Economics Bulletin

Volume 35, Issue 2

Benford's law for audit of public works: an analysis of overpricing in Maracanã soccer arena's renovation

Flavia C. Rodrigues da Cunha
Brazilian Federal Court of Accounts

Mauricio S. Bugarin
University of Brasilia

Abstract

Auditing of public works is a time consuming task because budget worksheets are often long and difficult to analyze. The present work illustrates the application of Newcomb-Benford Law (NB Law) to detecting overpricing in worksheets of public works. That law suggests that the frequency of the first digit in a multitude of non-manipulated numerical databases is decreasing from digit 1 to digit 9. The paper describes the relevant statistical tests of NB Law and applies these tests to the work of Brazil's Maracanã Soccer arena renovation for the 2014 FIFA World Cup. Next, it compares NB Law's results with those obtained with the analysis of prices conducted by the Brazilian Court of Accounts (TCU). The tests identified 17 items in the worksheet that did not comply with the Law and corresponded to 71.54% of the total overpricing uncovered by TCU.

The authors are grateful to two anonymous referees for their invaluable comments. The authors remain sole responsible for opinions expressed, remaining errors or omissions. Mauricio Bugarin gratefully acknowledges the research grant of the Brazilian Research Council, CNPq.
Citation: Flavia C. Rodrigues da Cunha and Mauricio S. Bugarin, (2015) ''Benford's law for audit of public works: an analysis of overpricing in Maracanã soccer arena's renovation'', *Economics Bulletin*, Volume 35, Issue 2, pages 1168-1176
Contact: Flavia C. Rodrigues da Cunha - flaviaceccato@hotmail.com, Mauricio S. Bugarin - bugarin.mauricio@gmail.com.
Submitted: August 14, 2014. **Published:** May 14, 2015.

1. Introduction

If we roll a fair dice, the chances of getting number 5, for example, is 1/6, since the dice has six sides. When we flip an unbiased coin, there is a 50% chance of head or tail. Thus, we tend to think that in a numeric database, the probability of randomly choosing a number that has first digit 1 is 1/9; the same applies to all other digits between 2 and 9.

However, 19^{th} century mathematician Simon Newcomb (1881) noted that the first pages of logarithm tables were more worn than the following ones, suggesting that the most commonly accessed value was 1. Since Newcomb did not gather numerical data or provide other evidences supporting his discovery, that fact only gained importance over half a century later, when the physicist Frank Benford (1938) reached the same conclusion. Benford published a seminal article in 1938, "The Law of Anomalous Numbers", which used data collected from numerous different sources. These data were random, and not related to each other, and ranged from numbers collected from the front pages of newspapers to river lengths and to mathematical tables and scientific constants. He recorded the first digit of the collected data and found that: 30.6% of the numbers had 1 as the first digit; the first digit 2 occurred in 18.5% of cases; and that, in contrast, only 4.7% of the numbers had 9 as the first digit. Such frequencies of first digits were shown to appear in a variety of databases, including energy bills, addresses, stock prices, city population values, and mortality rates, among others. That distribution is known as **Benford distribution** and the property discovered by Newcomb and Benford is known as **Newcomb-Benford's Law** or, more simply, **Benford's Law**.

In order to better understand the differences in frequency of the first digit, suppose you invest 10,000.00 dollars in a pension fund that will offer you a fixed return of 7% per annum. Then, your investment will double roughly every ten years. Therefore, after ten years with 1 as the first digit, the balance of your investment will eventually reach 20,000.00 dollars. After another 10 years, the balance will double to 40,000.00, so the numbers 2 (first part of the decade and 3 (second part) will appear in 10 years. After another decade, the amount will reach 80,000.00, so that the digits 4, 5, 6 and 7 will successively appear as first digits in the ten-year period. Eventually, the investment will reach the value of 100,000.00, with the first digit 1 materializing for another ten years, and so on. Thus, when choosing a random date, it is more likely that the balance of your investment's first digit is 1 than any other digit. This same logic applies to a multitude of data present in nature.

A database is more likely to represent a Benford distribution when data are collected from different sources (Hill, 1995). On the other hand, numbers assigned by human intervention, such as Social Security numbers, postal codes, bank accounts, phone numbers, or numbers produced by students in experiments usually do not conform to Benford's Law (Nigrini, 2000). This observation suggests that the "Law of Anomalous Numbers" may be used to detect evidence of human manipulation of data. Naturally, deviations from Benford's distribution do not constitute conclusive proof of manipulation, just as compliance does not ensure data reliability. However, nonconformity can be seen as a signal that the data need scrutiny. Thus, Benford's Law (NB Law) can be used in conjunction with other control mechanisms as a guide to check for possible manipulations.

The literature presents several empirical analyses based on the hypothesis that fabricated data do not follow Benford's distribution. Nigrini (2012), assuming that true financial data followed Benford distribution closely (as indicated by his previous research), argued that substantial deviations from this law suggest possible fraud or concocted data. Nigrini developed several tests to measure compliance with Benford's Law, and the Wall Street Journal (Berton, 1995) reported that the Attorney's office in Brooklyn, New York, detected fraud in seven companies in New York using these tests. The evidence found that fraudulent data reported too small frequencies of first digit 1 and too high frequencies of first digit 6. Based on the success cases, Nigrini became a consultant to internal revenue agencies of different countries and developed computer tests of NB Law to detect fraud that are currently being used by those agencies.

Rauch Göttsche, Brähler and Engel (2011) published an article in the German Economic Review showing that Benford's Law could be used to test manipulation in macroeconomic data, and suggesting which data needed a more rigorous inspection. They analyzed the first digit of macroeconomic data reported to the Statistical Office of the European Union (Eurostat) for EU countries and constructed a ranking of the 27 member countries according to the extent of the deviation from NB Law predictions. The country that presented the highest deviation was Greece, which manipulation of the data had been officially confirmed by the European Commission (2010).

University of Michigan professor Walter Mebane analyzed election data from several countries and discovered that the count of votes tended to follow Benford's Law for the second digit (Mebane, 2006) for the United States, Russia and Mexico. However, using data from the Iranian elections in 2009, Mebane (2009, 2010) found that in cities with few invalid votes, Ahmadinejad's votes strongly diverged from Benford distribution predictions and the official candidate, in these situations, had a large vote advantage.

The present paper aims to illustrate the use of Benford's Law to analyze public works, using the budget worksheets of the renovation of Brazil's Maracanã soccer arena, the iconic symbol of world soccer talent. That renovation work was selected, on one hand, due to the relevant volume of data available and, on the other hand, because it is possible to compare the test results with the price analysis undertaken by the Brazilian Court of Accounts (TCU).

2. Tests of Benford law based on the digits' frequencies

The tests used in the present study are carefully characterized in Nigrini (2012). We quickly present here their formats.

2.1. First-Two Digits Test

According to NB Law, the expected relative frequency of a number in which the first digit is $D_1 = d_1$ and the second digit is $D_2 = d_2$ is:

$$\text{Prob}(D_1 D_2 = d_1 d_2) = \log\left(1 + \frac{1}{d_1 d_2}\right) \quad \text{where} \quad D_1 D_2, d_1 d_2 \in \{10, 11, \dots, 99\}.$$

The test consists in comparing each two-digit's observed relative frequency with the (above) expected one by means of a typical *Z*-statistic. The *Z*-statistic is calculated as shown below, where $i \in \{10, 11, \dots, 99\}$ is the analyzed two-digit category, $n = 90$ is the number of two-digit categories, RF_i is the observed relative frequency of two-digits i, and ERF_i is the expected relative frequency of two-digits i.

$$z_i = \frac{|RF_i - ERF_i| - \frac{1}{2n}}{\sqrt{\frac{ERF_i(1 - ERF_i)}{n}}}$$

The 5% significance level threshold is 1.96. If the *Z*-statistic of a two-digit exceeds 1.96, the frequencies of the items starting with these two digits do not conform to the predicted ones. Therefore, these are the candidates for further inspection.

Nigrini (2012) suggests three criteria for overall compliance with MB Law based on the two-digit tests. Firstly, if no more than 5 two-digits among all 90 classes $\{10,11,\dots,99\}$ do not conform, there is no strong evidence of manipulation.

Secondly, a chi-squared statistic is also calculated as follows, where F_i is the observed frequency of two-digits i and EF_i is the expected frequency of two-digits i.

$$\chi^2 = \sum_{i=10}^{99} \frac{(F_i - EF_i)^2}{EF_i}$$

The 5% confidence threshold critical value for 89 degrees of freedom is 112.02. Therefore, if the chi-squared statistic exceeds 112.02 there is evidence of an overall non-conformity of the observed distribution with NB Law.

Finally, a mean absolute deviation (MAD) test is based on the absolute differences between observed and expected relative frequencies, according to the following statistic.

$$MAD = \frac{1}{n} \sum_{i=10}^{99} |R_i - E_i|$$

Nigrini (2012) proposes the following conformity criteria for the MAD test. If the MAD statistic is lower than 0.0012, there is *close conformity*; if it is higher than 0.012 but lower than 0.0018, there is *acceptable conformity*; if it lies in the interval (0.0018, 0.0022] there is *marginally acceptable conformity* and finally, if it exceeds 0.0022 there is *nonconformity.*

2.2. Summation Test

Nigrini (2012) simulated a Benford distribution and separated the resulting sample into 90 classes according to the first two digits $\{10,11,\dots,99\}$. Then, he added all number observations in each group and found evidence that all sums led to approximately the same amounts. In other words, the numbers in each class tended to sum up to 1/90=0.011 or 1.1% of the total sum of all numbers in the sample.

However, the author found that actual data rarely conformed completely to such a standard. The usefulness of this test is precisely to point out the nonconformities. There are no threshold explicitly suggested by Nigrini (2012); therefore, we take the 1.1% percentage as the upper bound for conformity in our analysis.

2.3. The confrontation between the two tests

Any two-digit category that falls into the nonconformity criteria range for either the *Z*-test or the summation test is a candidate for further scrutiny. However, some two-digit categories may fall into nonconformity simply because of their lack of frequency in the database. In that case, it might be an unrewarding task to dedicate time analyzing the corresponding items. Therefore, we propose to compare the frequencies of all categories that have been selected in at least one of the two tests. If one of them shows very little frequency according to both criteria, i.e., there are few observations in that category and the value of the summation of the category's items is low, then that category should be excluded from further scrutiny. We call this comparison the "confrontation" of the two tests. Our main point in doing the confrontation is that, in the case of public works' budget, the relevance of each group should be taken into account for selecting of the digits that need further auditing.

3. Analysis of Maracanã Soccer arena's renovation

The analysis of this study focused on the budget of Maracanã soccer arena's renovation originally presented to TCU by the procurement winning firm in the amount of R$ 931,885,382.19 (over US$412 million as of August 1st, 2014).

Subsequently, after the TCU auditing, alternative budgets that aimed at eliminating detected overpricing of most worksheet items were negotiated. We selected the initial budget for three main reasons. First, the subsequent budgets were changed after the TCU auditing; therefore, these budget sheets were not entirely formulated by the winning bidder. Since we wish to detect possible data manipulation from that bidder, the original bid should be used. Second, the first budget sheets were subject to careful TCU auditing that revealed significant overpricing. Therefore, we will be able to compare the results of our analysis based of NB Law with the results of TCU's auditing. Third, the TCU analysis is based on the ABC curve methodology, which consists of ordering the items in a budget sheet from most expensive to least expensive and selecting up to 20% of the most expensive items, until the total cost of those items adds up to about 80% of the total budget, and them compare those prices with market benchmarks. Therefore, the TCU did not make use of our proposed methodology in its analysis, which makes the comparison valuable.

The present study tests only the unit costs data, but it could alternatively test quantities of services or total prices (adding all multiples of the same service or item)[1]. We analyzed 828 items; the values smaller than R$ 10.00 were naturally excluded, because they do not have a second digit[2].

3.1. First-Two Digits Test

The results of the first-two digits test are reported in Table 1, where: "Dig." refers to the first two digits; "C" is the absolute frequency (count) of items staring with the corresponding first two-digits in the worksheet; "Actual" is the corresponding relative frequency; "LB" is the expected relative frequency according to NB Law; "Diff" is the difference between "Actual" and "LB"; "*Z*-Test" refers to the *Z*-statistic; "CS" is the Chi-Square statistic; and "MAD" is the Mean Absolute Deviation statistic. For the sake of space, we present only the tests for the two digits from 10 to 51, the range where all anomalies were detected; however, all tests were computed on the entire range 10 to 99.

According to Table 1, there are more intense peaks in the digits 11, 16, 25, 28 and 42 with respect to the proportions of the descending curve of NB Law. The results of the *Z*-statistic are 11 (2.954), 16 (2.105), 25 (2.524), 28 (2.303) and 42 (2.060).

Since only five digits among the 90 digits exceed the limit of 1.96, the proportions of the first digit of the unit costs of Maracanã soccer arena's renovation, in general, do not show inconsistency with NB Law if we follow Nigrini's (2012) suggestion to consider acceptable

---

[1] It is not clear whether both alternatives would lead to the same results in general. However, in the present application the two methodologies yielded similar results. In fact, the use of total prices rather than unit prices allowed us to identify 83.88% of total overpricing uncovered by TCU. Details can be obtained upon request to the authors.

[2] In some countries, such as the USA, businesses tend to set prices using the .99 strategy, i.e., instead of US$10, the price is set at US$9.99 or instead of US$50 it is set at 49.99. Such pricing strategy is, indeed, a possible reason for nonconformity with NB Law. However, in our original worksheet including 1028 items, only 12 items ended in .99, i.e., about 1% of the sample. Therefore, there appears to be no reason for concern with the .99-anomaly in the present study.

the occurrence of up to five peaks. Note, however, that we have reached the upper bound of Nigrini's threshold.

The Chi-Square statistic is 106.65. The critical value for 89 degrees of freedom and 5% significance level is 112.02. Therefore, we cannot reject the null hypothesis, suggesting accordance with NB Law.

The last test applied is MAD. The test statistic found for Maracanã soccer arena's renovation is 0.0031, which highly exceeds the 0.0022 threshold adopted by Nigrini (2012). This result suggests possible manipulation of data.

Table 1 – First-Two Digits Test for unit costs of Maracanã soccer arena's renovation

Dig.	C	Actual	LB	Diff.	Z-Test	CS	MAD	Dig.	C	Actual	LB	Diff.	Z-Test	CS	MAD
10	43	0.052	0.041	0.011	1.435	2.222	0.011	**31**	9	0.011	0.014	-0.003	0.571	0.512	0.003
11	48	0.058	0.038	0.020	2.954	8.925	0.020	**32**	13	0.016	0.013	0.002	0.434	0.338	0.002
12	28	0.034	0.035	-0.001	0.054	0.021	0.001	**33**	15	0.018	0.013	0.005	1.157	1.694	0.005
13	26	0.031	0.032	-0.001	0.029	0.016	0.001	**34**	15	0.018	0.013	0.006	1.271	2.009	0.006
14	28	0.034	0.030	0.004	0.548	0.410	0.004	**35**	8	0.010	0.012	-0.003	0.515	0.448	0.003
15	31	0.037	0.028	0.009	1.535	2.616	0.009	**36**	13	0.016	0.012	0.004	0.848	1.005	0.004
16	32	0.039	0.026	0.012	2.105	4.772	0.012	**37**	10	0.012	0.012	0.000	0.133	0.018	0.000
17	26	0.031	0.025	0.007	1.105	1.443	0.007	**38**	9	0.011	0.011	0.000	0.112	0.012	0.000
18	18	0.022	0.023	-0.002	0.216	0.107	0.002	**39**	10	0.012	0.011	0.001	0.132	0.088	0.001
19	16	0.019	0.022	-0.003	0.458	0.324	0.003	**40**	12	0.014	0.011	0.004	0.884	1.097	0.004
20	18	0.022	0.021	0.001	0.110	0.012	0.001	**41**	10	0.012	0.010	0.002	0.285	0.206	0.002
21	14	0.017	0.020	-0.003	0.550	0.445	0.003	**42**	2	0.002	0.010	-0.008	2.060	4.934	0.008
22	12	0.014	0.019	-0.005	0.880	0.993	0.005	**43**	7	0.008	0.010	-0.002	0.268	0.194	0.002
23	10	0.012	0.018	-0.006	1.240	1.838	0.006	**44**	8	0.010	0.010	0.000	0.029	0.001	0.000
24	12	0.014	0.018	-0.003	0.574	0.489	0.003	**45**	7	0.008	0.010	-0.001	0.144	0.103	0.001
25	24	0.029	0.017	0.012	2.524	6.944	0.012	**46**	4	0.005	0.009	-0.005	1.168	1.802	0.005
26	7	0.008	0.016	-0.008	1.662	3.182	0.008	**47**	4	0.005	0.009	-0.004	1.121	1.684	0.004
27	10	0.012	0.016	-0.004	0.718	0.724	0.004	**48**	6	0.007	0.009	-0.002	0.337	0.270	0.002
28	4	0.005	0.015	-0.010	2.303	5.887	0.010	**49**	6	0.007	0.009	-0.002	0.285	0.220	0.002
29	7	0.008	0.015	-0.006	1.353	2.210	0.006	**50**	9	0.011	0.009	0.002	0.519	0.496	0.002
30	13	0.016	0.014	0.001	0.208	0.124	0.001	**51**	8	0.010	0.008	0.001	0.197	0.148	0.001
N	**828**												**χ^2 Actual** **χ^2 threshold**	**106.65** **112.02**	**MAD** **0.0032**

Source: author's calculations

3.2. Summation Test

In order to assess the significance of each pair of digits in the budget worksheet we perform the complementary Summation Test. The results are shown in Table 2 below, where the 1st and 6th columns refer to the first two digits of the observations; the 2nd and 7th columns correspond to the sum of the items that have the first two digits indicated in the 1st and 6th columns, respectively; the 3rd and 8th columns show the proportions of the sums calculated in 2nd and 7th columns with respect to the sum of all unit costs of the worksheet; the 4th and 9th columns present the expected relative frequencies according to NB Law; and columns 5 and 10 compute the difference between the proportions of the sums and the relative frequencies of NB Law. To save space, we present only the tests for the two digits from 10 to 51, the range where all anomalies were detected.

Table 2 highlights peaks in the first two digits 11, 17, 18, 19, 20, 21, 22, 25, 32 and 48. It is noteworthy the very high ratio of appearance of digits 25, representing 48.3% of total unit costs. The test strongly suggests nonconformity to NB Law.

Table 2 - Summation Test for unit costs of Maracanã soccer arena's renovation.

Digit	Sum	Actual	Benford	Difference	Digit	Sum	Actual	Benford	Difference
10	1,117,783.14	0.002	0.011	-0.009	**31**	1,124.59	0.000	0.011	-0.011
11	35,228,545.85	0.060	0.011	0.049	**32**	32,871,298.35	0.056	0.011	0.045
12	9,996.90	0.000	0.011	-0.011	**33**	40,266.66	0.000	0.011	-0.011
13	1,427,480.57	0.002	0.011	-0.009	**34**	3,453,111.19	0.006	0.011	-0.005
14	149,926.09	0.000	0.011	-0.011	**35**	8,277.56	0.000	0.011	-0.011
15	7,327.10	0.000	0.011	-0.011	**36**	14,275.68	0.000	0.011	-0.011
16	12,400.03	0.000	0.011	-0.011	**37**	5,748.38	0.000	0.011	-0.011
17	19,334,196.90	0.033	0.011	0.022	**38**	40,635.97	0.000	0.011	-0.011
18	18,810,868.98	0.032	0.011	0.021	**39**	2,169.59	0.000	0.011	-0.011
19	22,983,744.00	0.039	0.011	0.028	**40**	6,314.11	0.000	0.011	-0.011
20	20,216,982.58	0.035	0.011	0.024	**41**	85,478.77	0.000	0.011	-0.011
21	23,621,379.30	0.040	0.011	0.029	**42**	8,514.19	0.000	0.011	-0.011
22	22,910,130.11	0.039	0.011	0.028	**43**	6,138.93	0.000	0.011	-0.011
23	3,365.19	0.000	0.011	-0.011	**44**	4,472,961.24	0.008	0.011	-0.003
24	2,441,496.65	0.004	0.011	-0.007	**45**	4,557,697.21	0.008	0.011	-0.003
25	282,240,352.41	0.483	0.011	0.472	**46**	605.18	0.000	0.011	-0.011
26	3,516.32	0.000	0.011	-0.011	**47**	5,356.92	0.000	0.011	-0.011
27	3,699.11	0.000	0.011	-0.011	**48**	48,849,864.06	0.084	0.011	0.073
28	629.84	0.000	0.011	-0.011	**49**	7,018.04	0.000	0.011	-0.011
29	9,100.20	0.000	0.011	-0.011	**50**	2,718.91	0.000	0.011	-0.011
30	3,041,197.41	0.005	0.011	-0.006	**51**	1,805.57	0.000	0.011	-0.011
TOTALSUM					**584,727,527.67**				

Source: author's calculations

3.3. Confrontation between the First-Two Digits Test and the Summation Test

Next, we select the digits detected as critical in the First-Two Digits Test and Summation Test. Then we carry out a confrontation between these tests to confirm the sample relevance of the selected digits, comparing their relative frequency in each one of the tests. The results are shown in Table 3.

Table 3 - Confrontation between the digits selected in the First-Two Digits Test and Summation Test.

Digits	First-Two Digits Test	Summation Test	Critical Digits	Digits	First-Two Digits Test	Summation Test	Critical Digits
11	0.058	0.060	Yes	22	0.014	0.039	Yes
16	0.039	0.000	Yes	25	0.029	0.483	Yes
17	0.031	0.033	Yes	28	0.005	0	No
18	0.022	0.032	Yes	32	0.016	0.056	Yes
19	0.019	0.039	Yes	42	0.002	0	No
20	0.022	0.035	Yes	48	0.007	0.084	Yes
21	0.017	0.040	Yes				

Source: author's calculations

Table 3 shows the digits that were selected by either one of the tests in column 1. Column 2 shows the relative frequencies of these digits according to column "Actual" in Table 1. Column 3 displays the proportions of the sum of items starting with these digits according to the "Actual" column in Table 2. Column 4 singles out the two-digits that have little significance in the spreadsheet according to both criteria: low percentage of items starting with those two digits and low sum of the corresponding values as percentage of the sum of all items (No).

The confrontation between the tests suggests excluding digits 28 and 42 from our analysis.

The results of the First-Two Digits Test points to the digits 11, 16 and 25, once we exclude digits 28 and 42. The Summation Test identifies excessive values for the proportions of 11, 17, 18, 19, 20, 21, 22, 25, 32 and 48 values. Note that the digits 11 and 25 were selected in both analyzes as excessive and therefore are more likely to be overpriced. As a guideline for the auditing, the tests would recommend first analyzing the items containing the first two digits 11 and 25, then examine the remainder.

3.4. Comparison with the Brazilian Court of Accounts' analysis

Table 4 details the items with first two digits 11 and 25 with the overpricing uncovered by TCU.

Table 4 - Confrontation between the results of tests of the NB Law and TCU's overpricing analysis for the digits 11 and 25

Digit	Service Description	Unit Cost (R$)	Overpricing detected by TCU
25	Tension roofing system, including insulated metal structure fully locked cables and "PTFE" membrane according to the Hightex project for Maracanã Stadium World Cup 2014's renovation	256,714,917.00	26,961,972.80
	Sports and auditorium furniture for the Maracanã Stadium World Cup 2014, as specified by Mackey Furniture Industry	25,518,649.48	4,057,726.58
	Demolition of reinforced concrete structures using compressed air equipment, except floors	258.38	699,201.19
11	Aluminum frames for Maracanã as proposed by Itefal, and additional frames	11,920,282.97	1,935,661.95
	Precast reinforced concrete including manufacturing and on-site installation, for grandstand of Maracanã	11,771,177.22	3,605,433.06
	PA system (internal and grandstand areas) for Maracanã	11,512,097.16	3,748,888.14
	High resistance Monolithic floor polyurethane flakes 3000	113.28	592,363.60
		Total	41,601,247.32

Source: Brazilian Court of Accounts, TCU

TCU's ABC Curve analysis identified overpricing in seven items that had 11 or 25 as the first two digits of the unit costs; the total overpricing for these services was R$ 41,601,247.32 (US$18,401,933.18 according to the August 1st, 2014 exchange rate). Note that the item "Tension Roof System" presented the highest overpricing of the ABC Curve and also represented the most expensive service (R$ 256,714,917.00, about US$113,555,507). Note that the first digits 25 where singled out by both the two-digit test and the summation test. TCU found a total overpricing of R$ 149,972,318.01 or US$66,338,890. The R$ 41,601,247.32 amount corresponds to 27.74% of total overpricing.

Consider next the other digits: 16, 17, 19, 21, 22, 32, 48. The corresponding items are described in Table 5.

TCU's ABC Curve analysis identified 10 additional overpriced items associated to the suggested digits. Only the first two digits 18 and 20 did not appear associated with overpricing. However, services 18052222-6 – "System information boards, displays (...)" and 18052259-6 – "Transformers, generators, UPS and (...)", which unit costs were respectively R$ 20,206,546.09 and R$ 18,600,382.98, were not analyzed by TCU. Therefore, we cannot clearly assess if they were overpriced.

The overpricing calculated for these items was R$ 65,692,812.51 (US$29,058,617), which represented 43.8% of total overpricing.

The sum of the overpricing found in the two-digit categories identified by the application of NB Law amounts to R$ 107,294,059.83, which corresponds 71.54% of total overpricing determined by the TCU.

Table 5 – Confrontation between the results of NB Law tests and TCU's analysis for the two-digits 16, 17, 19, 21, 22, 32 and 48

Digits	Service	Unit Cost	Overpricing TCU
16	Lining electro-metal grill 30x100mm, h = 20 cm, with main bar 20x2mm, carbon steel SAE 1006/1020	163.52	1,514,293.41
17	System of restricted access control and of public access control to Maracanã.	17,544,505.41	12,370,588.70
	Soil drilling for application root cutting, diam.410 mm.	177.96	373,048.34
19	Services (cables, software installation and interconnection, commissioning, startup and assisted operations)	19,081,957.09	4,241,578.74
	Floor covering in polar white granite polished slabs (60x60) cm, and c = 2.00 cm, fixed with cement grout on cement mortar, sand and gravel, trace 1:2:2 and white cement grout	196.72	435,709.44
	Drainage, top soil, natural and synthetic grass and irrigation system for the lawn of Maracanã Stadium - World Cup 2014, as proposed by the Campanelli Company	1,919,536.42	722,897.49
21	Demolition of reinforced concrete cover over the bleachers of Maracanã (sunroom), as in budget worksheet 002/2009-EMOP	21,439,722.66	13,464,684.58
22	Recovery of Maracanã's structure as in spreadsheet 004/2011 - EMOP	22,904,212.57	10,839,664.36
32	Local administration to works of renovation with changes and additions to the Maracanã Stadium as in spreadsheet No. 003/2011 EMOP	32,863,882.27	12,563,282.35
48	Central air conditioning and full mechanical ventilation System to the Maracanã Stadium, as proposed in PR-078 620 (Ambienter) supply and installation	48,844,340.89	9,167,065.10
	Total	R$ 65,692,812.51	

Source: Brazilian Court of Accounts, TCU

4. Conclusion

The present research tested the application of Newcomb-Benford's Law to the unit costs of the budget worksheet of Maracanã soccer arena's renovation, as a tool for data mining. It applied the First-Two Digits Test and the Summation Test, all based of Benford's Law, using the *Z*-statistic, Chi-Square and Mean Absolute Deviation tests with mixed results.

In the analysis of unit costs for the two-digits' categories, the first two digits 11 and 25 were detected in the First-Two Digits Test and the Summation Test. These consisted of seven items also found by TCU auditors, including a service with the highest overpricing, the "Tensioned Roofing System". The overpricing identified in these items totaled R$ 41,601,247.32 and represented 27.74% of total overpricing uncovered by TCU.

In addition, the Summation Test alone identified 10 other services also uncovered by TCU to have values above market benchmarks. The total overpricing of all items detected by NB Law was R$ 107,294,059.83, representing 71.54% of the total overpricing identified by TCU (R$ 149,972,318.01).

The present work represents a first step towards a practical implementation of NB Law to the audit of public works. We attempt to illustrate the applicability of this tool to budget worksheets for future research into developing effective methods to help selecting the items to be audited using the NB Law.

It is noteworthy to discuss the possible future of data manipulation in Brazilian procurements. Indeed, Newcomb-Benford's Law has not yet being applied as a regular, complementary tool by the Brazilian Court of Accounts, which, in part, may explain the highly significant result found in the present analysis. However, if this tool becomes standard, bidding firms might become aware of it and may devise more sophisticated ways to overprice their bids in other to try to avoid detection. In that case, new tests might be needed in order to effectively curb overpricing in procurements.

References

Benford, F. (1938) "The law of anomalous numbers". *Proceedings of the American Philosophical Society* **78** (4), 551-572.

Berton, L. (1995) "He's got their number: scholar uses math to foil financial fraud". *Wall Street Journal.* 10, B1, July 10.

European Commission (2010) "Report on Greek Government Deficit and Debt Statistics". Available at <http://epp.eurostat.ec.europa.eu/portal/page/portal/product_details/publication?p_product_code=COM_2010_report_greek>. Accessed on July 10, 2014.

Hill, T.P. (1995) "Base-Invariance Implies Benford's Law". *Proc. Amer. Math. Soc.* **123**, 887-895.

Mebane, W. R. (2006) "Election Forensics: vote counts and Benford's Law". Summer Meeting of the Political Methodology Society. Papers, Posters and Syllabi. Nº 620.

______ (2010) "Fraud in the 2009 presidential election in Iran?" *Chance* **23**(1), 6-15.

______ (2009) "Note on the presidential election In Iran". Updated notes on author's website. Available at http://www.benfordonline.net/fullreference/756. Accessed July 10, 2014.

Newcomb, S. (1881) "Note on the frequency of the different digits in natural numbers". *The American Journal of Mathematics* **4**(1), 39-40.

Nigrini, M. J. (1992) *The Detection of Income Tax Evasion Through an Analysis of Digital Frequencies*. Ph.D. thesis. Cincinnati, OH: University of Cincinnati.

______ (2000) *Digital analysis using Benford's Law: Tests Statistics for Auditors*. Global Audit Publications: Vancouver.

______ (2012) *Benford's Law. Applications for Forensic Accounting Auditing, and Fraud Detection*. John Wiley & Sons, Inc.: Hoboken, New Jersey.

Rauch, B.; G. Brähler and M. Göttsche (2011) "Fact and Fiction in EU-Governmental Economic Data". *German Economic Review* **12**(3), 243-255.