SOBER PROCEEDINGS

CIBES 2023 / 3rd Current Issues in Business and Economic Studies Conference

CAN BENFORD'S LAW REFLECT MAJOR ECONOMİC CHANGES?

Stanislava Hronova[a] Richard Hindls[b] Lubos Marek[c]

[a] Prof., Prague University of Economics and Business, Faculty of Informatics and Statistics, Prague, Czech Republic, hronova@vse.cz
[b] Prof., Prague University of Economics and Business, Faculty of Informatics and Statistics, Prague, Czech Republic
[c] Prof., Prague University of Economics and Business, Faculty of Informatics and Statistics, Prague, Czech Republic

Abstract

This paper examines a range of additional economic indicators in which Benford's law might apply. Specifically, it concerns the indicator of sales in the Czech Republic, monitored in three different sectors (manufacturing, trade, and hospitality, i.e., accommodation and catering). At the same time, the analysis has also been performed on the data mix created by aggregating data from these sectors. This paper also examines whether the validity of Benford's law in these sectors could have been affected by the COVID-19 pandemic or what the changes in the validity of Benford's law have been from the time before the pandemic (2019), during the pandemic (2020), and then after the partial relaxation of the strict restrictive measures (2021). It turns out that the validity of the law is only partial and varies across both sectors and years. On the other hand, it must also be acknowledged that, visually, the shape of the logarithmic curve governing the occurrence frequencies of the first and second digits from the left generally confirms the corresponding frequency distribution values, even if the goodness-of-fit tests do not prove the full validity of the law.

Keywords: Benford's law, goodness-of-fit test, revenues, COVID-19 pandemic

Cited: Hronova, S., Hindls, R., & Marek, L. (2023). Can Benford's law reflect major economic changes?. *Sustainability, Organization, Business and Economic Research (SOBER), 1*, 30-40. doi: 10.5281/zenodo.8060093

Selection and peer-review under responsibility of the 3rd Current Issues in Business and Economic Studies Conference.

1. INTRODUCTION

We have currently been dealing with big data sets on an increasing scale. In data analysis, we often ask ourselves whether or not a given data set follows specific rules, i.e., a particular probability distribution from a statistical point of view. This approach naturally raises another question about the type and shape of the underlying distribution. Are they always the same or different in each situation? Is it possible to find a rule (law of distribution) that can be used to verify our data's plausibility or unbiased character? Will such a rule vary over time and space, or will it be valid without restrictions? To a certain extent, these and similar questions can be answered in the case of Benford's law of distribution.

Benford's law says that probabilities of digits 1, 2, ..., 9 occurring as the first digit from the left in a given data set are not identical. Numbers starting with one have a higher frequency than those beginning with two, which have a higher frequency than those starting with three, and so on; finally, the least frequent are numbers starting with nine. The probabilities of the second, third, and fourth digits from the left are also different.[i] However, the differences between the occurrence probability values for a given digit gradually decrease until, in the case of the fifth digit from the left, the occurrence probability values for the digits 0, 1, ..., 9 are identical.[ii] The validity of Benford's law has been verified empirically (first in 1881 and again in 1938). New mathematical approaches to this empirical distribution applied at the end of the 20th century led to its incorporation into probability theory. The validity of the Benford distribution has not yet been mathematically proven. Still, its practical validation and use (in accounting, statistics, and natural sciences) demonstrate its validity to a certain degree.

In the context of the significant economic changes that have occurred due to the COVID-19 global pandemic, the question arises whether the rules of Benford's Law are maintained in the event of different economic conditions or after a fundamental change in economic conditions. Therefore, a comparison is proposed between the structures of the data set (by the first or second digit from the left) under conditions of a normally functioning economy on the one hand and under conditions of economic activity reduced due to the pandemic on the other hand. For such an analysis, it is proposed to compare the structure of business sales in selected sectors in 2019 (the year before the pandemic) concerning the values valid in 2020 and 2021 (the years of the pandemic).

2. THEORETICAL BACKGROUND AND APPLICABILITY SCOPE OF BENFORD'S LAW

The first to formulate the law (later called Benford's) was the American mathematician and astronomer Simon Newcomb (cf. Newcomb, 1881). S. Newcomb noticed in libraries that the first pages of logarithmic tables were much more worn than the remaining ones. From this observation, he concluded that students were much more likely to look for logarithms of numbers beginning with one than those beginning with higher digits. He then concluded that the probability of finding a number beginning with one is the highest; that is, higher than the probability of finding a number beginning with two, etc., and empirically derived[iii] that the following Formula gives the probability of finding a number beginning with the leftmost digit d:

$$P(d) = \log_{10}\left(1+\frac{1}{d}\right), \qquad \text{for } d = 1, 2, \ldots, 9. \qquad (1)$$

This rule means that the probability of a number starting with one is 0.3010, that of a number starting with two is 0.1761, and so on; finally, the probability of a number starting with nine is 0.0458. He also derived the occurrence probabilities for the second digit from the left (where it is necessary also to consider zero). The differences between the occurrence probabilities for the digits 0, 1, ..., 9 at the second position from the left are significantly lower than for the first digit from the left (the probability of occurrence for the digit 0 is 0.1197, and for the digit 9 it is 0.0850).

While Formula (1) governs the occurrence probability values for the digits 1, 2, ..., 9 at the first position from the left, such values for the digits 0, 1, ..., 9 at the second position from the left (assuming its independence from the occurrence of a certain first digit from the left)[iv] is given as

$$P(d) = \sum_{k=1}^{9}\log_{10}\left(1+\frac{1}{10k+d}\right), \text{for } d = 0, 1, \ldots, 9. \qquad (2)$$

Considering the independent occurrences of the digits 0, 1, ..., 9 at the third and subsequent positions, the governing Formula can be generalized to

$$P(d_k) = \sum_{d_1=1}^{9}\sum_{d_2=0}^{9}\cdots\sum_{d_{k-1}=0}^{9}\log_{10}\left(1+\frac{1}{\sum_{i=1}^{k} d_i.10^{k-i}}\right), \qquad \text{for } d_k = = 0, 1, \ldots, 9. \qquad (3)$$

Let us observe that the differences between the occurrence probability values for a given digit become significantly smaller at the second position from the left, and at the fifth position from the left (independent of the previous digits), the Benford distribution converges to a uniform multinomial distribution – see Hindls and Hronova (2015) for details.

Since, starting at the third valid digit from the left, the differences in probability values between digits become minimal, the practical use of the Benford distribution is restricted to verifying the independent occurrence of any digit at the first, or possibly the second position from the left.

The fact that we have at our disposal the Benford probability distribution of the first (or second) digit from the left[v] provides the possibility of checking any data set for compliance with the data structure given by this law. A χ^2 goodness-of-fit test is commonly used to check the agreement between the empirical and theoretical distributions. The hypothesis being tested assumes that the empirical distribution matches the theoretical Benford distribution; the alternative hypothesis denies this. The test criterion is given by the statistics

$$G = n\sum_{d=1}^{9}\frac{(p_d - \pi_d)^2}{\pi_d}, \tag{4}$$

whose distribution is approximately $\chi^2[8]$ under the hypothesis H_0;

here π_d – theoretical frequency under the Benford distribution;

p_d – empirical frequency; and

n – sample size.

The critical values are those of the respective quantiles $\chi^2[8]$; most often (at the 5% significance level) we use the 95% quantile of $\chi^2_{0.95}[8] = 15.5$. In the case of digit-matching testing at the second site, we would proceed by analogy, but the number of the groups equals ten, and the number of degrees of freedom equals nine.[vi]

Newcomb's paper did not attract much attention in its time. It was not until the 1930s that the American physicist Frank Benford noticed (as Newcomb had done before) the unequal wear on the different pages of logarithmic tables and arrived at the same Formulas for the first (see (1)) and second (see (2)) digits from the left. He examined many data sets from different fields (hydrology, chemistry, but also baseball and daily newspapers) and empirically verified the validity of the formulas derived by Newcomb – see Benford (1938).

In general, data sets have been identified (from various fields of natural science, economics, and everyday life) for which Benford's law holds, but at the same time it has always been possible to find data sets indicating the contrary.[vii] However, efforts to mathematically prove the validity of Benford's law appeared as late as the late 20th century. This effort was mainly the focus of T. P. Hill (Hill, 1995a; Hill, 1995b; Hill, 1998) and R. A. Raimi (Raimi, 1969a; Raimi, 1969b; Raimi, 1976), who published a number of papers on the subject but did not arrive at a mathematical proof of the law's validity.[viii] They at least succeeded in defining approximate rules for the validity of Benford's law, which can loosely be formulated as follows: if we make random selections from arbitrary distributions, then the set of these random selections approximates the Benford distribution.[ix]

The relationships given by the Benford distribution (see above) hold for arbitrary data sets and are invariant to changes in the base and the measurement units. In other words, data sets that conform to the Benford distribution will conform to it even if the data are in a system other than decimal, or if they are expressed in different units of measure (physical, monetary), or if the original data are multiplied by an arbitrary constant. In other words, any arithmetic operation performed on data obeying the Benford distribution results again in the Benford distribution.[x]

Even though no mathematical proof of the validity of Benford's law has been established, its properties are useful in a number of practical applications (see, e.g., Miller, 2015). In particular, Benford's law has been applied in attempts to prove the unbiasedness of large data sets, i.e., in accounting and macroeconomic statistics. A mismatch of empirical data with Benford's distribution is not yet conclusive evidence of data falsification, but it can be an indicator of certain deviations and a first step toward performing deeper analyses of the data. Therefore, verification of agreement with the Benford distribution should only be used as a complementary tool to traditional procedures or as a first step to finding possible problems. Such a recommendation has been put forth by all authors who deal with the application of Benford's law in economics.[xi]

Benford's law has served very well in the analysis of macroeconomic data.[xii] Rauch et al. (2011) tested the validity of Benford's distribution (by the first digit from the left) on selected national accounts data of the 27 EU Member States from 1999 through 2009.[xiii] Their analysis showed that the least reliable in terms of the validity of Benford's law was the national accounts data of not only Greece[xiv] but also Belgium, Romania, and Latvia. On the other hand, the national accounts data of Luxembourg, Portugal, the Netherlands, Hungary, Poland, and the Czech Republic showed the best agreement with the Benford distribution.

3. VALIDATION OF BENFORD'S LAW ON BUSINESS DATA FROM SELECTED INDUSTRIES IN THE CZECH REPUBLIC

In considering the possibilities of using and empirically verifying Benford's law, we ask whether or not a change in conditions (political, economic, etc.) will result in a significant difference in the structure of the data in terms of the validity of the Benford distribution. To test this hypothesis, we used sales data in selected sectors of the national economy of the Czech Republic from the years 2019 (the last year before the COVID-19 pandemic) and 2020 and 2021 (the years when the functioning of the national economy was most limited due to the pandemic-related measures). However, there was a certain "economic" difference between the two years of the pandemic: while 2020 was characterized by the maximum severity of the restrictions, in 2021, also under public pressure, some easing was already apparent.

The next question we ask is whether Benford's law actually has such a broad validity that it would be verifiable in different situations. Or whether, on the contrary, the arguments are relevant that the validity of Benford's law is essentially random and, consequently, a more general scheme cannot be established for the circumstances under which the rule can be used as a first step to subsequent analyses.

As an example of controversial attitudes, we can mention the auditing profession, where some auditors use Benford's law as a way to initially identify deviations in accounting systems, while others consider its use unnecessary (yet Benford's law is now a standard part of auditing and accounting software packages).

In order to enable a thorough discussion and commentary on this controversy, we have chosen data that meet the generally accepted requirements for the applicability of Benford's law:

1) maximum and minimum values do not limit the data;
2) the data does not contain randomly generated numbers or those used for identification;
3) the data cover at least three orders of the decimal system;
4) the data set is sufficiently large;
5) the data is not influenced by psychology and contains independent items.

For our analysis, we have chosen the results of a sales survey in three sectors in the Czech Republic manufacturing, trade, and hospitality (i.e., accommodation and catering).[xv] Businesses in these sectors were affected by the pandemic in different ways; the manufacturing industry was the least affected, while hotels, restaurants and other accommodation and recreational facilities were the most affected. The source of the data comes from the anonymized results of the annual surveys of the Czech Statistical Office, which we obtained in accordance with the conditions for providing data for research purposes. In terms of the number of units (sample size), the sets of sales corresponded to the requirements for the applicability of Benford's law. The same applies to the sales values reported by the surveyed units.

Table 1 shows the values of the G statistic for all the industries examined and the "compound" of all the sales data used together. As can be seen, the results do not confirm that Benford's law would dominantly hold for these selected indicators or more broadly. This observation is valid despite the fact that the conditions mentioned above for the law's applicability are met.

We now compare the values of the computed G statistics in the view of the χ^2 goodness-of-fit test described above – cf. Formula (4). By comparing with the usual critical quantile value ($\chi^2_{0,95}[8] = 15.5$ and $\chi^2_{0,95}[9] = 16.9$), we can see the conformity of the structure of the sales data with Benford's law for some situations, some industries, and some years, (both for the 1st and 2nd digits), while at other times, *ceteris paribus*, it is not. Just a cursory glance at the values shown in Table 1 clearly indicates no visible pattern in the G results compared to the critical quantile limit of 15.5 and 16.9, respectively.

Table 1. G statistics values

1st digit	Section	G=		
		2019	2020	2021
	C	26.667291	24.968637	30.173140
	G	11.696248	7.612169	5.820760
	I	109.712511	26.995850	31.390540
	All sections	36.174484	5.648677	6.905229

2nd digit	Section	G=		
		2019	2020	2021
	C	7.122656	15.995024	26.967501
	G	15.050749	11.366798	3.811156
	I	4.903436	4.903436	4.084663
	All sections	8.619214	9.526142	12.532839

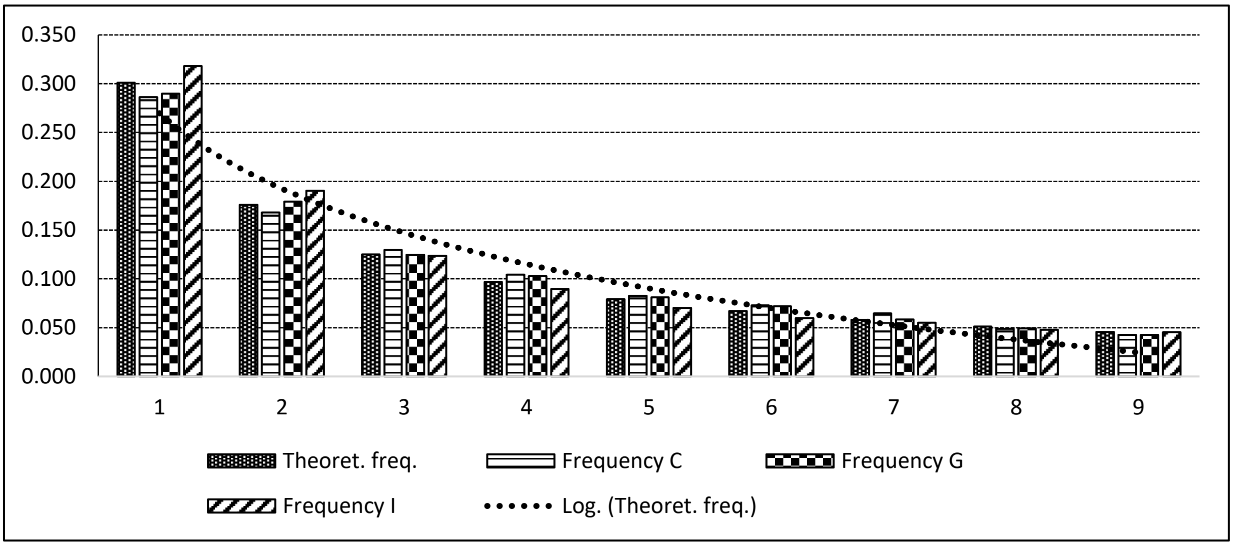
1st digit	Probability	All sections – absolute empirical frequency		
d	π_d	2019	2020	2021
1	0.301031	11 498	11 161	10 942
2	0.176091	6 905	6 610	6 548
3	0.124939	4 700	4 606	4 623
4	0.096910	3 574	3 642	3 605
5	0.079181	2 812	2 890	2 893
6	0.066947	2 434	2 467	2 456
7	0.057992	2 165	2 111	2 073
8	0.051153	1 819	1 849	1 810
9	0.045757	1 660	1 740	1 681
	1.000000	37 567	37 076	36 631

2nd digit	Probability	All sections – absolute empirical frequency		
d	π_d	2019	2020	2021
0	0.119679	4 573	4 468	4 374
1	0.113890	4 171	4 239	4 111
2	0.108821	4 139	4 127	3 991
3	0.104330	3 965	3 760	3 942
4	0.100308	3 778	3 694	3 634
5	0.096677	3 550	3 557	3 447
6	0.093375	3 480	3 447	3 391
7	0.090352	3 398	3 444	3 318
8	0.087570	3 261	3 189	3 191
9	0.084997	3 252	3 151	3 232
	1.000000	37 567	37 076	36 631

Source: the authors' own calculations

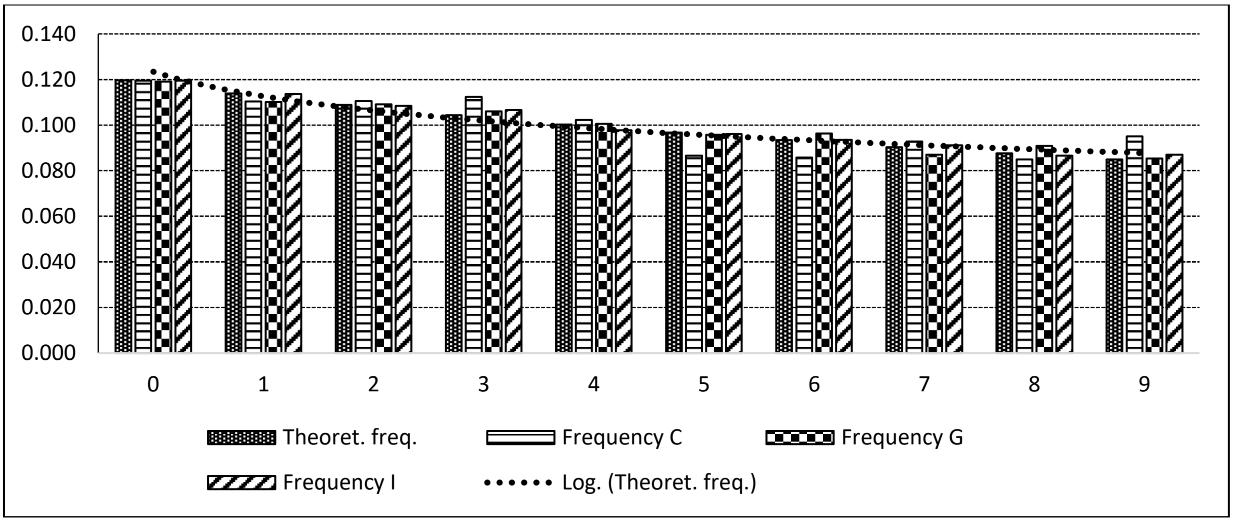
But there is another possibility: a graphical comparison of the occurrence of each digit (see Charts 1 and 2). In all cases analysed, there is present an optical gradient of empirical frequencies towards Benford frequencies, even where the χ^2 test fails. And this is even true for the occurrence of the first digit and that of the second digit. To illustrate, we present plots from two different periods (2019 and 2021) and for the occurrences of the first and second digits, respectively – cf. Formulas (1) and (2).

Chart 1. Theoretical and empirical frequencies of the first digit from the left – 2019



Source: the authors' own calculations

Chart 2. Theoretical and empirical frequencies of the second digit from the left – 2021



Source: the authors' own calculations

4. CONCLUSIONS

The validity of Benford's law in economic applications has been debated for some time. On the one hand, its proponents affirm the broad validity of the law. On the other hand, its opponents point out that its applicability in economics could be more extensive and that, even if it were valid, it would provide little additional stimulating information. In this paper, we have chosen sales indicators in the Czech economy's three different and economically vital sectors.

We have also tested the effect of the COVID-19 pandemic on the change in the data structure. Namely, we have selected those sectors (manufacturing, trade, and accommodation and catering) where the impacts of the restrictions during the pandemic were different; in the case of trade and accommodation, it was very noticeable, which could significantly affect the structure of the data about Benford's law. The validity of this law has only partially been proved; the compliance level varies across sectors and years. It should, however, be noted that the shape of the logarithmic curve of the occurrence frequencies for the first and second digits "optically" confirms the Benford distribution of the occurrence frequencies, even though the goodness-of-fit tests have not proved the full validity of the law.

These results lead us to conclude that Benford's law in economic data will need further, systematic attention; its validity has limits and cannot be trivially generalized. In other words, it is necessary to subject this law to many further analyses of a wide range of economic data.

ACKNOWLEDGEMENTS

This article has been prepared under the Institutional Support for Long-Term Conceptual Development of Research at the Faculty of Informatics and Statistics, Prague University of Economics and Business.

REFERENCES

Benford, F. (1938). The law of anomalous numbers. *Proceedings of the American Philosophical Society*, *78*(4), 551-572. https://www.jstor.org/stable/984802.

Carslaw, C. A. P. N. (1988). Anomalies in income numbers: Evidence of goal oriented behavior. *The Accounting Review*, *63*(2), 321-366. https://www.jstor.org/stable/248109.

Gonzales-Garcia, J., & Pastor, G. (2009). Benford's Law and Macroeconomic Data Quality. *IFM Working Paper 09/10*. https://www.imf.org/external/pubs/ft/wp/2009/wp0910.pdf.

Hill, T. P. (1995a). A Statistical derivation of the significant-digit law. *Statistical Science*, *10*(4), 354-363. https://doi.org/10.1214/ss/1177009869.

Hill, T. P. (1995b). Base-invariance implies Benford's Law. *Proceedings of the American Mathematical Society*, *123*(3), 887-895. https://doi.org/10.2307/2160815.

Hill, T. P. (1998). The first digit phenomenon. *American Scientist*, *86*, 358-363. https://www.jstor.org/stable/27857060.

Hindls, R., & Hronová, S. (2015). Benford's Law and possibilities for its use in governmental statistics. *Statistika*, *95*(2), 54-64. https://www.czso.cz/documents/10180/20550317/32019715q2054.pdf/ab281a2e-6a14-44c6-972a-be1d5fefd026?version=1.0.

Kossovsky, A. E. (2014). *Benford's Law*. World Scientific Publishing.

Miller, S. J. (Ed.). (2015). *Benford's Law. Theory and Applications*. Princeton University Press. http://assets.press.princeton.edu/chapters/s10527.pdf.

Newcomb, S. (1881). Note on the Frequency of use of the different digits in natural numbers. *American Journal of Mathematics*, *41*(1), 39-40. https://www.jstor.org/stable/i340583.

Nigrini, M. J. (1996). A taxpayer compliance application of Benford's Law: Tests and statistics for auditors. *Journal of the American Taxation Association*, *18*(1), 72-79. https://www.proquest.com/openview/25e7a047460619eadea963d5e63ae375/1?pq-origsite=gscholar&cbl=31656.

Nigrini, M. J. (2005). An assessment of the change in the incidence of earnings management around the enron-andersen episode. *Review of Accounting and Finance*, *4*(1), 92-110. https://doi.org/10.1108/eb043420.

Nigrini, M. J. (2011). *Forensic Analytics: Methods and Techniques for Forensic Accounting Investigations.* Wiley Corporate F&A. https://doi.org/10.1002/9781118386798.ch1.

Niskanen, J., & Keloharju, M. (2000). Earnings cosmetics in a tax-driven accounting environment: evidence from finnish public firms. *European Accounting Review*, *9*(3), 443-452. https://doi.org/10.1080/09638180020017159.

Raimi, R.A. (1969a). The peculiar distribution of first digits. *Scientific American*, *221*(6), 109-119. https://doi.org/10.1038/scientificamerican1269-109.

Raimi, R.A. (1969b). On distribution of first significant figures. *American Mathematical Monthly*, *76*(4), 342-348. https://doi.org/10.2307/2316424.

Raimi, R.A. (1976). The first digit problem. *American Mathematical Monthly*, *83*(7), 521-538. https://doi.org/10.1080/00029890.1976.11994162.

Rauch, B., Göttsche, M., Engel, S., & Brähler, G. (2011). Fact and fiction in EU-Governmental economic data. *German Economic Review*, *12*(3), 243-255. https://doi.org/10.1111/j.1468-0475.2011.00542.x.

Watrin, Ch., Struffert, R., & Ullmann, R. (2008). Benford's Law. An instrument for selecting tax audit targets. *Review of Managerial Science*, *2*, 219-237. https://doi.org/10.1007/s11846-008-0019-9.

[i] Naturally, the digit 0 must also be taken into consideration here.
[ii] Cf., e.g., Hindls and Hronová (2015).
[iii] Cf. Newcomb (1881).
[iv] If the occurrence probability of a digit from among 0, 1, ..., 9 at the second position from the left depends on the occurrence of a certain digit 1, 2, ..., 9 at the first position from the left, it is necessary to formulate a model of conditional probabilities – see Hindls and Hronova (2015) for details.
[v] For the reasons explained above, we will not consider the other positions from the left.
[vi] Another option to test the sample data for fit to the Benford distribution would be to use the Z statistic – see Kossovsky (2014); or to use the mean absolute difference (MAD) – see Nigrini (2011).
[vii] In particular, where the variables can only take values from a certain interval given by the factual nature of the data, e.g., shoe or clothing size, etc.
[viii] Cf. Raimi (1969b), p. 347.
[ix] Cf. Hill (1998) and Raimi (1969b).

[x] Cf., e.g., Watrin et al. (2008).

[xi] Cf., e.g., Carslaw (1988), Nigrini (2005), Nigrini (1996), Guan et al. (2006), Niskanen and Keloharju (2000), and Watrin et al. (2008).

[xii] Cf., for example, Nye and Moul (2007) or Gonzales-Garcia and Pastor (2009).

[xiii] In total, there were 36,691 values contained in 297 files.

[xiv] However, the problem with Greece's national accounts data had already been known. Eurostat refused to accept the Greek government's data twice, in 2002 and again in 2004, due to the unreliability issues (cf. Report by Eurostat on the Revision of the Greek Government Deficit and Debt Figures – http://ec.europa.eu/eurostat/documents/4187653/5765001/GREECE-EN.PDF).

[xv] The definitions of the sectors correspond to the sections of the CZ-NACE classification: Section C – Manufacturing, section G – Wholesale, retail trade, repair and maintenance of motor vehicles, section I – Accommodation, catering and restaurant services