Do Financial Items Determined Using Estimates and/or Professional Judgment Adhere to Benford's Law of First Digits?

Charles E. Jordan
Florida State University – Panama City

Stanley J. Clark
University of Alabama in Huntsville

Among other things, Benford's Law provides the rates at which the numerals one through nine should materialize as the first digit of naturally occurring figures and is often used as an analytical tool for identifying suspicious data sets. Yet, recent research in Romania suggests Benford's Law does not hold for financial statement items whose determination requires extensive use of estimates and/or professional judgment (e.g., depreciation expense or revenue) and, accordingly, should not be used for evaluating such items. The current article reports on similar research conducted in the U.S. and shows that all financial items examined, including those requiring significant use of estimates and/or professional judgment, conform to Benford's Law.

Keywords: Benford's Law, financial items, estimates, professional judgment

INTRODUCTION

With today's heavy emphasis on data analytics, Benford's Law (aka the first digit law) has gained significant traction as a means of identifying possible manipulations within large data sets. A simple keyword search using "Benford's Law" in the Business Source Complete platform of the EbscoHost online database of business/accounting articles reveals a total of 246 manuscripts addressing or using Benford's observations concerning the frequency distributions of leading digits in naturally occurring data. The theorem's popularity likely stems from its simplicity as well as its broad applicability.

The first digit law was originally postulated by mathematician and astronomer Simon Newcomb in 1881 who found that the initial few pages of books of logarithmic tables were more tattered than the last few pages (Newcomb, 1881). From this, he surmised that people looked up numbers in the tables that began with low digits (e.g., ones, twos, or threes) more frequently than numbers starting with high digits (e.g., sevens, eights, or nines) and that this characteristic reflected the nature of numbers occurring in practice. Newcomb developed mathematical formulas based on geometric progression for determining the rates at which the numerals one through nine would appear as the left leading digit in numbers in practice as well as the rates at which the numerals zero through nine would materialize in each digital position to the right of the leading digit.

About 50 years later and unaware of Newcomb's work in this area, General Electric physicist Frank Benford noticed the same phenomenon when he looked up numbers in logarithmic tables and similar to

Newcomb concluded that lower numerals occur more frequently in the leading digital positions of numbers appearing in practice than higher numerals (Benford, 1938). However, Benford went further than Newcomb and over a period of years collected 20 lists of numbers from data sets occurring in practice. His lists came from a variety of sources (e.g., topographical, scientific, demographic, etc.) and totaled 20,229 data points (see Nigrini, 1999). After analyzing the data sets, Benford found that ones occurred as the initial or first digit about 30.6 percent of the time and that the frequency of each successive number (i.e, two through nine) occurring as the initial digit declined (e.g., with nines materializing in this digital position at a frequency of only 4.7 percent). Benford then developed mathematical theorems equivalent to those Newcomb had earlier derived for determining the frequency distributions of the digits one through nine appearing in the left first digital position of data occurring in nature as well as the rates at which the numerals zero through nine should materialize in positions to the right of the first digit. Table 1 presents Benford's expected rates for each of the initial three digital positions.

TABLE 1
BENFORD'S DIGITAL RATES

Digital Position in Number			
Digit	1st	2nd	3rd
0		11.97%	10.18%
1	30.10%	11.39	10.14
2	17.61	10.88	10.10
3	12.49	10.43	10.06
4	9.69	10.03	10.02
5	7.92	9.67	9.98
6	6.70	9.34	9.94
7	5.80	9.04	9.90
8	5.12	8.76	9.86
9	4.58	8.50	9.83

Adapted from Nigrini & Mittermaier (1997)

In the number 47,369, four appears as the initial or first digit with seven as the second digit and so on. Interpreting Table 1 is quite forthright. For example, Benford's Law shows that twos should emerge as the first digit 17.61 percent of the time and as the third digit 10.10 percent of the time while nines should occur as the first digit at a rate of 4.58 percent and as the second digit at a frequency of 8.50 percent. The numerals zero through nine would be expected to materialize in digital positions beyond the third digit at approximately equal rates.

The logic behind Benford's Law is altogether intuitive and can be explained with a rudimentary example. Assume a company's net earnings for the present year is $100,000 and is growing at the rate of 6 percent annually. At this growth rate, the first digit will remain a one for 12 years until net income hits the $200,000 mark. However, it will only take seven more years for the first digit to increase again (i.e., net earnings of $300,000) and only five years after that for the first digit to go up again (i.e., net earnings of $400,000). As net income grows larger, it takes less and less time for the first digit to increase by one, until net income reaches the $1,000,000 mark, at which point the process starts anew. Thus, in a large cross-sectional sample of companies of varying ages, more companies would have lower numerals (e.g., ones or twos) as the first digit of net earnings than would have higher numerals (e.g., eights or nines) in this position.

As noted previously, Benford's Law has become a relatively common analytical tool for evaluating whether large sets of financial data appear natural (i.e., unmanipulated). However, Jianu and Jianu (2021) call into question the use of Benford's Law for examining financial statement items whose derivation requires the significant use of estimates and/or professional judgment. They argue that such numbers have been too heavily influenced by human intervention to expect them to exhibit the digital distributions found in naturally occurring numbers and offer proof from data obtained from Romanian publicly-traded companies in support of their hypothesis. By analyzing large data sets of various financial statement items for U.S. companies, however, the current study provides evidence that all financial statement items examined, even those subject to significant estimates and/or professional judgment, conform to Benford's Law.

The next section discusses some of the relevant literature related to Benford's Law and develops the research question for the study. This is followed by the methodology and results sections. The final section contains our conclusions and the implications of the research.

LITERATURE REVIEW AND DEVELOPMENT OF RESEARCH QUESTION

Varian (1972) notes that adherence of a data set to Benford's frequencies does not necessarily signify the data are authentic but nonconformity should spark some level of doubt about the data set's purity. Many studies have used Benford's Law to check for the manipulation of accounting data. For example, Carslaw (1988) examined net income numbers for New Zealand companies in testing for earnings rounding. He discovered that nines (zeros) occurred as the second digit of earnings numbers significantly less (more) often than anticipated according to Benford's frequencies. This led him to conclude that when the second earnings digit was high (i.e., a nine), managers manipulated income upward until the second digit just crossed the threshold (i.e., reached zero). The purpose of this manipulation was to enlarge the first digit in the income number by one. Similar studies were conducted and with analogous results in numerous countries (e.g., Jordan et. al, 2014, in Canada; Lebert et. al, 2018, in Germany; Niskanen and Keloharju, 2000, in Finland; Skousen et. al, 2004, in Japan; Thomas, 1989, in the U.S.; Van Caneghem, 2002, in the U.K.).

Christian and Gupta (1993) analyzed taxpayer data using Benford's Law and found signs of fraud as the examination indicated taxpayers attempted to reduce their taxable income from a higher to a lower tax bracket, thereby enabling them to pay less taxes. Nigrini (1994) applied Benford's Law to a fraudulent set of payroll data (i.e., after the fact) and discovered that the data did not conform to the expected frequencies, thus suggesting that Benford's Law could be useful in detecting the presence of fraudulent accounts or transactions in large data sets. Using Benford's Law, Nigrini (1996) examined the first and second digit distributions for interest paid and interest received amounts reported in about 200,000 tax returns filed in the U.S. in the mid-1980s. While in general the distributions conformed to Benford's expected frequencies, he did observe an inclination for an overabundance of low digits for interest revenue and high digits for interest expense.

Nigrini and Mittermaier (1997) show how digital analysis, including Benford's Law, could be employed as an analytical review method to aid auditors in identifying data sets that need closer examination during an audit. Similarly, Singh and Best (2020) also examine Benford's Law as a means of performing analytical review procedures in an audit context and conclude that "Benford's analysis, when used correctly, is a useful tool for identifying suspicious transactions for further analysis (p. 400)."

The key to using Benford's frequencies as an analytical tool for identifying suspicious data sets in an accounting environment is the belief that unmanipulated financial accounting numbers adhere to Benford's Law. Indeed, Singh and Best (2020) note that most accounting data sets conform to Benford's Law. Nigrini and Mittermaier (1997) conclude likewise and note that auditors, researchers and other interested parties can "assume that lists of [financial statement] items, such as accounts receivable or payable, inventory counts, fixed asset acquisitions, daily sales, and disbursements, should follow Benford's Law (p. 57)." They go on to state that deviations from Benford's expected distributions could indicate the presence of irregularities in financial data and that "If the human element is present, then the digit patterns could differ from those of Benford's Law. The human element refers to numbers that have deliberately been invented or estimated (Nigrini and Mittermaier, 1997, p. 57)."

Jianu and Jianu (2021) seem to interpret the above statement by Nigrini and Mittermaier (1997) concerning the "human element" to imply that financial statement items subject to estimates and/or professional judgment would not be expected to adhere to Benford's Law because they have been too heavily influenced by human thought. Jianu and Jianu (2021) test this hypothesis on a group of Romanian companies whose stock is traded on the Bucharest Stock Exchange (BSE) for two distinct time periods (i.e., before and after 2012, which is when Romania adopted International Financial Reporting Standards). The researchers examined the first digits for several financial statement items relative to their conformity with Benford's frequencies.

For items whose derivation is relatively clear cut and subject to little human thought (e.g., accounts payable, accounts receivable, operating cash flows, and income tax expense, which in Romania is determined based on the tax code), Jianu and Jianu (2021) discovered that the items complied with Benford's Law both before and after the adoption of IFRS. In addition, the chi-square test statistics were smaller for the post-IFRS samples than for the pre-IFRS samples, thus leading the researchers to conclude that IFRS improved financial reporting reliability in Romania (i.e., that items in the post-IFRS period conformed more closely to Benford's Law than items in the pre-IFRS era).

To test whether financial statement items subject to estimates conform to Benford's Law, Jianu and Jianu (2021) examined depreciation expense because it requires management or the accountant to estimate the useful life of the asset as well as its residual value. In addition, human thought is needed in selecting the depreciation method. The researchers found that the first digit for depreciation expense failed to adhere to Benford's frequencies in both the pre- and post-IFRS samples. This led the researchers to "conclude that the values reported in the financial statements based on estimates do not seem to conform to Benford's Law (Jianu and Jianu, 2021, p. 16)."

To evaluate if financial statement items requiring extensive use of professional judgment conform to Benford's Law, Jianu and Jianu (2021) tested the financial statement item "revenues." The researchers noted that prior to the implementation of IFRS in Romania, clear rules existed in the national accounting regulation concerning revenue recognition. Yet, the implementation of IFRS 15 (the current standard for revenue recognition) in Romania introduced complexities that require the use of professional judgment. In testing the first digit for revenues, Jianu and Jianu (2021) discovered that the pre-IFRS sample conformed to Benford's Law while the post-IFRS sample did not. This finding led them to conclude that Benford's Law is appropriate for evaluating financial statement items whose amounts are calculated using clearly defined rules but that it should not be used for examining amounts determined through the use of accountants' professional judgment.

Nonetheless, Jianu and Jianu (2021) identify a couple of limitations that may have affected their results. First is the small sample sizes that were available because of the relatively low number of publicly-traded companies in Romania. The researchers noted that, generally, a minimum sample size of 500 is needed for evaluating the conformity of a data set to Benford's Law. Aris et. al (2017) indicate that sample sizes between 2,500 and 5,000 should be used when employing statistical tests like chi-square or z-statistics for evaluating conformity with Benford's Law. Several of the financial items tested by Jianu and Jianu (2021) had sample sizes below 500, including the post-IFRS sample for depreciation expense (i.e., 443 observations). Second, Jianu and Jianu (2021) note that companies listed on the BSE would be considered members of an emerging market.

This leads to the research question for the present study. In particular, will the results from Jianu and Jianu (2021) hold true for companies in a mature capital market with larger sample sizes available for testing their adherence to Benford's Law? We examine the compliance, or lack thereof, with Benford's Law for a wide range of financial statement items for listed companies in the U.S., including items that are determined using rather definitive procedures as well as those requiring the use of estimates and/or professional judgment.

METHODOLOGY

To test whether financial statement items determined using estimates and/or professional judgment in the U.S. conform to Benford's Law, 2021 data are collected for publicly-traded companies in COMPUSTAT's annual fundamentals file. Since Jianu and Jianu (2021) show that balance sheet items (e.g., accounts payable and accounts receivable) and cash flows from operating activities adhere to Benford's Law, the current study does not examine balance sheet accounts or cash flows but instead concentrates on income statement items. Some accounts analyzed are calculated with virtually no use of estimates or professional judgment but instead are contractually determined (i.e., interest expense and rent expense). Other items examined require at least some use of estimates and/or professional judgment in their determination (i.e., cost of goods sold, pension and retirement expense, research and development expense, and selling, general and administrative expense).

Finally, three accounts calculated with significant use of estimates and/or professional judgment are examined. First is stock compensation expense, which under the rules of Codification Topic 718 is typically estimated through the use of fair value models. Second is depreciation expense, which as Jianu and Jianu (2021) point out is based heavily on the use of estimates and professional judgment (e.g., professional judgment is required in selecting the depreciation method and estimates are used in choosing useful lives and residual values). Third is sales revenue, which requires significant professional judgment in several areas under Codification Topic 606 (e.g., in determining whether a sales contract exists, identifying the separate performance obligations in the contract, ascertaining the transaction price, allocating the transaction price among the separate performance obligations, and discerning when the revenue should be recognized i.e., over time or at a point in time).

Jianu and Jianu (2021) examined only the first digit of their selected items for conformity with Benford's Law. As Brenner and Brenner (1982) show, because of their limited amount of memory, people place more emphasis on the first digit in a number with increasingly less interest put on the second, third, and so on digits. For this reason, and to be consistent with the Jianu and Jianu (2021) study, the current research analyzes only the first digit of the chosen items for ascertaining their conformity with Benford's Law. COMPUSTAT reports dollar amounts in millions and uses decimal points to delineate between millions and thousands. For example, an amount presented in COMPUSTAT as $1.433 represents an actual number of $1,433,000. To ensure, the first digits in the samples were not affected by rounding for inclusion in COMPUSTAT, only financial statement items greater than $10,000 (i.e., $.010 in COMPUSTAT) are collected.

A crucial issue in ascertaining a data set's adherence to Benford's frequencies is using the appropriate statistical test for evaluating differences between the actual rates observed and Benford's expected rates. Cleary and Thibodeau (2005) note that performing nine individual tests for evaluating each number's (i.e., one through nine) conformity with its expected frequency under Benford's Law likely increases the risk of Type I errors (i.e., identifying a data set as failing to adhere to Benford's Law when in fact it does conform). There are situations where z-statistics for individual numbers are warranted (e.g., in testing the rates that nines and zeros occur as the second income digit in earnings rounding research i.e., see Aono and Guan, 2008; Carslaw, 1988; Jordan and Clark, 2011; Kinnunen and Koskela, 2003; Lin and Wu, 2014). However, in the current study, there is no *a priori* reason to believe the actual rate for any particular numeral one through nine appearing as the first digit would deviate significantly from its expected frequency under Benford's Law. Instead, the issue here is determining whether the actual frequencies of the numbers one through nine in the first digit conform overall to Benford's Law. For this reason, as Sadaf (2017) indicates, most studies testing actual data for conformity with Benford's frequencies use chi-square tests for evaluating overall statistical significance. Accordingly, the present research uses chi-square tests as the primary means of determining conformity with Benford's Law, which is consistent with the Jianu and Jianu (2021) study. Also, like Jianu and Jianu (2021), the commonly used alpha level of .05 is employed for statistical testing.

Nigrini (2012) points out that statistical tests like chi-square and z-statistics become less useful at identifying discrepancies between actual frequencies and Benford's expected rates for sample sizes above 5,000. This is because the statistical tests for these large samples suffer from excess power, which can result in even minor differences between actual and expected rates being flagged as statistically significant. Drake and Nigrini (2000) developed a non-statistical method of evaluating a data set's conformity with Benford's Law that can be used for large samples (i.e., those where N exceeds 5,000). The test figure is known as the mean absolute deviation (MAD), which is computed for the first digit as the sum of the absolute values of the differences between the actual frequency and Benford's expected rate for each of the nine numerals materializing as the first digit divided by the number of numerals (i.e., nine). There is no cutoff test for statistical significance for MAD; nevertheless, Drake and Nigrini (2000) developed the following chart for using MAD in ascertaining whether the first digit frequencies for a data set adhere to Benford's Law.

Conformity decision:	Range for MAD:
Close conformity	0.00 % – 0.40 %
Acceptable conformity	0.40 % – 0.80 %
Marginally acceptable conformity	0.80 % – 1.20 %
Nonconformity	Greater than 1.20 %

For example, a data set producing a MAD of 0.61 percent would be considered in acceptable conformity with Benford's Law, while one generating a MAD of 1.30 percent would be viewed in nonconformity with the first digit law. Although none of the data sets in the current study exceed a sample size of 5,000, a few are above 4,000, and out of an abundance of caution the MAD number is computed and evaluated as a secondary test for each financial statement item examined.

RESULTS

Panels A and B of Table 2 present the results for interest expense and rent expense, respectively (i.e., the two items requiring only minimal, if any, use of estimates and/or professional judgment in their calculation). In Panel A, the first row of data presents the actual number of times the first digit of interest expense appeared as each of the numerals one through nine. The second row provides the actual percentage that each numeral occurred as the first digit. For example, ones materialized as the first digit 1,095 times, which represents 30.76 percent of the 3,560 observations for interest expense. The next row provides Benford's expected frequency for each numeral zero through nine occurring as the first digit, while the final row of data gives the mathematical difference between the actual percentage for each numeral and its anticipated rate. As an example, Benford's expected rate for ones as the first digit is 30.10 percent, which for interest expense differs from the actual frequency (30.76 percent) by only 0.66 percent.

TABLE 2
FIRST DIGIT RATES FOR INTEREST EXPENSE AND RENT EXPENSE

Panel A (Interest Expense) N=3560									
			Number Appearing as the First Digit						
	1	2	3	4	5	6	7	8	9
Observed count (n)	1095	649	474	324	292	222	179	155	170
Observed rate (%)	30.76	18.23	13.31	9.1	8.2	6.24	5.03	4.35	4.78
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	0.66	0.62	0.82	-0.59	0.28	-0.46	-0.77	-0.77	0.2
Chi-square & p-level	14.04	(.081)							
MAD % & category	0.57 (acceptable conformity)								

Panel B (Rent Expense) N=3748									
Observed count (n)	1088	637	484	361	299	270	251	175	183
Observed rate (%)	29.03	17	12.91	9.63	7.98	7.2	6.7	4.67	4.88
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	-1.07	-0.61	0.42	-0.06	0.06	0.5	0.9	-0.45	0.3
Chi-square & p-level	11.65	(.167)							
MAD % & category	0.49 (acceptable conformity)								

At a .05 significance level, the critical value for a chi-square test with eight degrees of freedom is 15.51. Panel A shows that the chi-square statistic for evaluating the differences between the number of times each numeral appeared as the first digit of interest expense and the frequency at which they were expected under Benford's Law is 14.04, which is less than the critical value and indicates the actual frequencies of the numerals appearing as the initial digit conform to Benford's Law. This is supported by the MAD rate of 0.57 percent, which falls in the 0.40 – 0.80 percent range for acceptable conformity.

Panel B of Table 2 presents similar results for rent expense. More specifically, the chi-square statistic of 11.65 (p-value of .167) and MAD rate of 0.49 percent, which falls in the acceptable conformity range, both indicate adherence to Benford's frequencies for the first digit. The findings in Table 2 come as no surprise since both interest expense and rent expense are contractually derived for the most part with very little human intervention involved. Notice that the sample sizes for interest expense (3,560) and rent expense (3,748) do not equal simply because some companies reporting rent expense provided no interest expense and vice versa.

Table 3 provides the first digit results for cost of goods sold (COGS), pension and retirement (P&R) expense, research and development (R&D) expense, and selling, general, and administrative (SG&A) expense in Panels A, B, C, and D, respectively. These four items each require at least some use of estimates and/or professional judgment in their derivation. As an example, when research and development, administrative (e.g., legal or payroll), and production activities all occur in various parts of the same building, an accountant would need to allocate utility cost incurred during the period on that building among R&D expense, SG&A expense, and the production function (i.e., COGS). Such an allocation would require the use of estimates and/or professional judgment.

TABLE 3
FIRST DIGIT RATES FOR COGS, P&R EXPENSE, R&D EXPENSE, AND SG&A EXPENSE

Panel A (COGS) N=4613									
	Number Appearing as the First Digit								
	1	2	3	4	5	6	7	8	9
Observed count (n)	1425	789	604	432	337	292	281	223	230
Observed rate (%)	30.89	17.1	13.09	9.37	7.31	6.33	6.09	4.83	4.99
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	0.79	-0.51	0.6	-0.32	-0.61	-0.37	0.29	-0.29	0.41
Chi-square & p-level	9.69	(.287)							
MAD % & category	0.47 (acceptable conformity)								

Panel B (P&R Expense) N=3130									
Observed count (n)	945	563	410	306	257	182	174	132	161
Observed rate (%)	30.19	17.99	13.1	9.78	8.21	5.81	5.56	4.22	5.14
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	0.09	0.38	0.61	0.09	0.29	-0.89	-0.24	-0.9	0.56
Chi-square & p-level	12.68	(.123)							
MAD % & category	0.45 (acceptable conformity)								
Panel C (R&D Expense) N=2245									
Observed count (n)	667	381	251	216	203	175	130	122	100
Observed rate (%)	29.71	16.97	11.18	9.62	9.04	7.8	5.79	5.43	4.46
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	-0.39	-.0.64	-1.31	-0.07	1.12	1.1	-0.01	0.31	-0.12
Chi-square & p-level	11.83	(.159)							
MAD % & category	0.56 (acceptable conformity)								
Panel D (SG&A Expense) N=4173									
Observed count (n)	1232	780	517	379	339	272	227	223	204
Observed rate (%)	29.52	18.69	12.39	9.08	8.12	6.52	5.44	5.34	4.89
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	-0.58	1.08	-0.1	-0.61	0.2	-0.18	-0.36	0.22	0.31
Chi-square & p-level	7.49	(.485)							
MAD % & category	0.41 (acceptable conformity)								

Despite the obvious presence of at least modest levels of human intervention in the calculation of the four financial statement items examined in Table 3, their chi-square statistics and MAD percentages suggest that each one conforms to Benford's frequencies for the first digit (see the key information from Table 3 reproduced below showing that, for each item, the chi-square statistic falls below the critical value of 15.51 and the MAD percentage lands comfortably in the acceptable conformity range).

Financial item	Chi-square	(p-value)	MAD percentage and category
COGS	9.69	(.287)	0.47 - acceptable conformity
P&R expense	12.68	(.123)	0.45 - acceptable conformity
R&D expense	11.83	(.159)	0.56 - acceptable conformity
SG&A expense	7.49	(.485)	0.41 - acceptable conformity

Table 4 presents the findings for the three financial statement items requiring significant use of estimates and/or professional judgment in their derivation (i.e., Panels A, B, and C contain the results for stock compensation expense, depreciation expense, and sales revenue, respectively).

TABLE 4
FIRST DIGIT RATES FOR STOCK COMPENSATION EXPENSE, DEPRECIATION EXPENSE AND SALES REVENUE

Panel A (Stock Compensation Expense) N=4374									
	Number Appearing as the First Digit								
	1	2	3	4	5	6	7	8	9
Observed count (n)	1318	836	536	402	351	286	236	212	197
Observed rate (%)	30.13	19.11	12.25	9.19	8.03	6.54	5.4	4.85	4.5
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	0.03	1.5	-0.24	-0.5	0.11	-0.16	-0.4	-0.27	-0.08
Chi-square & p-level	9.09	(.335)							
MAD % & category	0.37 (close conformity)								
Panel B (Depreciation Expense) N=4357									
Observed count (n)	1373	795	534	428	322	251	243	228	183
Observed rate (%)	31.51	18.25	12.26	9.82	7.39	5.76	5.58	5.23	4.2
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	1.41	0.64	-0.23	0.13	-0.53	-0.94	-0.22	0.11	-0.38
Chi-square & p-level	13.29	(.102)							
MAD % & category	0.51 (acceptable conformity)								
Panel C (Sales Revenue) N=4702									
Observed count (n)	1444	830	549	442	411	323	286	229	188
Observed rate (%)	30.71	17.65	11.68	9.4	8.74	6.87	6.08	4.87	4
Benford's rate (%)	30.1	17.61	12.49	9.69	7.92	6.7	5.8	5.12	4.58
Difference (%)	0.61	0.04	-0.81	-0.29	0.82	0.17	0.28	-0.25	-0.58
Chi-square & p-level	12.39	(.135)							
MAD % & category	0.43 (acceptable conformity)								

If the digital frequencies for any financial items fail to adhere to Benford's Law because of human intervention in the form of estimates and/or professional judgment, it is the three items in Table 4 for which that nonconformity would be the most likely to occur. The key measures from Table 4 for evaluating adherence to Benford's first digit rates for stock compensation expense, depreciation expense, and sales revenue are as follows:

Financial item	Chi-square	(p-value)	MAD percentage and category
Stk. comp. expense	9.09	(.335)	0.37 - close conformity
Dep. expense	13.29	(.102)	0.51 - acceptable conformity
Sales revenue	12.39	(.135)	0.43 - acceptable conformity

The summary results above show that all three of the items requiring significant use of estimates and/or professional judgment conform to Benford's rates for first digits. That is, each one produces a chi-square statistic below the critical value of 15.51 and a MAD percentage falling in either the close conformity or acceptable conformity ranges. This finding stands in marked contrast to the outcome in Jianu and Jianu (2021) where depreciation expense failed to adhere to Benford's Law for both the pre- and post-IFRS samples and sales revenue lacked conformity with Benford's rates for the post-IFRS sample. There is no way to know definitively what caused the discrepancies between the results in the current study and those found in Jianu and Jianu (2021). Nonetheless, one explanation could be the vastly different sample sizes in the two studies. For example, the post-IFRS sample size for revenue in the Jianu and Jianu (2021) research was only 516, while the sample size for revenue in the present study is 4,702. As mentioned previously, Aris et. al (2017) note that relatively large sample sizes (i.e., between 2,500 and 5,000 observations) are best for using statistical tests (like chi-square) to evaluate a data set's adherence to Benford's Law.

Another possible explanation for the discrepancies in the findings could be one alluded to by Jianu and Jianu (2021). In particular, they note that their sample is from Romanian companies traded on the BSE, which represents an emerging market. Jianu and Jianu (2021) identify this as a potential limitation of their study as they further state that the maturity of the capital market may impact the reliability of financial reporting within that market. The current study examines companies in a mature capital market (i.e., the U.S.); thus, this may drive the results found here (i.e., that all financial statement items examined appear to be reliably measured and reported, at least with respect to their conformity with Benford's Law).

SUMMARY AND CONCLUSION

Nigrini and Mittermaier (1997) indicate that the digit frequencies for data sets containing financial statement items or account balances would be expected to follow Benford's Law; however, they further note that the presence of a human element could cause the digit patterns for such items to deviate from Benford's rates. "The human element refers to numbers that have deliberately been invented or estimated (Nigrini and Mittermaier, 1997, p 57)." Jianu and Jianu (2021) seem to interpret the above statement to mean that financial items determined using estimates and/or accountants' professional judgment would not be expected to conform to Benford's Law. They tested this hypothesis on a number of financial statement items obtained from pre- and post-IFRS samples of Romanian companies and found that the majority of the accounts conformed to Benford's Law but that neither pre- nor post-IFRS depreciation expense nor post-IFRS revenue adhered to Benford's frequencies. Thus, Jianu and Jianu (2021) concluded that it is inappropriate to evaluate financial items determined with estimates and/or professional judgment against Benford's Law because such items should not be expected to follow Benford's digit patterns.

The above finding by Jianu and Jianu (2021) casts doubt on the application of Benford's Law as an analytical procedure for identifying possible irregularities within data sets of financial statement items. This is because under accrual basis accounting and the application of GAAP, the majority of financial account balances require at least some level of estimates and/or professional judgment in their determination. How does one ascertain whether an item's calculation requires the level of estimates and/or professional judgment that negates its conformity with Benford's Law? That is, if a data set for a financial item fails to comply with Benford's Law, does this lack of adherence stem from fraudulent activity or manipulation of the numbers or does it simply mean the data set was not expected to conform with Benford's frequencies in the first place because the derivation of the account balances required the use of estimates and/or professional judgment?

We take a different view than Jianu and Jianu (2021) concerning the meaning of the "human element" referenced by Nigrini and Mittermaier (1997) who state that financial items might not conform with Benford's Law if they "have deliberately been invented or estimated (Nigrini and Mittermaier, 1997, p. 57)." The key words here seem to be "deliberately" and "invented," which suggest the authors were referring to data items that are fabricated or made up with some desired goal in mind. It is unlikely Nigrini and Mittermaier (1997) meant this to include the benign use of estimates and/or professional judgment like choosing the useful life of an asset or the method used to depreciate it. In our view, using estimates and/or professional judgment in calculating financial items would not be expected to cause them to be in nonconformity with Benford's Law unless the estimates and/or professional judgment were being used to create specific or desired outcomes.

The results in the current study appear to support this notion. In particular, there was no *a priori* reason to believe that the numbers for any of the financial items examined here would have been manipulated or deliberately invented. Indeed, the data sets for all items analyzed conformed to Benford's Law, even those whose calculation necessitated the significant use of estimates and/or professional judgment (e.g., depreciation expense and revenue). The disparity between the results obtained in the current study and those found in Jianu and Jianu (2021) likely stems from the different samples used. In particular, the current project examined large samples of companies operating in a mature capital market (i.e., U.S), while Jianu and Jianu (2021) analyzed relatively small samples of firms taken from an emerging market (i.e., Romania).

The findings here suggest that Benford's Law is appropriate for evaluating the digital frequencies of financial items, without regard to whether those items are determined via the use of estimates and/or professional judgment. Nevertheless, to ensure the results transcend national boundaries, this study should be replicated in other countries with mature capital markets. Also, the present research provides general results for the U.S., but no industry analysis is made to ascertain whether the findings hold true within individual industries. Future research could address this as well.

REFERENCES

Aono, J., & Guan, L. (2008). The impact of the Sarbanes-Oxley Act on cosmetic earnings management. *Research in Accounting Regulation*, *20*, 205–215.

Aris, N., Othman, R., Bukhori, M., Arif, S., & Malek, M. (2017). Detecting accounting anomalies using Benford's Law: Evidence from the Malaysian public sector. *Management & Accounting Review*, *16*(2), 73–99.

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, *78*(4), 551–572.

Brenner, G., & Brenner, R. (1982). Memory and markets, or why are you paying $2.99 for a widget? *Journal of Business*, *55*(1), 147–158.

Carslaw, C. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review*, *63*(2), 321–327.

Christian, C., & Gupta, S. (1993). New evidence on secondary evasion. *The Journal of the American Taxation Association*, *15*(1), 72–87.

Cleary, R., & Thibodeau, J. (2005). Applying digital analysis using Benford's Law to detect fraud: The dangers of type I errors. *Auditing: A Journal of Practice & Theory*, *24*(1), 77–81.

Drake, D., & Nigrini, M. (2000). Computer assisted analytical procedures using Benford's Law. *Proceedings of Journal of Accounting Education*, *18*, 127–146.

Jianu, I., & Jianu, I. (2021). Reliability of financial information from the perspective of Benford's Law. *Entropy*, *23*(5), 1–22.

Jordan, C., & Clark, S. (2011). Detecting cosmetic earnings management using Benford's Law. *The CPA Journal*, *81*(2), 32–37.

Jordan, C., Clark, S., & Waldron, M. (2014). Cosmetic earnings management before and after corporate governance legislation in Canada. *Accounting and Finance Research*, *3*(4), 105–114.

Kinnunen, J., & Koskela, M. (2003). Who is Miss World in cosmetic earnings management? A cross-national comparison of small upward rounding of net income numbers among eighteen countries. *Journal of International Accounting Research*, *2*, 39–68.

Lebert, S., Mohrmann, U., & Stefani, U. (2018). Rounding up performance measures in German firms: Earnings cosmetics or earnings management on a larger scale? *Journal of Business Finance & Accounting*, *48*(3/4), 564–586.

Lin, F., & Wu, S. (2014). Comparison of cosmetic earnings management for the developed markets and emerging markets: Some empirical evidence from the United States and Taiwan. *Economic Modelling*, *36*, 466–473.

Newcomb, S. (1881). Note on the frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, *4*, 39–40.

Nigrini, M. (1994). Using digital frequencies to detect fraud. *The White Paper*, pp. 3–6.

Nigrini, M. (1996). A taxpayer compliance application of Benford's Law. *The Journal of the American Taxation Association*, *18*(1), 72–91.
Nigrini, M. (1999). I've got your number. *Journal of Accountancy*, *187*(5), 79–83.
Nigrini, M. (2012). *Benford's Law: Applications for forensic accounting, auditing, and fraud detection*. Hoboken, NJ: John Wiley & Sons.
Nigrini, M., & Mittermaier, L. (1997). The use of Benford's Law as an aid in analytical procedures. *Auditing: A Journal of Practice & Theory*, *16*(2), 52–67.
Niskanen, J., & Keloharju, M. (2000). Earnings management in a tax-driven accounting environment: Evidence from Finnish public firms. *European Accounting Review*, *9*(3), 443–452.
Sadaf, R. (2017). Advanced statistical techniques for testing Benford's Law. *The Annals of the University of Oradea, Economic Sciences*, *26*(2), 229–238.
Singh, K., & Best, P. (2020). Implementing Benford's Law in continuous monitoring applications. *Accounting and Management Information Systems*, *19*(1), 379–404.
Skousen, C., Guan, L., & Wetzel, T. (2004). Anomalies and unusual patterns in reported earnings: Japanese managers round earnings. *Journal of International Financial Management & Accounting*, *15*(3), 212–234.
Thomas, J. (1989). Unusual patterns in reported earnings. *The Accounting Review*, *64*(4), 773–787.
Van Caneghem, T. (2002). Earnings management induced by cognitive reference points. *British Accounting Review*, *34*(2), 167–178.
Varian, H. (1972). Benford's Law. *The American Statistician*, *26*, 65–66.