7-2016

Data mining journal entries for fraud detection: A replication of Debreceny and Gray's (2010) techniques

Poh Sun SEOW
Singapore Management University, psseow@smu.edu.sg

PAN, Gary
Singapore Management University, garypan@smu.edu.sg

Themin SUWARDY
Singapore Management University, tsuwardy@smu.edu.sg

Published in Journal of Forensic and Investigative Accounting, Volume 8, Issue 3, July 2016, pp. 501-514
http://web.nacva.com/JFIA/Issues/JFIA-2016-No3-7.pdf

DATA MINING JOURNAL ENTRIES FOR FRAUD DETECTION: A REPLICATION OF DEBRECENY AND GRAY'S (2010) TECHNIQUES

Poh-Sun SEOW [1]
Gary PAN [2]
Themin SUWARDY [3]

Singapore Management University
School of Accountancy
60 Stamford Road
Singapore 178900

December 2015

Acknowledgments: We are grateful to an anonymous Big 4 accounting firm for providing the data. This research was supported by the Singapore Ministry of Education (MOE) Academic Research Fund (AcRF) Tier 1 grant (12-C206-SMU-003).

[1] Email: psseow@smu.edu.sg, Tel: (65) 6828-0935, Fax: (65) 6828-0600 (Corresponding Author)
[2] Email: garypan@smu.edu.sg, Tel: (65) 6828-0983, Fax: (65) 6828-0600
[3] Email: tsuwardy@smu.edu.sg, Tel: (65) 6828-0399, Fax: (65) 6828-0600

Data Mining Journal Entries for Fraud Detection: A Replication of Debreceny and Gray’s (2010) Techniques

ABSTRACT

There is limited published research to detect financial statement fraud using digital analysis to analyse journal entry data. As far as we are aware, Debreceny and Gray’s (2010) study is the first and only such study. In this study, we replicated and extended Debreceny and Gray’s (2010) work by examining generalizability of their techniques beyond subjects from USA. Besides Chi-Square test, we also explored the use of mean absolute deviation method during digital analysis. We found Debreceny and Gray’s (2010) techniques useful in facilitating cross-sectional analysis for journal entry data sets that are based on multiple organizations. Our results confirmed that their techniques offered a comprehensive and systematic way of applying digital analysis on journal entries in a new setting. Our analysis also found that researchers should not rely solely on Benford’s Law during digital analysis because of potential false alarms.

Keywords: *fraud; journal entries; data mining; digital analysis; Benford’s Law.*

I. INTRODUCTION

The alarming frequency of fraud occurrences suggests that corporations continue to face persistent threat of fraud (Cecchini et al., 2010a; Summers and Sweeney, 1998). According to Association of Certified Fraud Examiner (ACFE)'s 2014 Report, a typical organization may lose five per cent of its revenue to fraud every year. As such, the consequences of fraud may impact the shareholders, creditors, auditors and the public's confidence in the integrity of corporations' financial systems (Rezaee, 2005).

Frauds can be classified into three primary categories: asset misappropriations, corruption and financial statement fraud (ACFE, 2014). The three different types of fraud would require different fraud detection techniques. For example, bank account reconciliation is a common control plan to detect incidents of employees stealing cash. For financial statement fraud, an emerging area that has attracted researchers' attention is utilizing data analytics to detect "red flags" in the financial statement (Jans et al., 2010; Fanning and Cogger, 1998).

This study focuses on financial statement fraud. In recent years, fraud detection has become a critical component of financial audits and audit standards have heightened emphasis on journal entries as part of fraud detection (Debreceny and Gray, 2010). SAS 122 (AU-C sec. 240) requires auditors to conduct direct assessment of journal entries for fraud risk so that the integrity and validity of the financial results associated with these journal entries will not be compromised. Examples of financial statement fraud involving journal entries include back-posting journal entries, hiding/obscuring entries, manipulating earnings, reserves and revenue, quarter-to-quarter timing issues and subverting approvals (Nieschwietz et al., 2000).

To date, there is very little published research in detecting financial statement fraud using digital analysis to analyse journal entry data. Debreceny and Gray's (2010) study is the first and only such study. Both Grabski (2010) and Kriel (2010) noted that the study was a valuable resource for auditors. To extend and contribute to this area of research, we have chosen to replicate and extend Debreceny and Gray's (2010) digital analysis technique in detecting journal entry fraud. The replication and extension in this study was highly motivated by the need to replicate the technique in another setting beyond one specific dataset from USA. This is important because countries differ in levels of maturity in terms of corporate and data governance. As a result, data quality may differ among countries which in turn, have impact on data analysis work during fraud detection. Therefore, it is useful to replicate whether Debreceny and Gray's (2010) technique can be applied in another dataset obtained from another setting outside USA. Another motivation for us that has methodological implication is to apply mean absolute deviation technique to our dataset during Benford analysis. This is motivated by Nigrini and Miller's (2009) recommendation that mean absolute deviation should be used instead of Chi-Square test because the latter may result in statistically significant differences even when there is no substantive difference in data. The data we used is based on journal entries of 12 Singapore organizations from an anonymous Big 4 accounting firm in Singapore.

II. BACKGROUND

Fraud Detection

Traditionally, fraud detection involves finding indicators of potential fraud, or red flags (Asare and Wright, 2004). To date, there have been many studies conducted in this research area. For instance, Pincus (1989) found that auditors using a standardized red flag program are less successful at correctly identifying fraud risk. Bell and Carcello (2000) developed a logistic regression model to estimate the likelihood of fraudulent financial reporting using red flag data.

In addition, Hansen et al. (1996) constructed a generalized qualitative-response model to analyse management fraud. While such studies are useful in many ways, one major challenge of relying on red flags to identify fraud is that the presence of such symptom is not necessarily indicative of fraud (Albrecht and Romney, 1986) and investigation of such anomalies usually results in a conclusion that fraud was not the underlying cause. Perpetrators may also attempt to conceal their acts and ‘red flags’ may be relatively few in frequency and minor in amount (Hogan et al., 2008). Given the growing complexity, there is a need to explore innovative and efficient way of identifying such red flags. As a consequence, researchers have started to explore the use of data analytics to detect fraud (Jans et al., 2010).

Forensic data analytics includes the use of digital analysis, data mining and data visualization (Ngai et al., 2011; Fanning and Cogger, 1998). Fraud detection, which increasingly relies on fraud data analytics, leverages on advanced technologies and techniques to extract and interpret information in order to uncover complex patterns and indicators of possible fraudulent activities (Spathis et al., 2002). Data analytics allows for analysis of 100 per cent of the data at a much shorter time as compared to manual review and real time red flags can be identified with the use of continuous monitoring. Data analytics can also help companies broaden the range of exposure and increase the ability to uncover new patterns of fraudulent behaviour.

Among various methods in forensic data analytics, one of the most commonly adopted methods is digital analysis using Benford's Law (Cleary and Thibodeau, 2005; Nigrini, 1999; Nigrini and Mittermaier, 1997; Tackett, 2007). Given its potential to identify data points (e.g., transaction amounts) that contain characteristics associated with fraudulent activity, digital analysis using Benford's Law has been proven to hold great promise in fraud detection process (Coderre, 1999). Durtschi et al. (2004) concluded that digital analysis can be a very useful tool for picking out potentially fraudulent accounts for further analysis when used correctly, but should not be overly dependent on such tests.

According to these researchers, this technique should only be applied to suitable accounts which conform to the Benford distribution and should exercise great care in interpreting the results. Reddy and Sebastin (2012) added that this technique is useful where occurrence of digits in the data set deviate significantly from the expected frequencies. Further use of the Benford's Law has also been discussed by Nigrini and Miller (2009), who suggested the use of second-order test. The second-order test involves the calculation of digit frequencies of the differences among the ordered values in a data set, which conforms to the digit frequencies of Benford's Law. There is sufficient empirical evidence that suggests the frequencies of first and second digits of a data set that contains credible numbers will indeed correspond to a Benford’s Law’s probability distribution (Nigrini and Mittermaier, 1997).

Financial Statement Fraud Involving Journal Entries

Fraud detection has become a critical component of financial audits and audit standards have heightened emphasis on journal entries as part of fraud detection (Debreceny and Gray, 2010). As accounting standards continue to evolve and grow in complexity, it is likely financial reporting will rely more on estimates and judgment. Although many corporations have invested in enterprise resource planning systems to automate financial reporting, these corporations still have hundreds – sometimes thousands – of manual journal entries being prepared, entered, reviewed and approved each month.

In general, journal entries can be either routine transactions, non-routine transactions or estimation transaction. Routine transactions are recurring business activities that are recorded in the normal flow of transaction. Non-routine transactions are entries made periodically and are usually not part of the normal course of business. Estimation transactions usually involve management’s decision and judgment due to precise measurements. Non-routine transactions can possibly be entered in a corporation’s accounting systems on any day of the year by its employee, which makes it difficult and tedious to identify any potential fraudulent activities through manual means. In the case of routine journal entries, it is easy to identify the person who made such an entry and thus, the focus ought to be on the motive, opportunity and pressure that the individual is undergoing, in order to identify possible red flags. To better detect fraud, Lanza et al. (2007) highlighted the usefulness of data analytics for journal entry testing, especially in the case of top-side journal entry. Debreceny and Gray (2010) believed that data mining could potentially improve both effectiveness and efficiency of auditors in their analysis of journal entries and fraud detection.

Given the important role of fraud detection, the increasing emphasis on financial statement fraud involving journal entries and the fact that Debreceny and Gray's (2010) study being the only paper published in this field so far, we believe there is a need to replicate Debreceny and Gray's (2010) digital analysis technique before future research in this area moves to the next phase of development - theorization stage, as suggested by Grabski (2010). It is our aim to narrow this gap in the fraud detection literature.

The rest of the paper is organized as follows: First, we present descriptive analysis of journal entries. This is followed by a summary of the statistical analysis of journal entries. We then conclude by highlighting our contribution and future research opportunities.

III. DESCRIPTIVE ANALYSIS OF JOURNAL ENTRIES

Getting access to real-world journal entries data set is a major difficulty for researchers (Debreceny and Gray, 2010). Debreceny and Gray (2010) indicated that they were fortunate to have access to journal entries for 27 organizations from an anonymous software vendor. We were also fortunate to obtain access to manual journal entries for 12 organizations from an anonymous Big 4 accounting firm. Due to clients' confidentiality reasons, the Big 4 accounting firm removed all identifying information before providing the data set. The 12 organizations are labelled sequentially as Entity A, B, and so forth. All monetary amounts are in Singapore Dollars (1SGD is approximately 0.80 USD).

Table 1 shows the descriptive statistics for the 12 set of charts of accounts. For this study, the average number of active accounts is 324, which is higher than the 164 reported in Debreceny and Gray (2010). This indicates that the charts of accounts in this study are more complex than Debreceny and Gray (2010). However, the maximum number of active accounts is much lower at 475, compared to 1,036 in Debreceny and Gray (2010). Table 2 shows the number of transactions per active account for the 12 organizations. There are four organizations which have more than 500 average transactions per account: Entity L (3,649), Entity K (904), Entity F (578), and Entity I (504).

Table 1: Active accounts in organizational chart of accounts.

Minimum active accounts	184
Maximum active accounts	475
Median active accounts	335
Average active accounts	324

Table 2: Transactions per account in organizational chart of accounts.

Entity	Minimum	Maximum	Median	Mean
A	1	12,506	31	281
B	1	12,229	51	328
C	1	2,283	29	93
D	1	23,528	47	310
E	1	906	13	41
F	1	66,607	47	578
G	1	4,570	43	159
H	1	906	43	78
I	1	14,706	96	504
J	1	2,839	44	171
K	1	56,396	48	904
L	1	81,675	339	3,649

Table 3 shows the descriptive statistics for the 12 organizations. The first column shows the number of line items and it varies widely. There is a big difference between Entity L which has the highest number of line items (671,487) and Entity E which has the lowest number of line items (7,828). The same four entities which have the most number of transactions per account in Table 2 (Entity F, I, K and L) also have the most number of line items in Table 3. However, Entity K has a relatively lower dollar values ($7.4 million) compared to the other three entities. Although Entity A does not have a high number of line items (98,870), it has the highest total dollar values ($521 billion). Table 4 shows the number of line items per journal entry. There are two entities which have very large number of line items per journal entry: Entity L (11,126) and Entity F (7,006), indicating presence of "mega entries". Debreceny and Gray (2010) suggested that such "mega entries" could be due to automatic journal entries that reverse prior-period adjusting journal entries or data transferred from subsidiary systems.

Table 3: Descriptive statistics for organizations.

Entity	Number of line items	Total line items $	Maximum line items $
A	95,870	521,482,242,735	1,882,075,421
B	127,417	69,022,877,555	743,269,971
C	25,903	4,313,833,127	38,508,133
D	147,411	74,388,468,924	300,000,000
E	7,828	1,709,511,347	73,168,000
F	203,497	3,430,793,755	100,000,000
G	52,135	45,286,821,791	145,014,787
H	18,122	657,281,116	12,267,261
I	200,465	14,674,615,034	52,827,698
J	46,954	63,687,777,318	334,124,141
K	397,021	3,356,572,444	7,392,861
L	671,487	198,297,856,683	5,868,498,113
Total	1,994,110	1,000,308,651,829	

Table 4: Line items per journal entry.

Entity	Journal entries	Mean	Std Dev	Min	Max
A	25,329	4	5	2	120
B	29,836	4	19	2	593
C	6,987	4	5	2	60
D	24,842	6	15	2	368
E	2,311	3	3	2	40
F	23,006	9	125	2	7,006
G	11,080	5	14	2	242
H	3,429	5	13	2	156
I	29,547	7	38	2	753
J	12,044	4	9	2	173
K	100,246	4	10	2	389
L	123,876	5	140	2	11,126
Total	392,533				

IV. STATISTICAL ANALYSIS OF JOURNAL ENTRIES

This section examines the statistical analysis of the data set.

First Digit

Benford's Law is one of the most widely-used digital analysis techniques to detect fraud (Benford, 1938; Debreceny and Gray, 2010; Nigrini and Miller, 2009; Nigrini and Mittermaier, 1997). Benford's Law proposes an expected probability (Expected%) for the first digit of numbers in data sets as shown in Table 5.

Table 5: Expected first digit distribution under Benford's Law.

Digit	Probability
1	30.1%
2	17.6%
3	12.5%
4	9.7%
5	7.9%
6	6.7%
7	5.8%
8	5.1%
9	4.6%

Table 6 shows the actual probability (Actual%) for the first digit of numbers and the difference between the actual and expected probability (Diff = Actual% - Expected%) for the 12 organizations. The Chi-square and p-value for each entity are also shown. The results indicate that each of the 12 organizations has an actual probability of first digits that differs significantly from the Benford's Law expected probability. This indicates many red flags which require further investigation by auditors. However, these results may not be useful practically to auditors as it does not address the "so-what" question. Even though there is a statistically significant difference between the actual and expected probability for each organization in this data set, how should auditors deal with practical concerns to extend the investigation? Does this mean that auditors should examine each of the 12 organizations? Which of these 12 organizations should auditors extend their audit procedures? This may cause practical constraints for auditors as this data set consists of 392,533 journal entries and 1,994,110 line items. In addition, most of the differences in this data set are below 1%. It may not be meaningful to conduct further investigation with such small differences. This suggests that the Benford’s Law may not be useful in detecting fraud for this data set.

Table 6: Observed first digit distributions in journal entry database.

Entity	Data	1	2	3	4	5	6	7	8	9
A	Count	29,907	16,625	11,552	9,386	7,570	6,345	5,759	4,511	4,215
χ^2=104.2	Act%	31.2%	17.3%	12.1%	9.8%	7.9%	6.6%	6.0%	4.7%	4.4%
P<0.0001	Diff	1.1%	-0.3%	-0.4%	0.1%	0%	-0.1%	0.2%	-0.4%	-0.2%
B	Count	38,498	22,074	15,516	12,187	10,187	8,315	7,625	7,044	5,971
χ^2=81.7	Act%	30.2%	17.3%	12.2%	9.6%	8.0%	6.5%	6.0%	5.5%	4.7%
P<0.0001	Diff	0.1%	-0.3%	-0.3%	-0.1%	0.1%	-0.2%	0.2%	0.4%	0.1%
C	Count	7,640	4,514	3,331	2,538	2,005	1,901	1,614	1,192	1,168
χ^2=44.5	Act%	29.5%	17.4%	12.9%	9.8%	7.7%	7.4%	6.2%	4.6%	4.5%
P<0.0001	Diff	-0.6%	-0.2%	0.4%	0.1%	-0.2%	0.6%	0.4%	-0.5%	-0.1%
D	Count	45,666	27,074	18,330	13,763	11,827	9,217	8,413	6,742	6,379
χ^2=260.6	Act%	31.0%	18.4%	12.4%	9.3%	8.0%	6.3%	5.7%	4.6%	4.3%
P<0.0001	Diff	0.9%	0.8%	-0.1%	-0.4%	0.1%	-0.4%	-0.1%	-0.5%	-0.3%
E	Count	1,849	1,432	1,272	811	683	590	405	443	343
χ^2=228.7	Act%	23.6%	18.3%	16.3%	10.4%	8.7%	7.5%	5.2%	5.7%	4.4%
P<0.0001	Diff	-6.5%	0.7%	3.8%	0.7%	0.8%	0.8%	-0.6%	0.5%	-0.2%
F	Count	61,891	34,900	27,952	18,918	16,601	14,064	11,369	9,736	8,066
χ^2=578.3	Act%	30.4%	17.2%	13.7%	9.3%	8.2%	6.9%	5.6%	4.8%	4.0%
P<0.0001	Diff	0.3%	-0.5%	1.2%	-0.4%	0.2%	0.2%	-0.2%	-0.3%	-0.6%
G	Count	15,809	9,589	6,097	5,377	4,204	3,523	2,752	2,622	2,162
χ^2=117.0	Act%	30.3%	18.4%	11.7%	10.3%	8.1%	6.8%	5.3%	5.0%	4.2%
P<0.0001	Diff	0.2%	0.8%	-0.8%	0.6%	0.2%	0.1%	-0.5%	-0.1%	-0.4%
H	Count	5,489	3,321	2,199	1,445	1,456	1,370	960	966	916
χ^2=101.6	Act%	30.3%	18.3%	12.1%	8.0%	8.0%	7.6%	5.3%	5.3%	5.1%
P<0.0001	Diff	0.2%	0.7%	-0.4%	-1.7%	0.1%	0.9%	-0.5%	0.2%	0.5%
I	Count	60,419	35,020	25,924	19,155	1,6075	13,218	11,866	9,979	8,809
χ^2=72.5	Act%	30.1%	17.5%	12.9%	9.6%	8.0%	6.6%	5.9%	5.0%	4.4%
P<0.0001	Diff	0%	-0.1%	0.4%	-0.1%	0.1%	-0.1%	0.1%	-0.1%	-0.2%
J	Count	14,976	8,131	5,260	4,348	3,640	3,483	2,644	2,426	2,046
χ^2=171.2	Act%	31.9%	17.3%	11.2%	9.3%	7.8%	7.4%	5.6%	5.2%	4.4%
P<0.0001	Diff	1.8%	-0.3%	-1.2%	-0.4%	-0.2%	0.7%	-0.2%	0.1%	-0.2%
K	Count	136,981	74,076	41,618	27,453	26,206	19,991	22,738	24,595	23,363
χ^2=12,128.3	Act%	34.5%	18.7%	10.5%	6.9%	6.6%	5.0%	5.7%	6.2%	5.9%
P<0.0001	Diff	4.4%	1.1%	-2.0%	-2.8%	-1.3%	-1.7%	-0.1%	1.1%	1.3%
L	Count	208,400	104,220	75,502	48,200	74,741	33,223	65,120	24,446	37,635
χ^2=40,911.7	Act%	31.0%	15.5%	11.2%	7.2%	11.1%	5.0%	9.7%	3.6%	5.6%
P<0.0001	Diff	0.9%	-2.1%	-1.3%	-2.5%	3.2%	-1.8%	3.9%	-1.5%	1.0%

As an alternative to the Chi-Square test employed in Debreceny and Gray (2010), this study examines the mean absolute deviation technique. Grabski (2010) argued that the mean absolute deviation technique should be used instead of the Chi-Square test. Grabski (2010) stated that it is expected for Chi-Square test to indicate statistical significance when there are many observations. Due to the large observations, it is likely that any small deviation from the expected probability will result in statistically significant results. Nigrini and Miller (2009) also recommended that the Chi-Square test should not be employed as it will result in statistically significant differences even when there is no substantive difference.

Table 7 presents the total deviation and mean absolute deviation for the 12 organizations. The results show that the mean absolute deviation value for each of the 12 organizations was zero, indicating that there were no substantive differences between the actual and expected probability. Relying on the Benford's Law alone in this study would have given rise to a false alarm. Researchers who are analysing journal entry data would be dealing with big data sets containing large observations. Thus they should not rely solely on the Benford's Law and also consider the mean absolute deviation technique as recommended by Grabski (2010) and Nigrini and Miller (2009).

Table 7: Mean absolute differences for first digit distributions in journal entry database.

Entity	Total deviation	Mean absolute deviation
A	0.03	0.00
B	0.02	0.00
C	0.03	0.00
D	0.03	0.00
E	0.15	0.00
F	0.04	0.00
G	0.04	0.00
H	0.05	0.00
I	0.01	0.00
J	0.05	0.00
K	0.16	0.00
L	0.18	0.00

Last Digits

Besides the first digit analysis, researchers also examined the last digits as fraudulent journal entries contain "round numbers or a consistent ending number" (CAQ 2008; Debreceny and Gray, 2010). Each number (0 to 9) is expected to have a uniform distribution of 10% for the last digits. Table 8 shows the actual probability of the fourth digit for all dollar amounts greater than $999. The results in Table 8 indicate that each of the 12 organizations has an actual probability of fourth digit that differs significantly from the expected uniform probability. For each of the 12 organizations, the biggest difference relates to the zero digit. Entity B has the biggest difference of 21.9% for the zero digit. Although this may suggests "round numbers or a consistent ending number" (CAQ 2008), it is also common for the fourth digit to be zero such as $1,000, $10,000, $100,000, etc. Thus, auditors need to identify situations in organizations that make certain digits appear more often, which may explain the higher actual probability.

Table 8: Observed fourth digit distribution in journal entry database.

Entity	Data	0	1	2	3	4	5	6	7	8	9
A	Count	10,013	3,993	4,607	4,254	4,113	4,778	4,234	4,099	4,555	4,092
χ^2=6,145.7	Act%	20.5%	8.2%	9.5%	8.7%	8.4%	9.8%	8.7%	8.4%	9.4%	8.4%
P<0.0001	Diff	10.5%	-1.8%	-0.5%	-1.3%	-1.6%	-0.2%	-1.3%	-1.6%	-0.6%	-1.6%
B	Count	20,876	4,764	4,789	4,769	4,952	6,779	4,852	4,734	4,749	4,267
χ^2=35,393.9	Act%	31.9%	7.3%	7.3%	7.3%	7.6%	10.3%	7.4%	7.2%	7.3%	6.5%
P<0.0001	Diff	21.9%	-2.7%	-2.7%	-2.7%	-2.4%	0.3%	-2.6%	-2.8%	-2.7%	-3.5%
C	Count	2,233	1,325	1,338	1,363	1,231	1,301	1,357	1,430	1,126	1,313
χ^2=591.0	Act%	15.9%	9.5%	9.6%	9.7%	8.8%	9.3%	9.7%	10.2%	8.0%	9.4%
P<0.0001	Diff	5.9%	-0.5%	-0.4%	-0.3%	-1.2%	-0.7%	-0.3%	0.2%	-2.0%	-0.6%
D	Count	9,720	4,455	4,682	4478	5,112	5,123	5,148	4,468	4,645	4,463
χ^2=4,430.7	Act%	18.6%	8.5%	9.0%	8.6%	9.8%	9.8%	9.8%	8.5%	8.9%	8.5%
P<0.0001	Diff	8.6%	-1.5%	-1.0%	-1.4%	-0.2%	-0.2%	-0.2%	-1.5%	-1.1%	-1.5%
E	Count	636	272	329	350	292	309	373	235	367	266
χ^2=332.3	Act%	18.5%	7.9%	9.6%	10.2%	8.5%	9.0%	10.9%	6.9%	10.7%	7.8%
P<0.0001	Diff	8.5%	-2.1%	-0.4%	0.2%	-1.5%	-1.0%	0.9%	-3.1%	0.7%	-2.2%
F	Count	16,337	4,960	4,418	4657	5,798	5,625	5,113	4,779	5,150	4,508
χ^2=19,148.8	Act%	26.6%	8.1%	7.2%	7.6%	9.5%	9.2%	8.3%	7.8%	8.4%	7.4%
P<0.0001	Diff	16.6%	-1.9%	-2.8%	-2.4%	-0.5%	-0.8%	-1.7%	-2.2%	-1.6%	-2.6%
G	Count	5,720	2,409	2,979	2,780	2,571	2,799	2,976	2,664	2,760	2,515
χ^2=2,792.0	Act%	19.0%	8.0%	9.9%	9.2%	8.5%	9.3%	9.9%	8.8%	9.2%	8.3%
P<0.0001	Diff	9.0%	-2.0%	-0.1%	-0.8%	-1.5%	-0.7%	-0.1%	-1.2%	-0.8%	-1.7%
H	Count	1,462	657	714	679	620	846	877	780	621	680
χ^2=715.2	Act%	18.4%	8.3%	9.0%	8.6%	7.8%	10.7%	11.0%	9.8%	7.8%	8.6%
P<0.0001	Diff	8.4%	-1.7%	-1.0%	-1.4%	-2.2%	0.7%	1.1%	-0.2%	-2.2%	-1.4%
I	Count	21,709	8,463	9,288	8,512	8,846	10,384	8,853	8,806	8,920	8,622
χ^2=14,539.9	Act%	21.2%	8.3%	9.1%	8.3%	8.6%	10.1%	8.7%	8.6%	8.7%	8.4%
P<0.0001	Diff	11.2%	-1.7%	-0.9%	-1.7%	-1.4%	0.1%	-1.3%	-1.4%	-1.3%	-1.6%
J	Count	2,724	2,508	2,545	2,352	2,322	2,434	2,714	2,438	2,526	2,502
χ^2=63.9	Act%	10.9%	10.0%	10.2%	9.4%	9.3%	9.7%	10.8%	9.7%	10.1%	10.0%
P<0.0001	Diff	0.9%	0%	0.2%	-0.6%	-0.7%	-0.3%	0.8%	-0.3%	0.1%	0%
K	Count	23,625	6,419	7,024	5,890	7,654	7,621	7,160	7,027	7,403	6,150
χ^2=29,564.1	Act%	27.5%	7.5%	8.2%	6.9%	8.9%	8.9%	8.3%	8.2%	8.6%	7.2%
P<0.0001	Diff	17.5%	-2.5%	-1.8%	-3.1%	-1.1%	-1.1%	-1.7%	-1.8%	-1.4%	-2.8%
L	Count	74,171	27,411	29,963	25,689	26334	27,680	24,788	24,132	25,982	24,704
χ^2=67,208.3	Act%	23.9%	8.8%	9.6%	8.3%	8.5%	8.9%	8.0%	7.8%	8.4%	8.0%
P<0.0001	Diff	13.9%	-1.2%	-0.4%	-1.7%	-1.5%	-1.1%	-2.0%	-2.2%	-1.6%	-2.0%

The Hartigan and Hartigan (1985) dip test of unimodality can also be used to investigate the last three digits (Debreceny and Gray, 2010). The last three digits range from zero to 999. Table 9 shows the dip test results for the 12 organizations. Except for Entity J (p-value = 0.099), each of the organizations had statistically significant dip test values where $p < 0.01$.

Table 9: Last three digits – dip test.

Entity	N	Dip	p	Low	High	Mean
A	48,738	0.006	0.000	0	0	0
B	65,531	0.008	0.000	0	0	0
C	14,017	0.008	0.000	0	0	0
D	52,294	0.008	0.000	0	0	0
E	3,429	0.017	0.001	247	301	276.3
F	61,345	0.012	0.000	0	0	0
G	30,173	0.006	0.001	0	0	0
H	7,936	0.010	0.001	0	200	99.0
I	102,403	0.008	0.000	0	0	0
J	25,065	0.004	0.099	0	0	0
K	85,973	0.010	0.000	0	0	0
L	310,854	0.013	0.000	0	0	0

Debreceny and Gray (2010) stated that the number of journal entries may be too high to be investigated using the Hartigan and Hartigan (1985) dip test. They recommended that the results have to be considered together with the number of accounts involved. They also indicated that investigators should focus on those journal entries that have high deviation from the expected distribution and involving a relatively small number of accounts. Otherwise, the investigation cost will be too high.

We followed the methods of Debreceny and Gray (2010). First, we selected the last three digits of all line items greater than $1,000. Second, we identified the five most common sets of last three digits. Next, we plotted the proportion of these five most common sets against the number of accounts involved (see Figure 1). Figure 2 includes only those journal entries totalling at least $1,000.

Figure 1: Proportion of journal entries to number of accounts line items greater than $1,000

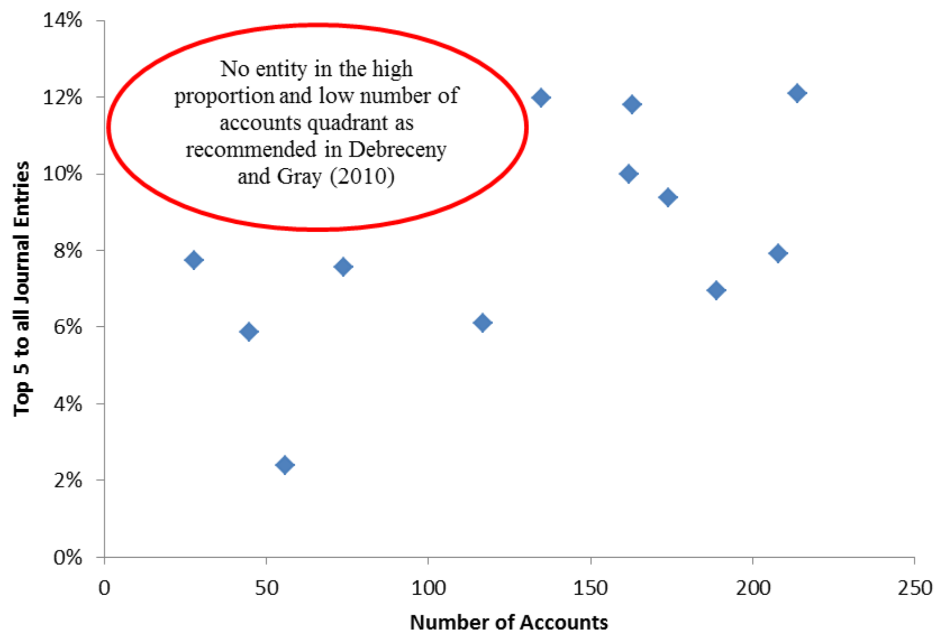
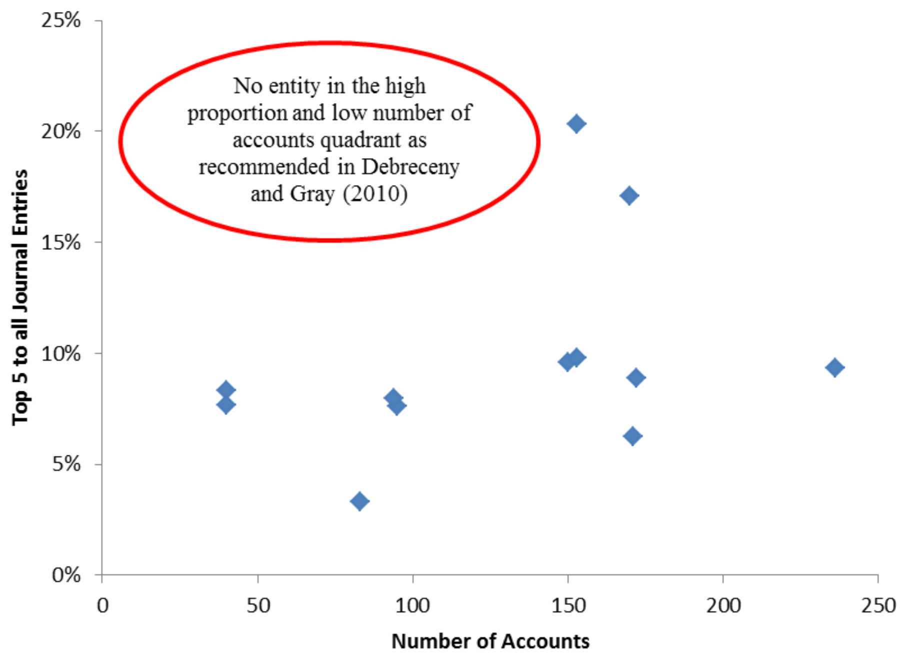


Figure 2: Proportion of journal entries to number of accounts journal entries greater than $1,000



Based on Figures 1 and 2, it appears that there is no entity which exhibits high deviation from the expected distribution and involving a relatively small number of accounts. This suggests that the investigation cost of checking the last three digits may be high for this study. This is unlike the findings in Debreceny and Gray (2010) which reported four entities exhibiting high deviation (30 to 60% deviation) and involving a relatively small number of accounts (less than 40 accounts).

V. CONCLUSION

In this study, we have replicated and extended Debreceny and Gray's (2010) work by examining generalizability of their technique beyond one specific dataset from USA. We believe it is important to replicate Debreceny and Gray's (2010) technique given that it is the first and only study published in this area so far. Besides Chi-Square test, we also explored the use of mean absolute deviation method during digital analysis. We found Debreceny and Gray's (2010) technique useful in facilitating cross-sectional analysis for journal entry data sets that are based on multiple organizations. Our results confirmed that their technique offered a comprehensive and systematic way of applying digital analysis on journal entries in a new setting. Our analysis also found that researchers should not rely solely on Benford's Law during digital analysis because of potential false alarms.

A limitation of this study is the generalizability of the findings to other data sets. This study involved journal entries of 12 organizations from an anonymous Big 4 accounting firm. A major difficulty for researchers in this research area is getting access to real-world journal entries data set. As the obtained data set will typically exclude all identifying information, researchers are unable to examine other variables such as the size of the company and industry. In addition, it is unlikely for external users such as financial analysts, investment managers and government regulators to obtain internal journal entries data set. Thus, external users need to explore other data mining techniques using publicly-available data such as the annual report. One possibility is to use text analytics to examine the annual report (Cecchini et al., 2010b; Goel et al., 2010; Li, 2008). Besides the few pages of financial statements, the rest of the annual report consists of text including the CEO and Chairman’s Letter to the Stockholders, Management Discussion and Analysis and the Notes to the Financial Statements. Li (2008) found that companies with lower earnings produce annual reports that are harder to read, suggesting that companies may be opportunistically structuring the annual reports to hide adverse information from investors. Goel et al. (2010) also analyzed the text in the annual report to detect fraud and their results indicated that using linguistic features is an effective mean for fraud detection. Cecchini et al. (2010b) analyzed Management Discussion and Analysis sections of annual reports and created a dictionary of discriminating concepts to detect fraud. Text analytics can also be applied on the tone of companies’ earnings press releases (Huang et al, 2014). Huang et al, 2014 found that companies use strategic tone management to mislead investors about firm fundamentals.

We hope other researchers will complement our work. There are some related research questions to be investigated. For example, future research should focus on theory development. With theorizing, it may offer insights to identifying what are the characteristics of a fraudulent event and what are the most suitable data mining techniques to detect fraud. Another limitation is that using digital analysis to analyse journal entry data may not be applicable for all types of frauds. Different types of fraud (asset misappropriations, corruption and financial statement fraud would require different fraud detection techniques. Essentially, the aim is to offer more information on how to mine journal entries in an optimal manner. Another future research may involve planting seeded errors in the journal entry data set and using digital analysis techniques to identify these seeded errors.

REFERENCES

ACFE Association of Certified Fraud Examiner (2014) ACFE report to the nations on occupational fraud and abuse, 2014 Global Fraud Study, Technical Report.

Albrecht, W., and Romney, M. (1986) Red flagging management fraud: a validation. *Advances in Accounting*, (3):323-333.

Asare, S., and Wright, A. (2004) The effectiveness of alternative risk assessment and programme planning tools in fraud setting. *Contemporary Accounting Research*, 21(2):325-352.

Bell, T., and Carcello, J. (2000) Research Notes, a decision aid for assessing the likelihood of fraudulent financial reporting. *Auditing: A Journal of Practice & Theory*, 19(1):169-175.

Benford, F. (1938) The law of anomalous numbers. *Proc Am Philos Soc*, (78):551-572.

CAQ Centre for Audit Quality (2008) Practice aid for testing journal entries and other adjustments pursuant to AY section 316. Washington, DC.

Cecchini, M., Aytug, H., Koehler, G., and Pathak, P. (2010a) Detecting management fraud in public companies. *Management Science*, 56(7):1146-1160.

Cecchini, M., Aytug, H., Koehler, G., and Pathak, P. (2010b) Making words work: Using financial text as a predictor of financial events. *Decision Support Systems*, 50(1):164-175.

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

Coderre, D. (1999) Computer-assisted techniques for fraud detection. *The CPA Journal*, (69):57-59.

Debreceny, R., and Gray, G. (2010) Data mining journal entries for fraud detection: an exploratory study. *International Journal of Accounting Information Systems*, (11):157-181.

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

Fanning, K., and Cogger, K. (1988) Neural network detection of management fraud using published financial data. *International Journal of Intelligent Systems Accounting, Finance and Management*, (7):21-41.

Grabski, S. (2010) Discussion of 'Data mining journal entries for fraud detection: an exploratory study'. *International Journal of Accounting Information Systems*, (11):182-185.

Goel, S., Gangolly, J. Faerman, S.R. and Uzuner, O. (2010) Can linguistic predictors detect fraudulent financial filings? *Journal of Emerging Technologies in Accounting*, 7(1):25-46.

Hansen, J., McDonald, J., Messier, W., and Bell, T. (1996) A generalized qualitative-response model and the analysis of management fraud. *Management Science*, 42(7):1022-1033.

Hartigan, J., and Hartigan, P. (1985) The dip test of unimodality. *The Annals of Statistics*, 13(1):70-84.

Hogan, C., Rezaee, Z., Riley Jr., R., and Velury, U. (2008) Financial statement fraud: insights from the academic literature. *Auditing: A Journal of Practice & Theory*. 27(2):231-252.

Huang, X., S.H. Teoh and Y. Zhang (2014) Tone management. *The Accounting Review*. 89(3):1083-1113.

Jans, M., Lybaert, N., and Vanhoof, K. (2010) Internal fraud risk reduction: results of a data mining case study. *International Journal of Accounting Information Systems*, (11):17-41.

Kriel, E. (2010) Discussion of 'Data mining journal entries for fraud detection: an exploratory study'. *International Journal of Accounting Information Systems*, (11):186-188.

Lanza, R. B., Gilbert, S., & Lamoreaux, M. (2007). A risk-based approach to journal entry testing. *Journal of Accountancy*, 204(1):32-35.

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

Ngai, E., Hu, Y., Wong, Y., Chen, Y., and Sun, X. (2011) The application of data mining techniques in financial fraud detection: a classification framework and an academic review of literature. *Decision Support Systems*, (50):559-569.

Nieschwietz, R., Schultz, J., and Zimbelman, M. (2000) Empirical research on external auditors' detection of financial statement fraud. *Journal of Accounting Literature*, (19):52-66,

Nigrini, M. (1999) I've got your number. *Journal of Accountancy*, 187 (May):79-83.

Nigrini, M., and Mittermaier, L. (1997) The use of Benford's Law as an aid in analytical procedures. *Auditing: A Journal of Practice & Theory*, 16(2):52-67.

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

Pincus, K. (1989) The efficacy of a red flags questionnaire for assessing the possibility of fraud. *Accounting, Organization and Society*, 14(1-2):153-163.

Reddy, Y. V., & Sebastin, A. (2012) Entropic Analysis in Financial Forensics. *Journal of Accounting Research and Audit Practices*, 11(3):42-57.

Rezaee, Z. (2005) Causes, consequences, and deterrence of financial statement fraud. *Critical Perspectives on Accounting*, 16(3):277-298.

Spathis, C., Doumpos, M., and Zopounidis, C. (2002) Detecting falsified financial statements: a comparative study using multicriteria analysis and multivariate statistical technique. *The European Accounting Review*, 11(3):509-535.

Summers, S., and Sweeney, J. (1998) Fraudulently misstated financial statements and insider trading: an empirical analysis. *The Accounting Review*, 73(1):131-146.

Tackett, J. (2007) Digital analysis: a better way to detect fraud. *Journal of Corporate Accounting Finance*, 18(4):328-333.