When summing a Benford random variable and a uniform random variable, we require one twentieth of the data to achieve the same effect. In contrast, in the case of replacing data following a Benford's distribution with values following either a normal distribution or a uniform distribution, the estimated MAD values based on the bootstrap often do not fall into the 95% confidence intervals.

As to the minimum number of data points required for fraud detection, assuming that a data set of size 5000 or larger is either free of manipulation or added by values which follow a normal distribution or a uniform distribution, we require 5% of the data set to detect potential manipulation based on the bootstrap method detailed in Section 5. This greatly reduces the auditor's need for full access to data in order to determine the Benfordness of a data set and cuts down the computational power required.

However, it is less straightforward when it comes to detect non-conformity to Benford's Law when data are partially replaced by values generated from either a normal distribution or a uniform distribution. Table 31 summarizes the results.

TABLE 31. Percentages of data required to detect non-conformity to Benford's Law

Replaced by data following normal distributions			
Size	Proportion replaced	% of data required	Comments
2500	0.7	4.0	All but $\sigma = 0.3$
2500	≥ 0.8	4.0	All scenarios
5000	0.8	9.2	All but $\sigma = 0.2, 0.3$
5000	0.9	9.2	All scenarios
10000	≥ 0.8	10.0	All scenarios
Replaced by data following uniform distributions			
Size	Proportion replaced	% of data required	Comments
2500	≥ 0.4	5.0	All scenarios
5000	≥ 0.4	5.0	All scenarios
10000	≥ 0.4	5.0	All scenarios

Note: We report the percentages of data required to apply the bootstrap method and detect non-conformity to Benford's Law. The original data conform to Benford's Law and a proportion of data have been replaced by values generated from either a normal distribution or a uniform distribution. The "Comments" column indidcates how well the method is able to detect non-conformity to Benford's Law. "All scenarios" means that the method is able to detect non-Benfordness of the data set in all specifications of the distributions studied at a significance level of 5%. See Section 5 for more details.

Assuming parital replacement by data following a normal distribution, for a data set of size 2500, we cannot effectively detect non-conformity to Benford's Law when the proportion of replacement is 0.6 or smaller. When the proportion of replacement is 0.7, we can use 4% of the data set to detect non-Benfordness except when the standard deviation of the normal distribution is 0.3. When the proportion of replacement is 0.8 or higher, using 4% of the data allow us to detect non-conformity to Benford's Law at a significance level of 5%.

For a data set of size 5000, when the proportion of replacement is 0.7 or smaller, the bootstrap method does not allow us to detect non-Benfordness effectively. When we use around 9.2% of the data set (464 out of 5000 data points), 80% of which are replaced by values generated by a normal distribution, the method detects 8 out of 10 different specifications of the normal distribution. It does not work effectively when the standard deviation of the normal distribution is 0.2 or 0.3. As the proportion of replacement reaches 0.9, we can use 9% of the data to effectively detect non-Benfordness of the data.

As the size of the data set becomes 10000, the bootstrap method allows us to use 10% of the data to detect potential non-conformity to Benford's Law when the proportion of replacement is 0.8 or higher.

Assuming partial replacement of data by values following a uniform distribution, when the proportion of replacement is 0.4 or larger, we can apply the bootstrap method with 5% of the data to detect non-Benfordness of a data set. However, the method does not work well when the proportion of replacement is 0.3 or smaller.

The settings of our paper are fairly general. Most of the results are basing on simulated data. One can use our results according to her own agenda. Oftentimes when auditors do not have full access to data, the bootstrap can be an alternative method to use a proportion of the data and estimate the extent to which the data deviate from the Benford's distribution.

Having said that, we are aware that there is much to be done in the area of applying the bootstrap in the context of Benford's Law. A few possible further research questions include

(1) Evaluate the threshold MAD values proposed by [DN]. The MAD values are a convenient way to detect deviation from the Benford's distribution. However, the exact cut-off values can be extensively studied with both generic and specific data sets.
(2) Extend the study here to the first two digits and investigate whether the same patterns observed for the first digits remain the same for the first two digits.
(3) Vary the mean of the normal distribution and the lower bound of the uniform distribution studied and investigate the relations among the amount of deviation, sample size, and attributes of the normal or uniform distribution involved.

(4) Use more data from Form 10-K or tax return information for empirical analysis.

8. APPENDICES

8.1. Table of convolution of a Benford's density and a normal density.

TABLE 32. Lower bounds of the 95% confidence interval of MAD when bootstrapping on the convolution of a Benford's density and a normal density (100 iterations)

Data Set Size = 2500				
Std. Dev.	Sample Size	Mean	MAD Std. Dev.	Lower
0.1	10	13.48	1.86	9.77
0.1	22	9.38	1.65	6.08
0.1	46	5.87	1.09	3.69
0.1	100	4.66	0.84	2.98
0.1	215	3.91	0.64	2.63
0.1	464	3.08	0.45	2.18
0.1	1000	2.51	0.28	1.95
0.1	2154	2.43	0.20	2.03
0.2	10	11.68	2.32	7.03
0.2	22	8.52	1.82	4.87
0.2	46	7.04	1.21	4.63
0.2	100	4.77	0.86	3.04
0.2	215	3.80	0.59	2.62
0.2	464	3.21	0.46	2.29
0.2	1000	2.94	0.28	2.38
0.2	2154	2.89	0.22	2.45
0.3	10	13.90	1.39	11.12
0.3	22	10.09	1.65	6.79
0.3	46	7.31	1.31	4.69
0.3	100	5.38	1.00	3.39
0.3	215	3.63	0.66	2.31
0.3	464	3.34	0.45	2.44
0.3	1000	2.61	0.29	2.02
0.3	2154	2.66	0.23	2.20
0.4	10	12.27	2.24	7.79
0.4	22	10.02	1.90	6.22
0.4	46	5.63	1.25	3.13
0.4	100	4.25	0.93	2.39
0.4	215	3.49	0.66	2.17
0.4	464	2.62	0.47	1.67
0.4	1000	2.38	0.32	1.73
0.4	2154	2.55	0.26	2.04
0.5	10	9.42	2.32	4.79
0.5	22	7.68	1.54	4.60
0.5	46	5.51	1.07	3.37
0.5	100	3.75	0.93	1.89
0.5	215	2.97	0.64	1.68

0.5	464	2.70	0.49	1.73
0.5	1000	2.44	0.36	1.72
0.5	2154	2.40	0.25	1.90
0.6	10	12.16	1.96	8.24
0.6	22	8.21	1.66	4.89
0.6	46	5.24	1.16	2.92
0.6	100	3.97	0.89	2.19
0.6	215	2.61	0.59	1.42
0.6	464	2.36	0.42	1.52
0.6	1000	2.04	0.31	1.42
0.6	2154	1.93	0.20	1.53
0.7	10	9.56	2.49	4.57
0.7	22	7.66	1.71	4.23
0.7	46	4.74	1.23	2.29
0.7	100	3.96	0.81	2.33
0.7	215	2.90	0.58	1.74
0.7	464	2.74	0.39	1.95
0.7	1000	2.34	0.28	1.78
0.7	2154	2.30	0.16	1.98
0.8	10	12.33	2.21	7.91
0.8	22	8.95	1.42	6.12
0.8	46	5.28	1.38	2.51
0.8	100	4.87	1.05	2.77
0.8	215	3.13	0.60	1.92
0.8	464	2.82	0.41	2.01
0.8	1000	2.80	0.31	2.18
0.8	2154	2.55	0.21	2.12
0.9	10	11.35	1.73	7.89
0.9	22	7.94	1.20	5.54
0.9	46	5.93	1.06	3.81
0.9	100	4.39	0.92	2.55
0.9	215	3.07	0.66	1.74
0.9	464	2.62	0.47	1.68
0.9	1000	2.70	0.32	2.05
0.9	2154	2.51	0.25	2.01
1.0	10	9.91	2.05	5.81
1.0	22	8.36	1.21	5.94
1.0	46	5.62	1.36	2.91
1.0	100	3.94	0.83	2.27
1.0	215	3.02	0.55	1.92
1.0	464	2.53	0.50	1.52
1.0	1000	2.88	0.27	2.34
1.0	2154	2.59	0.22	2.15

Data Set Size = 5000

Std. Dev.	Sample Size	Mean	MAD Std. Dev.	Lower

0.1	10	11.19		2.64	5.90
0.1	22	10.00		1.90	6.20
0.1	46	7.44		1.22	5.01
0.1	100	4.31		0.90	2.51
0.1	215	3.68		0.60	2.48
0.1	464	2.61		0.45	1.70
0.1	1000	2.35		0.27	1.81
0.1	2154	2.39		0.17	2.04
0.1	4642	2.54		0.13	2.28
0.2	10	11.47		2.47	6.54
0.2	22	9.20		1.65	5.91
0.2	46	7.42		1.33	4.77
0.2	100	4.69		0.89	2.91
0.2	215	4.36		0.71	2.94
0.2	464	3.53		0.49	2.55
0.2	1000	2.75		0.28	2.18
0.2	2154	2.91		0.22	2.46
0.2	4642	3.06		0.13	2.80
0.3	10	8.18		2.30	3.59
0.3	22	6.59		1.65	3.28
0.3	46	5.64		1.32	3.00
0.3	100	4.37		0.90	2.56
0.3	215	3.79		0.65	2.49
0.3	464	3.24		0.46	2.32
0.3	1000	3.12		0.34	2.43
0.3	2154	2.97		0.24	2.49
0.3	4642	2.94		0.15	2.64
0.4	10	10.77		2.03	6.71
0.4	22	7.11		1.65	3.81
0.4	46	6.16		1.21	3.75
0.4	100	3.44		0.85	1.74
0.4	215	2.76		0.60	1.56
0.4	464	2.57		0.44	1.69
0.4	1000	2.62		0.35	1.93
0.4	2154	2.55		0.22	2.12
0.4	4642	2.50		0.14	2.23
0.5	10	9.10		1.73	5.63
0.5	22	5.90		1.49	2.92
0.5	46	4.66		1.18	2.30
0.5	100	3.27		0.94	1.39
0.5	215	2.74		0.60	1.55
0.5	464	2.10		0.42	1.27
0.5	1000	2.23		0.37	1.49
0.5	2154	2.13		0.23	1.66

0.5	4642	2.15	0.14	1.87
0.6	10	9.50	1.97	5.57
0.6	22	7.42	1.49	4.43
0.6	46	5.77	1.31	3.15
0.6	100	4.57	0.85	2.88
0.6	215	3.20	0.56	2.08
0.6	464	2.29	0.45	1.38
0.6	1000	2.09	0.33	1.43
0.6	2154	1.91	0.22	1.47
0.6	4642	1.92	0.14	1.63
0.7	10	9.50	1.87	5.77
0.7	22	6.52	1.60	3.32
0.7	46	3.97	1.02	1.92
0.7	100	3.32	0.82	1.69
0.7	215	2.69	0.56	1.57
0.7	464	2.34	0.44	1.47
0.7	1000	1.83	0.28	1.27
0.7	2154	1.76	0.18	1.40
0.7	4642	2.12	0.13	1.86
0.8	10	8.27	1.93	4.41
0.8	22	5.36	1.39	2.57
0.8	46	5.11	1.20	2.72
0.8	100	4.64	0.86	2.92
0.8	215	3.66	0.53	2.60
0.8	464	2.82	0.42	1.97
0.8	1000	2.43	0.34	1.74
0.8	2154	2.34	0.20	1.95
0.8	4642	2.47	0.13	2.21
0.9	10	9.64	1.63	6.38
0.9	22	6.26	1.44	3.39
0.9	46	4.56	1.12	2.31
0.9	100	3.88	0.92	2.03
0.9	215	3.04	0.51	2.03
0.9	464	2.96	0.45	2.06
0.9	1000	2.50	0.33	1.84
0.9	2154	2.39	0.21	1.98
0.9	4642	2.48	0.14	2.21
1.0	10	11.54	1.46	8.62
1.0	22	7.56	1.39	4.78
1.0	46	6.09	1.02	4.05
1.0	100	3.88	0.90	2.07
1.0	215	2.38	0.61	1.15
1.0	464	2.51	0.51	1.50
1.0	1000	2.56	0.34	1.89
1.0	2154	2.38	0.24	1.90

1.0	4642	2.46	0.13	2.19
Data Set Size = 10000				
Std. Dev.	Sample Size	Mean	MAD Std. Dev.	Lower
0.1	10	10.49	1.90	6.69
0.1	22	6.92	1.63	3.65
0.1	46	6.09	1.19	3.71
0.1	100	4.77	0.86	3.05
0.1	215	3.53	0.55	2.43
0.1	464	3.15	0.39	2.37
0.1	1000	2.79	0.24	2.31
0.1	2154	2.65	0.20	2.25
0.1	4642	2.40	0.12	2.16
0.1	10000	2.44	0.10	2.24
0.2	10	12.59	2.36	7.86
0.2	22	8.24	2.07	4.09
0.2	46	5.56	1.32	2.91
0.2	100	3.34	0.83	1.69
0.2	215	2.98	0.61	1.76
0.2	464	3.04	0.51	2.03
0.2	1000	2.89	0.34	2.22
0.2	2154	2.68	0.23	2.21
0.2	4642	2.53	0.17	2.19
0.2	10000	2.77	0.10	2.58
0.3	10	14.53	1.21	12.11
0.3	22	9.23	1.78	5.67
0.3	46	7.72	1.39	4.94
0.3	100	5.35	0.98	3.39
0.3	215	4.41	0.60	3.21
0.3	464	4.01	0.46	3.08
0.3	1000	3.22	0.36	2.50
0.3	2154	2.99	0.24	2.51
0.3	4642	2.80	0.16	2.49
0.3	10000	2.83	0.11	2.61
0.4	10	12.22	1.66	8.91
0.4	22	7.77	1.62	4.53
0.4	46	6.51	1.21	4.08
0.4	100	4.28	0.83	2.63
0.4	215	3.56	0.63	2.29
0.4	464	3.49	0.52	2.46
0.4	1000	2.47	0.33	1.81
0.4	2154	2.46	0.21	2.03
0.4	4642	2.46	0.17	2.13
0.4	10000	2.42	0.11	2.20
0.5	10	11.22	2.19	6.84
0.5	22	7.50	1.73	4.04

0.5	46	6.12	1.31	3.50
0.5	100	3.99	0.86	2.28
0.5	215	3.24	0.67	1.90
0.5	464	3.34	0.50	2.34
0.5	1000	2.44	0.37	1.69
0.5	2154	2.42	0.23	1.97
0.5	4642	2.39	0.18	2.03
0.5	10000	2.24	0.09	2.06
0.6	10	10.49	2.40	5.69
0.6	22	7.55	1.86	3.83
0.6	46	5.67	1.24	3.18
0.6	100	3.46	0.85	1.76
0.6	215	3.40	0.61	2.18
0.6	464	3.09	0.39	2.31
0.6	1000	2.47	0.26	1.94
0.6	2154	2.14	0.17	1.79
0.6	4642	2.15	0.15	1.85
0.6	10000	1.98	0.09	1.79
0.7	10	11.99	2.24	7.51
0.7	22	7.22	1.78	3.67
0.7	46	6.17	1.40	3.36
0.7	100	3.48	0.82	1.85
0.7	215	3.02	0.59	1.83
0.7	464	2.67	0.43	1.81
0.7	1000	2.52	0.29	1.95
0.7	2154	2.34	0.17	2.00
0.7	4642	2.19	0.14	1.92
0.7	10000	2.21	0.10	2.01
0.8	10	12.13	2.27	7.59
0.8	22	7.65	1.85	3.95
0.8	46	7.16	1.19	4.78
0.8	100	3.90	0.97	1.96
0.8	215	3.68	0.66	2.35
0.8	464	3.04	0.46	2.11
0.8	1000	2.83	0.29	2.25
0.8	2154	2.37	0.20	1.97
0.8	4642	2.27	0.16	1.95
0.8	10000	2.25	0.09	2.06
0.9	10	11.57	2.16	7.25
0.9	22	6.43	1.56	3.30
0.9	46	5.25	1.08	3.08
0.9	100	3.96	0.76	2.44
0.9	215	3.84	0.72	2.41
0.9	464	3.20	0.43	2.34
0.9	1000	2.85	0.31	2.22

0.9	2154	2.55	0.20	2.15
0.9	4642	2.37	0.16	2.05
0.9	10000	2.35	0.10	2.15
1.0	10	9.73	2.25	5.24
1.0	22	6.68	1.72	3.24
1.0	46	6.02	1.10	3.83
1.0	100	4.39	0.85	2.69
1.0	215	4.16	0.62	2.92
1.0	464	3.18	0.45	2.28
1.0	1000	2.87	0.32	2.22
1.0	2154	2.56	0.21	2.14
1.0	4642	2.45	0.16	2.12
1.0	10000	2.46	0.11	2.24

Note: We report the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values when bootstrapping with different sample sizes and different standard deviations of the normal distribution which is one component of the convolution. The MAD values and estimates are reported in percentages. The data are the results of the sum of a Benford random variable and a normal random variable. The “Std. Dev.” column specifies the standard deviation of the normal distribution. The “Sample” Size column indicates the size of the bootstrap sample. The “Mean” column reports the mean estimates of the MAD values based on 100 times of bootstrapping. The “MAD Std. Dev.” column reports the standard deviation of the MAD values based on 100 times of bootstrapping. The “Lower” column reports the lower bound of the 95% confidence interval. [DN] state that data show non-conformity to the Benford’s distribution when the MAD value is greater than 1.2%.

8.2. Table of convolution of a Benford's density and a uniform density.

TABLE 33. Lower bounds of the 95% confidence interval of MAD when bootstrapping on the convolution of a Benford's density and a uniform density (100 iterations)

Data Set Size = 2500				
Upper Bound	Sample Size	Mean	Std. Dev.	Lower
0.1	10	8.94	1.50	5.93
0.1	22	8.07	1.87	4.33
0.1	46	4.71	1.27	2.17
0.1	100	4.69	0.89	2.92
0.1	215	3.65	0.62	2.42
0.1	464	2.98	0.46	2.07
0.1	1000	2.99	0.29	2.42
0.1	2154	2.66	0.18	2.30
0.2	10	9.81	2.25	5.30
0.2	22	9.09	1.64	5.81
0.2	46	6.15	1.23	3.70
0.2	100	5.34	0.87	3.60
0.2	215	4.83	0.54	3.76
0.2	464	4.31	0.35	3.61
0.2	1000	4.00	0.25	3.49
0.2	2154	3.96	0.16	3.64
0.3	10	11.98	2.44	7.09
0.3	22	10.42	1.76	6.90
0.3	46	6.71	1.23	4.24
0.3	100	6.98	0.91	5.15
0.3	215	5.85	0.71	4.43
0.3	464	5.34	0.39	4.56
0.3	1000	5.12	0.33	4.46
0.3	2154	4.97	0.20	4.57
0.4	10	11.66	2.44	6.78
0.4	22	9.97	1.99	6.00
0.4	46	7.02	1.35	4.32
0.4	100	6.93	0.94	5.06
0.4	215	6.16	0.69	4.79
0.4	464	5.81	0.41	4.98
0.4	1000	5.71	0.33	5.05
0.4	2154	5.51	0.20	5.12
0.5	10	13.56	2.43	8.71
0.5	22	10.16	1.93	6.30
0.5	46	7.16	1.35	4.46
0.5	100	6.58	1.03	4.52
0.5	215	5.93	0.79	4.35
0.5	464	6.07	0.48	5.12

0.5	1000	5.94	0.33	5.29
0.5	2154	5.60	0.22	5.16
0.6	10	11.60	2.59	6.43
0.6	22	9.92	1.92	6.09
0.6	46	7.50	1.45	4.61
0.6	100	6.67	1.09	4.50
0.6	215	6.19	0.81	4.58
0.6	464	6.23	0.48	5.27
0.6	1000	6.11	0.33	5.44
0.6	2154	5.72	0.22	5.28
0.7	10	12.12	2.11	7.91
0.7	22	11.51	1.74	8.03
0.7	46	7.53	1.41	4.71
0.7	100	6.53	1.05	4.42
0.7	215	5.90	0.75	4.40
0.7	464	5.96	0.48	5.01
0.7	1000	5.98	0.31	5.36
0.7	2154	5.71	0.20	5.31
0.8	10	12.76	2.17	8.42
0.8	22	8.80	1.71	5.37
0.8	46	7.00	1.53	3.95
0.8	100	6.66	0.99	4.68
0.8	215	5.96	0.79	4.39
0.8	464	5.90	0.45	5.00
0.8	1000	5.73	0.33	5.08
0.8	2154	5.40	0.20	5.00
0.9	10	12.29	2.27	7.76
0.9	22	8.82	1.72	5.38
0.9	46	6.50	1.35	3.80
0.9	100	5.97	0.85	4.27
0.9	215	5.59	0.58	4.43
0.9	464	5.60	0.37	4.86
0.9	1000	5.33	0.24	4.85
0.9	2154	5.17	0.17	4.83
1.0	10	13.36	2.06	9.25
1.0	22	10.04	1.35	7.35
1.0	46	6.95	1.22	4.51
1.0	100	6.26	0.81	4.64
1.0	215	6.12	0.58	4.97
1.0	464	5.97	0.37	5.22
1.0	1000	5.78	0.23	5.32
1.0	2154	5.60	0.16	5.28

Data Set Size = 5000

Upper Bound	Sample Size	Mean	Std. Dev.	Lower
0.1	10	9.64	2.37	4.90

0.1	22	6.24	1.63	2.99
0.1	46	4.84	1.38	2.07
0.1	100	3.38	0.77	1.85
0.1	215	2.73	0.60	1.53
0.1	464	2.70	0.42	1.86
0.1	1000	2.56	0.31	1.95
0.1	2154	2.61	0.21	2.20
0.1	4642	2.59	0.14	2.31
0.2	10	11.65	2.24	7.17
0.2	22	7.29	1.71	3.88
0.2	46	5.81	1.34	3.13
0.2	100	4.96	0.88	3.20
0.2	215	4.46	0.64	3.19
0.2	464	3.83	0.45	2.92
0.2	1000	3.69	0.24	3.21
0.2	2154	3.73	0.18	3.37
0.2	4642	3.79	0.12	3.55
0.3	10	10.37	2.42	5.53
0.3	22	8.05	1.70	4.65
0.3	46	8.09	1.54	5.02
0.3	100	7.01	0.95	5.10
0.3	215	5.28	0.65	3.97
0.3	464	4.55	0.54	3.48
0.3	1000	4.39	0.31	3.76
0.3	2154	4.66	0.24	4.18
0.3	4642	4.74	0.16	4.43
0.4	10	10.40	2.28	5.83
0.4	22	8.94	1.71	5.51
0.4	46	7.62	1.46	4.70
0.4	100	6.81	0.80	5.21
0.4	215	5.71	0.66	4.38
0.4	464	4.81	0.51	3.80
0.4	1000	4.79	0.28	4.23
0.4	2154	5.10	0.24	4.63
0.4	4642	5.15	0.16	4.84
0.5	10	11.17	2.13	6.91
0.5	22	9.47	1.81	5.86
0.5	46	8.43	1.43	5.57
0.5	100	6.73	0.92	4.89
0.5	215	5.85	0.61	4.62
0.5	464	5.13	0.46	4.21
0.5	1000	5.03	0.28	4.47
0.5	2154	5.53	0.24	5.05
0.5	4642	5.51	0.17	5.18
0.6	10	11.78	2.26	7.26

0.6	22	10.66	1.76	7.15
0.6	46	8.49	1.54	5.41
0.6	100	6.76	1.04	4.68
0.6	215	5.90	0.65	4.61
0.6	464	5.24	0.52	4.20
0.6	1000	5.13	0.32	4.48
0.6	2154	5.53	0.24	5.05
0.6	4642	5.56	0.16	5.25
0.7	10	11.97	2.33	7.31
0.7	22	11.15	1.82	7.51
0.7	46	8.33	1.55	5.23
0.7	100	6.63	1.08	4.46
0.7	215	5.80	0.67	4.45
0.7	464	5.05	0.51	4.02
0.7	1000	4.96	0.32	4.31
0.7	2154	5.47	0.25	4.97
0.7	4642	5.35	0.16	5.03
0.8	10	11.04	2.24	6.55
0.8	22	11.16	1.75	7.66
0.8	46	8.49	1.62	5.24
0.8	100	7.01	1.06	4.89
0.8	215	5.71	0.73	4.25
0.8	464	5.02	0.50	4.02
0.8	1000	4.52	0.30	3.92
0.8	2154	5.01	0.25	4.51
0.8	4642	5.06	0.16	4.74
0.9	10	10.79	1.94	6.91
0.9	22	10.28	2.01	6.25
0.9	46	7.38	1.47	4.44
0.9	100	6.28	1.08	4.12
0.9	215	4.90	0.63	3.65
0.9	464	4.44	0.39	3.67
0.9	1000	4.73	0.25	4.22
0.9	2154	4.89	0.20	4.50
0.9	4642	4.95	0.11	4.73
1.0	10	10.60	1.73	7.14
1.0	22	8.50	1.47	5.57
1.0	46	6.95	1.17	4.61
1.0	100	5.83	0.79	4.24
1.0	215	4.82	0.60	3.63
1.0	464	4.86	0.38	4.10
1.0	1000	5.14	0.26	4.61
1.0	2154	5.37	0.21	4.96
1.0	4642	5.45	0.12	5.20

Data Set Size = 10000

Upper Bound	Sample Size	Mean	Std. Dev.	Lower
0.1	10	9.30	1.85	5.60
0.1	22	6.33	1.33	3.67
0.1	46	6.66	1.50	3.67
0.1	100	5.04	0.91	3.21
0.1	215	3.24	0.58	2.07
0.1	464	2.96	0.40	2.16
0.1	1000	2.70	0.30	2.10
0.1	2154	2.53	0.20	2.13
0.1	4642	2.36	0.14	2.08
0.1	10000	2.51	0.09	2.33
0.2	10	9.34	2.08	5.18
0.2	22	7.68	1.48	4.73
0.2	46	6.40	1.06	4.29
0.2	100	5.40	0.86	3.68
0.2	215	4.44	0.52	3.40
0.2	464	4.44	0.39	3.65
0.2	1000	4.21	0.21	3.80
0.2	2154	3.95	0.13	3.68
0.2	4642	3.77	0.12	3.52
0.2	10000	3.89	0.07	3.74
0.3	10	8.75	2.27	4.21
0.3	22	7.69	1.55	4.58
0.3	46	6.62	1.36	3.90
0.3	100	6.00	0.88	4.24
0.3	215	5.54	0.64	4.26
0.3	464	4.72	0.47	3.77
0.3	1000	4.59	0.32	3.95
0.3	2154	4.52	0.21	4.11
0.3	4642	4.49	0.15	4.19
0.3	10000	4.79	0.09	4.61
0.4	10	10.17	2.27	5.63
0.4	22	9.47	1.63	6.20
0.4	46	7.41	1.30	4.81
0.4	100	5.98	0.99	4.01
0.4	215	5.32	0.66	4.00
0.4	464	5.03	0.45	4.12
0.4	1000	5.05	0.31	4.43
0.4	2154	5.14	0.18	4.79
0.4	4642	5.09	0.14	4.80
0.4	10000	5.26	0.08	5.09
0.5	10	9.32	2.29	4.75
0.5	22	8.35	1.88	4.59
0.5	46	7.98	1.30	5.37
0.5	100	6.24	1.02	4.20

0.5	215	5.50	0.63	4.24
0.5	464	5.58	0.47	4.64
0.5	1000	5.50	0.32	4.86
0.5	2154	5.43	0.21	5.01
0.5	4642	5.32	0.16	4.99
0.5	10000	5.44	0.09	5.26
0.6	10	10.47	2.36	5.75
0.6	22	8.95	1.88	5.18
0.6	46	8.35	1.21	5.93
0.6	100	6.45	1.09	4.27
0.6	215	5.65	0.73	4.18
0.6	464	5.66	0.51	4.64
0.6	1000	5.32	0.32	4.67
0.6	2154	5.26	0.19	4.88
0.6	4642	5.32	0.16	5.00
0.6	10000	5.43	0.09	5.24
0.7	10	10.79	2.32	6.15
0.7	22	8.72	2.02	4.68
0.7	46	8.21	1.27	5.66
0.7	100	6.43	1.04	4.35
0.7	215	5.62	0.73	4.16
0.7	464	5.64	0.54	4.57
0.7	1000	5.17	0.34	4.49
0.7	2154	5.06	0.21	4.65
0.7	4642	5.09	0.16	4.77
0.7	10000	5.18	0.10	4.99
0.8	10	10.30	2.19	5.92
0.8	22	9.12	1.72	5.68
0.8	46	8.18	1.29	5.60
0.8	100	6.18	1.03	4.13
0.8	215	5.52	0.70	4.13
0.8	464	5.29	0.56	4.18
0.8	1000	4.74	0.37	4.01
0.8	2154	4.64	0.18	4.28
0.8	4642	4.67	0.17	4.34
0.8	10000	4.72	0.11	4.50
0.9	10	10.85	2.11	6.63
0.9	22	8.81	1.83	5.16
0.9	46	7.84	1.36	5.11
0.9	100	6.39	0.99	4.40
0.9	215	5.45	0.63	4.18
0.9	464	4.85	0.41	4.04
0.9	1000	4.65	0.26	4.13
0.9	2154	4.82	0.16	4.49
0.9	4642	4.80	0.14	4.53

0.9	10000	4.93	0.09	4.76
1.0	10	11.47	2.30	6.87
1.0	22	9.25	1.38	6.49
1.0	46	6.69	1.16	4.36
1.0	100	5.65	0.88	3.90
1.0	215	5.77	0.62	4.53
1.0	464	5.20	0.38	4.43
1.0	1000	5.21	0.24	4.72
1.0	2154	5.34	0.15	5.04
1.0	4642	5.35	0.13	5.08
1.0	10000	5.43	0.09	5.25

Note: We report the mean, standard deviation, lower bound of the 95% confidence interval of the MAD values when bootstrapping with different sample sizes and different upper bounds of the uniform distribution which is one component of the convolution. The MAD values and estimates are reported in percentage. The data are the results of the sum of a Benford random variable and a uniform random variable with a lower bound of 0. The "Upper Bound" column specifies the upper bound of the uniform distribution. The "Sample Size" column indicates the size of the bootstrap sample. The "Mean" column reports the mean estimates of the MAD values based on 100 times of bootstrapping. The "Std. Dev." column reports the standard deviation of the MAD values based on 100 times of bootstrapping. The "Lower" column reports the lower bound of the 95% confidence interval. [DN] state that data show non-conformity to the Benford's distribution when the MAD value is greater than 1.2%.

8.3. **Tables of partial replacement with a normal random variable.**

TABLE 34. Chi-Square test statisitcs based on the distribution partially replaced by a normal random variable

Sample Size	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10% replaced										
10	8.0	7.8	7.9	8.0	8.1	8.2	8.1	8.1	8.1	8.1
22	8.1	7.9	7.9	7.9	8.2	8.3	8.1	8.1	8.1	8.1
46	8.1	7.9	7.9	8.0	8.1	8.0	8.1	8.1	8.0	8.0
100	8.2	8.0	8.0	8.1	8.1	8.0	8.1	8.1	8.1	8.0
215	8.1	8.1	8.0	8.1	8.1	8.0	8.2	8.3	8.1	8.1
464	8.2	8.1	8.2	8.3	8.3	8.4	8.3	8.4	8.2	8.2
1000	8.5	8.4	8.5	8.7	8.6	8.6	8.5	8.5	8.5	8.6
2154	9.2	9.0	9.0	9.1	9.1	9.1	9.1	9.1	9.0	9.1
4642	10.3	10.2	10.1	10.5	10.5	10.3	10.2	10.0	10.1	10.4
10000	12.9	12.8	12.9	13.2	13.3	12.9	12.5	12.4	12.6	13.0
20% replaced										
10	8.1	7.8	7.8	8.1	8.2	8.3	8.3	8.2	8.1	8.1
22	8.1	7.8	7.9	8.1	8.2	8.3	8.2	8.1	8.1	8.1
46	8.1	7.9	8.0	8.1	8.3	8.2	8.3	8.3	8.2	8.1
100	8.3	8.1	8.1	8.3	8.4	8.4	8.4	8.3	8.4	8.2
215	8.4	8.2	8.4	8.5	8.6	8.6	8.7	8.6	8.5	8.5
464	8.9	8.8	8.8	9.0	9.1	9.1	9.0	9.0	8.9	8.9
1000	9.9	9.9	9.7	10.3	10.3	10.2	9.9	9.8	10.0	10.0
2154	12.4	12.0	12.1	12.4	12.6	12.2	12.1	12.0	12.0	12.3
4642	17.1	16.9	17.1	17.8	18.0	17.0	16.3	16.2	16.5	17.2
10000	27.5	27.6	27.4	28.8	29.4	27.6	25.7	25.6	26.5	27.7
30% replaced										
10	8.2	7.6	7.9	8.2	8.3	8.4	8.4	8.3	8.2	8.1
22	8.3	7.8	7.9	8.1	8.4	8.4	8.3	8.3	8.1	8.2
46	8.2	7.9	7.9	8.2	8.4	8.4	8.5	8.4	8.3	8.2
100	8.7	8.1	8.4	8.6	8.7	8.8	8.7	8.6	8.7	8.5
215	9.0	8.7	8.7	9.1	9.2	9.1	9.2	9.0	9.0	9.1
464	10.1	9.8	10.0	10.2	10.4	10.3	10.2	10.1	10.0	10.1
1000	12.5	12.3	12.2	12.8	13.0	12.7	12.3	12.2	12.4	12.5
2154	17.7	17.3	17.4	18.0	18.6	17.7	16.9	16.8	17.0	17.6
4642	28.3	28.3	28.2	29.9	30.3	28.4	26.6	26.5	27.2	28.3
10000	52.0	52.1	51.9	55.2	55.8	51.6	47.5	47.5	49.8	52.3
40% replaced										
10	8.2	7.6	7.8	8.3	8.4	8.6	8.5	8.4	8.3	8.2
22	8.2	7.8	7.9	8.3	8.5	8.5	8.5	8.4	8.3	8.3
46	8.4	7.9	8.1	8.4	8.7	8.7	8.6	8.8	8.5	8.5
100	8.9	8.3	8.5	9.0	9.1	9.1	9.2	8.9	9.0	8.9
215	9.7	9.4	9.5	9.9	10.1	10.0	10.0	9.8	9.7	9.7
464	11.6	11.3	11.3	11.9	12.2	12.0	11.8	11.6	11.6	11.7
1000	16.0	15.5	15.8	16.6	16.9	16.1	15.4	15.2	15.6	15.9
2154	25.0	24.7	24.8	26.0	26.4	24.7	23.6	23.5	24.0	25.1
4642	44.2	44.3	44.2	46.7	47.6	44.1	41.0	40.8	42.4	44.3

10000	86.4	87.4	85.8	91.4	92.8	85.3	78.5	78.2	81.9	86.2
50% replaced										
10	8.3	7.6	7.7	8.2	8.6	8.8	8.6	8.5	8.5	8.2
22	8.5	7.8	8.0	8.3	8.7	8.8	8.7	8.6	8.5	8.3
46	8.6	8.0	8.3	8.7	9.2	9.0	9.0	9.0	8.8	8.8
100	9.5	8.7	9.0	9.5	9.8	9.7	9.6	9.4	9.5	9.4
215	10.6	10.2	10.3	10.8	11.3	10.9	10.8	10.7	10.6	10.7
464	13.8	13.2	13.4	14.1	14.6	14.0	13.6	13.5	13.5	13.6
1000	20.4	19.8	19.9	21.0	21.6	20.6	19.5	19.2	19.9	20.4
2154	34.6	34.0	33.9	36.1	36.7	34.2	32.3	32.2	32.9	34.4
4642	65.2	65.1	64.8	68.5	69.7	64.5	59.6	59.0	61.8	64.9
10000	130.3	131.9	130.1	138.7	140.4	128.7	118.2	117.5	123.9	130.6
60% replaced										
10	8.3	7.5	7.9	8.3	8.7	8.9	8.7	8.7	8.5	8.4
22	8.5	7.8	8.0	8.5	8.9	9.0	9.0	8.8	8.6	8.5
46	8.8	8.1	8.4	9.0	9.3	9.2	9.3	9.1	9.1	8.9
100	9.9	9.2	9.5	10.1	10.3	10.3	10.2	10.0	10.1	9.9
215	11.9	11.2	11.5	12.2	12.7	12.4	11.9	11.8	11.7	11.9
464	16.4	15.7	15.9	16.7	17.2	16.6	15.9	15.7	15.8	16.1
1000	25.7	25.4	25.2	26.8	27.5	25.9	24.3	24.0	25.1	25.9
2154	46.1	45.4	45.7	48.6	49.5	45.7	42.8	42.4	44.0	46.0
4642	89.7	90.3	89.5	95.2	96.9	88.8	81.6	81.5	85.3	89.8
10000	184.1	186.4	184.1	195.5	198.2	181.7	166.5	166.0	174.8	183.9
70% replaced										
10	8.3	7.6	7.8	8.5	8.8	9.2	9.0	8.8	8.7	8.5
22	8.7	7.7	8.0	8.7	9.1	9.0	9.1	9.0	8.8	8.6
46	9.2	8.3	8.7	9.2	9.7	9.6	9.6	9.4	9.5	9.2
100	10.5	9.6	10.0	10.7	11.2	11.0	10.9	10.6	10.8	10.6
215	13.3	12.5	12.7	13.7	14.0	13.6	13.2	13.0	13.1	13.2
464	19.4	18.5	18.9	19.8	20.6	19.4	18.7	18.5	18.7	19.4
1000	32.1	31.5	31.6	33.8	34.6	32.3	30.2	29.7	31.2	32.3
2154	59.9	59.2	59.2	63.1	64.1	59.5	55.1	54.7	57.2	59.8
4642	119.2	119.8	119.1	126.2	128.3	118.1	108.5	107.5	113.2	119.6
10000	247.6	251.1	247.5	263.2	266.3	243.9	223.4	223.0	234.4	247.1
80% replaced										
10	8.6	7.5	7.9	8.6	8.9	9.2	9.1	9.0	8.8	8.5
22	8.8	7.9	8.2	8.9	9.3	9.3	9.3	9.1	9.0	8.8
46	9.5	8.5	8.9	9.6	10.1	9.9	10.0	9.9	9.7	9.5
100	11.4	10.2	10.7	11.6	12.0	11.8	11.5	11.4	11.4	11.3
215	15.1	14.0	14.2	15.5	15.9	15.4	14.8	14.6	14.7	14.8
464	22.7	21.8	22.3	23.5	24.3	23.0	22.0	21.7	21.8	22.6
1000	39.4	38.8	38.9	41.5	42.5	39.8	36.7	36.2	38.0	39.4
2154	75.9	75.1	74.9	80.0	81.3	75.1	69.5	68.9	72.1	75.7
4642	153.2	154.5	152.9	162.7	165.7	151.0	139.3	138.5	145.5	153.2
10000	321.0	325.4	321.1	341.3	345.6	315.5	289.5	288.0	303.4	320.5
90% replaced										
10	8.7	7.5	7.8	8.6	9.1	9.4	9.4	9.1	8.9	8.6
22	9.0	7.9	8.3	9.1	9.5	9.6	9.6	9.4	9.2	8.9
46	9.9	8.8	9.3	10.1	10.7	10.4	10.5	10.2	10.1	10.0

100	12.1	11.0	11.4	12.4	13.1	12.7	12.4	12.2	12.3	12.2
215	16.7	15.6	16.0	17.3	17.8	17.2	16.4	16.2	16.3	16.6
464	26.6	25.6	26.2	27.5	28.6	26.8	25.3	25.2	25.6	26.6
1000	47.7	47.3	47.2	50.4	51.8	47.8	44.4	43.7	45.9	47.8
2154	93.3	93.4	93.1	99.1	100.8	92.8	85.6	85.0	89.4	93.5
4642	191.7	193.4	191.5	203.7	207.2	189.3	174.0	172.9	181.4	192.1
10000	404.2	409.7	403.8	430.2	435.4	397.5	363.4	363.0	382.8	403.3
					100% replaced					
10	8.8	7.5	7.9	8.8	9.3	9.6	9.5	9.3	9.0	8.8
22	9.3	8.0	8.5	9.5	10.0	10.0	9.9	9.8	9.5	9.3
46	10.4	9.1	9.7	10.6	11.2	11.0	11.0	10.7	10.7	10.5
100	13.1	11.8	12.3	13.4	14.1	13.7	13.3	13.1	13.3	13.1
215	18.9	17.5	18.0	19.4	20.1	19.3	18.3	18.0	18.2	18.7
464	31.2	29.8	30.5	32.1	33.4	31.2	29.4	29.2	29.8	30.9
1000	57.0	56.6	56.5	60.2	61.8	57.1	52.7	52.0	54.7	57.1
2154	113.4	113.2	113.0	120.8	122.5	112.6	103.7	103.0	108.2	113.6
4642	234.8	237.0	234.8	249.5	253.7	231.7	212.7	211.5	222.4	235.2
10000	497.5	504.4	497.1	529.0	535.8	488.8	447.3	446.0	470.5	496.1

Note: We report the Chi-Square test statistics on the frequency distribution of the first digit of the mixed distribution: $(1-m)\%$ of the data following a Benford's distribution and $m\%$ of the data following a normal distribution. The table reports different parameterizations of the normal distribution. The standard deviation of the normal distribution varies from 0.1 to 1.

TABLE 35. P-Values based on the Chi-Square test of the distribution partially replaced by a normal random variable

Sample Size	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
					10% replaced					
10	0.51	0.52	0.52	0.51	0.50	0.50	0.50	0.50	0.50	0.51
22	0.50	0.51	0.51	0.51	0.49	0.49	0.50	0.50	0.50	0.50
46	0.49	0.51	0.51	0.50	0.49	0.50	0.50	0.49	0.50	0.50
100	0.49	0.50	0.50	0.49	0.50	0.50	0.49	0.49	0.49	0.50
215	0.50	0.49	0.50	0.49	0.49	0.50	0.49	0.48	0.49	0.49
464	0.49	0.49	0.49	0.48	0.48	0.48	0.48	0.48	0.48	0.49
1000	0.47	0.47	0.47	0.46	0.46	0.46	0.47	0.47	0.47	0.46
2154	0.42	0.44	0.44	0.43	0.43	0.43	0.43	0.43	0.44	0.43
4642	0.36	0.36	0.37	0.35	0.35	0.37	0.36	0.38	0.37	0.35
10000	0.24	0.25	0.25	0.23	0.23	0.25	0.26	0.26	0.26	0.24
					20% replaced					
10	0.50	0.52	0.52	0.51	0.50	0.49	0.50	0.50	0.50	0.50
22	0.50	0.51	0.51	0.50	0.49	0.49	0.49	0.50	0.50	0.50
46	0.49	0.51	0.50	0.50	0.48	0.49	0.48	0.48	0.49	0.49
100	0.48	0.49	0.49	0.48	0.48	0.48	0.48	0.48	0.48	0.49
215	0.47	0.48	0.47	0.47	0.46	0.46	0.46	0.46	0.47	0.47
464	0.44	0.44	0.45	0.44	0.43	0.43	0.44	0.44	0.44	0.44
1000	0.38	0.38	0.39	0.37	0.36	0.37	0.39	0.39	0.38	0.38
2154	0.26	0.28	0.27	0.26	0.26	0.28	0.28	0.28	0.28	0.26
4642	0.12	0.12	0.12	0.11	0.11	0.12	0.14	0.14	0.13	0.12

10000	0.02	0.02	0.02	0.01	0.01	0.02	0.03	0.03	0.02	0.02
30% replaced										
10	0.50	0.53	0.51	0.50	0.49	0.48	0.49	0.49	0.49	0.50
22	0.49	0.51	0.51	0.50	0.48	0.48	0.49	0.49	0.50	0.49
46	0.49	0.51	0.50	0.49	0.48	0.48	0.47	0.48	0.48	0.49
100	0.46	0.49	0.47	0.46	0.46	0.46	0.46	0.47	0.46	0.47
215	0.43	0.45	0.45	0.43	0.43	0.43	0.43	0.44	0.44	0.43
464	0.37	0.39	0.37	0.37	0.36	0.36	0.37	0.37	0.38	0.37
1000	0.26	0.26	0.27	0.25	0.24	0.26	0.27	0.28	0.27	0.26
2154	0.11	0.11	0.11	0.11	0.09	0.12	0.13	0.13	0.13	0.11
4642	0.01	0.02	0.02	0.01	0.01	0.02	0.02	0.02	0.02	0.02
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
40% replaced										
10	0.49	0.53	0.52	0.49	0.49	0.47	0.48	0.49	0.49	0.49
22	0.49	0.51	0.51	0.48	0.47	0.48	0.47	0.48	0.48	0.49
46	0.47	0.50	0.49	0.48	0.46	0.46	0.46	0.46	0.47	0.48
100	0.45	0.47	0.47	0.44	0.43	0.43	0.43	0.44	0.44	0.44
215	0.39	0.41	0.40	0.38	0.37	0.38	0.39	0.39	0.39	0.40
464	0.30	0.31	0.31	0.28	0.27	0.29	0.29	0.30	0.29	0.29
1000	0.15	0.15	0.15	0.13	0.13	0.15	0.16	0.17	0.16	0.15
2154	0.03	0.03	0.03	0.02	0.02	0.03	0.04	0.04	0.04	0.03
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
50% replaced										
10	0.49	0.53	0.52	0.50	0.48	0.46	0.47	0.48	0.48	0.49
22	0.47	0.51	0.50	0.48	0.46	0.45	0.47	0.47	0.48	0.48
46	0.46	0.49	0.48	0.46	0.43	0.44	0.44	0.44	0.45	0.45
100	0.41	0.45	0.43	0.41	0.39	0.40	0.41	0.42	0.41	0.41
215	0.34	0.36	0.36	0.33	0.31	0.33	0.34	0.34	0.35	0.34
464	0.21	0.22	0.22	0.20	0.19	0.21	0.22	0.22	0.22	0.21
1000	0.07	0.07	0.07	0.06	0.06	0.07	0.08	0.08	0.08	0.07
2154	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.01	0.01	0.01
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
60% replaced										
10	0.49	0.53	0.52	0.49	0.47	0.46	0.47	0.47	0.48	0.48
22	0.47	0.51	0.50	0.47	0.45	0.45	0.44	0.45	0.46	0.47
46	0.45	0.49	0.47	0.44	0.43	0.43	0.43	0.43	0.44	0.45
100	0.38	0.42	0.40	0.37	0.37	0.37	0.38	0.38	0.38	0.38
215	0.28	0.31	0.30	0.27	0.26	0.27	0.29	0.29	0.29	0.28
464	0.14	0.15	0.14	0.13	0.12	0.14	0.16	0.16	0.15	0.14
1000	0.03	0.03	0.03	0.02	0.02	0.03	0.04	0.04	0.03	0.02
2154	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
70% replaced										
10	0.48	0.53	0.52	0.48	0.46	0.44	0.45	0.46	0.47	0.48
22	0.46	0.51	0.50	0.46	0.44	0.44	0.44	0.44	0.46	0.46
46	0.42	0.47	0.45	0.43	0.40	0.41	0.41	0.42	0.41	0.42

100	0.35	0.39	0.37	0.34	0.32	0.34	0.34	0.35	0.34	0.35
215	0.23	0.25	0.25	0.21	0.21	0.23	0.24	0.24	0.24	0.22
464	0.08	0.09	0.09	0.08	0.07	0.08	0.10	0.10	0.09	0.08
1000	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01
2154	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
80% replaced										
10	0.47	0.54	0.52	0.47	0.45	0.44	0.44	0.46	0.46	0.47
22	0.45	0.50	0.49	0.45	0.43	0.43	0.43	0.44	0.44	0.45
46	0.41	0.46	0.43	0.40	0.38	0.39	0.38	0.39	0.40	0.40
100	0.31	0.35	0.34	0.30	0.28	0.30	0.31	0.31	0.31	0.31
215	0.17	0.20	0.19	0.16	0.15	0.17	0.19	0.18	0.18	0.17
464	0.04	0.05	0.05	0.04	0.03	0.04	0.06	0.05	0.05	0.05
1000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2154	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
90% replaced										
10	0.46	0.54	0.52	0.47	0.44	0.43	0.43	0.45	0.45	0.47
22	0.44	0.50	0.48	0.43	0.42	0.41	0.42	0.42	0.43	0.44
46	0.38	0.44	0.41	0.38	0.35	0.37	0.36	0.38	0.38	0.38
100	0.27	0.31	0.30	0.26	0.24	0.26	0.27	0.28	0.27	0.27
215	0.13	0.15	0.14	0.12	0.11	0.13	0.15	0.14	0.14	0.13
464	0.02	0.02	0.02	0.02	0.02	0.02	0.03	0.03	0.02	0.02
1000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2154	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
100% replaced										
10	0.46	0.54	0.51	0.46	0.43	0.42	0.42	0.43	0.44	0.46
22	0.42	0.49	0.47	0.41	0.39	0.39	0.40	0.40	0.42	0.42
46	0.36	0.41	0.39	0.35	0.32	0.34	0.33	0.35	0.35	0.35
100	0.23	0.27	0.26	0.22	0.21	0.22	0.24	0.24	0.23	0.24
215	0.09	0.10	0.10	0.08	0.07	0.09	0.11	0.10	0.10	0.09
464	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
1000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2154	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4642	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
10000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Note: We report the p-values based on the Chi-Square tests of the first digit of the mixed distribution: $(1-m)\%$ of the data followng a Benford's distribution and $m\%$ of the data following a normal distribution. The table reports different parameterizations of the normal distribution. The standard deviation of the normal distribution varies from 0.1 to 1.

TABLE 36. MAD of the distribution partially replaced by a normal random variable

Sample Size	N(0, 0.1)	N(0, 0.2)	N(0, 0.3)	N(0, 0.4)	N(0, 0.5)	N(0, 0.6)	N(0, 0.7)	N(0, 0.8)	N(0, 0.9)	N(0, 1.0)
10% replaced										
10	0.076	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076	0.076
22	0.050	0.050	0.050	0.050	0.051	0.051	0.050	0.050	0.050	0.050
46	0.035	0.034	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035
100	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024
215	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016	0.016
464	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011	0.011
1000	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008	0.008
2154	0.006	0.005	0.005	0.006	0.006	0.005	0.005	0.005	0.005	0.006
4642	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004	0.004
10000	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003	0.003
20% replaced										
10	0.076	0.075	0.075	0.076	0.076	0.076	0.076	0.076	0.076	0.076
22	0.050	0.050	0.050	0.050	0.050	0.051	0.050	0.050	0.050	0.050
46	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035
100	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024	0.024
215	0.016	0.016	0.016	0.017	0.017	0.017	0.017	0.016	0.016	0.017
464	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012	0.012
1000	0.009	0.008	0.008	0.009	0.009	0.009	0.008	0.008	0.008	0.009
2154	0.007	0.006	0.006	0.007	0.007	0.007	0.006	0.006	0.006	0.007
4642	0.006	0.005	0.005	0.006	0.006	0.005	0.005	0.005	0.005	0.006
10000	0.005	0.005	0.005	0.005	0.005	0.005	0.004	0.004	0.005	0.005
30% replaced										
10	0.076	0.075	0.076	0.076	0.077	0.077	0.076	0.076	0.076	0.076
22	0.051	0.050	0.050	0.050	0.051	0.051	0.050	0.050	0.050	0.050
46	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035	0.035
100	0.025	0.024	0.024	0.025	0.025	0.025	0.024	0.024	0.024	0.024
215	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017	0.017
464	0.013	0.012	0.012	0.013	0.013	0.013	0.012	0.012	0.012	0.013
1000	0.010	0.010	0.009	0.010	0.010	0.010	0.009	0.009	0.010	0.010
2154	0.008	0.008	0.008	0.008	0.009	0.008	0.008	0.007	0.008	0.008
4642	0.008	0.007	0.007	0.008	0.008	0.007	0.007	0.006	0.007	0.008
10000	0.007	0.007	0.007	0.007	0.007	0.007	0.006	0.006	0.007	0.007
40% replaced										
10	0.076	0.075	0.076	0.077	0.077	0.077	0.077	0.077	0.077	0.076
22	0.050	0.050	0.050	0.051	0.051	0.051	0.051	0.051	0.051	0.051
46	0.036	0.035	0.035	0.036	0.036	0.036	0.036	0.036	0.035	0.036
100	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025	0.025
215	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018	0.018
464	0.014	0.013	0.013	0.014	0.014	0.014	0.013	0.013	0.013	0.014
1000	0.012	0.011	0.011	0.012	0.012	0.011	0.011	0.010	0.011	0.011
2154	0.010	0.010	0.010	0.010	0.010	0.010	0.009	0.009	0.010	0.010
4642	0.010	0.009	0.009	0.010	0.010	0.009	0.008	0.008	0.009	0.010
10000	0.009	0.009	0.009	0.009	0.010	0.009	0.008	0.008	0.009	0.009
50% replaced										
10	0.076	0.076	0.076	0.077	0.078	0.078	0.077	0.077	0.077	0.076

22	0.051	0.050	0.051	0.051	0.052	0.052	0.051	0.051	0.051	0.051
46	0.036	0.035	0.036	0.037	0.037	0.037	0.036	0.036	0.036	0.036
100	0.026	0.025	0.025	0.026	0.027	0.026	0.026	0.025	0.026	0.026
215	0.019	0.019	0.019	0.020	0.020	0.019	0.019	0.018	0.019	0.019
464	0.015	0.015	0.015	0.016	0.016	0.015	0.014	0.014	0.015	0.015
1000	0.013	0.013	0.013	0.014	0.014	0.013	0.012	0.012	0.013	0.013
2154	0.012	0.012	0.012	0.012	0.013	0.012	0.011	0.010	0.011	0.012
4642	0.012	0.012	0.011	0.012	0.012	0.012	0.010	0.010	0.011	0.012
10000	0.012	0.011	0.011	0.012	0.012	0.011	0.010	0.009	0.011	0.012
60% replaced										
10	0.077	0.075	0.076	0.077	0.078	0.078	0.077	0.078	0.077	0.077
22	0.051	0.050	0.051	0.052	0.052	0.052	0.052	0.052	0.052	0.051
46	0.037	0.036	0.036	0.037	0.037	0.037	0.037	0.037	0.037	0.037
100	0.027	0.026	0.026	0.027	0.027	0.027	0.026	0.026	0.026	0.027
215	0.021	0.020	0.020	0.021	0.021	0.021	0.020	0.019	0.020	0.021
464	0.017	0.016	0.016	0.017	0.018	0.017	0.016	0.016	0.016	0.017
1000	0.015	0.015	0.014	0.016	0.016	0.015	0.014	0.013	0.014	0.015
2154	0.014	0.014	0.014	0.015	0.015	0.014	0.013	0.012	0.013	0.014
4642	0.014	0.014	0.014	0.014	0.014	0.014	0.012	0.011	0.013	0.014
10000	0.014	0.014	0.014	0.014	0.014	0.014	0.012	0.011	0.013	0.014
70% replaced										
10	0.077	0.076	0.076	0.078	0.078	0.079	0.078	0.078	0.078	0.077
22	0.052	0.051	0.051	0.053	0.053	0.052	0.052	0.052	0.052	0.052
46	0.038	0.037	0.037	0.038	0.039	0.038	0.037	0.037	0.038	0.038
100	0.028	0.027	0.027	0.028	0.029	0.028	0.027	0.027	0.027	0.028
215	0.022	0.021	0.021	0.023	0.023	0.022	0.021	0.021	0.021	0.022
464	0.019	0.018	0.018	0.019	0.020	0.019	0.017	0.017	0.018	0.019
1000	0.017	0.017	0.017	0.018	0.018	0.017	0.015	0.015	0.016	0.017
2154	0.017	0.016	0.016	0.017	0.017	0.016	0.015	0.014	0.015	0.017
4642	0.016	0.016	0.016	0.016	0.017	0.016	0.014	0.013	0.015	0.016
10000	0.016	0.016	0.016	0.016	0.016	0.016	0.014	0.013	0.015	0.016
80% replaced										
10	0.078	0.076	0.077	0.079	0.079	0.079	0.078	0.078	0.078	0.078
22	0.053	0.051	0.052	0.053	0.054	0.053	0.053	0.052	0.053	0.052
46	0.038	0.037	0.038	0.039	0.040	0.038	0.038	0.038	0.038	0.038
100	0.029	0.028	0.028	0.030	0.030	0.029	0.028	0.028	0.028	0.029
215	0.024	0.023	0.022	0.024	0.025	0.023	0.022	0.022	0.023	0.024
464	0.021	0.020	0.020	0.021	0.021	0.021	0.019	0.018	0.020	0.021
1000	0.020	0.019	0.019	0.020	0.020	0.019	0.017	0.016	0.018	0.020
2154	0.019	0.018	0.018	0.019	0.019	0.019	0.017	0.015	0.018	0.019
4642	0.019	0.018	0.018	0.019	0.019	0.018	0.016	0.015	0.017	0.019
10000	0.018	0.018	0.018	0.018	0.019	0.018	0.016	0.015	0.017	0.018
90% replaced										
10	0.079	0.076	0.077	0.079	0.080	0.080	0.079	0.078	0.078	0.078
22	0.053	0.052	0.052	0.054	0.054	0.054	0.053	0.053	0.053	0.053
46	0.039	0.038	0.039	0.040	0.041	0.040	0.039	0.039	0.039	0.040
100	0.031	0.029	0.029	0.031	0.032	0.030	0.029	0.029	0.030	0.031
215	0.026	0.024	0.024	0.026	0.026	0.025	0.023	0.023	0.024	0.025
464	0.023	0.022	0.022	0.023	0.023	0.022	0.020	0.020	0.021	0.023
1000	0.022	0.021	0.021	0.022	0.022	0.021	0.019	0.018	0.020	0.022

2154	0.021	0.021	0.020	0.021	0.021	0.021	0.019	0.017	0.020	0.021
4642	0.021	0.020	0.020	0.021	0.021	0.021	0.018	0.017	0.019	0.021
10000	0.021	0.020	0.020	0.021	0.021	0.020	0.018	0.017	0.019	0.021
100% replaced										
10	0.079	0.076	0.077	0.080	0.080	0.080	0.080	0.079	0.079	0.079
22	0.054	0.052	0.053	0.056	0.056	0.055	0.054	0.054	0.054	0.054
46	0.041	0.039	0.040	0.042	0.042	0.041	0.040	0.039	0.040	0.041
100	0.032	0.030	0.031	0.033	0.033	0.032	0.030	0.030	0.031	0.032
215	0.027	0.026	0.026	0.028	0.028	0.027	0.025	0.024	0.026	0.027
464	0.025	0.024	0.024	0.025	0.025	0.025	0.022	0.021	0.023	0.025
1000	0.024	0.023	0.023	0.024	0.024	0.023	0.021	0.020	0.022	0.024
2154	0.023	0.023	0.023	0.023	0.024	0.023	0.021	0.019	0.022	0.023
4642	0.023	0.023	0.022	0.023	0.024	0.023	0.021	0.019	0.021	0.023
10000	0.023	0.023	0.022	0.023	0.023	0.023	0.020	0.018	0.021	0.023

Note: We report the MAD values of the mixed distribution: $(1-m)\%$ of the data following a Benford's distribution and $m\%$ of the data following a normal distribution. The table reports different parameterizations of the normal distribution. The values of the standard deviation of the normal distribution range from 0.1 to 1.

8.4. **Tables of bootstrapping on data replaced by a normal random variable.**

TABLE 37. MAD (in %) from partial replacement with a normal random variable

			Data Set Size = 2500									
Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.4	8.7	7.2	5.3	3.0	2.1	1.7	1.4	0.7		
0.1	0.2	0.5	9.4	7.0	5.4	3.1	2.4	1.7	1.1	0.6		
0.1	0.3	0.3	8.5	6.2	4.8	3.0	2.2	1.4	1.2	0.6		
0.1	0.4	0.4	8.7	7.2	5.0	2.9	2.1	1.4	1.1	0.6		
0.1	0.5	0.5	8.2	7.0	4.7	2.9	2.3	1.6	1.2	0.6		
0.1	0.6	0.3	8.2	7.3	5.1	3.1	2.3	1.5	1.1	0.6		
0.1	0.7	0.6	8.7	7.1	5.2	3.1	2.2	1.6	1.2	0.5		
0.1	0.8	0.4	8.7	6.4	5.3	3.0	2.3	1.6	1.4	0.6		
0.1	0.9	0.4	8.8	6.6	5.1	3.1	2.1	1.6	1.4	0.6		
0.1	1.0	0.5	8.7	7.2	5.3	3.0	2.1	1.7	1.4	0.7		
0.2	0.1	0.5	11.2	8.6	4.9	3.1	2.0	1.6	1.3	0.9		
0.2	0.2	0.6	11.2	7.9	5.2	3.3	2.0	1.5	1.2	0.9		
0.2	0.3	0.8	11.2	9.0	4.9	3.2	2.0	1.5	1.0	0.9		
0.2	0.4	0.7	11.2	8.4	4.9	3.5	2.1	1.6	1.1	0.8		
0.2	0.5	0.4	11.2	8.7	4.9	3.5	2.2	1.4	1.1	0.8		
0.2	0.6	1.2	11.2	8.2	5.2	3.5	2.2	1.6	1.1	0.7		
0.2	0.7	0.7	11.2	9.2	5.0	3.5	1.9	1.4	1.0	0.6		
0.2	0.8	0.6	11.2	7.7	4.5	3.4	2.0	1.5	1.1	0.8		
0.2	0.9	0.5	11.2	7.7	5.5	3.4	2.0	1.6	1.3	0.8		
0.2	1.0	0.8	11.2	8.6	4.9	3.1	2.0	1.6	1.3	0.9		
0.3	0.1	0.9	9.4	7.5	5.0	3.7	2.4	1.9	1.6	1.1		
0.3	0.2	1.0	8.0	5.1	5.0	2.8	1.8	1.5	1.3	0.9		
0.3	0.3	0.9	8.8	6.9	6.1	3.5	2.3	1.3	1.1	0.8		
0.3	0.4	0.9	8.9	7.2	6.3	4.2	2.5	1.4	1.1	0.8		
0.3	0.5	1.1	8.6	7.4	6.7	4.3	2.6	1.8	1.3	0.9		
0.3	0.6	0.9	9.4	6.9	6.0	4.1	2.7	1.6	1.3	0.9		
0.3	0.7	0.9	8.4	6.6	5.1	3.3	2.3	1.6	1.2	0.9		
0.3	0.8	0.6	8.5	6.1	4.6	3.7	2.6	1.8	1.4	0.9		
0.3	0.9	0.9	9.4	7.8	5.0	3.9	2.5	1.8	1.4	1.0		
0.3	1.0	1.1	9.4	7.5	5.0	3.7	2.4	1.9	1.6	1.1		
0.4	0.1	1.4	11.8	6.3	5.0	2.7	2.1	1.5	1.4	1.3		
0.4	0.2	0.7	10.8	6.5	4.5	4.4	2.8	1.7	1.1	1.2		
0.4	0.3	0.8	11.9	8.7	6.8	5.3	3.3	1.8	1.3	1.2		
0.4	0.4	1.1	11.7	8.2	7.4	5.6	3.4	2.2	1.5	1.2		
0.4	0.5	1.3	12.7	7.3	6.2	5.0	3.5	2.3	1.5	1.2		
0.4	0.6	1.3	13.1	7.6	5.9	4.5	3.2	2.2	1.6	1.2		
0.4	0.7	0.6	13.8	7.8	6.7	4.1	2.8	2.0	1.4	1.0		
0.4	0.8	0.8	12.6	6.9	5.6	3.7	2.6	1.7	1.3	1.0		
0.4	0.9	0.8	11.8	6.9	5.9	3.5	2.4	1.6	1.3	1.0		
0.4	1.0	0.6	11.8	6.3	5.0	2.7	2.1	1.5	1.4	1.3		
0.5	0.1	1.4	10.2	6.0	4.3	2.9	1.9	1.3	1.2	1.5		
0.5	0.2	1.2	9.2	6.6	3.8	2.9	2.3	1.4	1.3	1.1		
0.5	0.3	1.4	8.8	6.1	4.5	2.9	2.4	1.8	1.8	1.3		

0.5	0.4	1.4	11.0	5.7	4.7	3.1	2.2	1.5	1.7	1.4
0.5	0.5	1.4	11.7	6.2	5.5	2.9	2.0	1.6	1.7	1.6
0.5	0.6	1.1	9.8	5.7	5.0	2.6	2.1	1.6	1.7	1.5
0.5	0.7	0.9	9.4	5.1	4.8	2.9	2.2	1.5	1.4	1.3
0.5	0.8	1.0	9.2	5.4	4.4	3.7	2.1	1.3	1.4	1.3
0.5	0.9	1.0	9.7	5.9	4.5	3.2	1.9	1.3	1.2	1.3
0.5	1.0	1.3	10.2	6.0	4.3	2.9	1.9	1.3	1.2	1.5
0.6	0.1	1.6	9.5	6.7	5.3	3.1	2.4	2.2	1.7	1.6
0.6	0.2	1.6	8.5	6.6	4.8	3.4	2.3	1.7	1.7	1.4
0.6	0.3	1.5	11.5	8.4	4.6	3.7	2.5	2.4	1.8	1.5
0.6	0.4	1.1	8.4	6.8	4.6	3.7	2.8	2.1	1.9	1.6
0.6	0.5	1.8	9.5	7.4	5.7	4.0	3.0	2.2	1.8	1.7
0.6	0.6	1.5	9.8	7.1	4.4	3.3	2.1	1.8	1.5	1.6
0.6	0.7	1.4	10.1	5.5	4.2	2.9	2.0	1.9	1.4	1.4
0.6	0.8	1.2	9.0	6.0	4.1	2.9	2.0	2.0	1.4	1.4
0.6	0.9	1.4	9.3	6.8	4.5	3.3	2.0	1.9	1.4	1.5
0.6	1.0	1.6	9.5	6.7	5.3	3.1	2.4	2.2	1.7	1.6
0.7	0.1	1.7	10.1	5.5	4.4	3.3	2.6	2.4	2.2	1.7
0.7	0.2	1.7	10.8	5.7	3.9	3.2	2.2	2.0	1.9	1.5
0.7	0.3	1.7	9.4	5.6	4.5	2.8	2.2	2.1	2.0	1.8
0.7	0.4	1.6	8.5	6.1	5.1	3.0	2.4	2.1	1.8	1.7
0.7	0.5	1.5	8.4	5.8	5.2	4.1	2.9	2.4	2.1	1.9
0.7	0.6	1.5	9.8	6.5	4.9	4.1	2.8	2.3	2.2	1.8
0.7	0.7	1.5	9.4	5.4	4.7	3.5	3.1	2.2	2.0	1.6
0.7	0.8	1.3	11.6	6.4	4.7	3.7	2.7	2.1	2.0	1.5
0.7	0.9	1.3	11.0	6.0	4.8	3.2	2.7	2.4	2.2	1.6
0.7	1.0	1.7	10.1	5.5	4.4	3.3	2.6	2.4	2.2	1.7
0.8	0.1	2.2	9.8	7.5	5.6	3.5	3.5	2.7	2.4	2.0
0.8	0.2	1.8	9.1	5.7	4.5	2.8	2.2	2.1	2.1	1.9
0.8	0.3	1.6	8.6	6.2	4.2	3.3	2.6	2.3	2.0	1.9
0.8	0.4	2.2	11.1	7.8	5.6	3.7	3.4	2.4	2.1	2.0
0.8	0.5	1.9	11.5	8.3	7.2	4.0	3.0	2.8	2.3	2.1
0.8	0.6	2.0	9.7	7.1	5.8	3.8	3.1	2.6	2.2	2.0
0.8	0.7	1.3	9.0	6.6	5.1	3.8	3.4	2.4	2.1	1.8
0.8	0.8	1.5	8.3	6.6	4.6	3.2	2.6	2.2	1.7	1.7
0.8	0.9	1.9	8.6	7.4	5.3	3.5	2.8	2.4	2.0	1.7
0.8	1.0	1.9	9.8	7.5	5.6	3.5	3.5	2.7	2.4	2.0
0.9	0.1	2.3	10.0	6.3	4.8	3.6	3.7	2.6	2.2	2.2
0.9	0.2	2.0	8.9	5.0	4.1	3.3	3.3	2.5	2.0	2.2
0.9	0.3	2.1	8.1	6.2	4.5	2.8	3.6	2.9	2.1	2.3
0.9	0.4	2.1	8.6	5.7	5.2	3.7	3.9	2.8	2.2	2.3
0.9	0.5	1.9	10.3	6.5	4.8	3.7	4.0	3.0	2.3	2.3
0.9	0.6	2.2	9.0	6.2	5.4	3.7	3.7	2.8	2.1	2.2
0.9	0.7	1.6	9.6	6.1	5.4	3.7	3.2	2.4	2.0	1.9
0.9	0.8	1.5	8.8	6.1	4.3	3.2	2.7	2.1	1.7	1.8
0.9	0.9	2.1	8.4	6.0	4.3	3.2	3.0	2.4	1.8	2.0
0.9	1.0	2.2	10.0	6.3	4.8	3.6	3.7	2.6	2.2	2.2
1.0	0.1	2.9	12.0	9.5	5.9	5.0	3.8	3.7	2.9	2.4
1.0	0.2	2.2	10.3	7.9	6.4	5.0	4.0	3.7	2.8	2.3
1.0	0.3	2.3	9.6	7.3	6.2	4.3	2.8	3.0	2.7	2.2

1.0	0.4	2.2	11.4	9.1	7.3	5.5	4.0	3.8	2.8	2.4		
1.0	0.5	2.6	14.3	10.0	6.9	5.5	3.7	3.3	2.9	2.4		
1.0	0.6	2.3	11.2	10.2	7.0	5.3	4.0	3.7	2.8	2.3		
1.0	0.7	2.2	13.4	9.9	7.1	4.9	3.9	3.1	2.6	2.1		
1.0	0.8	2.2	13.0	10.6	7.9	4.9	3.7	3.3	2.4	2.0		
1.0	0.9	2.3	14.9	11.0	7.9	4.8	3.6	3.2	2.5	2.1		
1.0	1.0	2.4	12.0	9.5	5.9	5.0	3.8	3.7	2.9	2.4		

Data Set Size = 5000

Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.2	9.0	6.0	4.1	3.2	2.3	1.4	1.0	0.7	0.5	
0.1	0.2	0.4	8.6	5.6	3.8	3.0	2.2	1.5	1.1	0.8	0.4	
0.1	0.3	0.4	8.4	6.3	4.1	3.5	2.3	1.6	1.0	0.8	0.6	
0.1	0.4	0.3	8.9	6.2	4.1	3.3	2.2	1.6	1.0	0.9	0.6	
0.1	0.5	0.4	9.1	6.0	3.8	3.2	2.3	1.7	0.9	0.7	0.6	
0.1	0.6	0.6	8.8	6.2	4.3	3.2	2.3	1.6	0.9	0.7	0.5	
0.1	0.7	0.5	8.3	5.7	4.4	3.2	2.4	1.6	0.9	0.7	0.5	
0.1	0.8	0.3	8.8	5.4	4.0	3.0	2.2	1.5	0.9	0.7	0.4	
0.1	0.9	0.3	8.9	5.9	3.9	3.1	2.3	1.6	1.0	0.7	0.5	
0.1	1.0	0.3	9.0	6.0	4.1	3.2	2.3	1.4	1.0	0.7	0.5	
0.2	0.1	0.7	10.1	7.0	5.1	4.2	2.3	1.6	0.9	0.8	0.6	
0.2	0.2	0.6	10.4	6.1	4.4	4.0	2.2	1.9	1.2	0.7	0.6	
0.2	0.3	0.7	9.6	6.1	4.4	3.4	2.0	1.4	1.2	0.9	0.8	
0.2	0.4	0.7	9.6	6.3	4.5	3.8	2.1	1.4	1.4	1.1	0.9	
0.2	0.5	0.4	10.0	6.7	5.1	3.5	2.3	1.6	1.2	1.0	0.9	
0.2	0.6	0.5	10.3	5.8	5.4	4.3	2.2	1.7	1.2	1.0	0.8	
0.2	0.7	0.7	10.5	6.9	5.1	4.3	2.3	1.8	1.1	0.9	0.8	
0.2	0.8	0.2	10.0	6.2	4.8	3.9	2.2	1.8	1.0	0.9	0.7	
0.2	0.9	0.5	10.1	6.4	4.9	4.3	2.4	1.7	0.9	0.8	0.6	
0.2	1.0	0.8	10.1	7.0	5.1	4.2	2.3	1.6	0.9	0.8	0.6	
0.3	0.1	0.6	8.6	5.7	5.5	4.0	2.5	1.9	1.1	0.9	0.7	
0.3	0.2	0.6	8.6	5.9	4.1	3.0	2.2	1.9	1.1	0.8	0.8	
0.3	0.3	0.7	9.0	6.6	4.7	3.3	2.2	2.0	1.5	1.0	0.9	
0.3	0.4	1.0	8.1	6.0	5.4	4.5	2.6	2.2	1.7	1.1	0.9	
0.3	0.5	0.7	9.1	6.1	5.9	4.4	2.8	2.3	1.5	1.1	0.9	
0.3	0.6	0.9	9.5	5.9	5.6	4.2	3.1	2.8	1.6	1.1	0.8	
0.3	0.7	0.8	8.5	6.3	5.8	4.6	2.9	2.4	1.5	1.1	0.8	
0.3	0.8	0.6	8.3	5.9	6.1	4.7	2.5	2.3	1.3	0.9	0.7	
0.3	0.9	0.8	9.2	6.3	5.9	4.3	2.6	2.2	1.1	0.9	0.7	
0.3	1.0	0.7	8.6	5.7	5.5	4.0	2.5	1.9	1.1	0.9	0.7	
0.4	0.1	1.0	9.4	7.5	5.9	3.8	2.2	1.9	1.3	1.0	0.9	
0.4	0.2	1.0	11.4	9.7	4.4	3.2	2.2	1.8	1.2	0.8	0.8	
0.4	0.3	0.9	11.7	8.4	5.2	3.8	2.0	1.8	1.3	1.1	1.0	
0.4	0.4	1.0	11.0	7.2	5.4	3.5	2.7	2.0	1.3	1.4	1.2	
0.4	0.5	1.2	10.6	9.1	6.3	4.4	2.3	1.7	1.4	1.4	1.3	
0.4	0.6	0.8	9.8	7.8	6.0	4.3	2.2	1.8	1.4	1.4	1.2	
0.4	0.7	1.0	11.6	7.6	6.0	3.7	2.1	1.6	1.2	1.2	1.0	
0.4	0.8	1.0	10.0	6.5	5.7	3.3	2.4	1.7	1.2	1.2	0.9	
0.4	0.9	0.8	9.5	7.2	6.0	3.8	2.4	1.7	1.2	1.0	0.9	
0.4	1.0	0.6	9.4	7.5	5.9	3.8	2.2	1.9	1.3	1.0	0.9	
0.5	0.1	1.2	9.7	7.3	4.8	3.5	2.9	2.4	1.6	1.3	1.2	

0.5	0.2	1.1	11.4	8.1	5.3	3.9	3.0	2.2	1.5	1.4	1.1
0.5	0.3	1.3	10.8	7.3	5.1	3.9	2.6	1.9	1.5	1.7	1.4
0.5	0.4	1.5	9.8	8.1	6.1	4.1	2.2	1.8	1.5	1.6	1.5
0.5	0.5	1.2	10.4	7.3	5.5	4.0	2.4	1.9	1.7	1.7	1.5
0.5	0.6	1.5	10.8	6.5	4.4	2.9	1.8	1.6	1.4	1.4	1.4
0.5	0.7	1.0	10.9	6.3	4.6	3.2	2.0	1.7	1.5	1.3	1.3
0.5	0.8	0.8	10.0	7.2	5.5	3.5	2.6	2.1	1.5	1.2	1.2
0.5	0.9	1.3	9.8	7.1	5.3	3.4	3.0	2.4	1.6	1.3	1.1
0.5	1.0	1.0	9.7	7.3	4.8	3.5	2.9	2.4	1.6	1.3	1.2
0.6	0.1	1.5	9.4	6.1	3.8	3.4	2.1	1.6	1.4	1.2	1.4
0.6	0.2	1.4	10.4	7.7	4.0	3.5	2.5	1.5	1.3	1.2	1.3
0.6	0.3	1.8	10.7	7.6	4.6	3.5	2.2	1.9	1.6	1.4	1.6
0.6	0.4	1.3	8.8	6.0	4.4	4.0	2.0	2.2	1.8	1.4	1.7
0.6	0.5	1.2	10.6	6.6	4.5	3.7	2.4	1.8	1.7	1.3	1.7
0.6	0.6	1.6	10.7	5.8	4.2	3.3	2.5	2.0	1.7	1.2	1.6
0.6	0.7	1.1	12.2	7.0	4.9	3.3	2.4	2.0	1.6	1.2	1.4
0.6	0.8	1.2	10.6	7.4	5.5	3.3	2.8	2.0	1.6	1.1	1.3
0.6	0.9	1.1	9.9	6.5	4.1	3.3	2.2	1.5	1.4	1.1	1.2
0.6	1.0	1.4	9.4	6.1	3.8	3.4	2.1	1.6	1.4	1.2	1.4
0.7	0.1	1.7	11.0	7.5	4.1	3.2	2.9	1.8	1.6	1.5	1.7
0.7	0.2	1.6	9.2	7.3	4.3	3.4	2.6	2.1	1.6	1.5	1.6
0.7	0.3	1.7	11.5	6.8	4.9	3.7	2.8	2.2	1.7	1.5	1.7
0.7	0.4	1.6	10.9	7.6	4.7	4.2	3.2	2.3	1.8	1.7	1.8
0.7	0.5	1.5	11.2	6.9	4.9	3.6	2.9	2.1	1.7	1.5	1.8
0.7	0.6	1.5	11.2	7.7	4.6	3.3	2.8	2.1	1.9	1.6	1.7
0.7	0.7	1.5	12.4	8.1	5.0	3.6	2.8	1.8	1.7	1.3	1.5
0.7	0.8	1.1	12.3	8.6	4.5	3.0	2.8	1.6	1.4	1.3	1.5
0.7	0.9	1.5	11.0	6.7	4.8	3.2	3.0	1.6	1.3	1.4	1.5
0.7	1.0	1.6	11.0	7.5	4.1	3.2	2.9	1.8	1.6	1.5	1.7
0.8	0.1	2.1	9.0	7.9	6.1	5.1	3.0	2.6	1.9	1.9	1.8
0.8	0.2	2.0	9.3	7.1	5.0	3.9	2.3	2.1	1.7	1.7	1.7
0.8	0.3	1.5	8.1	7.4	5.6	3.3	2.2	2.0	1.5	1.6	1.7
0.8	0.4	1.8	9.9	7.9	6.0	3.8	2.6	2.2	1.8	1.9	2.0
0.8	0.5	1.6	8.8	7.9	6.0	4.4	2.8	2.4	2.0	1.9	1.9
0.8	0.6	1.9	9.7	5.9	4.3	3.6	2.5	2.1	1.7	2.0	1.9
0.8	0.7	1.6	10.2	7.2	4.8	3.7	2.2	2.1	1.5	1.8	1.7
0.8	0.8	1.4	10.6	6.6	4.7	3.8	2.5	2.2	1.6	1.7	1.7
0.8	0.9	1.5	9.4	7.4	4.4	4.3	2.6	2.6	1.8	1.8	1.7
0.8	1.0	1.8	9.0	7.9	6.1	5.1	3.0	2.6	1.9	1.9	1.8
0.9	0.1	1.9	9.1	6.9	4.7	4.6	3.6	3.5	2.9	2.7	2.2
0.9	0.2	2.3	9.7	6.7	5.2	4.7	3.6	3.1	2.5	2.4	2.1
0.9	0.3	1.9	9.5	6.9	5.6	5.2	3.5	3.4	2.7	2.4	2.1
0.9	0.4	2.3	9.3	5.9	5.3	4.7	3.5	3.4	2.6	2.5	2.2
0.9	0.5	2.2	9.7	7.1	5.3	5.0	3.7	3.5	2.8	2.6	2.2
0.9	0.6	2.0	10.2	5.7	4.5	4.4	3.1	3.0	2.7	2.5	2.1
0.9	0.7	1.6	9.0	6.2	4.5	3.8	2.9	2.8	2.5	2.2	2.0
0.9	0.8	1.4	8.3	5.8	4.9	3.6	3.1	2.9	2.3	2.2	1.9
0.9	0.9	1.7	9.7	6.8	4.7	4.5	3.9	3.4	2.7	2.5	2.1
0.9	1.0	2.2	9.1	6.9	4.7	4.6	3.6	3.5	2.9	2.7	2.2
1.0	0.1	2.3	9.9	7.1	4.9	3.5	2.9	3.4	2.8	2.7	2.5

1.0	0.2	2.1	9.0	6.8	5.0	3.1	2.6	3.1	2.6	2.7	2.4	
1.0	0.3	2.2	9.6	6.2	5.5	3.5	3.1	3.1	2.3	2.3	2.3	
1.0	0.4	2.2	10.8	7.4	4.9	3.9	3.5	3.2	2.7	2.7	2.4	
1.0	0.5	2.2	10.0	7.4	5.2	3.6	3.1	3.2	2.6	2.6	2.4	
1.0	0.6	1.8	10.6	7.2	4.6	3.5	2.8	2.9	2.5	2.6	2.3	
1.0	0.7	2.0	9.5	6.4	4.6	3.8	2.9	2.4	2.1	2.3	2.1	
1.0	0.8	1.7	10.2	6.7	4.4	3.2	2.7	2.7	2.3	2.3	2.0	
1.0	0.9	2.2	10.4	7.0	4.8	3.2	2.9	3.1	2.5	2.5	2.3	
1.0	1.0	2.3	9.9	7.1	4.9	3.5	2.9	3.4	2.8	2.7	2.5	

			Data Set Size = 5000									
Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.3	9.2	6.8	4.6	3.8	2.8	1.7	1.2	0.7	0.4	0.4
0.1	0.2	0.4	9.2	6.6	4.7	3.6	2.8	1.9	1.3	0.8	0.5	0.5
0.1	0.3	0.3	9.2	6.0	4.5	3.5	2.7	1.8	1.3	0.9	0.5	0.4
0.1	0.4	0.5	9.2	6.1	4.3	3.9	3.0	2.0	1.5	0.9	0.6	0.4
0.1	0.5	0.4	9.2	6.4	4.8	4.2	3.1	2.0	1.3	0.8	0.5	0.3
0.1	0.6	0.3	9.2	6.4	4.3	3.4	2.7	1.7	1.3	0.8	0.4	0.3
0.1	0.7	0.3	9.2	6.4	4.3	3.5	2.9	1.7	1.3	0.9	0.5	0.3
0.1	0.8	0.2	9.2	6.4	4.7	3.9	2.9	1.7	1.3	0.8	0.4	0.3
0.1	0.9	0.3	9.2	6.4	4.4	3.5	2.7	1.6	1.2	0.7	0.4	0.4
0.1	1.0	0.4	9.2	6.8	4.6	3.8	2.8	1.7	1.2	0.7	0.4	0.4
0.2	0.1	0.3	9.5	7.4	5.1	2.7	2.8	2.2	1.2	1.0	0.7	0.5
0.2	0.2	0.5	8.9	6.3	4.6	3.0	2.7	1.8	1.1	1.0	0.8	0.6
0.2	0.3	0.6	9.7	6.4	5.0	2.7	2.1	1.6	1.0	0.7	0.6	0.5
0.2	0.4	0.6	8.5	6.6	4.2	2.7	1.8	1.3	0.9	0.7	0.6	0.6
0.2	0.5	0.5	9.8	6.6	4.0	3.1	1.8	1.7	1.1	0.6	0.6	0.5
0.2	0.6	0.6	9.0	8.1	4.7	2.9	2.0	1.8	1.1	0.7	0.5	0.5
0.2	0.7	0.6	9.1	7.5	4.0	3.0	2.0	1.7	1.1	0.8	0.6	0.4
0.2	0.8	0.6	9.2	7.3	4.1	2.8	2.2	1.9	1.1	0.9	0.6	0.4
0.2	0.9	0.5	9.5	6.6	5.1	2.8	2.4	2.0	1.2	0.9	0.7	0.5
0.2	1.0	0.6	9.5	7.4	5.1	2.7	2.8	2.2	1.2	1.0	0.7	0.5
0.3	0.1	0.4	10.6	7.9	4.7	3.9	2.0	1.5	1.2	1.0	0.7	0.6
0.3	0.2	0.7	9.9	7.9	4.7	2.9	2.0	1.8	1.5	1.0	0.7	0.8
0.3	0.3	0.8	9.3	7.8	4.8	3.9	2.2	1.5	1.4	0.8	0.7	0.7
0.3	0.4	0.6	10.8	6.6	4.4	3.7	2.2	1.4	1.2	1.0	0.9	0.8
0.3	0.5	0.6	10.4	7.6	5.4	3.4	1.9	1.5	1.2	0.9	0.8	0.7
0.3	0.6	0.8	9.4	6.9	4.5	3.0	1.9	1.6	1.3	0.9	0.8	0.6
0.3	0.7	0.6	11.8	7.9	4.9	3.3	2.3	1.7	1.3	1.0	0.7	0.6
0.3	0.8	0.6	9.7	7.9	6.0	3.9	2.1	1.6	1.3	1.0	0.7	0.5
0.3	0.9	0.6	10.1	7.5	5.2	3.6	1.8	1.4	1.2	1.0	0.8	0.6
0.3	1.0	0.7	10.6	7.9	4.7	3.9	2.0	1.5	1.2	1.0	0.7	0.6
0.4	0.1	1.0	11.5	6.6	3.9	3.4	2.7	1.8	1.1	1.0	0.8	0.8
0.4	0.2	0.9	14.2	8.1	3.8	3.0	2.4	1.9	1.5	1.0	1.0	0.9
0.4	0.3	0.8	13.8	8.3	4.1	3.0	2.0	2.0	1.8	1.2	1.1	0.9
0.4	0.4	0.9	12.4	7.0	4.4	3.8	2.2	1.8	1.9	1.4	1.1	1.0
0.4	0.5	0.8	9.7	6.2	4.4	3.6	2.5	1.8	1.8	1.3	1.0	0.8
0.4	0.6	0.8	10.0	6.0	4.3	3.1	2.1	1.7	1.5	1.0	1.0	0.8
0.4	0.7	0.8	10.2	5.7	4.7	3.0	2.0	1.6	1.4	1.0	0.8	0.7
0.4	0.8	1.0	10.6	6.1	4.1	3.4	2.5	1.8	1.3	1.0	0.8	0.7
0.4	0.9	0.8	12.9	7.2	3.8	2.9	2.3	1.9	1.2	1.0	0.8	0.8

0.4	1.0	0.8	11.5	6.6	3.9	3.4	2.7	1.8	1.1	1.0	0.8	0.8
0.5	0.1	1.2	9.1	5.6	4.9	2.5	2.2	1.7	1.5	1.1	1.0	1.0
0.5	0.2	1.3	8.9	6.0	4.4	3.2	2.5	1.8	1.4	1.2	1.0	1.1
0.5	0.3	1.1	9.1	7.2	4.7	3.0	2.0	1.7	1.1	1.0	1.0	1.0
0.5	0.4	1.1	10.2	7.7	5.6	2.8	2.7	1.9	1.2	1.2	1.2	1.2
0.5	0.5	1.1	9.4	7.0	5.4	2.8	2.1	2.0	1.3	1.2	1.2	1.0
0.5	0.6	1.0	8.6	6.6	5.3	2.6	2.5	2.0	1.3	1.2	1.1	1.0
0.5	0.7	1.0	8.9	6.6	5.1	2.9	2.3	1.6	1.2	0.9	0.9	0.9
0.5	0.8	1.0	8.4	6.5	5.2	3.0	2.4	1.7	1.3	1.0	0.9	0.8
0.5	0.9	1.2	8.6	5.6	4.7	3.2	2.3	1.7	1.5	1.2	1.0	0.9
0.5	1.0	1.1	9.1	5.6	4.9	2.5	2.2	1.7	1.5	1.1	1.0	1.0
0.6	0.1	1.4	9.1	6.9	4.4	2.7	2.3	2.1	1.6	1.3	1.2	1.3
0.6	0.2	1.4	9.3	5.8	4.6	3.1	2.3	2.0	1.8	1.6	1.3	1.3
0.6	0.3	1.3	11.9	5.9	5.0	3.2	2.3	1.9	1.4	1.3	1.1	1.2
0.6	0.4	1.3	12.1	6.6	4.2	3.0	2.2	2.0	1.8	1.5	1.3	1.4
0.6	0.5	1.3	12.7	7.1	5.3	2.9	2.0	1.6	1.6	1.4	1.3	1.3
0.6	0.6	1.5	12.6	6.8	4.4	2.9	2.0	1.5	1.6	1.3	1.2	1.2
0.6	0.7	1.2	12.5	7.1	4.3	3.3	2.2	1.7	1.5	1.2	1.1	1.1
0.6	0.8	1.3	9.4	7.4	4.5	2.9	2.3	2.0	1.6	1.2	1.0	1.0
0.6	0.9	1.3	9.7	6.4	4.7	3.6	2.6	2.1	1.6	1.2	1.1	1.2
0.6	1.0	1.3	9.1	6.9	4.4	2.7	2.3	2.1	1.6	1.3	1.2	1.3
0.7	0.1	1.5	10.9	7.9	5.0	3.1	2.2	1.7	1.7	1.5	1.6	1.5
0.7	0.2	1.5	9.8	7.1	4.1	2.8	2.0	1.6	1.5	1.7	1.6	1.6
0.7	0.3	1.6	11.0	7.7	4.5	4.6	2.5	1.8	1.6	1.6	1.6	1.4
0.7	0.4	1.5	10.0	8.4	4.9	4.1	2.7	1.9	1.6	1.6	1.7	1.6
0.7	0.5	1.8	11.6	8.4	5.2	4.3	3.0	2.1	1.8	1.5	1.6	1.4
0.7	0.6	1.6	11.8	8.7	5.1	4.8	3.2	2.2	1.7	1.5	1.5	1.4
0.7	0.7	1.4	10.0	7.8	4.7	3.6	2.9	2.0	1.5	1.4	1.3	1.3
0.7	0.8	1.3	11.3	8.1	4.6	3.4	2.8	2.0	1.6	1.5	1.5	1.2
0.7	0.9	1.4	11.0	7.8	4.5	3.2	2.1	1.9	1.7	1.5	1.6	1.4
0.7	1.0	1.5	10.9	7.9	5.0	3.1	2.2	1.7	1.7	1.5	1.6	1.5
0.8	0.1	1.8	10.8	8.0	5.6	3.7	2.6	2.0	1.9	1.7	1.6	1.7
0.8	0.2	1.7	9.1	5.6	4.3	3.4	3.2	2.3	2.3	1.8	1.8	1.7
0.8	0.3	1.6	9.5	6.1	4.4	3.0	2.5	1.8	2.3	1.7	1.6	1.6
0.8	0.4	1.9	9.0	6.2	5.3	3.4	2.7	2.1	2.2	1.7	1.7	1.7
0.8	0.5	1.8	9.5	6.3	6.0	3.2	2.5	1.9	2.1	1.6	1.6	1.6
0.8	0.6	1.6	10.6	6.7	4.7	2.9	2.5	2.0	2.2	1.6	1.6	1.6
0.8	0.7	1.7	10.1	7.7	5.8	4.1	2.4	2.2	2.0	1.5	1.4	1.4
0.8	0.8	1.5	7.8	5.9	4.6	3.7	2.3	1.9	1.9	1.4	1.3	1.4
0.8	0.9	1.9	8.7	6.8	5.2	3.6	2.6	1.9	1.8	1.6	1.4	1.6
0.8	1.0	1.7	10.8	8.0	5.6	3.7	2.6	2.0	1.9	1.7	1.6	1.7
0.9	0.1	2.0	9.4	7.5	4.9	3.3	2.8	2.8	2.4	2.1	1.9	1.9
0.9	0.2	2.1	9.6	6.3	5.0	2.8	2.0	2.4	2.3	2.1	2.0	2.0
0.9	0.3	2.0	9.1	5.8	4.6	2.8	2.4	2.4	2.3	1.9	1.7	1.8
0.9	0.4	2.1	9.6	6.1	4.8	3.8	3.2	2.8	2.4	2.0	2.0	2.0
0.9	0.5	2.0	9.6	5.9	4.6	3.4	2.7	2.7	2.6	2.0	1.8	1.9
0.9	0.6	2.2	10.1	6.4	4.4	3.3	2.2	2.6	2.4	2.1	1.9	1.9
0.9	0.7	1.7	11.0	7.0	4.9	3.8	2.4	2.2	2.1	1.6	1.7	1.7
0.9	0.8	1.7	10.2	5.7	4.1	3.5	3.0	2.6	2.2	1.6	1.6	1.5
0.9	0.9	2.0	9.7	6.3	4.6	4.0	3.0	2.4	2.1	1.9	1.7	1.8

0.9	1.0	2.0	9.4	7.5	4.9	3.3	2.8	2.8	2.4	2.1	1.9	1.9
1.0	0.1	2.3	9.4	7.0	5.4	4.5	4.5	3.6	3.0	2.6	2.3	2.2
1.0	0.2	2.2	8.9	6.5	5.8	4.9	4.0	3.5	2.8	2.6	2.3	2.2
1.0	0.3	2.3	10.0	6.4	4.5	4.1	3.8	3.5	3.0	2.7	2.3	2.0
1.0	0.4	2.3	10.2	6.4	6.1	5.3	4.7	3.9	3.0	2.7	2.3	2.2
1.0	0.5	2.3	9.1	6.0	5.2	4.4	4.1	3.6	3.0	2.8	2.4	2.1
1.0	0.6	2.3	9.6	7.0	5.9	5.1	3.8	3.3	2.9	2.8	2.4	2.2
1.0	0.7	2.1	8.9	6.6	6.4	4.8	3.4	3.1	2.9	2.6	2.2	1.9
1.0	0.8	1.9	9.3	6.7	5.2	4.6	3.7	2.9	2.6	2.4	1.9	1.7
1.0	0.9	2.3	9.2	6.1	5.1	4.6	4.1	3.3	2.7	2.4	2.0	2.0
1.0	1.0	2.2	9.4	7.0	5.4	4.5	4.5	3.6	3.0	2.6	2.3	2.2

Note: We report the MAD values (in %) of data sets whose values follow a Benford's distribution partially replaced by data following a normal distribution. The normal distribution is of mean 0 and standard deviation ranging from 0.1 to 1. The values are based on the average of 100 bootstrap iterations.

TABLE 38. P-values based on Chi-Square tests from partial replacement with a normal random variable

			Data Set Size = 2500									
Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.2	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.3	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.4	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.5	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.6	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.7	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.8	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.9	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	1.0	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.1	1.000	0.975	0.989	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.2	1.000	0.975	0.989	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.3	1.000	0.975	0.987	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.4	1.000	0.975	0.988	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.5	1.000	0.975	0.981	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.6	1.000	0.975	0.989	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.7	1.000	0.975	0.978	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.8	1.000	0.975	0.990	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.9	1.000	0.975	0.989	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	1.0	1.000	0.975	0.989	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.1	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.2	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.3	1.000	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.4	1.000	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.5	1.000	0.989	0.998	0.999	1.000	1.000	1.000	1.000	1.000		
0.3	0.6	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.7	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.8	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000		

0.3	0.9	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.0	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.1	1.000	0.972	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.2	1.000	0.981	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.3	1.000	0.969	0.996	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.4	1.000	0.979	0.998	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.5	1.000	0.974	0.999	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.6	1.000	0.963	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.7	1.000	0.949	0.998	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.8	1.000	0.961	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.9	1.000	0.972	0.999	0.999	1.000	1.000	1.000	1.000	1.000
0.4	1.0	1.000	0.972	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.1	1.000	0.972	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.2	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.3	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.4	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.5	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.6	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.7	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.9	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.0	1.000	0.972	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.1	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.2	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	1.000	0.972	0.995	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.4	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.5	1.000	0.974	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.6	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.7	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.8	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.9	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.0	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.1	1.000	0.980	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.2	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.3	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.4	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.5	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.6	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.7	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.8	1.000	0.963	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.9	1.000	0.963	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.0	1.000	0.980	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.1	1.000	0.966	0.993	0.999	1.000	1.000	1.000	1.000	1.000
0.8	0.2	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.3	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.4	1.000	0.983	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.5	1.000	0.973	0.996	0.999	1.000	1.000	1.000	1.000	1.000
0.8	0.6	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.7	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.8	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000

0.8	0.9	1.000	0.988	0.997	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.0	1.000	0.966	0.993	0.999	1.000	1.000	1.000	1.000	1.000
0.9	0.1	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.2	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.3	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.4	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.5	1.000	0.980	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.6	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.7	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.8	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.9	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.9	1.0	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.1	1.000	0.937	0.946	0.998	1.000	1.000	1.000	1.000	1.000
1.0	0.2	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.3	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.4	1.000	0.978	0.994	0.999	1.000	1.000	1.000	1.000	1.000
1.0	0.5	1.000	0.940	0.983	0.999	1.000	1.000	1.000	1.000	1.000
1.0	0.6	1.000	0.962	0.980	0.999	1.000	1.000	1.000	1.000	1.000
1.0	0.7	1.000	0.951	0.979	0.999	1.000	1.000	1.000	1.000	1.000
1.0	0.8	1.000	0.937	0.965	0.998	1.000	1.000	1.000	1.000	1.000
1.0	0.9	1.000	0.883	0.949	0.996	1.000	1.000	1.000	1.000	1.000
1.0	1.0	1.000	0.937	0.946	0.998	1.000	1.000	1.000	1.000	1.000

						Data Set Size = 5000						
Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.2	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.3	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.4	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.5	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.6	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.7	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.8	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.9	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	1.0	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.1	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.2	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.3	1.000	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.4	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.5	1.000	0.989	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.6	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.7	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.8	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.9	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	1.0	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.1	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.2	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.3	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.4	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.5	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.6	1.000	0.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

0.3	0.7	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.8	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.9	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.0	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.1	1.000	0.985	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.2	1.000	0.978	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.3	1.000	0.969	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.4	1.000	0.986	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.5	1.000	0.982	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.6	1.000	0.983	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.7	1.000	0.972	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.8	1.000	0.989	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.9	1.000	0.983	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.0	1.000	0.985	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.1	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.2	1.000	0.984	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.3	1.000	0.982	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.4	1.000	0.981	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.5	1.000	0.961	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.6	1.000	0.959	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.7	1.000	0.965	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	0.967	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.9	1.000	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.0	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.1	1.000	0.982	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.2	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.4	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.5	1.000	0.966	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.6	1.000	0.969	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.7	1.000	0.916	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.8	1.000	0.979	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.9	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.0	1.000	0.982	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.1	1.000	0.966	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.2	1.000	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.3	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.4	1.000	0.977	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.5	1.000	0.974	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.6	1.000	0.954	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.7	1.000	0.937	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.8	1.000	0.933	0.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.9	1.000	0.954	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.0	1.000	0.966	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.1	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.2	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.3	1.000	0.997	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.4	1.000	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.5	1.000	0.989	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.6	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

0.8	0.7	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	0.8	1.000	0.980	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	0.9	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	1.0	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.1	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.2	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.3	1.000	0.988	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.4	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.5	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.6	1.000	0.974	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.7	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.8	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.9	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	1.0	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.1	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.2	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.3	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.4	1.000	0.988	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.5	1.000	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.6	1.000	0.978	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.7	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.8	1.000	0.983	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.9	1.000	0.981	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	1.0	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
						Data Set Size = 10000						
Proportion replaced	Std. Dev.	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.2	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.3	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.4	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.5	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.6	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.7	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.8	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.9	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1.0	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.1	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.2	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.3	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.4	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.5	1.000	0.982	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.6	1.000	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.7	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.8	1.000	0.990	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.9	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1.0	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.1	1.000	0.977	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.2	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.3	1.000	0.988	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.4	1.000	0.971	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	0.3	0.5	1.000	0.973	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	0.6	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	0.7	1.000	0.971	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	0.8	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	0.9	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.3	1.0	1.000	0.977	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.1	1.000	0.983	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.2	1.000	0.941	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.3	1.000	0.901	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.4	1.000	0.928	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.5	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.6	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.7	1.000	0.982	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.8	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	0.9	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.4	1.0	1.000	0.983	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.1	1.000	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.2	1.000	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.3	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.4	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.5	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.6	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.7	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.8	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	0.9	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.5	1.0	1.000	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.1	1.000	0.977	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.2	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.3	1.000	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.4	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.5	1.000	0.974	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.6	1.000	0.937	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.7	1.000	0.958	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.8	1.000	0.984	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	0.9	1.000	0.954	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.6	1.0	1.000	0.977	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.1	1.000	0.979	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.2	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.3	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.4	1.000	0.991	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.5	1.000	0.970	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.6	1.000	0.959	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.7	1.000	0.981	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.8	1.000	0.974	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	0.9	1.000	0.985	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.7	1.0	1.000	0.979	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.1	1.000	0.964	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.2	1.000	0.981	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.3	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.4	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

	0.8	0.5	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.6	1.000	0.979	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.7	1.000	0.983	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.8	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	0.9	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.8	1.0	1.000	0.964	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.1	1.000	0.995	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.2	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.3	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.4	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.5	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.6	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.7	1.000	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.8	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	0.9	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	0.9	1.0	1.000	0.995	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.1	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.2	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.3	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.4	1.000	0.977	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.5	1.000	0.988	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.6	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.7	1.000	0.991	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.8	1.000	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	0.9	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
	1.0	1.0	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: We report the p-values based on the Chi-Square tests performed on the data sets whose values follow a Benford's distribution partially replaced by data following a normal distribution. The normal distribution is of mean 0 and standard deviation ranging from 0.1 to 1. The values are based on the average of 100 bootstrap iterations.

8.5. **Tables of bootstrapping on data replaced by uniform random variables.**

TABLE 39. MAD from partial replacement with a uniform distribution

Data Set Size = 2500												
Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.4	11.7	6.0	4.3	3.4	2.0	1.6	1.0	0.6		
0.1	0.2	0.5	11.7	6.0	4.5	3.5	2.1	1.7	1.3	1.0		
0.1	0.3	0.3	11.7	6.0	4.0	2.7	1.9	1.2	1.0	1.0		
0.1	0.4	0.4	11.7	6.0	4.2	3.2	2.1	1.3	1.0	0.9		
0.1	0.5	0.5	11.7	6.0	4.3	3.2	2.0	1.5	0.9	0.8		
0.1	0.6	0.3	11.7	6.0	4.0	2.9	1.8	1.4	0.9	0.7		
0.1	0.7	0.6	11.7	6.0	4.1	3.1	2.0	1.4	0.9	0.6		
0.1	0.8	0.4	11.7	6.0	4.4	3.1	2.1	1.5	1.0	0.7		
0.1	0.9	0.4	11.7	6.0	4.5	3.1	2.0	1.5	1.0	0.6		
0.1	1.0	0.5	11.7	6.0	4.3	3.4	2.0	1.6	1.0	0.6		
0.2	0.1	0.5	11.6	8.2	5.5	4.1	2.8	1.7	1.5	1.1		
0.2	0.2	0.6	10.1	7.5	5.7	3.9	2.5	2.4	1.8	1.4		
0.2	0.3	0.8	9.7	7.0	4.8	4.1	2.6	2.3	1.5	1.1		
0.2	0.4	0.7	11.0	8.0	5.2	4.8	2.7	2.0	1.3	1.2		
0.2	0.5	0.4	11.2	7.2	5.0	4.9	3.0	2.2	1.5	1.2		
0.2	0.6	1.2	11.5	7.1	5.0	4.6	2.9	2.1	1.6	1.3		
0.2	0.7	0.7	11.3	8.0	5.0	4.3	2.9	2.2	1.8	1.2		
0.2	0.8	0.6	12.0	7.8	5.3	4.4	3.1	2.0	1.6	1.2		
0.2	0.9	0.5	11.1	7.2	5.2	4.3	2.9	1.7	1.5	1.2		
0.2	1.0	0.8	11.6	8.2	5.5	4.1	2.8	1.7	1.5	1.1		
0.3	0.1	0.9	10.6	7.6	7.4	5.2	3.4	2.4	1.9	1.7		
0.3	0.2	1.0	8.5	5.9	5.9	4.0	2.9	2.6	2.0	1.7		
0.3	0.3	0.9	8.3	5.9	4.7	3.3	2.2	1.7	1.6	1.6		
0.3	0.4	0.9	9.4	6.0	5.5	3.8	2.8	2.3	1.9	1.7		
0.3	0.5	1.1	9.0	7.0	6.5	4.7	3.4	2.5	2.0	1.7		
0.3	0.6	0.9	9.0	7.1	6.2	4.2	3.3	2.9	2.1	1.8		
0.3	0.7	0.9	10.7	7.6	7.2	4.6	3.5	2.7	1.9	1.7		
0.3	0.8	0.6	10.5	7.6	6.0	4.5	3.5	2.5	1.7	1.7		
0.3	0.9	0.9	10.8	7.3	6.3	4.7	3.2	2.3	1.7	1.7		
0.3	1.0	1.1	10.6	7.6	7.4	5.2	3.4	2.4	1.9	1.7		
0.4	0.1	1.4	10.2	8.7	6.3	3.6	3.3	2.8	2.5	2.3		
0.4	0.2	0.7	10.7	7.4	5.0	3.8	4.0	3.1	2.3	2.4		
0.4	0.3	0.8	10.1	7.0	5.0	2.8	3.5	2.7	1.9	2.2		
0.4	0.4	1.1	10.1	6.1	4.3	3.2	3.2	2.5	2.3	2.2		
0.4	0.5	1.3	10.8	7.1	6.2	3.2	3.0	2.9	2.6	2.3		
0.4	0.6	1.3	11.6	8.0	5.9	3.5	2.7	2.6	2.4	2.3		
0.4	0.7	0.6	10.2	7.1	5.9	4.0	3.3	3.1	2.7	2.4		
0.4	0.8	0.8	10.9	7.1	5.8	3.9	3.2	2.8	2.5	2.4		
0.4	0.9	0.8	10.3	7.1	5.7	3.8	3.6	2.9	2.5	2.4		
0.4	1.0	0.6	10.2	8.7	6.3	3.6	3.3	2.8	2.5	2.3		
0.5	0.1	1.4	11.6	8.1	5.1	4.6	4.2	3.4	2.7	3.0		
0.5	0.2	1.2	9.8	7.2	4.6	4.2	4.3	3.8	3.0	3.0		
0.5	0.3	1.4	9.5	5.8	4.8	3.8	3.9	4.0	2.9	2.8		

0.5	0.4	1.4	9.5	6.1	4.0	2.9	3.1	3.3	2.8	2.7
0.5	0.5	1.4	10.9	6.7	4.1	3.8	4.1	3.5	3.0	2.9
0.5	0.6	1.1	9.4	7.2	4.8	4.2	4.1	3.2	3.0	3.0
0.5	0.7	0.9	9.0	7.4	4.6	4.5	4.4	3.5	3.0	3.1
0.5	0.8	1.0	9.8	8.3	5.0	4.8	4.2	3.3	2.8	3.1
0.5	0.9	1.0	10.1	8.3	5.1	4.7	4.1	3.3	2.8	3.1
0.5	1.0	1.3	11.6	8.1	5.1	4.6	4.2	3.4	2.7	3.0
0.6	0.1	1.6	8.4	6.0	5.5	4.9	4.9	4.0	3.6	3.7
0.6	0.2	1.6	9.3	6.3	6.0	6.8	5.2	3.7	3.1	3.6
0.6	0.3	1.5	8.8	7.5	6.1	5.4	4.6	3.9	3.1	3.3
0.6	0.4	1.1	9.4	6.3	5.1	4.3	3.8	3.6	3.0	3.2
0.6	0.5	1.8	9.4	7.2	5.2	4.2	3.8	3.9	3.4	3.4
0.6	0.6	1.5	9.5	6.8	4.5	4.0	4.3	4.2	3.5	3.5
0.6	0.7	1.4	9.2	6.6	4.7	4.1	4.8	4.3	3.6	3.7
0.6	0.8	1.2	8.2	5.9	3.7	4.7	5.0	4.1	3.6	3.8
0.6	0.9	1.4	8.4	6.1	4.1	4.4	5.0	4.0	3.6	3.7
0.6	1.0	1.6	8.4	6.0	5.5	4.9	4.9	4.0	3.6	3.7
0.7	0.1	1.7	11.8	8.9	6.0	5.8	5.5	4.7	3.9	4.3
0.7	0.2	1.7	9.7	7.8	6.3	5.4	5.1	4.8	4.1	4.2
0.7	0.3	1.7	9.6	7.0	5.2	4.5	3.9	4.2	4.0	3.9
0.7	0.4	1.6	12.2	10.0	7.0	5.1	4.7	4.0	3.7	3.8
0.7	0.5	1.5	12.4	9.2	6.8	5.6	5.0	4.1	3.8	4.1
0.7	0.6	1.5	12.8	9.5	6.3	5.7	5.5	4.2	4.0	4.2
0.7	0.7	1.5	12.7	9.8	7.2	5.8	5.9	4.5	4.0	4.4
0.7	0.8	1.3	12.1	9.7	7.0	5.4	5.3	4.4	3.9	4.3
0.7	0.9	1.3	12.0	8.6	6.3	5.6	5.5	4.7	3.9	4.3
0.7	1.0	1.7	11.8	8.9	6.0	5.8	5.5	4.7	3.9	4.3
0.8	0.1	2.2	11.9	7.0	5.3	5.0	4.8	5.2	5.1	4.8
0.8	0.2	1.8	11.6	9.2	6.3	5.3	5.3	5.1	5.0	4.8
0.8	0.3	1.6	10.8	9.2	5.6	5.0	4.7	4.6	4.7	4.6
0.8	0.4	2.2	9.7	7.8	5.4	4.9	4.5	4.6	4.4	4.4
0.8	0.5	1.9	9.6	6.6	5.3	4.3	3.9	4.1	4.5	4.5
0.8	0.6	2.0	8.5	6.8	5.6	4.2	4.2	4.6	4.8	4.7
0.8	0.7	1.3	9.7	6.5	5.2	4.0	4.4	4.6	4.9	4.8
0.8	0.8	1.5	10.3	6.3	4.9	4.4	4.4	4.8	4.9	4.8
0.8	0.9	1.9	10.7	7.1	5.6	4.5	4.4	4.8	4.9	4.8
0.8	1.0	1.9	11.9	7.0	5.3	5.0	4.8	5.2	5.1	4.8
0.9	0.1	2.3	11.3	9.4	6.2	5.4	4.4	4.5	4.7	5.1
0.9	0.2	2.0	10.0	7.5	6.1	5.6	4.6	4.6	4.9	5.3
0.9	0.3	2.1	10.1	7.7	5.7	5.4	4.3	4.4	4.7	5.2
0.9	0.4	2.1	11.7	8.0	5.5	5.3	4.1	4.2	4.2	4.8
0.9	0.5	1.9	11.1	8.5	5.3	4.5	3.7	4.1	4.3	5.0
0.9	0.6	2.2	11.0	7.6	5.3	5.0	4.2	4.4	4.4	5.1
0.9	0.7	1.6	11.9	9.5	6.6	5.2	4.4	4.7	4.5	5.1
0.9	0.8	1.5	10.3	7.5	6.1	5.2	4.4	4.3	4.4	5.2
0.9	0.9	2.1	11.9	9.0	6.1	5.2	4.3	4.3	4.6	5.2
0.9	1.0	2.2	11.3	9.4	6.2	5.4	4.4	4.5	4.7	5.1
1.0	0.1	2.9	12.1	9.4	7.5	6.5	5.9	5.7	5.8	5.9
1.0	0.2	2.2	10.1	9.6	7.5	6.6	6.2	6.0	5.9	6.0
1.0	0.3	2.3	10.0	9.5	7.6	6.8	5.9	5.9	5.8	6.0

1.0	0.4	2.2	9.2	7.7	6.9	6.0	5.8	5.6	5.4	5.6		
1.0	0.5	2.6	10.1	6.8	6.1	5.4	5.2	5.3	5.6	5.9		
1.0	0.6	2.3	8.7	8.4	6.2	6.1	5.8	5.7	5.8	6.0		
1.0	0.7	2.2	9.9	8.8	6.5	6.3	5.6	5.6	5.7	5.9		
1.0	0.8	2.2	10.4	8.7	6.7	6.1	5.7	5.5	5.7	6.0		
1.0	0.9	2.3	11.6	10.2	8.0	6.8	5.8	5.8	5.8	6.0		
1.0	1.0	2.4	12.1	9.4	7.5	6.5	5.9	5.7	5.8	5.9		

Data Set Size = 5000

Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.2	10.8	6.2	4.8	3.3	2.0	1.4	1.0	0.8	0.8	
0.1	0.2	0.4	10.7	6.5	4.3	3.0	2.0	1.5	1.0	0.8	0.6	
0.1	0.3	0.4	10.3	6.4	4.3	3.2	2.3	1.4	0.9	0.9	0.6	
0.1	0.4	0.3	11.5	6.6	4.5	3.0	2.3	1.4	0.9	0.8	0.6	
0.1	0.5	0.4	10.1	6.1	4.4	3.1	2.1	1.3	0.8	0.8	0.7	
0.1	0.6	0.6	10.4	6.3	4.7	3.2	2.3	1.4	0.9	0.8	0.8	
0.1	0.7	0.5	10.6	5.9	4.5	3.1	2.0	1.3	0.9	0.8	0.8	
0.1	0.8	0.3	10.7	6.4	4.6	3.1	2.1	1.3	1.0	0.8	0.8	
0.1	0.9	0.3	10.7	6.5	4.7	3.1	2.0	1.2	1.0	0.7	0.8	
0.1	1.0	0.3	10.8	6.2	4.8	3.3	2.0	1.4	1.0	0.8	0.8	
0.2	0.1	0.7	13.1	8.8	5.7	3.7	2.4	2.0	1.4	1.3	1.3	
0.2	0.2	0.6	13.1	7.8	4.7	3.0	2.9	1.9	1.7	1.3	1.1	
0.2	0.3	0.7	13.1	8.8	5.5	3.9	2.7	1.9	1.1	1.2	1.2	
0.2	0.4	0.7	13.1	8.1	5.4	2.9	1.8	1.5	1.1	1.1	1.3	
0.2	0.5	0.4	13.1	8.6	5.0	3.2	2.5	1.7	1.3	1.2	1.4	
0.2	0.6	0.5	13.1	8.5	5.3	3.0	2.2	1.5	1.2	1.4	1.5	
0.2	0.7	0.7	13.1	8.5	5.5	3.1	2.7	1.6	1.1	1.4	1.4	
0.2	0.8	0.2	13.1	8.9	5.5	3.6	2.6	1.9	1.2	1.4	1.4	
0.2	0.9	0.5	13.1	8.9	5.6	3.6	2.3	1.9	1.2	1.3	1.3	
0.2	1.0	0.8	13.1	8.8	5.7	3.7	2.4	2.0	1.4	1.3	1.3	
0.3	0.1	0.6	13.5	9.7	6.3	4.9	2.8	2.1	2.1	1.6	1.5	
0.3	0.2	0.6	10.3	7.8	5.3	3.9	2.8	2.1	1.9	1.9	1.6	
0.3	0.3	0.7	8.9	6.9	5.1	3.9	2.2	1.7	1.6	1.9	1.8	
0.3	0.4	1.0	9.1	7.9	5.2	4.8	2.6	1.9	1.7	1.9	1.8	
0.3	0.5	0.7	9.8	7.0	4.9	4.0	2.5	1.5	1.7	1.7	1.9	
0.3	0.6	0.9	10.7	8.2	4.9	4.0	2.3	1.9	1.9	1.8	1.9	
0.3	0.7	0.8	11.0	8.5	4.8	3.7	2.5	1.9	2.1	1.7	1.8	
0.3	0.8	0.6	11.4	8.6	4.9	3.4	2.4	2.2	2.2	1.7	1.7	
0.3	0.9	0.8	12.4	7.9	5.0	3.5	2.4	2.1	2.1	1.6	1.6	
0.3	1.0	0.7	13.5	9.7	6.3	4.9	2.8	2.1	2.1	1.6	1.5	
0.4	0.1	1.0	13.2	10.5	7.3	4.3	3.3	3.5	3.5	2.8	2.3	
0.4	0.2	1.0	10.5	6.5	4.4	3.5	3.1	3.0	3.0	2.4	2.2	
0.4	0.3	0.9	9.1	6.3	5.5	4.6	3.5	2.9	2.3	2.1	2.2	
0.4	0.4	1.0	9.0	6.4	4.9	3.9	2.7	2.7	2.2	2.1	2.2	
0.4	0.5	1.2	10.0	8.6	6.5	4.9	3.2	3.0	2.5	2.4	2.4	
0.4	0.6	0.8	11.0	8.9	6.4	4.9	3.6	3.3	2.9	2.5	2.4	
0.4	0.7	1.0	12.5	10.1	6.8	4.9	3.5	3.3	3.0	2.7	2.4	
0.4	0.8	1.0	12.9	10.6	6.6	4.4	3.3	3.4	3.1	2.8	2.4	
0.4	0.9	0.8	12.5	10.8	6.9	4.7	3.6	3.5	3.3	2.7	2.3	
0.4	1.0	0.6	13.2	10.5	7.3	4.3	3.3	3.5	3.5	2.8	2.3	
0.5	0.1	1.2	12.2	7.9	5.4	3.5	3.1	2.8	2.8	2.6	2.8	

0.5	0.2	1.1	11.4	9.5	7.6	4.8	4.3	3.8	3.2	2.7	2.7
0.5	0.3	1.3	8.7	6.9	5.4	4.5	3.5	3.0	2.9	2.6	2.8
0.5	0.4	1.5	9.6	7.1	5.3	4.2	3.3	2.9	2.6	2.3	2.6
0.5	0.5	1.2	11.9	8.4	6.4	3.7	3.1	2.4	2.1	2.4	2.8
0.5	0.6	1.5	10.9	7.6	6.1	4.0	3.3	2.6	2.4	2.4	2.9
0.5	0.7	1.0	11.4	8.2	6.3	4.6	4.0	2.9	2.7	2.5	2.8
0.5	0.8	0.8	11.6	8.6	6.4	4.9	4.1	3.0	2.8	2.6	2.9
0.5	0.9	1.3	12.6	8.3	5.6	3.9	3.6	2.8	2.7	2.6	2.8
0.5	1.0	1.0	12.2	7.9	5.4	3.5	3.1	2.8	2.8	2.6	2.8
0.6	0.1	1.5	12.1	8.5	6.6	5.1	5.1	4.1	4.1	3.7	3.5
0.6	0.2	1.4	9.4	8.1	5.9	4.4	4.5	3.7	3.8	3.5	3.3
0.6	0.3	1.8	9.5	7.6	5.3	4.2	3.7	3.4	3.7	3.4	3.4
0.6	0.4	1.3	8.6	6.5	5.9	4.3	3.7	3.1	3.5	3.4	3.2
0.6	0.5	1.2	12.9	8.0	6.0	4.6	4.0	3.6	4.1	3.6	3.5
0.6	0.6	1.6	10.5	7.5	5.9	4.6	4.3	3.7	4.2	3.6	3.5
0.6	0.7	1.1	10.6	8.9	6.8	5.0	4.7	3.7	4.3	3.7	3.4
0.6	0.8	1.2	10.7	8.1	7.0	5.0	4.9	3.8	4.4	3.9	3.5
0.6	0.9	1.1	10.4	8.2	6.8	5.0	4.8	3.9	4.2	3.7	3.5
0.6	1.0	1.4	12.1	8.5	6.6	5.1	5.1	4.1	4.1	3.7	3.5
0.7	0.1	1.7	11.3	8.3	7.0	5.1	4.3	4.6	3.9	4.2	4.1
0.7	0.2	1.6	10.6	8.4	6.5	4.3	4.1	3.9	3.9	4.2	3.8
0.7	0.3	1.7	9.5	7.2	5.2	3.8	4.0	4.3	4.1	4.2	3.9
0.7	0.4	1.6	9.5	8.1	6.7	4.4	5.2	4.8	4.2	4.2	3.9
0.7	0.5	1.5	9.7	7.3	6.1	4.9	5.2	5.2	4.4	4.3	4.1
0.7	0.6	1.5	10.4	8.9	8.0	5.5	5.2	5.2	4.4	4.3	4.1
0.7	0.7	1.5	10.0	8.1	7.7	5.0	4.9	5.0	4.1	4.2	4.0
0.7	0.8	1.1	11.2	9.1	7.9	5.1	4.7	5.0	4.1	4.2	4.1
0.7	0.9	1.5	11.6	7.8	7.1	5.2	4.4	4.9	4.0	4.2	4.1
0.7	1.0	1.6	11.3	8.3	7.0	5.1	4.3	4.6	3.9	4.2	4.1
0.8	0.1	2.1	9.9	9.3	5.2	4.3	4.5	4.1	4.6	4.6	4.6
0.8	0.2	2.0	9.8	8.3	6.1	5.3	4.9	4.5	4.5	4.5	4.5
0.8	0.3	1.5	10.3	8.1	5.7	4.6	4.2	4.1	4.4	4.6	4.6
0.8	0.4	1.8	9.4	7.5	5.8	4.6	3.6	3.9	4.4	4.5	4.4
0.8	0.5	1.6	11.6	8.9	6.5	4.8	4.2	4.1	4.5	4.7	4.7
0.8	0.6	1.9	10.6	7.9	6.1	4.4	4.5	4.3	4.8	4.7	4.6
0.8	0.7	1.6	10.0	7.5	5.3	4.3	4.2	4.1	4.6	4.6	4.6
0.8	0.8	1.4	10.4	8.9	5.7	4.4	4.7	4.3	4.7	4.7	4.7
0.8	0.9	1.5	9.7	7.9	4.9	4.1	4.6	4.1	4.7	4.6	4.6
0.8	1.0	1.8	9.9	9.3	5.2	4.3	4.5	4.1	4.6	4.6	4.6
0.9	0.1	1.9	13.4	10.2	7.5	6.4	5.7	5.4	5.3	5.4	5.3
0.9	0.2	2.3	12.4	9.5	7.7	6.5	5.7	5.4	5.3	5.4	5.3
0.9	0.3	1.9	11.0	8.6	6.7	5.6	5.7	5.2	5.5	5.4	5.3
0.9	0.4	2.3	11.1	8.8	7.5	5.7	5.2	4.8	5.1	5.2	5.0
0.9	0.5	2.2	10.0	8.6	6.0	5.6	5.5	5.2	5.7	5.6	5.4
0.9	0.6	2.0	10.5	9.4	6.3	5.3	5.9	5.2	5.5	5.4	5.3
0.9	0.7	1.6	11.7	9.0	6.2	5.7	5.5	5.1	5.2	5.4	5.2
0.9	0.8	1.4	12.0	10.6	6.7	5.5	5.2	5.1	5.3	5.5	5.4
0.9	0.9	1.7	13.8	11.3	7.0	6.0	5.6	5.3	5.3	5.4	5.3
0.9	1.0	2.2	13.4	10.2	7.5	6.4	5.7	5.4	5.3	5.4	5.3
1.0	0.1	2.3	12.3	8.1	7.3	6.0	6.1	6.1	5.9	5.8	5.9

1.0	0.2	2.1	12.2	8.7	8.4	5.9	5.8	6.3	5.9	5.8	5.9
1.0	0.3	2.2	10.9	7.9	7.5	5.8	5.6	5.8	5.4	5.6	5.7
1.0	0.4	2.2	11.2	8.0	6.9	5.2	5.2	5.9	5.5	5.4	5.5
1.0	0.5	2.2	11.4	7.7	6.9	5.1	5.7	6.1	5.8	5.8	5.9
1.0	0.6	1.8	12.5	7.6	6.6	5.6	5.4	6.0	5.6	5.6	5.7
1.0	0.7	2.0	12.5	8.3	7.2	5.6	5.6	6.0	5.8	5.6	5.8
1.0	0.8	1.7	11.7	8.4	7.3	5.9	5.8	6.0	5.7	5.7	5.9
1.0	0.9	2.2	12.1	7.4	6.8	5.8	5.6	6.1	5.8	5.7	5.8
1.0	1.0	2.3	12.3	8.1	7.3	6.0	6.1	6.1	5.9	5.8	5.9

Data Set Size = 10000

Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	0.3	11.6	9.0	6.3	4.5	2.8	1.5	1.3	0.9	0.8	0.6
0.1	0.2	0.4	10.8	8.3	4.7	3.3	2.1	2.1	1.8	1.2	0.8	0.8
0.1	0.3	0.3	10.8	8.3	4.7	3.3	2.2	1.7	1.2	0.9	0.7	0.7
0.1	0.4	0.5	12.0	8.6	5.3	4.0	2.7	1.9	1.3	0.9	0.8	0.7
0.1	0.5	0.4	12.5	9.1	5.7	4.4	3.0	1.9	1.4	0.9	0.8	0.6
0.1	0.6	0.3	12.5	9.1	6.3	4.7	3.0	1.9	1.6	0.9	0.9	0.6
0.1	0.7	0.3	11.2	9.0	6.2	4.7	3.1	1.7	1.5	0.9	0.8	0.6
0.1	0.8	0.2	12.6	9.2	6.3	4.8	3.1	1.8	1.5	0.9	1.0	0.7
0.1	0.9	0.3	12.6	9.2	6.4	4.7	2.9	1.6	1.3	0.9	0.8	0.6
0.1	1.0	0.4	11.6	9.0	6.3	4.5	2.8	1.5	1.3	0.9	0.8	0.6
0.2	0.1	0.3	8.1	6.6	4.4	3.3	2.7	1.8	1.5	1.4	1.2	1.2
0.2	0.2	0.5	9.0	6.2	4.8	3.1	2.8	2.3	2.1	1.6	1.4	1.3
0.2	0.3	0.6	9.3	6.0	4.8	3.2	2.8	1.8	1.6	1.3	1.4	1.2
0.2	0.4	0.6	9.4	5.9	4.4	3.4	2.8	1.4	1.3	1.2	1.2	1.1
0.2	0.5	0.5	9.1	6.3	4.5	3.4	2.9	1.7	1.4	1.2	1.2	1.1
0.2	0.6	0.6	8.1	6.4	4.4	2.9	2.6	1.6	1.2	1.0	1.1	1.1
0.2	0.7	0.6	8.1	6.4	5.0	3.2	2.7	1.5	1.0	1.0	1.0	1.1
0.2	0.8	0.6	8.1	6.8	4.5	2.9	2.2	1.7	1.2	1.1	1.2	1.2
0.2	0.9	0.5	8.1	6.5	4.6	3.3	2.8	1.9	1.4	1.3	1.1	1.2
0.2	1.0	0.6	8.1	6.6	4.4	3.3	2.7	1.8	1.5	1.4	1.2	1.2
0.3	0.1	0.4	9.8	7.2	4.4	2.8	2.1	1.8	2.0	1.6	1.7	1.7
0.3	0.2	0.7	9.3	7.3	5.7	3.5	2.3	2.2	2.2	2.0	1.8	1.8
0.3	0.3	0.8	8.8	7.1	5.7	3.7	2.4	1.9	2.0	2.1	1.9	1.8
0.3	0.4	0.6	8.5	6.5	4.8	3.5	2.3	1.8	1.8	2.0	1.8	1.7
0.3	0.5	0.6	8.4	6.4	4.9	3.4	2.3	1.9	1.8	2.0	1.9	1.8
0.3	0.6	0.8	9.6	6.9	4.8	3.3	2.2	1.9	1.7	1.8	1.7	1.7
0.3	0.7	0.6	9.5	6.4	4.4	2.9	2.3	2.0	1.8	1.8	1.7	1.7
0.3	0.8	0.6	9.3	6.5	4.6	2.9	2.0	1.7	1.9	1.8	1.7	1.8
0.3	0.9	0.6	10.2	6.4	4.4	3.5	2.1	1.9	2.1	1.7	1.7	1.7
0.3	1.0	0.7	9.8	7.2	4.4	2.8	2.1	1.8	2.0	1.6	1.7	1.7
0.4	0.1	1.0	9.5	8.7	7.8	5.5	3.9	2.7	2.3	2.1	2.2	2.2
0.4	0.2	0.9	9.0	7.2	5.4	3.9	3.9	3.1	2.8	2.4	2.4	2.3
0.4	0.3	0.8	9.9	7.4	7.4	5.9	3.8	2.8	2.7	2.5	2.3	2.3
0.4	0.4	0.9	10.1	7.9	7.5	5.7	4.1	2.8	2.5	2.4	2.3	2.3
0.4	0.5	0.8	10.9	8.0	7.6	6.1	4.4	2.9	2.7	2.5	2.4	2.3
0.4	0.6	0.8	8.7	7.3	7.7	6.3	4.5	2.8	2.8	2.6	2.4	2.3
0.4	0.7	0.8	9.0	8.5	7.9	6.3	4.1	2.6	2.4	2.4	2.2	2.3
0.4	0.8	1.0	9.8	8.9	8.5	6.6	4.4	2.8	2.6	2.4	2.3	2.4
0.4	0.9	0.8	9.9	9.1	8.4	6.3	4.1	2.8	2.7	2.4	2.3	2.3

	0.4	1.0	0.8	9.5	8.7	7.8	5.5	3.9	2.7	2.3	2.1	2.2	2.2
	0.5	0.1	1.2	8.4	6.4	4.3	3.8	3.9	3.7	3.0	3.2	2.9	2.9
	0.5	0.2	1.3	9.2	6.8	5.3	4.2	4.0	3.7	3.5	3.2	3.0	2.9
	0.5	0.3	1.1	9.4	6.0	4.5	3.9	3.2	2.8	3.0	3.0	3.0	2.9
	0.5	0.4	1.1	9.8	6.1	4.5	3.5	3.2	2.6	3.0	3.1	2.9	2.9
	0.5	0.5	1.1	9.1	6.2	4.3	3.4	2.6	2.1	2.7	3.3	3.1	3.0
	0.5	0.6	1.0	10.2	6.7	4.2	3.6	3.1	2.9	3.0	3.2	3.0	2.9
	0.5	0.7	1.0	8.9	6.7	5.0	3.7	2.7	2.5	2.8	3.3	3.0	2.9
	0.5	0.8	1.0	9.4	6.8	4.3	4.0	3.4	3.2	3.1	3.3	3.0	3.0
	0.5	0.9	1.2	9.5	6.6	4.2	3.4	3.4	3.3	3.0	3.3	3.0	2.9
	0.5	1.0	1.1	8.4	6.4	4.3	3.8	3.9	3.7	3.0	3.2	2.9	2.9
	0.6	0.1	1.4	11.5	6.6	4.9	3.5	4.2	3.8	3.2	3.2	3.3	3.4
	0.6	0.2	1.4	10.0	8.1	6.9	5.4	5.4	4.6	4.0	3.6	3.5	3.5
	0.6	0.3	1.3	10.4	8.1	5.9	5.3	5.1	4.8	4.2	3.7	3.4	3.5
	0.6	0.4	1.3	9.9	6.7	5.1	4.2	4.1	4.1	3.5	3.3	3.2	3.5
	0.6	0.5	1.3	8.0	5.4	4.5	3.6	3.4	3.3	3.2	3.3	3.4	3.6
	0.6	0.6	1.5	9.5	6.0	4.4	3.3	3.9	3.8	3.4	3.3	3.4	3.5
	0.6	0.7	1.2	9.0	6.7	4.9	4.1	3.6	3.7	3.2	3.1	3.4	3.4
	0.6	0.8	1.3	10.3	6.7	4.5	4.3	3.7	3.9	3.3	3.2	3.4	3.5
	0.6	0.9	1.3	9.8	6.5	4.8	3.4	3.9	3.7	3.2	3.2	3.3	3.4
	0.6	1.0	1.3	11.5	6.6	4.9	3.5	4.2	3.8	3.2	3.2	3.3	3.4
	0.7	0.1	1.5	9.1	6.8	6.6	5.8	6.2	5.3	4.2	4.0	4.2	4.1
	0.7	0.2	1.5	9.8	7.6	6.5	6.2	6.4	5.9	4.8	4.5	4.2	4.0
	0.7	0.3	1.6	10.3	7.3	6.9	4.7	5.4	5.1	4.6	4.3	4.0	4.0
	0.7	0.4	1.5	9.1	6.6	6.1	5.4	6.1	5.0	4.2	4.0	4.0	4.0
	0.7	0.5	1.8	9.3	6.5	6.8	5.0	5.6	4.6	4.1	4.0	4.2	4.1
	0.7	0.6	1.6	10.2	6.7	7.0	6.1	6.3	5.1	4.1	4.0	4.1	4.0
	0.7	0.7	1.4	10.5	6.1	6.3	5.5	5.9	4.7	3.8	3.9	4.1	4.1
	0.7	0.8	1.3	9.8	6.7	6.7	5.8	6.4	5.3	4.3	4.2	4.2	4.2
	0.7	0.9	1.4	10.0	7.9	7.2	5.7	6.2	5.2	4.2	4.1	4.2	4.1
	0.7	1.0	1.5	9.1	6.8	6.6	5.8	6.2	5.3	4.2	4.0	4.2	4.1
	0.8	0.1	1.8	14.5	10.9	7.9	5.5	5.6	5.1	4.7	4.5	4.6	4.6
	0.8	0.2	1.7	13.2	9.0	7.6	6.2	5.9	5.6	5.2	5.0	4.9	4.6
	0.8	0.3	1.6	11.3	8.2	7.2	6.7	5.6	5.4	4.9	5.1	4.9	4.6
	0.8	0.4	1.9	10.3	8.8	6.2	5.7	5.6	5.1	4.7	4.5	4.4	4.5
	0.8	0.5	1.8	12.6	11.2	7.4	6.6	5.8	5.4	4.8	4.7	4.6	4.7
	0.8	0.6	1.6	11.3	9.7	6.5	5.4	5.3	5.0	4.6	4.7	4.6	4.6
	0.8	0.7	1.7	14.0	10.1	7.2	5.5	5.5	5.2	4.9	4.5	4.6	4.7
	0.8	0.8	1.5	12.8	10.3	6.8	4.8	5.2	4.9	4.7	4.6	4.7	4.7
	0.8	0.9	1.9	14.6	11.1	8.0	5.8	5.7	5.1	4.8	4.6	4.6	4.6
	0.8	1.0	1.7	14.5	10.9	7.9	5.5	5.6	5.1	4.7	4.5	4.6	4.6
	0.9	0.1	2.0	9.7	7.3	6.3	5.2	5.0	4.9	4.8	5.0	5.2	5.2
	0.9	0.2	2.1	10.0	7.5	6.4	6.1	5.6	5.3	5.0	5.0	5.2	5.2
	0.9	0.3	2.0	9.7	7.8	5.5	5.9	5.4	5.2	5.2	5.2	5.2	5.2
	0.9	0.4	2.1	13.2	9.1	6.1	5.6	5.5	5.4	4.5	5.0	5.0	5.1
	0.9	0.5	2.0	10.8	8.3	5.7	5.2	5.2	5.0	5.1	5.2	5.3	5.3
	0.9	0.6	2.2	10.9	7.6	5.9	5.4	5.4	5.0	5.0	5.3	5.3	5.2
	0.9	0.7	1.7	10.6	7.5	5.7	5.4	5.1	4.9	4.8	5.1	5.2	5.2
	0.9	0.8	1.7	12.6	7.6	6.2	5.3	5.0	4.9	4.9	5.1	5.3	5.3
	0.9	0.9	2.0	10.9	7.4	5.7	5.1	5.1	4.8	4.9	5.2	5.3	5.2

0.9	1.0	2.0	9.7	7.3	6.3	5.2	5.0	4.9	4.8	5.0	5.2	5.2
1.0	0.1	2.3	9.7	6.9	7.0	6.4	6.5	6.4	6.0	5.8	6.0	5.9
1.0	0.2	2.2	8.6	7.3	7.1	6.8	6.7	6.4	6.0	5.8	6.0	5.9
1.0	0.3	2.3	11.0	6.1	6.8	5.7	6.4	6.2	5.9	5.7	6.0	5.8
1.0	0.4	2.3	10.4	7.1	6.3	6.2	6.1	5.5	5.5	5.5	5.5	5.6
1.0	0.5	2.3	9.5	7.1	7.3	6.2	6.5	6.3	6.0	5.8	6.0	5.9
1.0	0.6	2.3	10.0	7.2	7.7	6.3	6.5	6.2	6.0	5.7	6.0	5.8
1.0	0.7	2.1	9.7	6.7	7.7	6.3	6.3	6.1	5.9	5.7	5.8	5.9
1.0	0.8	1.9	10.3	7.7	7.2	6.2	6.8	6.4	5.9	5.8	6.0	5.9
1.0	0.9	2.3	10.0	7.5	7.1	6.1	6.5	6.3	6.0	5.7	6.0	5.9
1.0	1.0	2.2	9.7	6.9	7.0	6.4	6.5	6.4	6.0	5.8	6.0	5.9

Note: We report the MAD values (in %) of data sets whose values follow a Benford's distribution partially replaced by data following a uniform distribution. The uniform distribution has a lower bound 0 and an upper bound ranging from 0.1 to 1. The values are based on the average of 100 bootstrap iterations.

TABLE 40. P-values from partial replacement with a uniform distribution

			Data Set Size = 2500									
Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.2	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.3	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.4	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.5	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.6	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.7	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.8	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	0.9	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.1	1.0	1.000	0.920	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.1	1.000	0.960	0.995	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.2	1.000	0.985	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.3	1.000	0.981	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.4	1.000	0.973	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.5	1.000	0.976	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.6	1.000	0.973	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.7	1.000	0.953	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.8	1.000	0.898	0.995	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	0.9	1.000	0.966	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.2	1.0	1.000	0.960	0.995	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.1	1.000	0.959	0.997	0.999	1.000	1.000	1.000	1.000	1.000		
0.3	0.2	1.000	0.986	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.3	1.000	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.4	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.5	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.6	1.000	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.7	1.000	0.982	0.999	0.999	1.000	1.000	1.000	1.000	1.000		
0.3	0.8	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.3	0.9	1.000	0.969	0.999	1.000	1.000	1.000	1.000	1.000	1.000		

0.3	1.0	1.000	0.959	0.997	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.1	1.000	0.939	0.991	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.2	1.000	0.962	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.3	1.000	0.965	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.4	1.000	0.961	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.5	1.000	0.956	0.998	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.6	1.000	0.910	0.996	0.999	1.000	1.000	1.000	1.000	1.000
0.4	0.7	1.000	0.962	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.8	1.000	0.926	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.9	1.000	0.929	0.996	0.999	1.000	1.000	1.000	1.000	1.000
0.4	1.0	1.000	0.939	0.991	0.999	1.000	1.000	1.000	1.000	1.000
0.5	0.1	1.000	0.955	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.2	1.000	0.968	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.3	1.000	0.990	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.4	1.000	0.985	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.5	1.000	0.980	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.6	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.7	1.000	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	0.984	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.9	1.000	0.981	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.0	1.000	0.955	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.1	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.2	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	1.000	0.988	0.986	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.4	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.5	1.000	0.990	0.997	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.6	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.7	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.8	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.9	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.0	1.000	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.1	1.000	0.891	0.987	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.2	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.3	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.4	1.000	0.959	0.989	0.999	1.000	1.000	1.000	1.000	1.000
0.7	0.5	1.000	0.978	0.996	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.6	1.000	0.941	0.990	0.999	1.000	1.000	1.000	1.000	1.000
0.7	0.7	1.000	0.935	0.987	0.999	1.000	1.000	1.000	1.000	1.000
0.7	0.8	1.000	0.931	0.989	0.999	1.000	1.000	1.000	1.000	1.000
0.7	0.9	1.000	0.912	0.991	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.0	1.000	0.891	0.987	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.1	1.000	0.955	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.2	1.000	0.988	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.3	1.000	0.985	0.998	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.4	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.5	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.6	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.7	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.8	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.9	1.000	0.977	0.997	1.000	1.000	1.000	1.000	1.000	1.000

0.8	1.0	1.000	0.955	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	0.1	1.000	0.951	0.989	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	0.2	1.000	0.978	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	0.3	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	0.4	1.000	0.985	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	0.5	1.000	0.976	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
0.9	0.6	1.000	0.973	0.997	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	0.7	1.000	0.971	0.995	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	0.8	1.000	0.976	0.996	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	0.9	1.000	0.959	0.992	0.999	1.000	1.000	1.000	1.000	1.000		
0.9	1.0	1.000	0.951	0.989	0.999	1.000	1.000	1.000	1.000	1.000		
1.0	0.1	1.000	0.915	0.992	0.999	1.000	1.000	1.000	1.000	1.000		
1.0	0.2	1.000	0.978	0.990	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.3	1.000	0.986	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.4	1.000	0.990	0.998	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.5	1.000	0.986	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.6	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.7	1.000	0.973	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.8	1.000	0.957	0.996	1.000	1.000	1.000	1.000	1.000	1.000		
1.0	0.9	1.000	0.938	0.992	0.999	1.000	1.000	1.000	1.000	1.000		
1.0	1.0	1.000	0.915	0.992	0.999	1.000	1.000	1.000	1.000	1.000		
						Data Set Size = 5000						
Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.951	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.2	1.000	0.958	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.3	1.000	0.959	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.4	1.000	0.939	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.5	1.000	0.958	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.6	1.000	0.957	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.7	1.000	0.955	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.8	1.000	0.953	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	0.9	1.000	0.953	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.1	1.0	1.000	0.951	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.1	1.000	0.919	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.2	1.000	0.919	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.3	1.000	0.919	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.4	1.000	0.919	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.5	1.000	0.919	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.6	1.000	0.919	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.7	1.000	0.919	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.8	1.000	0.919	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	0.9	1.000	0.919	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.2	1.0	1.000	0.919	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.1	1.000	0.947	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.2	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.3	1.000	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.4	1.000	0.965	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.5	1.000	0.977	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.6	1.000	0.992	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.3	0.7	1.000	0.985	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	

0.3	0.8	1.000	0.986	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.9	1.000	0.969	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.0	1.000	0.947	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.1	1.000	0.949	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.2	1.000	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.3	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.4	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.5	1.000	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.6	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.7	1.000	0.983	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.8	1.000	0.972	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.9	1.000	0.967	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.0	1.000	0.949	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.1	1.000	0.909	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.2	1.000	0.989	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.3	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.4	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.5	1.000	0.967	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.6	1.000	0.985	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.7	1.000	0.973	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	0.941	0.991	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.9	1.000	0.931	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.0	1.000	0.909	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.1	1.000	0.961	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.2	1.000	0.986	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.4	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.5	1.000	0.962	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.6	1.000	0.988	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.7	1.000	0.980	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.8	1.000	0.978	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.9	1.000	0.982	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.0	1.000	0.961	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.1	1.000	0.963	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.2	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.3	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.4	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.5	1.000	0.986	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.6	1.000	0.982	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.7	1.000	0.986	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.8	1.000	0.973	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.9	1.000	0.951	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.0	1.000	0.963	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.1	1.000	0.967	0.983	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.2	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.3	1.000	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.4	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.5	1.000	0.962	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.6	1.000	0.975	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.7	1.000	0.962	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000

0.8	0.8	1.000	0.975	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	0.9	1.000	0.976	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.8	1.0	1.000	0.967	0.983	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.1	1.000	0.940	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.2	1.000	0.989	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.3	1.000	0.980	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.4	1.000	0.987	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.5	1.000	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.6	1.000	0.983	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.7	1.000	0.973	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.8	1.000	0.953	0.987	1.000	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	0.9	1.000	0.945	0.973	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
0.9	1.0	1.000	0.940	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.1	1.000	0.949	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.2	1.000	0.990	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.3	1.000	0.990	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.4	1.000	0.967	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.5	1.000	0.979	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.6	1.000	0.964	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.7	1.000	0.943	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.8	1.000	0.947	0.991	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	0.9	1.000	0.965	0.997	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
1.0	1.0	1.000	0.949	0.994	0.998	1.000	1.000	1.000	1.000	1.000	1.000	
						Data Set Size = 10000						
Proportion replaced	Upper bound	Original	10	22	46	100	215	464	1000	2154	4642	10000
0.1	0.1	1.000	0.972	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.2	1.000	0.975	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.3	1.000	0.975	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.4	1.000	0.972	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.5	1.000	0.960	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.6	1.000	0.960	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.7	1.000	0.973	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.8	1.000	0.939	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	0.9	1.000	0.939	0.993	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.1	1.0	1.000	0.972	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.1	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.2	1.000	0.992	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.3	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.4	1.000	0.981	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.5	1.000	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.6	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.7	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.8	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	0.9	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.2	1.0	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.1	1.000	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.2	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.3	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.4	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.5	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000

0.3	0.6	1.000	0.989	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.7	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.8	1.000	0.987	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	0.9	1.000	0.975	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.3	1.0	1.000	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.1	1.000	0.977	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.2	1.000	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.3	1.000	0.992	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.4	1.000	0.988	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.5	1.000	0.974	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.6	1.000	0.994	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.7	1.000	0.994	0.995	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.8	1.000	0.984	0.990	0.993	0.999	1.000	1.000	1.000	1.000	1.000	1.000
0.4	0.9	1.000	0.968	0.988	0.989	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.4	1.0	1.000	0.977	0.993	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.1	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.2	1.000	0.995	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.3	1.000	0.986	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.4	1.000	0.994	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.5	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.6	1.000	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.7	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.8	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	0.9	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.5	1.0	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.1	1.000	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.2	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.3	1.000	0.992	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.4	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.5	1.000	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.6	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.7	1.000	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.8	1.000	0.994	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	0.9	1.000	0.995	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.6	1.0	1.000	0.983	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.1	1.000	0.980	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.2	1.000	0.990	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.3	1.000	0.969	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.4	1.000	0.991	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.5	1.000	0.989	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.6	1.000	0.968	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.7	1.000	0.976	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.8	1.000	0.987	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	0.9	1.000	0.983	0.995	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.7	1.0	1.000	0.980	0.998	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.1	1.000	0.917	0.982	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.2	1.000	0.984	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.3	1.000	0.993	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.4	1.000	0.992	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.5	1.000	0.973	0.978	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000

0.8	0.6	1.000	0.983	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.7	1.000	0.922	0.992	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.8	1.000	0.962	0.986	0.997	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	0.9	1.000	0.898	0.980	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.8	1.0	1.000	0.917	0.982	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.1	1.000	0.985	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.2	1.000	0.996	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.3	1.000	0.992	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.4	1.000	0.974	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.5	1.000	0.986	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.6	1.000	0.980	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.7	1.000	0.985	0.998	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.8	1.000	0.961	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	0.9	1.000	0.983	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
0.9	1.0	1.000	0.985	0.995	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.1	1.000	0.971	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.2	1.000	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.3	1.000	0.984	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.4	1.000	0.983	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.5	1.000	0.987	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.6	1.000	0.987	0.998	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.7	1.000	0.979	0.999	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.8	1.000	0.966	0.997	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	0.9	1.000	0.967	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
1.0	1.0	1.000	0.971	0.996	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000

Note: We report the p-values based on the Chi-Square tests performed on the data sets whose values follow a Benford's distribution partially replaced by data following a uniform distribution. The uniform distribution has a lower bound 0 and an upper bound ranging from 0.1 to 1. The values are based on the average of 100 bootstrap iterations.

REFERENCES

[AMH] Ader, H. J., G.J Mellenbergh, and D. J. Hand. Advising on Research Methods: A Consultant's Companion. *Johannes van Kessel Publishing*, 2008

[CLTF] Costas, Eduardo, Victoria Lopez-Rodas, F. Javier Toro, and Antonio Flores-Moya. The Number of Cells in Colonies of the Cyanobacterium Microcystis Aeruginosa Satisfies Benford's Law. *Aquatic Botany*, 89(3):341-343, Oct. 2008

[DN] Drake, Philip D., and Mark J. Nigrini. Computer Assisted Analytical Procedures Using Benford's Law. *Journal of Accounting Education*, 18(2): 127-146, Spring 2000

[Ef] Efron, Bradley. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics*, 7(1):1-26, Jan. 1979

[FLP] Fleiss, Joseph L., Bruce Levin, and Myunghee Cho Paik. Statistical Methods for Rates and Proportions (Third Edition). *Wiley*, Sept. 5, 2003

[GM] Geyer, A., and J. Marti. Applying Benford's Law to Volcanology. *Geology*, 40(4):327-330, April 2012

[Ne] Newcomb, Simon. Note on the Frequency of Use of the Different Digits in Natural Numbers. *American Journal of Mathematics*, 4(1):39-40, 1881

[NiMil] Nigrini, Mark J., and Steven J. Miller. Benford's Law Applied to Hydrology Data - Results and Relevance to Other Geophysical Data. *Mathematical Geology*, 39(5):469-490, July 2007

[NiMit] Nigrini, Mark J., and Linda J. Mittermaier. The Use of Benford's Law as an Aid in Analytical Procedures. *Accounting: A Journal of Practice & Theory*, 16(2):52-67, Fall 1997

[Pi] Pike, Daniel P. Testing for the Benford's Property. *SIAM Undergraduate Research Online*, 1(1):10-19, July 2008