Chaos, Solitons and Fractals 107 (2018) 186–194

Contents lists available at ScienceDirect

Chaos, Solitons and Fractals

Nonlinear Science, and Nonequilibrium and Complex Phenomena

journal homepage: www.elsevier.com/locate/chaos

Regular paths in financial markets: Investigating the Benford's law

Jessica Riccioni, Roy Cerqueti*

University of Macerata, Department of Economics and Law, Via Crescimbeni 20, Macerata I-62100, Italy

ARTICLE INFO

Article history:
Received 20 October 2017
Revised 21 December 2017
Accepted 4 January 2018

Keywords:
Data science
Benford law
Stock markets
Finance

ABSTRACT

This paper aims at verifying whenever the Benford's Law is valid in the context of global stock markets, which are here viewed as complex systems. In so doing, we pursue the scope of assessing the presence of data regularities and interpret obtained discrepancies.

Specifically, we check the reliability of Benford's Law for all the indexes listed on the stock exchanges of several countries, with a particular reference to prices and volumes of stocks.

To pursue our scope, we adopt comparison criteria grounded on statistical theory, like the Chi-squared test for both the distributions of the first and the second meaningful digits.

Evidence of violations is provided and some insights taken from the historical facts and economic shocks are carried out.

© 2018 Published by Elsevier Ltd.

1. Introduction

The evolution of prices and volumes in the international stock markets has a remarkable informative content. In fact, financial markets and their components are highly correlated with the surrounding socio-political environment and a shock of noneconomic nature might have an effect on the economic system.

This paper moves from this scientific ground. We aim at analyzing the context of global stock markets with a specific focus on all the indexes composing them. In particular, we deal with the nature of complex system of the markets, and discuss whenever the available time series associated to prices and volumes of the stocks exhibit a sort of "regularity".

Indeed, datasets exhibit often deviations from a pure uniform law, even if they are not constructed in an "ad-hoc" manner. Think about countries' GDPs, with a great percentage of low levels and a few data with high values.

In this respect, the so-called Benford's Law (BL), firstly empirically observed by Newcomb [21] and lately formalized by Benford [6], has been proved to hold in a wide set of contexts.

BL is an evident regularity of the distribution of the digits of a number of large datasets. It states that the frequency of the first and second digits of the values of a set of data decreases with the value of the digit, and achieves its maximum when the digit is "one".

More in details, Benford [6] provided an extensive study of 20,149 data from 20 different areas, and calculated the frequency of each first digit from 1 to 9 and of each second digit from 0 to 9.

The outcome of this effort has led to the statement of an exact rule given by a logarithmic relationship.

For the first digit, BL says that:

$$P(first\ digit = n) = \log_{10}\left(1 + \frac{1}{n}\right). \tag{1}$$

where $P\ (first\ digit = n)$ is the probability that a number has the first digit equals to n, with $n = 1, \ldots, 9$.

As for the probability of the second digit, BL can be written as follows:

$$P(second\ digit = n) = \sum_{k=1}^{9} log_{10}\left(1 + \frac{1}{10k + n}\right). \tag{2}$$

where $n = 0, 1, \ldots, 9$.

The theoretical frequencies are shown in Table 1.

Generalizations of formulas (1) and (2) have been also provided, leading to the distributions of the digits subsequent to the second and obtaining frequencies more equally distributed. From the fifth digit the distribution no longer follows the BL but becomes a uniform distribution.

Not all dataset are eligible to be controlled for agreement with the BL. In particular, data must be not chosen under an "ad hoc" criterion and the cardinality of the dataset must be large. Such requirements are fulfilled in our specific context.

We check the validity of the BL for extensive daily series of prices and volumes of 4166 stocks, listed in the main international financial markets. Several clusterings of such series are performed at the level of the related geographic areas (Italy, Europe, USA, Asia

* Corresponding author.
E-mail address: roy.cerqueti@unimc.it (R. Cerqueti).

https://doi.org/10.1016/j.chaos.2018.01.008
0960-0779/© 2018 Published by Elsevier Ltd.

Table 1
Relative frequencies of the first and second significant digit.

1st digit		1	2	3	4	5	6	7	8	9
Frequency		0.301	0.176	0.125	0.097	0.079	0.067	0.058	0.051	0.044
2nd digit	0	1	2	3	4	5	6	7	8	9
Frequency	0.119	0.114	0.109	0.104	0.100	0.097	0.093	0.090	0.088	0.085

and Pacific area and a generic “other relevant markets”) and of the individual stock markets.

The beginning of the reference period varies with the stock, on the basis of the free availability of the data on the yahoo.com/finance website. The ending date is for all November 14th, 2014.

Both first and second digits distributions are explored.

The analysis is performed by the application of a Chi-squared procedure, to check whenever the empirical data obeys with the theoretical frequencies of BL.

The analysis reveals the presence of deviations of the data digits from the statements of the BL. In accord to the literature (see the next Section), they might be caused by shocks of different nature which are able to influence the dynamics of the international stock markets components. A discussion of such deviations and also of the regularities is provided, along with a related socio-economic-political contextualization.

The rest of the paper is organized as follows: Section 2 contains a critical review of the main contributions of the literature. Section 3 is devoted to a detailed description of the data used and of the methodological instruments employed to perform the analysis. Section 4 outlines and comments the obtained results, and proposes a reading of the deviations from the statements of the BL through the occurred events. Section 5 offers some conclusive remarks and gives some suggestions for future research lines.

2. Relevant literature

Here are briefly described the cases in which this law has been applied with success over the years.

In the context of social science, Cho and Gaines [27] took care of fraud in funding for charity campaigns. See also [3,5,18,19], who tested the consistency of social data and assessed the presence of some sort of manipulation.

In the political environment Mebane [15] used the law for verify the trustworthiness of the election results. In this respect, the law was also used by the Iranian minister Boudewijn F. Roukema to ensure the accuracy of the electoral results in the presidential election of 2009, as a consequence of the complaint about electoral fraud made by Moussavi (opposition leader) against the election of President Ahmadinejad. City after city, the minister analyzed the relative frequencies of the first significant digit of the number of votes and compared the frequencies obtained with the theoretical frequencies of BL. As a result Roukema remarked that while all the frequencies of the other digits were very similar to those theoretical, the number 7 appeared too many times, hence supporting the truthfulness of the charges of Moussavi. Decided to settle the matter, the minister remarked that the anomaly concerned three out of the six larger areas of Iran. In these areas the winner Ahmadinejad had won a share of the vote much higher than in other areas.

Also Varian [28] suggested the possibility of using BL to identify any falsification in the collection of data used to support political decisions. The author proposed to compare the relative frequency of the first digit of the numbers used to the support political decisions with the theoretical frequencies of BL, so as to highlight any abnormal results. He applied the law in a set of land planning for 777 tracts in San Francisco Bay. The law was respected.

Nigrini [22–24] was able to show the benefits of the possible application of BL to detect the falsification and the frauds in accounting and activity auditing (which of course may not be intentional, such as arithmetic errors and, more generally, errors in calculation or misapplication of the applicable accounting standards; or it could be outright intentional frauds, such as the alteration of records or documents, the lack of enforcement of accounting standards, the omission of some results or finally recording non existent transactions). This system is now used in the majority of American States to support the identification of tax fraud.

In the economic field BL can also be applied to check the efficiency of financial indexes, and in general financial quantities. Ley [14] calculated if the distribution of the relative frequencies of the first significant digit of daily returns of two American stock indexes (i.e. the S & P for the period from 1926 to 1993 and the Dow Jones for the period from 1900 to 1993), using the BL: the result obtained that both stock indexes respected this law and it was a demonstration of the current index efficiency.

Realdon [26] and Ausloos et al. [2] dealt with the assessment of the validity of the BL on the sovereign credit default swap markets. Carrera [7] analyzed exchange rates to check whether there exists the possibility of manipulation for policy management. In support of macroeconomic data, Nye and Moul [25] studied the GDP data of the Penn World Tables. Günnel and Tödter [10] applied the BL on the forecast of GDP growth and inflation of German consumer prices. Holz [11] evaluated the quality of the Chinese GDP advancing suspicions on the statistics published by the National Bureau of Statistics of China. Michalski and Stoltz [16] examined the balance of payments data for 103 countries between 1989 and 2007. Mir [17] studied the illicit financial outflows from developing countries. Finally, Mir et al. [20] and Ausloos et al. [4] have instead applied the BL to the aggregate of income taxes of municipalities and Italian regions for the period between 2007 and 2011.

In this work we have tried to interpret the efficacy and functionality of financial markets worldwide. We checked the reliability of BL for all stock indexes listed on the stock exchanges of the various countries but not to apply it on daily returns but on prices and average volumes of stocks comprising indexes.

We are in the line traced by other scholars dealing with the assessment of BL in financial markets (see e.g. [8,9,13]). However, this is the first paper dealing with a so global context, with all the international stock markets and both the key stock variables (volumes and prices).

It will be interesting to see how this application of BL analyses and comments consistent results; in fact, the cases in which the Law will be verified are various, and we provide a brief discussion on the technical and historical reasons leading to the failure of this law.

3. Data analysis and methodology

3.1. Data

All processed data are freely available on the web (see e.g. https://finance.yahoo.com). The dataset is of scientific reliability, as also witnessed by authoritative financial studies who had referred to such a dataset for collecting empirical data (see e.g. [1,12]).

Download English Version:

https://daneshyari.com/en/article/8254133

Download Persian Version:

https://daneshyari.com/article/8254133

Daneshyari.com