Rev Account Stud (2015) 20:1540–1593
DOI 10.1007/s11142-015-9333-z

Financial statement errors: evidence from the distributional properties of financial statement numbers

Dan Amiram[1] · Zahn Bozanic[2] · Ethan Rouen[1]

Published online: 2 August 2015

Abstract Motivated by methods used to evaluate the quality of data, we create a novel firm-year measure to estimate the level of error in financial statements. The measure, which has several conceptual and statistical advantages over available alternatives, assesses the extent to which features of the distribution of a firm's financial statement numbers diverge from a theoretical distribution posited by Benford's Law. After providing intuition for the theory underlying the measure, we use numerical methods to demonstrate that certain error types in financial statement numbers increase the deviation from the theoretical distribution. We corroborate the numerical analysis with simulation analysis that reveals that the introduction of errors to reported revenue also increases the deviation. We then provide empirical evidence that the measure captures financial statement data quality. We first show the measure's association with commonly used measures of accruals-based earnings management and earnings manipulation. Next, we demonstrate that (1) restated financial statements more closely conform to Benford's Law than the misstated versions in the same firm-year and (2) as divergence from Benford's Law increases, earnings persistence decreases. Finally, we show that our measure predicts material misstatements as identified by SEC Accounting and Auditing Enforcement Releases and can be used as a leading indicator to identify misstatements.

盈余持续性

✉ Zahn Bozanic
bozanic.1@fisher.osu.edu

Dan Amiram
da2477@columbia.edu

Ethan Rouen
ethanrouen@columbia.edu

[1] Columbia Business School, Columbia University, New York, NY, USA

[2] Fisher College of Business, The Ohio State University, Columbus, OH, USA

Keywords Benford's Law · Financial statements errors · Accounting quality · Earnings management

JEL Classification M41

1 Introduction

Financial statement data that are free of error—whether in the form of misestimations, mistakes, biases, or manipulation—are crucial for well-functioning capital markets. Accurate financial reports enable efficient resource allocation and efficient contracting (Bushman and Smith 2003). Therefore, assessing the errors in financial statements is an important task for investors, analysts, auditors, regulators, and researchers. Prior literature has taken important steps in creating and validating methods to assess different constructs of errors in firm-level financial statement information, such as accruals quality or earnings quality. However, despite substantial progress in this area, available methods have deficiencies that limit their usefulness. Prior accounting literature outlines the limitations of current measures of financial statement errors, such as their correlation with underlying firm characteristics and their reliance on time-series, cross-sectional, or forward-looking data, to name a few (Dechow et al. 2010; Owens et al. 2013). We build on a statistical method developed by researchers in a variety of disciplines to assess the level of error in data. We construct a parsimonious, *firm-year* measure to assess the level of error in financial statements that overcomes some of the concerns surrounding existing measures.

Literature in mathematics, statistics, and economics suggests that examining the distribution of the first or leading digits (e.g., the leading digit of the number 217.95 is 2) of the numbers contained in a dataset allows users to assess the level of error within the underlying data. The theoretical foundation of prior research using this method is based, implicitly or explicitly, on the theorem proved by Hill (1995), which states that if distributions are selected at random and random samples of varying magnitudes are then taken from each of these distributions, the leading digits of the combined mixture distribution will converge to the logarithmic or Benford distribution, otherwise known as Benford's Law.[1] Specifically, Benford's Law states that the first digits of all numbers in an empirical dataset will appear with decreasing frequency (that is, 1 will appear as the first digit 30.1 % of the time, 2 will appear 17.6 % of the time, and so forth).[2] Methods based on the law have been used to detect errors in published scientific studies, questionable election data in Iran, suspicious macroeconomic data, internal accounts receivables data, and misreported tax returns. However, we are unaware of any attempt to apply it to the entire population of numbers contained in a firm's annual financial statements in

[1] Distributions need to be nontruncated or uncensored to conform to Benford's Law. For example, a petty cash account with a reimbursement limit of $25 would not be expected to follow Benford's Law.

[2] Please see Appendix 1 for the full theoretical distribution specified by Benford's Law.

order to ascertain whether it can be used as a firm-year measure of the degree of errors in financial reporting.

The intuition behind why empirical data follow Benford's Law can be distilled into two mathematical facts.[3] The first fact relies on using a mathematical approach to determine the first digit of any number *N*, which is to take its base 10 log and find the fraction behind the integer (i.e., the remainder or mantissa). If the fraction is between 0 and 0.301, the original number *N* will start with one, if the fraction is between 0.301 and 0.477 (interval of 0.176), the number *N* will start with 2, and so forth. Hence, the intervals between the fractions after the decimal point of the log number that determine its first digit are the same as the probabilities defined by Benford's Law. The second fact is that, if the probability distribution function of the log of the original number *N* is smooth and symmetric, the probability that a number will be in the interval between n and $n + 0.301$, where n is any integer in this logarithmic distribution, is 30.1 %. Similarly, the probability that a number will be in the intervals between $n + 0.301$ and $n + 0.477$ is 17.6 %, and so forth. Because distributions in nature tend to be smooth and symmetric due to the Central Limit Theorem, datasets tend to follow Benford's Law (Pimbley 2014). For a distribution that generally follows Benford's Law to diverge from it, certain types of errors have to be introduced to the data in a way that makes the distribution of the base 10 log less smooth or less symmetric.

The intuition outlined above likely applies to financial statement data. The true (unobservable) realizations of all cash flows, both present and future, which the items in the financial statements are intended to represent, are determined by many interactions during and after a given period. Therefore, the financial statements' line items are estimates of the realizations of cash flows from unknown random distributions. Since the true realization of every item in the financial statement is likely to be created by a different distribution (for example, the distribution of cash flows from sales that occurred during the year is likely to be different than that of administrative costs), the mixture distribution of the cash flows realization of these data may follow the criteria in Hill's (1995) theorem and therefore will be distributed according to Benford's Law. Specifically, the cash flows realization of revenue of a certain year, together with the cash flows realization of the payments to suppliers, employees, tax authorities, etc., may follow Benford's Law. However, since these realizations are unobservable in the reporting year, the preparers of the financial statements have to estimate them, a process that introduces error, whether in the form of mistakes, biases, or manipulation.[4]

[3] An intuitive but mathematically inaccurate way to briefly describe the intuition for Benford's Law is as follows. When cumulating numbers from 0, we will reach 100 before we reach 200 and 200 before we reach 300 and so forth. In the same way, the concept is scale independent, i.e., we are also going to reach 1,000,000 before we get to 9,000,000 and so forth. Given that we will stop at a random point each time we cumulate, the process will reach lower first digits (e.g., 1's and 2's) more often than higher leading or first digits (e.g., 8's and 9's).

[4] For example, the preparer needs to estimate what the returns and rebates on sales will be, as well as sales bonuses, tax payments, and so forth. If there is no error (intentional or otherwise) in the reported numbers, these items should follow Benford's Law. However, if these estimates contain certain type of errors, as the error increases, the estimates will likely diverge further from Benford's Law.

We construct a measure, the Financial Statement Divergence Score (FSD Score for short) based on the mean absolute deviation statistic as applied to the distribution of the leading digits of the numbers in annual financial statement data. The FSD Score allows us to compare the empirical distribution of the leading digits of the numbers in a firm's annual financial statements to that of the theoretical or expected distribution defined by Benford's Law. As alluded to above and detailed in the next section, the FSD Score overcomes many of the disadvantages of existing measures of accounting or earnings quality. For example, it does not require time-series or cross-sectional data to estimate, does not require forward-looking information, does not require returns or price information, and, by construction, is not likely to be correlated with firm-level characteristics or firms' business models ex ante.[5]

To provide intuition on the mathematical and statistical foundations behind the measure, we first use numerical methods to demonstrate that introducing errors to line items in financial statements will increase the divergence of the financial statements from Benford's Law. This divergence occurs because introducing errors of different size to different items in the financial statements makes the distribution less smooth and less symmetric, which, as noted above, is a condition for a distribution to follow Benford's Law.[6] We then perform a simple simulation to show that introducing errors into actual financial statement data creates deviations from Benford's Law. Since our numerical analysis suggests that deviations from the law should increase when errors are introduced to accounting numbers, we introduce errors for a typical firm in our sample by randomly manipulating its revenue. In this simple simulation, we demonstrate that the manipulation induces an increase in the FSD Score 95 % of the time.

We next assess whether the realized empirical distribution of the first digits of firms' financial statement numbers follows Benford's Law. This is a critical step in our empirical inquiry, as no study has examined whether annual financial statements are distributed according to Benford's Law. We show that, whether in aggregate, by year, by industry, or by firm-year, firms' financial statements generally conform to Benford's Law. We also demonstrate that the income statement is the most susceptible to errors while the cash flow statement is the least.

Once initial conformity is established, we continue by examining the relation between Benford's Law and commonly used measures of accruals-based earnings management and earnings manipulation. We show that the FSD Score is significantly positively related with the Dechow–Dichev measure, discretionary accruals measures, and Beneish's M-Score, which is consistent with the FSD Score capturing some of the underlying forces measured by those tools. We also corroborate this analysis by investigating the FSD Scores of firms reporting annual

[5] Our claim that there is not likely to be an ex ante relation with underlying firm characteristics or business models does not imply the absence of a spurious correlation ex post. For example, because firms with lower profitability may be more likely to manipulate their financial statements, our measure may be spuriously correlated with profitability, despite our claim that it is not theoretically related to a firm's profitability ex ante. Unfortunately, like any other measure that bears resemblance to an exogenous instrumental variable, this lack of correlation cannot be tested (Wooldridge 2010).

[6] Please refer to Sect. 3.4 and Appendix 2 for further detail.

income near zero in the spirit of Burgstahler and Dichev (1997). We find that firms just below zero have significantly lower FSD Scores than those just above zero, where the latter set of firms are more likely to be managing their earnings.

We expand our validation of the FSD Score's ability to reflect financial statement errors by conducting an experiment. Specifically, we identify a sample of firms that restated their financial statements and compare the FSD Score for the restated and unrestated numbers. This test provides a novel empirical setting to examine the usefulness of the FSD Score, since we compare the same firm-year to itself, thus keeping all else equal (e.g., economic conditions, firm performance, etc.) except for the reported numbers. We show that the restated numbers have significantly lower divergence (lower FSD Score) from Benford's Law than the same firm-year's unrestated numbers. These results provide strong evidence that divergence from Benford's Law is a useful tool for detecting errors.

Next, we explore the informational implications of divergence from Benford's Law by examining the relation between the level of conformity to the law and earnings persistence. If a higher FSD Score captures a higher degree of financial statement error, it is likely that current earnings are less likely to explain future earnings for such firms (Richardson et al. 2005). Li (2008) provides support from a qualitative disclosure perspective for this argument by showing a negative relation between financial report readability and earnings persistence. Consistent with those studies, we find that the FSD Score is negatively related to earnings persistence.

Given that our findings suggest that divergence from Benford's Law may be a useful tool for assessing financial statement errors, a natural question that arises is whether Benford's Law is predictive of material misstatements.[7] We show that, while the contemporaneous FSD Score negatively predicts material misstatements, lagged versions positively predict material misstatements, which suggests that the FSD Score can be used as a leading indicator of material misstatements. The rationale is that, compared to firms that did not receive AAERs, firms receiving AAERs have higher FSD Scores one and two years *before* the period in which these firms were identified by the Securities and Exchange Commission as having materially misstated their financial statements.[8] The decrease in the FSD Score leading up to the AAER period is consistent with the argument that the SEC only pursues firms that engage in the most egregious misstatements and only once those firms run out of room to manipulate their accounting numbers (Dechow et al. 2010, 2011). Taken together, our results suggest that some firms engage in activities that allow their financial statement errors to remain undetected by the SEC, yet such activities leave a trace of the errors in features of the distributional properties of financial statement numbers.

The remainder of the paper proceeds as follows. Section 2 discusses the paper's motivation and contribution. Section 3 describes the foundations of Benford's Law and provides intuition for its use in accounting. Section 4 discusses our sample and

---

7 We follow Dechow et al. (2011) and use the term "material misstatement" to refer to SEC AAERs, which the SEC itself refers to as "alleged fraud."

8 This pattern is consistent with Dechow et al. (2011), who show an increase in abnormal accruals and a higher probability of manipulation in the years leading up to a material misstatement.

presents descriptive statistics. Section 5 establishes financial statement conformity and provides descriptive evidence on the relation between Benford's Law and commonly used measures of accounting quality. Section 6 examines the relation between Benford's Law and ex post measures of accounting errors, and Sect. 7 concludes.

2 Motivation and contribution

The level of errors in financial statement data has a first-order impact in capital markets (Bushman and Smith 2003). Literature in accounting, finance, and economics has highlighted the importance of financial statements for efficient resource allocation, financial development, employment contracts, debt contracts, cost of capital, and efficiency of equity and debt market prices (e.g., Rajan and Zingales 1998; Rajan and Zingales 2003a, b; La Porta et al. 2000; Duffie and Lando 2001; Francis et al. 2004; Francis et al. 2005). Consequently, prior research in accounting and finance has spent significant effort constructing and evaluating measures of accounting quality (e.g., Jones 1991; Beneish 1999; Dechow and Dichev 2002).

However, prior literature also outlines the limitations of existing measures (e.g., Dechow et al. 2010). We contribute to this literature by implementing a measure that overcomes many of these limitations. First, the FSD Score does not require time-series or cross-sectional data to estimate and does not model the error as a residual from a prediction model. Estimating residuals in time-series or cross-sectional models (e.g., the Jones model or Dechow and Dichev 2002) assumes that the estimated coefficients are identical over time or in the cross-section. Therefore, any unobserved change in those coefficients caused by underlying firm changes will also change the estimated financial statement error. As such, these estimation techniques may bias inferences since the measures will inherently be correlated with the underlying economic reasons that caused the estimated model to deviate in the time series or cross section.

Second, based on its theoretical derivation, the measure is unlikely to have an ex ante relation with underlying firm characteristics or business models since those characteristics or models do not theoretically cause firms to have financial statement items that start with 1, 2, or any other digit. For example, theoretically, a loss firm is as likely as a profitable firm to have a revenue realization that starts with 1. It may, however, be the case that loss firms are more likely to have errors—which is exactly what the measure aims to capture. If, on the other hand, a loss firm does not have errors, there will not be a deviation. This aspect of the measure is a significant advantage in that, unlike accruals measures, a deviation is not caused by firm characteristics or business models. That is, the levels or changes in operating performance are not expected to change the distribution of the first digits as long as financial statements reflect these changes or levels accurately. Correlations with firm characteristics or business models are a major limitation of the accruals-based models in that that they inherently depend on firm performance (Dechow et al. 2010; Owens et al. 2013).

Third, the measure does not require forward-looking information. Using forward-looking information, such as future realizations of cash flows (e.g., Dechow and Dichev 2002), reduces the usefulness of certain measures in settings where relying on such information is infeasible. For example, it is not possible to use these measures, as originally developed, for trading strategies that rely on timely identification of errors. While using lagged values of these measures can give significant insights into certain questions (such as identifying risk factors), they cannot answer questions related to the information content of disclosures. Using these measures with perfect foresight is also a challenge because, in addition to facing look-ahead bias, if the realization of forward-looking information is correlated with current information, then their use may create bias in inferences.

Fourth, the measure does not require returns or price information. This requirement limits the usefulness of other measures and creates selection bias that may be acute in certain settings. Fifth, it does not require identifying managerial incentives to manipulate earnings like other measures (Beneish 1999; Dechow and Skinner 2000). Identifying managerial incentives ex ante to model errors limits the usefulness of these measures as they assume knowledge of the incentives. Sixth, certain measures, such as Beneish's M-Score, are constructed as a linear combination of firm-level performance variables, such as gross margin and sales growth. While these measures are very useful in many settings, they are, by construction, correlated with firm performance, making it difficult to draw conclusions about errors that are separate from firm performance. Seventh, the measure is scale independent and thus fits to every currency or size. Eighth, it is available to essentially every firm with accounting information, even private companies where such information exists.

We are not the first researchers to use Benford's Law as an error detection tool. The idea that Benford's Law could be used to detect errors in economic data was first suggested by Varian (1972) with relation to economic forecasts. More recently, Michalski and Stoltz (2013) showed that this method can be used to detect errors in macroeconomic data. Carslaw (1988) used a variant of Benford's Law to argue that firms in New Zealand whose earnings did not conform to the law were rounding up their earnings numbers. While Thomas (1989) showed similar results for US firms, he further found that the relation inverts for loss firms by demonstrating a greater (lower) than expected frequency of 9's (0's) for such firms. Since Carslaw (1988) and Thomas (1989) are interested in showing that pooled earnings numbers are rounded to a reference point, they focus strictly on the distribution of the second digit of the distribution of earnings and do not make firm-year inferences.[9] The advancement in the use and development of Benford's Law in accounting, and particularly in tax settings, can be found in inquiries by Mark Nigrini and his coauthors. His work has largely focused on internal transactional data from individual financial accounts and personal income tax return data. For example, Nigrini (1996) uses Benford's Law to examine items such as the interest received and interest paid on individual tax returns and finds a higher (lower) than expected frequency of 1's (9's) on interest received (paid). Nigrini and Miller (2009) provide a guide to auditors for how to use Benford's Law to detect errors in transactional

[9] Strategic rounding has also been documented by Grundfest and Malenko (2009).

data, and Nigrini (2012) demonstrates how Benford's Law can be used to assess errors within the accounts receivables of a firm when one has access to invoice-level data. Relatedly, Durtschi et al. (2004) provide a practitioner's guide for auditors on potential uses of Benford's Law to uncover fraud in transactions from individual, internal financial accounts.[10]

Considering the extant literature on Benford's Law, we are unaware of any large-scale application of it to detect errors in the firm-year data found in external corporate financial reports. The literature has largely restricted itself to the auditing of internal transactional data from individual accounts, tax returns, or deviations of one account across several firms (such as earnings per share). Distinct from prior literature, we employ a measure of annual financial statement conformity to Benford's Law on a firm-year basis for the composite distribution of the leading digits from all numbers contained in a firm's annual financial statements. Our measure can be created solely using publicly available information, making it available to anyone interested in analyzing the level of errors in firm-year financial data.

3 Foundations of Benford's Law

3.1 Background

Benford's Law is a mathematical property discovered in 1881 by astronomer Simon Newcomb, who noticed that the earlier pages in books of logarithms were more worn than the latter pages. He inferred from this observation that scientists looked up smaller digits more often than larger digits and determined that the probability that a number has a first digit, d, is:

$$\mathrm{P}(\text{the first digit is d}) = \mathrm{Log}_{10}(\mathrm{d}+1) - \mathrm{Log}_{10}(\mathrm{d}), \quad \text{where } \mathrm{d} = 1, 2, \ldots, 9.$$

This equation gives us the theoretical distribution of what is now commonly referred to as Benford's Law, or the expected frequency of the first digits 1 through 9. And the distribution resulting from this equation is:

1	2	3	4	5	6	7	8	9
0.301	0.1761	0.1249	0.0969	0.0792	0.0669	0.058	0.0512	0.0458

In 1938, physicist Frank Benford tested Newcomb's discovery on a variety of datasets, including the surface areas of rivers, molecular weights, death rates, and the numbers contained in an issue of *Reader's Digest* and found that the law held in each

[10] Benford's Law has also been employed in auditing software, such as ACL. However, similar to prior research, its use has been limited to internal transactional data on a digit-by-digit (not distributional) basis. To our knowledge, prior to this paper, no commercial auditing software computes the conformity of the entire distribution of first digits, nor assesses firm-year conformity from external corporate financial reports.

dataset (Benford 1938). Some years later, Hill (1995) provided a formal derivation of Benford's Law. Hill's theorem states that, if distributions are selected at random and random samples are then taken from each of these distributions, the first digits of the combined mixture distribution will converge to the logarithmic or Benford distribution.

For a distribution to deviate from Benford's Law, certain types of errors must be introduced. For example, evidence suggests that stock indices' returns conform to Benford's Law (Ley 1996), which allows us to compare the law with the empirical distribution of the first digits from the monthly returns of the *Fairfield Sentry Fund*, a fund-of-funds that invested solely with Bernie Madoff, during the 215 months in which it reported returns (Blodget 2008):

1	2	3	4	5	6	7	8	9
0.396	0.142	0.104	0.071	0.075	0.066	0.061	0.066	0.019

One would expect unaltered returns to conform to Benford's Law, but this distribution differs significantly from the theoretical distribution above, indicating that non-zero mean errors were added to the returns data.

3.2 Measuring conformity and deviation from Benford's Law

Measuring whether a dataset conforms to Benford's Law has been the subject of some debate in the field of mathematics (Pike 2008; Morrow 2010). Test statistics can be strongly influenced by the pool of digits used, with some statistics requiring near-perfect adherence to the distribution as the pool becomes large (Nigrini 2012). We employ two statistics when measuring conformity to Benford's Law—the Kolmogorov–Smirnov (KS) statistic and the Mean Absolute Deviation (MAD) statistic. The KS statistic uses the maximum deviation from Benford's distribution, determined by the cumulative difference between the empirical distribution of the digits from 1 to 9 and the theoretical distribution. (See Appendix 1 for the distribution and calculation methods.) This statistic is useful for firm-level examinations of conformity to Benford's distribution since there exists a critical value against which to test, that is, the critical value at the 5 % level $= 1.36/\sqrt{P}$, where P is the total number, or pool, of digits used.[11]

The KS statistic becomes less useful as P increases, however. To establish (fail to reject) the null hypothesis of distributional conformity at the 5 % level, the statistic requires near perfect conformity of the underlying empirical distribution to Benford's Law for large pools of digits (Nigrini 2012). As a result, the KS statistic tends toward over-rejection as the pool of digits increases. The MAD statistic, on the other hand, does not take P into account. The MAD statistic is calculated as the

[11] Another method to examine conformity relies on the expected distribution of the first two digits (from 10 to 99) of a number (Nigrini 2012). We cannot employ the first two digits in our setting because the number of buckets required to generate the distribution is 90 (i.e., leading two digits, 10–99) instead of 9 (i.e., leading first digits 1–9).

sum of the absolute difference between the empirical frequency of each digit, from 1 to 9, and the theoretical frequency found in Benford's Law, divided by the number of leading digits used. The scale invariance of the MAD statistic makes it useful when examining large pools of digits, as well as when comparing financial statements across firms and through time, since the number of line items in an annual report can vary across industries and through time. Consequently, we use the FSD Score based on the KS statistic only in our descriptive tests when we examine the number of individual firm-years that conform to Benford's distribution, that is, where we require a critical value to assess conformity. In all other tests throughout the paper, we rely exclusively on the FSD Score based on the MAD statistic to assess the shift in the empirical distribution.[12]

3.3 Theoretical underpinnings of the FSD score

Appendix 2 details the theoretical underpinnings of the FSD Score. In summary, the intuition behind why empirical data follow Benford's Law can be distilled into two mathematical facts. The first fact relies on using a mathematical approach to determine the first digit of any number N, which is to take its base 10 log and find the fraction behind the integer (i.e., the remainder or mantissa). If the fraction is between 0 and 0.301, the original number N will start with one, if the fraction is between 0.301 and 0.477 (interval of 0.176), the number N will start with 2, and so forth. Hence, the intervals between the fractions after the decimal point of the log number that determine its first digit are the same as the probabilities defined by Benford's Law. The second fact is that, if the probability distribution function of the log of the original number N is smooth and symmetric, the probability that a number will be in the interval between n and $n + 0.301$, where n is any integer in this logarithmic distribution, is 30.1 %. Similarly, the probability that a number will be in the intervals between $n + 0.301$ and $n + 0.477$ is 17.6 %, and so forth. Because distributions in nature tend to be smooth and symmetric due to the Central Limit Theorem, datasets tend to follow Benford's Law (Pimbley 2014). For a distribution that generally follows Benford's Law to diverge from the law, certain types of errors have to be introduced to the data in a way that makes the distribution of the base 10 log less smooth or less symmetric.

3.4 Numerical and simulation analyses

3.4.1 Numerical analysis

Since accounting data are a series of estimations of the true cash flow realizations of the underlying items (for example, cash flows from sales, cash flows from payments

[12] While two other statistics, the Z-statistic and the Chi square statistic, were widely used in the early stages of the forensic accounting literature, in this area researchers have progressed to using the MAD statistic (Cleary and Thibodeau 2005; Nigrini 2012). The main deficiency of using the Z-statistic to examine Benford's Law is that it examines conformity of only a single digit at a time, rather than the composite distribution of digits. The main deficiencies of using the Chi-square statistic is that, unlike the MAD statistic, it assumes observational independence and, similar to the KS statistic, is sensitive to the pool of digits used.

to employees, etc.), the resulting distribution of the mixture of these cash flow realizations may fulfill the conditions of Hill's theorem and hence follow Benford's Law. In Appendix 3, we show under certain assumptions that this is the case with accounting data using a stylized numerical model. We also show that, if the accounting estimates of the true cash flow realizations are without error, the distribution of the accounting estimates (the financial statements) will follow Benford's Law exactly. While we cannot prove or empirically demonstrate that the actual cash flow realizations of accounting data will follow Benford's Law (as they are unobservable), Fig. 1 (Appendix 3) reveals that the actual estimates of these realizations (the accounting line items), which include errors and manipulations, follow Benford's Law for the whole sample and for the typical firm, and these distributions in the log scale are symmetric and smooth (and near normal).

We then numerically characterize the types of errors in accounting data that are likely to create deviations from Benford's Law. For brevity, we summarize the results from the numerical analysis here. In sum, we show that, under certain parameters, the FSD Score is increasing with the size of the error. However, not all errors create deviations from Benford's Law; the error needs to be applied in different rates to different items in the distribution. That is, mean-zero errors will not create deviations from Benford's Law, and neither will an error that is constant across all items. Empirically, it is unlikely that the errors will be identical in all line items in the financial statements. If we introduce non-zero mean errors to some of the underlying distributions in the mixture distribution (i.e., errors to some of the line items in the financial statements) or errors of different size to different items, then the larger the error, the larger is the deviation from Benford's Law. For example, overestimating revenue, underestimating expenses, meet-or-beat behavior, or a combination of these are likely to introduce deviation from the law. More specifically, the analysis indicates that overestimating revenue by itself (Case 3A) or together with the associated cost of goods sold (Case 3B) will create deviations from Benford's Law. Furthermore, an error that is correlated with the size of the item (Case 3C) will create deviations in the financial statements. The reason is that introducing an error to the underlying distributions in the mixture creates asymmetries and lack of smoothness in the mixture distribution. This, in turn, creates measurable deviations from Benford's Law. To provide intuition using real-world data, we show in Appendix 4 that when we introduce errors into observable realizations of equity prices, the distribution of market values of equity begin to deviate from Benford's Law as the errors increase. The advantage of this simulation is that, unlike cash flows realizations, stock price realizations are observable, so we can compare the realized distribution to the distribution with error.

3.4.2 Simulation analysis

To further demonstrate how errors could alter conformity to Benford's Law, we run a simple simulation that involves changing the value of a single line item in a firm's income statement and calculate how that change affected the financial statements overall. Because we need a firm that is unlikely to have a manipulated financial statement, we choose to manipulate sales for Alcoa's 2011 financial statements, which is a firm that generally, but not perfectly, conforms to Benford's Law. We manipulate revenues since revenue is an item that managers may be tempted to change to mask poor performance and is interconnected with many other financial statement items. As a result of the sales manipulation, a firm likely needs to adjust cost of goods sold and tax expense accordingly. Therefore, we add three journal entries to the original numbers:

1. Increase accounts receivables	Increase revenue
2. Increase cost of goods sold	Decrease inventory
3. Increase tax expense	Increase tax payable

These three journal entries affect more than 30 line items in Alcoa's financial reports (see Appendix 5). We then re-measure the FSD Score based on the manipulation and the changes the manipulation induced in the financial statements. The results of this simulation, when run 1,000 times, show that the random revenue manipulation increased the FSD Score 95 % of the time. The evidence from the simple simulation suggests that revenue manipulation in firms that conform to Benford's Law is likely to result in an increase in the deviation from Benford's Law. These results support the implications of our numerical example in the prior section.

4 Sample selection, variable measurement, and descriptive statistics

4.1 Sample selection and variable measurement

Our sample consists of all annual financial statement data from Compustat for the period 2001–2011. For simplicity and objectivity, we use all Compustat variables that appear in the balance sheet, income statement, and statement of cash flow to calculate the FSD Score.[13] For variables reported with an absolute value of less than 1, we take the first non-zero digit. Variables with missing values are ignored. We remove any firm-years from the sample where the total number of line items used to calculate the FSD Score for a given firm-year is less than 100.[14] We also remove firms with negative total assets. All non-indicator control variables in the total sample of 43,332 firm-years are then winsorized at the 1 and 99 % levels to eliminate the influence of outliers. (See Appendix 6 for further details, as well as for the definitions of the control variables.)

[13] We do, however, exclude data items provided by Compustat that do not appear on firms' financial statements, e.g., price data. Furthermore, while we would prefer to use the Edgar 10-K filing itself to overcome possible Compustat shortcomings (e.g., missing variables, modified definitions, etc.), extracting the current year's financial statements from a given 10-K presents technological obstacles that make automated extraction infeasible as well as susceptible to its own biases.

[14] The rationale for doing so is to ensure we do not mechanically create measurement error. Including firm-years with fewer than 100 line items does not alter our results.

As previously discussed, the primary measure we use throughout the paper to assess the conformity of the empirical distribution of annual financial statements to Benford's theoretical distribution is the FSD Score based on the MAD statistic, as it is insensitive to the size of the pool of first digits used (i.e., the number of financial statement line items). While the FSD Score based on the KS statistic also tests conformity to the law and, unlike the FSD Score based on the MAD statistic, has established critical values against which to test, it becomes unreliable as the pool of digits increases. We therefore only rely on the FSD Score based on the KS statistic when gauging the conformity of individual firm-years.

We use several proxies for accruals-based earnings management and earnings manipulation. For accruals-based earnings management, we calculate the 5-year moving standard deviation of the Dechow-Dichev residual (STD_DD_RESID) from Dechow and Dichev (2002), as suggested by Francis et al. (2005); the absolute value of the accruals quality residual (ABS_JONES_RESID) from the modified Jones model (Jones 1991), as suggested by Kothari et al. (2005); the absolute value of working capital accruals (ABS_WCACC); and the absolute value of working capital accruals (ABS_RSST), as defined by Richardson et al. (2005). For earnings manipulation, we calculate the M-Score following Beneish (1999) and create an indicator variable (MANIPULATOR) equal to 1 if the M-Score is greater than -1.78, indicating that a firm may be manipulating its earnings. We also calculate the F_SCORE, the scaled probability of earnings management or a misstatement for a firm-year based on firm financial characteristics, following Dechow et al. (2011).

As for other variables of interest for our tests, RESTATED_NUMS is an indicator variable assigned to all firms that have both restated and originally reported numbers in a year available through Compustat and, for the sake of materiality, at least 10 restated variables available in that year.[15] RESTATED_NUMS is equal to 1 if the reported numbers are restated and zero if the numbers are what was originally reported. Following Shumway's (2001) hazard model in a logit regression setting, AAER is an indicator variable equal to 1 for the first year in which a firm was identified by the SEC as having materially misstated its financial statements.

4.2 Descriptive statistics

Table 1 Panel A provides descriptive statistics for the full sample of firms from 2001 to 2011. The FSD Score's mean is 0.03 with a standard deviation of 0.009. Table 1 Panel B presents Spearman correlations above the diagonal and Pearson correlations below the diagonal. In untabulated results, autocorrelations between the contemporaneous FSD Score and prior year's FSD Score is 0.26 for the Pearson correlation and 0.23 for the Spearman correlation. These correlations are significant but also suggest that the measure is not too sticky over time. Table 1 Panel C groups firm-years by the number of line items available to calculate the FSD Score and examines the average FSD Score for firm-years in the top and bottom 1 % of line items available, as well as the average FSD Score for firm-years by tercile. Panel C

[15] Removing the materiality condition does not alter our inferences.

suggests that the FSD Score is decreasing with the number of line items in the financial statements. In Table 1 Panel D, we group firm-years by total assets to calculate FSD Score and examine the average FSD Score for firm-years in the top and bottom 1 % based on firm size, as well as the average FSD Score for firm-years by tercile. Panel D suggests that the FSD Score is decreasing in firm size.

5 Establishing conformity

5.1 Aggregate and firm-year conformity to Benford's Law

Table 2 investigates how the aggregate empirical distribution of numbers reported in financial statements conforms to Benford's Law. That is, the FSD Score is calculated by measuring the frequencies of the first digits from all firm-years in the sample. We begin with an aggregate analysis of the numbers reported in financial statements before proceeding to a firm-year analysis to determine whether the numbers contained in financial statements generally conform to Benford's Law. In the aggregate, the FSD Score is 0.0009, well below 0.006, which can be considered close conformity to the law in very large samples (Nigrini 2012).[16] This result can also be seen graphically in Fig. 2 (Appendix 7). Panels B and C of Table 2 show similar results when examining aggregate financial results by industry based on the Fama–French 17-industry classification and by fiscal year. This table supports the conjecture that the empirical distribution of the frequency of first digits in aggregate financial results conforms to Benford's Law.

Table 3 examines individual firm-year conformity to Benford's Law. Here, the FSD Score based on the KS statistic is used because it enables us to assess whether the financial statements for a given firm-year adhere to the law. Of the 43,332 firm-years in our sample, 37,104, or 86 %, conform to the law at the 5 % level or better, as shown in Panel A.[17] Figure 3 (Appendix 7) provides examples of the empirical distributions for two firm-years, one that conforms to Benford's distribution at the 5 % level (Verizon Communications 2001) and one that does not conform (Sprint Nextel 2001). While there are some kinks in Verizon's distribution, the overall divergence from Benford's distribution is visually apparent for Sprint Nextel, which experienced a restatement. We examine the firm-year FSD Scores for individual financial statements in Panel B of Table 3, which shows that 91, 79, and 98 % of firm-year balance sheets, income statements, and cash flow statements conform, respectively. Panel C sorts firms by industry and shows a minimum conformity of 82 % for all firms in a given industry and a maximum conformity of 91 %. Panel D

[16] As noted previously, unlike the FSD Score based on the KS statistic, the FSD Score based on the MAD statistic has no critical value against which to test. However, based on simulation analysis, Nigrini (2012) suggests, when using the MAD statistic, a value of 0.006 or lower can be considered as close conformity to Benford's Law.

[17] While we do not claim that all 16 % of the firms that deviate from Benford's Law engage in material misreporting, this estimate is consistent with Dyck, Morse, and Zingales (2013), who report that the probability of a firm committing fraud is 14.5 % a year.

sorts firms by fiscal year and shows all years exhibit between 85 and 87 % conformity. Overall, Table 3 supports our conjecture that a significant majority of firm-year empirical distributions conform to Benford's Law.[18]

5.2 Conformity to Benford's Law: financial statement partitions and firm characteristics

We next examine the relation between Benford's Law and partitions of firms' financial statements. Table 4, Panel A individually measures the aggregate FSD Score for the balance sheet, income statement, and statement of cash flows for all firm-years in our sample. To measure the aggregate FSD Score, we calculate the FSD Score using all available variables that make up each financial statement for all firm-years. We find that, on average, the aggregate FSD Score for the income statement diverges the most from Benford's Law, which suggests that the income statement contains more errors as compared to the balance sheet and statement of cash flows.

Panel B of Table 4 provides an additional financial statement partition by financial statement subcategories, that is, assets versus liabilities, income versus expenses. As in Panel A, we compute the aggregate FSD Score; however, we now categorize line items based on where they fall in the appropriate subsections of the financial statements. The evidence reveals that, on average, equity and liability accounts have higher FSD Scores than asset accounts, suggesting that equity and liability accounts contain a relatively higher level of errors. Furthermore, income accounts, on average, have higher FSD Scores than the expense accounts, suggesting that income accounts contain a relatively higher level of errors.

Having established the types of financial statements and financial statement accounts that exhibit greater conformity, we enhance our understanding of the types of firms that are more likely to conform to Benford's Law in Table 5. To do so, we divide firm-years into terciles based on their FSD Scores and calculate the means of several firm characteristics based on this segmentation. We find that firms with high FSD Scores tend to be smaller, younger, more volatile, and growing.

5.3 The relation between Benford's Law and existing measures of reporting quality

To shed light on the types of firm behavior associated with the FSD Score, we examine the relation between the FSD Score and proxies for accruals-based earnings management and earnings manipulation in Table 6. Panel A presents univariate analysis by dividing firm-years into terciles based on the FSD Score and calculating the means of these proxies for each tercile. In examining the accrual quality measures, firms with higher FSD Scores tend to have more working capital accruals, more discretionary accruals and higher values of the Dechow–Dichev

[18] These results further imply that the pre-errors financial statements follow Benford's Law because, if most financial statements follow Benford's Law after-errors, it is likely that firms follow Benford's Law before errors were introduced.

measure, and are more likely to be a manipulator according to Beneish's M-Score. The result pertaining to the F_Score suggests that firms with higher FSD Scores are less likely to be accused by the SEC of making material misstatements, which we explore in detail in Sect. 6. Finally, inspecting the earnings quality measures, firms with higher FSD Scores tend to have less persistent earnings and are more likely to have a loss. Panel B presents multivariate analysis on the relation between the FSD Score and proxies for accruals-based earnings management and earnings manipulation. The results of the multivariate regression generally mirror the evidence provided in Panel A.[19]

6 Benford's Law and ex post measures of accounting errors

6.1 Univariate evidence

In our final set of analyses, to understand whether Benford's Law captures firms with a higher propensity for errors in their accounting results ex post, we examine the relation between the FSD Score and several ex post measures of earning management. In Table 7, we conduct univariate analyses on restated data, loss firms, firms that just beat the zero earnings threshold, and SEC Accounting and Auditing Enforcement Releases (AAERs). We find that the FSD Score is lower after firms restate their misstated data, higher for loss firms, higher for firms that just beat the zero-earnings benchmark, and lower for firms that receive an AAER. As alluded to previously, we explore the latter finding regarding AAERs in further detail in Sect. 6.2.3.

6.2 Multivariate evidence

6.2.1 Restated data

Firms that misstate their financial results by manipulating select accounts may report numbers with first digits that are not, in expectation, driven from the same interactions of random distributions that create conformity to Benford's Law. Given the nature of double-entry bookkeeping, this lack of consistency should trickle through several of the financial statement line items. For example, a firm that is trying to increase earnings for the current period may underreport depreciation. This manipulation will affect net property, plant, and equipment, accumulated depreciation, depreciation expense, operating income, taxable income, and net income. After a restatement, however, financial reports should more closely represent the "true" nature of the distribution of leading digits found within the financial statements. As such, we conjecture that the empirical distribution of restated

[19] The modified Jones model becomes statistically insignificant in explaining FSD only after including ABS_WCACC and ABS_RSST as they capture similar constructs.

financial statements will more closely follow Benford's distribution than the empirical distribution of misstated financial reports.

To examine our conjecture, we investigate all firm-years in Compustat from 2001 to 2011 where both misstated and restated financial results are available (in Compustat, datafmt = STD for original and datafmt = SUMM_STD for restated). We then create an indicator variable, RESTATED_NUMS, which is equal to 1 for results that have been restated and 0 for the originally reported results and regress it on the FSD Score. Since the regression compares the firm to itself, we do not include additional firm control variables in this specification.

Table 8 presents the results of our test. The coefficient on RESTATED_NUMS in Column (1) is −0.0009 and is statistically significant at the 1 % level. To ensure that our measure of conformity isn't merely a proxy for existing measures of accounting discretion, in Column (2) we control for accruals levels, accruals-based earnings management, and earnings manipulation. When adding these additional measures, we find similar results. Consequently, we find that the FSD Score is lower for restated financial results, which implies that the empirical distribution of restated financials more closely conforms to Benford's Law. In terms of economic significance, a 0.001 decrease in the FSD Score represents a 3.5 % reduction in the mean value of the FSD Score.

6.2.2 *Earnings persistence*

If deviations from Benford's Law reflect a decrease in the quality of reported financial results, greater divergence from the law may signal lower earnings persistence. The idea is based on the notion that it is less likely that current earnings will be as informative about future earnings in firms with lower accounting quality (Richardson et al. 2005). Li (2008) provides support for this argument from a qualitative disclosure perspective by showing a negative relation between low financial report readability and earnings persistence. As such, we conjecture that current earnings will exhibit less persistence for firms with greater divergence from Benford's Law.

To test our conjecture, we regress the interaction between net income and the FSD Score in year t on net income in year $t + 1$. In addition to the controls used in prior tests, following Li (2008), we control for sales growth (SALES_GROWTH), whether the firm pays a dividend (DIV), the log of the market value of equity (SIZE), growth (MTB), special items (SI), survivorship (AGE), return volatility (RET_VOL), and the volatility of net earnings (NI_VOL). The results of this test are presented in Table 9. The coefficient of −2.6277 on the interaction between the FSD Score and net income in year t is significant at the 1 % level, which suggests that divergence from Benford's Law reflects the informational quality of financial disclosures.[20]

[20] The addition of an interaction term to the regression between the FSD score, size, and net income does not alter our results.

6.2.3 Accounting and Auditing Enforcement Releases

Lastly, in response to the recent debate and renewed efforts by the SEC surrounding accounting fraud and detection, we examine whether the FSD Score predicts material misstatements. Based on the criticism leveled at the SEC after disbanding the accounting fraud group, and the severe underfunding of the SEC's unit to pursue detection and enforcement (Gallu 2013; McKenna 2012, 2013; Whalen et al. 2013), our expectations regarding the FSD Score's predictive ability of material misstatements depend on two key factors: the SEC's own ability to detect material misstatements and the types of firms that make these misstatements.[21] If the SEC does indeed detect and prosecute all firms that make material misstatements, then a positive coefficient is expected. However, prior research suggests that SEC AAERs reflect only firms that experience significant declines in their ability to hide the misstatements (Dechow et al. 2011), which may result in a nonpositive relation as FSD Scores decrease.

In our last set of analyses, we follow Shumway (2001) and use the logit regression equivalent to a hazard model to examine whether the FSD Score predicts material misstatements. Material misstatements are proxied for by the variable AAER, which is an indicator variable equal to 1 for the initial year in which a firm was identified by the SEC as having materially misstated its financial statements through an Accounting and Auditing Enforcement Release. As before, the regressions control for accruals levels, accruals-based earnings management, earnings manipulation, and firm characteristics. In addition to the firm-characteristic control variables used in prior tables, we also control for change in cash sales (CH_CS), change in ROA (CH_ROA), soft assets (SOFT_ASSETS), and whether the firm issued debt or equity (ISSUE).

There are four groups of firms to consider in our analysis: (a) those that do not manipulate, (b) those that manipulate but do not get caught, (c) those that manipulate, are caught, but are not prosecuted, and (d) those that manipulate, are caught, and are prosecuted. The latter three groups will have higher FSD Scores than the first because they manipulate. Prior literature suggests that only the most egregious manipulators are prosecuted and only after those firms can no longer sustain the manipulation (Dechow et al. 2010, 2011). Contingent on having the ability to manipulate, the FSD Scores of AAER firms should be positively associated with material misstatements. Inconsistent with our conjectures, Column (1) of Table 10 shows that the coefficient on the contemporaneous FSD Score is negative, which may imply that, as AAER firms run out of room to manipulate, their FSD Scores decrease.

To shed light on this result, Fig. 4 (Appendix 7) shows that, compared to firms that did not receive an AAER, firms that received an AAER have higher FSD Scores in the years *before* the misstatement period. However, there is a striking difference in the trends of these two types of firms. While firms that were not prosecuted by the

[21] According to Dechow et al. (2010): "… the SEC has limited resources that constrain its ability to detect and prosecute misstatements. Thus, the SEC may not pursue cases that involve ambiguity and that it does not expect to win. As a result, the AAER sample is likely to contain the most egregious misstatements and exclude firms that are aggressive but manage earnings within GAAP."

SEC can maintain their FSD Scores at fairly constant levels, firms that were prosecuted appear to have a significant decline in their FSD Scores in the years before the misstatement period. This evidence is consistent with the assertion that AAER firms are precisely those firms that are unable to sustain the financial statement manipulation and suggests that FSD Scores in the years leading up to the SEC recognizing material misstatements may predict material misstatements.

Consistent with our conjectures, Columns (2) and (3) reveal that the FSD Score from both 1 and 2 years before the misstatement period positively predict material misstatements. Our prediction results collectively suggest that AAER firms are prosecuted for making material misstatements only once they run out of room to manipulate their numbers, forcing them to report numbers that more accurately reflect their underlying business activities and more closely reflect the theoretical distribution posited by Benford's Law. In addition, consistent with critics' views that the SEC should ramp up its efforts to detect accounting fraud, these results provide evidence that firms may be able to evade detection of financial statements errors, but their manipulations will still leave traces in the distributional properties of their financial statements in the form of deviations from Benford's Law.

7 Summary and conclusion

Building on a method used in a variety of disciplines, we propose that firm stakeholders may find a firm-year measure of financial reporting errors to be a useful tool to augment existing techniques to assess accounting data quality. Our measure, the FSD Score, relies on the divergence from Benford's Law, which states that the first digits of all numbers in a dataset containing numbers of varying magnitude will follow a particular theoretical and mathematically derived distribution where the leading digits 1 through 9 appear with decreasing frequency. This measure has significant advantages over alternative measures of accounting quality currently used in the literature. For example, it does not require time-series, cross-sectional, or forward-looking information, is available for essentially every firm with accounting information, and is uncorrelated ex ante with firms' operating performance and business models.

After providing intuition for the theory behind the measure, we use numerical methods to demonstrate that financial statements without error are distributed according to Benford's Law. We then provide several scenarios to demonstrate the types of financial statement errors that are likely to create divergence from the law. For example, overestimating revenue, underestimating expenses, meet-or-beat behavior, or a combination of these are likely to introduce deviation from the law. To corroborate the results from the numerical analysis, we provide a simple simulation to demonstrate that when accounting numbers are manipulated, there is a high likelihood of an increase in the divergence from the law.

Next, to establish whether the law applies to actual financial statement data, we show that at the aggregate level, financial statement numbers conform to Benford's Law in all industries and years. When assessing the conformity of individual firm-years, we find that roughly 86 % of firm-years conform to the law as well. In examining the financial statements individually, we find that the income statement has the greatest divergence from Benford's Law. In examining the financial statements by account types, we find that equity and liability (in contrast to asset) accounts, as well as income (in contrast to expense) accounts, have the greatest divergence from the law.

Turning to firm characteristics, we find that firms that diverge from Benford's Law tend to be smaller, younger, more volatile, and growing. To shed light on the types of firm behavior associated with the FSD Score, we find that proxies for accruals-based earnings management and earnings manipulation are related to divergence from Benford's Law. However, multivariate empirical analysis indicates that the FSD Score is incremental to these proxies. In addition, firms reporting losses have weaker conformity to the law, and firms that report just above the zero earnings threshold have weaker conformity than firms reporting just below zero.

We conclude by examining the relation between divergence from Benford's Law and several ex post measures of earning management. Our findings suggest that when restatements occur, the restated numbers are significantly closer to Benford's Law relative to the misstated numbers. Furthermore, as firms' financial statements diverge from the law, their earnings persistence decreases. The negative relations between the FSD Score and these ex post measures of earning management support our claim that there exists a relation between the level of divergence from Benford's Law and the informational quality of reported financial results. Finally, we provide evidence that the FSD Score may serve investors, auditors, regulators, and researchers by providing a leading indicator of material misstatements as identified by SEC AAERs.

To our knowledge, this paper is the first to document whether firms' annual financial reports conform to Benford's Law, how firms' reports are likely to exhibit divergence, and the implications for those firms that diverge. In today's environment of increasingly electronic, machine-readable disclosures, where information overload has become the norm, we provide the investment and regulatory communities with an easily implementable, parsimonious approach for assessing errors in financial reports.

Acknowledgments We would like to thank Patty Dechow (the editor), an anonymous referee, Anil Arya, Rob Bloomfield, Qiang Cheng, Ilia Dichev, Dick Dietrich, Peter Easton, Paul Fischer, Ken French, Joseph Gerakos (CFEA discussant), Jon Glover, Trevor Harris, Colleen Honigsberg, Jeff Hoopes, Gur Huberman, Amy Hutton, Bret Johnson, Steve Kachelmeier, Alon Kalay, Bin Ke, Bill Kinney, Alastair Lawrence, Melissa Lewis-Western, Scott Liao, Sarah McVay (FARS discussant), Rick Mergenthaler, Brian Miller, Brian Mittendorf, Suzanne Morsfield, Suresh Nallareddy, Jeff Ng, Craig Nichols, Mark Nigrini, Doron Nissim, Ed Owens, Bugra Ozel, Oded Rozenbaum, Gil Sadka, Richard Sansing, Richard Sloan, Steve Smith, Steve Stubben, Alireza Tahbaz-Salehi, Dan Taylor, Andy Van Buskirk, Kyle Welch, Jenny Zha (TADC discussant), Amir Ziv, conference participants at the 2014 AAA FARS Midyear Meeting, and workshop participants at Columbia University, Baruch College, Dartmouth College, Florida Atlantic University, George Washington University, Georgetown University, the London Trans-Atlantic Doctoral Conference, Nanyang Technological University, Singapore Management University, Syracuse University, UC Berkeley, UCLA, UNC, The University of Texas—Austin, and The University of Utah for their helpful comments and suggestions. We would also like to thank the PCAOB and the SEC for their insights.

Appendix 1

How to calculate conformity to Benford's Law, an empirical example

Assets		Liabilities	
Cash	**1**364	Accounts payable	**1**005
Accounts receivable	**9**31	Short-term loans	**7**80
Inventory	**2**054	Income taxes payable	**3**1
Prepaid expenses	**1**200	Accrued salaries and wages	**3**7
Short-term investments	**3**8	Unearned revenue	**4**05
Total short-term assets	**5**587	Current portion of long-term debt	**2**97
		Total short-term liabilities	**2**555
Long-term investments	**1**674		
Property, plant, and equipment	**4**355	Long-term debt	**6**507
(Less accumulated depreciation)	**2**215	Deferred income tax	**1**89
Intangible assets	**6**08	Other	**5**87
Other	**8**4		
		Total liabilities	**9**838
Total assets	**1**4,523		
		Equity	
		Owner's investment	**1**118
		Retained earnings	**2**732
		Other	**8**35
		Total equity	**4**685
		Total liabilities and equity	**1**4,523

Above is a sample balance sheet. To test its conformity to Benford's Law, take the first digit of each number (in bold) and calculate the frequency of the occurrence of each digit. In this case, there are 28 total numbers and eight appearances of the number 1, so its frequency is 8/28 = 0.2857.

Next, compare the empirical distribution to Benford's theoretical distribution:

Digit	1	2	3	4	5	6	7	8	9
Total occurrences	8	5	3	3	2	2	1	2	2
Empirical distribution	0.2857	0.1786	0.1071	0.1071	0.0714	0.0714	0.0357	0.0714	0.0714
Theoretical distribution	0.3010	0.1761	0.1249	0.0969	0.0792	0.0669	0.0580	0.0512	0.0458

The *Mean Absolute Deviation (MAD) statistic* and the *Kolmogorov–Smirnov (KS) statistic* can be computed to test the conformity of the empirical distribution to Benford's distribution.

1. The KS statistic is calculated as follows:

$$\mathrm{KS} = Max(|\mathrm{AD}_1 - \mathrm{ED}_1|, |(\mathrm{AD}_1 + \mathrm{AD}_2) - (\mathrm{ED}_1 + \mathrm{ED}_2)|, \ldots, |(\mathrm{AD}_1 + \mathrm{AD}_2 + \cdots + \mathrm{AD}_9) - (\mathrm{ED}_1 + \mathrm{ED}_2 + \cdots + \mathrm{ED}_9)|$$

where AD (actual distribution) is the empirical frequency of the number and ED (expected distribution) is the theoretical frequency expected by Benford's distribution.

In this example,

$$\begin{aligned}&Max(|0.2857 - 0.3010|,\ |(0.2857 + 0.1786) - (0.3010 + 0.1761)|, \ldots,\\&(|(0.2857 + 0.1786 + 0.1071 + 0.1071 + 0.0714 + 0.0714 + 0.0357 + 0.0714 + 0.0714)\\&-(0.3010 + 0.1761 + 0.1249 + 0.0969 + 0.0792 + 0.0669 + 0.0580 + 0.0512 + 0.0458)|) = 0.0459\end{aligned}$$

To test conformity to Benford's distribution at the 5 % level based on the KS statistic, the test value is calculated as $1.36/\sqrt{\mathrm{P}}$, where P is the total number, or pool, of first digits used. The test value for the sample balance sheet is $1.36/\sqrt{28} = 0.2570$. Since the calculated KS statistic of 0.0459 is less than the test value, we cannot reject the null hypothesis that the empirical distribution follows Benford's theoretical distribution.

2. The MAD statistic is calculated as follows:

$\mathrm{MAD} = (\sum_{i=1}^{K}|\mathrm{AD\text{-}ED}|)/\mathrm{K}$, where K is the number of leading digits being analyzed.

In this example,

$(|0.2857 - 0.3010| + |0.1786 - 0.1761| + |0.1071 - 0.1249| + |0.1071 - 0.0969| + |0.0714 - 0.0792| + |0.0714 - 0.0669| + |0.0357 - 0.0580| + |0.0714 - 0.0580| + |0.0714 - 0.0458|)/9 = 0.0140.$

Since the denominator in MAD is K, this statistic is insensitive to scale (the pool of digits used, or P). This statistic becomes more useful as the total pool of first digits increases, while the KS statistic become more sensitive as P increases.

Note that there are no determined critical values to test the distribution using MAD.

Appendix 2: Theoretical underpinnings of the FSD Score

There are two mathematical facts that explain the prevalence of Benford's Law in empirical data. First, it can be shown that the mantissa (the fraction behind the decimal point of an integer) of the log 10 of a number is what determines the first digit of that number. If the mantissa is between $\log(d + 1)$ and $\log(d)$, where d is an integer between 1 and 9, then the original number will start with d. Second, since many distributions observed in nature and all of those that are characterized by Hill's (1995) theorem, are smooth and symmetric in the log scale (because of variations of the Central Limit Theorem), the probability of being in a region between $n + \log(d + 1)$ and $n + \log(d)$, where n is any integer in the logarithmic distribution, is exactly $\log(d + 1) - \log(d)$. This is precisely the probability given by Benford's Law. We detail this intuition in the following subsections.

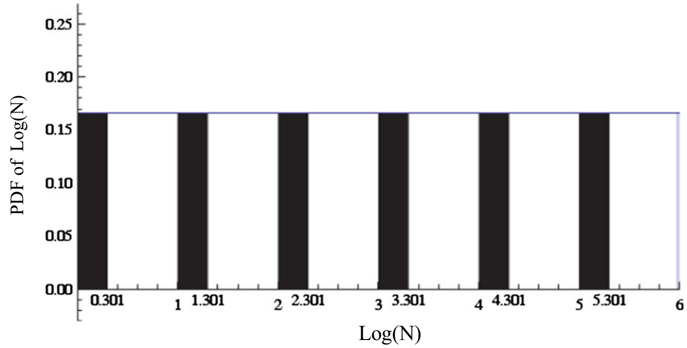
Determining the first digit of a number

The first fact that mathematically explains the prevalence of Benford's Law is that we can obtain the leftmost (or first) digit of a positive number by using the following algorithm (Smith 1997; Pimbley 2014). First, calculate the base 10 log of the number. For example, the base 10 log of 7823.22 is 3.893. Second, isolate the mantissa, that is, the part of the number to the right of the decimal point; in our example, it will be 0.893. Third, raise 10 to the power of the mantissa found in the prior step; in our example, $10^{0.893}$ is 7.81. Fourth, the integer of the number found in the prior step is the first digit of the original number. In our case, the integer of 7.81 is 7, which is indeed the first digit of our original number 7823.22.

This algorithm shows that the first digit of a number can be recovered from the remainder (or mantissa) of its base 10 log. More formally, any number N will start with the digit d (where d is between 1 and 9) if and only if the mantissa of $\log(N)$ is between $\log(d + 1)$ and $\log(d)$. This means that N will start with 1 if the mantissa of the log of N is between $\log(2) = 0.301$ and $\log(1) = 0$. The number N will start with 2 if the mantissa of the log of N is between $\log(3) = 0.477$ and $\log(2) = 0.301$, and so forth. The advantage of this algorithm is that it takes numbers with any length and isolates them to a length of only one digit. Furthermore, as this example shows, the differences of $\log(d + 1) - \log(d)$ for digits 1 through 9, which determine the intervals between the first digits, are exactly the probabilities that a first digit will be d as defined by Benford's Law, which leads us to the second mathematical fact.

Probability distribution functions and the area under the curve: uniform distributions

The second mathematical fact that empirically determines the prevalence of the first digit 1 and the rarity of the first digit 9 is that the area under the curve of a probability density function (PDF) is the probability that a number drawn from this distribution will be in this range. To demonstrate the mechanics of this fact, it is convenient to examine the first digits on the log 10 scale rather than the linear scale. Therefore, we initially consider a uniform distribution between 0 and 6 on the log scale (which implies that the distribution ranges from 1 to 1 million on a linear scale). The PDF of this distribution is PDF(log(N)) = 1/6, and the graphical representation is:



The solid black bars in above figure are the areas under the curve between every integer n in the distribution and $n + \log(2) = n + 0.301$. If N is a random number drawn from this distribution and falls in any of these areas, it will begin with the number 1 in the linear scale. The reason is that, according to the algorithm discussed in the previous section, any number that is between an integer n and $n + 0.301$ in the log scale will start with 1 in the linear scale because its mantissa is between 0 and 0.301.

To obtain the probability that a number from this distribution (uniform in the log scale) will start with the digit 1 in the linear scale, we must find the area under the curve between n and $n + \log(2)$. We can obtain this by taking the integral of the PDF between n and $n + \log(2)$. Thus, the probability that a first digit, d, is 1 can be expressed as:

$$\begin{aligned}\sum_{n=0}^{5}\int_{n}^{n+\log(2)}\frac{1}{6}dN &= 1/6*(0.301-0)+1/6*(1.301-1)+1/6*(2.301-2)\\&\quad+1/6*(3.301-3)+1/6*(4.301-4)+1/6*(5.301-5)\\&=0.301=\mathbf{log(2)}-\mathbf{log(1)}\end{aligned}$$

The same rationale applies for every first digit d where d can equal 1 to 9. That is, if N is distributed uniformly in the log scale, it will follow Benford's Law because t probability of obtaining the first digit d is exactly $\log(d + 1) - \log(d)$, which is Benford's Law. More formally, in the case of our uniform distribution:

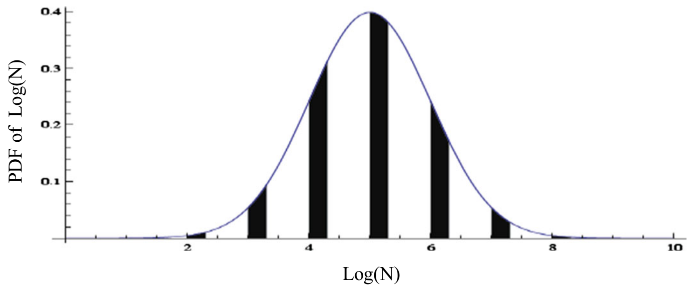
$$\sum_{n=0}^{5}\int_{n+\log(d)}^{n+\log(d+1)}\frac{1}{6}dN = \mathbf{log(d+1)}-\mathbf{log(d)}$$

Probability distribution functions and the area under the curve: normal distributions

While the uniform distribution is useful in explaining the intuition, it is not as useful when applying the intuition to empirical data. Two types of distributions arise naturally in many processes because of variations of the Central Limit Theorem, the normal and log-normal distributions. The intuition above applies in these cases as well. As long as these distributions are spread across a few orders of magnitudes in the log scale (e.g., range between 2 and 4 in the log-scale, which translates to 100–10,000 in the linear scale), they will follow Benford's Law.

To see this clearly, we need to examine a distribution that is distributed normally on the log scale, which means it is log normal in the linear scale. (The distinction between natural log or base 10 log is not crucial here for the shape of the distribution.) Consider a normal distribution with a mean of 5 and standard deviation of 1 in the log scale.

$$\mathrm{PDF}(\mathrm{Log}(N), \mu = 5, \sigma = 1) = \frac{1}{\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-5)^2}{2}}$$



The shaded area in above figure represents all the areas between any integer n and $n + 0.301$. While it is not clear to the naked eye as it was in the case of the uniform distribution above, the area under the curve in all sections between n and $n + 0.301$ is the probability of a number in a linear scale starting with 1. Here, the probability that a first digit is 1 is:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log(1)}^{n+\log(2)} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-5)^2}{2}} dN \cong \mathbf{0.301} = \mathbf{log(2)} - \mathbf{log(1)}$$

Similarly, we can find the probability of any digit for this normal distribution in the following way:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log(d)}^{n+\log(d+1)} \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-5)^2}{2}} dN \cong \mathbf{log(d+1)} - \mathbf{log(d)}.$$

Probability distribution functions and the area under the curve: generic distributions

More generally, for any given probability distribution function, the probability that a first digit begins with d can be found by obtaining the area under the curve for the function specified:

$$\sum_{n=-\infty}^{\infty} \int_{n+\log(d)}^{n+\log(d+1)} PDF(\log(N))dN$$

For a given digit d, if the area under the curve is equal to $\log(d + 1) - \log(d)$, then the probability that the first digit for the numbers drawn from this distribution is d will follow Benford's Law. Stated differently, if a distribution is smooth and symmetric in the log scale over several orders of magnitude, it will follow Benford's Law (Smith 1997; Pimbley 2014). This happens because the area under the curve from $n + \log(d)$ to $n + \log(d + 1)$ is equal to $\log(d + 1) - \log(d)$, which is equal to the probability that a first digit is d under Benford's Law. Since many empirical distributions tend to be smooth and symmetric in the log scale, it is not surprising that first digits are empirically distributed following Benford's Law.

Mean absolute deviation and financial statement deviation

It is not sufficient to examine only a single digit in isolation to detect deviation from Benford's Law (Smith 1997). A natural measure to examine the distance of all leading digits from Benford's Law is the Mean Absolute Deviation (MAD), which takes the mean of the absolute value of the difference between the empirical frequency of each leading digit that appears in the distribution and the theoretical frequency specified by Benford's Law. We can now construct the Financial Statement Deviation (FSD) Score based on the Mean Absolute Deviation (MAD) statistic:

$$\text{FSD Score} = \frac{\sum_{d=1}^{9} ABS[(\overbrace{\sum_{n=-\infty}^{\infty} \int_{n+\log(d)}^{n+\log(d+1)} PDF\,(\log(N))dN}^{\text{The probability that the first digit will be } d \text{ for a given PDF}}) - (\overbrace{Log(d+1) - Log(d)}^{\text{The probability that the first digit will be } d \text{ following Benford's Law}})]}{9}$$

The FSD Scores of the uniform and log-normal scale PDFs above are equal to zero. This occurs because, as shown above, since these distributions are smooth and symmetric, the probability that a number drawn from any of these distributions begins with a digit d is $\log(d + 1) - \log(d)$, which is exactly the probabilities given by Benford's Law. Therefore, for each first digit d, there is no deviation from Benford's Law, which implies that the mean of the absolute deviation, as captured by the FSD Score, is equal to zero.

Appendix 3: A stylized numerical model

To strengthen the intuition regarding the way Benford's Law can be used to detect errors in accounting data, consider the following setting. A manager starts a project at year 1 that has a vector X with K $\{1,2,\ldots K\}$ different random cash flow streams X_k $\{X_1, X_{2\ldots}\ X_K\}$. All cash flow streams will be realized in year 2 and are constructed to be positive (i.e., we take the absolute value of the cash flow streams). X_1 is the random flow of cash from activity 1 (say, cash flow from revenue from activity 1), X_2 is the random flow of cash from activity 2 (say, cash outflow for payment to suppliers), and X_k is the random inflow of cash from activity k. X_K is the last cash flow stream.[22] Assume that the K cash flows are all log normal (base 10) distributed with mean μ_k and standard deviation of σ_k (in the log scale), which implies that $\log(X_k)$ is distributed normal (μ_k, σ_k). For simplicity, we will assume all cash flows and error terms are uncorrelated with each other and we will modify this assumption later in our illustration.

At the end of year 1, the manager needs to report financial statements that include his estimate of the cash flow stream X. This report could be the manager's best estimate, could be strategically manipulated, or could be constrained by correct application of accounting methods; we do not distinguish between these possibilities. The report is a vector Y with K different estimates for each of the K cash flows. To make the calculation tractable, assume that $Y_k = X_k{*}Z_k$, where Z is a vector of the estimation errors for each of the X_k. If $Z_k = 1$, there is no error in the estimation. If $Z_k > 1$, there is over-estimation of the true X_k, and if $Z_k < 1$, there is under-estimation of X_k. The reason for the multiplicative error structure, rather than the more common additive error structure, is that we can now easily recast the example in log scale as $\log(Y_k) = \log(X_k) + \log(Z_k)$, that is, there is an additive error in the log scale, which makes the problem more tractable. Since log(1) is zero, it is clear that if there is no error, $Z_k = 1$, and $\log(Y_k) = \log(X_k)$.

Since we showed above that normal distributions in the log scale follow Benford's Law, adding an error term Z_k that is distributed log normal with a mean $\mu_{\varepsilon k}$ and standard deviation $\sigma_{\varepsilon k}$ does not create deviation from the law. The reason is that the convolution in the log scale of Y (i.e., the distribution of $\log(X_k) + \log(Z_k)$) will be distributed normal ($\mu_{k\ +}\ \mu_{\varepsilon k}$, $\sigma_{k\ +}\ \sigma_{\varepsilon k}$). This distribution will also follow Benford's Law, even if there is a nonzero mean error ($\mu_{\varepsilon k} \neq 0$) or decreased precision ($\sigma_{\varepsilon k} > 0$).

However, the example becomes more interesting when we look at the errors in the report in a specific year (i.e., when we look at the distribution of the cross section of all the X_ks in 1 year). The reason is that, despite the fact that all X_ks in a given year are distributed normally, the mixture distribution of the vector X for that year will not be normally distributed unless the means of the underlying distributions are equal. The distribution of the vector X in the cross section is a mixture distribution, and its density function is given by the following formula:

[22] The example can be constructed to include balance sheet and cash flows statement and include multiple periods.

$$\mathrm{PDF}(X) = \sum\nolimits_{k=1}^{K} W_k * PDF(X_k),$$

where W_k is the weight of each of the individual distributions that comprise the mixture distribution. In our case, since the X_ks are distributed normally in the log scale, the mixture distribution is given by the following expression:

$$\mathrm{PDF}(\log(X)) = \sum_{k=1}^{K} \frac{1}{K}\left(\frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}}\right)$$

The theoretical FSD Score of X (in the cross section) in this case is therefore:

$$\text{FSD Score} = \frac{\sum_{d=1}^{9} \mathrm{ABS}\left[\left(\sum_{n=-\infty}^{\infty} \int\limits_{n+\log(d)}^{n+\log(d+1)} \sum_{k=1}^{K} \frac{1}{K}\left(\frac{1}{\sigma_k\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{2\sigma_k^2}}\right) dX_k\right) - (\log(d+1) - \log(d))\right]}{9}.$$

A mathematically interesting fact about the mixture of normal distributions is that when the means of the distributions are less than two standard deviations apart, the resulting distribution has a single peak, and it looks exactly like a normal distribution (Ray and Lindsay 2005). Therefore, it will follow Benford's Law. More importantly, Hill (1995) provides a proof that mixtures of distributions that do not contain error will follow Benford's Law under certain conditions. However, there is no analytical or empirical way to show that these conditions are met in the context of financial accounting. We do, however, show that the distribution of Y in the log scale appears to be relatively smooth and symmetric (and looks similar to a normal distribution). Figure 1a plots the empirical density function of all numbers from all financial statements from 2001 to 2011 in the log scale, which suggests that the underlying no-error distribution follows Benford's Law as well. Figure 1b shows the distribution in the log scale for a typical firm, Alcoa in 2011.

Solving for a general closed-form solution of how the FSD Score is changing with the error term Z is beyond the scope of this paper and therefore we leave this

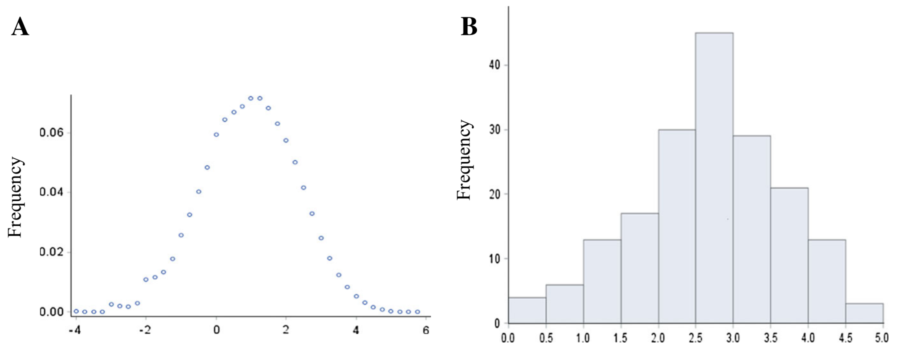


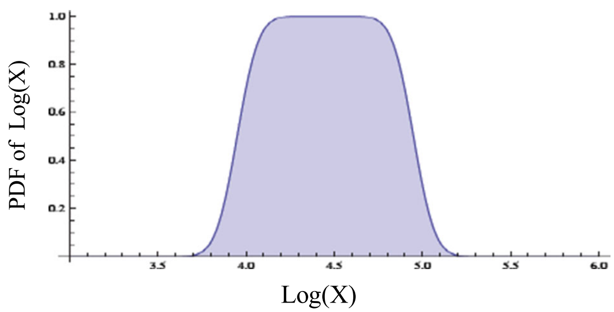
Fig. 1 **a** Distribution of log of numbers in all financial statements (Log(Y)), **b** Distribution of log of numbers for Alcoa's 2011 financial statements (Log(Y))

question for future analytical research. However, we now extend the analysis and use numerical parameters for specific cases to show the intuition of how FSD changes.

A special numerical solution

Assume there are 10 groups of cash flow streams (i.e., $K = 10$, so we have X_1 to X_{10} cash flow streams) and that each of the cash flow streams has a different mean in the log scale, starting from 4 to 4.9, separated by 0.1 (i.e., $\mu_1 = 4$, $\mu_2 = 4.1..$, $\mu_{10} = 4.9$), which means the numbers range from 10,000 to 100,000 in the linear scale. Finally, assume that the standard deviation of each of the X_ks in log scale is $\sigma_k = 1$.

The probability density function of X, that is, the mixture distribution in this year, is therefore the following: PDF (log(X)) $= \sum_{k=1}^{10} \frac{1}{10}\left(\frac{1}{\sqrt{2\pi}} e^{-\frac{(x-\mu_k)^2}{2}}\right)$. As can be seen in the figure below, this distribution is smooth and symmetric and looks similar to a normal distribution:



Furthermore, this distribution follows Benford's Law, and the FSD Score for this distribution under those parameters is FSD Score = 0.

The problem is that X is unobservable to an outsider (and may also be unobservable to the manager). The outsider is observing only the Y_ks where $Y_k = X_k{*}Z_k$. The conclusions about the errors that outsiders can make must come from the distribution of the reported vector of numbers Y.[23] If Z_k is distributed log normal, which means it is distributed normal in the log scale with $\mu_{\varepsilon k}$ and $\sigma_{\varepsilon k}$, then each Y_k is also distributed normal in the log scale with parameters $\mu_{yk} = \mu_k + \mu_{\varepsilon k}$ and $\sigma_{yk} = \sigma_k + \sigma_{\varepsilon k}$. This is essentially the distribution of the sum of two normal variables. Now consider the following three cases.

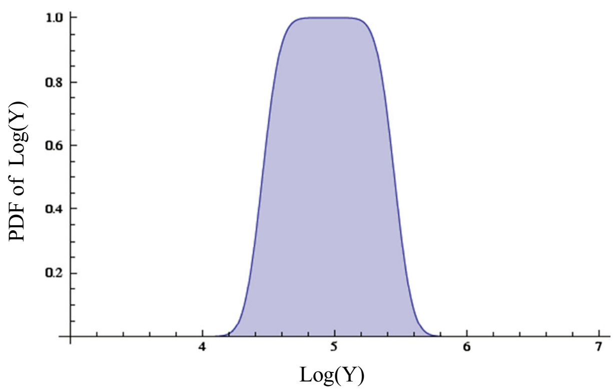
[23] Insider and outsiders do not need to know the means and standard deviations of the original distributions or the error term. They simply need to know that the distribution follows Benford's Law.

Error distributions with equal means and equal standard deviations (Case 1)

In this case, $\mu_{\varepsilon 1} = \mu_{\varepsilon 2} = \ldots = \mu_{\varepsilon 10}$ = Constant C and $\sigma_{\varepsilon 1} = \sigma_{\varepsilon 2} = \ldots = \sigma_{\varepsilon 10}$ = Constant S. In this case, $\mu_{yk} = \mu_k + C$ and $\sigma_{yk} = \sigma_k + S$. The resulting mixture distribution of Y in the log scale will again look like the distribution of X but shifted to the right by a constant C and flatter because of the increased standard deviation, that is,

$$\mathrm{PDF}(\log(Y)) = \sum_{k=1}^{K} \frac{1}{K} \left(\frac{1}{(\sigma_k + s)\sqrt{2\pi}} e^{-\frac{(x-\mu_k+C)^2}{2(\sigma_k+S)^2}} \right),$$

which will follow Benford's Law to a similar degree as the distribution of X. This is because multiplying a distribution that follows Benford's Law in the linear scale by a constant creates a distribution that follows Benford's Law (Hill 1995). With parameters C = 0.5 and S = 0.01, the FSD Score of the resulting distribution is zero, and its PDF is shown in the figure below:



In conclusion, adding identical error terms to all the X_ks does not create deviations from Benford's Law.

Error distributions with equal means but different standard deviations (Case 2)

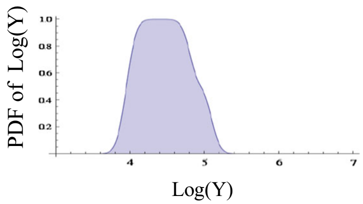
In this case, $\mu_{\varepsilon 1} = \mu_{\varepsilon 2} = \ldots = \mu_{\varepsilon 10}$ = Constant C and $\sigma_{\varepsilon k}$ varies across the ks. Therefore, $\mu_{yk} = \mu_k + C$ and $\sigma_{yk} = \sigma_k + \sigma_{\varepsilon k}$. The resulting mixture distribution of Y in log scale will again look like the distribution of X but wider because of the increased standard deviation. Still, it will closely follow Benford's Law. Here again the FSD Score is zero.

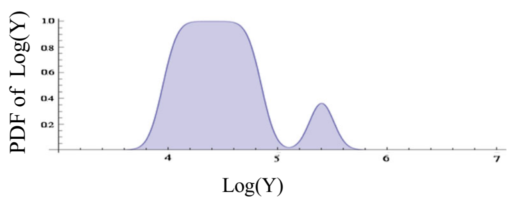
Error distributions with different means but constant standard deviation (Case 3)

In this case, $\mu_{\varepsilon k}$ varies across the ks and $\sigma_{\varepsilon 1} = \sigma_{\varepsilon 2} = \ldots = \sigma_{\varepsilon 10} =$ Constant S. This is the interesting case as it will create deviations from Benford's Law. We consider three different subcases.

Error in the estimation of a single element in the cash flow streams (Case 3A)

We start with the simple case where we change only the $\mu_{\varepsilon 10}$ to add error to X_{10}, which is the highest number in our cash flow streams. We will start increasing $\mu_{\varepsilon 10}$ by increments of 0.1. Therefore μ_{yk} will grow from 4.9 to 5 in the first iteration, to 5.1 in the next iteration, and so on. This situation could be an example of overestimating revenues. The graphical evidence on the way the mixture distribution changes and the resulting FSD Scores is striking for the case of $S = 0.01$. In the case of $\mu_{\varepsilon 10} = 0.1$ and $S = 0.01$, the FSD Score is 0.008, and the resulting distribution is shown in the figure below (left) . In the case of $\mu_{\varepsilon 10} = 0.5$ and $S = 0.01$, the FSD Score is 0.017, and the resulting distribution is shown in the figure below (right).



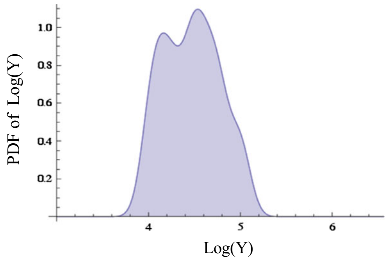


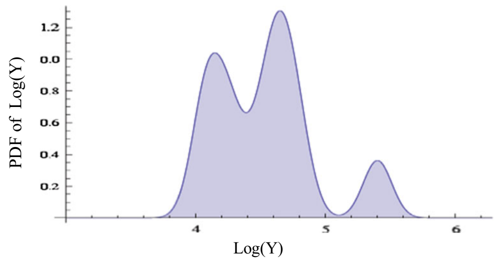
As we increase the mean of the error, under these parameters, the distribution monotonically moves further away from Benford's Law and reaches a limit. This case is consistent with managing revenue upward (or overestimating revenue compared to the actual distribution) leading to deviations in Benford's Law and an increase in the FSD Score.

The case where the errors are correlated with each other (Case 3B)

The case above represents an error in one element of the report. However, a feature of the accounting system is that an error in one element leads to errors in other elements as well. For example, if the manager overestimates revenue, he is also likely to overestimate cost of goods sold (in an amount less than revenue) to match the revenue and will overestimate the related tax payment (in an amount less than revenue). In the terms of our example, there will be a mean error in several of the Z_ks. For example, let us assume $\mu_{\varepsilon 10}$ is increasing by increments of 0.1 as before, but now $\mu_{\varepsilon 5} = 0.5\mu_{\varepsilon 10}$ and $\mu_{\varepsilon 1} = 0.1\mu_{\varepsilon 10}$. Again, it is clear from the shape of the graph and the change in FSD that this will cause a significant deviation from Benford's Law.

In the case of $\mu_{\varepsilon 10} = 0.1$, $\mu_{\varepsilon 5} = 0.5\mu_{\varepsilon 10}$, $\mu_{\varepsilon 1} = 0.1\mu_{\varepsilon 10}$, and $S = 0.01$, the FSD Score is 0.009, and the resulting distribution is shown in the figure below (left). In the case of $\mu_{\varepsilon 10} = 0.5$, $\mu_{\varepsilon 5} = 0.5\mu_{\varepsilon 10}$, $\mu_{\varepsilon 1} = 0.1\mu_{\varepsilon 10}$, and $S = 0.01$, the FSD Score is 0.017, and the resulting distribution is shown in the figure below (right).

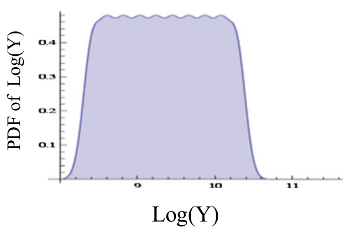


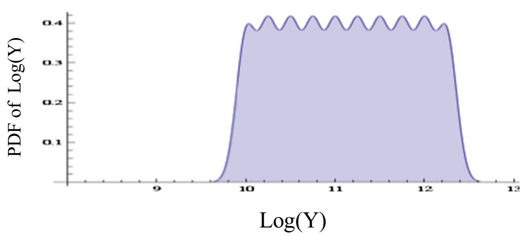


Once again, the point to make from this exercise is that deviation from Benford's Law, under these parameters, is monotonically increasing with the error and reaches a limit, even when the errors are correlated with each other.

The case where the errors are correlated with the mean of the cash flow streams (Case 3C)

It also possible that the estimation errors may be larger for items that are larger. In terms of our example, $\mu_{\varepsilon k}$ is a function of μ_k. For the sake of simplicity, assume $\mu_{\varepsilon k} = \mu_k * B$, where B is a constant multiplier that determines the error size (the larger is B, the larger is the error). It is clear that, if B is zero, we revert to Case 1, and the distribution follows Benford's Law exactly with FSD Score equal to 0. However, when we start increasing B by increments of 0.1, the distributions start to change. In the case of $\mu_{\varepsilon k} = \mu_k * B$, $B = 1.1$, and $S = 0.01$, the FSD Score is 0.004, and the resulting distribution is shown in the figure below (left). In the case of $\mu_{\varepsilon k} = \mu_k * B$, $B = 1.5$, and $S = 0.01$, the FSD Score is 0.016, and the resulting distribution is shown in the figure below (right).





In this case, uneven errors across accounts create deviations from Benford' Law that, under these parameters, monotonically increase the FSD Score before reaching a limit.

Appendix 4: Numerical example when realizations are observable

To see the intuition for why deviations from Benford's Law can be used to assess data quality using real world data, consider the following example. The market value of equity at the end of a trading day is one realization of a random distribution. A sample of different firms in a random day is likely to fit the criteria in Hill (1995). Indeed, consistent with Hill (1996), when examining a random sample of the market value of equity of companies traded in the United States, the distribution follows Benford's Law. Now assume that instead of measuring the market value of equity accurately by transaction price (where we can observe true realizations), the actual realizations are unknown. Therefore, the data provider has to use estimation techniques (for example, using last year's prices times the average return from 2 years ago, or just randomly choosing based on a possible distribution of prices). Errors in the estimation techniques or fabricated data (random or human) are likely to create a very different dataset from the true realized distribution and hence create a deviation from Benford's Law.[24] Therefore, the deviation from Benford's Law can be used as a proxy for how divergent a dataset is from the true, unobservable realizations. If the realization is known and can be measured with complete accuracy, then there is obviously no need to use Benford's Law to validate the data. However, in this case, since the realizations are known, we can observe the actual deviation from the true distribution. Below, we illustrate this with real data.

We look at the market value of equity (MVE) for all firms with available data in CRSP's monthly file (price and shares outstanding) for a random day, August 31, 2011, to build intuition for why Benford's Law can be used to assess data quality. MVE (price * shares outstanding) is a random distribution, and as expected, the FSD Score for MVE for all firms (created using the distribution of the first digits of all firms with available data) is 0.00295, which can be considered close conformity to Benford's Law.

Next, we ask, what if the true market price is unknown, and instead, MVE needs to be estimated or is fabricated? To answer this question, we introduce a noise term that changes MVE, where firm-level MVE is equal to MVE * (1 + a randomly generated number from a normal distribution) and then re-measure the FSD Score. We manipulate the mean of the random number (i.e., the estimation error) first, with the expectation that, as the size of the noise increases, deviation from Benford's Law should also increase. We next keep the mean constant and manipulate the variance, expecting the FSD Score to remain constant since we are no longer changing the magnitude of the noise.

[24] Not all misestimated or fabricated data create deviations from Benford's Law. For example, if the misestimation simply multiplies all true realizations by a constant, the new erroneous data will still follow Benford's Law.

As can be seen below, holding the variance constant, when we increase the mean noise term, the FSD Score increases.

Constant variance	MVE FSD score
Mean = 1, var = 1	0.00294
Mean = 2, var = 1	0.00304
Mean = 3, var = 1	0.00320
Mean = 4, var = 1	0.00322

In contrast, holding the mean noise term constant, when we increase the variance, the FSD Score remains stable.

Constant mean	MVE FSD
Mean = 1, var = 2	0.00292
Mean = 1, var = 3	0.00293
Mean = 1, var = 4	0.00292

These results provide insights into why Benford's Law and the FSD Score can be used to assess the quality of data in financial statements. Financial statement numbers require significant estimation by management. Investors (and even possibly managers) do not observe the true realization of these numbers. Much like changing the mean around the noise term in the MVE example, as estimation error increases in estimating financial statement numbers, we expect the FSD Score to increase as well.

Appendix 5: Simulation analysis

To demonstrate how a firm's potential manipulation of its financial results could alter its conformity to Benford's Law, we ran a simulation that involved changing the value of a single line item in a firm's income statement and calculated how that change affected the financial statements overall. We then re-measured the FSD Score based on the manipulation and the changes the manipulation induced in the financial statements.

We chose to manipulate sales since it is an item that managers may be tempted to change to mask poor performance and is interconnected with many other financial statement items. As a result of the sales manipulation, a firm likely needs to adjust cost of goods sold and tax expense accordingly. Our simulation randomly (from a uniform distribution) seeded a journal entry to increase sales by between 5 and 10 % to make the change material. COGS were increased by between 20 and 90 % as a percent of the increase of sales manipulation, and taxes payable were increased by between 0 and 35 % of the difference between the previous two calculations. Put more simply, we added three journal entries to the original numbers:

1. Increase accounts receivables	Increase revenue
2. Increase cost of goods sold	Decrease inventory
3. Increase tax expense	Increase tax Payable

As a result of the journal entries, we list below the line items that changed in our simulation when sales changed as described above.

Income statement
Sales
Cost of goods sold
Gross profit (Loss)
Operating income after depreciation
Operating income before depreciation
Pretax income
Pretax income–domestic
Income taxes—federal
Income taxes—total
Income before extraordinary items
Income before extraordinary items—adjusted for common stock equivalents
Income before extraordinary items—available for common
Income before extraordinary items and noncontrolling interests
Net income adjusted for common/ordinary stock (Capital) equivalents
Balance sheet
Receivables—Trade
Receivables—Total
Inventories—finished goods
Inventories—total
Current assets—total
Assets—total
Income taxes payable
Current liabilities—total
Liabilities—total
Retained earnings
Stockholders equity—total
Liabilities and stockholders equity—total
Statement of cash flow
Income before extraordinary items (cash flow)
Accounts receivable—decrease(increase)
Inventory—decrease (increase)
Income taxes—accrued—increase/(decrease)

In our simulation, we chose to manipulate a firm with a set of financial numbers that generally, but not perfectly, conforms to Benford’s Law. We therefore chose Alcoa’s 2011 financial results since the results not only conform to Benford’s Law but also contain a large number of line items, ensuring that a single number does not have an undue impact on our measurements. In running the simulation 1000 times, Alcoa's FSD Score increases 950 times (95 %). We interpret the findings from our simulation to imply that divergence from Benford's Law could signal that a firm is intentionally manipulating its financial numbers.

Appendix 6: Variable definitions

Variable	Description	Definition
FSD_Score based on the MAD statistic	Mean absolute deviation statistic for annual financial statement data	The sum of the absolute difference between the empirical distribution of leading digits in annual financial statements and their theoretical Benford distribution, divided by the number of leading digits. See Appendix 1 for a sample calculation
FSD_Score based on the KS statistic	Kolmogorov–Smirnov statistic for annual financial statement data	The maximum deviation of the cumulative differences between the empirical distribution of leading digits in annual financial statements and their theoretical Benford distribution. See Appendix 1 for a sample calculation
AAER	Indicator equal to 1 for the year in which a firm was first identified by the SEC as having materially misstated its financial statements	Firms that were included in the annual SEC Accounting and Auditing Enforcement Releases (AAER) database (Dechow et al. 2011)
ABS_JONES_RESID	Absolute value of the residual from the modified Jones model, following Kothari et al. (2005)	The following regression is estimated for each industry year: $\text{tca} = \Delta\text{sales} + \text{net PPE} + \text{ROA}$, where $\text{tca} = (\Delta\text{current assets} - \Delta\text{cash} - \Delta\text{current liabilities} + \Delta$ debt in current liabilities − depreciation and amortization), ROA is defined as below, and all variables are scaled by beginning-of-period total assets
STD_DD_RESID	Five-year moving standard deviation of the Dechow-Dichev residual, following Francis et al. (2005)	The following regression is estimated for each industry year: $\text{tca} = \text{cfo}_{t-1} + \text{cfo} + \text{cfo}_{t+1}$, where tca is defined as above, and cfo = (income before extraordinary items − (wc_acc—depreciation and amortization)). All variables are scaled by average total assets. The five-year rolling standard deviations of the residuals are then calculated
MANIPULATOR	Indicator variable equal to 1 if the M-Score is greater than −1.78	M-Score is calculated following Beneish (1999)

Variable	Description	Definition
F_SCORE	The scaled probability of earnings management or a misstatement for a firm-year based on firm financial characteristics	Calculated using the coefficients in Table 7 Model 2 of Dechow et al. (2011)
ABS_WCACC	The absolute value of working capital accruals	Calculated as (Δcurrent assets − Δcash − Δcurrent liabilities + Δdebt in current liabilities + Δtaxes paid) scaled by average total assets
ABS_RSST	The absolute value of working capital accruals as defined by Richardson et al. (2005)	Calculated as (ΔWC + ΔNCO + ΔFIN) scaled by average total assets. WC = (current assets − cash and short-term investments) − (current liabilities − debt in current liabilities). NCO = (total assets − current assets − investments and advances) − (total liabilities − current liabilities − long-term debt). FIN = (short-term investments + long-term investments) − (long-term debt + debt in current liabilities + preferred stock)
LOSS	Indicator if firm-year had negative net income	Equal to 1 if net income < 0, 0 otherwise
CH_CS	Change in cash sales	Cash sales$_t$ − cash sales$_{t-1}$/cash sales$_{t-1}$, where cash sales = total revenue − Δtotal receivables
ROA	Return on assets	Income before extraordinary items$_t$/total assets$_{t-1}$
CH_ROA	Change in ROA	ROA$_t$ − ROA$_{t-1}$
SOFT_ASSETS	Soft assets	(Total assets—net PPE—cash)/total assets$_{t-1}$
ISSUE	Indicator variable that equals 1 if the company issued debt or equity in that year	When long-term debt issuance (Compustat DLTIS) > 0 or sale of common or preferred stock (SSTK) > 0, then ISSUE = 1
MKT_VAL	Market value of equity	Closing price at the end of the fiscal year * common shares outstanding
MTB	Market-to-book ratio	MKT_VAL/book value of total stockholders' equity.
NI_VOL	Earnings volatility	Standard deviation of net income for the last five years.
RET_VOL	Return volatility	Standard deviation of monthly stock returns in the last year
PE	Price-to-earnings ratio	Closing stock price at the end of the fiscal year/earnings per share
AT	Total assets	Compustat AT

Variable	Description	Definition
EARNINGS PERSISTENCE		Correlation between net income and net income in the following year
SALES_GROWTH	Year-over-year percentage change in revenue	$(\text{Revenue}_t - \text{Revenue}_{t-1})/\text{Revenue}_{t-1}$
DIV	Dividend indicator	Equal to 1 if a firm issued dividends, 0 otherwise.
SIZE	Log of market value of equity	Log(common shares outstanding * price at the end of the fiscal year)
SI	Special items	Total special items/total assets
AGE	Age of the firm	Number of years the firm appears in the CRSP monthly stock return file
RESTATED_NUMS	Indicator variable that equals 1 if reported numbers are restated	For all firms from 2001 to 2011 where both restated and original financial numbers are available in Compustat (datafmt = STD for original and datafmt = SUMM_STD for restated) and at least 10 variables have changed, we separate the original from the restated financial numbers and create an indicator equaling 1 for restated numbers
INDUSTRY	Industry classification	Groups companies into 17 industry portfolios based on the Fama–French industry classification

Appendix 7

See Figs. 2, 3, 4 and Tables 1, 2, 3, 4, 5, 6, 7, 8, 9, 10.

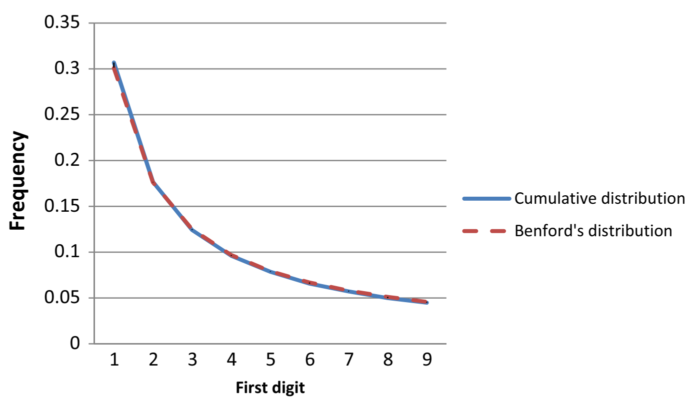


Fig. 2 Aggregate Distribution and Benford's Distribution. Figure 2 shows the similarity between Benford's distribution and the aggregate distribution of all financial statement variables available on Compustat for the period 2001–2011. Not shown are distributions by industry and year, which similarly conform to Benford's Law

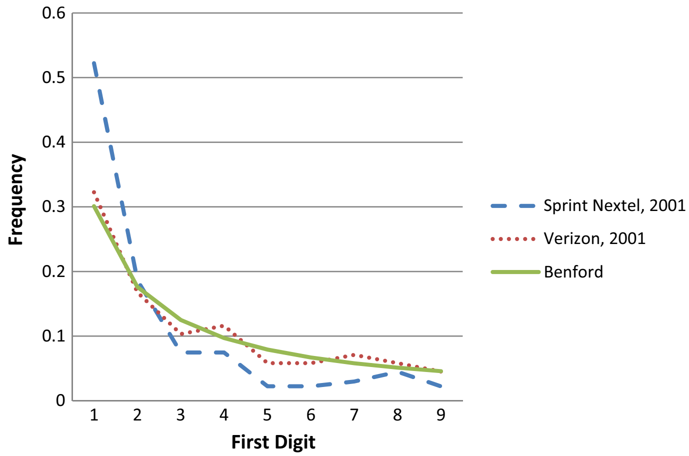


Fig. 3 Conformity to Benford's distribution, firm examples Figure 3 shows the conformity to Benford's distribution for two firm years, Sprint Nextel 2001, which does not conform to Benford's Law (FSD Scored based on the KS statistic = 0.224, FSD Score based on the MAD statistic = 0.052) and restated its financial results for that year, and Verizon Communications 2001, which does conform to Benford's Law (FSD Score based on the KS statistic = 0.056, FSD Score based on the MAD statistic = 0.017)

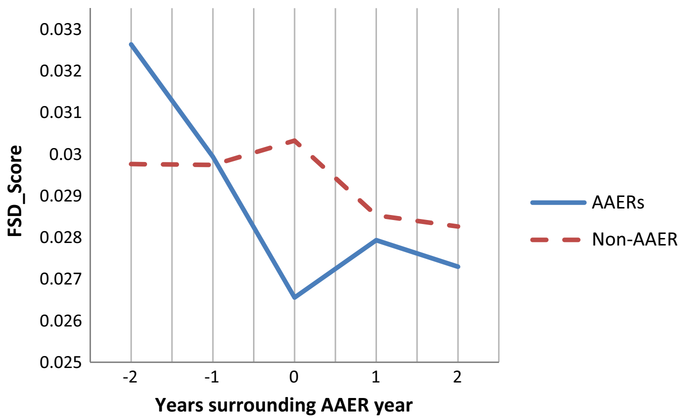


Fig. 4 Trend in FSD_Score for AAER and non-AAER firms. Figure 4 depicts the time trend in FSD_Score for firms identified by the SEC, through an Accounting and Auditing Enforcement Release, as having materially misstated their financial statements. Year 0 on the x axis is centered on the year the misstatement began to demonstrate how FSD_Score changes before and after the misstatement

Table 1 Descriptive statistics

Variable	N	Mean	Std Dev	Q1	Median	Q3
Panel A: Descriptive statistics						
FSD_Score	43,332	0.0296	0.0087	0.0234	0.0288	0.0351
ABS_JONES_RESID	43,332	0.1835	0.3622	0.0281	0.0707	0.1734
STD_DD_RESID	43,332	0.1232	0.1523	0.0383	0.0794	0.1558
MANIPULATOR	39,895	0.1425	0.3496	0	0	0
F_SCORE	38,590	0.4005	0.2593	0.2067	0.3431	0.5287
ABS_WCACC	37,693	0.0540	0.0679	0.0124	0.0305	0.0668
ABS_RSST	38,590	0.1383	0.1713	0.0328	0.0772	0.1697
LOSS	43,332	0.3598	0.4799	0	0	1
AAER	43,332	0.0085	0.0919	0	0	0
CH_CS	38,590	0.1461	0.5017	−0.0481	0.0718	0.2174
CH_ROA	38,590	−0.0019	0.1958	−0.0418	0.0004	0.0360
SOFT_ASSETS	38,590	0.5446	0.2349	0.3661	0.5650	0.7352
ISSUE	38,590	0.9150	0.2790	1	1	1
SALES_GROWTH	42,771	0.1575	0.5710	−0.0415	0.0753	0.2160
DIV	35,818	0.4027	0.4905	0	0	1
SIZE	35,818	6.20	2.12	4.72	6.18	7.61
SI	35,453	−0.0240	0.0794	−0.0148	−0.0011	0
AGE	35,818	20	11.89	12	17	27
MKT_VAL	43,332	3645.67	11,088.37	67.34	369.05	1745.56
MTB	43,332	1.36	1.42	0.4940	0.9283	1.68
NI_VOL	35,818	148.28	409.23	7.00	22.21	82.55
RET_VOL	35,364	0.1456	0.0927	0.0827	0.1216	0.1791
PE	43,055	11.53	107.37	−2.19	11.48	21.42
AT	43,332	3228.38	7797.25	85.40	383.91	1883.85

Table 1 continued

	FSD_Score	ABS_JONES_RESID	STD_DD_RESID	MANIPULATOR	F_SCORE	ABS_WCACC	ABS_RSST	LOSS
Panel B: Correlation matrix								
FSD_Score		0.0593*	0.0889*	0.0385*	−0.1130*	0.0594*	0.0646*	0.0814*
ABS_JONES_RESID	0.0563*		0.3685*	0.0842*	−0.0982*	0.2694*	0.2521*	0.1384*
STD_DD_RESID	0.1179*	0.2519*		0.0735*	−0.0603*	0.1876*	0.2312*	0.2027*
MANIPULATOR	0.0609*	0.0703*	0.1280*		0.1155*	0.2147*	0.1732*	0.0219*
F_SCORE	−0.0841*	−0.0099	0.0119*	0.1617*		0.0819*	0.0149*	−0.1548*
ABS_WCACC	0.0955*	0.2196*	0.3054*	0.2463*	0.1491*		0.2530*	0.2040*
ABS_RSST	0.1058*	0.2263*	0.3518*	0.1986*	0.1229*	0.3927*		0.2634*
LOSS	0.1144*	0.0909*	0.2281*	0.0528*	−0.0929*	0.1472*	0.1914*	
AAER	−0.0166*	0.0035	0.0070	0.0016	0.0380*	0.0031	0.0084	−0.0025
PE	−0.0042	−0.0103*	−0.0316*	−0.0044	0.0107*	−0.1172*	−0.1227*	−0.2652*
AT	−0.1484*	−0.0643*	−0.1349*	−0.0840*	−0.0417*	−0.1722*	−0.1488*	−0.1807*
MKT_VAL	−0.1165*	−0.0282*	−0.0793*	−0.0700*	−0.0177*	−0.1363*	−0.1075*	−0.1905*
NI_VOL	−0.1122*	−0.0284*	−0.0345*	−0.0433*	−0.0516*	−0.1035*	−0.0268*	−0.0350*
RET_VOL	0.1052*	0.1285*	0.2337*	0.1014*	−0.0567*	0.2465*	0.2711*	0.4024*
MTB	0.0981*	0.1385*	0.2471*	0.1009*	0.0296*	0.1180*	0.2125*	−0.0654*

	AAER	PE	AT	MKT_VAL	NI_VOL	RET_VOL	MTB
Panel B: Correlation matrix							
FSD_Score	−0.0199*	−0.0604*	−0.2502*	−0.1997*	−0.1938*	0.1092*	0.0734*
ABS_JONES_RESID	−0.0010	−0.0792*	−0.1976*	−0.1238*	−0.0760*	0.1838*	0.1365*
STD_DD_RESID	0.0025	−0.1161*	−0.2866*	−0.1892*	−0.0605*	0.2170*	0.1821*
MANIPULATOR	0.0003	−0.0343*	−0.1093*	−0.0722*	−0.0575*	0.1074*	0.0666*
F_SCORE	0.0426*	0.1236*	0.0665*	0.0782*	−0.0585*	−0.1260*	0.0340*
ABS_WCACC	−0.0033	−0.0312*	−0.3124*	−0.2765*	−0.1848*	0.2487*	0.0208*

Table 1 continued

	AAER	PE	AT	MKT_VAL	NI_VOL	RET_VOL	MTB
ABS_RSST	0.0082	−0.0281*	−0.2130*	−0.1221*	−0.0255*	0.2221*	0.1855*
LOSS	−0.0056	−0.8099*	−0.2961*	−0.3901*	−0.0177*	0.4060*	−0.2355*
AAER		0.0154*	0.0304*	0.0370*	0.0309*	0.0067	0.0196*
PE	0.0024		0.2062*	0.3190*	−0.0039	−0.3095*	0.2805*
AT	0.0070	0.0075		0.8920*	0.8041*	−0.3920*	−0.0570*
MKT_VAL	0.0062	0.0186*	0.8286*		0.7188*	−0.4448*	0.3564*
NI_VOL	0.0104*	−0.0072	0.7322*	0.6828*		−0.1380*	−0.0298*
RET_VOL	0.0091	−0.0445*	−0.2304*	−0.2181*	−0.0893*		−0.1620*
MTB	0.0171*	0.0372*	−0.0098*	0.1856*	0.0334*	−0.0069	

	Average number of line items	Number of firm-years	Average digit distributions									FSD_Score
			1	2	3	4	5	6	7	8	9	
Panel C: FSD_Score by number of financial statement line items												
Top 1 % of line items	169	420	0.3019	0.1798	0.1249	0.0980	0.0776	0.0672	0.0558	0.0506	0.0442	0.0231
Top tercile	144	14,730	0.3037	0.1781	0.1247	0.0969	0.0789	0.0660	0.0572	0.0498	0.0448	0.0259
Middle tercile	124	14,941	0.3041	0.1769	0.1248	0.0958	0.0786	0.0661	0.0579	0.0506	0.0454	0.0292
Bottom tercile	108	13,661	0.3042	0.1771	0.1241	0.0966	0.0788	0.0661	0.0577	0.0502	0.0453	0.0335
Bottom 1 % of line items	100	736	0.3017	0.1764	0.1254	0.0957	0.0799	0.0666	0.0597	0.0499	0.0447	0.0362
Overall	125	43,332	0.3040	0.1773	0.1245	0.0964	0.0787	0.0661	0.0576	0.0502	0.0452	0.0296

	Average Assets ($M)	Number of firm-years	Average digit distributions									FSD_Score
			1	2	3	4	5	6	7	8	9	
Panel D: FSD_Score by firm assets												
Largest 1 % of firms	38,710	433	0.2940	0.1775	0.1311	0.0996	0.0797	0.0671	0.0560	0.0500	0.0450	0.0260

Table 1 continued

	Average Assets ($M)	Number of firm-years	Average digit distributions									FSD_Score
			1	2	3	4	5	6	7	8	9	
Top tercile	54	14,444	0.3066	0.1775	0.1238	0.0952	0.0780	0.0658	0.0576	0.0503	0.0452	0.0271
Middle tercile	448	14,444	0.3059	0.1795	0.1243	0.0964	0.0775	0.0652	0.0567	0.0497	0.0448	0.0293
Bottom tercile	9183	14,444	0.2994	0.1749	0.1255	0.0976	0.0807	0.0672	0.0585	0.0506	0.0456	0.0325
Smallest 1 % of firms	1.72	434	0.3048	0.1841	0.1206	0.0971	0.0786	0.0642	0.0580	0.0496	0.0430	0.0355
Overall	3228	43,332	0.3040	0.1773	0.1245	0.0964	0.0787	0.0661	0.0576	0.0502	0.0452	0.0296

Table 1 reports several sets of descriptive statistics. Panel A reports descriptive statistics for all sample variables. In Panel B, Pearson (Spearman) correlations are below (above) the diagonal where * indicates significance at the 5 % level. Panel C segments firm-years by the number of nonmissing financial statement line items (i.e., number of Compustat variables) available to calculate the FSD_Score. Reported are the average number of line items, the number of firm-years, the average empirical digit distribution of the first digits 1 through 9, and the average FSD_Score for firms in the top 1 %, top tercile, middle tercile, bottom tercile, and bottom 1 % of number of available line items. Panel D segments firm-years by total assets (AT). Reported are the average number of line items, the number of firm-years, the average empirical digit distribution of the first digits 1 through 9, and the average FSD_Score for firms in the top 1 %, top tercile, middle tercile, bottom tercile, and bottom 1 % of total assets. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. See Appendix 6 for variable definitions

Table 2 Aggregate conformity to Benford's Law

Number of firm-years	Aggregate FSD_Score
Panel A: FSD_Score for all firm-years' financial statement numbers	
43,332	0.0009

FF Industry	Number of firm-years	Aggregate FSD_Score
Panel B: FSD_Score for all financial statement numbers, by industry		
1	1410	0.0009
2	689	0.0018
3	1931	0.0009
4	888	0.0013
5	1176	0.0010
6	1046	0.0013
7	1901	0.0012
8	1185	0.0011
9	728	0.0012
10	362	0.0011
11	7056	0.0007
12	725	0.0011
13	1971	0.0020
14	1215	0.0017
15	2819	0.0020
17	18,230	0.0010

Fiscal year	Number of firm-years	Aggregate FSD_Score
Panel C: FSD_Score for all financial statement numbers, by year		
2001	4418	0.0008
2002	4345	0.0011
2003	4177	0.0013
2004	4153	0.0010
2005	4072	0.0012
2006	3955	0.0009
2007	3854	0.0007
2008	3747	0.0008
2009	3643	0.0010
2010	3531	0.0010
2011	3437	0.0009

Table 2 computes the aggregate FSD_Score from all financial statement variables available for all firm-years in the sample. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in all firms' financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. Panel A reports the aggregate FSD_Score for all variables for all firm-years. Panel B reports the aggregate FSD_Score by Fama–French industry portfolios. Panel C reports the aggregate FSD_Score by fiscal years

Table 3 Firm-year conformity to Benford's Law

Firm-years conforming		Percent conforming
Panel A: Number of firm-years conforming to Benford's Law		
37,104		85.63

Financial statement	Firm-years conforming	Percent conforming
Panel B: Number of firm-years conforming to Benford's Law by financial statement		
Balance sheet	39,274	90.64
Income statement	34,138	78.78
Cash flow statement	42,259	97.52

FF Industry	Firm-years conforming	Percent conforming
Panel C: Number of firm-years conforming to Benford's Law by industry		
1	1218	86.38
2	580	84.18
3	1680	87.00
4	765	86.15
5	987	83.93
6	910	87.00
7	1556	81.85
8	1054	88.95
9	642	88.19
10	318	87.85
11	6030	85.46
12	627	86.48
13	1706	86.56
14	1100	90.53
15	2382	84.50
17	15,549	85.29

Fiscal year	Firm-years conforming	Percent conforming
Panel D: Number of firm-years conforming to Benford's Law by fiscal year		
2001	3795	85.90
2002	3753	86.38
2003	3562	85.28
2004	3538	85.19
2005	3492	85.76
2006	3360	84.96
2007	3301	86.65
2008	3221	85.96
2009	3110	85.37
2010	3017	85.44

Table 3 continued

Fiscal year	Firm-years conforming	Percent conforming
2011	2955	85.98

Table 3 computes FSD_Score based on the KS statistic for each individual firm-year in the sample and reports the percentage of firm-years that conform to Benford's Law, where conformity is assessed as having a KS statistic that is not significantly different from zero at the 5 % level. In Panel A, 86 % of all firm-years are not different from zero at the 5 % level. Panel B computes FSD_Score based on the KS statistic for the variables in each of the three financial statements for each individual firm-year and reports the percentage of firm-years that conform to Benford's Law for each statement. Panel C (Panel D) reports conformity to Benford's Law based on the KS statistic across industries (fiscal years). See Appendix 1 for the calculation of FSD_Score based on the KS statistic

Table 4 Aggregate FSD_Score by financial statement characteristics and line items

Financial statement	Number of accounts	Aggregate FSD_Score
Panel A: Aggregate FSD_Score by financial statement		
Balance sheet	111	0.0005
Income statement	101	0.0020
Cash flow statement	38	0.0005
Panel B: Aggregate FSD_Score by financial statement subcategory		
Balance sheet		
Assets	37	0.0004
Liabilities	43	0.0007
Equity	28	0.0010
Income statement		
Expenses	11	0.0012
Income	32	0.0027

Table 4 computes the aggregate FSD_Score from all financial statement variables available for all firm-years in the sample by financial statement characteristics and line items. Panel A reports the aggregate FSD_Score for all variables for all firm-years for each of the financial statements. Panel B reports the aggregate FSD_Score for subcategories of each of the financial statements. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score

Table 5 Firm characteristics based on FSD_Score

Variable	Top tercile	Middle tercile	Bottom tercile
Terciles by FSD_Score			
AGE	18.68	20.26	21.61***
CH_CS	0.1688	0.1416	0.1279***
CH_ROA	−0.0011	−0.0028	−0.0017
DIV	0.3403	0.4067	0.4612***
ISSUE	0.8924	0.9199	0.9328***
MKT_VAL	2199.64	3633.94	5103.43***
MTB	1.52	1.35	1.22***
PE	11.79	9.80	13.00
RET_VOL	0.1564	0.1444	0.1359***
SALES_GROWTH	0.1838	0.1567	0.1322***

Table 5 segments firm-years into terciles based on FSD_Score, calculates the means of various firm characteristics based on this segmentation, and reports the significance level of the difference between terciles 1 and 3. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. See Appendix 6 for variable definitions. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 6 FSD_Score and reporting quality

Variable	Tercile by FSD_Score		
	Top tercile	Middle tercile	Bottom tercile
Panel A: Univariate evidence			
Accrual Quality			
ABS_JONES_RESID	0.2076	0.1787	0.1644***
STD_DD_RESID	0.1434	0.1195	0.1067***
MANIPULATOR	0.1646	0.1407	0.1223***
F_SCORE	0.3733	0.4071	0.4212***
ABS_WCACC	0.0611	0.0532	0.0479***
ABS_RSST	0.1579	0.1374	0.1198***
Earnings Quality			
EARNINGS PERSISTENCE	0.6094	0.6831	0.6921***
LOSS	0.4212	0.3483	0.3097***

Panel B: Multivariate evidence

$$\text{FSD\_Score}_{i,t} = \alpha + \beta_1 \text{ABS\_JONES\_RESID}_{i,t} + \beta_2 \text{STD\_DD\_RESID}_{i,t} + \beta_3 \text{MANIPULATOR}_{i,t} + \beta_4 \text{F\_SCORE}_{i,t} + \beta_5 \text{ABS\_WCACC}_{i,t} + \beta_6 \text{ABS\_RSST}_{i,t} + \beta_7 \text{LOSS}_{i,t} + \varepsilon_{i,t}$$

Variable	FSD_Score
ABS_JONES_RESID	0.0002
	(1.45)
STD_DD_RESID	0.0042***
	(12.62)
MANIPULATOR	0.0010***

Table 6 continued

Panel B: Multivariate evidence

$$FSD\_Score_{i,t} = \alpha + \beta_1 ABS\_JONES\_RESID_{i,t} + \beta_2 STD\_DD\_RESID_{i,t} + \beta_3 MANIPULATOR_{i,t} + \beta_4 F\_SCORE_{i,t} + \beta_5 ABS\_WCACC_{i,t} + \beta_6 ABS\_RSST_{i,t} + \beta_7 LOSS_{i,t} + \varepsilon_{i,t}$$

Variable	FSD_Score
	(7.23)
F_SCORE	−0.0031***
	(−17.06)
ABS_WCACC	0.0051***
	(6.72)
ABS_RSST	0.0023***
	(7.63)
LOSS	0.0012***
	(11.72)
Constant	0.0292***
	(291.92)
Observations	34,351
R-squared	0.036

Table 6 examines the relation between Benford's Law and proxies for accruals-based earnings management and earnings manipulation. Panel A segments firm-years into terciles based on FSD_Score, calculates the means of various proxies for reporting quality based on this segmentation, and reports the significance level of the difference between terciles 1 and 3. Panel B reports an OLS regression on the association between these variables and the FSD_Score. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. See Appendix 6 for variable definitions. *t* statistics are reported in parentheses in the table. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 7 FSD_Score and Ex post measures of earnings manipulation

	FSD_Score	Number of firm-years	*t* statistic
RESTATED_NUMS = 0	0.0289	4935	5.36***
RESTATED_NUMS = 1	0.0280	4935	
LOSS = 0	0.0289	27,743	23.98***
LOSS = 1	0.0310	15,589	
$-0.005 \leq NI_t/MKT\_VAL_{t-1} < 0$	0.0283	588	2.32**
$0 \leq NI_t/MKT\_VAL_{t-1} \leq 0.005$	0.0296	426	
AAER = 1	0.0270	82	2.75***
AAER = 0	0.0296	42,963	

Table 7 provides univariate evidence on the relation between Benford's Law and ex post measures of earnings manipulation. RESTATED_NUMS is an indicator that equals 1 for when the FSD_Score is calculated based on restated numbers and 0 for unrestated numbers. LOSS is an indicator equal to 1 for firms reporting negative net income. We follow Burgstahler and Dichev (1997) and examine firms reporting just below and just above the zero net income mark, scaled by market value of equity (NI/MKT_VAL). AAER is an indicator equal to 1 for the year in which a firm was first identified by the SEC as having materially misstated its financial statements. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 8 FSD_Score and restated data

$$\begin{aligned}\mathrm{FSD}_{\mathrm{Score}\,i,t} = \alpha &+ \beta_1\mathrm{RESTATED\_NUMS}_{\mathrm{i,t}} + \beta_2\mathrm{ABS\_JONES\_RESID}_{\mathrm{i,t}}\\ &+ \beta_3\mathrm{STD\_DD\_RESID}_{\mathrm{i,t}} + \beta_4\mathrm{MANIPULATOR}_{i,t} + \beta_5\mathrm{F\_SCORE}_{i,t}\\ &+ \beta_6\mathrm{ABS\_WCACC}_{i,t} + \beta_7\mathrm{ABS\_RSST}_{i,t} + \varepsilon_{i,t}\end{aligned}$$

Variable	FSD_Score	
	(1)	(2)
RESTATED_NUMS	−0.0009***	−0.0009***
	(−5.26)	(−5.33)
ABS_JONES_RESID		0.0000
		(0.06)
STD_DD_RESID		0.0049***
		(7.88)
MANIPULATOR		0.0005*
		(1.89)
F_SCORE		−0.0033***
		(−10.11)
ABS_WCACC		0.0024*
		(1.65)
ABS_RSST		0.0032***
		(5.66)
Constant	0.0288***	0.0289***
	(244.85)	(149.34)
Observations	10,192	10,192
R-squared	0.003	0.030

Table 8 examines the relation between Benford's Law and restated data. The OLS regressions use financial statement data from firms that restated their financial statements for the period 2001–2011. We require that firms have both restated and original financial data available in Compustat. RESTATED_NUMS is an indicator that equals 1 for restated numbers and 0 for misstated numbers used in the calculation of FSD_Score. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. See Appendix 6 for definitions of the control variables. t-statistics are reported in parentheses in the table. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

Table 9 FSD_Score and earnings persistence

$$\begin{aligned}NI_{i,t+1} = {} & \alpha + \beta_1 NI_{i,t} + \beta_2 FSD\_Score_{i,t} + \beta_3 FSD\_Score \times NI_{i,t} + \beta_4 ABS\_JONES\_RESID_{i,t} \\ & + \beta_5 STD\_DD\_RESID_{i,t} + \beta_6 MANIPULATOR_{i,t} + \beta_7 F\_SCORE_{i,t} + \beta_8 ABS\_WCACC_{i,t} \\ & + \beta_9 ABS\_RSST_{i,t} + \beta_{10} LOSS_{i,t} + \beta_{11} SALES\_GROWTH_{i,t} + \beta_{12} DIV_{i,t} + \beta_{13} SIZE_{i,t} \\ & + \beta_{14} MTB_{i,t} + \beta_{15} SI_{i,t} + \beta_{16} AGE_{i,t} + \beta_{17} RET\_VOL_{i,t} + \beta_{18} NI\_VOL_{i,t} + \varepsilon_{i,t}\end{aligned}$$

Variable	NI_{t+1}
NI	0.3268***
	(16.70)
FSD_Score	−1.0403***
	(−5.58)
FSD × NI	−2.6277***
	(−5.26)
ABS_JONES_RESID	0.0005
	(0.11)
STD_DD_RESID	−0.1625***
	(−12.00)
MANIPULATOR	−0.0018
	(−0.37)
F_SCORE	0.0071
	(1.09)
ABS_WCACC	−0.0393
	(−1.33)
ABS_RSST	−0.1151***
	(−9.45)
LOSS	−0.0957***
	(−23.16)
SALES_GROWTH	−0.0105***
	(−3.31)
DIV	−0.0233***
	(−6.59)
SIZE	0.0137***
	(13.00)
MTB	−0.0032**
	(−2.47)
SI	−0.1116***
	(−4.95)
AGE	0.0002
	(1.29)
RET_VOL	−0.1388***
	(−7.22)
NI_VOL	−0.0000
	(−0.52)
Constant	0.0362***

Table 9 continued

$$NI_{i,t+1} = \alpha + \beta_1 NI_{i,t} + \beta_2 FSD\_Score_{i,t} + \beta_3 FSD\_Score \times NI_{i,t} + \beta_4 ABS\_JONES\_RESID_{i,t} + \beta_5 STD\_DD\_RESID_{i,t} + \beta_6 MANIPULATOR_{i,t} + \beta_7 F\_SCORE_{i,t} + \beta_8 ABS\_WCACC_{i,t} + \beta_9 ABS\_RSST_{i,t} + \beta_{10} LOSS_{i,t} + \beta_{11} SALES\_GROWTH_{i,t} + \beta_{12} DIV_{i,t} + \beta_{13} SIZE_{i,t} + \beta_{14} MTB_{i,t} + \beta_{15} SI_{i,t} + \beta_{16} AGE_{i,t} + \beta_{17} RET\_VOL_{i,t} + \beta_{18} NI\_VOL_{i,t} + \varepsilon_{i,t}$$

Variable	NI_{t+1}
	(3.29)
Observations	28,042
R-squared	0.225

Table 9 examines the relation between Benford's Law and earnings persistence. NI is reported net income scaled by total assets. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. Control variables are based on those used in Li (2008). See Appendix 6 for definitions of the control variables. t-statistics are reported in parentheses in the table. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively. Standard errors are clustered by firm and year

Table 10 FSD_Score and material misstatements

$$AAER_{i,t} = \alpha + \beta_1 FSD\_Score + \beta_2 ABS\_JONES\_RESID_{i,t} + \beta_3 STD\_DD\_RESID_{i,t} + \beta_4 MANIPULATOR_{i,t} + \beta_5 F\_SCORE_{i,t} + \beta_6 ABS\_WCACC_{i,t} + \beta_7 ABS\_RSST_{i,t} + \beta_8 CH\_CS_{i,t} + \beta_9 CH\_ROA_{i,t} + \beta_{10} SOFT\_ASSETS_{i,t} + \beta_{11} ISSUE_{i,t} + \beta_{12} MTB_{i,t} + \beta_{13} AT_{i,t} + \varepsilon_{i,t}$$

Variable	AAER		
	(1)	(2)	(3)
FSD_Score	−40.691***		
	(−3.87)		
FSD_Score_{t-1}		21.963*	
		(1.80)	
FSD_Score_{t-2}			39.222***
			(7.34)
ABS_JONES_RESID	−1.078	−1.074	−1.059
	(−1.38)	(−1.33)	(−1.32)
STD_DD_RESID	0.011	−0.171	−0.191
	(0.02)	(−0.27)	(−0.32)
MANIPULATOR	0.122	0.116	0.109
	(0.48)	(0.45)	(0.44)
F_SCORE	1.980***	1.978***	1.994***
	(5.88)	(5.80)	(5.58)
ABS_WCACC	−1.233	−1.613	−1.702
	(−0.78)	(−1.02)	(−1.09)
ABS_RSST	0.401	0.356	0.274
	(0.83)	(0.75)	(0.57)
CH_CS	0.004	0.001	−0.042

Table 10 continued

$$\begin{aligned}\mathrm{AAER_{i,t}} = {} & \alpha + \beta_1 \mathrm{FSD\_Score} + \beta_2 \mathrm{ABS\_JONES\_RESID_{i,t}} + \beta_3 \mathrm{STD\_DD\_RESID_{i,t}} \\ & + \beta_4 \mathrm{MANIPULATOR_{i,t}} + \beta_5 \mathrm{F\_SCORE_{i,t}} + \beta_6 \mathrm{ABS\_WCACC_{i,t}} + \beta_7 \mathrm{ABS\_RSST_{i,t}} \\ & + \beta_8 \mathrm{CH\_CS_{i,t}} + \beta_9 \mathrm{CH\_ROA_{i,t}} + \beta_{10} \mathrm{SOFT\_ASSETS_{i,t}} + \beta_{11} ISSUE_{i,t} + \beta_{12} MTB_{i,t} \\ & + \beta_{13} AT_{i,t} + \varepsilon_{i,t}\end{aligned}$$

Variable	AAER		
	(1)	(2)	(3)
	(0.03)	(0.01)	(−0.26)
CH_ROA	1.339*	1.328*	1.325**
	(1.80)	(1.89)	(1.99)
SOFT_ASSETS	−0.121	0.071	0.100
	(−0.27)	(0.16)	(0.22)
ISSUE	−0.341	−0.249	−0.231
	(−0.51)	(−0.38)	(−0.34)
MTB	0.166*	0.146*	0.145
	(1.83)	(1.65)	(1.61)
AT	−0.000	−0.000	0.000
	(−0.46)	(−0.11)	(0.02)
Constant	−5.686***	−7.620***	−8.198***
	(−5.67)	(−9.15)	(−8.73)
Observations	27,805	27,805	27,805

Table 10 examines the relation between Benford's Law and material misstatements as identified by the SEC. The logistic regressions use all financial statement data for the period 2001–2011. AAER is an indicator variable equal to 1 for the year in which a firm was first identified by the SEC as having materially misstated its financial statements. Column (1) examines the relation between the contemporaneous FSD_Score and material misstatements. Columns (2) and (3) examine the relation between material misstatements and the FSD Scores in years t−1 and t−2, respectively. FSD_Score is the mean absolute deviation between the empirical distribution of leading digits contained in a firm's financial statements and Benford's Law. See Appendix 1 for the calculation of FSD_Score. See Appendix 6 for variable definitions. t-statistics are reported in parentheses in the table. *, **, and *** indicate significance at the 0.10, 0.05, and 0.01 levels, respectively

References

Beneish, M. (1999). The detection of earnings manipulation. *Financial Analyst Journal, 55*(1999), 24–36.

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

Blodget, H. (2008). Bernie Madoff's miraculous returns: Month by month. *Business Insider*, December 12. Accessed August 10, 2013.

Burgstahler, D., & Dichev, I. (1997). Earnings Management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24*(1), 99–126.

Bushman, R., & Smith, A (2003). Transparency, financial accounting information, and corporate governance. *Federal Reserve Bank of New York Policy Review* (April), pp. 65–87.

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

Cleary, R., & Thibodeau, J. C. (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.

Dechow, P., & Dichev, I. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*(supplement), 35–59.

Dechow, P., Ge, W., Larson, C., & Sloan, R. (2011). Predicting material accounting misstatements. *Contemporary Accounting Research, 28*(1), 17–82.

Dechow, P., Ge, W., & Schrand, C. (2010). Understanding earnings quality: A review of the proxies, their determinants and their consequences. *Journal of Accounting and Economics, 50*, 344–401.

Dechow, P., & Skinner, D. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons, 14*(2), 235–250.

Duffie, D., & Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica, 69*(3), 633–664.

Durtschi, C., Hillison, W., & Pacini, C. (2004). The effective use of Benford's law to assist in the detecting of fraud in accounting data. *Journal of Forensic Accounting, 5*, 17–34.

Dyck, A., Morse, A., & Zingales, L. (2013). *How pervasive is corporate fraud*? Working paper.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2004). Cost of equity and earnings attributes. *The Accounting Review, 79*, 967–1010.

Francis, J., LaFond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accruals quality. *Journal of Accounting and Economics, 39*, 295–327.

Gallu, J. (2013). SEC to move past financial crisis cases under chairman white. *Bloomberg*, April 18. Accessed July 15, 2013.

Grundfest, J., and N. Malenko (2009). Quadrophobia: Strategic Rounding of EPS Data. Working paper.

Hill, T. (1995). A statistical derivation of the significant digit law. *Statistical Science, 10*, 354–363.

Hill, T. (1996). A note on the distribution of true versus fabricated data. *Perceptual and Motor Skills, 83*, 776–778.

Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*, 193–228.

Kothari, S. P., Leone, A., & Wasley, C. (2005). Performance matched discretionary accrual measures. *Journal of Accounting and Economics, 39*, 163–197.

La Porta, R., Lopez-De-Silanes, F., Shleifer, A., & Vishny, R. (2000). Legal determinants of external capital. *Journal of Finance, 55*(1), 1–33.

Ley, E. (1996). On the peculiar distribution of the US Stock Indexes' Digits. *The American Statistician, 50*(4), 311–313.

Li, F. (2008). Annual report readability, current earnings, and earnings persistence. *Journal of Accounting and Economics, 45*, 221–247.

McKenna, F. (2012). Is the SEC's Ponzi Crusade Enabling Companies to Cook the Books, Enron-Style? *Forbes*, October 18. Accessed July 15, 2013.

McKenna, F. (2013). Where should SEC start a fraud crack down? Maybe look at fake restatements. *Forbes.com*, June 18. Accessed July 15, 2013.

Michalski, T., & Stoltz, G. (2013). Do countries falsify economic data strategically? Some evidence that they might. *The Review of Economics and Statistics, 95*(2), 591–616.

Morrow, J. (2010). *Benford's law, families of distributions and a test basis*. Working paper.

Nigrini, M. (1996). Taxpayer compliance application of Benford's law. *Journal of American Taxation Association, 18*(1), 72–92.

Nigrini, M. (2012). *Benford's law: Applications for forensic accounting, auditing, and fraud detection*. Hoboken, N.J.: Wiley.

Nigrini, M., & Miller, S. (2009). Data diagnostics using second-order test of Benford's law. *Auditing: A Journal of Practice and Theory, 28*(2), 305–324.

Owens, E., Wu, J., & Zimmerman, J. (2013). *Business model shocks and abnormal accrual models*. Working paper.

Pike, D. Testing for the benford property. Working paper (2008).

Pimbley, J. M. (2014) Benford's law and the risk of financial fraud. *Risk Professional* (May), 1–7.

Rajan, R., & Zingales, L. (2003a). Financial dependence and growth. *American Economic Review, 88*(3), 559–586.

Rajan, R., & Zingales, L. (2003b). The great reversals: The politics of financial development in the twentieth century. *Journal of Financial Economics, 69*(1), 5–50.

Ray, S., & Lindsay, B. (2005). The topography of multivariate normal mixtures. *The Annals of Statistics, 33*(5), 2042–2065.

Richardson, S., Sloan, R., Soliman, M., & Tuna, I. (2005). Accrual reliability, earnings persistence and stock prices. *Journal of Accounting and Economics, 39*, 437–485.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business, 74*, 101–124.

Smith, S. (1997). *The scientist and Engineer's guide to digital signal processing*. California: California Technical Publications.

Thomas, J. (1989). Unusual patterns in reported earnings. *The Accounting Review, 5*(4), 773–787.

Varian, H. (1972). Benford's Law. *American Statistician, 23*, 65–66.

Whalen, D., Cheffers, M., & Usvyatsky, O (2013). 2012 Financial Restatements: A Twelve Year Comparison. *Audit Analytics*.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT press.