The Truth about Honesty in the Nonprofit Sector*

Bruce D. McDonald, III† North Carolina State University
Christopher B. Goodman Northern Illinois University

The nonprofit sector aims to provide services with a public benefit, but how honest is it? Since the nonprofit sector relies on fundraising efforts to support its administration and program costs, and since poor financial performance can scare potential contributors away, nonprofit organizations have an incentive to appear fiscally healthy regardless of their true condition. We examine the factors associated with the honesty of organizations in the nonprofit sector using Benford’s Law, which tests for abnormalities in data that result from intentional falsification. Using the 990 tax filings for 51,010 nonprofits in the United States from 2012 and 2013, we find evidence of problems in the accuracy of their financial reporting. Those organizations with more external users of their financial information tend to conform more closely with Benford’s Law, suggesting more external monitoring of non-profit organizations may decrease the likelihood of misreported financial information.

Keywords: Not-for-profit Organizations, IRS 990, Benford’s Law

Introduction

The nonprofit sector is a key provider of public and social services in the United States (Carroll and Calabrese 2017; Salamon 2015). Its role has become more pivotal as many state and local governments have reduced funding for services due to fiscal stress (Lamothe and Lamothe 2015; Lin and Wang 2016). Currently, there are more than 1.57 million registered nonprofit organizations in the United States, providing a wide variety of services (National Center for Charitable Statistics 2017). In exchange for providing their services, nonprofit organizations are often granted tax-exempt status and their donors can deduct their financial contributions on their annual tax returns. Accompanying the tax benefits for the nonprofit sector are loose controls and oversight of the sector, partly due to the common perception that the sector is inherently honest. As organizations focused on providing a public benefit rather than generating revenue for owners or investors, they have a different incentive structure than for-profit businesses (Child, Witesman, and Braudt 2015; Schatteman and Bingle 2015). The sector’s continued ability to deliver services, however, depends upon the financial resources it receives from fundraising. Past research has noted that this dependency strongly drives nonprofit behavior, with organizations adjusting their missions and policies to reflect revenue opportunities (Eger, McDonald, and Wilsker 2015). It also raises concerns regarding how nonprofits report their finances. Since the adoption of the Taxpayers Bill of Rights 2 in 1996 and its subsequent interpretation in 1999, most nonprofits have had to make their financial statements publicly available (Hofman and McSwain 2013; Luecke, Shortill, and Meeting 1999). Potential donors can use these statements to determine the financial soundness of

---

*Forthcoming at Public Administration Quarterly
†August 18, 2020
an organization before contributing. Unfortunately, this requirement could have an unintended consequence, as it also creates an incentive for nonprofits to engage in accounting manipulation to improve their perceived financial position, since large donors are more likely to contribute to financially sound entities (Carroll and Stater 2009; Frumkin and Kim 2001; Trussel and Parsons 2007).

In this paper, we explore the incentive to engage in possible accounting manipulation of the financial position of a nonprofit organization to understand the issue of honesty in the third sector better. Previous researchers into the financial reporting of nonprofit organizations have argued that nonprofit organizations misreport their expenses (see Keating, Parsons, and Roberts 2008; Krishnan and Yetman 2011; Krishnan, Yetman, and Yetman 2006; Thornton and Belski 2010; Yetman and Yetman 2012), but the inaccuracy of reporting is frequently attributable to the lack of expertise of those who are managing the organizations financial records (Gordon, Greenlee, and Nitterhouse 1999) or the absence of the resources to do a better job (Abramson 1995; Froelich and Knoepfle 1996). We aim to move the literature forward by investigating whether the inaccuracies and misreporting in the financial reports of nonprofit organizations are intentional. We accomplish this by applying Benford’s Law to the annual 990 tax filings that nonprofits submit to the Internal Revenue Service. Benford’s Law is an observation about the frequency with which numbers naturally appear in data sets (Judge and Schechter 2009; Nigrini and Mittermaier 1997). Data that are outside the expected frequency suggest that numbers are likely false due to intentional manipulation or accidental misreporting (Cho and Gaines 2007; Nigrini 1999; Varian 1972).

While Benford's Law has a long tradition within the auditing literature as a tool for identifying financial fraud and tax evasion by businesses (Durtschi, Hillison, and Pacini 2004; Etteridge and Srivastava 1999; Nigrini 1996), it has rarely been used in the study of public finance. Only two studies to date has considered the use of Benford's Law to understand the accuracy of nonprofit financial statements. In the first study, van Caneghem (2016) applied the idea to nonprofit organizations in Belgium. More recently, in the second study Dang and Owens’s 2020 looked to understand transparency of British charities use the law. Given the unique set of laws and regulations regarding the reporting of a nonprofits financial status in the United States and the differences between the underlying accounting systems of the United States and the rest of the world, a U.S.-specific application of the law is necessary to understand our accuracy. This understanding is important as the Internal Revenue Service has attempted to address the concerns of the previous literature on 990 inaccuracies by adjusting the reporting forms and filing system several times since the 1990s. As a result of these changes, we are no longer certain whether the reporting is accurate.

Utilizing a panel of 51,010 nonprofit organizations for the fiscal years 2012 and 2013, we search for that understanding by testing the accuracy of nonprofit reporting. According to Benford’s Law, the presence of inaccuracies may result from the intentional falsification of financial records. The finding of possible manipulation in the reporting brings concerns regarding the honesty of the organization. We find, however, that the likelihood of manipulation varied with organizational size and age, as older and larger nonprofit organizations had a higher probability of dishonesty. On the other hand, internal characteristics of nonprofits that are more likely to bring attention to the financial reports, such as the presence of an endowment, were associated with a reduced probability of manipulation.

Having provided an overview of the intent of the study in the introduction, we structure the remainder of this paper as follows. Following an overview of nonprofit organizations and the
trust placed in them, we discuss in greater detail why they may have an incentive to lie or misrepresent their financial position in their financial statements. Next, we introduce Benford’s Law, its mechanics, and how it can detect potential accounting manipulation or fraud. We connect the two issues in an overview of our data and methodological approach in the third section. The following section provides our findings. A discussion of the findings and their implications for both nonprofits and academics researching the sector concludes the paper.

An Expectation of Honesty

The perception of trust and an expectation of honesty play a central role in the daily operation of nonprofit organizations (Hager and Hedberg 2016; O’Neill 2009; Salamon 1999). The importance of this expectation stems from the reliance of nonprofits on the financial support of people outside the organization to achieve their missions. Nonprofit organizations often provide services to the less fortunate, earning a reputation as noble and honorable contributors to social welfare (Berry and Arons 2003). Although there has been some anecdotal evidence of a crisis of confidence in American nonprofit organizations (Estes, Binney, and Bergthold 1989; O’Neill 2009; Salamon 1999), overall trust in the sector has remained relatively high and consistent since the 1980s (O’Neill 2009). Reasons for this high level of trust include (a) the incentive structure prevalent in the nonprofit sector (Hager and Hedberg 2016; O’Neill 2009), (b) the reliance of nonprofit boards of directors on experts to provide organizational oversight (Handy 1995), and (c) financial accountability through the public reporting of annual financial reports (Frumkin and Kim 2001; Lee 2004; Trussel 2003).

The first reason that nonprofits command a perception of honesty and trust stems from the nature of what the organizations do and how they are structured. Nonprofits differ from their for-profit counterparts in their preeminent focus on the public interest-delivering services to the needy, advocating for important social causes, and pursuing community building, among other activities (Child, Witesman, and Braudt 2015; Schatteman and Bingle 2015). Researchers have argued that the pursuit of public interest over self-interest commands trust due to the selfless nature of the work (see Hager 2004; Hager and Hedberg 2016; Lee 2004). This perception is carried further by the financial structure of the organizations. Unlike their for-profit counterparts, nonprofits are restricted from distributing earned income to their leadership or governing board as for-profit businesses do (Handy 1995; Hansmann 1980; Yetman and Yetman 2013). As they prohibit the leadership from accessing organization revenues, nonprofits are viewed in a more favorable light, since their resources are directed toward accomplishing their mission and public purpose.

The second root of trust in the nonprofit sector stems from the sector’s heavy reliance upon experts to guide its organizations. All nonprofits in the United States are overseen by internally appointed boards of directors or trustees (Duca 1996; Fischer, Vadapalli, and Coulton 2017). Although boards determine the direction of and establish the mission for the organization, from a sociological perspective they also function as legitimizers (Brody 2002; Handy 1995). To enhance public confidence in the organization, nonprofits appoint individuals who have a strong public reputation and/or expertise in the area in which the nonprofits operate to their boards (Bai 2013, Handy 1995). Potential donors can then assume that those on the board have too much to lose by being associated with a less-than-reputable organization. As a result, organizations with well-formed boards project an expectation of legitimacy and honesty in their operations.
The third source of trust in nonprofits is forced accountability. The sector has historically lacked transparency; instead, it was expected that nonprofits, of their own accord, would continue to pursue the public good for which they were established (Lee 2004). This changed after a series of scandals surrounding the management and improper use of resources by nonprofits in the 1980s and early 1990s (Hager 2004; Salamon 1999; Skolnik 1993). To restore public confidence and to ensure accountability within the sector, the Taxpayers’ Bill of Rights 2, was adopted into law in 1996 (Trussel 2003). With the passage of the bill and its subsequent interpretation in 1999, most nonprofit organizations were forced to make their financial statements publicly available (Luecke, Shortill, and Meeting 1999; Trussel 2003). The dissemination of these statements came through the 990 tax filings the organizations submit to the Internal Revenue Service as a requirement of maintaining tax-exempt status.1 This reporting obligation was expanded in 2010 to cover nearly all nonprofit organizations in the United States (McKeever and Pettijohn 2014).

Donors can now use these required filings to access information about an organization when making contribution decisions (Behn, DeVries, and Lin 2010; Calabrese 2011; van Caneghem 2016). In this way, nonprofits that use their resources responsibly are rewarded with additional donations (Lee 2004). Conversely, when nonprofits are not effective with their resources or use their funds in ways that are inconsistent with donors’ interests, they may be punished by receiving fewer contributions in the future (Yetman and Yetman 2013). Even if donors do not actively use the financial reports when making their giving decisions, the possibility that they could be used in this way forces nonprofits into a position of accountability, which adds to the sense of trust between the public and the nonprofit sector (Frumkin and Kim 2001; Hager and Hedberg 2016; Skolnik 1993).

I Incentives for Accounting Manipulation

We may expect that the behavior of organizations and individuals in the nonprofit sector will be honest. But is this expectation misguided? To answer this question, we turn to the possibility that nonprofit organizations may engage in accounting manipulation.

Accounting manipulation, also referred to as fraudulent financial reporting, is the intentional manipulation of an organization’s financial statements to create a favorable but inaccurate picture of the organization’s fiscal health (M. D. Beneish 1999; Trussel 2003). It typically involves an incentive to commit fraud, the perception of an opportunity to do so, and a rationalization for it (American Institute of Certified Public Accountants 2002). An organization may engage in accounting manipulation by the intentional

Manipulation, falsification, or alteration of accounting records or supporting documents from which the financial statements are prepared. Misrepresentation in, or intentional omission from, the financial statements of events, transactions, or other significant information. Intentional misapplication of accounting principles relating to accounts, classification, manner of presentation, or disclosure (p. 1722)

Previous research has found that accounting manipulation activity is correlated with the financial condition of the organization (Dechow et al. 2011; Trussel 2003). That is, when organizations begin

1. There are differences between audited financial statements and IRS 990 information. See Keating and Frumkin (2003) for more information.
to experience periods of fiscal stress, they are more likely to engage in fraudulent reporting (Messod D. Beneish 1997; Dechow, Sloan, and Sweeney 1995). Several models have been developed within the literature to assist in the detection of fraudulent accounting by organizations (Messod D. Beneish 1997; M. D. Beneish 1999; Dechow et al. 2011; Summers and Sweeney 1998). Rather than focusing on the raw data of the organization, most of these models look to financial ratios that reflect the financial condition of the organization to place those data in context. When an organization’s financial prospects are low, it tends to engage in manipulation to make itself appear more appealing to investors and loan managers (Trussel 2003).

Most research on accounting manipulation has concentrated on for-profit businesses. Nonprofit organizations, however, do not have the same incentive to give an appearance of profit maximization in their reporting. Instead, the behavior of nonprofits is centered on achieving a charitable mission and maximizing the services provided (Dang and Owens 2020; van Caneghem 2016). The public reporting of a nonprofit’s financial statements allows potential donors to evaluate the organization with regard to accountability and to gauge the entity’s performance at meeting its mission and its overall financial condition. One metric that is often used to interpret the success of a nonprofit is its program-spending ratio (Finkler et al. 2017). Also referred to as the program-expense ratio or simply the program ratio within the literature, the program-spending ratio is measured as the share of total expenditures devoted to the actual delivery of programs and services. This ratio is viewed as a reflection of the effort and efficiency the organization demonstrates in fulfilling its charitable mission. Previous research on donor behavior has also indicated that the program spending ratio is a key determinant of giving decisions (Yetman and Yetman 2013).

The accounting rules that direct the allocation of costs between program expenses, fundraising expenses, and management and general expenses leave this ratio susceptible to reporting error (Tinkleman 2007). Errors in accounting records can be exacerbated by a lack of expertise among those overseeing the books. Nonprofits typically rely upon volunteers to fill relevant financial accountability roles in the organization (Jang, Wang, and Yoshioka 2016), including the role of treasurer. A significant body of literature has pointed to this factor as a potential cause of inaccuracies in nonprofit financial records (see Frumkin and Kim 2001; Trussel 2003; van Caneghem 2016). It is also possible that such misrepresentations may result from intentional accounting manipulation. Nonprofit organizations are reliant upon the financial assistance of others to fund their operations, and most are required to report on their financial status annually (Yetman and Yetman 2013). Relatively few nonprofits, however, have their financial reports formally audited on their own accord, and fewer still are audited by the IRS (Calabrese 2011; Tate 2007). This situation may tempt nonprofits to manipulate their financial records to paint themselves in a better light, as their breach of integrity has a low risk of being discovered. As a result, we hypothesize that despite the general perception that nonprofit organizations are honest, not all of them are behaving accordingly.

**Benford’s Law**

To explore the factors associated with potential nonprofit financial misrepresentation, we use Benford’s Law. Benford’s Law is an observation about the distribution of first significant digits as they naturally appear in data sets (Nigrini and Mittermaier 1997). The origins of the law date back to the work of Simon Newcomb, an astronomer and mathematician, in 1881 (Judge and Schechter
Newcomb observed that the books of logarithmic tables in the library were more worn in the beginning pages, which dealt with low digits, and that the degree of wear in the books progressively lessened as the tables moved on to deal with higher digits. From this pattern, he inferred that scientists used the tables to look up numbers that started with one more often than numbers that started with two. As the first digit of a number became larger, the frequency of looking up that number decreased. From his observations, Newcomb calculated the probability that a number has a non-zero first digit. The probability is:

\[ P(d_i) = \log_{10}(1 + \frac{1}{d_i}) \quad (1) \]

where \( d_i \) is the first digit of a number 1, 2, \ldots, 9 that occurs at probability \( P \). Using the formula, the probability that the first digit is 1 is about 30 percent. The measurement of the probable distribution can also be applied to all subsequent digits in a number. The probabilities of appearing as each of the first five digits of any number for all digits between 0 and 9 are provided in table 1.

Newcomb (1881) published the results of his observations and the resulting probability calculations in the *American Journal of Mathematics*. Half a century later, unaware of Newcomb’s work, Benford (1938), also a physicist, noticed a pattern of wear in his own logarithmic tables. Based on the pattern of wear, Benford reached the same conclusion as Newcomb, but he sought to test the hypothesis with real data. To demonstrate the probability pattern, he collected more than 20,000 observations from a variety of sources, including the numbers appearing in the Reader’s Digest, the atomic weights of elements, and the areas of rivers. Benford found that the numbers consistently fell into a pattern in which low digits occurred more frequently in the first position than higher digits. Although Newcomb was the first to observe the pattern of first significant digits, it was Benford’s rediscovery and testing of the relationship that led it to become widely known has and that led the relationship to be identified as Benford’s Law (Hill 1995).

Since its rediscovery by Benford, the law has undergone rigorous testing using a variety of data sources, leading Furlan (1948) to conclude that it reflects a profound truth in nature. Varian (1972), for instance, argued that it can be used to test the honesty or validity of purportedly random data in the social sciences. Not until the research of Carslaw (1988) and Thomas (1989), however, did

---

2. Calculated by: \( P(1) = \log_{10}(2) = 0.3010 \).
the law begin to be considered as a tool for detecting inconsistencies in financial data.\textsuperscript{3} If financial data generally adhered to the frequencies of a number as established by Benford’s Law, both Carslaw and Thomas hypothesized that any significant deviation from this distribution would be associated with an error in reporting. Carslaw investigated this using the earnings data of businesses in New Zealand, ultimately concluding that the earnings did not conform to the expected distribution. In looking at the data, Carslaw found more zeros in the second position than expected and fewer nines, leading him to conclude that firms altered their earnings by rounding them upwards. Thomas found a similar pattern in the earnings of U.S. firms.

Nigrini (1996) took Carslaw and Thomas work another step forward to become the first to apply Benford’s Law to detect explicit accounting fraud. He hypothesized that when accounting data are truthfully reported, the frequency of numbers should converge to the logarithmic distribution of the digits. When data are falsified, however, the numbers are expected to occur outside their natural order and thus not adhere to Benford’s Law. Nigrini found support for his hypothesis with a higher than expected frequency of ones and a lower than expected frequency of nines in a sample of federal tax returns. Nigrini and Mittermaier (1997) further tested the hypothesis, concluding that the law is a useful tool for auditors investigating fraud in financial statements, because statements with distributions that are significantly different from the expected (or natural) distribution likely contain errors. This finding has subsequently been replicated in subsequent studies by other researchers (also see Durtschi, Hillison, and Pacini 2004).

Data, Methodology, and Hypotheses

The data for this analysis came from IRS 990 tax filings by U.S. nonprofit organizations for tax years 2012 and 2013 (394,822 organizations). They were collected from the National Center for Charitable Statistics (NCCS) raw data files. We excluded nonprofits that filed an IRS 990-EZ, as they lack the financial complexity necessary for analysis.\textsuperscript{4} This operation excludes 173,862 organizations with revenues of less than $250,000 and assets less than $500,000. The use of raw data files rather than the cleaned NCCS data commonly seen in the nonprofit finance literature is intentional. Data cleaning is useful for analysis of nonprofit financial information; however, during the cleaning process, corrections to the data are made that may obscure the misrepresentation of nonprofit accounting data. The cleaning may also wipe away some or all of the useful information on nonprofit filings that we need to conduct our analysis. Following Dang and Owens (2020), we eliminated unnecessary outliers by removing any organizations with negative total assets (this eliminated 549 organizations). We further reduced the sample by eliminating 3,273 organizations with missing data. These two tax years provided a data set of 217,138 nonprofit organizations.

The unit of analysis is the individual organization. From the 990s, we utilized the financial information associated with the balance sheet and income statement to construct our dependent variable. There are 254 such potential entries in these data for each organization. Problematically, many of these entries are zero for a large portion of the organizations in our data set. Nigrini (2012) suggested that a minimum of 100 nonzero observations should be used in calculating the probable distribution predicted by Benford’s Law. Following Michalski and Stoltz (2013) and Dang

\textsuperscript{3} For a more thorough discussion of the application of Benfords Law to financial report and accounting, see Nigrini (2012).

\textsuperscript{4} This is a common data-cleaning tactic, see Searing (2015) for an explanation for this relative to small non-profit organizations and Calabrese (2011) for an overall explanation.
Owens (2020), we pooled the data across the two tax years contained in our data set to attain the minimum 100 nonzero observations. This practice is common in the literature as a means to get to the minimum number of nonzero observations (Dang and Owens 2020). The dependent variables are constructed using the individual nonzero financial observations from the 990s over the two tax years. The independent variables are averaged across those same years. After we eliminated the organizations with fewer than 100 nonzero financial observations, the final data set consisted of 51,010 organizations.

In the aggregate, most large data sets conform to Benford’s Law. Our approach involved constructing proxies of deviations from Benford’s Law for each organization and correlating these deviations with variables associated with external monitoring of an organization’s financial information. This approach requires the selection and construction of proxies for deviations from the distribution suggested under Benford’s Law. There are two general approaches: (a) measures of statistical deviation such as mean absolute deviation (Amiram, Bozanic, and Rouen 2015; Dang and Owens 2020) and (b) test statistics and critical values to determine whether a given set of numbers conforms to Benford’s Law (Amiram, Bozanic, and Rouen 2015; Marrow 2014; Michalski and Stoltz 2013; Nye and Moul 2007; Dang and Owens 2020). However, the second approach tends to over-reject the null hypothesis of conformity with Benford’s Law as the number of digits used grows large Dang and Owens (2020). As the critical values are a function of sample size, this problem becomes larger as the sample size grows. The first approach does not suffer from this issue and is our preferred proxy. Nevertheless, we also calculated results from two additional proxies from the second family as robustness checks.

The mean absolute difference (MAD) statistic is calculated as the difference between the observed proportions of leading digits and the theoretical proportions suggested by Benford’s Law using the pooled financial data from our data set. Mathematically, the equation is as follows:

$$MAD_j = \frac{1}{9} \sum_{i=1}^{9} |P_o(d_i) - P_e(d_i)|$$  \hspace{1cm} (2)

where $j$ is the organization, $d_i = 1, 2, \ldots, 9$ represent the leading digits of each observation, $P_o(d_i)$ is the observed proportion of each digit, $d_i$, and $P_e(d_i)$ is the theoretical proportion of each digit, $d_i$. For instance, when $i = 1$, $P_o(d_1)$ is equal to the proportion of financial observations, pooled over two tax years, with a first digit of one for organization $j$. This is compared against the theoretical proportion according to Benford’s law for leading first digits equal to one of 0.301. This process is repeated for all other digits. The absolute difference between the observed (actual) proportion of leading digits and the theoretically expected proportion of leading digits are summed and divided by nine. The greater deviation from the theoretical proportions of each digit, the larger $MAD_j$ becomes. Therefore, larger values for $MAD_j$ indicate more deviation from theoretical exceptions and a higher likelihood of manipulation.\(^5\)

In addition to our primary proxy, we utilize two additional proxies as robustness checks against our $MAD_j$ measure. The first is based on $\chi^2$ and is derived from Michalski and Stoltz (2013). The proxy is shown as:

$$D^2_j = N \sum_{i=1}^{9} \frac{[P_o(d_i) - P_e(d_i)]^2}{P_e(d_i)}$$  \hspace{1cm} (3)

\(^5\) The actual length of the number in question is irrelevant to the calculation of MAD or any of the other proxies of deviation from Benford’s Law. In our analysis, only the leading or first digit is of consequence.
As with equation 2, \( j \) denotes the organization, \( d_i = 1, 2, \ldots, 9 \) representing the leading digits of each observation, \( P_o(d_i) \) and \( P_e(d_i) \) are defined in the same way as before, and \( N \) is the total number of observations for each nonprofit organization. A second measure is the Kolmogorov-Smirnov (KS) statistic defined as follows:

\[
KS_j = \max_{d_i \in \{1, 2, \ldots, 9\}} | \sum_{i=1}^{d_i} P_o(d_i) - P_e(d_i) |
\]

(4)

where the definitions of the various variables are again the same as defined above.

There are a number of caveats about the use of proxies in analyses of this type (Dang and Owens 2020). A comparison of the observed distribution of first digits to the theoretical expectation provided by Benford’s Law does not reveal manipulations that happen at the end of numbers, such as rounding up. Second, organizations may not manipulate every statement, or they may be strategic about their manipulations. Finally, organizations may inflate or deflate their financial statements by a common factor, and Benford’s Law would not detect this alteration. However, this last point is of less concern, Dang and Owens (2020) explained, as most individuals are not aware of Benford’s Law and are unlikely to be able to replicate the theoretical distribution of first digits effectively.

Following Dang and Owens (2020), the following regression is specified.

\[
Y_j = \beta_0 + \beta_1 M_j + \beta_2 X_j + \epsilon_j
\]

(5)

where \( Y_j \) is a proxy of conformity to Benford’s Law for organization \( j \), \( M_j \) is a vector of variables pertaining to the monitoring on financial information for organization \( j \), \( X_j \) is a vector of control variables, and \( \epsilon_j \) is the usual error term.

Our primary hypothesis is that organizations with higher levels of external monitoring are less likely to engage in potential manipulation (Yetman and Yetman 2012). That is, the external governance structure of a nonprofit influences its ability to manipulate the reporting of its finances. The underlying assumption is that as the number of users of financial information increases, the likelihood of detection of financial irregularities increases and the incentives for manipulation decline. External monitoring is broken into two categories: external monitoring coming from capital providers/donors and monitoring coming from regulators. The former is operationalized with three dichotomous variables and one continuous one: whether an organization engages in fundraising, whether it has an endowment, whether it issues tax-exempt (municipal) bonds, and the proportion of net assets that are temporarily or permanently restricted. In each case, the capital provider or donor has an incentive to monitor the activities of the organization. As Yetman and Yetman (2012) explained, donors often act like stockholders and see their donations as investments in future outcomes generated by the nonprofit organizations. They have an incentive to monitor whether their donations are being properly used by the organizations to that end. Financial statements are one type of information donors can monitor. Bondholders act as lenders to an organization and have an incentive to monitor the investments they make. More importantly, bondholders are more likely to be sophisticated consumers of financial information and are more likely to engage in monitoring. Across all four of these variables, we hypothesize that having these characteristics (or having a larger proportion of restricted net assets) is negatively related to our proxy of conformity to Benford’s Law.
Monitoring coming from regulators is operationalized as whether an organization engages in unrelated business (taxable) activities. According to Yetman and Yetman (2012), the audit rate of nonprofit organizations is roughly similar to all other organizations; however, the audit rate of nonprofit organizations with unrelated business activities or taxable activities is eight times as large as nonprofits that do not have such activities. Therefore, organizations with unrelated business activities have an incentive to be more careful in their reporting to the IRS as their probability of audit is high. We hypothesize that organization with unrelated business activities is negatively related to our proxy of conformity to Benford’s Law. In each case, engaging in one of these activities attracts more attention to the entity’s financial information from people who are well versed in finance and accounting practices and better able to detect financial irregularities than the average person. Therefore, engaging in each of these activities should reduce the likelihood of potential manipulation.

We include a number of control variables as well. We control for the size of the organization using total assets and the number of employees of the organization, as well as the age of the organization. To control for the financial complexity of an organization, the total number of financial observations on an organization’s 990 is included. Finally, we include a dichotomous variable indicating one if an organization’s total liabilities are greater than its total assets and zero otherwise. Having this characteristic increases the risk of organizational insolvency (Keating et al. 2005).

Results

Overall, organizations in our data deviate from the expected distribution of first digits about 3 percent of the time (See MAD statistic in Table 2). This is similar to the average MAD statistic in UK nonprofit organizations (0.038) as reported by Dang and Owens (2020) and significantly more deviation on average than in Belgian nonprofits as reported by van Caneghem (2016). It is important to note this finding pertains to the financial information contained in individual 990s rather than the proportion of organizations in the sample. Nigrini (2012) explains that MAD values larger than 0.015 suggest nonconformity with Benford’s Law, indicating that a potentially large proportion of our data are out of conformity with Benford’s Law. This potentially large deviation is not uncommon when applying Benford’s Law to financial statements or 990s (Nigrini 2012). In an examination of the $\chi^2$ proxy and their associated critical values, we fail to reject the null hypothesis of conformity with Benford’s Law for all organizations in our sample.

As Table 2 shows, these five financial monitoring variables range in their importance to the nonprofit sector. A large percentage of the organizations in our sample (40 percent) have a fundraising operation and about one quarter have an endowment. Approximately 18 percent of an organization’s net assets are restricted in some way on average. There is significant variation in this measure, with some organizations having no restricted assets and some having nearly all their assets restricted. The remaining two variables are less common, with about 15 percent having unrelated business activities and 9 percent receiving proceeds from tax-exempt bonds. The other control variables indicate that the organizations in the sample, when compared to the total universe of nonprofits, tend to be older, larger in terms of employees, holders of vast amounts of assets, at relatively low risk of insolvency, and producers of more complex 990s. This profile of organizations is largely a function of the decision to limit the analysis to organizations with a min-

---

7. In some specifications, the square of age is included.
Table 2: Summary Statistics

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean</th>
<th>St. Dev.</th>
<th>Min.</th>
<th>Max.</th>
</tr>
</thead>
<tbody>
<tr>
<td>MAD</td>
<td>0.0327</td>
<td>0.0095</td>
<td>0.0065</td>
<td>0.0819</td>
</tr>
<tr>
<td>$D^2$</td>
<td>19.9302</td>
<td>11.8168</td>
<td>0.8912</td>
<td>124.8003</td>
</tr>
<tr>
<td>KS</td>
<td>0.0816</td>
<td>0.0303</td>
<td>0.0141</td>
<td>0.2930</td>
</tr>
<tr>
<td>Has endowment (1=Yes)</td>
<td>0.2424</td>
<td>0.4286</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Has professional fundraising (1=Yes)</td>
<td>0.4089</td>
<td>0.4916</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Has tax exempt bonds (1=Yes)</td>
<td>0.0879</td>
<td>0.2831</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Has unrelated business expenses (1=Yes)</td>
<td>0.1494</td>
<td>0.3565</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Restricted assets</td>
<td>0.1803</td>
<td>0.2565</td>
<td>0.0000</td>
<td>0.9999</td>
</tr>
<tr>
<td>Age of NPO</td>
<td>32.5354</td>
<td>19.1241</td>
<td>0.0000</td>
<td>113.0000</td>
</tr>
<tr>
<td>Number of employees</td>
<td>252.0694</td>
<td>1234.5460</td>
<td>0.0000</td>
<td>67088.0000</td>
</tr>
<tr>
<td>Total assets ($M)</td>
<td>46.8754</td>
<td>499.4478</td>
<td>0.0015</td>
<td>64567.1500</td>
</tr>
<tr>
<td>Insolvency risk</td>
<td>0.0457</td>
<td>0.2089</td>
<td>0.0000</td>
<td>1.0000</td>
</tr>
<tr>
<td>Total statement observations</td>
<td>121.5898</td>
<td>18.6881</td>
<td>100.0000</td>
<td>291.0000</td>
</tr>
</tbody>
</table>

$n = 51,010$

The minimum of 100 financial information observations. Our robustness checks confirmed, however, that the limiting of our sample in this way had no appreciable effect on the results.8

Table 3 shows the results of our regression analysis, as modeled in Equation 5. Due to the pooled nature of the data, we do not attempt to interpret the point estimates of these results. Rather, the signs of the coefficients are our concern, as they provide an indication of positive or negative influence on the likelihood of manipulation of an organization’s 990. Positive signs are associated with a larger likelihood of manipulation (greater distance between observed and expected distributions of first digits) and negative signs are associated with a smaller likelihood of manipulation (smaller distance between observed and expected distributions of first digits).

Turning first to our five variables of interest, three of the four dichotomous variables were negative and statistically significant. Having an endowment, having a fundraising operation, and having unrelated business expenses were associated with a lower probability of manipulation of financial information. Additionally, as the proportion of net assets that are restricted increases, the probability of manipulation declines. All these activities place a spotlight on an organizations financial documents (from donors, capital providers, or regulators), making the discovery of potential manipulation more likely. The spotlight can consist of closer attention to the documents by internal actors (such as fundraising staff or consultants and fund managers) or external actors (such as donors or regulators), or some combination by all actors may be the source of the influence. The remaining variable related to monitoring, issuing tax-exempt bonds, was not statistically different from zero in either specification. Taken as a group, however, the five variables were jointly significant. These results are qualitatively similar to the results of Dang and Owens (2020) as they pertain to monitoring. Organizations with more external monitoring pressure tend to conform more closely with Benford’s Law.

8. Similar to Dang and Owens (2020), we specify a Heckman selection model to determine if procedure to limit organization to only those with 100 non-zero financial observations has introduced any selection bias into our results. The finding from the Heckman selection model are qualitatively similar to the primary results presented here and are available upon request.
Table 3: Primary Regression Results

<table>
<thead>
<tr>
<th>Variable</th>
<th>(1) Coef.</th>
<th>(2) Coef.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has endowment (1=Yes)</td>
<td>-0.4504**</td>
<td>-0.4605**</td>
</tr>
<tr>
<td></td>
<td>(-4.38)</td>
<td>(-4.48)</td>
</tr>
<tr>
<td>Has fundraising activities (1=Yes)</td>
<td>-0.9837**</td>
<td>-0.9959**</td>
</tr>
<tr>
<td></td>
<td>(-11.12)</td>
<td>(-11.26)</td>
</tr>
<tr>
<td>Has tax exempt bonds (1=Yes)</td>
<td>0.1565</td>
<td>0.2157</td>
</tr>
<tr>
<td></td>
<td>(0.99)</td>
<td>(1.36)</td>
</tr>
<tr>
<td>Has unrelated business expenses (1=Yes)</td>
<td>-0.3080*</td>
<td>-0.2998*</td>
</tr>
<tr>
<td></td>
<td>(-2.56)</td>
<td>(-2.49)</td>
</tr>
<tr>
<td>Restricted assets</td>
<td>-2.7141**</td>
<td>-2.6773**</td>
</tr>
<tr>
<td></td>
<td>(-15.93)</td>
<td>(-15.73)</td>
</tr>
<tr>
<td>Age of NPO</td>
<td>0.0192**</td>
<td>0.0812**</td>
</tr>
<tr>
<td></td>
<td>(8.27)</td>
<td>(10.39)</td>
</tr>
<tr>
<td>Age of NPO(^2)</td>
<td>-</td>
<td>-0.0008**</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(-8.41)</td>
</tr>
<tr>
<td>Number of employees</td>
<td>0.0001**</td>
<td>0.0001**</td>
</tr>
<tr>
<td></td>
<td>(3.58)</td>
<td>(3.80)</td>
</tr>
<tr>
<td>Total assets (Ln)</td>
<td>-0.3724**</td>
<td>-0.3704**</td>
</tr>
<tr>
<td></td>
<td>(-12.31)</td>
<td>(-12.25)</td>
</tr>
<tr>
<td>Insolvency risk</td>
<td>-0.1365</td>
<td>-0.0727</td>
</tr>
<tr>
<td></td>
<td>(-0.64)</td>
<td>(-0.34)</td>
</tr>
<tr>
<td>Total statement observations</td>
<td>-0.0785**</td>
<td>-0.0783**</td>
</tr>
<tr>
<td></td>
<td>(-28.05)</td>
<td>(-28.02)</td>
</tr>
<tr>
<td>Observations</td>
<td>51,010</td>
<td>51,010</td>
</tr>
<tr>
<td>State dummies</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>NTEE major group dummies</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Robust t-statistics in parentheses; ** p < 0.01, * p < 0.05, All coefficients are multiplied by 1,000 to aid in interpretation.
The control variables tell a somewhat confusing story. Both older organizations and organizations with large numbers of employees were associated with a higher likelihood of manipulation. However, when specified as a quadratic, the age of the organization displays an inverse-U shaped relationship with MAD. The likelihood of manipulation increases as age increases until an organizational age of 48 and beyond this age the likelihood of manipulation declines. Organizations with extensive assets had a lower likelihood of manipulation. It does not appear that organizations at risk of insolvency manipulate their reporting, suggesting that the organizations with potentially the most to gain from manipulation, declining organizations, are no more likely than any others to manipulate. This result is at odds with van Caneghem (2016), who found that smaller organizations (those in the first decile of total assets) are more likely to deviate from Benford’s Law than large organizations (those in the top decile of total assets).

Dang and Owens (2020) also found a positive relationship between size by assets and likelihood of manipulation in a regression context among UK nonprofits. Given the similarities in approaches between this analysis and Dang and Owens (2020), we must conclude that the differences arise from differences in the two samples. The organizations included here are uniformly larger than those in Dang and Owens. This finding is similar to that of Krishnan, Yetman, and Yetman (2006), where size was negatively correlated to misrepresented expenses (but did not deviate from Benford’s Law). Last, organizations with more complex 990s (as indicated by a greater number of financial observations) had a lower likelihood of manipulation, similar to Dang and Owens (2020). This result is not surprising, and it can be explained in a similar way to the results for our financial monitoring variables: large, complex 990s are unlikely to be completed by a single individual, making the potential for manipulation much lower than if a single individual were doing the work.

As we mentioned in the previous section, we present a series of secondary results to test the robustness of our primary findings, since the literature on the application of Benford’s Law is somewhat split as to the most appropriate methodology and these robustness checks represent the test statistics vein of research. As with the primary results in Table 3, the point estimates are not interpreted; rather, the sign of the coefficients is the primary concern.

Table 4 shows that the results of these robustness checks are similar to those in Table 3. Again, both having an endowment and having fundraising activities lowered the potential for manipulation. Receiving proceeds from tax-exempt bonds was statistically indistinguishable from zero, as was the case in the original calculations. Having unrelated business expenses was statistically significant and negative for the $D^2$ model, but statistically nonsignificant for the KS model (although the sign was still negative). Last, increased proportions of restricted assets are associated with lowered potential for manipulation. As in the previous results, the five primary variables of interest were jointly significant. The control variables in the robustness checks generally supported the results in Table 3; however, the sign changed with regard to total financial observations for the $D^2$ model, contrary to our original finding. Rather than suggesting that more complex 990s were associated with lower levels of potential manipulation, this particular result suggests that more research is necessary to tease out the relationship between 990 complexity and the potential for manipulation.

A final concern regarding the construction of our data set was the exclusion of organizations with fewer than 100 financial observations. By doing so, selection issues could cause biases our primary results and robustness checks. To investigate this possibility, we specified a Heckman selection model estimating the probability of being included in our final data set and included
Table 4: Robustness Checks

<table>
<thead>
<tr>
<th>Variable</th>
<th>D²</th>
<th>KS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Has endowment (1=Yes)</td>
<td>-0.7863**</td>
<td>-1.1195**</td>
</tr>
<tr>
<td></td>
<td>(-6.12)</td>
<td>(-3.38)</td>
</tr>
<tr>
<td>Has fundraising activities (1=Yes)</td>
<td>-1.8563**</td>
<td>-1.7684**</td>
</tr>
<tr>
<td></td>
<td>(-16.64)</td>
<td>(-6.17)</td>
</tr>
<tr>
<td>Has tax exempt bonds (1=Yes)</td>
<td>0.1079</td>
<td>0.3316</td>
</tr>
<tr>
<td></td>
<td>(0.52)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Has unrelated business expenses (1=Yes)</td>
<td>-0.5019**</td>
<td>-0.1696</td>
</tr>
<tr>
<td></td>
<td>(-3.21)</td>
<td>(-0.44)</td>
</tr>
<tr>
<td>Restricted assets</td>
<td>-3.3982**</td>
<td>-6.7503**</td>
</tr>
<tr>
<td></td>
<td>(-16.56)</td>
<td>(-12.24)</td>
</tr>
<tr>
<td>Age of NPO</td>
<td>0.0862**</td>
<td>0.2047**</td>
</tr>
<tr>
<td></td>
<td>(8.63)</td>
<td>(8.08)</td>
</tr>
<tr>
<td>Age of NPO²</td>
<td>-0.0009**</td>
<td>-0.0021**</td>
</tr>
<tr>
<td></td>
<td>(-6.88)</td>
<td>(-6.79)</td>
</tr>
<tr>
<td>Number of employees</td>
<td>0.0001**</td>
<td>0.0004**</td>
</tr>
<tr>
<td></td>
<td>(2.70)</td>
<td>(3.53)</td>
</tr>
<tr>
<td>Total assets (Ln)</td>
<td>-0.6793**</td>
<td>-0.9125**</td>
</tr>
<tr>
<td></td>
<td>(-17.45)</td>
<td>(-9.40)</td>
</tr>
<tr>
<td>Insolvency risk</td>
<td>0.1585</td>
<td>-0.9000</td>
</tr>
<tr>
<td></td>
<td>(0.58)</td>
<td>(-1.34)</td>
</tr>
<tr>
<td>Total statement observations</td>
<td>0.0990**</td>
<td>-0.2298**</td>
</tr>
<tr>
<td></td>
<td>(25.15)</td>
<td>(-25.41)</td>
</tr>
<tr>
<td>Observations</td>
<td>51,010</td>
<td>51,010</td>
</tr>
<tr>
<td>State dummies</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>NTEE major group dummies</td>
<td>X</td>
<td>X</td>
</tr>
</tbody>
</table>

Robust t-statistics in parentheses; ** p < 0.01, * p < 0.05, KS coefficients are multiplied by 1,000 to aid in interpretation.
this probability in the form of an inverse Mills ratio in the estimation of Equation 5. The results from this regression suggested that selection bias is not a concern in our primary models, as the inverse Mills ratio in the second stage was not statistically significant and the main results did not exhibit any qualitative change.

Discussion and Conclusion

We have a general expectation that the focus of nonprofit organizations on serving the public indicates their adherence to a greater good (Lee 2004). The leaders of nonprofit organizations, after all, lack the incentive of their private sector counterparts (Hager 2004; Hager and Hedberg 2016). A perverse incentive exists within the sector, however, in that nonprofits may be rewarded for quietly setting integrity aside and engaging in accounting manipulation. The reliance of the sector on fundraising to cover expenses, coupled with the limited oversight of the sector, provides both the incentive and the opportunity to engage in potentially malicious activity.

To explore the issue of honesty within individual nonprofit organizations, we used Benford’s Law to examine the 990 tax filings of 51,010 nonprofit organizations in the United States from fiscal years 2012 and 2013. According to Benford’s Law, there is a natural frequency with which numbers should appear in a data set. By comparing at the frequency at which numbers appear in the tax filings to the natural frequency with which they should appear, we were able to measure the probability that each nonprofit misrepresented its data in completing its 990s. Overall, organizations in our data deviate from the expected distribution of first digits about 3 percent of the time. This finding suggests that for the average non-profit organization chosen from our sample, their non-zero financial observations will deviate from what is expected under Benford’s Law approximately 3 percent of the time. To understand the instances of probable manipulation further, we explored potential institutional drivers of this behavior. In doing so, we found that the data from organizations with more assets and those with more employees were less likely to conform to Benford’s Law, indicating larger deviations from the expected distribution of first digits. On the other hand, organizations with more people engaged in the financial monitoring of the organization, whether through restricted donations or regulations, had a reduced likelihood of manipulation.

We expect nonprofit organizations and governmental agencies to act in an ethical manner. We often even go so far as to expect a higher standard for those who lead the organizations than we ourselves adhere to. Based on our study, this attitude is potentially misguided. Previous literature on nonprofit finance has suggested that the 990s submitted by nonprofits are often inaccurate, but researchers have attributed this problem to unintentional, random errors by bookkeepers (Yetman, Yetman, and Badertscher 2009). Closer evaluation using auditing techniques such as Benford’s Law suggests there are factors associated with closer conformity, thereby reducing the potential for manipulation.

We provide no evidence of widespread falsification in the financial reports of nonprofit organizations. Our findings merely demonstrate several factors associated with closer conformity with Benford’s Law. What our findings allow us to conclude is that further steps are necessary to ensure proper financial oversight of the nonprofit sector. The responsibility for this oversight cannot be left to the federal government, but should also be shared by donors, who can request audits and evidence of suitable administrative controls be put into place before making contributions to an organization. To the extent that large donors and sources of capital are requesting further financial
information, our results suggest that the probability of financial manipulation declines, providing a key additional source of external oversight.

The results of our study are certainly interesting and have implications for how we should look to the nonprofit sector, but they do have some limitations. In our analysis, we excluded nonprofit organizations for which we were not able to obtain at least 100 data points across the two years of 990 filings. In doing so, we created a bias toward larger nonprofit organizations, as smaller nonprofits are more likely to have filed a 990-N or 990-EZ or had filings too simple to provide the required minimum set of data points. We demonstrated that this exclusion had no bearing on the results, but it is still possible that smaller organizations may differ systematically from larger ones in some relevant fashion. Previous literature on auditing has noted that organizational fraud and employee theft are more likely in smaller organizations (see Singleton et al. 2006). Given our evidence, a more systematic examination of smaller nonprofit organizations is warranted.

References


