NPO FINANCIAL STATEMENT QUALITY:

AN EMPIRICAL ANALYSIS BASED ON BENFORD'S LAW

Tom Van Caneghem

Associate Professor, KU Leuven

Warmoesberg 26

1000 Brussel

BELGIUM

tom.vancaneghem@kuleuven.be

NPO FINANCIAL STATEMENT QUALITY:
AN EMPIRICAL ANALYSIS BASED ON BENFORD'S LAW

ABSTRACT

In order to assess the accuracy of the figures reported in NPOs' financial statements, I perform a digital analysis on Belgian non-profit organizations' financial statements for accounting years 2007 up to 2012. Specifically, I compare observed frequencies for digits in the second-from-the-left position with expected frequencies based on Benford's Law. Results based on the full sample indicate that observed frequencies strongly conform to Benford's Law (and thus suggest a high degree of accuracy of the figures reported in NPOs' financial statements). Nevertheless, I note statistically significant deviations from Benford's Law (both for the entire distribution and at the individual digit level). The largest deviation is noted for zeroes in the second position (i.e., a significantly positive deviation), which can be explained based on humans' reliance upon so-called cognitive reference points. Considering different sub-samples, I note that observed deviations from Benford's Law are largest for the smallest non-profits and those non-profits that rely most heavily on grants and/or donations.

KEY WORDS: non-profit organizations, financial reporting, financial statements, Benford's Law, digital analysis

1. Introduction

As argued by Privett and Erhun (2011), non-profit organizations (henceforth NPOs) are a significant and growing segment of the economy. The growing importance of the non-profit sector and financial reporting problems with NPOs have led legislators worldwide to examine governance and accountability issues in this sector (Vermeer et al., 2009). For example, there has been an international trend to increase financial disclosure regulation for NPOs as a means of improving public accountability (Calabrese, 2011). Unlike their for-profit counterparts, NPOs quite often rely heavily on donations and/or grants to finance their operations. NPOs typically also enjoy a variety of tax benefits (Yetman and Yetman, 2012). NPOs are therefore accountable to the general public, who are, in effect, financing these donations, grants, and fiscal benefits (Yetman and Yetman, 2012). For resource providers (i.e., donors, governments and volunteers), there is an increasing and pressing need to ensure that society reaps the highest social benefit from their funding and/or labour. As argued by Yetman and Yetman (2003), financial reports play an important role as they provide a means for funders and other stakeholders to assess whether the NPO is using obtained funding towards its mission in an efficient manner. Prior research indicates that donors and governments do indeed rely on financial statement information in their decision to provide funding to an NPO (see e.g., Parsons, 2003; Kitching, 2009; Thornton and Belski, 2010; Feng et al., 2011; Verbruggen, Reheul et al., 2011; Amin and Harris, 2012). As a result, the reliability and quality of NPOs' financial statements (henceforth FS) is important to, and relevant for, various NPO stakeholder groups, including donors, governments and volunteers (Verbruggen and Christiaens, 2012).

In the current paper, I assess the quality (or accuracy) of the figures reported in Belgian NPOs' FS by means of a digital analysis. More specifically, I compare observed frequencies for digits in the second-from-the-left position (henceforth second digits) with expected frequencies based on Benford's Law. As discussed in Durtschi et al. (2004), various authors have promoted Benford's Law as a simple and effective tool to uncover irregularities in accounting numbers. Whereas prior studies already explore FS quality among NPOs, they typically focus on very specific items (e.g., the reported earnings figure (see e.g., Verbruggen and Christiaens, 2012; Jegers, 2013); taxable income (see e.g., Omer and Yetman, 2003); and fundraising expenses (see e.g., Krishnan et al., 2006)) or the formal quality of the FS[a] (see e.g., Verbruggen, Christiaens et al., 2011). Unlike prior studies, I do not focus on one or more specific figures reported in the FS, rather I assess the overall accuracy of figures reported in the balance sheet and the income statement. This is highly relevant, as it is very unlikely that all FS users focus on exactly the same figure (e.g., the earnings figure) in NPOs' FS. As discussed in Verbruggen and Christiaens (2012), subsidizing governments do not merely consider the reported earnings figure, but consider a mixture of FS items (e.g., accumulated reserves, the presence (and magnitude) of other types of revenues, the (relative) magnitude of operating expenses, etc.) in order to make their funding decisions. In a similar vein, NPOs' creditors are likely to consider a variety of FS items in order to assess the creditworthiness of the organization. Creditworthiness is typically assessed in terms of solvency and liquidity, which is mainly based on balance sheet information, rather than information from the income statement. As such, it is important to assess the overall accuracy of the figures reported in NPOs' FS.

The remainder of the paper is organized as follows. In Section 2, I briefly discuss relevant prior literature and develop my hypotheses, while the sample and research method are introduced in Section 3. Results are presented in Section 4. Finally, I summarize my main conclusions in Section 5.

2. Review of the literature and development of hypotheses

2.1. Benford's Law

Contrary to what one might logically expect, the occurrence of each digit, as a first (or second) digit in a particular number, is not equally likely. For example, ones (nines) are most (least) likely to be observed as a first digit by chance. This is referred to as Benford's Law. Inspired by the observation that the first pages of a table of common logarithms show more wear than do the last pages, Benford (1938) assumes that more numbers begin with digit one than digit nine[b]. In order to empirically verify this apparent oddity, Benford examines the occurrence of each digit as a first digit based on several datasets of numerical data (total number of observations = 23,197)[c]. Employing integral calculus, Benford developed formulae to estimate expected frequencies of first digits, second digits and digit combinations in lists of numbers. Table 1 presents an overview of expected frequencies for first and second digits based on Benford's Law. The law applies to lists of numbers that describe the relative sizes of similar phenomena (i.e., if these numbers are not influenced by human thought) (Nigrini and Mittermaier, 1997).

[INSERT TABLE 1 HERE]

Ever since Benford's article, various empirical studies have aimed at ascertaining that specific numerical datasets are Benford sets (i.e., datasets that conform Benford's Law). Examples include Wlodarski (1971); Sentance (1973); Becker (1982); Burke and Kincanon (1991) and Ley

(1996).[d] In general, results lend support for the validity of Benford's Law. The application of digital analysis (i.e., based on Benford's Law) to various types of financial numbers have led to the promotion of this tool for detecting fraudulent practices (see e.g., Nigrini and Mittermaier, 1997; Busta and Weinberg, 1998; Durtschi et al., 2004). Rauch et al. (2011), for example, apply Benford's Law to macroeconomic data reported to Eurostat by EU member states. They find that data reported by Greece, for which issues with regard to data quality have been known for a long time, show the largest deviations from Benford's Law. Greece's data were particularly deviant in 2000, just before Greece joined the Euro. Nigrini and Mittermaier (1997) provide a case-based example to illustrate the usefulness of digital analysis (i.e., examining patterns in digits and numbers) as an analytical review procedure[e]. Busta and Weinberg (1998), on the other hand, empirically assess the performance of digital analysis as an analytical review procedure. Relying on simulated datasets, Busta and Weinberg conclude that digital analysis outperforms most other analytical review procedures and that its major strength lies in the fact that it allows detection of small manipulations, whereas other analytical review procedures are typically designed to signal larger oddities. The latter is attributable to the fact that digital analysis is insensitive to an error's magnitude, whereas this is not true for most other analytical review procedures. Durtschi et al. (2004) further discuss how Benford's Law can be effectively used by auditors in detecting fraud in accounting data.

2.2. Hypotheses development

2.2.1. Learning curve effect

As discussed in Reheul et al. (2014), Belgian NPOs have been faced with quite extensive reforms of accounting and reporting requirements in the recent past. The original Belgian non-profit legislation of 1921 introduced hardly any accounting and reporting rules. Some form of cash accounting was obligatory and an undefined overview of cash flows and financial budget had to be filed with the local court of justice. However, this information was not made public. This ‘free-of-obligations situation’ has changed through the introduction of a new law on 2 May 2002 (Belgisch Staatsblad, 2002) that is applicable to FS as from 2006 onwards. As a result, Belgian NPOs are confronted with (i) new accounting legislation (i.e., accrual-based instead of cash-based accounting); and (ii) new standardized formats of the FS, amongst other. Belgian NPOs therefore had to familiarize themselves with a new legal framework and start reporting in a completely different way. As argued by Lande and Rocher (2011), the introduction of accrual-based accounting (i.e., one of the important changes in the Belgian setting, cf. supra) has many obstacles and can lead to a variety of technical and conceptual difficulties. Because of these obstacles and difficulties, the introduction of new accounting requirements may have resulted in inaccuracies in the figures disclosed by Belgian NPOs. Because of learning effects, such inaccuracies are likely to erode over time. That is, as the organization’s accountant learns more, ‘teething problems’ will disappear (Owusu-Ansah, 2000). Accordingly, based on *learning curve theory*, I expect that FS figures for the initial years of the sample period are less compliant with Benford’s Law than those for the later years. I therefore hypothesize:

H1: *Frequencies for second digits follow Benford’s Law more closely for the later years in my sample period than for the initial years.*

2.2.2. Size of the NPO

As argued by Vermeer et al. (2006), larger NPOs are faced with more stakeholders and are therefore more likely to be scrutinized by regulatory bodies, the media and the (general) public. In addition, large organizations are characterized by more resources (e.g., more accounting staff and/or more advanced accounting systems) (Owusu-Ansah, 2000; Ismail and Chandler, 2003). Trussel and Parsons (2008), for example, argue that smaller NPOs are less likely to have expertise in cost allocations and FS preparation than larger ones. Several prior studies report findings that are in line with these arguments. In the Belgian context, Verbruggen, Christiaens et al. (2011) document a positive relationship between NPO size and the degree of compliance with financial reporting standards.[f] Also in the Belgian setting, Reheul et al. (2014) document a significantly negative relationship between NPO size and the financial reporting lag[g]. In other words, results presented by Reheul et al. (2014) indicate that larger NPOs report significantly more quickly than small(er) NPOs. Consistent with the argument that larger organizations are characterized by more resources, Mook et al. (2007) observe that larger NPOs are significantly more likely to keep records on volunteer value (and quantify them). Several other studies document a positive relationship between NPO size and the amount of (financial) information disclosed (see e.g., Christensen and Mohr, 2003; Crawford et al., 2009; Zainon et al., 2012). In sum, based on the aforementioned arguments and empirical findings, I predict that the combination of increased public scrutiny and more resources for large NPOs is likely to result in more accurate financial reporting (compared to small(er) NPOs). In other words, I hypothesize:

H2: Frequencies for second digits follow Benford's Law more closely for large(r) NPOs than for small(er) NPOs.

2.2.3. Resource dependence

The main sources of nonprofit financing include donations, grants, and sales of products and services (Yetman and Yetman, 2003). In the literature, it is quite common to classify NPOs into 'donative' versus 'commercial' organizations[h], depending upon their main source of financing (see e.g., Hansmann, 1987; Calabrese, 2011; Balsam and Harris, 2014; Reheul et al., 2014). Based on *resource dependence theory*, the aforementioned distinction is highly relevant, because an organization's need for resources is argued to determine its decisions and actions (see e.g., Pfeffer and Salancik, 1978; Vermeer et al., 2009). A major difference between donative and commercial NPOs is that donative NPOs' resource providers (i.e., private donors and/or governments) are generally not direct consumers of an organization's programs or services and are therefore unable to directly evaluate the quality of the NPOs' output (Saxton et al., 2011). Accordingly, FS play an important monitoring role for donative organizations (Yetman and Yetman, 2012). Based on interviews among stakeholders of Scottish charities, Crawford et al. (2009) find that funders are the only group that interviewees consider would use charity FS. Consistent with these arguments, Amin and Harris (2012) find that donors' contributions are negatively related to a going concern audit opinion, but that decisions of those who receive goods or services in return for their support are not affected by a going concern opinion. In this respect it might be interesting to add that results from an online survey among Belgian NPOs' FS users confirms that members (i.e., fund providers) and volunteers (i.e., providers of labor) are the most frequent users of these reports (Mortier and Baten, 2010). In addition, as discussed in Verbruggen, Reheul et al. (2011), governments also rely on NPOs' FS in a Belgian setting (i.e., for both funding and supervisory purposes). Based on these considerations, it could be argued that in order to guarantee the stream of funding, which is less stable and less predictable than in commercial NPOs (Gronjberg, 1991), donative NPOs are likely to report more accurately. This is especially true in the current context of increased competition between NPOs in the 'market' for donations and in the context of financially limited governments. Accurate financial reporting can lead to greater confidence in the sector, which in turn can lead to willingness by donors to increase funding (Hyndman and McMahon, 2011). In line with these arguments, Verbruggen, Christiaens et al. (2011) document a positive relationship between reliance on government funding and the degree of compliance with financial reporting standards. In addition, Reheul et al. (2014) document that NPOs that rely more heavily on donations and/or grants are significantly less likely to file their FS late[i]. Nevertheless, the fact that FS serve a more prominent monitoring role for donative NPOs also gives rise to certain financial reporting incentives. As argued by Verbruggen and Christiaens (2012), when trying to increase their funding, theory and evidence suggests that NPOs will manage earnings downward in order to demonstrate a need for funding. Various studies document the prevalence of earnings management practices among NPOs (see e.g., Leone and Van Horn, 2001; Krishnan et al., 2006; Verbruggen and Christiaens, 2012; Jegers, 2013). In line with the aforementioned argument, Verbruggen and Christiaens (2012) find that NPOs that rely more heavily on governmental funding are more likely to manage earnings downward towards a zero profit. Because of stronger financial reporting incentives, it could be argued that the FS figures of NPOs that rely heavily upon donations and/or grants are likely to be less accurate. In sum, I face two competing views and I therefore state the third hypothesis in the null form:

H3: The extent of reliance upon donations and/or grants does not affect the degree of compliance of frequencies for second digits with Benford's Law.

Financial debt may also represent an important financial resource for NPOs. Just as for reliance upon donations and/or grants, I face two competing views regarding the relationship between reliance upon financial debt and financial reporting accuracy. That is, given their non-profit characteristics, NPOs often attempt to negotiate below-market interest rates (Verbruggen, Christiaens et al., 2011). To obtain financial loans (and particularly at beneficial conditions), NPOs have to be able to present reliable financial information. Financial institutions are professional FS users, who have the knowledge, the ability and the experience as well as the custom to scrutinize FS before making investment decisions. Accordingly, it could be argued that NPOs that rely more heavily on financial debt are likely to report more accurately. Nevertheless, reliance upon financial debt also gives rise to financial reporting incentives. As argued by Verbruggen and Christiaens (2012), it is important for NPOs with financial debt to show positive results and financial strength in order to convince financial institutions of their creditworthiness. In line with this argument, Bouwens et al. (2004) show that Dutch non-profit hospitals manage earnings upwards both in the year prior to and the year in which additional financial debt is obtained. In a similar vein, Jegers (2013) observes a positive relationship between reliance upon financial debt and earnings management. Because of stronger financial reporting incentives, it could be argued that the FS figures of NPOs that rely heavily upon financial debt are likely to be less accurate. In sum, I face two competing views and I therefore state the fourth hypothesis in the null form:

H4: *The extent of reliance upon financial debt does not affect the degree of compliance of frequencies for second digits with Benford's Law.*

3. Data collection and research method

3.1. Data collection

From Belfirst[j], I select all organizations that filed the complete format[k] of the FS. I only consider NPOs that file the complete format of the FS for two reasons. First, NPOs that file the abbreviated format of the FS do not need to disclose the amount of donations and/or grants they receive (i.e., information I need in order to test H3) and I would therefore be unable to test all my hypotheses for NPOs filing the abbreviated format. Second, by only considering NPOs that file the complete format of the FS, I ensure homogeneity among the sample in terms of (i) the information that is (or needs to be) disclosed in the FS; and (ii) the presence of an external FS audit (i.e., while the complete format of the FS is subject to a mandatory external FS audit, this is not the case for the abbreviated format). For NPOs filing the complete format of the FS, I export all figures disclosed in both the balance sheet and income statement for the accounting years 2007 up to 2012. I discard FS figures for accounting year 2006 (being the first year for which Belgian NPOs were obliged to file their FS with the NBB), because I noted inaccuracies in the Belfirst database with regard to the unit (i.e., EUR vs. thousands of EUR) that is used to prepare the FS. As from 2007 onwards, all FS have to be filed in EUR (and the aforementioned problem is therefore solved). Doing so, I obtain a sample of 8,012 NPO-year observations (ranging between 1,267 for 2007 and 1,373 for 2012).

3.2. Research method

In order to assess the accuracy of the figures reported in NPOs' FS, I perform a so-called 'digital analysis' on second digits (i.e., digits in the second-from-the-left position). The phrase 'digital analysis' refers to the fact that I examine patterns in digits in numbers. Specifically, I compare observed digital frequencies with expected frequencies based on Benford's Law. The research method is thus based on the assumption that accurate FS data conform to Benford's Law, while manipulated and/or fabricated data deviate from Benford's Law. My focus on second digits (as opposed to first digits) is induced by prior empirical evidence indicating that anomalies in fabricated data are much more explicit among second digits than among first digits (see e.g., Mosimann et al., 1995; Diekmann, 2007). As I focus on second digits, reported figures smaller than ten are excluded from the analyses. In addition, I exclude negative figures because prior studies have demonstrated that deviations for negative figures may exhibit exactly the opposite pattern as positive figures (and, as a result, deviations for negative figures may (partially) offset deviations for positive figures) (see e.g., Thomas, 1989). The statistical significance of observed deviations is assessed using the normalized *Z*-statistic and the Chi-squared test. Because goodness of fit tests usually produce statistically significant results when based on large sample sizes (Rauch et al., 2011), I also consider the Mean Absolute Deviation (henceforth MAD) (see e.g., Reddy and Sebastin, 2012; Haynes, 2012), which is independent of sample size.

In order to test my hypotheses (where I link NPO characteristics to the likelihood of observing deviations from Benford's Law), I create sub-samples of NPOs and compare digital frequencies for these sub-samples both with each other and with Benford's Law. In order to test H1 (learning curve effect), I compare digital frequencies for the period 2007 up to 2009 with digital frequencies for the period 2010 up to 2012. In order to test the other hypotheses, I consider NPOs that belong to the top and bottom decile based on the variable of interest.[l] In order to test H2 (NPO size), sub-samples are created based on the natural logarithm of total assets (in thousands of EUR) (to be denoted by *SIZE*) (see e.g., Verbruggen, Christiaens et al., 2011; Jegers, 2013). In order to test H3 (reliance upon donations and/or grants), sub-samples are created based on the ratio of 'membership contributions, other contributions, bequests and grants' over total assets (to be denoted by *DONSUBS*) (see e.g., Calabrese, 2011; Reheul et al., 2014). In order to test H4, sub-samples are created based on the ratio of financial debt over total assets (to be denoted by *FINLEV*) (see e.g., Jegers, 2013). Because the latter two ratios exhibit a large number of zero observations[m] (i.e., more observations than those that belong to a decile), I compare the top decile with the sub-sample of NPOs that report a zero value for these variables (i.e., instead of the bottom decile).

4. Results

4.1. Descriptive statistics and correlation matrix

Table 2 presents descriptive statistics for the sample under study, while Table 3 presents a correlation matrix for the variables related to the hypotheses.[n] Based on the descriptive statistics, I note that the distribution of NPO size is heavily skewed (i.e., the mean value is about three times the median value). I further note that NPOs in the sample are quite heterogeneous with respect to reliance upon donations and/or grants (as displayed by the reported values for the 25th and the 75th percentile). Overall, reliance upon financial debt is modest (cf. the mean (median) value of 6.62 (16.67) percent for *FINLEV*), but the reported values for the 25th and the 75th percentile clearly indicate heterogeneity among the sample. Based on the correlation matrix, I note that correlations among the 'grouping' variables are rather modest.

[INSERT TABLE 2 HERE]

[INSERT TABLE 3 HERE]

4.2. Main result

4.2.1. Full sample

Table 4 presents results for the digital analysis based on the full sample. Based on the MAD, I note that observed frequencies strongly conform to Benford's Law (i.e., for second digits, a MAD below 0.008 indicates strong conformity (see e.g., Haynes, 2012)). Nevertheless, the Chi-squared statistic is highly significant (i.e., at the one percent level) and I therefore reject the null hypothesis that observed frequencies conform to expected frequencies based on Benford's Law.[o] At the individual digit level, I note that the large majority of percentage unit deviations attain statistical significance at the conventional levels (i.e., at the one or five percent level). The largest deviation is noted for zeroes (i.e., about one percentage point), while deviations for all other digits are rather modest. To put the observed deviations in perspective, it might be interesting to add that the 0.98 percentage point deviation for zeroes implies that about 8.2 percent (i.e., 0.98 / 11.97) of the observed zeroes in the second position is ‘unexpected’, which is quite substantial.

Looking for a pattern in the observed deviations, I note that both zeroes and fives occur significantly more often than would normally be expected, while all other significant deviations are negative. This pattern is consistent with rounding-up behaviour. Rosch (1975) finds that multiples of ten serve as *cognitive reference*[p] points with a view to perceiving and evaluating numbers[q]. Worded differently, humans tend to summarize multi-digit numbers as multiples of ten: $N*10^{k-1}$, where N equals the first digit of the multi-digit number and k equals the number of digits. Thus, other digits than the first digit are typically ignored in order to assess the magnitude of a given number and any number’s first digit is therefore of vital importance.[r] Based on the aforementioned line of thought, a reported figure of five million Euro will be perceived as being abnormally larger than a figure of 4,998,500 Euro, while the actual difference only amounts to a marginal 1,500 Euro (Van Caneghem, 2002). The observed deviation for zeroes in the second position (and the observation that all other significant deviations, except for fives, are negative) is therefore consistent with managers rounding-up reported figures in order to influence FS users’ perceptions. That is, observing significantly more zeroes in the second position is consistent with an abnormally large amount figures just exceeding a cognitive reference point and thus rounding-up behavior. While the observed deviation for fives in the second position appears to conflict with the aforementioned arguments, it is consistent with prior empirical evidence reported by Aerts et al. (2008). That is, in line with results reported in the current paper, Aerts et al. (2008) observe significantly more zeroes and fives in the second position of dividends per share for a sample of US firms. As discussed by Aerts et al. (2008), the observed deviation for fives in the second position suggests that (certain) managers believe that FS users adopt a rounding rule to assess the magnitude of a given number (see e.g., Ashworth et al., 2003) instead of the generally documented truncation strategy. Adopting a rounding strategy, a FS user will first round a reported figure of 1,516 Euro (1,487 Euro) to 2,000 Euro (1,000 Euro) instead of merely considering the first digit in order to assess its magnitude.

[INSERT TABLE 4 HERE]

4.2.2. Learning curve effect (H1)

In order to test H1, Table 5 presents results for a digital analysis based on two sub-periods (i.e., being 2007-2009 and 2010-2012). Considering the different test statistics, no differences in digital frequencies are noted between both periods and results are therefore inconsistent with H1 (and a learning curve effect). Findings for both sub-periods are consistent with those based on the full sample (and I therefore do not discuss results in more detail).

[INSERT TABLE 5 HERE]

4.2.3. Size of the NPO (H2)

In order to test H2, Table 6 presents results for the digital analysis performed on the top and bottom decile based on NPO size (i.e., natural logarithm of total assets (in thousands of EUR)). Considering the different test statistics, I note statistically significant differences between both sub-samples. First, the MAD for the bottom decile (i.e., the smallest NPOs) is twice the MAD for the top decile (i.e., the largest NPOs). Results therefore indicate that digital frequencies for the largest NPOs follow Benford's Law more closely than for the smaller NPOs. Second, the Chi-squared statistic for the difference between both sub-samples attains statistical significance at the one percent level. Third, at the individual digit level, I also note statistically significant differences between both sub-samples. The most striking difference is noted for zeroes in the second position. While both sub-samples exhibit significantly more zeroes in the second position than would be expected based on Benford's Law, the deviation is about three times as large for the bottom decile (i.e., 0.57 percentage point for the top decile vs. 1.67 percentage point for the bottom decile). Results therefore indicate that rounding-up behavior is more prevalent among small NPOs. In sum, results presented in Table 6 are consistent with H2.

[INSERT TABLE 6 HERE]

4.2.4. Resource dependence (H3 and H4)

In order to test H3, Table 7 presents results for the digital analysis performed on the top decile based on *DONSUBS* (i.e., the ratio of 'membership contributions, other contributions, bequests and grants' over total assets) and the sub-sample of NPOs for which *DONSUBS* equals zero. Considering the different test statistics, I note statistically significant differences between both sub-samples. First, the MAD for the top decile (i.e., the sub-sample of NPOs that rely most heavily on donations and/or grants) is about twice the MAD for the sub-sample of NPOs that do not rely on donations and/or grants. Results therefore indicate that digital frequencies for the sub-sample of NPOs that do not rely on donations and/or grants follow Benford's Law more closely than for those NPOs that rely most heavily on donations and/or grants. Second, the Chi-squared statistic for the difference between both sub-samples attains statistical significance at the one percent level. Third, at the individual digit level, I also note statistically significant differences between both sub-samples. The largest difference is noted for zeroes in the second position. While both sub-samples exhibit significantly more zeroes in the second position than would be expected based on Benford's Law, the deviation is substantially larger for the top decile (i.e., 1.89 percentage point for the top decile vs. 0.91 percentage point for the sub-sample of NPOs that do not rely on donations and/or grants). Results therefore indicate that rounding-up behavior is more prevalent among the sub-sample of NPOs that rely heavily on donations and/or grants. In sum, results presented in Table 7 are inconsistent with H3 (presented in the null form). Results therefore support the view (and prior empirical evidence) that NPOs that rely more heavily on donations and/or grants face stronger financial reporting incentives and are therefore more likely to manage reported financial statement figures.

[INSERT TABLE 7 HERE]

Table 8 presents results for a digital analysis performed on the top decile based on *FINLEV* (i.e., the ratio of financial debt over total assets) and the sub-sample of NPOs for which *FINLEV* equals zero. Based on the Chi-squared statistic for both sub-samples, I note that digital frequencies for both sub-samples deviate significantly (at the one percent level) from Benford's Law. Comparing the MAD for both sub-samples, I note that they are exactly the same (being 0.003). Accordingly, reported results do not indicate differences between both sub-samples with regard to the extent to which digital frequencies for both sub-samples conform to Benford's Law. In other words, results presented in Table 8 are consistent with H4 (presented in the null form). Nevertheless, I note statistically significant differences in digital frequencies between both sub-samples (i.e., the Chi-squared statistic attains statistical significance at the one percent level and based on the Z-statistic I note statistically significant differences for particular digits).

[INSERT TABLE 8 HERE]

4.3. Sensitivity analyses

In order to assess the robustness of my findings, I also compared top and bottom quartiles (instead of deciles) based on the variables of interest. In addition, I compared digital frequencies for 2007 with those for 2012 in order to assess the robustness of my results for H1. These robustness checks do not affect reported findings.

5. Conclusions

NPOs' FS are an important source of information for various stakeholders (e.g., governments, donors, volunteers). As argued by Yetman and Yetman (2012), inaccurate FS can lead to suboptimal decisions and potential misallocation of resources. In the current paper, I perform a digital analysis on Belgian NPOs' FS in order to assess the accuracy of the reported figures. Unlike prior studies, I do not focus on one or more specific figures reported in the FS, rather I assess the overall accuracy of the figures reported in the balance sheet and the income statement. This is highly relevant, as it is very unlikely that all FS users focus on exactly the same figure (e.g., the reported earnings figure) in NPOs' FS. Overall, results based on the full sample indicate that observed frequencies strongly conform to Benford's Law (and thus suggest a high degree of accuracy of reported FS figures). Nevertheless, I note statistically significant deviations from Benford's Law (both for the entire distribution and at the individual digit level). The largest deviation is noted for zeroes in the second position (i.e., a significantly positive deviation), which can be explained based on humans' reliance upon cognitive reference points. Considering different sub-samples, I note that observed deviations from Benford's Law are largest for the smallest NPOs and those NPOs that rely most heavily on grants and/or donations.

As argued by Jegers (2013), contrary to the large amount of research on FS quality of for-profits, the literature on FS quality of NPOs is rather scant. Results obtained in the current study add to the literature on financial reporting quality of NPOs and are therefore of interest to NPOs' stakeholders. Reported results are relevant for external FS auditors because they document a specific type of financial reporting behavior that auditors should be aware of. While the observed behavior might not be quantitatively material[s], it might affect FS users decisions and thus be qualitatively material. If FS users only consider the first digit of reported FS figures (as suggested by the psychological literature), the latter seems plausible. Auditors should therefore be made aware of this phenomenon and its potential impact on FS users' decisions. They might, for example, consider the impact of proposed adjusting entries on the first digit of reported figures when deciding on whether or not to waive them. Moreover, auditors might also apply digital analysis to clients' accounting data in an attempt to bare irregularities. Other FS users should also be aware of the phenomenon in order to avoid psychological biases in interpreting FS figures (i.e., not putting too much emphasis on the first digit of reported FS figures). They could avoid doing so by considering FS ratios instead of absolute FS figures. If the observed rounding behavior is quantitatively immaterial, its impact on FS ratios will be minor.

Endnotes

[a] Verbruggen, Christiaens et al. (2011) rely on an index to assess compliance with existing accounting regulation. Specifically, their compliance index consists of quantifiable measures related to four accounting principles (i.e., objectivity, quality of information, periodicity, and prudence) that are directly observable in the FS. As an illustration, they check the presence of debtors and creditors in the balance sheet, as this is typical under accrual accounting (and can thus be considered as being related to quality of information). As such, Verbruggen, Christiaens et al. (2011) assess the formal compliance of the FS. That is, the mere presence of debtors and creditors does not imply that the reported figures are accurate.

[b] It is important to acknowledge that Newcomb (1881) reports the same factual observation several years prior to Benford (1938). However, Benford does not refer to Newcomb's article and ever since the publication of Benford's article, the *law of anomalous numbers* is known as Benford's Law.

[c] More specifically, Benford examined 20 different datasets that were neither too restricted in numerical range nor too conditioned in some way. Examples of datasets studied by Benford include: Arabic figures appearing in front page news item of a newspaper; all figures (except for dates and page numbers) appearing in an issue of Reader's Digest; figures appearing in mathematical tables from engineering handbooks, etc.

[d] See www.benfordonline.net for a comprehensive bibliography with regard to (empirical applications of) Benford's Law.

[e] *"(…) analytical review procedures compare expected relationships among data items to actual observed relationships. If the actual relationships are not consistent with the expected relationships further audit investigation is required to explain the unexpected results."* (Busta and Weinberg, 1998: 356) Examples of traditional analytical review procedures include ratio analysis and trend analysis.

[f] It is important to add that NPO size is no longer significant in the model when controlling for the presence of an external FS audit. However, because NPO size and the presence of an external FS audit are strongly correlated (i.e., I refer to the correlation matrix presented in Verbruggen, Christiaens et al. (2011)), the presence of an external FS audit (partially) captures the size effect.

[g] The financial reporting lag (*or* financial reporting delay) is the period that elapses between the closing date of the accounting year and the date of making the FS public. Because timeliness is recognized to be of vital importance to the usefulness of FS information (Ismail and Chandler, 2003), a short(er) financial reporting lag implies more useful FS.

[h] Balsam and Harris (2014) distinguish between 'service oriented' and 'charitable' NPOs, which essentially comes down to the same classification.

[i] Belgian NPOs have to file their FS within 7 months after the closing date of the accounting year (subject to legal sanctions).

[j] Bureau van Dijk's Belfirst database contains FS data for Belgian and Luxembourgian firms and organizations.

[k] Belgian accounting legislation imposes a standardized format for the FS. Three models exist: (i) the complete format; (ii) the abbreviated format; and (iii) some alternative formats (that are used/prescribed in specific industries). A limitation of the Belfirst database is that it only includes data from FS that have been filed according to a format prescribed by law, being (i) and (ii) (because information in the Belfirst database is presented according to these templates). As the name already suggests, the complete format of the FS is more detailed and provides more information than the abbreviated format of the FS. For example, while the complete format counts 52 pages, the abbreviated format counts 30 pages. Specific examples of differences between both formats are that, on the balance sheet, the abbreviated format contains less detailed information with respect to financial fixed assets, inventories, investments, and long-term debt. In the abbreviated format of the income statement, operating revenues (e.g., turnover) may be merely expressed as a gross margin, whereas detailed information on both operating revenues and expenses are mandatory in the complete format. Finally, far less information (and detail) is required in the notes for the abbreviated format of the FS. Importantly, regardless of the FS format used, the organization is obliged to disclose all information contained in that type of format.

Size criteria determine whether the complete or the abbreviated format of the FS has to be filed. The complete format is mandatory for very large NPOs and allowed for large and small NPOs on a voluntary basis. The abbreviated format is mandatory for large NPOs, unless they choose to file the complete format, and allowed for small NPOs on a voluntary basis. Very large NPOs exceed at least two of the following criteria: (i) total assets of 3,125,000 EUR, (ii) total incoming resources of 6,250,000 EUR, and/or (iii) 50 employees (expressed as full-time equivalents). NPOs with at least 100 full-time equivalent employees are always considered to be very large. Large NPOs exceed at least two of the following characteristics: (i) five full-time equivalent employees; (ii) total assets of 1,000,000 EUR, and/or (iii) total incoming resources of 250,000 EUR. All other NPOs are considered to be small. Note that for small NPOs there is no obligation to file FS.

[l] NPOs are assigned to deciles for each sample year separately.

[m] I refer to Table 2 Descriptive statistics for additional detail.

[n] Reported figures are based on winsorized data (at the 1 and 99 percent level).

[o] Results based on the MAD and the Chi-squared statistic might appear to contradict each other. That is, the MAD indicates strong conformity, whereas the Chi-squared statistic is highly significant (indicating non-conformity). As discussed earlier, goodness of fit tests usually produce statistically significant results when based on large sample sizes, while the MAD is not affected by sample size. Given the large sample employed in the current study, small deviations attain statistical significance because of sample size. While the MAD indicates strong conformity, it is important to note that the observed deviation for zeroes in the second position is quite substantial (regardless of sample size). That is, about 8.2 percent of the observed zeroes in the second position is 'unexpected' (cf. infra). Nevertheless, deviations for all other digits are modest, which explains the low value for the MAD.

[p] A cognitive reference point can then be defined *as "(...) a stimulus (...) which other stimuli are seen in relation to"* (Rosch, 1975: 532). Worded differently, a cognitive reference point is a member of a category that acts as a natural benchmark for comparing other members of that category (Bowdle and Gentner, 1997).

[q] Results reported by Hinrichs et al. (1982) and Poltrock and Schwartz (1984) support the idea that multiples of ten serve as cognitive reference points. Both studies examine the way humans process multi-digit numbers.

[r] This is clearly expressed in the so-called '$1.99'or 'odd pricing' phenomenon (i.e., the observation that an abnormally large quantity of retail prices fall just below a round number). Several studies in marketing (see e.g., Twedt, 1965; Holdershaw et al., 1997) demonstrate the prevalence of this phenomenon. Whereas several alternative explanations have been offered for this phenomenon (see e.g., Schindler and Kibarian, 1996; Holdershaw et al., 1997), results lend support for the 'underestimation mechanism' hypothesis (see e.g., Schindler and Wiman, 1989; Schindler and Kibarian, 1996). That is, Schindler and Wiman (1989) show that leftmost digits are most likely to be accurately recalled and recall errors on odd prices are therefore more likely to be underestimates than on even prices (i.e., prices ending in zeroes). Results by Schindler and Kibarian (1996) show that the odd pricing phenomenon really induces a sales effect and that this effect is attributable to the fact that humans tend to underestimate odd prices. Brenner and Brenner (1982) argue that this is due to the fact that humans are flooded with numerical data (e.g., in advertising

brochures, financial statements, etc.) and that first rounding numbers would result in an additional inordinate load on humans' information processing capabilities. Accordingly, humans merely store the most important bits of information (i.e., first digits of large numbers).

s Auditing standards require auditors to provide reasonable assurance that FS are free of material misstatements. Nevertheless, auditing standards do not provide detailed materiality guidelines (i.e., the assessment of materiality is considered to be a matter of professional judgment). Frequently mentioned cut-off levels for assessing materiality in the auditing literature are 0.50 percent of total assets; 5 percent of net income; and 0.50 percent of operating income. From this, it should be clear that materiality is typically determined in terms of quantitative measures. As such, it is possible that the observed rounding behavior is not quantitatively material.

References

Aerts, W., Van Campenhout, G., and Van Caneghem, T. (2008). Clustering in dividends: Do managers rely on cognitive reference points?. Journal of Economic Psychology, 29, 276--284.

Amin, K., and Harris, E. (2012). The Impact of Going Concern Audit Opinions on Nonprofit Revenues and Expenses. Available at SSRN: http://ssrn.com/abstract=2015486.

Ashworth, J., Heyndels, B., and Smolders, C. (2003). Psychological taxing in Flemish municipalities. Journal of Economic Psychology, 24, 741--762.

Balsam, S., and Harris, E. E. (2014). The Impact of CEO Compensation on Nonprofit Donations. Accounting Review, 89, 425--450.

Becker, P. (1982). Patterns in Listings of Failure-Rate & MTTF Values and Listings of Other Data. IEEE Transactions on Reliability, 31, 132--134.

Belgisch Staatsblad (2002). Wet betreffende de verenigingen zonder winstoogmerk, de international verenigingen en de stichtingen. Brussels: Belgisch Staatsblad. [Law concerning non-profit organizations, international organizations, and foundations]

Benford, F. (1938). The law of anomalous numbers. Proceedings of the American Psychological Society, 78, 551--572.

Bouwens, J., Hollander, S., and Schaepkens, F. (2004). Accountability in (not-for-profit) hospitals: Evidence of calculative management. Working paper, Tilburg University.

Bowdle, B. F., and Gentner, D. (1997). Informativity and Asymmetry in Comparisons. Cognitive Psychology, 34, 244--286.

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

Burke, J., and Kincanon, E. (1991). Benford's Law and Physical Constants: The Distribution of Initial Digits. American Journal of Physics, 59, 952.

Busta, B., and Weinberg, R. (1998). Using Benford's Law and Neural Networks as a Review Procedure. Managerial Auditing Journal, 13, 356--366.

Calabrese, T. (2011). Public mandates, market monitoring, and nonprofit financial disclosures. Journal of Accounting and Public Policy, 30, 71--88.

Carslaw, C. A. P. N. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. Accounting Review, 63, 321--327.

Christensen, A. L., and Mohr, R. M. (2003). Not-for-profit annual reports: What do museum managers communicate? Financial Accountability & Management, 19, 139--158.

Crawford, L., Dunne, T., Hannah, G., and Stevenson, L. (2009). An Exploration of Scottish Charities' Governance and Accountability. Institute of Chartered Accountants of Scotland, Edinburgh.

Diekmann, A.. (2007). Not the First Digit. Using Benford's Law to Detect Fraudulent Scientific Data. Journal of Applied Statistics, 34, 321--329.

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

Feng, N., Ling, Q., Neely, D., and Roberts, A. (2011). Using Archival Data Sources to Conduct Nonprofit Accounting Research. Working Paper.

Gronjberg, K.. (1991). Managing Grants and Contracts: The Case of Four Nonprofit Social Service Organizations. Nonprofit and Voluntary Sector Quarterly, 20, 5--24.

Hansmann, H. (1987). Economic Theories of Nonprofit Organization. In P. Walker, The nonprofit sector: a research handbook. Yale University Press, New Haven.

Haynes, A. H. (2012). Detecting fraud in bankrupt municipalities using Benford's Law. Scripps Senior Theses.

Hinrichs, J. V., Berie, J. L., and Mosell, M. K. (1982). Place information in multidigit comparison. Memory and Cognition, 10, 487--495.

Holdershaw, J., Gendall, P., and Garland, R. (1997). The widespread use of odd pricing in the retail sector. Marketing Bulletin, 8, 53--58.

Hyndman, N., and McMahon, D. (2011). The hand of government in shaping accounting and reporting in the UK charity sector. Public Money & Management, 31, 167--174.

Ismail, K., and Chandler, R. (2003). The timeliness of quarterly financial reports of companies in Malaysia. Working paper.

Jegers, M. (2013). Do Nonprofit Organisations Manage Earnings? An Empirical Study. Voluntas, 24, 953--968.

Kitching, K.. (2009). Audit value and charitable organizations. Journal of Accounting and Public Policy, 28, 510--524.

Krishnan, R., Yetman, M., and Yetman, R. (2006). Expense Misreporting in Nonprofit Organizations. Accounting Review, 81, 399--420.

Lande, E., and Rocher, S. (2011). Prerequisites for applying accrual accounting in the public sector. Public Money & Management, 31, 219--222.

Leone, A., and Van Horn, L. (2001). Financial reporting incentives in not-for-profit organizations: Evidence from hospitals. Working Paper.

Ley, E. (1996). On the Peculiar Distribution of the U.S. Stock Indexes' Digits. The American Statistician, 50, 311--313.

Mook, L., Handy, F., and Quarter, J. (2007). Reporting volunteer labour at the organizational level: A study of Canadian nonprofits. Voluntas, 18, 55--71.

Mortier, T., and Baten, T. (2010). Empirisch gebruiksonderzoek van jaarrekeningen van verenigingen zonder winstoogmerk. Masterproef UGent. [An Empirical Study of Non-Profit Organizations' Financial Statement Users]

Mosimann, J. E., Wiseman, C. V., and Edelman, R. E. (1995). Data fabrication: Can people generate random digits? Accountability in Research: Policies and Quality Assurance, 4, 31--55.

Newcomb, S. (1881). Note on the Frequency of Use of the Different Digits in Natural Numbers. American Journal of Mathematics, 4, 39--40.

Newton, J., and Ashton, R. (1989). The association between audit technology and audit delay. Auditing: A Journal of Practice and Theory, 8, 22--37.

Nigrini, M. J., and Mittermaier, L. J. (1997). The use of Benford's Law as an aid in analytical procedures. Auditing: A Journal of Practice & Theory, 16, 52--67.

Omer, T. C., and Yetman, R. J. (2003). Near Zero Taxable Income Reporting by Nonprofit Organizations. Journal of the American Taxation Association, 25, 19--34.

Owusu-Ansah, S. (2000). Timeliness of corporate financial reporting in emerging capital markets: Empirical evidence from the Zimbabwe Stock Exchange. Accounting and Business Research, 30, 241--254.

Parsons, L. M. (2003). Is accounting information from nonprofit organizations useful to donors? A review of charitable giving and value-relevance. Journal of Accounting Literature, 22, 104--129.

Pfeffer, J., and Salancik, G. (1978). The External Control of Organizations: A Resource Dependence Perspective. Harper and Row, New York.

Poltrock, S. E., and Schwartz, D. R. (1984). Comparative judgments of multidigit numbers. Journal of Experimental Psychology: Learning, Memory and Cognition, 10, 32--45.

Privett, N., and Erhun, F. (2011). Efficient Funding: Auditing in the Nonprofit Sector. Manufacturing & Service Operations Management, 13, 471--488.

Rauch, B., Göttsche, M., Brähler, G., and Engel, S. (2011). Fact and Fiction in EU-Governmental Economic Data. German Economic Review, 12, 243--255.

Reddy, Y. V., and Sebastin, A. (2012). Entropic analysis in financial forensics. IUP Journal of Accounting Research & Audit Practices, 11, 42--57.

Reheul, A. - M., Van Caneghem, T., and Verbruggen, S. (2014). Financial Reporting Lags in the Non-Profit Sector: An Empirical Analysis. Voluntas, 25, 352--377.

Rosch, E. (1975). Cognitive reference points. Cognitive Psychology, 7, 532--547.

Saxton, G., Neely, D., and Guo, C. (2011). Web Disclosure and the Market for Charitable Contributions. Working Paper.

Schindler, R. M., and Kibarian, T. M. (1996). Increased consumer sales response through use of 99-ending prices. Journal of Retailing, 72, 187--199.

Schindler, R. M., and Wiman, A. R. (1989). Effects of odd pricing on price recall. Journal of Business Research, 19, 165--177.

Sentance, W. A.. (1973). A Further Analysis of Benford's Law. Fibonacci Quarterly, 11, 490--494.

Thomas, J. K.. (1989). Unusual patterns in reported earnings. Accounting Review, 64, 773--787.

Thornton, J., and Belski, W. (2010). Financial Reporting Quality and Price Competition among Nonprofit Firms. Applied Economics, 42, 2699--2713.

Trussel, J. M., and Parsons, L. M. (2008). Financial reporting factors affecting donations to charitable organizations. Advances in Accounting, 23, 263--285.

Twedt, D. W. (1965). Does the “9 fixation” in retail pricing really promote sales?. Journal of Marketing, 29, 54--55.

Van Caneghem, T. (2002). Earnings Management Induced by Cognitive Reference Points. British Accounting Review, 34, 167--178.

Verbruggen, S., Christiaens, J., and Milis, K. (2011). Can Resource Dependence and Coercive Isomorphism Explain Nonprofit Organizations' Compliance With Reporting Standards. Nonprofit and Voluntary Sector Quarterly, 40, 5--32.

Verbruggen, S., Reheul, A. - M., Van Caneghem, T., Dierick, J., Christiaens, J., and Vanhee, C. (2011). Het bedrijfsrevisoraat in de verenigingssector/Le révisorat d'entreprises dans le secteur associatif. Antwerpen/Apeldoorn : ICCI & Maklu. [External Auditing in the Non-Profit Sector]

Verbruggen, S., and Christiaens, J. (2012). Do Non-Profit Organizations Manage Earnings toward Zero Profit and Does Governmental Financing Play a Role?. Canadian Journal of Administrative Sciences, 29, 205--217.

Vermeer, T., Raghunandan, K., and Forgione, D. (2006). The Composition of Nonprofit Audit Committees. Accounting Horizons, 20, 75--90.

Vermeer, T., Raghunandan, K., and Forgione, D. (2009). Audit Fees at U.S. Non-Profit Organizations. Auditing: A Journal of Practice & Theory, 28, 289--303.

Wlodarski, J. (1971). Fibonacci and Lucas Numbers Tend to Obey Benford’s Law. Fibonacci Quarterly, 9, 87--88.

Yetman, M. H., and Yetman, R. J. (2003). The Effect of Nonprofits’ Taxable Activities on the Supply of Private Donations. National Tax Journal, LVI, 243--258.

Yetman, M., and Yetman R. (2012). The Effects of Governance on the Accuracy of Charitable Expenses Reported by Nonprofit Organizations. Contemporary Accounting Research, 29, 738--767.

Zainon, S., Atan, R., Wah, Y. B., and Ahmad, R. A. R. (2012). Information Disclosure by Charity Organizations. In: M.K., Jha, M. Lazard, A. Zaharim, and K. Sopian (eds.), Recent Advances in Business Administration, Harvard, Cambridge.

TABLE 1 : Expected frequencies based on Benford's Law

Digit i	Expected frequency (%) of digit i in the first position	Expected frequency (%) of digit i in the second position
0	-	11.97
1	30.10	11.39
2	17.61	10.88
3	12.49	10.43
4	9.69	10.03
5	7.92	9.67
6	6.70	9.34
7	5.80	9.04
8	5.12	8.76
9	4.58	8.50

TABLE 2 : Descriptive statistics

	Mean	Median	St.Dev.	25th Percentile	75th Percentile
SIZE	8.6732	8.6349	1.3930	7.8742	9.4711
Total assets (th EUR)	16,906	5,625	39,667	2,628	12,979
DONSUBS	.7327	.2278	1.0689	.0000	1.0671
FINLEV	.1667	.0662	.2113	.0000	.2830

SIZE = natural logarithm of total assets (in thousands of EUR); *DONSUBS* = ratio of ‘membership contributions, other contributions, bequests and grants’ over total assets; *FINLEV* = financial debt over total assets

TABLE 3 : Correlation matrix

	(1)	(2)	(3)
SIZE **(1)**	1.0000		
DONSUBS **(2)**	-.3656 *(.000)*	1.0000	
FINLEV **(3)**	.0326 *(.004)*	.0550 *(.000)*	1.0000

SIZE = natural logarithm of total assets (in thousands of EUR); *DONSUBS* = ratio of 'membership contributions, other contributions, bequests and grants' over total assets; *FINLEV* = financial debt over total assets

TABLE 4 : Digital analysis on full sample

Digit i (2nd position)	**0**		**1**		**2**	**3**		**4**		**5**		**6**		**7**		**8**		**9**	
Obs Freq	63,165		54,391		53,158	49,583		47,922		48,217		45,062		43,669		41,859		40,841	
Obs %	12.95		11.15		10.90	10.16		9.82		9.88		9.24		8.95		8.58		8.37	
Exp %	11.97		11.39		10.88	10.43		10.03		9.67		9.34		9.04		8.76		8.50	
% Unit Dev	0.98		-0.24		0.02	-0.27		-0.21		0.21		-0.10		-0.09		-0.18		-0.13	
Z-statistic	21.07	**	-5.28	**	0.31	-6.16	**	-4.84	**	5.09	**	-2.41	*	-2.05	*	-4.37	**	-3.22	**
Chi-squared	530.20	**																	
MAD	0.002																		

Obs Freq = total number of observed digits i in the second position; Obs % = observed frequency (%) of digit i; Exp % = expected frequency (%) of digit i based on Benford's Law; % Unit Dev = percentage unit deviation (being the difference between Obs % and Exp %); ** = statistically significant at the 1% level; * = statistically significant at the 5% level

TABLE 5 : Digital analysis on sub-periods (i.e., 2007-2009 vs. 2010-2012)

Digit i (2nd position)	**0**		**1**		**2**		**3**		**4**		**5**		**6**		**7**		**8**		**9**	
2007 up to 2009																				
Obs Freq	30,550		26,040		25,704		23,774		22,836		23,088		21,831		20,893		20,168		19,495	
Obs %	13.03		11.11		10.97		10.14		9.74		9.85		9.31		8.91		8.60		8.32	
Exp %	11.97		11.39		10.88		10.43		10.03		9.67		9.34		9.04		8.76		8.50	
% Unit Dev	1.07		-0.28		0.08		-0.29		-0.29		0.18		-0.02		-0.12		-0.15		-0.18	
Z-statistic	15.91	**	-4.25	**	1.32		-4.59	**	-4.64	**	2.99	**	-0.38		-2.04	*	-2.61	**	-3.16	**
Chi-squared	305.83	**																		
MAD	0.003																			
2010 up to 2012																				
Obs Freq	32,615		28,351		27,454		25,809		25,086		25,129		23,231		22,776		21,691		21,346	
Obs %	12.87		11.18		10.83		10.18		9.90		9.91		9.16		8.99		8.56		8.42	
Exp %	11.97		11.39		10.88		10.43		10.03		9.67		9.34		9.04		8.76		8.50	
% Unit Dev	0.90		-0.20		-0.05		-0.25		-0.13		0.25		-0.17		-0.05		-0.20		-0.08	
Z-statistic	13.94	**	-3.24	**	-0.83		-4.14	**	-2.26	*	4.18	**	-2.98	**	-0.88		-3.56	**	1.43	
Chi-squared	238.68	**																		
MAD	0.002																			
Difference between both sub-periods																				
% Unit Dev	0.17		-0.07		0.14		-0.04		-0.15		-0.06		0.15		-0.07		0.05		-0.10	
Z-statistic	1.75		-0.82		1.53		-0.44		-1.80		-0.73		1.81		-0.87		0.60		-1.30	
Chi-squared	14.41																			
MAD	0.001																			

Obs Freq = total number of observed digits i in the second position; Obs % = observed frequency (%) of digit i; Exp % = expected frequency (%) of digit i based on Benford's Law; % Unit Dev = percentage unit deviation (being the difference between Obs % and Exp %); ** = statistically significant at the 1% level; * = statistically significant at the 5% level

TABLE 6: Digital analysis on top and bottom decile based on *SIZE* (i.e., natural logarithm total assets in thousands of EUR)

Digit i (2nd position)	0		1	2		3		4	5	6		7	8	9
Top decile														
Obs Freq	6,950		6,243	6,174		5,704		5,529	5,390	5,036		4,969	4,811	4,615
Obs %	12.54		11.26	11.14		10.29		9.98	9.73	9.09		8.97	8.68	8.33
Exp %	11.97		11.39	10.88		10.43		10.03	9.67	9.34		9.04	8.76	8.50
% Unit Dev	0.57		-0.12	0.26		-0.14		-0.05	0.06	-0.25		-0.07	-0.08	-0.17
Z-statistic	4.15	**	-0.92	1.95		-1.08		-0.43	0.46	-2.02*		-0.57	-0.63	-1.46
Chi-squared	27.05	**												
MAD	0.002													
Bottom decile														
Obs Freq	5,104		4,276	3,984		3,700		3,688	3,674	3,325		3,329	3,225	3,121
Obs %	13.64		11.43	10.65		9.89		9.85	9.82	8.88		8.89	8.62	8.34
Exp %	11.97		11.39	10.88		10.43		10.03	9.67	9.34		9.04	8.76	8.50
% Unit Dev	1.67		0.04	-0.24		-0.55		-0.18	0.15	-0.45		-0.14	-0.14	-0.16
Z-statistic	9.95	**	0.22	-1.47		-3.46	**	-1.13	0.97	-3.01	**	-0.95	-0.96	-1.12
Chi-squared	112.90	**												
MAD	0.004													
Difference between both sub-samples														
% Unit Dev	-1.10		-0.16	0.50		0.41		0.12	-0.09	0.20		0.07	0.06	-0.01
Z-statistic	-4.88	**	-0.75	2.37	*	2.01	*	0.61	-0.46	1.06		0.37	0.34	-0.06
Chi-squared	31.65	**												
MAD	0.003													

Obs Freq = total number of observed digits i in the second position; Obs % = observed frequency (%) of digit i; Exp % = expected frequency (%) of digit i based on Benford's Law; % Unit Dev = percentage unit deviation (being the difference between Obs % and Exp %); ** = statistically significant at the 1% level; * = statistically significant at the 5% level

TABLE 7 : Digital analysis on top decile and zero values based on *DONSUBS* (i.e., the ratio of 'membership contributions, other contributions, bequests and grants' over total assets)

Digit i (2nd position)	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
Top decile										
Obs Freq	6,461	5,199	5,193	4,691	4,499	4,586	4,201	4,058	3,911	3,813
Obs %	13.86	11.15	11.14	10.06	9.65	9.84	9.01	8.71	8.39	8.18
Exp %	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
% Unit Dev	1.89	-0.24	0.26	-0.37	-0.38	0.17	-0.32	-0.33	-0.37	-0.32
Z-statistic	12.59 **	-1.60	1.79	-2.61 **	-2.72 **	1.25	-2.41 *	-2.48 *	-2.80 **	-2.48 *
Chi-squared	182.49 **									
MAD	0.005									
Zero values										
Obs Freq	14,797	12,800	12,359	11,661	11,359	11,551	10,624	10,300	9,809	9,664
Obs %	12.88	11.14	10.75	10.15	9.88	10.05	9.24	8.96	8.54	8.41
Exp %	11.97	11.39	10.88	10.43	10.03	9.67	9.34	9.04	8.76	8.50
% Unit Dev	0.91	-0.25	-0.13	-0.29	-0.15	0.38	-0.09	-0.07	-0.22	-0.09
Z-statistic	9.48 **	-2.68 **	-1.39	-3.17 **	-1.66	4.39 **	-1.08	-0.86	-2.66 **	-1.11
Chi-squared	125.42 **									
MAD	0.003									
Difference between both sub-samples										
% Unit Dev	0.99	0.02	0.39	-0.08	-0.23	-0.21	-0.23	-0.26	-0.14	-0.23
Z-statistic	5.31 **	0.09	2.26 *	-0.50	-1.42	-1.29	-1.46	-1.64	-0.94	-1.51
Chi-squared	39.89 **									
MAD	0.003									

Obs Freq = total number of observed digits i in the second position; Obs % = observed frequency (%) of digit i; Exp % = expected frequency (%) of digit i based on Benford's Law; % Unit Dev = percentage unit deviation (being the difference between Obs % and Exp %); ** = statistically significant at the 1% level; * = statistically significant at the 5% level

TABLE 8: Digital analysis top decile and zero values based on *FINLEV* (i.e., the ratio of financial debt over total assets)

Digit i (2nd position)	**0**		**1**		**2**		**3**		**4**		**5**		**6**		**7**		**8**		**9**	
Top decile																				
Obs Freq	6,524		5,375		5,344		5,032		5,042		5,000		4,707		4,583		4,275		4,113	
Obs %	13.05		10.75		10.69		10.07		10.09		10.00		9.41		9.17		8.55		8.23	
Exp %	11.97		11.39		10.88		10.43		10.03		9.67		9.34		9.04		8.76		8.50	
% Unit Dev	1.08		-0.64		-0.19		-0.37		0.05		0.33		0.08		0.13		-0.21		-0.27	
Z-statistic	7.45	**	-4.49	**	-1.39		-2.69	**	0.40		2.52	*	0.60		1.03		-1.63		-2.19	*
Chi-squared	88.89	**																		
MAD	0.003																			
Zero values																				
Obs Freq	17,764		15,297		14,586		13,409		12,998		13,076		12,600		12,015		11,409		11,298	
Obs %	13.21		11.38		10.85		9.97		9.67		9.73		9.37		8.94		8.49		8.40	
Exp %	11.97		11.39		10.88		10.43		10.03		9.67		9.34		9.04		8.76		8.50	
% Unit Dev	1.24		-0.01		-0.03		-0.46		-0.36		0.06		0.03		-0.10		-0.27		-0.10	
Z-statistic	14.05	**	-0.14		-0.39		-5.52	**	-4.44	**	0.71		0.43		-1.26		-3.52	**	-1.28	
Chi-squared	233.91	**																		
MAD	0.003																			
Difference between both sub-samples																				
% Unit Dev	-0.16		-0.63		-0.16		0.09		0.42		0.28		0.04		0.23		0.07		-0.18	
Z-statistic	-0.92		-3.79	**	-0.98		0.59		2.68	**	1.77		0.29		1.54		0.45		-1.22	
Chi-squared	27.74	**																		
MAD	0.002																			

Obs Freq = total number of observed digits i in the second position; Obs % = observed frequency (%) of digit i; Exp % = expected frequency (%) of digit i based on Benford's Law; % Unit Dev = percentage unit deviation (being the difference between Obs % and Exp %); ** = statistically significant at the 1% level; * = statistically significant at the 5% level