IMF Working Paper

Benford's Law and Macroeconomic Data Quality

Jesus Gonzalez-Garcia and Gonzalo Pastor

INTERNATIONAL MONETARY FUND

WP/09/10

IMF Working Paper

Statistics Department

Benford's Law and Macroeconomic Data Quality

Prepared by Jesus Gonzalez-Garcia and Gonzalo Pastor♦

Authorized for distribution by J.R. Rosales

January 2009

Abstract

This Working Paper should not be reported as representing the views of the IMF.
The views expressed in this Working Paper are those of the authors and do not necessarily represent those of the IMF or IMF policy. Working Papers describe research in progress by the authors and are published to elicit comments and to further debate.

This paper examines the usefulness of testing the conformity of macroeconomic data with Benford's law as indicator of data quality. Most of the macroeconomic data series tested conform with Benford's law. However, questions emerge on the reliability of such tests as indicators of data quality once conformity with Benford's law is contrasted with the data quality ratings included in the data module of the Reports on the Observance of Standards and Codes (data ROSCs). Furthermore, the analysis shows that rejection of Benford's law may be unrelated to the quality of statistics, and instead may result from marked structural shifts in the data series. Hence, nonconformity with Benford's law should not be interpreted as a reliable indication of poor quality in macroeconomic data.

Contents

Page

I. INTRODUCTION

Data and information are the basis for decision making. Recent events underscore the crucial importance of the availability of reliable, timely and accurate economic and financial data to monitor and assess market conditions and the adequacy of policies. Notable examples include the financial crisis that erupted in Mexico in December 1994, the series of financial crises in East Asia and Russia in 1997–98, the corporate bankruptcies related to governance failures in the U.S. in 2000–01, and the ongoing financial crisis in the U.S. and its reverberations in global financial markets.

In recent decades, concerns about the quality of macroeconomic data have prompted significant efforts to encourage countries to observe international best practices. These efforts have been articulated within the Standards and Codes Initiative, in which several international standard-setting bodies participate; for instance, the Financial Stability Forum, the International Monetary Fund (IMF), the World Bank, and the Basel Committee on Banking Supervision. On dissemination, the IMF's Data Dissemination Initiative launched in the mid-1990s has aimed at strengthening transparency by establishing internationally-recognized data dissemination standards and codes. Concerning data quality, the IMF's Data Quality Assessment Framework (DQAF) has provided a structure for systematic data quality analysis in which experts, using best international practices as benchmarks, evaluate the statistical practices in a particular country. The results from these assessments are published in the Data Dissemination Reports on the Observance of Standard and Codes (also called ROSC data modules or data ROSCs). As of July 2008, data ROSCs (and updates) containing data quality ratings have been published for some 80 IMF member countries.[1] [2]

This paper examines the usefulness of Benford's law to ascertain overall data quality. To this end, it revisits and expands the work by Nye and Moul (2007) who tested certain data for conformity with Benford's law, focusing on the gross domestic product (GDP) series for members of the Organization for Economic Cooperation and Development (OECD) and African countries. Benford's law, also called the first digit law, refers to an empirical regularity commonly found in sets of data describing naturally-occurring phenomena. The regularity observed is that the frequency of the digits appearing as the first or leading digit in the numbers collected declines as the value of the digits increase. As a result, 1 is the digit more often observed and 9 is the least frequent. That is, not all digits have the same chance of appearing as the first digit in the numbers collected. The basis for Nye and Moul's analysis was the expectation that the first digits of good quality macroeconomic data should follow Benford's law. Otherwise, they argued, further research on the quality of the data series is warranted.

---

[1] Information on the Standards and Codes and Data Dissemination Standards, respectively, can be found in the following web pages: http://www.imf.org/external/standards/index.htm and http://dsbb.imf.org/Applications/web/dsbbhome/. A detailed description of the DQAF is in http://dsbb.imf.org/Applications/web/dqrs/dqrsdqaf/. The ROSC data modules can be found in: http://.imf.org/esternal/np/rosc/rosc.asp.

[2] A review of the Data Dissemination Initiative can be found in Alexander, Cady, and Gonzalez-Garcia (2008).

Nye and Moul analyzed six versions of GDP series for OECD and African countries: nominal and real GDP in domestic currency, nominal and real GDP in U.S. dollars, and nominal and real GDP in U.S. dollars adjusted for purchasing power parity. They found that domestic currency nominal GDP series for OECD and African countries, as well as the GDP series in current U.S. dollars for African countries and the GDP series at constant prices for OECD countries conformed with Benford's law. The remaining eight data series did not conform with Benford's law, leading the authors to conclude that more research was needed on the quality of the data and/or on the source data transformations used to generate the series.

This paper broadens the work described above in three ways. First, it widens the set of macroeconomic data analyzed to cover the main sectors of the economy. Second, the number of countries is enlarged to cover 80 countries (nested in six country groups) for which data ROSCs were available as of mid-2008. Third, the results of the goodness of fit tests used to assess conformity of the data with Benford's law are compared with the data quality ratings in the data ROSCs in order to assess the utility of the former as a data quality metric.

The paper concludes that, in general, the frequency distribution of first digits of macroeconomic data tends to conform with the distribution implied by Benford's law. However, interpreting the rejection of the first digit law as an indicator of poor data quality is not supported by a closer examination of the results. A solid pattern of consistency between the results of goodness of fit tests and the data quality ratings in data ROSC cannot be established. Also, the analysis shows that deviations from Benford's law may be unrelated to issues of data quality and may instead result from marked economic fluctuations and structural shifts/breaks captured in the data series. Hence, non conformity with Benford's law should not be interpreted as signaling poor quality in macroeconomic data.

The rest of the paper is structured as follows. Section II describes Benford's law and the goodness of fit test used to check whether the macroeconomic data conforms with the frequency distribution associated with Benford's law. Section III shows the results of the goodness of fit tests conducted for an ample array of datasets obtained from the *International Financial Statistics* (*IFS*). Section IV compares the results of goodness of fit tests with the assessments in the data ROSCs. Sections V and VI caution against using the goodness of fit tests as indicators of data quality and offer concluding remarks.

II. BENFORD'S LAW AND TESTING METHODOLOGY

A. Benford's Law

The Benford's law, or the first digit law, refers to an empirical regularity named after Frank Benford (1938), who documented that the first or leading digits of numbers in many datasets follow a frequency distribution with the following characteristics: digit 1 occurs more often than any other digit and the frequency of the remaining digits (2 to 9) decreases as the value of the digits increases. The distribution of leading digits implied by Benford's law is described by the following expressions:

$$f(i) = \log_{10}(1 + i^{-1}); \quad \text{and} \quad \sum_{i=1}^{9} f(i) = 1$$

where $f(i)$ is the frequency of digit i being the first or leading digit, $\log_{10}$ denotes a base 10 logarithm, and $i = 1, \dots, 9$.[3]

This empirical regularity has elicited considerable interest, and even some fascination, for the following reasons. First, Benford's law contradicts what seems to be a natural intuition suggesting that any digit has the same probability of appearing as the first digit in the numbers collected. In other words, that digits have a uniform distribution when appearing as leading digits in numbers. Second, observance of Benford's law has been confirmed in several datasets, as noted below. Third, the statistical literature has taken a long time to provide solid formal explanations for the often observed presence of Benford's law (see Hill and Schurger (2005), Hill (1995), and Raimi (1976)).

A non-technical illustration of the presence of Benford's law in sets of series that grow geometrically can be outlined as follows (see Browne (1998)). Assuming random pairs of growth rates and initial values, at some point in time, each series will show figures having one as the first digit and will grow for a certain number of periods until reaching figures having two as the first digit. The number of times in which the digit one appears as the leading digit will be greater than the number of times figures starting with digit two occur. The reason is that in the latter case the absolute increments are larger than in the former, and the next leading digit (i.e., digit three in this case) is reached at an earlier stage. Hence, as the value of the first digit increases, the number of periods necessary to reach numbers starting with the next digit decreases. As a result, the digit 1 will appear more often as the leading digit, and 9 will be the least frequent first digit.[4]

Benford's law has been used to study different hypotheses about datasets. For example, Nigrini (1996 and 1999) and Nigrini and Mittermaier (1997) used Benford's law to detect fraudulent data in tax payments and accounting data. Hales, Sridharan, Radhakrishnan, Chakravorty, and Siha (2008) studied the reliability of employee-reported operational data by testing the presence of Benford's law in those datasets; and Diekmann (2007) analyzed whether the first digits of regression coefficients published in scientific literature tend to be distributed according to the first digit-law. In economics, Varian (1972) suggested that Benford's law could help detect anomalous information in datasets used for public planning decisions and Ley (1996) verified that the first digits of one-day returns of stock indexes'

---

[3] The resulting probabilities are (digit/probability): 1/0.3010; 2/0.1761; 3/0.1249; 4/0.0969; 5/0.0792; 6/0.0669; 7/0.0580; 8/0.0512; and 9/0.0458.

[4] It is much more difficult to distill simple illustrations of other well established results in the literature indicating that Benford's law can be observed also in sets of numbers that are constructed using random drawings from randomly chosen distributions or in other sophisticated sets of numbers. See for instance, Berger, Bunimovich and Hill (2005); and Janvrese and De la Rue (2004).

follow the first digit law. More recently, Judge and Schechter (2007) used Benford's law to detect problems in survey data.

B. Testing Methodology

Goodness of fit tests are used to compare the distribution of the first digits in different sets of macroeconomic data with that implied by Benford's law. Specifically, the Pearson's chi-square (χ^2) test statistic is used, which is written as follows:

$$\chi^2_{(8)} = N \times \sum_{i=1}^{9} \frac{(\hat{\theta}_i - f(i))^2}{f(i)}$$

where N is the number of observations, $\hat{\theta}_i$ is the observed frequency of digit i, and $f(i)$ is the frequency of digit i implied by Benford's law. Under the null hypothesis, the first digits in the set of macroeconomic data follow Benford's law. A rejection of the null hypothesis implies that the data do not follow the Benford distribution, thus suggesting that the data may not represent accurately the true behavior of the economic variables measured. The core idea of this paper is that such an interpretation is not necessarily warranted and that its validity should be tested by contrasting results with other data quality metrics.

As noted above, this paper applies goodness of fit test to a large dataset in order to cover a wider set of macroeconomic statistics for a larger set of countries. The dataset includes annual time series for the period 1945–2007 covering variables for the following sectors: banking, balance of payments, national accounts, and government finance statistics.[5] Countries were classified in six different groups in line with the April 2008 World Economic Outlook (WEO) country classification: Advanced, African, Central and Eastern Europe, Commonwealth of Independent States and Mongolia, Developing Asia and the Middle East, and Western Hemisphere. The dataset includes 80 countries for which the data ROSCs published in the IMF website as of end-July 2008 show data quality ratings.[6,7]

III. CONFORMITY OF MACROECONOMIC DATA WITH BENFORD'S LAW

For each country group, goodness of fit tests were applied to five variables within each of the four economic sectors mentioned above.[8] The selection of variables for each economic sector

---

[5] The samples used are unbalanced panels as the number of observations over time varies across countries. The number of observations used in the goodness of fit tests is satisfactory, ranging in the hundreds of data points.

[6] See Table A.1 in the Appendix for the lists of countries in each group. Developing Asia and Middle East are two different groups in the WEO, but they are combined in this paper, otherwise too few observations are available for those groups.

[7] The ratings of the ROSC data modules can be found in: http://www.imf.org/external/np/rosc/rosc.asp.

[8] See Table A.2 in the Appendix for the lists of macroeconomic aggregates considered.

was based on data availability and relevance for economic analysis. In general, the tested data series are aggregate concepts considered relevant for economic surveillance. For each variable, the test is applied to a dataset containing the pooled series for the countries in each group. The data series are used as reported in the *IFS*, that is, they are denominated in the country's national currency, except in the case of balance of payments where the series are reported in U.S dollars as in the *IFS*. The aim is exclusively to test whether the *first digits* of those numbers (independently on their sign and size) conform with Benford's law. Using higher frequency data to increase the number of observations would not provide more information on the frequency distribution of the leading digits, as the more frequent data points would span across the same orders of magnitude as the lower frequency data series.

The results from the goodness of fit tests for each aggregate and country group—presented in Tables 1–4—show that macroeconomic data conform well with Benford's law. Using a significance level of 5 percent as benchmark, only 12 out of the 120 tests reject the frequency distribution of first digits implied by Benford's law. The Africa group shows the most cases of rejection (eight cases), followed by the Western Hemisphere group (two cases) and the Advanced countries and Developing Asia and Middle East groups (one rejection each).

Across sectors, the banking survey and balance of payments statistics show the less rejections followed by national accounts and government finance statistics. More specifically, it is found that the banking survey statistics conform very well with Benford's law, only one aggregate (i.e., claims on central or general government) in the Africa group rejects Benford's law (Table 1). Balance of payments statistics data conform well with Benford's law, with only two aggregates (the current and capital account balances) rejecting conformity in the African countries group (Table 2). Concerning national accounts, three aggregates (i.e., nominal GDP, household consumption, and gross fixed capital formation) do not conform to Benford's law in the Africa group. Surprisingly, Benford's law is also rejected for the nominal GDP data of the advanced economies group (Table 3). For government finance statistics, there are two rejections (i.e., cash receipts and payments, and cash surplus/deficit) in two groups: Africa and Western Hemisphere, and one rejection (i.e., cash surplus/deficit) in the Developing Asia and Middle East group (Table 4).

IV. COMPARISON OF GOODNESS OF FIT TESTS RESULTS WITH ROSC ASSESSMENTS OF DATA QUALITY

In this section, the results of goodness of fit tests are compared with the data quality ratings in data ROSCs. ROSC assessments are prepared by experts using the DQAF, which provides for a systematic evaluation of the degree of adherence of a country's statistical practices to international best practices. The DQAF covers six data quality dimensions. The first dimension reviews the legal and institutional environment supporting the production of data, while the other dimensions include: assurance of integrity, such as staff's professionalism, transparency, and ethical standards; methodological soundness; data accuracy and reliability; serviceability (including data periodicity and timeliness, consistency, and revision policies in place); and data accessibility to users.

The results from the data ROSC assessments are summarized in the form of ratings assigned to each of the DQAF data quality dimensions for the various datasets reviewed. Four different ratings are assigned depending on the degree of adherence to internationally-accepted best practices. An "O" rating indicates that best practices are observed in the country; "LO" means that best practices are largely observed; "LNO" indicates that best practices are largely not observed; and "NO" that best practices are not observed.

Among the various data quality dimensions assessed in ROSCs, the one on data accuracy and reliability is the most relevant for this paper as it focuses on whether the statistical outputs sufficiently portray the reality of the economic concept measured. For this dimension, ratings are assigned for five elements: (i) the adequacy of source data to compile statistics; (ii) the completeness of source data to adequately approximate the economic concepts at hand; (iii) the use of sound statistical and aggregation techniques to deal with data sources; (iv) the availability of data consistency checks to assess and validate intermediate results and statistical outputs; and (v) the existence of regular practices to track data revisions over time.

Table 5 compares data ROSC ratings on data accuracy and reliability with the results of goodness of fit tests to assess the consistency of these two metrics for each country group and economic sector. The first column shows the number of series in which conformity with Benford's law was rejected and the second column summarizes the evaluation of data accuracy and reliability according to the ratings assigned in the data ROSCs. More specifically, the second column shows the number of LNO and NO ratings as a percentage of the total ratings assigned. That is, the share of cases in which international best practices are either not observed or largely not observed; these are called the "non observance ratings" in our analysis. Higher shares of "non observance ratings" indicate less data quality.

In general, there is only one point of consistency between the two sets of results, and a clear pattern of consistency between them cannot be identified. There is country group concurrence between the highest shares of "non observance ratings" and the largest number of rejections of Benford's law, which occurs for all economic sectors in the African country group.[9] However, a striking result is that Benford's law is rejected for the GDP series of advanced countries, which have by far the lowest share of "non observance ratings" in national accounts according to the data ROSCs (only four percent). Also puzzling is the fact that for national accounts statistics, which present very high percentages of "non observance ratings" across the board, conformity with Benford's law is rejected only for the Africa group. Indeed, Benford' law is not rejected for the Asia and the Middle East and Western Hemisphere groups, notwithstanding the weaknesses identified by experts in data accuracy and reliability (the share of "non observance ratings" is about one third in both cases).

---

[9] In the same vein, for cases in which Benford law is not rejected, relatively lower shares of non-observance ratings are registered.

V. CAN TESTS OF CONFORMITY WITH BENFORD'S LAW BE CONSIDERED RELIABLE INDICATORS OF MACROECONOMIC DATA QUALITY ?

The results in section IV show lack of a pattern of consistency between the results of goodness of fit tests and data ROSC assessments, and some contradictions between them. On the latter, the most striking result is the rejection of Benford's law for the national accounts statistics of advanced countries, despite the high standards of data compilation reported in the data ROSCs, which show very low shares of "non observance ratings".

A graphical presentation of the GDP series of advanced countries shows that almost all of them follow trajectories resembling geometric growth patterns. However, the series of two countries, Japan and Finland, show marked divergences from that pattern (Figure 1). Japan's nominal GPD data reflect the protracted period of low nominal GDP growth that started in the early 1990s following a long period of very rapid growth. On the other hand, Finland's nominal GDP series show a brief period of output contraction in the aftermath of the financial crisis of 1991. Both economic events are well identified in the economic history of these countries, and the data ROSCs ratings for these countries indicate that best international compilation practices for data accuracy and reliability are indeed closely followed. For Japan, four out of five ratings for data accuracy and reliability suggest that international best practices are observed or largely observed. For Finland all ratings assigned indicate that those practices are at least largely observed.

A revised goodness of fit test that excluded the nominal GDP series for Japan and Finland from the sample of advanced countries sheds some light. The objective was to check whether the structural shifts in the Japanese and Finnish GDP series may have caused the rejection of Benford's law; which indeed turned out to be the case. Excluding these series, the distribution implied by Benford's law cannot be rejected any longer for the advanced country group, which suggests that interpreting a rejection of Benford' law as a signal of low data quality can be misleading at least in certain cases.[10]

An analysis of the national accounts data for the African group also argues against using nonconformity with Benford's law as a reliable indication of low data quality. As noted above, the only point of consistency between the goodness of fit tests and the data ROSC ratings appears in the Africa group, which shows both the highest shares of "non observance ratings" and the largest number of rejections of Benford's law across all data sets analyzed. However, the number of Benford's law rejections in the national accounts statistics in the African group can be reduced from 3 to zero by removing from the sample a country whose data series show marked structural shifts: Niger. [11] Figure 2 shows that Niger's data series for

---

[10] The revised test involving 681 observations showed a test statistic of 8.570 with a corresponding p-value of 0.380.

[11] The details of the tests excluding Niger are as follows. For GDP, the test involves 699 observations, the test statistic is 9.560, and its p-value is 0.297; for household consumption expenditure the test involves 552 observations, the tests statistic is 15.317 and its p-value is 0.053; and for gross fixed capital formation the test statistic was calculated with 558 observations, the test statistic is 8.352 and the p-value is 0.400.

nominal GDP, household consumption expenditure and gross fixed capital formation have recorded structural shifts that led to the rejection of Benford's law.

Closer examination of the results of the goodness of fit tests for the advanced and African country groups show that testing for conformity with Benford's law is far from being a reliable first cut in the analysis of data quality. Specifically, there is evidence that nonobservance of Benford's law may be associated with the presence of structural shifts in the series rather than with poor data quality. These results militate against some propositions in the study of Nye and Moul, who advocate for testing conformity with Benford's law as "a technique for assessing data quality," and suggest that nonconformity with Benford's law is "consistent with deliberate manipulation of the underlying series."

VI. CONCLUSION

This paper has reviewed the reliability of testing conformity of macroeconomic data with Benford's law to assess data quality. It revisited the analysis by Nye and Moul (2007) using a wider set of macroeconomic time series published in the *IFS* for some 80 countries and compared the results with those of the data quality ratings in published data ROSCs.

Most of the reviewed macroeconomic data series tend to show a distribution of first digits consistent with Benford's Law, which confirms the robustness of this empirical regularity. However, interpreting the rejection of Benford's law as a reliable indication of poor data quality is not supported by the analysis of the results. First, it is not possible to find a solid pattern of consistency between the results of goodness of fit tests and data quality ratings in data ROSCs. Second, it is shown that rejections of Benford's law may be unrelated to the quality of statistics. Rather, it is found that economic variables showing marked structural shifts can result in the rejection of Benford's law regardless of the observance of best international statistical practices. Hence, nonconformity with Benford's law should not be interpreted as reliably signaling poor quality of macroeconomic data.

Table 1. Banking Survey Statistics: Goodness of Fit Tests[1,2]
(Data in domestic currencies)

	Advanced	African	Central and Eastern Europe	CIS and Mongolia	Dev. Asia and Middle East	Western Hemisphere
Foreign Assets						
Countries	7	16	11	10	10	18
Observations	398	667	233	144	521	950
χ^2 statistic	8.041	11.500	7.840	7.739	14.324	11.331
p-value	0.429	0.175	0.449	0.459	*0.074*	0.184
Domestic Credit						
Countries	7	16	11	10	10	18
Observations	403	663	233	144	517	951
χ^2 statistic	9.802	12.807	12.886	4.266	7.279	5.733
p-value	0.279	0.119	0.116	0.832	0.507	0.677
Claims on Central or General Govermnent						
Countries	7	16	11	10	10	18
Observations	401	663	233	144	506	950
χ^2 statistic	7.040	16.597	8.466	12.590	15.480	15.319
p-value	0.532	**0.035**	0.389	0.127	*0.050*	*0.053*
Money						
Countries	6	17	11	10	10	18
Observations	346	715	233	144	522	954
χ^2 statistic	14.962	10.609	2.822	5.002	4.713	6.329
p-value	*0.060*	0.225	0.945	0.757	0.788	0.610
Money and Quasi-money						
Countries	7	16	11	10	10	18
Observations	401	668	233	144	520	954
χ^2 statistic	7.581	7.657	8.122	1.970	3.322	12.170
p-value	0.475	0.468	0.422	0.982	0.913	0.144

Source: *IFS* database and authors' calculations.
[1] Under the null hypothesis, the first digits follow Benford's law.
[2] Bold and italic fonts indicate rejection of the null hypothesis at 5 and 10 percent significance levels, respectively.

Table 2. Balance of Payments Statistics: Goodness of Fit Tests[1,2]
(Data in U.S. dollars)

	Advanced	African	Central and Eastern Europe	CIS and Mongolia	Dev. Asia and Middle East	Western Hemisphere
Current Account Balance						
Countries	13	17	11	10	10	14
Observations	469	498	220	135	314	448
χ^2 statistic	9.597	17.628	10.065	6.272	8.565	7.112
p-value	0.294	**0.024**	0.260	0.617	0.380	0.525
Capital Account Balance						
Countries	13	15	11	9	10	12
Observations	411	328	116	102	110	149
χ^2 statistic	8.748	22.142	4.338	3.256	2.463	11.554
p-value	0.364	**0.005**	0.825	0.917	0.963	0.172
Financial Account Balance						
Countries	13	17	11	10	10	14
Observations	472	466	222	138	315	451
χ^2 statistic	7.722	6.510	7.266	5.976	8.481	8.495
p-value	0.461	0.590	0.508	0.650	0.388	0.387
Net Errors and Omissions						
Countries	13	17	11	10	10	14
Observations	474	500	222	138	315	451
χ^2 statistic	7.069	10.358	13.871	10.054	5.840	3.450
p-value	0.529	0.241	*0.085*	0.261	0.665	0.903
Reserves and Related Items						
Countries	13	17	11	10	10	14
Observations	474	499	222	138	315	451
χ^2 statistic	10.715	14.281	4.649	13.089	11.164	9.076
p-value	0.218	*0.075*	0.794	0.109	0.193	0.336

Source: *IFS* database and authors' calculations.
[1] Under the null hypothesis, the first digits follow Benford's law.
[2] Bold and italic fonts indicate rejection of the null hypothesis at 5 and 10 percent significance levels, respectively.

Table 3. National Accounts Statistics: Goodness of Fit Tests[1,2]
(Data in domestic currencies)

	Advanced	African	Central and Eastern Europe	CIS and Mongolia	Dev. Asia and Middle East	Western Hemisphere
Nominal GDP						
Countries	14	17	11	9	10	14
Observations	794	748	233	143	504	726
χ^2 statistic	16.580	17.079	2.773	4.979	3.217	10.042
p-value	**0.035**	**0.029**	0.948	0.760	0.920	0.262
Private Consumption or Household Consumption Expenditure						
Countries	5	14	10	10	10	11
Observations	286	595	203	162	486	628
χ^2 statistic	5.725	24.121	5.625	15.043	8.947	4.585
p-value	0.678	**0.002**	0.689	*0.058*	0.347	0.801
Government Consumption Expenditure						
Countries	8	14	10	9	10	13
Observations	444	595	200	150	497	697
χ^2 statistic	1.882	8.196	12.761	13.877	5.015	10.337
p-value	0.984	0.415	0.120	*0.085*	0.756	0.242
Gross Fixed Capital Formation						
Countries	5	14	10	9	9	11
Observations	288	599	210	148	447	592
χ^2 statistic	8.711	15.962	9.260	3.966	5.312	11.561
p-value	0.367	**0.043**	0.321	0.860	0.724	0.172
Change in Stocks						
Countries	8	15	10	9	9	11
Observations	435	566	205	141	402	570
χ^2 statistic	7.700	3.857	4.872	14.109	4.948	9.942
p-value	0.463	0.870	0.771	*0.079*	0.763	0.269

Source: *IFS* database and authors' calculations.
[1] Under the null hypothesis, the first digits follow Benford's law.
[2] Bold and italic fonts indicate rejection of the null hypothesis at 5 and 10 percent significance levels, respectively.

Table 4. Government Finance Statistics: Goodness of Fit Tests[1,2]
(Data in domestic currencies)

	Advanced	African	Central and Eastern Europe	CIS and Mongolia	Dev. Asia and Middle East	Western Hemisphere
Cash Receipts (operating activities)						
Countries	14	16	8	8	10	14
Observations	492	395	77	75	248	578
χ^2 statistic	4.495	4.660	12.550	3.398	4.948	16.164
p-value	0.810	0.793	0.128	0.907	0.763	**0.040**
Cash Payments (operating activities)						
Countries	14	16	8	8	10	14
Observations	511	383	73	72	256	574
χ^2 statistic	4.110	1.771	14.185	3.435	9.700	17.894
p-value	0.847	0.987	*0.077*	0.904	0.287	**0.022**
Net Cash Inflow (operating activities)[3]						
Countries	6	7	5	6	8	3
Observations	47	52	34	32	96	29
χ^2 statistic	9.273	6.230	10.712	8.459	10.788	6.018
p-value	0.320	0.621	0.219	0.390	0.214	0.645
Net Cash Outflow[3]						
Countries	6	7	6	6	8	3
Observations	47	52	46	32	96	29
χ^2 statistic	2.040	16.539	9.173	13.563	10.165	13.908
p-value	0.980	**0.035**	0.328	*0.094*	0.254	*0.084*
Cash Surplus / Deficit						
Countries	14	16	7	8	10	14
Observations	513	396	62	75	256	577
χ^2 statistic	13.090	16.383	11.770	3.415	24.707	8.251
p-value	0.109	**0.037**	0.162	0.906	**0.002**	0.409

Source: *IFS* database and authors' calculations.
[1] Under the null hypothesis, the first digits follow Benford's law.
[2] Bold and italic fonts indicate rejection of the null hypothesis at 5 and 10 percent significance levels, respectively.
[3] In the case of the net cash inflow and net cash outflow series, samples sizes were small. Therefore, these tests should be taken with some caution.

Table 5. Rejections of Benford's Law and Shares of LNO and NO Data ROSC Ratings
(per economic sector and country group)

Country Groups	Rejections of Benford's Law	Shares of LNO and NO Ratings (in percent) [1]
	Number of series	Data Accuracy and Reliability
Banking Survey Statistics		
Advanced	0	0
African	1	4
Central and Eastern Europe	0	0
CIS and Mongolia	0	4
Dev. Asia and Middle East	0	2
Western Hemisphere	0	1
Balance of Payments Statistics		
Advanced	0	2
African	2	23
Central and Eastern Europe	0	9
CIS and Mongolia	0	10
Dev. Asia and Middle East	0	8
Western Hemisphere	0	7
National Accounts Statistics		
Advanced	1	4
African	3	44
Central and Eastern Europe	0	11
CIS and Mongolia	0	13
Dev. Asia and Middle East	0	27
Western Hemisphere	0	33
Government Finance Statistics		
Advanced	0	2
African	2	16
Central and Eastern Europe	0	6
CIS and Mongolia	0	4
Dev. Asia and Middle East	1	4
Western Hemisphere	2	14

Source: Authors' calculations.
[1] Number of LNO and NO ratings as percent of total number of ratings assigned for the whole country group.

APPENDIX : COUNTRY GROUPS AND DATA SERIES USED

Table A1 shows the groups of countries. The actual groups used for the tests may be smaller due to lack of data in the *IFS* or because some statistical sector was not assessed in the data ROSC.

Table A1. Groups of Countries.

Advanced	Central and Eastern Europe	Western Hemisphere
Canada		
Finland	Albania	Bolivia
France	Bosnia & Herzegovina	Chile
Germany	Bulgaria	Colombia
Greece	Estonia	Costa Rica
Iceland	Hungary	Dominica
Israel	Latvia	Dominican Republic
Italy	Lithuania	Ecuador
Japan	Macedonia, FYR	El Salvador
Korea	Poland	Grenada
Malta	Slovak Republic	Guatemala
Netherlands	Turkey	Honduras
Norway	**Commonwealth of**	Mexico
Sweden	**Independent States**	Nicaragua
Africa	**and Mongolia**	Panama
Botswana	Azerbaijan, Rep. of	Paraguay
Burkina Faso	Belarus	Peru
Chad	Georgia	St. Kitts and Nevis
Gambia, The	Kazakhstan	St. Vincent and the
Kenya	Kyrgyz Republic	Grenadines
Malawi	Moldova	
Mauritius	Mongolia	
Morocco	Russia	
Mozambique	Tajikistan	
Namibia	Ukraine	
Niger	**Developing Asia**	
Senegal	**and Middle East**	
South Africa	Bangladesh	
Tanzania	India	
Tunisia	Indonesia	
Uganda	Pakistan	
Zambia	Philippines	
	Sri Lanka	
	Thailand	
	Egypt	
	Jordan	
	Oman	

Source: Based on the classification of the April 2008 World Economic Outlook.

The statistical sectors covered are: balance of payments statistics, banking survey statistics, national accounts and government finance statistics. For each variable, the available annual data for the period 1945-2006 was obtained from the *IFS* database. Table A.2 shows the variables of each sector used in the analysis.

Table A2. Data Series Tested.

Banking Survey Statistics (domestic currency)

Foreign Assets
Domestic Credit
Claims on Central or General Government
Money
Money and Quasi-money

Balance of Payments Statistics (U.S. dollars)

Current Account Balance
Capital Account Balance
Financial Account Balance
Errors and Omissions (net)
Reserves and Related Items

National Accounts Statistics (domestic currency)

Nominal GDP
Private Consumption or Household Consumption Expenditure
Government Consumption Expenditure
Gross Fixed Capital Formation
Change in Stocks

Government Finance Statistics (domestic currency)

Cash Receipts (operating activities)
Cash Payments (operating activities)
Net Cash Inflow (operating activities)
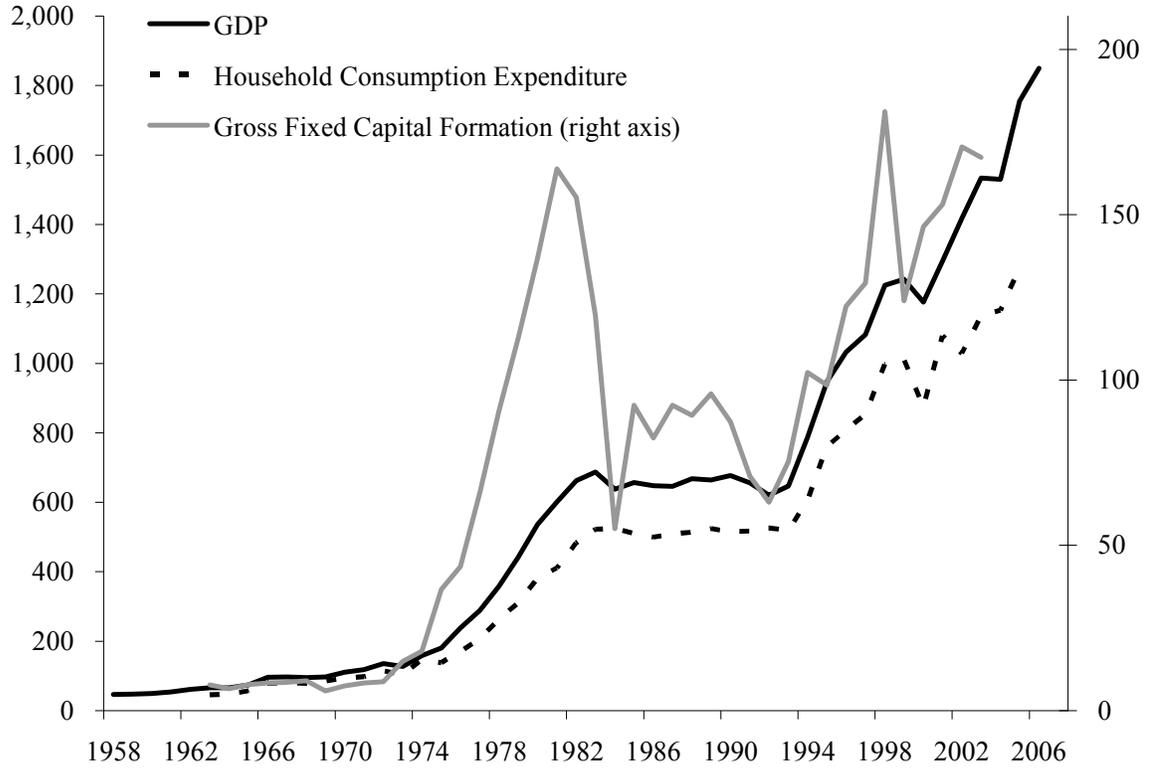Net Cash Outflow
Cash Surplus/Deficit

For detailed data definitions see the *IFS*.

Figure 1. Finland and Japan Nominal GDP
(domestic currencies, billions)
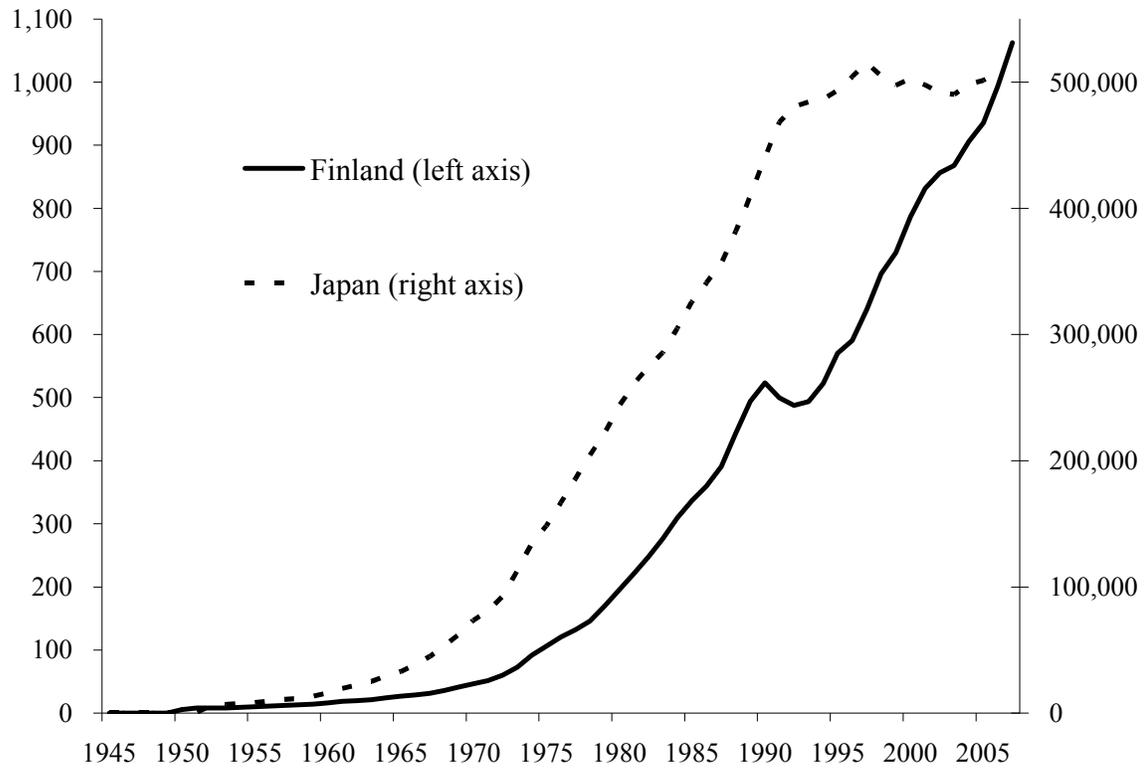
Source: *IFS*.

Figure 2. Niger Nominal GDP, Household Consumption Expenditure and Gross Fixed
Capital Formation
(domestic currency, billions)

Source: *IFS*.

References

Alexander, William, John Cady, and Jesus Gonzalez-Garcia (editors) (2008), *The IMF's Data Dissemination Initiative After 10 Years*, Washington D.C., International Monetary Fund, 130 pages.

Benford, F. (1938), "The Law of Anomalous Numbers," *Proceedings of the American Philosophical Society*, 78: pp. 551-72.

Berger, Arno, Bunimovich, Leonid, and Hill Theodore, 2005, "One-dimensional Dynamical Systems and Benford's Law," *Transactions of the American Mathematical Society*, Vol. 357, pp. 197-219.

Browne, Malcolm, W., 1998, "Following Benford's Law, or Looking Out for No. 1," *The New York Times*, Tuesday, August 4, 1998.

Diekmann, Andreas, 2007, "Not the First Digit! Using Benford's Law to Detect Fraudulent Scientific Data," *Journal of Applied Statistics*, Vol. 34, Issue 3, pp. 321-29.

Hales, Douglas; Sridharan, V.; Radhakrishnan, Abirami; Chakravorty, Satya; and Siha, Samia, 2006, "Testing the Accuracy of Employee-reported Data: An Inexpensive Alternative Approach to Traditional Methods," *European Journal of Operational Research*, Vol. 189, Issue 3, pp. 583-93.

Hill, Theodore, 1995, "A Statistical Derivation of the Significant-digit Law," *Statistical Science*, Vol., 10, No. 4, pp. 354-363.

Hill, Theodore, and Schurger, Claus, 2005, "Regularity of Digits and Significant Digits of Random Variables," *Stochastic Processes and their Applications*, Vol. 15, Issue 10, pp. 1723-43.

Janvresse, Elise, and De la Rue, Thierry, 2004, "From Uniform Distributions to Benford's Law," *Journal of Applied Probability*, Vol. 41, pp. 1203-10.

Judge, George and Schechter, Laura, 2007, "Detecting Problems in Survey Data using Benford's Law", forthcoming, Journal of Human Resources.

Ley, Eduardo, 1996, "On the Peculiar Distribution of the U.S. Stock Indexes' Digits," *The American Statistician*, Vol. 50, No. 4, pp. 311-13.

Nigrini, M.J., 1996, "A Taxpayer Compliance Application of the Benford's Law," *The Journal of the American Taxation Association*, Vol. 18, pp. 72-81.

Nigrini, M.J., 1999, "I've Got Your Number," *Journal of Accountancy*, May, pp. 79-83.

Nigrini, M.J. and Mittermaier, L.J., 1997, "The Use of Benford's Law as an Aid in Analytical Procedures," *Auditing: A Journal of Practice and Theory*, Vol. 16, pp. 52-67.

Nye, John and Moul, Charles, 2007, "The Political Economy of Numbers: On the Application of Benford's Law to International Macroeconomic Statistics," *The Berkeley Electronic Journal of Macroeconomics*, Vol. 7, Issue 1 (Topics), Article 17 http://www.bepress.com/bejm/vol7/iss1/art17/.

Raimi, Ralph, 1976, "The First Digit Problem," *American Mathematical Monthly*, Vol. 83, pp. 521-38.

Varian, Hal, 1972, "Benford's Law," *The American Statistician*, Vol. 26, No. 3, pp. 65-6.