UDC 004.42:005
Original Scientific Paper

Analysis of Top 500 Central and East European Companies Net Income Using Benford's Law

Mario Žgela *mario.zgela@hnb.hr*
Croatian National Bank
Zagreb

Jasminka Dobša *jasminka.dobsa@foi.hr*
University of Zagreb
Faculty of Organization and Informatics Varaždin

Abstract

There are numerous useful methods that can be conducted in data analysis in order to check data correctness and authenticity. One of contemporary and efficient methods is application of so-called Benford's Law. In this paper we examine ways of application of this law in investigation of specific net income number set. Our aim is to make a conclusion if this number set conforms to Benford's Law. An examination target focus is set on values of top 500 central and east European companies according to their income. Data set contains 1,500 records and spans through 3 years (2007, 2008 and 2009) including 500 net incomes per year. Research is based on net income profit and loss subsets as well as absolute values of net income. Analysis covers first digit Benford's Law test and proves conformance to Benford's Law of all observed subsets.

Keywords: Benford's Law, first digit test, net income values analysis, data auditing

1. Introduction

Aim of this research is to perform data analysis of net income (profit or loss) of top 500 central and east European companies according to their business results in three years (2007, 2008 and 2009). Research is based on Deloitte yearly reports [25]. Our hypothesis is that observed data sets conform Benford’s law. If such hypothesis will be confirmed that will allow application of Benford’s law to similar data sets in order to discover possible irregularities and frauds present in data.

Benford's Law is extremely valuable method for data analysis. It proved its efficiency in number of business cases and fields of human interest as presented in [5], [9], [12], [13], [16], [17] and [18].

As it is stated in [25], underlying Deloitte document "Central Europe (CE) Top 500" ranks the 500 largest companies according to their income and net income. It draws on the knowledge and insights of Deloitte’s professionals along with renowned economists and academics to provide a valuable commentary on the CE markets and current trends. With individual country reports and special focus on seven industry sectors and analysis of the dramatically differing corporate results seen between 1st quarter of 2008 and 1st quarter of 2009 it provides an in-depth and wide-ranging insight of the latest Central European developments. So, the purpose of the "Central Europe Top 500" is to rank the 500 largest companies and provide commentary and insights from Deloitte’s professionals along with renowned economists and top executives from some of the most prominent business across the region.

Changes in economy are often described and explained by changes in income and net income values of business companies. Not bearing in mind considerable dependency of national and world economies on that parameter, it is especially interesting to answer the question if it is possible to note whether there are some irregularities or deviations in income and net income values of companies.

In this paper, net income values in euros are analyzed for top 500 companies within each year in time span 2007-2009. Analysis focuses on profit but also on loss values and absolute values.

Examination of conformity with Benford's law according to first digit i.e. analysis of frequencies of certain first digit in net income/loss/absolute value is in focus. Intuitively, it may be assumed that payment amounts are uniformly distributed which means that probability of appearance of each digit on leading (most important, most valuable) position is equal. However, many data sets (in nature, business etc.) do not conform to uniform, but to Benford's law distribution. In the next section properties necessary for data set to conform to Benford's law are listed. According to our estimation observed data sets obey necessary properties and conformation to Benford's law is proven.

2. Research Method

2.1. Benford's Law

American astronomer Simon Newcomb was first who found out that numbers more frequently begin with smaller digits than with greater digits. Newcomb noticed that pages in logarithm tables were dirtier at the start, i. e. more used, and progressively cleaner as approaching to the end. He concluded that numbers more often begin with digit 1 than with any other digit, and in addition, that probability of each following digit (up to 9) at the most significant position in number progressively decreases.

Frank Benford gathered more than 20,000 observations from different sources (geographical area, population, river areas, physical constants etc.). He analysed frequencies of first digits for each number set. After he summarized all individual analyses he concluded that probability of first digit being 1 is 0.30103 which equals $\log_{10}2$, probability of first digit being 2 is 0.17609 which equals $\log_{10}3/2$ etc.

There is a rather extensive literature on various fields of usage of Benford's Law. Also, there are numerous works carried out on application of Benford's Law in information systems auditing i.e. data analysis for auditing purposes ([9], [12], [16], [18], [23]).

Benford's Law is used as a primary data auditing method in above mentioned setting. It defines expected digit frequencies in certain number sets. It is noticeable that in sets of numbers from many data sources, certain digits are distributed in a particular way which significantly differs from uniform distribution. According to the first digit Benford's Law, digit "1" appears as the first digit in number for almost one third of the time, and larger digits appear on the leading number position with lower and lower frequencies. E.g. digit "8" appears as a first digit in slightly more than 5% of values, while digit "9" appears as a first digit in slightly more than 4.5% numbers. The basis for Benford's Law lies in the fact that values of real world data sources are often distributed logarithmically, while the logarithms of these real world data sources are distributed uniformly. Benford's Law may be applied to any position in number and to n first digits (where n is less than total number of digits). However, the most often used are the first, second, first two and first three digit tests. Thus, there are in fact four common methods based on Benford's Law: the first digit test, the second digit test, the first two digits test and the first three digits test.

However, the most often used, although not the most appropriate for all cases, is the first digit Benford's Law test. That is why Benford's Law is also called "First digit law", "First Digit Phenomenon" and "Leading Digit Phenomenon".

Benford's Law of the first digit i.e. probability P of appearance of digit d_1 in number system with base 10 on leftmost position in number is expressed by the following formula [18, page 54]:

$$P(d_1)=\log_{10}(1+1/d_1),\ d_1 \in [1..9] \qquad (1)$$

Formulas for probabilities of appearance of the second, first two and first three digits in number system with base 10 are [18, page 54]:

$$P(d_2) = \sum_{d1=1}^{9} \log_{10}(1+1/d_1 d_2), d_2 \in [0..9] \qquad (2)$$

$$P(d_1d_2)=\log_{10}(1+1/d_1d_2),\ d_1d_2 \in [10..99] \qquad (3)$$

$$P(d_1d_2d_3)=\log_{10}(1+1/d_1d_2d_3), d_1d_2d_3 \in [100..999] \qquad (4)$$

So, Benford's Law states that in certain number sets, for example digit "1" will occur on leftmost position with probability of around 30.1%. If number set conforms to uniform distribution that probability would be much smaller (11.1% i.e., one digit out of 9). It is confirmed that this counter intuitive result can be applied to a wide variety of data sets and that it even holds to any base of numeric system (base invariance). Of course, when changing number bases, actual digit distributions will change. Benford's law states that the leading digit d ($d \in [1..b-1]$) in base b ($b \geq 2$) occurs with probability

$$P(d)=\log_b(d + 1) - \log_b d = \log_b((d + 1)/d) \qquad (5)$$

Probabilities (P) of each digit (d_1) on the most significant position in number are shown in Table 1.

digit (d_1)	probability $P(d_1)$
1	0.30103
2	0.17609
3	0.12494
4	0.09691
5	0.07918
6	0.06695
7	0.05799
8	0.05115
9	0.04576

Table 1. Probabilities of each digit on the first position in number according to Benford's Law (base b=10)

2.2. Prerequisites for Application and Properties of Benford's Law

This law starts from assumption that number set sorted ascending forms geometric series. Intuitive explanation of Benford's Law is pretty clear. If town with population of 10,000 is observed, first digit is 1. Digit 1 will stay on first position of population number until population rise for 100%, which is 20,000 inhabitants. After this, only rise of 50% is needed in order to change first digit from 2 to digit 3. It is clear that town will have digit 1 the most of the time because the most time is needed to change first digit from 1 to 2.

In [16] and [18] prerequisites are set for number series to conform to Benford's Law:

P1. Number series must describe values of same or similar phenomenon, e.g. lake area, heights of mountains, total yearly revenue of companies, total daily turnover on stock exchange etc.

P2. Number series should not have defined minimal and maximal values. If minimal commission on foreign currency exchange in exchange office is 3 kunas, then set of commission values will not fit to Benford's Law, because large number of commission values will have digit 3 as a first digit. Digit 0 is allowed minimum.

P3. Number series should not comprise of so called assigned numbers. These numbers are assigned to various phenomena instead of description, and their important attribute is that there is no sense to perform mathematical operations on these numbers. Examples are citizens' identification numbers, bank account numbers, telephone numbers, numbers on car registration plates etc.

P4. This law does not apply to numbers which creation is influenced by psychological factors, like prices in supermarket or Automatic Teller Machine (ATM) cash withdrawals.

Very important feature of Benford's Law is scale invariance. If certain number set fits Benford's Law, then the set will follow the law independently on measurement unit in which it is expressed ([10] and [15]). Consequently, if all numbers in a set that conforms to Benford's Law are multiplied by a constant, then new set will also conform to the law. For example, if the law is followed by set of total yearly companies' turnover, then the law will be followed independently on currency in which turnovers are expressed. Invariance rule also holds for reciprocal number sets ([10] and [15]). For example, if the law is followed by set of prices in kunas per stock, it will hold for numbers of stocks per kuna.

2.3. Measures of Evaluation

We used chi-square (χ^2) and average relative deviation (ARD) tests in order to evaluate conformance of payment amounts with Benford's Law. With these tests we wanted to show if eventual deviation of frequencies of observed phenomena from Benford's Law frequencies is incidental or not i.e. if frequencies follow Benford's Law. Our goal was to compare practical usefulness and note differences between these conformity tests in auditing environment.

Chi-square test is very often used in statistics in order to make conclusion if empirical data set conforms certain theoretical probability distribution. . In our research it is used for checking if the first digits in our data sets conform to Benford's Law.

Average relative deviation (ARD) is used in [18, page 59] and [16, page 79] for intuitive explanation if certain number set conforms to Benford's Law. In [16] and [18] test is named Mean absolute deviation (MAD), but we think expression average relative deviation (ARD) is more appropriate. This measure does not have strictly defined limit values i.e. range in which it can be stated whether deviation of value sets is significant or not. As a consequence, there are no strict rules by which auditor can state if amounts conform to Benford's Law after applying ARD test. However, in [16] it is stated that ARD (MAD) is the best conformity test in an auditing context. Also, certain guidelines are set for ARD (MAD) cut off levels [16, pages 118-122]. Unlike the chi-square and Z-statistics tests, ARD (MAD) is not affected by the size of observed data set. That means the result is independent on the number of observations itself.

As it is explained in [16] and [18], only positive significant deviations (surpluses) of specific first digits deserve special and additional attention. Negative deviations are only the effect of positive deviations, so they are usually not in the focus of analysis.

3. Experiment

Our hypothesis is that top 500 CEE companies' net income amounts follow Benford's Law first digit distribution. This hypothesis is based on our estimation that this data set fulfils prerequisites P1 (values are of similar phenomenon), P2 (there are no minimal or maximal values), P3 (numbers are not assigned). If companies net income (profit or loss) amounts do not conform to Benford's Law first digit distribution it can be assumed net income category is in certain cases product of psychological influence (so that data set does not confirm P4). In those cases, net income category may be consequence of intentional accounting adjustments which is not direct outcome of operating results. In business environments, it is well-known that net income, either profit or loss, is accounting category which sometimes can be easily manipulated and set according to business goals of owner and/or management.

Our research focused on data of top 500 central and east European companies according to their income in years 2007, 2008 and 2009. The data are extracted from Deloitte Central Europe ranking reports created for each business year according to [25]. For each above mentioned year, 500 companies with largest income in corresponding year are listed. That totals with 1.500 records for three years from 2007 to 2009.

3.1. Data Set

An examination target focus is set on values of top 500 central and east European companies according to their income. Data set contains 1,500 records and spans through 3 years (2007, 2008 and 2009) including 500 net incomes per year**.**

For the purpose of this research and also within Deloitte analyses, central and east Europe geographically includes following countries: Albania, Bosnia and Herzegovina, Croatia, Slovenia, Montenegro, Serbia, Macedonia, Bulgaria, Romania, Moldova, Czech Republic, Slovakia, Ukraine, Latvia, Lithuania, Estonia, Hungary and Poland.

Methodology of top 500 rankings is explained in [26], with following main principles:

"The Central Europe Top 500 rankings are based on company revenues.

As much as possible, we have used consolidated reported revenue for the fiscal year ending in 2007. The information has been sourced from publicly available databases, data from the companies themselves, and estimates based on our research.

In a few instances where revenue for fiscal year 2007 was not available we used the reported 2006 revenue as a proxy for 2007. For Ukrainian companies we used revenue data provided by "Delo."

The revenue for subsidiaries of large groups has been reported as part of the consolidated revenue and shown separately for those subsidiaries.

The rankings are based on revenues reported by a particular legal entity operating in Central and Eastern Europe."

Data on profits or losses (net income or net revenue) are also presented within each company, according to availability of figures. That results in total of 1,194 figures about net income out of 1,500 total revenue data in three observed years. Total of 994 figures correspond to profits, while 200 values designate losses. Profits, losses and absolute net income values are analysed in this paper. No grouping according to industry or country is applied, so all of data are analyzed either on year level or for complete observed period.

3.2. Results

Firstly we conducted Benford's Law first digit test of positive net income values for three year period (years 2007-2009). The results are shown in Table 2.

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	299.22	300	30.103	30.181	0.08	0.26	0.00
2	175.03	165	17.609	16.600	-1.01	-5.73	0.57
3	124.19	134	12.494	13.481	0.99	7.90	0.77
4	96.33	100	9.691	10.060	0.37	3.81	0.14
5	78.71	87	7.918	8.753	0.83	10.54	0.87
6	66.55	68	6.695	6.841	0.15	2.18	0.03
7	57.64	51	5.799	5.131	-0.67	-11.52	0.76
8	50.85	49	5.115	4.930	-0.19	-3.63	0.07
9	45.48	40	4.576	4.024	-0.55	-12.06	0.66
Total:	994.00	994					3.89
Mean:					0.54	6.40	

Table 2: Benford's Law first digit (d) test analysis of positive net income values for three year period

Explanation of symbols used:

d - observed first digit

f_e – expected frequency – number of observations expected according to Benford's Law

f_a – actual frequency - number of actual observations

PBL - Benford's Law probability
P_a – actual probability
AD – absolute deviation - difference between actual and expected probability (=P_a-PBL)
RD - relative deviation - difference between actual and expected probability in percents (=(P_a/PBL-1)*100)
χ^2 chi-square value

As it is already mentioned, Average Relative Deviation (ARD) is average deviation from percentages of deviation (sum of absolute values RD divided by number of frequency categories i.e. 9). As noted in Table 2, ARD value is 6.40.

In Table 2 it can be noticed that for first digit 1 the actual frequency (f_a) is 300, while expected frequency is 299.22 i.e. 299. Deviation is denoted in absolute (AD) and relative (RD) percentage values. The largest individual positive deviation according to RD is for digit 5, which is in surplus for almost 11% when compared to Benford’s Law expected frequency. This surplus may be explained by influence of psychological factors when creating companies' balance sheets specifically profit categories. It is much more sound and ear-catching when profit starts with 5 then with e.g. 4 or even 6. For example, it is much welcomed if profit is 502 than 498. Interestingly, even profit of 589 is much more noticeable by observer than profit of 602 although latter is of a greater value. Digits 7 and 9 are in deficit when compared to expected Benford's Law frequencies. However, in total, these deviations are not significant, as it is shown by chi-square test.

In order to perform chi-square test, a null hypothesis is set and significance level is tested on 5%. According to 8 degrees of freedom and testing on significance level of 5%, in order to confirm our hypothesis, in this case for three year profits, sum of individual χ^2 values should be less than 15.507. Since it is clearly the case (3.89 < 15.507), we should accept assumption that number set of positive net income amounts conform to Benford's Law according to chi-square test.

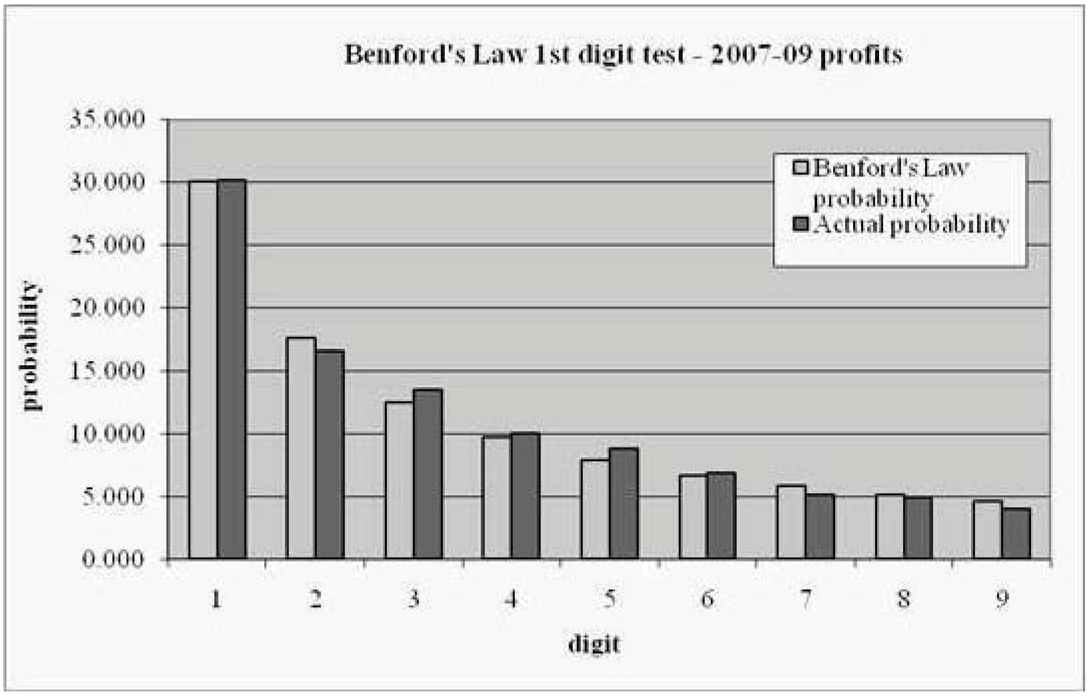

Diagram 1: Benford's Law first digit test analysis of positive net income values for three year period

Also, it is worth to mention it is usual in financial data auditing context that digits in surplus according to conformance tests deserve additional attention. Auditors should carefully and furtherly investigate what is in a background of surpluses. Digits in deficiency usually do not deserve to much additional work because their shortage is only the reflection of before mentioned surpluses.

Secondly, Benford's Law first digit test of negative net income values (losses) for three year period (years 2007-2009) is performed. The results are shown in Table 3.

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	60.21	60	30.103	30.000	-0.10	-0.34	0.00
2	35.22	39	17.609	19.500	1.89	10.74	0.41
3	24.99	17	12.494	8.500	-3.99	-31.97	2.55
4	19.38	22	9.691	11.000	1.31	13.51	0.35
5	15.84	17	7.918	8.500	0.58	7.35	0.08
6	13.39	15	6.695	7.500	0.81	12.02	0.19
7	11.6	8	5.799	4.000	-1.80	-31.02	1.12
8	10.23	7	5.115	3.500	-1.62	-31.57	1.02
9	9.15	15	4.576	7.500	2.92	63.90	3.74
Total:		200					9.47
Mean:					1.67	22.49	

Table 3: First digit (d) analysis of negative net income values for three year period

Similarly as in three year profit analysis, chi-square value is smaller than threshold value (9.47 < 15.507) what is basis for conclusion that losses are also in accordance to Benford's Law first digit test distribution. However, one may notice that losses starting with digit 9 have the greatest deviation (almost 64%) from Benford's Law distribution. Instead of expected 9 there are 15 losses with first digit 9. This may be outcome of psychological influence in accounting process of loss calculation. Therefore, it is well known that companies' financial results may be manipulated in order to present larger or smaller profits and losses. In this case, it may be assumed that deviation between expected (9) and actual frequencies (15) of losses starting with digit 9 is outcome of manipulation.

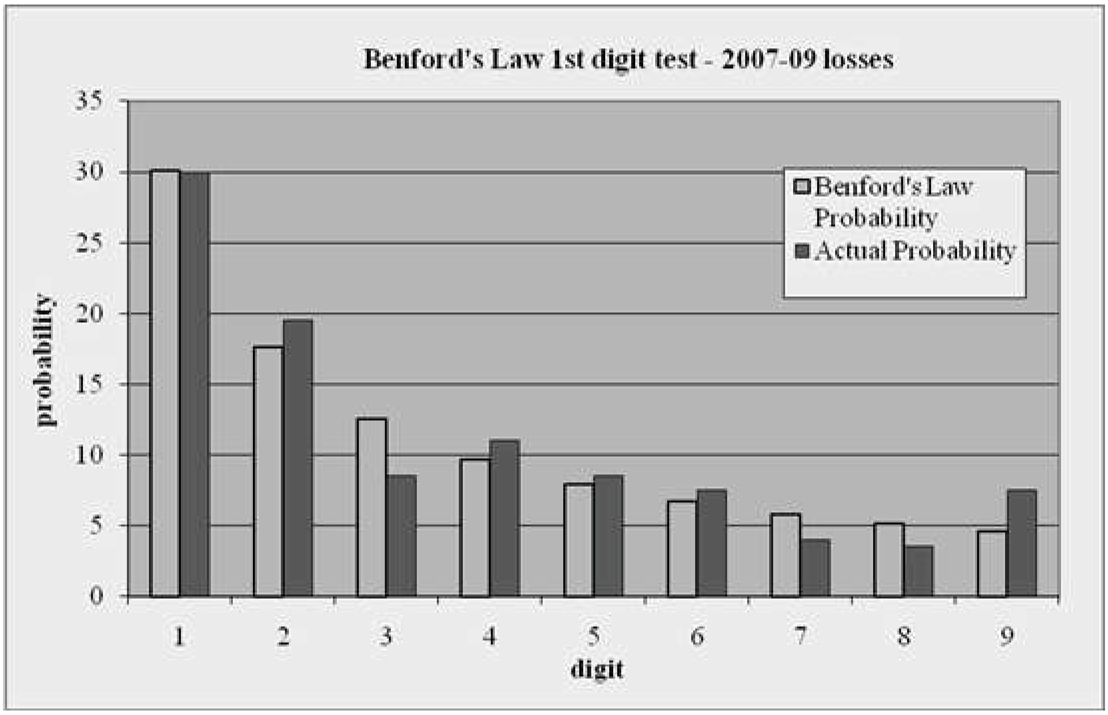

Diagram 2: Benford's Law first digit test analysis of negative net income values for three year period

The aim of the manipulative number correction is quite obvious: it is easier to accept loss of 980 than of 1,030. So, it is possible that when some companies calculated their financial result they tried to lower loss from e.g. 1,030 to 980 in order to have better view on their performance. However, it may be concluded that negative net income values for three year period are in accordance with Benford's Law first digit distributions.

Also, all absolute values for observed three-year period are in put focus. The results are shown in Table 4 and Diagram 3.

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	359.43	360	30.103	30.151	0.05	0.16	0.00
2	210.25	204	17.609	17.085	-0.52	-2.97	0.19
3	149.18	151	12.494	12.647	0.15	1.22	0.02
4	115.71	122	9.691	10.218	0.53	5.44	0.34
5	94.54	104	7.918	8.710	0.79	10.01	0.95
6	79.93	83	6.695	6.951	0.26	3.83	0.12
7	69.24	59	5.799	4.941	-0.86	-14.79	1.51
8	61.08	56	5.115	4.690	-0.42	-8.31	0.42
9	54.63	55	4.576	4.606	0.03	0.66	0.00
Total:		1,194					3.55
Mean:					0.40	5.26	

Table 4: First digit (d) analysis of absolute net income values for three year period

Analysis is performed on all 1,194 net income values in 2007-09 three year period. All data are submitted to absolute values function so first digit distributions are calculated for both profits and losses.

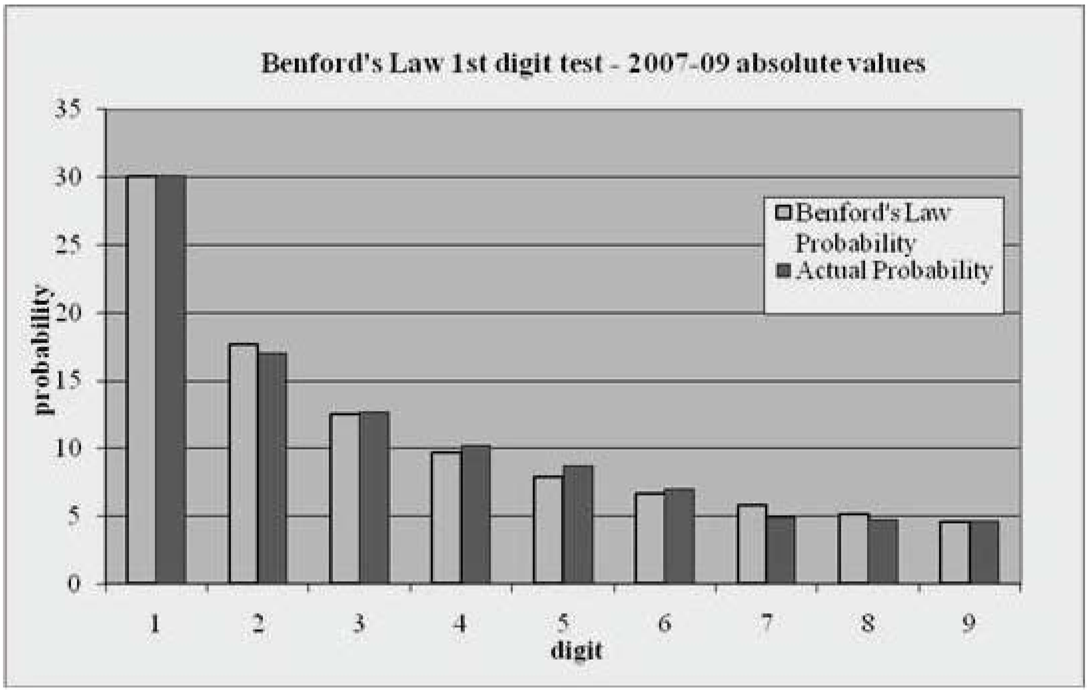

Diagram 3: Benford's Law first digit test analysis of absolute net income values for three year period

Although each test for three year period (profits, losses and absolute values) show no significant deviation from Benford's Law first digit test, test on absolute values for each financial year are performed. This is done because there is possibility that some tests on individual financial year may show deviations from Benford's Law although three year test is compliant.

Results of analysis performed on absolute values for 2007 are shown in Table 5.

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	121.32	123	30.103	30.521	0.42	1.39	0.02
2	70.96	66	17.609	16.377	-1.23	-7.00	0.35
3	50.35	54	12.494	13.400	0.91	7.25	0.26
4	39.05	41	9.691	10.174	0.48	4.98	0.10
5	31.91	39	7.918	9.677	1.76	22.22	1.58
6	26.98	29	6.695	7.196	0.50	7.48	0.15
7	23.37	17	5.799	4.218	-1.58	-27.26	1.74
8	20.61	20	5.115	4.963	-0.15	-2.98	0.02
9	18.44	14	4.576	3.474	-1.10	-24.08	1.07
Total:		403					5.28
Mean:					0.90	11.63	

Table 5: First digit (d) analysis of absolute net income values for 2007

Visual representation is shown on Diagram 4:

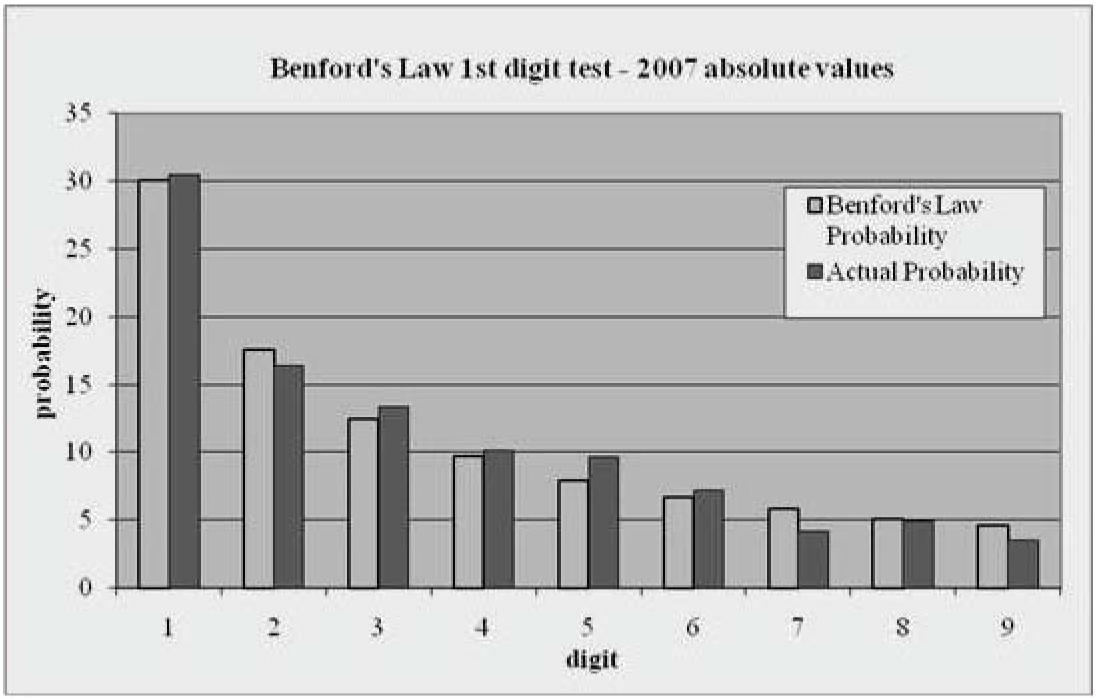

Diagram 4: Benford's Law first digit test analysis of absolute net income values for 2007

It is clear that deviations are not significant (chi-square 5.28<15.507), so it may be concluded that absolute net income values in 2007 are compliant to Benford's Law first digit test. Inspection of Diagram 4 shows some deviation for digits 5 (which are in surplus), 7 and 9 (which are both in shortage) although chi-square for complete set shows no significant deviations.

Data for 2008 are also checked against Benford's Law first digit test and results are presented in Table 6.

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	117.40	117	30.103	30.000	-0.10	-0.34	0.00
2	68.68	66	17.609	16.923	2.68	-3.90	0.10
3	48.73	47	12.494	12.051	1.73	-3.54	0.06
4	37.79	41	9.691	10.513	-3.21	8.48	0.27
5	30.88	37	7.918	9.487	-6.12	19.82	1.21
6	26.11	28	6.695	7.179	-1.89	7.24	0.14
7	22.62	20	5.799	5.128	2.62	-11.57	0.30
8	19.95	18	5.115	4.615	1.95	-9.77	0.19
9	17.85	16	4.576	4.103	1.85	-10.35	0.19
Total:		390					2.48
Mean:					2.46	8.33	

Table 6: First digit (d) analysis of absolute net income values for 2008

Although some deviations exist, especially for digits 5 (surplus), 7 and 9 (both in shortage) as it is case with data for 2007, they are not significant (chi-square 2.48<15.507).

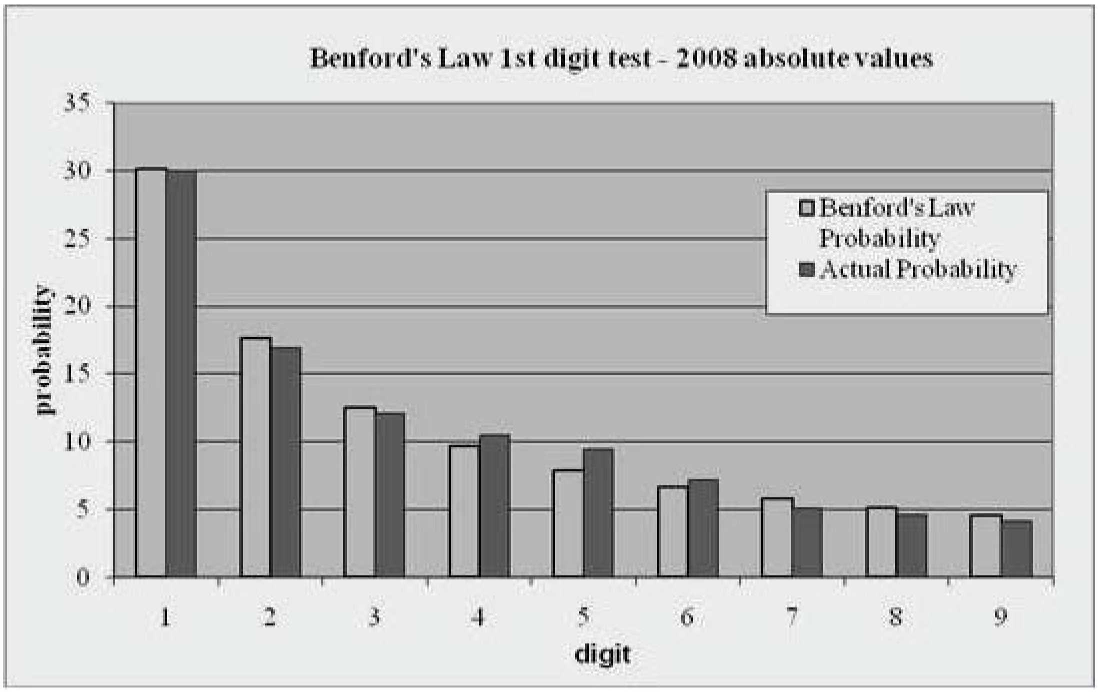

Diagram 5: Benford's Law first digit test analysis of absolute net income values for 2008

However significant surpluses of first digit 5 in 2007 and 2008 have the effect on complete three year number set and its corresponding surplus of 5 as a first digit (RD of 10.01%).

Corresponding visual representation of 2007 data analysis is shown on Diagram 5.

Finally, results of analysis for absolute net income values for 2009 are shown in Table 7 and Diagram 6.

d	f_e	f_a	PBL	Pa	AD	RD	χ^2
1	120.71	120	30.103	29.925	-0.18	-0.59	0.00
2	70.61	72	17.609	17.955	0.35	1.97	0.03
3	50.10	50	12.494	12.469	-0.03	-0.20	0.00
4	38.86	40	9.691	9.975	0.28	2.93	0.03
5	31.75	28	7.918	6.983	-0.94	-11.81	0.44
6	26.85	26	6.695	6.484	-0.21	-3.15	0.03
7	23.25	22	5.799	5.486	-0.31	-5.39	0.07
8	20.51	18	5.115	4.489	-0.63	-12.24	0.31
9	18.35	25	4.576	6.234	1.66	36.24	2.41
Total:		401					3.32
Mean:					0.51	8.28	

Table 7: First digit (d) analysis of absolute net income values for 2009

It is obvious that only positive significant deviation is for digit 9 (RD of 36.24%) which is not the case for data belonging to 2007 and 2008. Deviation is outcome of difference between expected (18.35) and actual (25) frequencies which produces surplus of 7 net income values starting with digit 9.

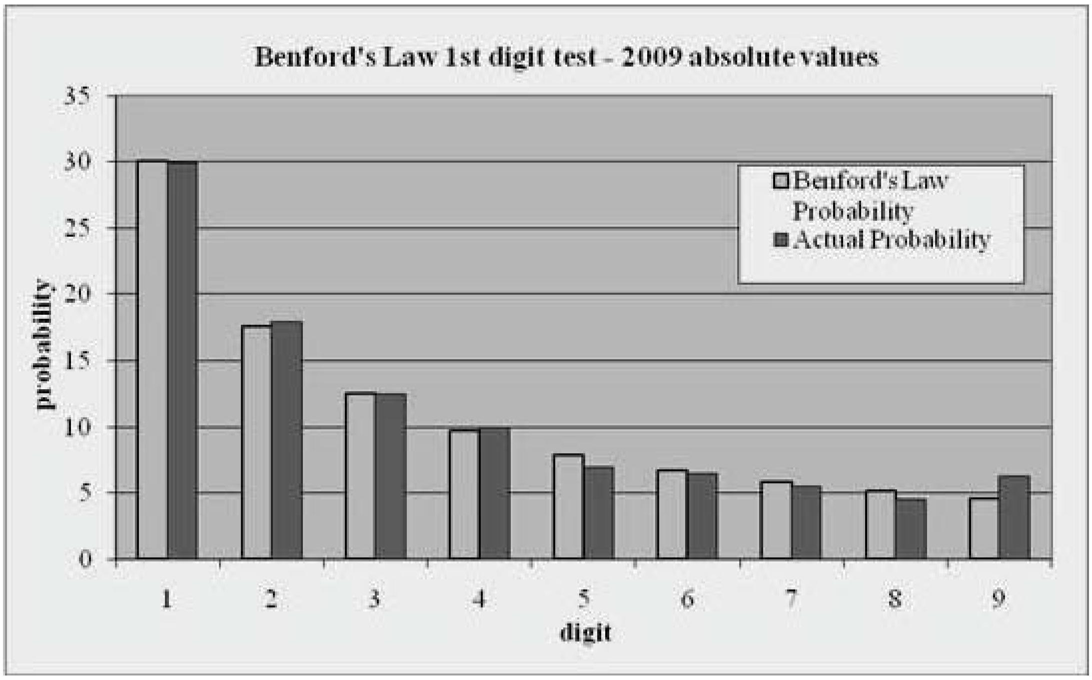

Diagram 6: Benford's Law first digit test analysis of absolute net income values for 2009

This demands additional research, so we analysed positive (profit) and negative (loss) net income values for 2009. Analysis of positive net income values for 2009 show no discrepancies for first digit 9. Expected frequency is 14.37 while actual frequency is 15. Resulting chi-square is only 1.17.

Since there is no deviation for digit 9 within profit values it is reasonable to check losses i.e. negative net income values (Table 8).

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	26.19	25	30.103	28.736	-1.37	-4.54	0.05
2	15.32	16	17.609	18.391	0.78	4.44	0.03
3	10.87	9	12.494	10.345	-2.15	-17.20	0.32
4	8.43	9	9.691	10.345	0.65	6.75	0.04
5	6.89	7	7.918	8.046	0.13	1.62	0.00
6	5.82	5	6.695	5.747	-0.95	-14.16	0.12
7	5.05	2	5.799	2.299	-3.50	-60.36	1.84
8	4.45	4	5.115	4.598	-0.52	-10.11	0.05
9	3.98	10	4.576	11.494	6.92	151.19	9.11
Total:		87					11.55
Mean:					1.88	30.04	

Table 8: First digit (d) analysis of negative (losses) net income values for 2009

Although chi-square value for negative net income values in 2009 is still under the threshold value (11.55<15.507) there is deviation on first digit 9. Expected frequency is 4 while actual frequency is 10 which is surplus of 151%. Since actual frequency (15) for first digit 9 in set of positive net incomes (profits) does not differ from expected frequency (14.37), it is clearly noticeable that difference in set of absolute net income values is caused by deviation in negative net income values. This discrepancy can be explained by the fact that profit or loss is sometimes true accounting category. That means that skilled and sophisticated accounting and financial expert can produce balance sheet according to wishes of its management or owner. Likewise, it is possible that in some cases experts slightly changed values of losses in order to look slightly better than they are in reality. That could have been done by modification of losses starting with digit 1 to digit 9.

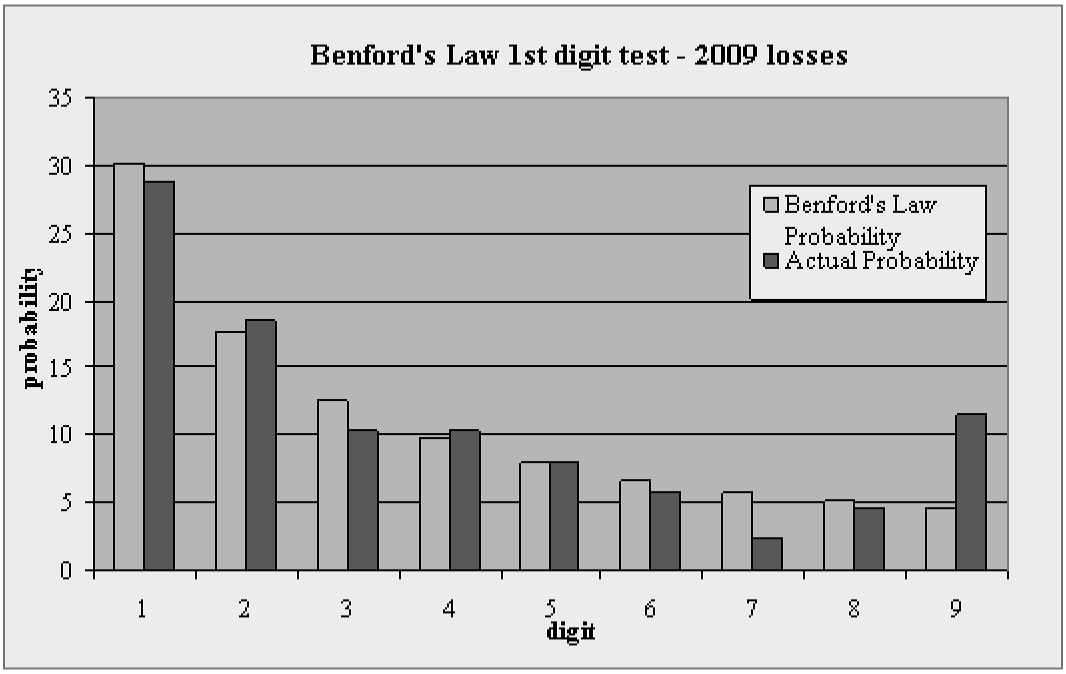

Diagram 7: Benford's Law first digit test analysis of negative net income values for 2009

As it is already explained for some deviations on first digit 5, such modifications can have certain influence on psychology of observer. For some, it is considerably easier to accept the loss of 98 than the loss of 101 although the difference is only about 3%.

Deviation of 2009 losses set on first digit 9 is visible on Diagram 7.

Discrepancy between Benford's Law first digit test and actual probability for first digit 9 in set of 2009 negative net income values is obvious.

4. Conclusion

Basic objective of this paper is to examine if Benford's Law first digit test applies to net income of top 500 central and east European companies in period of three financial years (2007-09). Hypothesis set in the paper is:

"Top 500 CEE companies' net income amounts follow Benford's Law first digit distribution".

For purpose of this work net income values are collected from Deloitte's research "Deloitte Central Europe Top 500" as explained in [25] and [26]. Data includes 1,500 rows on companies' income with 1,194 rows on companies' net income.

Results show that net positive, negative and absolute income values when analyzed for complete observed period conform to Benford's Law first digit test. We conducted additional examinations on profits, losses and absolute net income values.

We spotted some discrepancies for some first digits, notably 5 and 9, and we explained possible reasons for such phenomenon. We believe deviations exist because accounting experts may influence final financial results which may result in first digit distributions.

Regardless of that and according to chi-square test all data subsets we examined are in conformance to Benford's Law first digit test. As a final conclusion, hypothesis must be accepted as true. That means Benford's Law can be used as a method for finding out eventual deviations and manipulations in similar data sets. Nevertheless, if certain company's income data set does not conform to Benford's Law, one may suspect data is manipulated and does not represent true and/or valid business facts.

References

[1] Aggarwal, R. *Persistent Puzzles in International Finance and Economics*. The Economic and Social Review, pages 241-250, Vol. 35, No. 3, 2004.

[2] Albrecht, W. S; Albrecht C. C. *Root out Financial Deception: Detect and Eliminate Fraud or Suffer the Consequences*. Journal of Accountancy, pages 30-34, Vol. 193. Issue 4, 2002.

[3] Brooks, D. *War, Politics & Customer Loyalty: forecasting Using Benford's Law*. Frequencies – The Journal of Size Law Applications, pages 1-15, Ekaros Analytical Inc., Vol. 1. No 1, 2001.

[4] Coderre, D. G. *CAATs and Other BEASTs for Auditors*, Global Audit Publications, Vancouver, Canada, 2001.

[5] Coderre, D. G. *Fraud Detection – A Revealing Look at Fraud*, Ekaros Analytical Inc., Vancouver, Canada, 2003.

[6] Collins, G. W. *Fundamental Numerical Methods and Data Analysis*, Case Western Reserve University, Cleveland, 2003.

[7] DeCeuster, M. J. K; Dhaene, G; Schatteman, T. *On the hypothesis of psychological barriers in stock markets and Benford's law*, Journal of Empirical Finance 5(3), pages 263-279, 1998.

[8] Doucouliagos, C. *Number preference in Australian stock prices*, School Working Papers – Series 2003, Deakin University, 2003.

[9] Durtschi, C; Hillison, W; Pacini, C. *The Effective Use of Benford's Law to Assist in Detecting Fraud in Accounting Data*, Journal of Forensic Accounting 1524-5586, pages 17-34, Vol. V, 2004.

[10] Hill, T. P. *The First Digital Phenomenon*, American Scientist, pages 1-6, Vol 86. No 4, 1998.

[11] Koedijk, K. G; Stork, P. A. *Should we care? Psychological barriers in stock markets*, Economic Letters 44, pages 427-432, 1994.

[12] Krakar, Z; Žgela, M. *Evaluation of Benford's Law application in stock prices and stock turnover*, Society and Technology 2008, 26th to 28th June 2008, Zadar, Croatia.

[13] Lanza, R. B. *Proactively Detecting Occupational Fraud Using Computer Audit Reports*, The IIA Research Foundation, Altamonte Springs, USA, 2003.

[14] Ley, E. *On the peculiar distribution of the U.S. Stock Indices Digits*, Resources for the Future, Washington DC, USA, 1994.

[15] Matthews, R. *The Power of One*, New Scientist Vol. 163 Issue 2194, 1999., pages 26-30.

[16] Nigrini, M. J. *Digital Analysis Using Benford's Law*, Global Audit Publications, Vancouver, Canada, 2000.

[17] Nigrini, M. J. *I've Got Your Number*, Journal of Accountancy, pages 79-84, Vol. 187. Issue 5, 1999.

[18] Nigrini, M. J.; Mittermaier, L. J. *The Use of Benford's Law as an Aid in Analytical Procedures*, Auditing - A Journal of Practice & Theory, pages 52-67, Vol. 16, No 2, 1997.

[19] Opaska, W. *Survey of Accounts Payable Disbursements*, available at http://www.fcc.gov/Bureaus/Inspector_General/Reports/surveymemo.txt, Accessed: March, 27th 2008.

[20] Paukowits, F. *Mainstreaming CAATs - promoting wide usage of computer assisted audit tools*, Internal Auditor, downloaded from www.findarticles.com: August, 13th 2008.

[21] Shinder, D. L. *Scene of the Cybercrime: Computer Forensics Handbook*, Syngress Publishing, Inc., Rockland, 2002.

[22] Silverstone, H; Davia, H. R. *Fraud 101 – Techniques and Strategies for Detection*, John Wiley & Sons, Hoboken, 2005.

[23] Turnbull, C. S. *Fraud Investigation Using IDEA*, Ekaros Analytical Inc., Vancouver, Canada, 2003.

[24] Wells, J. T. *Corporate Fraud Handbook – Prevention and Detection*, John Wiley & Sons, Hoboken, 2007.

[25] *...Deloitte Central Europe Top 500*, Deloitte, http://www.deloitte.com/cetop500, accessed on February, 25th 2011.

[26] *...Central Europe Top 500 Methodology*, Deloitte, http://www.deloitte.com/view/en_LB/lb/insights-ideas/deloitte-research/strategy-cross-industry-research/article/5ebfa0be83ffd110VgnVCM100000ba42f00aRCRD.htm, accessed on February, 23rd 2011.