Financial Statement Errors:
Evidence from the Distributional Properties of Financial Statement Numbers

Dan Amiram
Columbia Business School
Columbia University
da2477@columbia.edu

Zahn Bozanic*
Fisher College of Business
The Ohio State University
bozanic.1@fisher.osu.edu

Ethan Rouen
Columbia Business School
Columbia University
ethanrouen@columbia.edu

December 2014

Abstract: Motivated by methods used to evaluate erroneous 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. Using numerical methods, we first demonstrate that errors in financial statement numbers increase the deviation from the theoretical distribution. We then corroborate this analysis using simulation analysis that reveals that the introduction of errors to reported revenue also increases the deviation. Next, we provide archival-empirical evidence that the measure captures financial statement data quality by showing that i) the measure is positively associated with commonly used accruals-based earnings management proxies, ii) firms just above the zero earnings benchmark have significantly higher deviation, and iii) the restated financial statements of misstating firms exhibit less divergence. Lastly, we highlight an advantage of the measure by investigating a question in a context that prior measures, due to their design, could not: whether equity market participants immediately impound the implications of erroneous financial statements into prices. Using an event study, relative to low error firms, we show that firms with high financial statement errors lose 1% of their equity value upon information release.

JEL Classification: M41
Keywords: Benford's Law, Earnings Management, Financial Statements Errors, Market Efficiency

* Corresponding author. We would like to thank Anil Arya, Rob Bloomfield, Qiang Cheng, Patty Dechow, Ilia Dichev, Dick Dietrich, Peter Easton, Ken French, Jon Glover, Trevor Harris, Colleen Honigsberg, Gur Huberman, Amy Hutton, Bret Johnson, Steve Kachelmeier, Alon Kalay, Bin Ke, Bill Kinney, Alastair Lawrence, 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 Sloan, Steve Smith, Alireza Tahbaz-Salehi, 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, the London Trans-Atlantic Doctoral Conference, Nanyang Technological University, Singapore Management University, Syracuse University, UC Berkeley, UCLA, UNC, and The University of Texas - Austin for their helpful comments and suggestions. We would also like to thank the PCAOB and the SEC for their insights. All errors are our own.

1. Introduction

Financial statement data that is free of error—whether in the form of mistakes, biases, or manipulation—is crucial for well-functioning capital markets. Accurate financial reports enable efficient resource allocation and efficient contracting. 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 significant deficiencies that limit their usefulness. In this study, 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 that can be applied by financial statement users to assess the level of error in financial statements which overcomes some of the concerns surrounding existing measures. We first provide intuition for the mathematical and statistical mechanics behind the measure and then proceed to validate it in several empirical contexts. To highlight an advantage of the measure over prior measures, we apply it in a setting to answer a question that prior measures, dues to the nature of their design, could not. Namely, using event study methodology, we examine whether equity market investors immediately impound the implications of financial statement errors into prices upon the release of financial statement information.

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, Ge, and Schrand, 2010; Owens, Wu, and Zimmerman, 2013). In parallel, 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 errors of 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). Methods based on the law have been used to detect errors in published scientific studies, questionable election data in Iran, suspicious macroeconomic data, accounts receivables errors, and tax returns misreporting. 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 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. The first fact relies on 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 .176 or 17.6%), 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

[1] Distributions need to be non-truncated or uncensored in order 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.

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 variances of the central limit theorem, datasets tend to follow Benford's Law (Pimbley, 2014). In order 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.

The same 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 an unknown random distribution. 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), it is possible that the mixture distribution of the cash flows realization of these data will follow the criteria in Hill's (1995) theorem, and therefore will be distributed according to Benford's Law. Specifically, it is possible that 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., 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 which introduces error into the financial statements, whether in the form of mistakes, biases, or manipulation.[2]

[2] 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, they should follow Benford's Law. However, if these estimates contain certain type of errors, as the error increases, they will likely diverge further from Benford's Law.

To provide intuition on the mathematical and statistical foundations of Benford's Law, 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 occurs because introducing different size errors to different items in the financial statements make the distribution less smooth and less symmetric, which, as noted above, is a condition for a distribution to follow Benford's Law.

From this analysis, we construct our measure, the Financial Statement Divergence Score, or FSD Score for short, which is 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 specified by Benford's Law. As 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*.[3]

After providing theoretical intuition for the measure, we 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 to accounting numbers are introduced, we introduce errors for a typical firm in our sample

[3] 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).

by randomly manipulating its revenue. In this simple simulation, we are able to demonstrate that the manipulation induces an increase in the FSD Score more than 87% of the time.

We next assess whether the realized 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.

Once initial conformity is established, we continue by examining the relation between Benford's Law and commonly-used measures of accruals-based earnings management, earnings manipulation, and real activities earnings management. We show that the FSD Score is significantly positively related with the Dechow-Dichev measure, the modified Jones model measure, and Beneish's M Score, which is consistent with the FSD Score capturing some of the underlying forces measured by those tools. However, the FSD Score is unrelated to proxies for real activities earnings management, which is consistent with the claim that our measure is unrelated to firms' underlying operating environment and economic performance. We also corroborate the preceding analysis by investigating the FSD Scores of firms reporting annual income near zero in the spirt 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 next expand our validation of the FSD Score's ability to reflect financial statement error by conducting an "experiment" to directly examine our conjecture. Specifically, we identify a sample of firms that restated their financial statements and compare the FSD Score for the restated and unrestated numbers. Because of its similarity to a natural experiment, 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 as compared to 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.

We conclude by applying the measure to answer one out of presumably many important research questions in a setting for which its advantages over alternate measures allows us to gain new insights. Namely, we examine whether equity market participants immediately impound the implications of erroneous financial statements into prices at the time of information release. This question has been the subject of debate in the literature and significant disagreement has arisen among practitioners, regulators, and academics (e.g., Dechow and Skinner, 2000). On the one hand, there is a stream of theoretical and empirical research that suggests that markets are efficient and should adjust for any observable error. On the other hand, there is another stream of research that claims that markets do not always take into account the properties of financial statements (e.g., Sloan, 1996). Given its timeliness, non-reliance on cross-sectional or time-series data, and that it is not likely to be correlated with a firm's performance or business model, the FSD Score is well suited to examine this question and setting in that it allows researchers to assess how the market reacts to erroneous financial statement data upon the release of the report.

We find that buying (selling) a portfolio of the lowest (highest) quintile of FSD firms yields abnormal return of 0.3% (0.6%) in the 10 days following the earnings announcement. When we double sort the portfolios to quintiles first based on earnings surprise and then based on the FSD Score, we find that for the most negative (positive) earnings surprises, high FSD firms have 1.3% (1.3%) lower abnormal returns than low FSD firms in the 10 days following the earnings announcement. This result is inconsistent with the argument that FSD returns are driven by the drift of the earnings surprise. Lastly, we show that there are similar abnormal returns around the 10-K release date, which is consistent with investor learning. However, we do not find any abnormal returns starting 10 days after the 10-K release date, suggesting that the market impounds information related to financial statement errors in a timely manner.

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. In Section 4, we detail how Benford's Law can be used to detect errors in financial statements and form our predictions. In Section 5, we present our sample and descriptive statistics. Finally, Section 6 presents our empirical findings prior to concluding in Section 7.

2. Motivation and Contribution

The level of errors in financial statement data has a first-order impact in capital markets (Bushman and Smith, 2003). Existing 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, 2003; La Porta, Lopez-De-Silanes, Shleifer, and Vishny, 2000; Duffie and Lando, 2001; Francis, LaFond, Olsson, and Schipper, 2004; Francis, LaFond, Olsson, and Schipper, 2005). 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. It also does not model the error as a residual from a prediction model. Estimating residuals in time-series or cross-sectional models (such as 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 changes will also change the estimated financial statement error. This issue may create estimation error and lack of timelines. More importantly, 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., the Dechow and Dichev (2002) model), 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 or for 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. Fifth, 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. Sixth, the measure is scale independent and thus fits to every currency or size. Seventh, it is available to essentially every firm with accounting information, even private companies where such information exists.

We do not claim to be 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 U.S. 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 round reference point, they focus strictly on the distribution of the second digit of the distribution of earnings and do not make firm-year inferences.[4] 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 various 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 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, Hillison, and Pacini (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.

In light of prior literature, we are unaware of any large-scale application of Benford's Law 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, therefore, is unique in its ability to

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

explore and answer questions distinct from prior literature.[5] Importantly, unlike extant prior literature, which has relied largely on internal and private data, our measure can be created using solely publicly available information, making it available to anyone interested in analyzing the level of errors in firm-year financial data.

We also contribute to the literature by examining whether equity market participants can detect and impound the implications of errors in the financial statements. For the market to detect errors, those errors need to be observable to the market at relatively low cost and market participants need to understand their implications. A short-window event study is a useful way to examine this question. However, available measures of the level of error in financial statements lack the timeliness, data availability, and are more likely to be correlated with firm performance. Using the newly developed measure, we can provide new evidence on this question and setting.

3. Foundations of Benford's Law

3.1 Historical 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:

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

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.

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

[5] 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, no commercial auditing software computes the conformity of the entire distribution of first digits, nor assesses firm-year conformity from external corporate financial reports.

an issue of *Reader's Digest,* and found that the law held in each 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 significant digits of the combined mixture distribution will converge to the logarithmic or Benford distribution.

In order for a distribution to deviate from Benford's Law, certain types of errors need to be introduced. For example, evidence suggests that stock indices' returns conform to Benford's Law (Ley, 1996), which allows 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, 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 A 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.[6]

The KS statistic becomes less useful as P increases, however. In order 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 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 to compare 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.[7]

3.3 Understanding the prevalence of Benford's Law

---

[6] An alternate method to examine conformity relies on the expected distribution of the first two digits (from 10 to 99) of a number (Nigrini, 2006). We are unable to employ the first two digits in our setting since the pool of digits required to generate the distribution is 90 (instead of 9) buckets.

[7] While two other statistics were widely used in the early stages of the forensic accounting literature in this area, the Z-statistic and the Chi-square statistic, researchers have progressed to using the MAD statistic (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.

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) to log(d), where d is an integer between 1 to 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 theorem, are smooth and symmetric in the log scale because of the central limit theorem, the probability of being in a region between n+log (d+1) to 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.

3.3.1 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, 2007; 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, i.e., 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 what determines the first digit of a number is 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 to 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. Further, 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.

3.3.2 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:

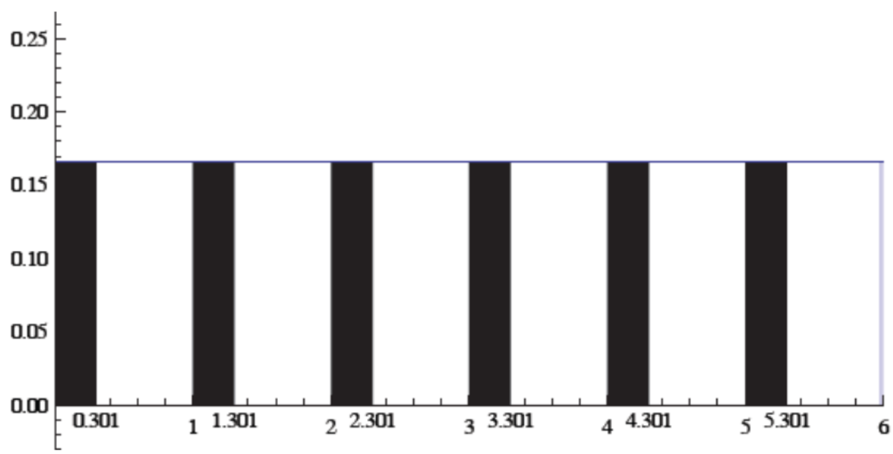

Figure 1

The solid black bars in Figure 1 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:

$$\sum_{n=0}^{5} \int_{n}^{n+\log(2)} \frac{1}{6} dN$$

= 1/6 * (0.301-0) + 1/6 * (1.301-1) + 1/6 * (2.301-2) + 1/6 * (3.301-3) + 1/6 * (4.301-4) + 1/6 * (5.301-5)

= 0.301 = **log(2) – log(1)**

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 the 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:

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

= **log(d+1) – log(d)**

3.3.3 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. So long as these distributions are spread across a few orders of magnitudes in the log scale (e.g., range between 2 to 4 in the log-scale, which means 100 to 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.

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

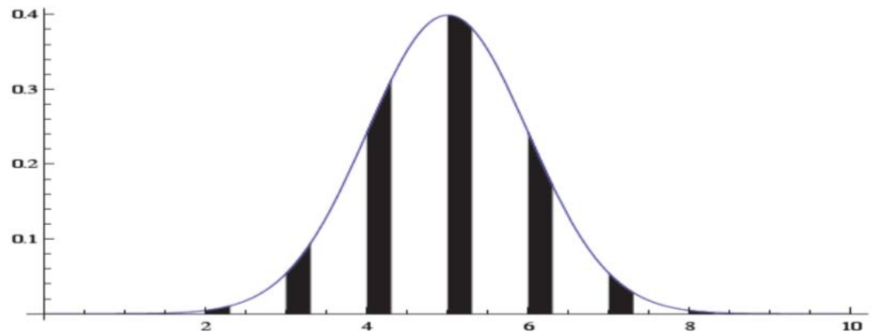

Figure 2

The shaded area in Figure 2 represents all the areas between any integer *n* to *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 =

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

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

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

3.3.4 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, 2007; 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.

3.3.5 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, 2007). A natural measure to examine the distance all leading digits are 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. With this knowledge, we can mathematically construct our Financial Statement Deviation (FSD) Score based on the Mean Absolute Deviation (MAD) statistic:

$$\text{FSD Score} = \frac{\sum_{d=1}^{9} ABS[(\sum_{n=-\infty}^{\infty} \int_{n+\log(d)}^{n+\log(d+1)} PDF(\log(N))dN) - (Log(d+1) - Log(d))]}{9}$$

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 zero.

4. Detecting errors in financial reporting using Benford's Law

4.1 Motivating Intuition

Since accounting data are a series of estimations of the true cash flows realizations of the underlying items (for example, cash flows from sales, cash flows from payments to employees, etc.), the resulting underlying distribution of the mixture of distributions of these cash flows realizations may fulfil the conditions of Hill's theorem and follow Benford's Law. In our example below, we show numerically that under certain assumptions, this is the case with accounting data. We also show that if the accounting estimates of the true cash flows realization are without error, the distribution of the accounting estimates (the financial statements) will follow Benford's Law exactly. While we cannot prove or empirically show that that actual cash flow realizations of accounting data will follow Benford's Law (as they are unobservable), we show 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).

Using numerical methods, we then characterize the type of errors in accounting data that are likely to create deviations from Benford's Law. We show that introducing a zero mean error or multiplying all numbers in the distribution by the same number will not create deviations from Benford's Law. However, if we introduce errors with non-zero mean to the 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, and/or meet-or-beat behavior are likely to introduce deviation from the law. 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 further intuition, we show in Appendix B that when we introduce errors into observable realizations of equity prices, they 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.

4.2 A stylized numerical example

In order 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 for suppliers), and X_k is the random inflow of cash from activity k. X_K is the last cash flow stream.[8] 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;

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

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 $\log(Y_k) = \log(X_k) + \log(Z_k)$, i.e., 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 non-zero 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 one 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 normal unless their means 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:

$$\text{PDF}(X) = \sum_{k=1}^{K} W_k * PDF(X_k),$$

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:

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

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

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

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 3 plots the empirical density function of all numbers from all financial statements from 2001-2011 in the log scale, which suggests that the underlying no-error distribution follows Benford's Law as well. Figure 4 shows the distribution in the log scale for a typical firm, Alcoa in 2011.

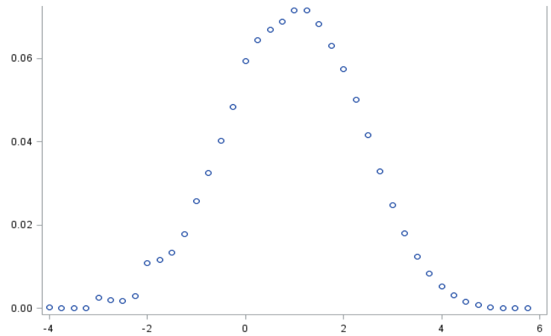

Figure 3

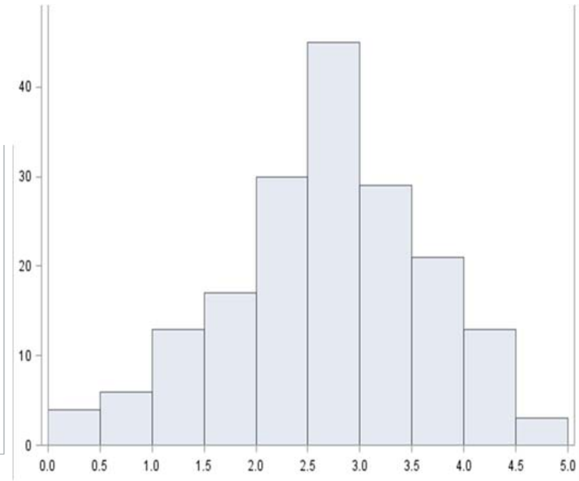

Figure 4

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 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.

4.3 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, i.e., 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 Figure 5, this distribution is smooth and symmetric and looks similar to a normal distribution:

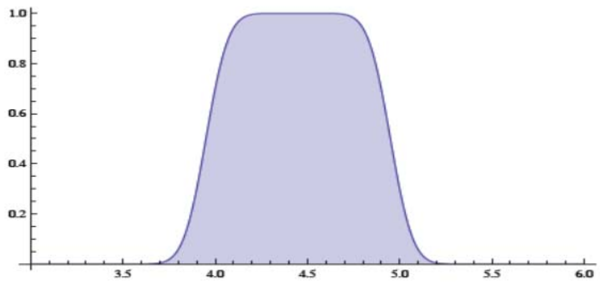

Figure 5

Further, 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.[9] 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.

4.3.1 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, 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

[9] 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.

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 Figure 6:

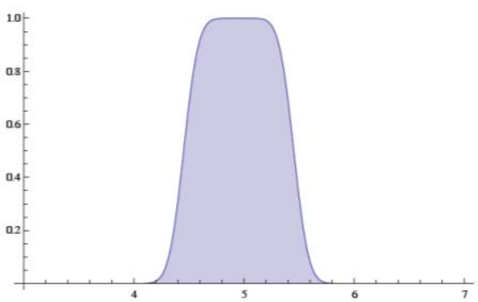

Figure 6

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

4.3.2 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.

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

In this case, $\mu_{\varepsilon k}$ varies across the ks, $\sigma_{\varepsilon k}$ 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.

4.3.3.1 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 Figure 7. 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 Figure 8.

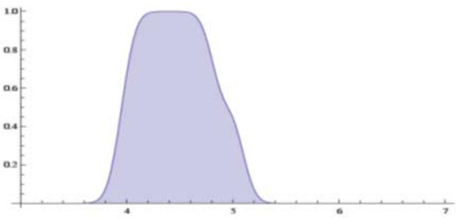

Figure 7

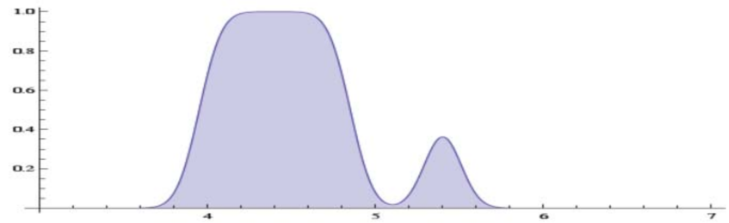

Figure 8

As we increase the mean of the error, 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.

4.3.3.2 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 Figure 9. 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 Figure 10.

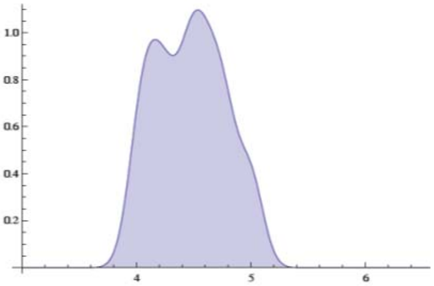

Figure 9

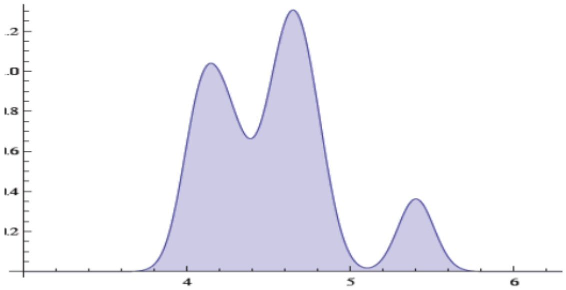

Figure 10

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

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

It also possible that the estimation errors will 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 Figure 11. 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 Figure 12.

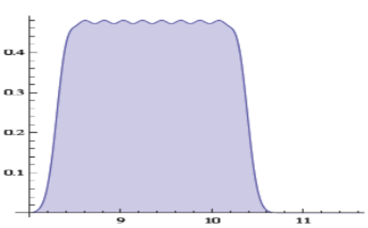

Figure 11

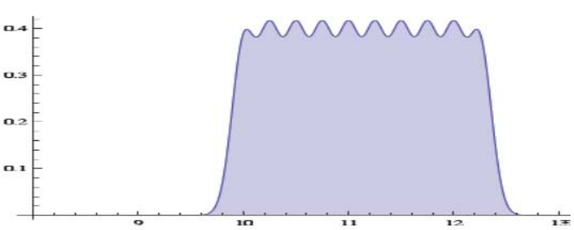

Figure 12

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

4.3.4 Summary of the insights from the numerical results

We have showed that, under certain parameters, the FSD Score (or the deviation from Benford's Law) 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. In reality, it is unlikely that the errors will be identical in all line items in the financial statements. Our numerical analysis indicates that, for example, overestimating revenue by itself (Case 3A) or together with the associated cost of goods sold (Case 3B) will create deviations from Benford's Law. Further, an error that is correlated with the size of the item (Case 3C) will create deviations in the financial statements.

4.4 Simulation analysis

To further demonstrate how errors could alter conformity to Benford's Law, we ran a simple 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. Because we need a firm that is unlikely to have an already manipulated financial statement, we chose to manipulate sales for Alcoa's 2011 financial statements, which is a firm that generally, but not perfectly, conforms to Benford's Law. We chose to 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, consistent with Case 3B, 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

These three journal entries affect more than 30 line items in Alcoa's financial reports (see Appendix C for further detail). We then re-measured the FSD Score based on the manipulation and the changes the manipulation induced in the financial statements. The results of this simulation show that the random revenue manipulation increased the FSD Score 87% 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.

5. Sample selection, variable measurement, and descriptive statistics

5.1 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.[10] For variables reported with an absolute value of less than 1, we take the first non-zero digit. We set missing variables to 0, as this process do not affect our calculations of the FSD Score, which requires only digits 1 through 9. We remove any firm-years from the sample where the total number of first digits used to calculate the FSD Score for a given firm-year is less than 100 in order to increase the power of the test statistics.[11] We also remove firms with negative total assets. All non-indicator control variables in the total sample of 46,674 firm-years are then winsorized at the 1% and 99% levels to eliminate the influence of outliers. See Appendix D for further details, as well as for the definitions of the control variables.

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

[10] We do, however, exclude data items provided by Compustat that do not appear on firms' financial statements, i.e., price data. Further 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.

[11] Including firm-years with less than 100 first digits does not alter our results.

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, earnings manipulation, and real activities earnings management. For accruals-based earnings management, we calculate the five-year moving standard deviation of the Dechow-Dichev residual (STD_DD_RESID) from Dechow and Dichev (2002), as suggested by Kothari, Leone, and Wasley (2005), and the absolute value of the accruals quality residual (ABS_JONES_RESID) from the modified Jones model (Jones, 1991), as suggested by Francis 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. For real activities earnings management, we calculate three measures of real activities, abnormal level of cash flows from operations (R_CFO), abnormal level of production costs (R_PROD), and abnormal level of discretionary expenses (R_DISX), as defined in Roychowdhury (2006) and following Cohen, Dey, and Lis (2008).

In terms of 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.[12] RESTATED_NUMS is equal to 1 if the reported numbers are restated and zero if the numbers are what was originally reported. Returns data from CRSP is used to create our abnormal returns measures. ABN_RET_EARN is the 10-day market-adjusted abnormal return

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

starting from the earnings release date. ABN_RET_10K is the 10-day market-adjusted abnormal return starting from the 10-K release date. ABN_RET_LONG is the market-adjusted abnormal return from 11 days after the 10-K release date to 90 days after the release date.

5.2 Descriptive statistics

Table 1 provides descriptive statistics for the full sample of firms from 2001-2011. The FSD Score's mean is 0.030 with a standard deviation of 0.009. Table 2 presents Spearman correlations above the diagonal and Pearson correlations below the diagonal. The Pearson correlations between the FSD Score and ABN_RET_EARN and ABN_RET_10K are -0.0248 and -0.0206, respectively, and both are significant at the 5% level. The relations are negative and significant at the 5% level (-0.0325 and -0.0238, respectively) with the Spearman calculations as well. We find positive correlations that are significant at the 5% level when examining the relation between the FSD Score and our measures of accounting quality as well. The Pearson correlations between FSD Score and ABS_JONES_RESID, STD_DD_RESID, and MANIPULATOR are 0.0726, 0.1486, and 0.0638, respectively. The Spearman correlations are similar, with significant (at the 5% level) correlations of 0.0701, 0.1083, and 0.0521 for the above relations. 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.

6. Methodology and empirical results

6.1 Investigating the distribution of first digits in financial reports

Table 3 shows 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. In the aggregate, the FSD Score is 0.0014, well below 0.006, which can be considered close conformity to the law in very large samples (Nigrini, 2012).[13] This result can also be seen graphically in Figure 13 in Appendix E. Panels B and C of Table 3 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 4 examines individual firm-year conformity to Benford's Law. Here, we must use the FSD Score based on the KS statistic because it enables us to assess whether the financial statements for a given firm-year adhere to the law. Of the 46,674 firm-years in our sample, 38,983, or 84%, conform to the law at the 5% level or better, as shown in Panel A.[14] Figure 14 in Appendix E provides examples of the empirical distributions for two firm-years, one that conforms to Benford's distribution at the 5% level (AT&T, 2003) and one that does not conform (Sprint Nextel, 2001). While there are some kinks in AT&T's distribution, the overall divergence from Benford's distribution is visually apparent for Sprint Nextel, which experienced a restatement for, amongst other things, understating interest expense. Panel B of Table 4 shows similar results when firms are sorted by industry, with a minimum conformity of 79% of all firms in a given industry and a maximum conformity of 87%. Panel C shows similar results when firms are sorted by fiscal year, with all years exhibiting between 82% and 85% conformity. This table supports our conjecture that a significant majority of firm-year empirical distributions conform to Benford's Law. 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.

[13] 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.

[14] 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.

6.2 Benford's Law and existing measures of accounting quality

To understand what types of firm behavior are associated with the FSD Score, we examine the relation between it and known measures of accounting errors. Table 5 examines the relation between the FSD Score and proxies for accruals-based earnings management, earnings manipulation, and real activities earnings management. The FSD Score is significantly positively associated with two commonly used measures of accruals-based earnings management, ABS_JONES_RESID and STD_DD_RESID. The coefficients on these variables are 0.001 (significant at the 5% level) and 0.006 (significant at the 1% level), respectively. The coefficient on MANIPULATOR, our indicator that signals firms may be manipulating their earnings, is 0.001 and significant at the 1% level. Finally, the coefficients on our three measures of real activities earnings management – abnormal levels of cash flow from operations (R_CFO), abnormal levels of production costs (R_PROD), and abnormal levels of discretionary expenses (R_DISX) – are all insignificant. These results suggest that the FSD Score is more likely to be associated with accruals-based earnings management than real activities that change the actual cash flow realizations and the accounting numbers accordingly. These results also provide support for our claim that the FSD Score is not likely to be correlated with firm operating performance.

6.3 Benford's Law and earnings management around zero

To understand whether Benford's Law captures firms that manipulate their accounting results, in Table 6 we examine the difference in the FSD Scores for firms that are just above and just below the zero-earnings threshold. Burgstahler and Dichev (1997) find inconsistencies in the distribution of net income for firms that are around this threshold, with fewer firms just below zero than expected and more firms just above zero, suggesting that earnings management is used to avoid reporting losses. Following Burgstahler and Dichev (1997), we scale net income by the beginning-of-year market value of equity for all firms. We then divide firms into scaled net income bins of 0.005.[15] As compared to the firms just below zero, we find 38 percent more firms in the bin just above zero, and that these firms have FSD Scores that are, on average, 4.5 percent higher as well. This difference is statistically significant at the 1% level, which suggests that firms that are more likely to have manipulated their earnings have greater divergence from Benford's Law.[16]

6.4 Misstated versus restated financial statements

We have established that the empirical distributions of most firms' financial results conform to Benford's Law, revenue manipulations tend to increase the divergence from Benford's Law, and the FSD score is related to earnings management. When firms restate their financial results, the empirical distribution of the restated results should more closely conform to Benford's Law than the misstated results. To test this prediction, we investigate a sample of firms that have restated their financial results and compare the FSD Scores of the misstated financial results with those of the restated results. Consequently, we expect that the FSD Score will decrease, or more closely conform to Benford's Law, for the restated results.

To conduct our test, we examine firm-years in Compustat from 2001-2011 where both misstated and restated financial results are available (in Compustat, datafmt=STD for original and datafmt=SUMM_STD for restated). To increase the materiality of the restatement, we require that at least 10 variables change between the unrestated and restated numbers. 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. We regress this indicator variable on the FSD Score and include several variables to control for accruals-based earnings management, earnings

[15] We conduct similar analysis using bins of size 0.001 and 0.0025 and find similar results.
[16] We corroborate the preceding analysis by examining the relation between the FSD Score and earnings persistence. Richardson, Sloan, Soliman, and Tuna (2005) show that current earnings are less informative about future earnings for firms with low accounting quality. Consistent with the FSD Score capturing accounting quality, in untabulated results, we find that the interaction between contemporaneous net income and the FSD Score is significantly negatively related to one-year-ahead net income, suggesting that firms with greater divergence from Benford's Law have lower earnings persistence.

manipulation, and real activities earnings management. Since the regression compares the firm to itself, we do not include additional firm control variables in this specification.

Table 7 presents the results of our test of our prediction that the conformity of firms' annual financial reports to Benford's Law is lower for misstated versus restated results in the same firm-year. Consistent with this prediction, the coefficient on RESTATED_NUMS in Column (1) is -0.001, which is statistically significant at the 1% level. To ensure that our measure of conformity to Benford's Law isn't merely a proxy for existing measures of accounting discretion, in Column (2) we control for accruals-based earnings management, earnings manipulation, and real activities earnings management. When adding these additional measures, we find similar results, with the coefficient on RESTATED_NUMS equal to -0.001 and significant at the 1% level. Consequently, for the sample of firms that have both original and restated financial results available through Compustat from 2001-2011, the FSD Score is lower for the restated 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. As our result is incremental to standard accounting quality proxies, this result also implies that our measure, while somewhat correlated with these proxies, is distinct from it. Consequently, our measure of conformity may be useful in augmenting existing accounting quality models.[17]

6.5 Event study examining equity market returns of portfolios based on the FSD Score

[17] We chose not to tabulate and discuss results from a prediction model of SEC Accounting and Auditing Enforcement Releases (AAER) due to research design issues. In brief, the issues, identified at length by prior literature (see, for example, Dechow et a. (2011)), include, but are not limited to: i) the existence of firms that manipulate but do not get caught, ii) the inability to model the SEC decision to issue AAERs to certain firms and not others, iii) statute of limitation laws that prevent us from identifying when the manipulation began, and iv) problems with the dataset (Karpoff, Koester, Lee, and Martin, 2014). While we plan to address these challenges in a future study, we do find that the lagged FSD Score and change in FSD Score from year t-1 to t are predictive of AAERs.

Table 8 provides results on the relation between firms' level of conformity to Benford's Law and equity market returns. Portfolio returns are based on the Fama-Macbeth method, i.e., we construct portfolios every year and report the mean of these portfolios.

In Panel A, we classify firms into quintiles based on their FSD Scores corresponding to the annual earnings release date. We find that in the 10 days following the annual earnings announcement date, the abnormal return of buying the lowest FSD Score quintile of firms is 0.33%, while the abnormal return of shorting the highest FSD Score quintile of firms is 0.62%. This trading strategy yields abnormal returns of 1.0% in 10 days.[18] This result suggests that the market punishes firms with high probability of erroneous data and rewards firms with low probability.

In Panel B we take into account the possibility that the FSD Score is correlated with the earnings surprise. To deal with this concern, we first sort the data into five portfolios based on unexpected earnings (UE) defined as the difference between actual earnings per share and the mean of each analyst's latest forecasts taken from the IBES Detail dataset, scaled by price. We then sort each of these portfolios into five portfolios based on the FSD Score. We find that, within the lowest quintile of earnings surprise, buying the lowest FSD Score quintile of firms and shorting the highest FSD Score quintile of firms in the 10 days around the earnings announcement yields a return of 1.3%. The results also show that within the quintile of highest earnings surprises, buying the lowest FSD Score quintile of firms and shorting the highest FSD Score quintile of firms in the 10 days around the earnings announcement yields 1.3% return. Taken together, the

[18] There are two concerns with this classification scheme. First, it may be that some firms do not provide all financial statements at the earnings release date. In this case, our classification will also suffer from look-ahead bias. While evidence suggests that this is not the case for many firms, in our tests, we provide similar evidence using the 10-K release date, as opposed to the earnings release date. Second, in order to classify firms, we need to know where they rank in the distribution of FSD compared to other firms-years. Technically, for this classification we need for all firms to report before classification. Practically, however, since the thresholds of classifying firms into quintiles based on the FSD Score do not vary significantly over the years, this is not a significant issue. Nevertheless, when we use prior distributions to determine whether the current year's FSD Score is high or low, the results are unchanged.

results from Panels A and B suggest that the market punishes firms with higher probability of error, regardless of whether they disclose good news or bad news. These results also reduce the possibility that our measure captures post-earnings announcement drift.

In Panel C, we repeat the analysis in Panel A, but here abnormal returns are formed from the 10-K release date to 10 days after that date. We show that in the 10 days following the 10-K release date, the abnormal return of buying the lowest quintile of FSD Score firms is 0.19% while the abnormal return of shorting the highest FSD Score firm is 0.49%. This trading strategy can yield 0.68% of abnormal returns in 10 days. This result suggests that the market learns additional information beyond the annual earnings announcement about the probability of errors in the financial statements from the 10-K disclosure.

In Panel D we show that, starting 10 days after the 10-K release date, significant abnormal returns based on FSD Score quintiles disappear (0.53% for 80 days). This suggests that the market impounds all information in the days following the information release. Given the lack of persistence in returns, this result also suggests that the FSD Score is unlikely to be correlated with risk factors.

Lastly, in Panel E we show using multivariate analysis that the FSD Score is correlated with earnings announcement returns incremental to available accounting quality measures. Column (1) presents the results with no control variables. The coefficient on FSD is -0.412 and is significant at the 1% level. Column (2) presents the results after controlling for the earnings surprise, available measures of financial statement errors, and firm performance measures. The coefficient on FSD is -0.286 and is significant at the 1% level. The negligible reduction in the coefficient size and statistical significance suggest that the measure has incremental explanatory power over available measures of financial statement errors.

The results from Table 8 collectively suggest that market participants punish (reward) firms with high (low) probability of errors in their financial statements. The market reacts on a timely basis, with most of the abnormal returns concentrated around information release dates.[19] The results provide evidence of the ability of market participants to detect financial statement errors and price protect against their implications.

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 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 other measures of accounting quality that are 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.

Based on Hill's theorem, we construct several scenarios using numerical methods which reveal that financial statements without error are distributed according to Benford's Law and, the larger the error, the larger the deviation from the Benford's Law. To corroborate the results from the numerical analysis, we provide a simple simulation to demonstrate that when accounting

[19] The seemingly efficient market reaction to financial statements errors begs the question, why do firms have errors in the first place? While we do not take a position on this question, we suggest a few possible answers. First, the errors may be accidental. Second, firms may attempt to bias or manipulate financial statements to deceive parties that are not as efficient as equity markets (such as regulators, auditors, compensation committees, etc.). Third, certain parties, such as debt holders, may expect managers to manage their financial statements. Therefore, managers have to manipulate their financial statement because these manipulations are already taken into account by these parties (for this signal jamming model we refer the reader to Stein (1989)).

numbers are manipulated, there is a high likelihood of an increase in the divergence from Benford's Law.

To ascertain 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 84% of firm-years conform to the law as well. Next, we find that proxies for accruals-based earnings management and earnings manipulation are related to divergence from Benford's Law while proxies for real activities earnings management are not. We corroborate this analysis by finding that firms more likely to manage their earnings have significantly lower FSD Scores. We then show that when restatements occur, the restated numbers are significantly closer to Benford's Law relative to the misstated numbers.

Lastly, we provide evidence on whether market participants can impound the possibility of errors in financial statements into prices in a timely manner. Our evidence shows that the market punishes firms for erroneous financial reporting, regardless of whether the announcement contains good or bad news. We find that the market response is concentrated in the days around the earnings announcement and the 10-K release date. However, the short-window market response does not persist in longer windows. These results support the argument that the market is efficient in detecting observable errors.

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, our paper provides the investment community (investors, regulators, auditors, and researchers) with an easily implementable, parsimonious approach for assessing errors in financial reports.

References

Beneish, M. "The Detection of Earnings Manipulation." *Financial Analyst Journal* 55 (1999): 24-36.

Benford, F. "The Law of Anomalous Numbers." *Proceedings of the American Philosophical Society* 78 (1938): 551-572.

Blodget, H. "Bernie Madoff's Miraculous Returns: Month By Month." *Business Insider*, December 12, 2008, accessed August 10, 2013.

Burgstahler, D. and I. Dichev. "Earnings Management to Avoid Earnings Decreases and Losses." *Journal of Accounting and Economics* 24(1) (1997): 99–126.

Bushman, R. and A. Smith. "Transparency, Financial Accounting Information, and Corporate Governance." *Federal Reserve Bank of New York Policy Review* (April 2003): 65-87.

Carslaw, C. "Anomalies in Income Numbers: Evidence of Goal Oriented Behavior." *The Accounting Review* LXIII(2) (1988): 321-327.

Cohen, D.; A. Dey; and T. Lys. "Real and Accrual-Based Earnings Management in the Pre- and Post-Sarbanes Oxley Periods." *The Accounting Review* 83(3) 2008: 757-787.

Dechow, P., and I. Dichev. "The Quality of Accruals and Earnings: The Role of Accrual Estimation Errors." *The Accounting Review* 77 (Supplement): 35-59.

Dechow, P.; W. Ge; C. Larson; and R. Sloan. "Predicting Material Accounting Misstatements." *Contemporary Accounting Research* 28 (1) (2011): 17-82.

Dechow, P.; W. Ge; and C. Schrand. "Understanding Earnings Quality: A Review of the Proxies, their Determinants and their Consequences." *Journal of Accounting and Economics* 50 (2010): 344-401.

Dechow, P. and D. Skinner. "Earnings Management: Reconciling the Views of Accounting Academics, Practitioners, and Regulators." *Accounting Horizons* 14(2) (2000): 235-250.

Duffie, D. and D. Lando. "Term Structures of Credit Spreads with Incomplete Accounting Information." *Econometrica* 69(3) (2001): 633-664.

Durtschi, C. and P. Easton. "Earnings Management? Erroneous Inferences Based on Earnings Frequency Distributions." *Journal of Accounting Research* 47(5) (2009): 1249–1281.

Durtschi, C.; W. Hillison; and C. Pacini. "The Effective Use of Benford's Law to Assist in the Detecting of Fraud in Accounting Data." *Journal of Forensic Accounting* V (2004): 17-34.

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

Fama, E., and E. French. "Permanent and Temporary Components of Stock Price." *Journal of Political Economy* 96 (1988): 246-273.

Francis, J.; R. LaFond; P. Olsson; and K. Schipper. "Cost of Equity and Earnings Attributes." *Journal of Accounting Research* 79 (2003): 967–1010.

Francis, J.; R. LaFond; P. Olsson; and K. Schipper. "The Market Pricing of Accruals Quality." *Journal of Accounting and Economics* 39 (2005): 295–327.

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

Hill, T. "A Statistical Derivation of the Significant Digit Law." *Statistical Science* 10 (1995): 354-363.

Hill, T. "A Note on the Distribution of True versus Fabricated Data." *Perceptual and Motor Skills* 83 (1996): 776-778.

Jones, J. "Earnings Management During Import Relief Investigations." *Journal of Accounting Research* 29 (1991): 193-228.

Karpoff, J.; A. Koester; D. Scott Lee; and G. Martin. "Database Challenges in Financial Misconduct Research." *Working paper* (2014).

Kothari, S.; A. Leone; and C. Wasley. "Performance Matched Discretionary Accrual Measures." *Journal of Accounting and Economics* 39 (2005): 163-197.

La Porta, R.; F. Lopez-De-Silanes; A. Shleifer; and R. Vishny. "Legal Determinants of External Capital." *Journal of Finance* 55(1) (2000): 1-33.

Ley, E. "On the Peculiar Distribution of the US Stock Indexes' Digits." *The American Statistician* 50.4 (1996): 311-313.

Michalski, T., and G. Stoltz. "Do Countries Falsify Economic Data Strategically? Some Evidence that they Might." *The Review of Economics and Statistics* 95(2) 2013: 591-616.

Morrow, J. "Benford's Law, Families of Distributions and a Test Basis." *Working paper* (2010).

Nigrini, M. "Taxpayer Compliance Application of Benford's Law." *Journal of American Taxation Association* 18(1) (1996): 72-92.

Nigrini, M. *Benford's Law: Applications for Forensic Accounting, Auditing, and Fraud Detection* (Hoboken, N.J.: John Wiley & Sons, 2012).

Owens, E.; J. Wu; and J. Zimmerman. "Business Model Shocks and Abnormal Accrual Models." *Working paper* (2013).

Pimbley, J.M. "Benford's Law and the Risk of Financial Fraud." *Risk Professional* (May 2014): 1-7.

Pike, D. "Testing for the Benford Property." *Working paper* (2008).

Rajan, R. and L. Zingales. "Financial Dependence and Growth." *American Economic Review* 88(3) (2003): 559-586.

Rajan, R. and L. Zingales. "The Great Reversals: The Politics of Financial Development in the Twentieth Century." *Journal of Financial Economics* 69(1) (2003): 5-50.

Ray, S., and B. Lindsay. "The Topography of Multivariate Normal Mixtures." *The Annals of Statistics* 33(5) (2005): 2042-2065.

Richardson, S.; R. Sloan; M. Soliman; and I. Tuna. "Accrual Reliability, Earnings Persistence and Stock Prices." *Journal of Accounting and Economics* 39 (2005): 437-485.

Rouen, E. "Machine Illegible: Evidence of Firms' Strategic Use of Regulatory Technology." Working paper (2014).

Roychowdhury, S. "Earnings Management through Real Activities Manipulation." *Journal of Accounting and Economics* 42 (2006): 335-370.

Sloan, R. "Do Stock Prices Fully Reflect Information in Accruals and Cash Flows about Future Earnings?" *The Accounting Review* 71 (1996): 289-315.

Smith, S. *The Scientist and Engineer's Guide to Digital Signal Processing* (California Technical Publications, 1997).

Stein, J. "Efficient Capital Markets, Inefficient Firms: A Model of Myopic Corporate Behavior." *The Quarterly Journal of Economics* (1989): 655-669.

Thomas, J. "Unusual Patterns in Reported Earnings." *The Accounting Review* LXIV(4) (1989): 773-787.

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

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

APPENDIX A: How to calculate conformity to Benford's Law, an 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 1's frequency is 8/28=.2857.

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

Digit	1	2	3	4	5	6	7	8	9
Total occurences	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:

$KS=Max(|AD_1-ED_1|, |(AD_1+AD_2)-(ED_1+ED_2)|, \ldots, |(AD_1+AD_2+\ldots+AD_9)-(ED_1+ED_2+\ldots+ED_9)|$

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

In this example,

Max(|0.2857-0.3010|, |(0.2857+0.1786)-(0.3010+0.1761)|, …,
(|(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

To test conformity to Benford's distribution at the 5% level based on the KS statistic, the test value is calculated as $1.36/\sqrt{P}$, 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:

$MAD=(\sum_{i=1}^{K}|AD-ED|)/K$, 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 B

Numerical example when realizations are observable

To see the intuition for why deviations from Benford's Law can be used to assess data quality, 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 two 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.[20] 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.

[20] Not all misestimated or fabricated data create deviations from Benford's Law. For example, if the mis-estimation simply multiplies all true realizations by a constant, the new erroneous data will still follow 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 consistent and manipulate the variance, expecting the FSD Score to remain constant since we are no longer changing the magnitude of the noise.

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 on the behalf of 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 C: 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) increased sales by between 5% and 50% to make the change material. COGS were increased by between 20% and 90% of the manipulated sales, 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 1,000 times, Alcoa's FSD Score increases 870 times (87%). 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 D: 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 A for a sample calculation of MAD.
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 A for a sample calculation of KS.
ABN_RET_EARN	Short-window abnormal return after the release date of a firm's annual earnings	The 10-day market-adjusted abnormal return is measured from the release date of a firm's annual earnings to 10 days after the release. Release dates are taken from Compustat. Returns information is retrieved from CRSP.
ABN_RET_10K	Short-window abnormal return after the release date of a firm's 10-K	The 10-day market-adjusted abnormal return is measured from the release date of a firm's 10-K to 10 days after the release. Release dates are taken from Compustat. Returns information is retrieved from CRSP.
ABN_RET_LONG	Long-window abnormal return after the release date of a firm's 10-K	The market-adjusted abnormal return from 11 days after the release date of a firm's 10-K to 90 days after the release date. Release dates are taken from Compustat. Returns information is retrieved from CRSP.
UE	Unexpected earnings	Analyst earnings per share forecast errors proxy for earnings surprise and are taken from the IBES Details file. Forecast errors are calculated as actual earnings per share minus the mean of the last forecast of the period for every analyst reported by IBES, scaled by price.
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: tca = Δsales + net PPE + ROA, where tca = (Δcurrent assets - Δcash - Δcurrent liabilities + Δ 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: tca = cfo_{t-1} + cfo + cfo_{t+1}, where tca is defined as above, and cfo = (interest before extraordinary items - (wcacc - depreciation and amortization)). All variables are scaled by average total assets. The five-year rolling standard deviations of the residuals are then calculated.
INDUSTRY	Industry classification	Groups companies into 17 industry portfolios based on the Fama-French.
MANIPULATOR	Indicator variable equal to 1 if the M Score is greater than -1.78	M Score is calculated following Beneish (1999).
R_CFO	Level of abnormal cash flows from operations	Abnormal cash flows are measured as defined in Roychowdhury (2006) following Cohen et al. (2008).
R_PROD	Level of abnormal production costs	Abnormal production costs are measured as defined in Roychowdhury (2006) following Cohen et al. (2008).

R_DISX	Level of abnormal discretionary expenses	Abnormal discretionary expenses are measured as defined in Roychowdhury (2006) following Cohen et al. (2008).
RESTATED_NUMS	Indicator variable that equals 1 if reported numbers are restated	For all firms from 2001-2011 with at least 10 restated variables, where RESTATED=1 and both restated and original financial numbers are available in Compustat (datafmt=STD for original and datafmt=SUMM_STD for restated), we separate the original from the restated financial numbers and create an indicator equaling 1 for restated numbers.
WCACC	Working capital accruals	Calculated as (Δcurrent assets - Δcash - Δcurrent liabilities + Δ debt in current liabilities) scaled by average total assets.
CHCSALE	Change in cash sales	Cash sales $_{t}$ - cash sales $_{t-1}$/cash sales $_{t-1}$, where cash sales = total revenue - Δtotal receivables.
SOFTAT	Soft assets	(Total assets - net PPE - cash)/total assets $_{t-1.}$
ISSUANCE	Indicator variable that equals 1 if the company issued debt or equity in that year	When long-term debt issuance (Compustat DLTIS) > 1 or sale of common or preferred stock (SSTK) > 1, then issuance = 1.
BTM	Book-to-market	Total stock holders' equity (Compustat SEQ)/(closing price at the end of the fiscal year (Compustat PRCC_F) * common shares outstanding (Compustat CSHO).
AT	Total assets	Compustat AT.
ROA	Return on assets	Income before extraordinary items $_{t}$/total assets $_{t-1.}$

APPENDIX E

Figure 13: Aggregate Distribution and Benford's Distribution

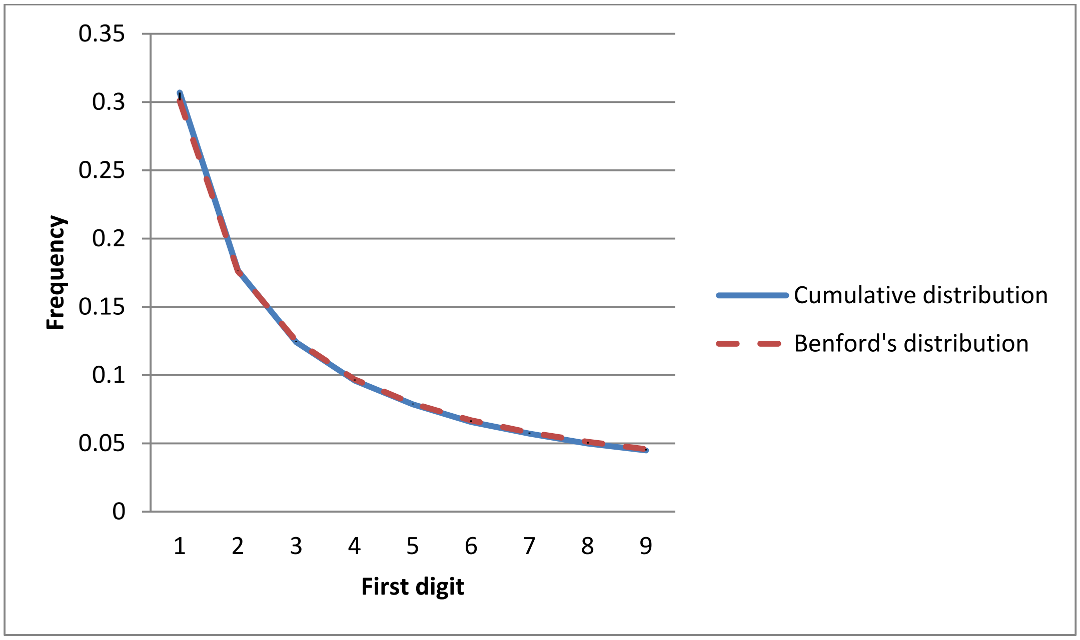

Figure 13 shows the similarity between Benford's distribution and the aggregate distribution of all financial statement variables available on Compustat for the period 2001-2011. Not shown are distributions by industry and year, which similarly conform to Benford's Law.

Figure 14: Conformity to Benford's Distribution, Firm Examples

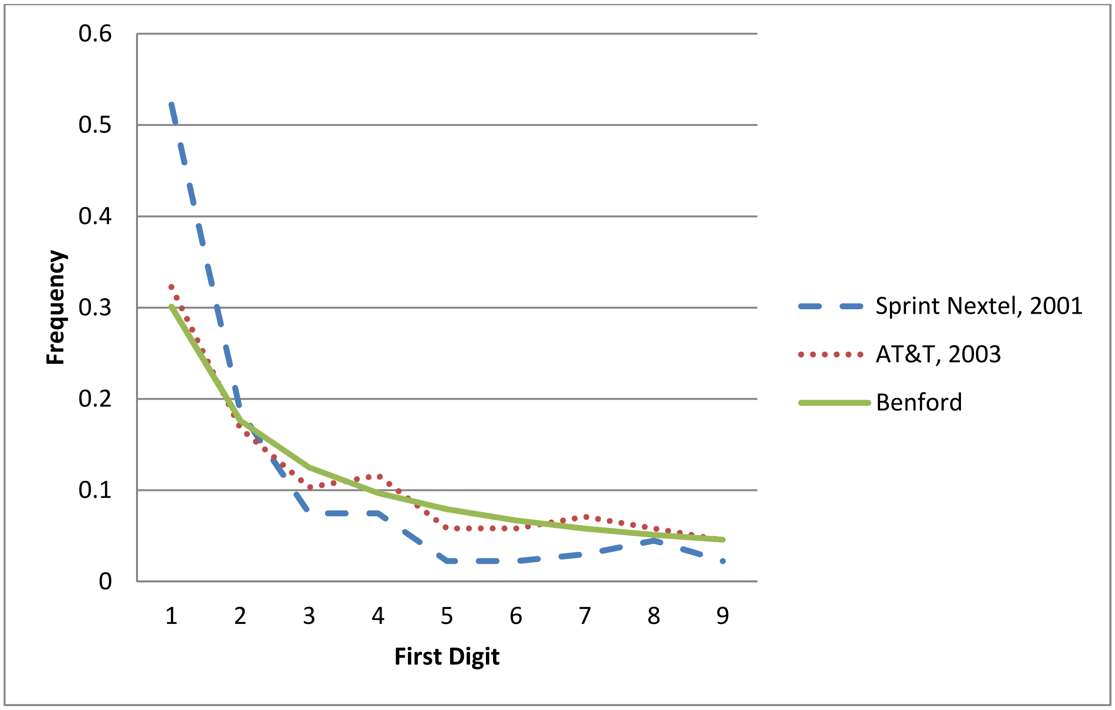

Figure 14 shows the conformity to Benford's distribution for two firm years, Sprint Nextel, 2001, which does not conform to Benford's Law (FSD Scored based on the KS statistic = 0.224, FSD Score based on the MAD statistic = 0.052) and restated its financial results for that year, and AT&T, 2003, which does conform to Benford's Law (FSD Scored based on the KS statistic = 0.028, FSD Score based on the MAD statistic = 0.013).

Table 1
Descriptive Statistics

VARIABLE	N	MEAN	SD	P25	P50	P75
FSD_Score	46,574	0.030	0.009	0.024	0.029	0.036
ABN_RET_EARN (%)	39,896	0.002	0.154	-0.068	-0.001	0.065
ABN_RET_10K (%)	35,240	0.030	0.133	-0.051	-0.002	0.047
ABN_RET_LONG (%)	35,168	0.302	0.398	-0.176	0.002	0.169
UE	30,380	-0.035	0.219	-0.005	0.000	0.003
ABS_JONES_RESID	46,574	0.191	0.374	0.029	0.073	0.181
STD_DD_RESID	46,574	0.131	0.166	0.039	0.083	0.162
MANIPULATOR	42914	0.148	0.355	0	0	0
R_CFO	46008	-0.102	4.066	-0.227	0.036	0.305
R_PROD	45599	-0.034	0.380	-0.173	-0.039	0.081
R_DISX	41705	0.401	2.550	-0.194	0.010	0.272
WCACC	46,574	0.004	0.139	-0.036	0.003	0.045
CHCSALE	46,574	0.164	0.685	-0.052	0.072	0.222
ISSUANCE	46,574	0.905	0.293	1.000	1.000	1.000
BTM	46,574	0.531	1.192	0.253	0.479	0.812
AT ($M)	46,574	3,049	7,608	67	327	1,692
ROA	46,574	-0.223	13.376	-0.067	0.028	0.081

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 A for the calculation of FSD_Score. ABN_RET_EARN is the 10-day abnormal return around the release date of a firm's earnings. ABN_RET_10K is the 10-day abnormal return around the release date of a firm's 10-K. ABN_RET_LONG is the abnormal return from 11 days after the release date of a firm's 10-K to 90 days after the release date. UE is unexpected earnings, proxied for by mean analyst forecast error. ABS_JONES_RESID is the absolute value of the residual from the modified Jones model. STD_DD_RESID is the five-year moving standard deviation of the Dechow-Dichev residual. MANIPULATOR is an indicator variable equal to 1 if the M Score is greater than -1.78 (Beneish, 1999). R_CFO is the level of abnormal cash flows from operations. R_PROD is the level of abnormal production costs. R_DISX is the level of abnormal discretionary expenses. See Appendix D for further detail. WCACC is working capital accruals. CHCSALE is the change in cash sales. SOFTAT is total assets less net PPE and cash, scaled by beginning of period total assets. BTM is the book-to-market ratio. AT is total assets. ROA is return on assets. See Appendix D for further details.

Table 2
Correlations

	FSD_Score	ABN_RET_EARN	ABN_RET_10K	ABN_RET_LONG	UE	ABS_JONES_RESID	STD_DD_RESID	MANIPULATOR
FSD_Score		-0.0325*	-0.0238*	-0.0007	-0.0148*	0.0793*	0.1083*	0.0521*
ABN_RET_EARN	-0.0248*		0.2914*	0.0715*	0.1097*	-0.0346*	-0.0375*	-0.0560*
ABN_RET_10K	-0.0206*	0.5066*		0.1249*	0.0043	-0.0325*	-0.0130*	-0.0364*
ABN_RET_LONG	-0.0044	0.0858*	0.1038*		-0.0756*	0.0018	0.0174*	-0.0454*
UE	-0.0218*	0.0145*	-0.0111	-0.0915*		-0.0110	-0.0115	0.0203*
ABS_JONES_RESID	0.0726*	-0.0187*	-0.0372*	-0.0198*	-0.0254*		0.3998*	0.1083*
STD_DD_RESID	0.1486*	-0.0343*	-0.0122*	0.0082	-0.0574*	0.2708*		0.0784*
MANIPULATOR	0.0638*	-0.0353*	-0.0294*	-0.0444*	-0.0038	0.0749*	0.1318*	

Pearson (Spearman) correlations are below (above) the diagonal. * indicates significance at the 5% level. All variables are defined in Appendix D.

Table 3

Aggregate Conformity to Benford's Distribution: All Firm Years

Panel A: All financial statement numbers

NUMBER OF FIRM YEARS	FSD_ Score
46,574	0.0014

Panel B:
All financial statement numbers by industry

INDUSTRY	NUMBER OF FIRM YEARS	FSD_Score
Food	1,463	0.0015
Mining	751	0.0019
Oil	2,116	0.0010
Clothes	900	0.0018
Durables	1,225	0.0011
Chemicals	1,103	0.0018
Consumer goods	2,108	0.0016
Construction	1,236	0.0017
Steel	756	0.0016
Fabricated products	367	0.0015
Machinery	7,500	0.0014
Cars	749	0.0011
Transportation	2,088	0.0023
Utilities	1,331	0.0018
Retail	2,953	0.0024
Others	19,573	0.0015

Panel C: All financial statement numbers by fiscal year

FISCAL YEAR	NUMBER OF FIRMS	FSD_Score
2001	4,788	0.0013
2002	4,674	0.0016
2003	4,472	0.0019
2004	4,416	0.0015
2005	4,348	0.0017
2006	4,196	0.0014
2007	4,093	0.0012
2008	3,942	0.0012
2009	3,865	0.0014
2010	3,751	0.0014
2011	3,638	0.0013

Table 3 computes the aggregate FSD_Score from all financial statement variables available on Compustat for the period 2001-2011. See Appendix A for the calculation of FSD_Score. Panel A shows the distribution for the entire sample. Panel B calculates the distributions by Fama-French industry portfolios. Panel C calculates the distribution by fiscal years. In all instances, the FSD Score is well below 0.006, which can be considered close conformity to the law (Nigrini, 2012).

Table 4

Conformity to Benford's Distribution: By Individual Firm Year

Panel A: Total number of firm-years that follow Benford's distribution

NUMBER	PERCENT
38,983	83.70

Panel B: Total number of firm-years by industry that follow Benford's distribution

INDUSTRY	NUMBER	PERCENT
Food	1,251	84.99
Mining	613	81.08
Oil	1,842	85.71
Clothes	767	84.85
Durables	1,007	82.54
Chemicals	941	84.93
Consumer goods	1,689	78.85
Construction	1,008	87.11
Steel	651	85.66
Fabricated products	316	86.10
Machinery	6,281	83.34
Cars	646	85.79
Transportation	1,778	84.55
Utilities	1,200	88.89
Retail	2,448	83.12
Others	16,465	83.33

Panel C: Total number of firm-years by year that follow Benford's distribution

FIRM YEAR	NUMBER	PERCENT
2001	4,081	84.41
2002	4,004	84.92
2003	3,769	83.57
2004	3,706	83.39
2005	3,720	84.93
2006	3,530	83.43
2007	3,485	84.36
2008	3,343	84.06
2009	3,205	82.10
2010	3,115	82.26
2011	3,025	82.65

Table 4 computes FSD_Score based on the KS statistic for each firm-year from 2001-2011 and shows the percentage of individual 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, 84% of all firm-years are not different from zero at the 5% level. Panel B (Panel C) shows similar conformity to Benford's Law across industries (years). See Appendix A for the calculation of FSD_Score based on the KS statistic.

Table 5

Benford's Distribution and Accounting Discretion

$$FSD\_Score_{i,t} = \alpha + \beta_1 ABS\_JONES\_RESID_{i,t} + \beta_2 STD\_DD\_RESID_{i,t} + \beta_3 MANIPULATOR_{i,t} + \beta_4 R\_CFO_{i,t} + \beta_5 R\_PROD_{i,t} + \beta_6 R\_DISX_{i,t} + \varepsilon_{i,t}$$

VARIABLES	FSD_Score
	(1)
ABS_JONES_RESID	0.001**
	(2.10)
STD_DD_RESID	0.006***
	(11.59)
MANIPULATOR	0.001***
	(6.79)
R_CFO	-0.000
	(-0.26)
R_PROD	-0.000
	(-1.22)
R_DISX	0.000
	(1.48)
Observations	36,789
R-squared	0.019

Table 5 examines the relation between Benford's Law and proxies for accruals-based earnings management, earnings manipulation, and real activities earnings management. The OLS regressions use all financial statement data for the period 2001-2011. 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 A for the calculation of FSD_Score. ABS_JONES_RESID is the absolute value of the residual from the modified Jones model. STD_DD_RESID is the five-year moving standard deviation of the Dechow-Dichev residual. MANIPULATOR is an indicator variable equal to 1 if the M Score is greater than -1.78 (Beneish, 1999). R_CFO is the level of abnormal cash flows from operations. R_PROD is the level of abnormal production costs. R_DISX is the level of abnormal discretionary expenses. See Appendix D 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 6

FSD Score of Firms Just Above and Below zero Net Income

	Number of Observations	FSD Score	Difference (%)	t-statistic
-0.005<=NI<0	460	0.0285	4.56	2.3878***
0<=NI<=0.005	635	0.0298		

Table 6 examines mean FSD Scores for firms just above and below 0 net income. Following Burgstahler and Dichev (1997), net income is scaled by the beginning-of-period market value of equity. 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 A for the calculation of FSD_Score. See Appendix D 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 7

Misstated Versus Restated Financial Statements

$$FSD\_Score_{i,t} = \alpha + \beta_1 RESTATED\_NUMS_{i,t} + \beta_2 ABS\_JONES\_RESID_{i,t} + \beta_3 STD\_DD\_RESID_{i,t} + \beta_4 MANIPULATOR_{i,t} + \beta_5 R\_CFO_{i,t} + \beta_6 R\_PROD_{i,t} + \beta_7 R\_DISX_{i,t} + \varepsilon_{i,t}$$

VARIABLES	FSD_Score	FSD_Score
	(1)	(2)
RESTATED_NUMS	-0.001*	-0.001***
	(-5.63)	(-5.68)
ABS_JONES_RESID		0.000
		(0.70)
STD_DD_RESID		0.006***
		(11.76)
MANIPULATOR		0.001***
		(2.72)
R_CFO		-0.000
		(-1.07)
R_PROD		-0.000
		(-1.46)
R_DISX		0.000**
		(2.31)
Observations	11,228	11,228
R-squared	0.003	0.019

Table 7 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 and that at least 10 variables were changed in the restated numbers. 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 A for the calculation of FSD_Score. ABS_JONES_RESID is the absolute value of the residual from the modified Jones model. STD_DD_RESID is the five-year moving standard deviation of the Dechow-Dichev residual. MANIPULATOR is an indicator variable equal to 1 if the M Score is greater than -1.78 (Beneish, 1999). R_CFO is the level of abnormal cash flows from operations. R_PROD is the level of abnormal production costs. R_DISX is the level of abnormal discretionary expenses. See Appendix D 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 8
FSD Score and Stock Market Returns

Panel A: 10-day abnormal returns by quintile around the *earnings release date*

Quintiles by FSD_Score	N	ABN_RET_EARN	Standard Deviation	t-statistic
1	7,982	0.3345	0.1426	2.10**
5	7,976	-0.6161	0.1680	-3.28***
Difference between 1 and 5		0.9506		3.85***

Panel B: 10-day abnormal returns around the *earnings release date* sorted by change in earnings surprise and FSD Score

		Quintiles by FSD_Score			
		1	5	Difference between FSD Quintiles 1 and 5 (%)	t-statistic
Quintiles by UE	1	-0.5284	-1.8220	1.2935	1.68*
	5	1.8254	0.5559	1.2695	2.04**

Panel C: 10-day abnormal returns by quintile around the *10-K release date*

Quintiles by FSD_Score	N	ABN_RET_10K	Standard Deviation	t-statistic
1	7,052	0.1904	0.1226	1.30
5	7,041	-0.4869	0.1397	-2.92***
Difference between 1 and 5		0.6773		3.06***

Panel D: Long-window abnormal returns by quintile from 10 days after *10-K release date* to 90 days after the 10-K release date

Quintiles by FSD_Score	N	ABN_RET_LONG	Standard Deviation	t-statistic
1	7,039	0.2696	0.3474	0.65
5	7,029	-0.2642	0.4305	-0.51
Difference between 1 and 5		0.5338		0.81

Panel E: Multivariate regressions

$$ABN\_RET\_EARN_{i,t} = \alpha + \beta_1 FSD\_Score_{i,t} + \beta_2 UE_{i,t} + \beta_3 ABS\_JONES\_RESID_{i,t} + \beta_4 STD\_DD\_RESID_{i,t} + \beta_5 MANIPULATOR_{i,t} + \beta_6 R\_CFO_{i,t} + \beta_7 R\_PROD_{i,t} + \beta_8 R\_DISX_{i,t} + \beta_9 WCACC_{i,t} + \beta_{10} CHCSALE_{i,t} + \beta_{11} R\_SOFTAT_{i,t} + \beta_{12} ISSUANCE_{i,t} + \beta_{13} BTM_{i,t} + \beta_{14} AT_{i,t} + \beta_{15} ROA_{i,t} + \varepsilon_{i,t}$$

VARIABLES	ABN_RET_EARN	ABN_RET_EARN
	(1)	(2)
FSD_Score	-0.412***	-0.286***
	(-3.89)	(-2.65)
UE		0.011***
		(2.63)
ABS_JONES_RESID		-0.004
		(-1.55)
STD_DD_RESID		-0.014*
		(-1.83)
MANIPULATOR		-0.016***
		(-5.68)
R_CFO		-0.000
		(-0.81)
R_PROD		0.002
		(0.72)
R_DISX		0.000
		(1.27)
WCACC		-0.000
		(-0.63)
CHCSALE		-0.008***
		(-4.33)
SOFTAT		-0.002
		(-0.83)
ISSUANCE		-0.002
		(-0.34)
BTM		0.018***
		(12.21)
AT		0.000
		(0.21)
ROA		0.021***
		(5.47)
Observations	24,434	24,434
R-squared	0.001	0.013

Table 8 examines the relation between firms' level of conformity to Benford's Law and equity market returns. Portfolio returns are based on the Fama-Macbeth method, i.e., we construct portfolios every year and report the mean of these portfolios. Panel A divides firm-years into quintiles based on their FSD Score with the lowest (highest) quintile having the strongest (weakest) conformity to the Law. ABN_RET_EARN is the 10-day abnormal return around the release date of a firm's earnings. Panels B uses the same 10-day return windows and first sorts on mean analyst forecast error scaled by stock price (UE) then based on the FSD Score. Panel C divides firm-years into quintiles based on their FSD Score with the lowest (highest) quintile having the strongest (weakest) conformity to the Law. ABN_RET_10K is the 10-day abnormal return around the release date of a firm's 10-K. Panel D divides firm-years into quintiles based on their FSD Score with the lowest (highest) quintile having the strongest (weakest) conformity to the Law. ABN_RET_LONG is the abnormal return from 11 days after the release date of a firm's 10-K to 90 days after the release date. Panel E conducts multivariate analysis of the relation between FSD Score and abnormal returns. ABN_RET_EARN, the 10-day abnormal return measured from the date of the earnings release to 10 days after the release, is the dependent variable. ABS_JONES_RESID is the absolute value of the residual from the modified Jones model. STD_DD_RESID is the five-year moving standard deviation of the Dechow-Dichev residual. MANIPULATOR is an indicator variable equal to 1 if the M Score is greater than -1.78 (Beneish, 1999). R_CFO is the level of abnormal cash flows from operations. R_PROD is the level of abnormal production costs. R_DISX is the level of abnormal discretionary expenses. See Appendix D 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.