BULGARIAN ACADEMY OF SCIENCES

CYBERNETICS AND INFORMATION TECHNOLOGIES • Volume **11**, No 4

Sofia • 2011

Application of Benford's Law in Analysis of DAX Percentage Changes

Mario Žgela

Croatian National Bank, Trg hrvatskih velikana 3, 10000 Zagreb, Croatia
Email: Mario.Zgela@hnb.hr

Abstract: *The application of Benford's Law is very rarely covered in the field of stock market analysis, especially in percentage change of stock market indices. Deutscher Aktien IndeX (DAX) was chosen as an index of interest. As stated in [1], DAX measures the development of the 30 largest and best-performing companies on the German equities market and represents around 80% of the market capital authorized in Germany. DAX is a very important stock market index of Frankfurt Deutsche Börse which serves as underlying basis for a large number of financial instruments. It is calculated for 30 selected German blue chips stocks. In this paper Benford's Law first digit test is applied on 10-years DAX daily percentage changes in order to check the compliance. Deviations of the 10-years DAX percentage changes set, as well as distortions of certain subsets from Benford's Law distribution are detected. The possibility that the deviations are an outcome of speculations and psychological influence, should not be eliminated.*

Keywords: *Benford's Law, DAX, index percentage changes, stock market.*

I. Introduction

It was the curiosity of the author who wanted to determine if stock market performance has certain deviations from Benford's Law distributions. As one of the primary stock markets in Europe and worldwide, according to the influence of

German economy and status of equity trade in Europe as a whole, and particularly in Central and Eastern Europe, Deutsche Börse came into focus of research. Particularly, Deutscher Aktien IndeX (DAX) was chosen as an index of interest. As it is stated in [1], DAX measures the development of the 30 largest and best-performing companies on the German equities market and represents around 80% of the market capital authorized in Germany. DAX as underlying basis for a large number of financial instruments and the index calculated for 30 German blue chips stocks will be analysed and checked against Benford's Law distribution.

Previously, in [5], it was shown that the closing stock prices on Zagreb Stock Exchange do not follow Benford's Law distribution according to its first digit test. Also, some other works (e.g., [1]) proved that the stock indices FTSE 100, DJIA and Nikkei are in accordance with Benford's Law.

In this paper it is examined if the daily closing percentage changes of DAX values are compliant to Benford's Law. If the examination proves non-conformance, it may be assumed that there exists psychological and/or speculative influence in the creation of DAX percentage changes, which results in a deviation of the observed number set from Benford's Law.

II. Description of the problem

Further to the importance of Deutsche Börse and curiosity, concerning the changes in index values, DAX was chosen as an index of interest. As stated in [1], DAX measures the development of the 30 largest and best-performing companies on the German equities market and represents around 80% of the market capital authorized in Germany. DAX, as one of the most important equity indices in the world, serves as an underlying basis for more than 40,000 financial products and is the third largest underlying index for derivatives.

According to Bloomberg [12], DAX or the German Stock Index is a total return index of 30 selected German blue chip stocks traded on the Frankfurt Stock Exchange (FSE). The equities use free float shares in the index calculation. The DAX has a base value of 1000 as of December 31, 1987.

Nowadays, DAX includes the following companies: Adidas AG, Allianz SE-REG, BASF SE, Bayer AG-REG, Bayer Motoren WK, Beiersdorf AG, Commerzbank, Daimler AG, Deutsche Bank-RG, Deutsche Boerse, Deutsche Luft-RG, Deutsche Post-RG, Deutsche Telekom, E.On AG, Fresenius Medica, Fresenius Se & C, Heidelbergcement, Henkel AG-PFD, Infineon Tech, K+S AG, Linde AG, Man SE, Merck KGaA, Metro AG, Muenchener Rue-R, Rwe AG, Sap AG, Siemens AG-REG, Thyssenkrupp AG and Volkswagen AG-PFD.

In this paper, publicly available data on DAX percentage changes from Deutsche Börse are used as input values, and a collected set of values is analyzed by use of Benford's Law first digit test. After the research it was possible to make conclusions about the eventual deviations of DAX percentage changes from Benford's Law.

Doubtlessly, stock markets have very important influence on world economy, and consequently on the European economy. The changes in economy are often described and explained by changes in stock prices and/or stock turnover. Not bearing in mind the considerable dependency of the national and world economies on these parameters, it is especially interesting to answer the question if it is possible to note whether there are some (*i*) regularities in the creation of certain index changes. Some interesting questions arise: are the index changes uniformly distributed, do the index changes follow certain rules that are they applicable for some number sets, whether the index changes are an outcome of the psychology of stock buyers and sellers. In this paper it is examined if DAX percentage changes will show discrepancy with respect to Benford's Law first digit test.

Intuitively, one could assume that for any stock index, its daily percentage changes are uniformly distributed, which means that the probability of each digit on the most important (leading, leftmost, most valuable) position in DAX percentage changes, is equal.

All stock exchanges in the world use the so called HLOC (high-low-open-close) prices in order to show the range of daily stock prices changes. The greatest achieved price for a certain stock is marked as *high*, while the smallest achieved price for a certain stock during the working day hours is marked as *low*. The open and close stock price mark the first and the last agreed price for a certain stock. Sometimes HLOC values show the prices for a week, a month, a year or a specific period of time.

The analysis in this paper is based on the daily closing (*close*) DAX value (or price). With regard to the available data, it could also be possible to conduct analysis against the opening (*open*), lowest (*low*) and highest (*high*) daily index prices against the volume, i.e., the total number of stocks traded during the day which are included in a DAX index. Also, it could be interesting to perform analysis according to the volatility of each and every stock included in DAX with respect to the closing price on a previous day and the closing price of the observed day (that is, comparison between the closing prices in two consecutive days).

Further, the total daily turnovers of all stocks included in DAX could be analyzed. The author believes that conclusions on this subject would be very interesting.

The data source for a historical DAX was the web site **www.deutsche-boerse.com**. The selected data included DAX values for a period from February 1, 2001 up to February 1, 2011. The total number of records observed in this period is 2,521[1].

Preliminary investigation proved that the absolute DAX values are not in accordance with the first digit Benford's Law distribution. The primary reason lies in the fact that the observed time span of ten years was too short, so the sample was not complete. Basically, in that period of time, DAX had limits between 2,202.96 (on March 12, 2003) and 8,105.69 (on July 16, 2007). It means that DAX phenomenon had minimum and maximum values during the observed period. And as it is shown in [5, 8, 9] the number sets with defined minimum and maximum

[1] Although total number of observed records is 2521, total number of records in period of interest is 2536. The difference between observed (2521) and records in period of interest (2536) is outcome of absolute index percentage changes less than 0.01. There are 15 index changes having such values.

values will usually not comply with Benford's Law distribution. Even more important, the cut-off values are such that they do not span values having a leading digit "1" and "9". That outcome with gaps for the first digits "1" and "9", consequently distorting frequencies of the set when compared to Benford's Law distribution.

However, with DAX percentage changes the situation is not so intuitively clear. The aim of this paper is to investigate how DAX percentage changes relate to the first digit Benford's Law test. The answer to that question is particularly important, since if DAX percentages changes are not in accordance with Benford's Law, one can assume psychological and/or speculative factors are responsible for the deviation. One should have no doubt that the differentiation between psychological and speculative factors can be very complex and even hidden from the external researcher. The aim of this paper is to check the conformance of DAX percentage changes with Benford's Law, to analyze eventual distortions and to give possible explanations of the eventual distortions. Finally, the investigation if the values of such important stock market index under non-transparent and non-market influence, like speculation of some individual or an interest group, is beyond the scope of this paper.

III. Research method: Benford's Law

Benford's Law defines the 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 a number for almost one third of the time, and the larger digits appear at the leading number position with lower and lower frequencies. E.g., the digit "8" appears as a first digit in slightly more than 5% of values, while the 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 the number and to n first digits[2]. However, the most often used are the first, the second, the 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 the number system with base 10 on the leftmost position in the number is expressed by the following formula [8, p. 54]:

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

[2] Applicable when n is less than or equal to the total number of digits.

The formulas for probabilities of appearance of the second, first two and first three digits in a number system with base 10 are [8, p. 54]:

(2) $$P(d_2)=\sum_{d_1=1}^{9}\log_{10}(1+1/d_1d_2),\ d_2\in[0;9],$$

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

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

Each of these methods uses a certain numeric attribute (a number field) of the observed data set as an input and counts the frequencies of certain combination of digits, depending on the specific method. The output is a list of all digits combinations and their respective frequency in the observed number set.

So, Benford's Law opposes the uniform distribution since it states that in certain number sets, for example digit "1" will occur on the leftmost position with a probability of around 30.1% which is much greater than the expected 11.1% (i.e., one digit out of 9), according to the uniform distribution. 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 the numeric system (base invariance). Of course, when changing the number bases, the 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

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

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

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

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

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

In [8, 9] prerequisites are set for the number series to conform to Benford's Law:

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

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

3. Number series should not comprise of the 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.

4. This law does not apply to numbers which creation is influenced by psychological factors, like prices in a supermarket or ATM cash withdrawals.

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

Historical background. The American astronomer Simon Newcomb was the first who found out that the numbers begin more frequently with smaller digits than with greater digits. Newcomb noticed that the 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 the digit 1 than with any other digit, and in addition, that the probability of each following digit (up to 9), at the most significant position in a number progressively decreases.

Frank Benford gathered more than 20,000 observations from different sources (geographical area, population, river areas, physical constants, etc.). He analyzed the frequencies of the first digits for each number set. After he summarized all individual analyses he concluded that the probability of the first digit being 1 is 0.30103, which equals $\log_{10}2$, the probability of the 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 the application of Benford's Law in information systems auditing, i.e., data analysis for auditing purposes.

IV. Status of the problem in existing literature

Some authors intuitively claim that the frequencies of stock prices in a certain period conform to Benford's Law, while others claim stock prices can not fit Benford's Law. In [4, p. 1], it is stated that certain digits in stock-market prices should occur more often than others. Unfortunately, the author does not give detailed explanation of this assertion or prove its truth. The reason for inherently non conformance of stock prices with Benford's Law some authors find in the fact that stock prices are formed to a significant extent on the basis of market psychology, as stated in [8], or under the influence of financially powerful groups, when some other rules are valid.

On the basis of this, [5] analyzed if the number set of closing daily stock prices conform to Benford's Law. The assumption was not confirmed in the case of Zagreb Stock Exchange closing prices. The basic objective of this paper was to examine if Benford's Law applies to changes of daily closing stock prices and daily stock turnovers. Consequently, these parameters were observed on Zagreb stock exchange in the period from January 1st 1998 to February 26th 2008. The observation included 82,134 data rows. The results show that the closing daily stock prices in the observed ten-year period do not conform to Benford's Law. The same conclusion is valid for more detailed observation of these data, performed on two-year periods. The same is valid for stratification of stock prices (less than 100 kunas, 100-1000 kunas, over 1000 kunas). However, the total daily stock turnovers on Zagreb Stock Exchange, observed in a ten-year period, completely fit to Benford's Law. The same conclusion is valid for stratification of this data on five-year and two-year periods. The results confirm the assumptions of some other authors who claim that Benford's Law does not hold for number sets, on which psychology has important influence. It is confirmed in the example of closing daily stock prices. However, the general expectation that number sets on which some creation psychological factors does not have direct influence, fit to Benford's Law, is confirmed by the investigation of daily stock turnovers analysis. Furthermore, it is noted that certain stock prices appear far seldom than it could be expected according to Benford's, and especially according to uniform distribution. It can be concluded that certain psychological barriers exist, i.e., (un)willingness of investors to buy stocks on certain prices.

Although the author tried to find as much as possible papers on the subject relating to DAX percentage changes and relations to Benford's Law in contemporary scientific research, he did not manage to find any dealing with that subject. There are only two papers which analyze certain DAX data and their relation to Benford's Law. One analyzes the annual financial statements and their relation to Benford's Law [7], while another [2] investigates four European stock indices and prices of eight German stocks. However, neither of these papers analyzed DAX percentage changes and their correlation to Benford's Law distribution.

Also, it may be noticed that the former research of Benford's Law is mainly focused on the analysis of stock market indices or their conformance to Benford's

Law according to the first, first two and first three digits. On the basis of the investigations of certain digit frequencies on the first, first two and first three positions in daily market indices, some authors make conclusion about the existence of psychological barriers in stock markets. In [6] it is shown that the number sets of one-day returns on the Dow-Jones Industrial Average Index (DJIA) and the Standard and Poor's Index (S&P) agree with Benford's Law.

In [7] it is explained that Benford's Law is being recently discussed as an audit instrument in order to gain insights into possible conscious and inadvertent errors in data records. The paper analyzes the closing accounts of 1373 annual financial statements of companies listed on the German DAX and compares these with the Benford's distribution hypothesis. It can be seen in the process that the data pertaining to financial accounts, which have not been audited by the so-called "Big 4" audit firms, but by smaller external auditors, deviate to a significant extent from the regularity of Benford's Law, which appears to make use of Benford's Law suitable as a benchmark to assess the quality of the audit.

In [2] four European stock indices and the prices of eight major German stocks are examined for indications of psychological barriers. The frequency (expected) returns, intraday volatility and trading volume of these assets are studied contingent on whether the prices lie within a certain range around round numbers. The results indicate that psychological barriers do not exist on a consistent basis. The authors assume that some barriers have disappeared after these anomalies have been published. However, the authors did not investigate DAX percentage changes and their correlation with Benford's Law distributions.

Research carried out in [3] shows that the stock prices on Australian stock market do not follow uniform distribution. The stock indices are also analyzed and it is stated that for NASDAQ the psychological barrier was positioned on the value 5000 in March of 2000. Dow Jones Industrial Average Index had an important psychological limit on the value 10,000, as Nikkei on the value 25,000.

In [1] it is explained that the research of eventual psychological barriers in the change of stock indices should not be based on comparison with uniform distribution, but on comparison with Benford's Law distribution. Starting from changed hypothesis, the authors have shown that there are no reasonable arguments on the existence of psychological barriers for stock indices. Particularly, the authors assume that the stock indices must fit to Benford's Law, with deviation from the uniform distribution as a consequence. In order to prove if the psychological barriers in stock indices values exist, it is necessary to check the existence of the relevant deviation of indices number set from Benford's Law distribution. The authors studied the stock indices FTSE 100, DJIA and Nikkei and concluded that there are no reasonable proofs for the existence of psychological levels of certain indices if Benford's Law is defined as a basis of comparison.

According to the previously mentioned researches, it can be concluded that the values of stock indices are not behaving in conformance with the uniform distribution, but are close to Benford's Law distribution. If significant deviations against uniform distribution of stock indices exist, i.e., the indices are following Benford's Law, with regard to invariance of number sets for which Benford's Law holds, we may assume that the percentage index changes are diverging from the uniform distribution and conform to Benford's Law. This assumption is also proved in [1] for FTSE 100, DJIA and Nikkei. Since no author proved such assumption for DAX percentage changes, this paper's objective is to check it. If the analysis shows non-conformance of DAX percentage changes to Benford's Law, one may conclude that DAX index changes are also under psychological influence and/or result of speculative activities. Such conclusion would surely be relevant for everyone who invests in stocks on FSE, since it could be assumed that even such economically important and widely recognized stock index is not a result of transparent calculation, but also speculations and/or psychology of the market.

V. Analysis and results

According to the problem explanations given in the previous chapters, we set the following hypotheses:

1. The number set of DAX percentage changes for a 10-years period (2001-2011) conform to the first digit Benford's Law distribution.

2. Certain number subsets of DAX percentage changes for a 10-years period (2001-2011) may not be in conformance with the first digit Benford's Law distribution.

As stated in the introduction, the observed data includes all DAX percentage changes from February 1, 2001 up to February 1, 2011. Thus, the research is focused on 2521 DAX values.

DAX percentage change (DAXPC) is calculated as a percentage of change between two consecutive DAX Indices (DAXI), as it is stated in (6):

$$DAXPC = (DAXI_{date}/DAXI_{(date-1)} - 1) \times 100. \tag{6}$$

The first hypothesis should be checked on complete DAX percentage changes set during the selected 10-years time span.

In order to check the second hypothesis, six data subsets are set. Each subset spans all possible nine first digits, otherwise it will not make sense because subsets with cut-off value will not be in accordance with Benford's Law first digit test by default. Also, we separated subsets for negative and positive values, since we assume that increase in the index value (positive index change) reflects very different characteristics than the decrease (negative index change). So, the subsets are defined as follows:

$$[0.01; 0.1), \tag{7}$$

$$[0.1; 1), \tag{8}$$

$$[1; 10), \tag{9}$$

$$(-0.1; -0.01], \tag{10}$$

$$(-1; -0.1], \tag{11}$$

$$[-10; -1). \tag{12}$$

There are 1335 positive and 1186 negative DAX changes in the observed interval.

As it is noted in the hypotheses set, in this paper only the first digit test of Benford's Law (1) will be performed on the data.

We used Chi-square (χ^2) test in order to evaluate the conformance of DAX percentage changes with Benford's Law first digit distribution. That test will indicate if eventual deviation of frequencies of the observed phenomenon from Benford's Law frequencies is incidental or not. We set two null hypotheses and a tested significance level on 5%.

Firstly we conducted analysis of absolute DAX percentage changes in a complete 10-years period. The results are shown in Table 2.

Table 2. Absolute DAX percentage change distributions and Benford's Law

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	758.90	895	30.103	35.502	5.40	17.93	24.41
2	443.93	453	17.609	17.969	0.36	2.04	0.19
3	314.97	272	12.494	10.789	–1.70	–13.64	5.86
4	244.31	199	9.691	7.894	–1.80	–18.55	8.40
5	199.62	185	7.918	7.338	–0.58	–7.32	1.07
6	168.77	132	6.695	5.236	–1.46	–21.79	8.01
7	146.20	146	5.799	5.791	–0.01	–0.13	0.00
8	128.96	122	5.115	4.839	–0.28	–5.39	0.38
9	115.35	117	4.576	4.641	0.07	1.42	0.02
	Σ	2521		Σ/9	1.29	9.80	48.34

Explanation of the symbols used:

d – the first digit;

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

f_a – actual frequency – number of actual observations;

PBL – Probability of a digit according to Benford's Law;

P_a – actual probability of a digit in absolute DAX percentage change number set;

AD – Absolute Deviation, i.e., the difference between the actual and expected frequency ($f_a - f_e$);

RD – Relative Deviation, i.e., percentage of the deviation of actual from expected frequency;

χ^2 – Chi square.

Sum of AD/9 – average absolute deviation (sum of absolute values AD divided by the number of frequency categories, i.e., 9);

Sum of RD/9 – average relative deviation from the percentages of deviation (sum of absolute values RD divided by the number of frequency categories, i.e., 9);

Fig. 1 shows the actual and Benford's Law (expected) probabilities of the absolute values of DAX percentage changes from Table 1.

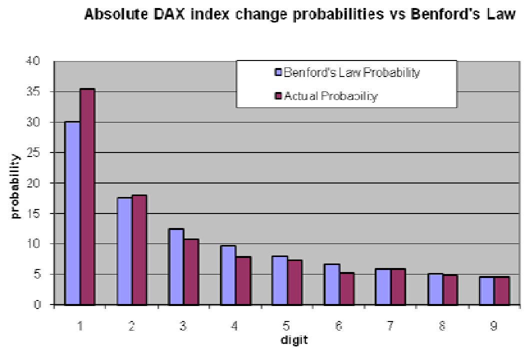

Fig. 1. Absolute DAX percentage change probabilities vs. Benford's Law

The average relative deviation (the sum of absolute relative deviations, divided by 9) is used in [8] for intuitive explanation if a certain number set conforms to Benford's Law. This measure does not have defined limit values, i.e., the range in which it can be stated whether the deviation of value sets is significant or not. Furthermore, by means of the average relative deviation one can not clearly judge if the set of DAX percentage changes conforms to Benford's Law.

According to 8 degrees of freedom and testing on significance level of 5%, in order to confirm the first null hypothesis, the sum value of χ^2 should be less than 15,507. Since it is not the case, i.e., the sum of χ^2 is 48.34, we have to reject the assumption that the number set of DAX percentage changes for 10-years period conforms to Benford's Law. Consequently, it is proved that the first hypothesis is false.

However, it can be noted that the greatest deviation relates to digit 1 (24.41). This deviation deserves more attention and additional explanation. When analyzing significant deviations (digits 1, 3, 4 and 6) one can notice the following:

1. The deviation for digit 1 is the only significant deviation caused by surplus of actual frequency. There are 136 DAX percentage changes with a leftmost digit 1 more than it is expected according to Benford's Law first digit test distributions.

2. All other significant deviations (digits 3, 4 and 6) are an outcome of actual frequencies shortage in comparison to Benford's Law first digit test distributions. The sum of these significantly deviated frequencies is 123 (for digit 3 the shortage is 45, for digit 4 it is 43 and for digit 6 it is 35).

As it is noted in [8, 9], and [5, Application of BL in Payment Systems Auditing, p. 5], only significant surpluses should be furtherly analyzed. In the auditing and business context, usually only digits that are in surplus according to conformance tests deserve additional attention. The auditors should carefully and furtherly investigate what is in the background of surpluses. The digits in deficiency usually do not deserve too much additional work because their shortage is only the reflection of the before mentioned surpluses. Having this in mind, we conclude that the only deviation deserving further analysis is the distribution of the first digit 1.

According to the second hypothesis, we also analyzed the values of DAX percentage changes within three pairs of separate sets. Each of these sets is thoroughly analyzed according to positive and negative DAX percentage change values, as noted in (7) through (12). That means we covered 6 number sets.

It is clearly understandable that significant surpluses of the first digit 1 must be detected in some (or even all) subsets defined in (7) through (12), because the surplus of digit 1 is discovered when a complete number set was analyzed.

On the contrary, the analysis of positive DAX percentage changes within set (7) shows significant lack of digits 1 at first position. The actual frequency is 10, which significantly deviates from 27, which is the expected frequency. It means that the increase in DAX percentage change values between [0.01; 0.02) is a very uncommon phenomenon, much rare than expected according to Benford's Law first digit test distribution. It must also be stressed that this subset contains only 3.6% (90 out of total of 2521) instances of DAX percentage changes in 10 observed years. The value of χ^2 is 31.98 which again means that subset (7) of DAX percentage changes does not follow Benford's Law first digit distribution.

The analysis of positive DAX percentage changes within the set (8) proves significant lack of digits 1 at first position (Table 3). The actual frequency is 111, which significantly and negatively deviates from the expected frequency of 214. On the contrary, the frequency of a first digit 9 is in a significant surplus. DAX percentage changes falling in [0.9; 1) are 81% more often than they should be according to Benford's Law which indicates clustering towards the value 1 from lower values. Similarly it may be concluded for values in [0.8; 0.9) which appears 54% more than they should when compared to Benford's Law. A question that emerges is whether the clustering is accidental or is a result of psychological (speculative) factors. To support the notion of psychological (speculative) influence on DAX percentage changes formation, it may be implied that in the eyes of market participants, an increase of 1% is "far greater" than an increase of let say 0.92% (or 0.86%). However, an increase of 0.92% is usually rounded to 1%. It may be concluded that for stock market optimism and speculative reasons the clustering in subset [0.9; 1), or even [0.8; 1), has much sense. The value of χ^2 is 114.61, which again means that the subset of DAX percentage changes limited to [0.1; 1) does not follow Benford's Law first digit distribution.

Table 3. DAX percentage change distributions within the set [0.1; 1) and Benford's Law

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	214.03	111	30.103	15.612	–14.49	–48.14	49.60
2	125.2	105	17.609	14.768	–2.84	–16.13	3.26
3	88.83	91	12.494	12.799	0.30	2.44	0.05
4	68.9	84	9.691	11.814	2.12	21.91	3.31
5	56.3	86	7.918	12.096	4.18	52.76	15.67
6	47.6	61	6.695	8.579	1.88	28.15	3.77
7	41.23	58	5.799	8.158	2.36	40.67	6.82
8	36.37	56	5.115	7.876	2.76	53.98	10.59
9	32.53	59	4.576	8.298	3.72	81.34	21.54
	Σ	711		Σ/9	3.85	38.39	114.61

The analysis of subset (9) shows significant surplus of digit 1 on the leftmost position in DAX percentage changes (Table 4). The actual frequency is 335, which significantly positively deviates from the expected frequency of 160. Digit 2 also positively deviates, while all other digits are in a significant shortage compared to the expected Benford's Law frequencies. The value of χ^2 reveals that the subset of DAX percentage changes limited to [1; 10) does not follow Benford's Law first digit distribution.

Table 4. DAX percentage change distributions within set [1; 10) and Benford's Law

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	159.85	335	30.103	63.089	32.99	109.58	191.91
2	93.50	125	17.609	23.540	5.93	33.68	10.61
3	66.34	39	12.494	7.345	–5.15	–41.21	11.27
4	51.46	11	9.691	2.072	–7.62	–78.62	31.81
5	42.05	11	7.918	2.072	–5.85	–73.84	22.93
6	35.55	5	6.695	0.942	–5.75	–85.94	26.25
7	30.79	5	5.799	0.942	-4.86	–83.76	21.60
8	27.16	0	5.115	0.000	–5.12	–100.00	27.16
9	24.30	0	4.576	0.000	–4.58	–100.00	24.30
	∑	531		∑/9	8.65	78.51	367.85

It is strange that 25% of all DAX increasing (positive) values fall between 1 and 2%. Also, DAX percentage changes between 1% and 2% account for over 13% of the indices during the observed 10-years time span. Moreover, within subset (9), the values with leftmost digit 1 account for 63%, which is more than twice than it could be expected according to Benford's Law distribution. That means that DAX percentage change clustering between 1% and 2% is extremely significant and importantly influences the surplus of digit 1 in the absolute DAX percentage change within the observed set, indicated in Table 2. Further, the deviations in interval (9) are very close to the deviations of Standard and Poor's Index and Dow-Jones Industrial Average Index during the periods 1926-1993 and 1900-1993 respectively, as pointed out in [6, p. 7]. It seems that clustering of the stock market index changes around these values is something which appears regularly and is not a result of a chance alone. Further research should give the answer whether the reasons lie in speculative character of stock markets based on (false) optimism, constraints of capital growth or something else.

The analysis of the negative DAX percentage changes within set (10) shows lack of digits 1 at the first position. The actual frequency is 10, which significantly deviates from 24, what is the expected frequency. It means that the increase in DAX percentage values between (–0.02; –0.01] is a very uncommon phenomenon, much rare than expected according to Benford's Law first digit test distribution. It also must be stressed that this subset contains only 3.2% (80 out of total of 2521) instances of DAX percentage changes in 10 observed years. The actual frequency of indices starting with digit 1 is 10, which is exactly the same as in (7). Also, cardinality of subset (10) is 80, what is similar to cardinality of subset (7) which is 90. As an outcome, the percentages of the number of indices in these two subsets compared to the total set are very close – 3.2% in (10) vs. 3.6% in (7). The value of χ^2 is 47.34 which again means that subset (7) of DAX percentage changes does not follow Benford's Law first digit distribution.

The analysis of negative DAX percentage changes in set (11) indicates significant lack of digits 1 at the leftmost position. The actual frequency is 95, which significantly and negatively deviates from the expected frequency of 171. On the contrary, the frequency of first digit 9 is in a significant surplus. DAX percentage changes falling in (–1; –0.9] are 61% more often than they should be according to Benford's Law which indicates clustering toward value –1. Similar may be concluded for values in (–0.9; –0.8] and (–0.8; 0.7], which appear 55% and 64% more than they should when compared to Benford's Law. Interestingly, similar deviations are revealed in the corresponding positive subset (8). A question that appears is whether the clustering is accidental or a result of psychological (speculative) factors. To support the notion of psychological (speculative) influence on DAX percentage changes formation, it may be implied that in the eyes of market participants a decrease of 1% is "far less" than a decrease of let say, 1.22%. The value of χ^2 is 74.80, which again means that the subset of DAX percentage changes limited to (–1; –0.1] does not follow Benford's Law first digit distribution.

The analysis of subset (12) shows significant surplus of digit 1 at the leftmost position in DAX percentage changes (Table 5). The actual frequency is 331, which significantly positively deviates from the expected frequency of 162. Digit 2 also positively deviates, while all other digits are in a significant shortage compared to the expected Benford's Law frequencies. The value of χ^2 (324.10) signifies that the subset of DAX percentage changes limited to (–10; –1] does not follow Benford's Law first digit distribution.

Table 5. DAX percentage change distributions within set (–10; –1] and Benford's Law

d	f_e	f_a	PBL	P_a	AD	RD	χ^2
1	161.65	331	30.103	61.64	31.54	104.76	177.42
2	94.56	114	17.609	21.23	3.62	20.56	4.00
3	67.09	42	12.494	7.82	–4.67	–37.40	9.38
4	52.04	27	9.691	5.03	–4.66	–48.12	12.05
5	42.52	14	7.918	2.61	–5.31	–67.07	19.13
6	35.95	4	6.695	0.74	–5.95	–88.87	28.40
7	31.14	4	5.799	0.74	–5.05	–87.16	23.65
8	27.47	1	5.115	0.19	–4.93	–96.36	25.51
9	24.57	0	4.576	0.00	–4.58	–100.00	24.57
	Σ	537		Σ/9	7.81	72.26	324.10

It is peculiar that 28% of all DAX decreasing (negative) values fall between –1 and –2%. Also, DAX changes between –1% and –2% account for over 13% of indices during the observed 10-years time span. Moreover, within subset (12), the values with leftmost digit 1 account for 62%, which is more than twice than it could be expected according to Benford's Law distribution. This means that DAX percentage change clustering between –1% and –2% is extremely significant and importantly influences the surplus of digit 1 in the absolute DAX percentage change within the observed set, indicated in Table 2. It is interesting that DAX changes in subsets (9) and (12) have some similar characteristics. Obviously, there is a relation between subsets which contain the largest positive ([1; 10)) and negative ((–10; –1]) percentage of index changes (Table 6).

Table 6. Similarities between DAX changes contained in subsets [1; 10) and (–10; –1]

Characteristics/subsets	[1; 10)	(–10; –1]
Cardinality	531	537
Percentage in total number of instances (2521)	21	21
Cardinality of subsets [1; 2) and (–2; –1]	335	331
Percentage of subsets [1; 2) and (–2; –1]	63	62
Percentage of subsets [1; 2) and (–2; –1] in all positive (1335) and negative (1186) indices	25	28
Percentage of subsets [1; 2) and (–2; –1] in total number of indices (2521)	13	13
Absolute difference between actual and expected frequency for first digit 1	175	169

It may be concluded that the surplus of leftmost digit 1 in absolute DAX percentage changes (Table 2) is an outcome of surpluses of DAX percentage changes with leftmost digits 1 in subsets [1; 10) and (–10; –1].

As with interval (9), the deviations in interval (12) are similar to the deviations of Standard and Poor's Index and Dow-Jones Industrial Average Index during the periods 1926-1993 and 1900-1993 respectively, as indicated in [6, p. 7]. This is additional similarity between DAX percentage changes in subsets (9) and (12). It seems that the clustering of stock market index changes around these values is something which appears regularly, being a fact not only for DAX percentage changes, meaning it is not a result of a chance only. Additional and more focused research should give the answer whether the reasons lie in speculative character of stock markets based on forced optimism or suppressing pessimism, constraints of companies' capital growth, or something else. However, a clustering different from the expected Benford's Law distribution could be a warning that certain speculation and/or psychological influence is present within DAX percentage changes, which especially holds for percentage changes starting with the digit 1. While it may be stated that pushing a positive percentage change otherwise less than 1, slightly below the value 1 and after two first decimal digits rounding producing digit 1 (e.g., a percentage change 0.996 rounded to 1.000) on the leftmost position can have positive psychological influence on the market individuals and groups. It lies in human psychology that accepts a positive change (rise) of 1% much better than 0.99 which is, mathematically speaking, almost neglecting. Further, such situation and possibility to push the "invention" of positive 1's could be very interesting for market speculators and could result in significant stock market investment growth.

However, mirrored situation with significant surplus of negative numbers less or equal than –1 and greater than –2, so starting with the digit 1, could be explained similarly. For example, the daily DAX decrease of 1.9% in the eyes of investors and market as a whole is much less than a decrease of 2%. Such situation again creates clustering within the number set (–2; –1], producing a surplus of digits 1 at the leftmost position.

Although the above explanations of clustering of DAX percentage changes on the leftmost digit 1 in sets [1; 2) and (–2; –1] may seem logical, the author does not have a proof for those assumptions. What is the reason of such clustering remains unclear and hidden. Further curiosity, availability of additional data and knowledge on specific Deutsche Börse stock market and stocks constituting DAX, could probably give satisfactory answers.

It must be stressed here that DAX percentage changes are often even a more important dimension than the absolute DAX values. It is so because the majority of market players watch the index change more closely than its absolute value. The reason is very clear: the direction of the index trend (negative or positive) is very important, i.e., whether the stock prices of the companies composing DAX are decreasing or increasing. This information is extremely valuable since the investors use it when determining the market stability and the position of its bull-bear line. If one interest group or individual can indirectly influence the values of some (major) stock market indices, like DAX, it/he can have major influence on the stock market status and development, as well as stability of the state or even global economy. This is a clear consequence of the stock indices importance in contemporary state and global economies.

Basically, DAX percentage change may be influenced in two ways. Firstly, it is feasible by influencing certain absolute stock prices, thus setting their percentage change on a desired value. While setting certain stock prices of 30 German companies included in DAX calculation, DAX absolute value could be influenced. Thus, the desired percentage changes of DAX would be indirectly attained. The second possibility is to intervene in DAX calculation algorithm and thus directly change the absolute DAX and percentage change values. Since the calculation algorithm is public and subject to a simple check and public control, this possibility is not feasible.

This paper, as well some other research works like [5, 8-10], shows the immense strength of Benford's Law in data analysis. As a consequence, Benford's Law is often used as a principal method for information systems (data) auditing, and represents important capability of Computer Assisted Auditing Tools (CAAT). One of its principal strength is the fact that the analyst (auditor) does not have to define in advance what type of a fraud, error or omission should be investigated. Only after the specific Benford's Law test is executed and as a consequence, the deviations are noted, the analyst should question the results and compare eventual deviations with previously set business norms, data creation, deletion and modification rules. Also, as it is shown previously, the first digit Benford's Law test can not be a final proof that the data is erroneous, fraudulent or flawed, nor that some irregular activities have taken place. Some more detailed and focused tests, together with analysis of business processes and rules, must be undertaken. However, the first digit Benford's Law test may be a first warning that some additional audit actions should be performed and that integrity and accuracy of some data should be questioned.

VI. Conclusion

In this work two hypotheses are set:

1. The number set of DAX percentage changes for a 10-years period (2001-2011) conform to the first digit Benford's Law distribution.

2. Certain number subsets of DAX percentage changes for the 10-years period (2001-2011) may not be in conformance with the first digit Benford's Law distribution.

As it is indicated by means of χ^2 test, the first hypothesis that the number set of DAX percentage changes for a 10-years period conform to Benford's Law proved to be false.

After analysis of each interval it may be concluded that the second hypothesis is true: A certain number of subsets of DAX percentage changes may not be in conformance with the first digit Benford's Law distribution. More precisely, none of the number subsets defined in (7) through (12) are conformed to Benford's Law first digit test.

However, it can be noted that the greatest deviation relates to digit 1 (χ^2 value of 24.41). Without that deviation, DAX percentage changes would be Benford's Law compliant. The observed number set was further analyzed through subsets defined in (7) through (12) with a focus on instances with digit 1 at the leftmost position. Clustering of DAX percentage changes around values 1 and –1 is discovered in two number subsets: [1; 10) and (–10; –1]. In fact, the excess of cardinality in subsets [1; 2) and (–2; –1] efficiently causes clustering of digit 1 (1 or –1) and results in non-conformity of DAX percentage changes to Benford's Law. It is unclear if the excess of digit 1 in DAX percentage changes is caused by speculation, i.e., speculative character of stock markets, bull market trends based on forced optimism (since DAX is increasing), possible natural constraints of companies' capital growth, or simply speculation. Similarly, it remains unconfirmed if the excess of digit –1 on the leftmost position in a number, i.e., significant excess of cardinality in subset (–2; –1] is an outcome of suppressing pessimism (since DAX is decreasing), speculative character of stock price formation, or something else.

It is interesting to note that the subsets of the corresponding negative and positive DAX percentage changes, i.e., (7) and (10), (8) and (11), (9) and (12), have very similar characteristics concerning frequencies, distributions, compliance to Benford's Law first digit test, digits that significantly deviate, as well as deviation figures.

Whatever the reason for non-compliance of DAX percentage changes to Benford's Law distribution, a fact leading to the conclusion that the speculative forces on the stock market influence even such important market indicator, like DAX, it is clear that Benford's Law is a very powerful method for data analysis and thus information systems auditing. Its very important strength lies in the fact that the auditor does not have to know what type of omissions, errors or fraud based on data manipulation should be investigated. The auditor should only apply Benford's Law test on the data and check their conformance to Benford's Law distribution. What requires most knowledge, is making a conclusion on eventual non-conformance.

References

1. DeCeuster, M. J. K., G. Dhaene, T. Schatteman. On the Hypothesis of Psychological Barriers in Stock Markets and Benford's Law. – Journal of Empirical Finance, Vol. **5**, 1998., No 3, 263-279.
2. Dorfleitner, G., C. Klein. Psychological Barriers in European Stock Markets: Where Are They? – Global Finance Journal, Vol. **19**, 2009, Issue 3, 268-285.
3. Doucouliagos, C. Number Preference in Australian Stock Prices. School Working Papers – Series 2003. Deakin University, 2003.
4. Hill, T. P. The First Digital Phenomenon. – American Scientist, Vol. **86**, 1998, No 4, p. 1.
5. Krakar, Z., M. Žgela. Evaluation of Benford's Law Application in Stock Prices and Stock Turnover. – Informatologija 42, Vol. **3**, 2009, 158-165.
6. Ley, E. On the Peculiar Distribution of the U.S. Stock Indices Digits. Resources for the Future, Washington DC, 1994.
7. Moeller, M. Measuring the Quality of Auditing Services with the Help of Benford's Law? – An Empirical Analysis and Discussion of this Methodical Approach. Zurich, Institute for Accounting and Controlling at the University of Zurich, 2009.
8. Nigrini, M., L. Mittermaier. The Use of Benford's Law as an Aid in Analytical Procedures. Auditing – A Journal of Practice & Theory, Vol. **16**, 1997, No 2, 52-67.
9. Nigrini, M. Digital Analysis Using Benford's Law. Vancouver, Canada, Global Audit Publications, 2000.
10. Žgela, M. Development of Methodology for Computer Assisted Information Systems Auditing. Doctoral Dissertation, Faculty of Organization and Informatics – University of Zagreb, Varaždin, 2010.
11. **http://deutsche-boerse.com/dbag/dispatch/en/isg/gdb_navigation/lc?module=InOverview_Index&wp=DE0008469008&foldertype=_Index**
12. **http://www.bloomberg.com/markets/stocks/movers/dax/**