Do countries falsify economic data strategically? Some evidence that they do.

Gilles Stoltz, Tomasz Michalski

DO COUNTRIES FALSIFY ECONOMIC DATA STRATEGICALLY? SOME EVIDENCE THAT THEY DO.*

Tomasz Michalski

HEC Paris
Jouy-en-Josas, France

michalski@hec.fr

Gilles Stoltz

Ecole Normale Supérieure,† CNRS
Paris, France
&
HEC Paris, CNRS,
Jouy-en-Josas, France

stoltz@math.ens.fr

May 8, 2010

Abstract

We find evidence supporting the hypothesis that countries at times misreport their economic data in a strategic manner. Among those suspected are countries with fixed exchange rate regimes, high negative net foreign asset positions or negative current account balances, which corroborates the intuition developed with a simple economic model. We also find that countries with bad institutional quality rankings and those in Africa, Middle East, Eastern Europe and Latin America release economic data of questionable veracity. Our evidence calls for models with public signals to consider strategic misinformation and for establishing independent statistical agencies to assure the delivery of high quality economic data.

JEL codes: F31, F32, F33, F34, D82
Keywords: capital flows, public information provision, misinformation, Benford's law, transparency

"I almost died when I had to pretend for 1 1/2 years [...] as if we were governing. Instead, we lied in the morning, we lied in the evening. [...] It was absolutely clear that what we were saying was not true... And all this time we have done nothing for four years. Nothing. [...] No country in Europe has acted as brazenly as we did. [...] The moment of truth will come swiftly. It was divine intervention, the abundance of cash in the world economy and hundreds of tricks – you obviously don't need to know about – which helped us survive so far."
— Hungarian Prime Minister Ferenc Gyurcsany in a speech to his party's MPs on 26 May, 2006.

"Our first order of business is transparency everywhere. We are recording the real situation in cooperation with the Bank of Greece. To put a final end to the obscurity of finances, the statistics service will become truly independent."
— Greek Prime Minister George Papandreou's speech in the Greek parliament, 16 October 2009.

*We would like to thank Evren Ors, Romain Rancière, Tristan Tomala and the participants at the ESSEC-HEC-INSEAD-PSE Workshop on Financial Economics for helpful comments. All remaining errors are ours.

†This research was carried out within the INRIA project CLASSIC hosted by Ecole normale supérieure and CNRS.

1 Introduction

Conventional wisdom has it that governments may lie strategically to the public about economic data which they collect and provide. Some anecdotal examples may be the accusations of Greece or Italy of tinkering with their budget deficit figures before joining the Eurozone. Argentina has been suspected of understating inflation figures since mid–2007. The Hungarian government, according to its prime minister in a statement that leaked out, lied to the general public about the state of the economy to win the elections in 2006. China is believed to embellish its GDP growth numbers. Even the United States came under scrutiny of the market after GDP growth revisions were consistently negative in the crisis that started in 2008.[1] At times governments are caught red-handed (Hungary), but most of the time it is simply unclear whether the data that is provided to the public is just inaccurate (because of, say, measurement errors or bad data collection methods) or suffers from deliberate alterations. In some cases, misinforming economic agents may bring tangible (possibly short-term) gains for a government; for example, Argentina by misstating inflation figures avoided paying out higher interest on government bonds indexed to inflation (which constituted in fact a partial default) and raising the wages in the public sector. Greece enjoyed lower borrowing rates (close to Germany's) on its government debt because of its Eurozone membership and because investors did not know the entire extent of Greek budget troubles.

In this paper we use a statistical test based on the distribution of first digits of economic accounting data (also known as Benford's law) to test whether countries falsify the economic figures that they report or not.[2] Benford's distribution of the first digits arises naturally for processes with exponential growth; that is, in applications, for many economic data due to inflation or economic growth. It is preserved under multiplication by a common factor, so will survive conversions into different currencies. Also, this distribution arises when many different economic processes are considered together. For these reasons, economic data stemming from many countries should adhere to Benford's law, and any deviations from Benford's distribution may flag data reporting irregularities. Before putting our data to test, we develop a simple model that gives us insights about why and when countries would strategically misinform investors (economic agents). Our story is simple: a country may want to hide the state of the world to prevent capital outflows or incite inflows. Then, we group countries in different relevant categories and investigate whether we can reject the hypothesis that the distribution of the first digits of data that they provide complies with Benford's distribution. Using balance of payments data we find evidence that countries with fixed exchange rate regimes, those with high negative net foreign asset positions or negative current account balances misreport economic data. On top of that, countries with bad institutional quality scores or those in Africa, Middle East, Eastern Europe and Latin America (at least the ones in our sample) have unusual distributions of first digits in the reported accounts that raise suspicions that these countries may be altering data that they report. There is also evidence that some countries misinformed investors in the wake of the Asian financial crisis in 1997.

These interesting patterns are observed for many countries grouped based on sharing similar characteristics. For example, if a country changes its exchange rate regime within the range of our sample, its data will enter different categories. Therefore, the rejections of the null hypothesis of conformity to Benford's law are not country specific, but rather category specific. Relying on our evidence supporting our economic model we conclude that countries, at times, falsify economic data and do it strategically, when it suits their interest. This calls for the inclusion of strategic information provision by governments in models where public signals play an important role. We doubt that the rejection of Benford's distribution for the first digits of the data for many groups of countries is due to just poor data collection and noise in the data in the less developed countries.[3] Large positions in the errors and omissions section of the balance of payments data may indicate,

[1] For Greece's actions before joining the Euro zone, see The Wall Street Journal (2004). For the more recent account of releasing dodgy budget statistics, see the Financial Times (2010). For Italy, see the Financial Times (2001). Argentina's story is described for example in the Financial Times (2008) and The Wall Street Journal (2008). The Hungarian case is treated in the Financial Times (2007). For accounts for China and the USA see the Financial Times (2009).

[2] An alternative test would be to test whether the last digits of the data follow a uniform distribution. But we do not have the data on these as many figures are rounded to their thousands. With more data, we could also apply our test to the first two digits of the data.

[3] This was the conclusion of Nye and Moul (2007) that analyzed Penn World Tables data.

of course, that the data collection process in a given country is of low quality or that there is a great amount of illicit or unrecorded transactions; but if there are no manipulations involved, the first digits of the series (even incompletely registering transactions) should still adhere to Benford's law. Also, because we find evidence for countries which would appear (at least in the short run, according to our model) to benefit from misinforming investors we do not think that the data suffers from rounding errors; if it were so, we would find rejections of the null hypothesis all across the board.

In our tests we use quarterly balance of payments data from the International Monetary Fund for years 1989–2007, for several reasons. Firstly, this is accounting data that is available for many countries from the same source. Secondly, the data is somewhat standardized as it should be prepared (categorized) by providers according to an IMF manual. It offers a lot of datapoints relatively to other data on economic fundamentals. Also, it is not transformed many times (and if the balance is recorded in U.S. dollars it is not transformed at all) which could worsen its statistical properties.[4] Moreover, as the data is quarterly, it includes series that are closely watched by investors so that countries may have an incentive to misrepresent them.

The tests we employ are used by accounting practitioners to detect fraud in company accounts (see Nigrini [1999] for examples). Varian (1972) suggested to use Benford's Law to test the naturalness of the data and adequacy of the numbers produced by forecasting models. These tests, to our knowledge, were not used widely up to date in international (macroeconomic) comparisons (an exception is Nye and Moul [2007]); there is typically a lack of enough suitable data to boost the statistical power of such tests that render these methods impractical and until recently (for example at the time of writing of Varian [1972]) the theoretical processes generating this distribution were not well understood. Unfortunately, our methods do not allow us to detect which country at what point provided false information. Also, we cannot claim how grave the possible infractions were; we cannot exclude that these would be just some price tag effects, though the latter are likely to appear rather in nominal currency rather than in U.S. dollars. For example, a country might report instead of a deficit on the trade balance of 2.01 billion in local currency a one that is slightly lower, say at 1.99 billion. Our task is complicated by the fact that countries may not cheat all of the time (and serial cheaters might be found out quickly); detecting fraud in one or two quarters when many quarters of the data are present is then unlikely.

The only study that extensively uses Benford's distribution to inspect macroeconomic data is that of Nye and Moul (2007). These authors provide evidence (and also simulations of economic quantities) that widely used international macroeconomic data such as the GDP series in the Penn World Tables exhibits first digits that do in general conform to Benford's distribution. The data, however, does not conform to Benford's law for non-OECD countries, posing questions mainly about data quality; the alteration of data might come from the source (falsification or simply from an inefficient data collection infrastructure in the considered countries) or from transformations like conversion of the nominal values into U.S. dollars. As such, however, they do not provide any model why countries would misreport data and do it strategically as their focus is rather about assessing the quality of data used in empirical work. It also appears that they do not correct for the possible persistence of the economic series which could lead to false rejections of the conformity of the realized first-digit distribution with Benford's law, while we address this issue.

From the vast literature it is clear that the public signal is important for economic agents. The literature on the truthfulness of information policies of countries, however, to date is meagre. The problem is that it is not readily verifiable by economic agents whether the information issued by the government is of low quality (i.e., imprecise) or whether the government in question is supplying deliberately false information.[5] Each government or statistical authority has some leeway in providing the various numbers describing the state of the economy as long as these are not too far off from the public expectations; and as long as these can be subsequently put in the errors or omissions sections. Related to the topic is a literature on public signals and transparency under strategic uncertainty recently represented inter alia by Morris and Shin (2002) or Angeletos and Pavan (2007) and explicitly in the context of a currency attack by Heinemann and Illing (2002).

[4]Nye and Moul (2007) found that economic series transformed many times with imprecise procedures (e.g., relying on inaccurate exchange rates) may not obey Benford's law. They advise to work with the raw, nominal data.

[5]Institutions like the International Monetary Fund or the Eurostat can verify at least partially the information that is given out by the country, but this does not solve the problem in its entirety.

The notion of strategically misinforming investors has not been considered in this literature to our knowledge; the authority can control at best the precision (the variance) of the public signal that the investors get, which is not the same as what we mean here.[6] Note, however, that changing the precision of the signal would not constitute data manipulation and our tests would not pick it up, as the data that would be generated should still adhere to Benford's law (see the arguments in Section 4.1 below). Angeletos et al. (2006) and Tarashev (2007) offer models where the government chooses a signal (the interest rate) to inform the public about the state of the economy. This affects also the cost of the agent's later decision. But providing false information is a different issue; the agents may not know well the information policy of the government and (provided that the agents do pay attention to what the government says) this information may affect the belief formation in one way or the other. Analyzing a second generation currency crisis model without any strategic uncertainty, extending the work of Obstfeld (1986) and Obstfeld (1996), Sbracia and Zaghini (2001) notes (without engaging into modeling the strategic choices of a central authority) that greater precision of public information may stop an otherwise imminent attack when the state of the economy is bad and also that the contrary is true when the state is good. This opens up the question whether the government would like to use its information policy to its advantage. On a different note, Sandleris (2008) argues that a country may repay sovereign debt in order not to reveal bad information about the state of its economy. The literature on the strategic delay of information release is related but does not apply here. The large literature on dynamic inconsistency and limited commitment started by Kydland and Prescott (1977) is not relevant in this context; and neither does the literature on moral hazard with ex-post verifiable actions.

Successful misinformation on the side of the government may seem puzzling and is rarely taken into account, as it is at odds with the assumption of rational expectations and the economic agents' knowledge about the true state of the world employed in many models.[7] In some models with strategic uncertainty, information is aggregated through individual trades into prices; and hence the governments' influence over the information known to market participants may be extremely limited. This, however, may be not true all the time; for example in a fixed exchange regime where an important price – the exchange rate – is fixed and the shadow price (the real exchange rate) must be calculated based on the statistics provided by governments. A fundamental question is why individuals (knowing that the government may reveal information strategically) would pay attention to public signals at all; however the basic conclusions of the literature on one-sided private information when signalling is not possible (see for example Aumann et al. [1995]) claims this can be the case, and partial information revelation may be an equilibrium. Another question is whether the ability of governments to provide false information is permanent or just temporary; in the long run serial cheaters may be found out. This paper, by providing some evidence on the strategic character of government release of public signals, provides an incentive to continue and expand theoretical research in these fields.

The paper is constructed as follows. First, we discuss the reasons for which countries would want to misreport economic data; in Section 2 we lay out a model of financial flows and of country incentives to misinform investors about the state of the economy. Then in Section 3 we discuss the data we use while in Section 4 we describe the statistical methodology. Section 5 reports the evidence and Section 6 concludes. Tables with statistical documentation and results of our study are included in the Appendix.

2 Incentives to misreport economic data

To fix ideas, we provide a basic open economy model where we study the incentives of countries to provide false information to investors. Suppose there are N identical countries with a population of workers $L_i = 1$ and endowed with respective immobile capital installed $C_i > 0$, with $i \in \{1, \dots, N\}$;

[6] A different notion is that of vagueness, when a policymaker would release a set of signals containing the true one to be informative in the equilibrium; this notion is studied for example by Stein (1989) based on the model of Crawford and Sobel (1982).

[7] The Billion Prices Project by Alberto Cavallo and Roberto Rigobon is an initiative by private agents to gather information in lieu of government agencies (in this case on the evolution of prices). This way of substituting the government, however, is limited by what can be observed by economic agents themselves. In the case of the balance of payments or gross domestic product data, a similar project is hard to envisage.

we assume in particular that $C_1 = \ldots = C_N$. Consider also an additional $(N+1)$th country with population of $L_{N+1} = 1$ and immobile capital C_{N+1} such that $C_{N+1} < C_1$. There are also $\overline{K}$ units of mobile capital in the world; for simplicity suppose these are owned by some agents which have no domicile, but the capital has to be parked in at least one of the aforementioned countries. There are free capital flows between all countries.

The state of each country's economy can be either High (h) or Low (ℓ); for each country $m = 1, \ldots, N+1$, we denote by $s_m \in \{h, \ell\}$ its state.[8] In a high state of the world a country is more productive, that is, denoting by A the common productivity of all countries, one has $A(h) > A(\ell)$. Each country has the same technology and the statistical process generating a periods' productivity.

The capital (mobile and immobile) installed in country $m = 1, \ldots, N+1$ in a given state of the world's economy is a function of all the states $s_1, \ldots, s_{N+1}$, which we denote by $K_m(s_1, \ldots, s_{N+1})$. With free capital flows, the end level of installed capital will not depend on the country's initial endowment.

The production function in country m, given all countries' states

$$Y_m = A(s_m)\; F\big(K_m(s_1, \ldots, s_{N+1}),\, L_m\big) \tag{1}$$

is Cobb–Douglas with the capital intensity $0 < \alpha < 1$. We will deal with the production function in the intensive form in per capita terms, that is, $y_m = A(s_m)\big(k_m(s_1, \ldots, s_{N+1})\big)^{\alpha}$, where $y_m = Y_m/L_m$ and $k_m = K_m/L_m$.

Suppose first that the states of the world are easily observable among potential investors prior to making the investment. Then, in equilibrium, the stocks of capital employed in each economy are equal as the returns to capital are equalized between countries. That is, the functions K_m are equal in the following sense. The value of $K_m(s_1, \ldots, s_{N+1})$ only depends on s_m and on the total number of countries that have high states of the world $s_j = h$, when $j \in \{1, \ldots, N+1\}$, a number that we denote by $n(h)$. This is why we can simply write in the sequel

$$K_m(s_1, \ldots, s_{N+1}) = K\big(s_m, n(h)\big) \tag{2}$$

for all countries m (and the same for k_m and k). We assume that the endowments of capital C_i and C_{N+1} are such that even when being in the low state ℓ each economy m would receive some international flows.[9] Further when considering two countries m and m' in different states by the equilibrium condition on the equality of capital returns it has to be true that

$$\alpha A(h)\Big(k\big(h, n(h)\big)\Big)^{(\alpha-1)} = \alpha A(\ell)\Big(k\big(\ell, n(h)\big)\Big)^{(\alpha-1)}, \tag{3}$$

which shows that for all economies, we would have in effect $x\,K\big(h, n(h)\big) = K\big(\ell, n(h)\big)$ where the parameter $x = \big(A(h)/A(\ell)\big)^{1/(\alpha-1)}$ is independent of $n(h)$ and satisfies $x < 1$.

Consider the country $(N+1)$ and the total capital employed there $K\big(s_{N+1}, n(h)\big)$. The equation on global mobile capital demand and supply indicates that by conditioning on the possible values h or ℓ of s_{N+1},

$$\left\{ \begin{array}{l} K\big(h, n(h)\big) - C_{N+1} + \big(n(h)-1\big)\Big(K\big(h, n(h)\big) - C_1\Big) + \big(N - n(h) + 1\big)\Big(K\big(\ell, n(h)\big) - C_1\Big) = \overline{K}\,; \\ K\big(\ell, n(h)\big) - C_{N+1} + n(h)\Big(K\big(h, n(h)\big) - C_1\Big) + \big(N - n(h)\big)\Big(K\big(\ell, n(h)\big) - C_1\Big) = \overline{K}\,. \end{array} \right. \tag{4}$$

We get after substitutions the capital values that would be employed in country $N+1$ respectively

[8] We make no assumptions whether the realization of the states of the world are correlated between countries or not.

[9] More precisely, that even if a country is the only one to receive a bad shock, it would still have some inflows of the international mobile capital.

in the high and low states of the world:

$$K\big(h,n(h)\big) = \frac{\overline{K} + NC_1 + C_{N+1}}{n(h) + \big(N+1-n(h)\big)x}$$

$$\text{and} \qquad K\big(\ell,n(h)\big) = x\,K\big(h,n(h)\big) = x\,\frac{\overline{K} + NC_1 + C_{N+1}}{n(h) + \big(N+1-n(h)\big)x}\,, \tag{5}$$

which means that the capital employed in a country falls as the number $n(h)$ of countries with high productivity in the given period increases.

Each country m's gross national product (the remuneration of domestic factors) is $wL_m + rC_m$. Since $L_m = 1$, the wage of country m given the states of the world is $w = (1-\alpha)A(s_m)\Big(k\big(s_m,n(h)\big)\Big)^{\alpha}$ and the interest rate equals $r = \alpha A(s_m)\Big(k\big(s_m,n(h)\big)\Big)^{\alpha-1}$, so that the payoff to country m is

$$R_m\big(s_m,n(h)\big) = A(s_m)\bigg((1-\alpha)\Big(k\big(s_m,n(h)\big)\Big)^{\alpha} + \alpha\Big(k\big(s_m,n(h)\big)\Big)^{\alpha-1}C_m\bigg). \tag{6}$$

It is increasing in the state of the world s_m and in the capital C_m.

Suppose now that the states of the world are not verifiable by capital owners but each state of the world is private information of each country. What are the incentives for say country $N+1$ to reveal the state of the world truthfully? Here we ignore what is the nature of the game (for example, any potential punishments for faulty disclosure of the state of the world); we are interested only in one country's payoffs from claiming a high instead of a low state of the world if investors believed the statement and other countries report their states truthfully. In other words, we consider a unilateral deviation by one country. Then investors, expecting a high state of the world, would invest $K\big(h,n'(h)+1\big)$ instead of $K\big(\ell,n'(h)\big)$, where $n'(h)$ denotes the number of countries $m = 1,\ldots,N$ (country $N+1$ excluded) reporting a high state. Now, in view of (6), the payoff for country $N+1$ from misinforming investors is equal to the difference of payoffs when investors are fooled and invest despite the state of the country being low and the productivity being $A(\ell)$; formally, it equals

$$\begin{aligned} & A(\ell)\bigg((1-\alpha)\Big(k\big(h,n'(h)+1\big)\Big)^{\alpha} + \alpha\Big(k\big(h,n'(h)+1\big)\Big)^{\alpha-1}C_m\bigg) \\ - \quad & A(\ell)\bigg((1-\alpha)\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha} + \alpha\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha-1}C_m\bigg). \end{aligned} \tag{7}$$

Proposition 1. *The payoff (7) from misinforming investors is positive.*

Proof. See Appendix B. □

Countries that suffer a low state of the world (low productivity) have thus incentives to misreport their true economic conditions. One can readily verify that this incentive is higher the more countries are in the same low state (as then a country reporting a high $K\big(h,n(h)\big)$ receives a relatively large inflow) and the lower capital C_{N+1} is installed in the country.

When we turn to data, we will interpret the last two observations in the following way. Countries with high negative asset positions (i.e., endowed with little domestic capital), open to capital flows and/or having a need of financing their negative current account should have higher incentives to misreport their true state of the economy; and in a time of a crisis (defined as a period with many countries suffering a bad state of the world) countries will have higher incentives to cheat. Such countries are cautious not to experience large capital outflows and/or face problems in refinancing their debts. Based on this, we suppose that countries that have a negative net asset position and that have liquid portfolio markets are more likely to defend themselves during a crisis using all means possible – including misreporting information.

It is possible to incorporate more features in the above model easily.

Fixed exchange rate regimes. Suppose there is a cost of θ of abandoning a fixed exchange rate regime, and the country holds some reserves R of capital (or this is their borrowing capacity in case of an attack). If the desired capital outflows exceed R the exchange rate regime is assumed to be abandoned (or a devaluation occurs). Then, this could lead to the realignment of the exchange rate when a Low state of the world would succeed a High state (if the R held is low enough). A truthful declaration of the state of the world would cost the country θ apart from the effects of having a lower installed capital stock as in (5). Hence, ceteris paribus, we expect that countries with fixed exchange rates are going to have more incentives to misreport economic data so as to defend their exchange regime.

Countries open to capital flows. In what we assumed above, capital flows were unrestricted for the mobile international capital. Suppose now that there are some bounds on capital flows, so that say you have to pay a tax τ in order to place capital into a country. Then, given the same net foreign asset position, the same state of a world, a country with high openness (low τ) will have higher incentives to misinform investors in the low state of the world as the resulting capital flows will be more important.

3 Data

For our tests to have rejection power as described in Section 4, we require a large amount of comparable data. Furthermore, we seek economic series which could be manipulated by the government and at the same time easily observed and followed by a large group of investors. For our purpose, we use quarterly balance of payments data from the International Monetary Fund BOPS database for the period 1989–2007.[10] This data is provided by statistical agencies of the IMF members and should be prepared in a standard way, using the Balance of Payments manual of the IMF (2005). It is provided in U.S. dollars, which means that the BOPS data was transformed in a minimal way (see the importance of this in Nye and Moul [2007]) given the fact that most international transactions are conducted in this currency. The initial date of 1989 is chosen for two reasons: before 1989 the data is available for relatively few countries and international capital flows were relatively small as the financial markets were still closed; also after 1989 many newly established countries (for example, from the break-up of USSR, Yugoslavia or Czechoslovakia) emerge that start providing statistics. For this reason, in the middle of 1990s we are able to have over 600 datapoints for each quarter. We use quarterly data because this is the data that investors often take into account while making their decisions. Yearly data occurs with a sufficient lag and may not say much about future developments (and, as a result, may not be actively falsified); moreover we would get four times fewer datapoints. Monthly data, on the other hand, is available only for a few countries on a regular basis. We take all the net figures available for each country from the Balance of Payments so as to lower the problem of persistence (see Section 4.3 for a discussion). All in all, we have at maximum 13 entries[11] per country for each quarter, with 76 quarters of data. This gives us at a maximum 988 observations for each country. More detailed statistics for the balance of payments data are not available for a large set of countries and often come with a considerable lag.[12]

Data on the balance of payments is relevant for investors in the view of our model as it provides information about the state of productivity of capital in a country in numerous ways. Firstly, changes in the current account balance, inter alia, are a measure of the country's competitiveness relative to the rest of the world. In a fixed exchange regime, the value of the information coming from the balance of payments is even higher to the agents as the exchange rate (which would contain aggregate information about the underlying transactions) does not change. A sustained

[10] The Balance of Payments data was accessed online at the IMF website on 31 December 2009.

[11] These are: Current Account: net; Goods: net; Services: net; Income: net; Current transfers: net; Capital Account: net; Financial Account: net; Direct Investment: net; Portfolio Investment: net; Other Investment: net; Financial Derivatives: net; Reserve Assets: net; Net Errors and Omissions.

[12] We are thus not using the data that is supplied directly to investors in real time, but the data that is given to the IMF with some lag. Nevertheless, the IMF is an important lending institution which has an influence on the financial conditions in many, especially non-OECD, countries. The data submitted to the IMF is also used by market participants in their analyzes.

current account imbalance may point to a misalignment of the real exchange rate. Moreover, investors obtain information from the balance of payments about the solvency of the country and its firms, so directly on the possible returns to capital within the country in the nearest future. For example, changes in reserves, or changes in the volume and composition of the current account items have implications for the ability to retire capital (for possible capital flow restrictions in the future), the probability of a speculative attack (about the valuation of the capital installed in the country) or the solvency (or possible illiquidity problems) of the country and the firms located there (which affects the value of their assets). Reinhart and Rogoff (2009) include among the best few predictors of banking crises and currency crashes (compiling from the rich literature on early crisis indicators) respectively such indicators as the short-term capital inflows/GDP, current account balance/investment, current account balance/GDP and exports.

We could also (but chose not to) add into consideration other government issued data describing the economy that are closely watched by investors such as inflation, quarterly GDP or industrial production figures. We did not do it, reluctantly, for the following important reasons. There are considerably fewer statistics for GDP or industrial production available for each quarter for a wide range of countries. Inflation figures may be very persistent. All these measures are also calculated using different methodologies from country to country and typically involve many transformations of the data. On the other hand, balance of payments data comes for each country always from the same provider, and it is fairly standardized accounting data.

To categorize countries according to their international investment position, we use the updated and extended version of the External Wealth of Nations Mark II database developed by Lane and Milesi-Ferretti (2007). These authors provide yearly estimates of foreign assets and liabilities in the period 1970–2007 for many countries.

The exchange rate regimes were classified based on information of the International Monetary Fund in various issues of the "Annual report on exchange arrangements and exchange restrictions" from the period 1989–2007 and the data provided by the Fund staff themselves; see IMF (1989–2008). We opted for the de facto classifications of the IMF. What matters for us is the end-of-period (quarter) status of the regime. Any exchange rate regime where the intervention of the central bank was substantial (a fixed regime, a crawling peg, a crawling band) was considered as heavily managed and labeled in our dataset as "fixed."[13] Such a wide range of regimes was classified together to have a large number of different countries. We label a regime as "floating" when it was deemed by the IMF as being independently floating.[14] It is to note that the currencies that were a part of the European Exchange Rate Mechanism I or II in a given quarter are considered to belong to the fixed category prior to the introduction of the Euro and as floating afterwards.

Data for institutional quality measures were taken from the World Bank's World Governance Indicators database (Kaufmann et al. [2009]) while those for country ratings from the Institutional Investor.

All in all, we have data for 103 countries for the period 1989–2007. We also considered a subset of the data set described above; it is obtained by considering only the countries reporting balance of payments data for at least once quarter and one series a year between 1995 and 2007, keeping only the observations corresponding to the latter period of time. This subset will be referred to as the stable sample in the sequel while the notion of full sample will refer to the whole data set. We have 66 countries in the stable sample for the time period 1995–2007.[15] We inspect this sample for the following reasons. Some additional data that we use to provide conditions for slicing the samples (data on country investment positions, World Governance Indicators etc.) is available for many countries for this later period. Importantly, when we scrutinize yearly data only, the

[13] Consequently, for example crawling band regimes prior to 1998, when the IMF adopted a new classification, were coded as being fixed from the category of managed floating.

[14] The remainder of the countries have regimes which are difficult to classify into either category. For some heavily managed exchange rate regimes it is difficult to make a call whether it is a floating regime or not; and we preferred only to make conservative assessments. Secondly, we classify here also countries when they do not have any own legal tender (like Panama, El Salvador or Ecuador). Indeed, in our tests, countries with regimes that are neither fixed or floating exhibit different behavior than the two "ideal" regimes (available upon request).

[15] Choosing 1994 as the initial year would give us only 61 countries with more incomplete data, while picking 1996 would only raise the number of countries in the sample to 70. In both cases the alternative choice would lead to fewer observations being included in the stable sample. As we need as many countries and quarters as possible given the arguments advanced in Section 4, the choice of the sample 1995–2007 is a natural one.

number of observations per year remains roughly constant[16] while using this stable subsample. Moreover, the countries included therein may provide higher quality data. For example, Poland is excluded from this sample as it did not report the data between 1996–2000 when it was changing and improving the statistical methodology of balance of payments reporting. Some countries may have not reported the data also for strategic motives. For these reasons, the stable sample is our preferred one.

Our dataset contains data from heterogenous countries from all continents (cf. Section A). The characteristics of the data are given in Table II. The basic characteristics of the variables in the full and stable samples do not differ substantially. The major exception is in the exchange rate regimes: a higher fraction of the observations is available for the purely floating regimes in the stable sample than in the full sample (36.3 % vs. 28.2 % respectively) and a lower fraction for fixed exchange rate regimes (37.5 % vs. 45.8 % respectively). The countries in the stable sample have on average slightly higher liabilities but have higher net foreign assets (are less indebted). On average, they also have better institutional rankings.

4 Statistical methodology

In this section we first recall why it is expected in theory that the numbers contained in the balance of payments data would follow a specific distribution called Benford's law and then show that this is empirically the case. We then demonstrate using chi-square tests of goodness of fit that this conformity to Benford's law extends to most of the randomly selected subsets of this data set with either many countries involved or few quarters per country.

A persistence issue appears for the subsets containing data from few countries for many (consecutive) quarters; these usually do not conform to Benford's law. The reason is that the same first digits may be present for several quarters, and since we do not have a sufficiently large number of time periods, the distribution of the first digits may appear to be unusual even when no data manipulation is in place. To overcome this persistence problem and study such subsets, we create a new assessment criterion based on choosing sub-subsets of these subsets at random a large number of times.

4.1 Benford's law

Simon Newcomb (1881) noted, by looking at the most worn pages of logarithmic tables, that first digits of real-life numbers were not uniformly distributed but that the smallest values were more likely. He postulated that this could be explained by a "law of probability of the occurrence of numbers [...] such that all mantissas of their logarithms are equally probable." A similar observation was made again, independently, by the physicist Franck Benford (1938). He collected a huge amount of data from various sources (street addresses, numbers appearing in a given issue of the *Reader's Digest*, surface area of rivers, etc.) to estimate the corresponding universal distribution of the first digits of these numbers, leading to the same distribution as the one proposed by Newcomb and named since then Benford's law. The predicted occurrences of first digits are given, for all $j \in \{1, \ldots, 9\}$, by

$$\theta_j = \log_{10}\left(1 + \frac{1}{j}\right) . \qquad (8)$$

The values of these frequencies θ_j are given in Table I. This law also describes the frequencies of occurrences of the next digits, which we do not use here for the lack of a sufficient number of observations.

4.1.1 Mathematical foundations of Benford's law

Three natural processes or characterizations lead the data to have this distribution, which, as described below, is important for our purposes. These are processes with exponential growth; scale-invariant (unit-invariant) processes; and random samples of random distributions. We label these properties F1, F2, and F3 and this is how we refer to them in the remainder of the text.

[16] This is because some net items are not provided at all times even by countries in this sample.

j	θ_j	j	θ_j
1	30.1 %	6	6.7 %
2	17.6 %	7	5.8 %
3	12.5 %	8	5.1 %
4	9.7 %	9	4.6 %
5	7.9 %		

Table I: Benford's law $(\theta_1, \ldots, \theta_9)$ for the first digits.

(F1) Processes with exponential growth. Diaconis (1977) shows that geometric sequences $\{a, a^2, \ldots, a^n\}$ with ratio a such that $\log_{10} a$ is irrational[17] lead to sets of data that conform more and more closely to Benford's law as n grows. That is, almost all geometric sequences lead to Benford's law as the set of rational numbers has a null measure within the set of all real numbers. This implies that *in practice* all geometric sequences linked to real data lead the data to obey Benford's law. The result can be extended to certain random geometric sequences $\{X_1, X_1X_2, \ldots, X_1X_2 \cdots X_n\}$.

Another generating process: log–uniform mantissas. Another generating process that is less interpretable in terms of economic modeling but is useful to gain intuition is to pick numbers $Y_1, \ldots, Y_n$ at random in the interval $[1, 10)$ such that the random variables $\log_{10} Y_1, \ldots, \log_{10} Y_n$ are uniformly distributed over the interval $[0, 1)$; that is, we consider log–uniform mantissas. Then, for all natural integers $k_1, \ldots, k_n$, the distribution of the first digits of the numbers $Y_1 \times 10^{k_1}, \ldots, Y_n \times 10^{k_n}$ converges to Benford's law as n grows.

(F2) Scale invariance. Pinkham (1961) and Hill (1995a) proved that Benford's law was the *unique* law such that the distribution of the first digits of number drawn at random according to it was stable when the drawn numbers are multiplied by a common multiplicative factor.

(F3) Random samples of random distributions. Hill (1995b) considered the case of data sets with numbers chosen at random according to a two-step procedure choosing first at random a distribution over the positive real numbers, then, drawing k numbers according to it, and finally repeating the whole procedure a large number n of times; he provided natural conditions on the distribution of the random probability distribution for Benford's law to occur when n grows. Janvresse and de la Rue (2004) showed that it suffices to consider random probability distributions taking values in the family of uniform distributions over intervals. This mathematical phenomenon shows in particular that data coming from many different distributions is much more likely to conform to Benford's law, as observed[18] by Benford himself in his attempt to collect a huge amount of data.

4.1.2 Application to macroeconomic data

Based on the above properties F1–F3, and on the properties of artificially generated sequences in the existing literature, we argue that macroeconomic data should lead to Benford's law; the discussion that follows refers heavily to Nye and Moul (2007).

Theoretical grounds. The sequence of macroeconomic statistics for a given country should, if observed for a long enough time, result in a collection of numbers with distribution of first digits

[17] Otherwise the sequence of the first digits is cyclic and has asymptotic proportions that are rational and thus are different from the ones of Benford's law; consider the simple example when $\log_{10} a = 2/3$, in which case the sequence is $10^{2k/3}$ and contains only elements with first digits equal to the first digits of $10^{2/3}$, $10^{4/3}$, and 100, that is, 4, 2, 1; the asymptotic repartition of the first digits is uniform between these three numbers.

[18] Raimi (1969) points out that whereas many tables of data considered by Benford did not conform individually to the prescribed law, their union did: "what came closest of all, however, was the union of all his tables."

abiding Benford's law; this is due to economic growth and inflation, in virtue of the mathematical foundation F1. For example, in a stationary economy with constant real GDP, the first digits of the nominal GDP figures should adhere to Benford's distribution in the long run when the rate of inflation is for example 5 % because $\log_{10} 1.05$ is irrational. The presence of some errors in the data does not impact per se its conformity to Benford's law, provided that these errors can be described in terms of random multiplicative factors whose logarithms only put mass on the irrational numbers, as in foundation F1. Property F2 states that the distribution of first digits according to Benford's law is preserved even if data for a country was provided in home currency and then converted into U.S. dollars (which is the case for the balance of payments data that we have). Finally, foundation F3 means that aggregating the data of several countries also preserves the conformity to Benford's law – and even should increase it; this is important to remember especially when the number of data points per country is too small (i.e., the series are too short) for foundation F1 to lead to the conformity to Benford's law per se.

Previous studies. Varian (1972) suggested assessing the quality of data generated by forecasting models by testing whether the distributions of the first digits of the generated data adhere to Benford's law. He generated series from some economic forecasting models and confirmed that indeed the generated data appears to be natural. At the time of his writing, however, the theoretical foundations F1–F3 were not known. The most thorough study of conformity to Benford's law of data in macroeconomics to date was pursued by Nye and Moul (2007). They showed by means of simulations that long enough sequences of nominal GDPs for a set of fictitious countries formed sequences of numbers whose first digits obey Benford's law, provided that the random economic factors (various growth rates) were set in a proper way. They then illustrated on the Penn World Tables dataset that some subgroups of countries (OECD countries on the one hand, African countries on the other hand) are such that the data set obtained by aggregating their GDP data was conform to the predictions of Benford's law. They also argued that various transformations of data that are operated on nominal data (and require inflation adjustments, purchasing power indices, etc.) may alter the quality of data, since they lead to sets of numbers that do not conform anymore to Benford's law. That is, they used Benford's law as a test of naturalness of the data, as is the case in other fields, e.g., accounting. All in all, they came with the conclusion that "broadly speaking, country GDP figures should be consistent with Benford's distribution when countries are heterogeneous in their initial levels (e.g., population, currency relative to U.S. dollar, per-capita income) and then grow."

4.1.3 Use of Benford's law in other fields

The idea of detecting manipulations in the data by tests of conformity to Benford's law is now well established is auditing and accounting, after the seminal article of Nigrini (1996) and the successful use of his methodology by the authorities of the city of New York, leading to the detection of frauds in seven companies; see Nigrini (1999) for an extended review of many other successful stories. The methodology there is to use as the data set to be tested all quantities appearing in accounts (the values of all individual transactions); the theoretical justification behind this is idea is mainly the foundation F3 discussed above.

Other occurrences and practical uses of Benford's law can be found, e.g., in Hill (1995b); one can cite, among others, stock market data and census statistics as occurrences and the design of more efficient computers as a possible use.

4.2 Conformity of our proposed data set to Benford's law

The theoretical properties F1–F3, the previous empirical studies as Nye and Moul (2007) or the accounting literature imply that the balance of payments data (which is accounting data) should conform in its entirety to Benford's law; hence we do not provide any simulations of artificial balance of payments series, but we turn directly to the empirical assessment below.

We consider here the aggregated data set obtained by considering all available series of the balance of payments data for all countries and all quarters (henceforth referred to as the full sample); we study the conformity of the first digits of the numbers thus obtained to Benford's law via the chi-square test of goodness of fit against this distribution.

Such a goodness-of-fit test compares the empirical frequencies $\widehat{\theta}_j$ of the digits $j \in \{1,\ldots,9\}$ to the theoretical frequencies θ_j prescribed by Benford's law, via the statistic

$$D^2 = N \sum_{j=1}^{9} \frac{\left(\theta_j - \widehat{\theta}_j\right)^2}{\theta_j}\,, \tag{9}$$

where N denotes the total number of numbers (of first digits) available in the data set. The statistic D^2 converges in distribution to a chi-square distribution with 8 degrees of freedom as $N \to \infty$; in practice, the distribution of D^2 is close enough to this limit whenever $N \geqslant 30$ and $N\theta_j \geqslant 5$ for all $j \in \{1,\ldots,9\}$, which translates in our case to

$$N \geqslant \frac{5}{\min_j \theta_j} = \frac{5}{\theta_9} \geqslant 110\,. \tag{10}$$

This approximation is used to report P–values associated with values of the statistic D^2 in the tables.

In the first line of the first four columns of Table III, we assess the conformity of the data set considered in its entirety to Benford's law; we report in the third column the corresponding P–value and in the fourth column the value of N. The first and second columns indicate the number of countries encompassed by the economic condition considered (e.g., no condition, whether a country was an OECD member in 1989 or not), as well as the average number of quarters per country. In the case of the full sample, these two quantities equal respectively 103 and 76 as indicated above.

Since the P–value is much larger than any conventional threshold for rejecting a null hypothesis, we see that the data set considered globally conforms to Benford's law. This conclusion extends to some subsets of the data, in particular to those formed by OECD and non-OECD countries or the stable sample discussed at the end of Section 3.

4.3 Conformity of subsets of the data set to Benford's law: An issue of persistence

In this paper, we consider subsets of the entire balance of payments data set defined by some economic conditions and show that some of them, for appropriately chosen conditions, lead to subsets whose distributions of the first digits do not conform to Benford's law.

Before doing this we need to show that in the BOPS data set typical randomly drawn subsets do conform to the latter law. These subsets are the ones corresponding to the choice of a large number of countries and/or a few number of quarters per country. (This statement is to be made precise below.)

Before presenting this detailed study, we motivate it by the perhaps deceiving results that are obtained when testing whether the data corresponding to each country is Benford distributed or not: the test accepts the hypothesis of conformity to Benford's law at the level 5 % for 58 countries but rejects it for 44 other countries (whereas no conclusion can be drawn for the remaining country, Serbia, for which fewer than $N = 110$ observations are available). That is, if one chooses one country with enough observations at random, the obtained data set will not conform to Benford's law with probability $44/(44 + 58) \approx 43.1\,\%$. There is no striking pattern in countries for which we obtain these rejections. For example, we reject the null hypothesis for the USA, France, Switzerland, Japan or Canada while for many non-OECD countries we cannot.

As already noted and illustrated by simulations in Nye and Moul (2007, Section I), it seems necessary to consider several countries for the corresponding data subset to conform to Benford's law; this is to increase heterogeneity at the initial levels and match the requirement provided by the foundation F3 that the number n of separately drawn subsets should be large. For example, although for many individual OECD countries the distribution of the first digits does not conform to Benford's law (as noted above), for the entire set of the OECD countries for which we have data (which includes 22 countries) Benford's distribution cannot be rejected (see Table III). We further illustrate this fact on the top part of Table IV, where we estimated the probability that when choosing a given number C of countries at random from our balance of payments data set, the data subset formed by these C countries passes the test of goodness of fit against Benford's law at the 5 % level.

This table indicates that a significant fraction of the subsets formed by C countries, with $C \leqslant 70$, cannot be said to be distributed according to Benford's law. Only when taking a large number of countries (e.g., $C = 80$), one gets subsets that typically can be said to conform to Benford's law.

4.3.1 Persistence issues sidestepped: considering fewer quarters per country

In the subsets grouped by economic conditions, however, fewer countries will be typically involved; fewer quarters too, and this will be the key. Indeed, as is shown in Table IV (bottom table, first lines of each group of lines), the rejection rates become smaller as either the number C of involved countries increases or the number M of quarters picked at random for each country decreases.

This phenomenon can be explained by persistence of data from one quarter to another for individual countries; it is especially crucial to deal with it when a small number C of countries is involved for otherwise, enough heterogeneity is introduced for the aggregated data to conform to Benford's law. We use net items from the International Monetary Fund BOPS database to reduce persistence as much as possible, but obviously an extra treatment is needed. Before showing this treatment, we give a second illustration of persistence.

4.3.2 Persistence issues sidestepped: selecting some series from the balance of payments data

All previous methods relied on considering for each country–quarter pair all the 13 series that occur in the balance of payments data we obtained from the IMF. We now study what happens when only some series are selected. To that end we considered two choices, the *independent* series and the *less persistent* series.

The first subset of series is formed by taking into account the identities that occur in the balance of payments data and by removing a series for each such identity; several choices were possible and the one we made is the following: Current account; Goods; Services; Income; Financial account; Direct investment; Portfolio investment; Financial derivatives; Reserve assets; Net errors and omissions. That is, we dropped the series: Current transfers; Capital account; Other investment.

The second subset of series is formed by taking, out of the 13 series, the 6 series which showed on average the smallest persistence from one quarter to another. For each series and each country, we computed the number of breaks in the sequence of the first digits indexed by quarters and then considered the average of these results with respect to countries. The series included in the resulting subset are then: Financial account; Portfolio investment; Other investment; Financial derivatives; Reserve assets; Net errors and omissions.

We first study what happens for individual countries. We recall that using all series, the data corresponding to 58 countries could be said to conform to Benford's law whereas for 44 countries, it did not pass the test (whereas for 1 country there was not enough data). For independent series, the respective figures are 54, 44, and 5 while for the less persistent series, we obtain 78, 15, and 10. Clearly, the persistence of the series does matter for rejections of Benford's distribution for individual countries.

In addition to this comparison, we also used the same random drawing methodology as above to obtain Table IV. We only reported therein the results for all series and the less persistent ones since the consideration of the independent series instead of all series almost does not change the picture for any pair (C, M) (detailed results provided upon request). On the other hand, the restriction to series chosen as being the less persistent ones ensures that for almost all pairs, the rejection rates get close to or smaller than 5 %.

This illustrates once again that the observed persistence is due to a lack of independence in the quarter-to-quarter values of the series and is not linked to an intra-quarter dependence caused by the identities between the series of the balance of payments.

4.3.3 Remedy: Considering random sub-subsets

A remedy could have been to select only the less persistent series; however, doing so one may not detect the manipulations on the most crucial series (that are watched and analyzed by investors), which can be the most persistent ones as well (for example, the balance on the current account or goods entries). Some persistent series in fact may also be such *because* they are manipulated. This is why we focus on the fact that conformity to Benford's law is highly likely when the data subset is formed randomly by the consideration of few quarters only (relatively to the number of countries) as seen in Table IV.

In particular, Table IV shows that when at most 10 quarters are selected at random for each country of the data set (that is, by taking a random fraction of about 10 % of the data subset), the resulting data subset typically conforms to Benford's law.

We use this observation in the following way. While assessing the conformity of a subset, we do not consider all quarters of the corresponding subset, but take a small fraction of them at random and test for the conformity to Benford's law, a procedure we repeat a large number of times. The average rejection rate will be an indicator of conformity to Benford's law without the bias due to persistence.

The method formally here is, given a subset D of the data, to pick at random a given fraction f of its country–quarter pairs, form the corresponding sub-subset, and test whether the latter can be said distributed according to Benford's law at the level 5 %; a fact that we denote by R_D, which is therefore a Bernoulli random variable. We repeat this procedure a large number of times, say, $1,000$ times, by fixing D but by choosing different sub-subsets of it at random. This gives rise to the random variables $R_{D,1}, R_{D,2}, \ldots, R_{D,1000}$, which, conditionally to D, are independent and identically distributed according to a Bernoulli distribution with parameter denoted by q_D. We consider the empirical mean $\overline{R}_{D,1000}$ of these random variables as the statistic of interest.

The heuristics behind this procedure is to reduce efficiently the average number of consecutive quarters at hand per country, thus weakening persistence issues, and to consider q_D instead of the P–value quantifying the goodness of fit of D against Benford's law. The repetition of the procedure aims at obtaining a stable result given D. Formally, me mean that

$$\left[\overline{R}_{D,1000} \pm 1.96\sqrt{\frac{\overline{R}_{D,1000}\big(1-\overline{R}_{D,1000}\big)}{1\,000}}\right] \tag{11}$$

is a confidence interval at a 95 % confidence level for q_D, where, given that the typical realized values of $\overline{R}_{D,1000}$ are around 0.10, the precision of the estimation of the rejection rate q_D is about $\pm 2\,\%$.

To be able to use this new criterion, we first determine the typical values of the statistic $\overline{R}_{D,1000}$ when $f \in \{5\,\%, 10\,\%, 20\,\%\}$ and D is a set drawn at random[19] from Benford's distribution with a given size N. Results are reported in Table V, which provides estimates of the quantiles of the underlying distribution according to the values of f. These estimates were constructed by running $1,000$ times the above procedure on randomly generated sets D. This involved computing 10^6 tests of goodness of fit per cell of the table; because of the computational cost, no sharper estimates based on more repetitions are provided.

This method provides us with another test of conformity to Benford's law of subsets D; this test, given D, compares the obtained value of $\overline{R}_{D,1000}$ to the above quantiles and is then able to associate a P–value with the hypothesis of conformity to Benford's law. This is done by identifying in which interval of the last line of Table V (corresponding to the chosen f) this value lies in. For instance, if the realized value of the statistic $\overline{R}_{D,1000}$ is 8.5 and f was chosen equal to 10 %, the corresponding P–value is between 1 % and 5 %; or if the realized value is 5.4 and f was chosen equal to 5 %, the the corresponding P–value is larger than 5 %.

[19] Of course, it is immediate that in this case, the expectation of $\overline{R}_{D,1000}$ (with respect to the choice of D and the random subset) is 5 % but we want a sharper idea of its distribution, namely, good estimates of its tail.

4.4 Conclusions of the initial statistical study

4.4.1 How to assess the conformity to Benfords's law

The aim of this preliminary statistical study was first to show that the considered data set globally conforms to Benford's law and secondly to exhibit conditions under which the subsets of this data set are or are not typically Benford distributed. For the latter, a trade-off between

– the consideration of many consecutive or close in time quarters (issue of persistence)

– and the number of countries at hand (which increases the heterogeneity)

needs to be set: if there are too few countries, not more than a small number of quarters per country can be taken, otherwise the behavior of subset is doomed to be far from Benford's law.

This restriction to a small number of quarters can be done at random, leading to the proposed remedy against the persistence issue, which relies on subsampling in a proportion f, with $f \in \{5\,\%, 10\,\%, 20\,\%\}$ in the experiments below. We call these, respectively, 1-in-20 (when $f = 5\,\%$), 1-in-10 (when $f = 10\,\%$), and 1-in-5 (when $f = 20\,\%$) criteria. We are going to include the 1-in-10 criterion in our main results as this already delivers good randomness of the sampled data, but is not as demanding in terms of the number of required observations as the 1-in-20 criterion.[20]

Alternatively, one can also discard the most persistent series in the balance of payments data; the remaining series typically exhibit behavior close to what is predicted by Benford's law (as shown by the small rejection rates indicated in Table IV).

In the rest of the study, we will identify the subsets of the data defined by some economic conditions that are detected not to conform to Benford's law whereas the subsets of their sizes should typically do. This will flag the conditions under which manipulations or at least embellishing of the data was performed. The resulting practical criteria to assess the conformity of a subset D to Benford's law of our balance of payments data in this respect are summarized below.

(C1) The plain P–value of the chi-square test of goodness of fit on D the is a relevant statistic if and only if the average number of quarters per country is small (of the order of 10); or if the number of involved countries is large (of the order of 80); or if the number of countries is moderate (larger than 60) and the average number of quarters is not too large (less or equal to 20). These statements are made precise by the rejection rates exhibited in Table IV.

(C2) The criterion based on random sub-subsets of D (1-in-10) is an efficient way to reduce the bias due to persistence when few countries are involved with many quarters of data each. The obtained rejection rates should be compared to the quantiles provided in Table V.

(C3) The consideration of the less persistent series brings auxiliary information on whether the possible non-conformity to Benford's law indicated by criteria C1 and/or C2 may or may not be due to the persistence of the balance of payments data. However, finding a conformity of the first digits of these less persistent entries with Benford's distribution does not indicate that there is no data falsification going on at all.

4.4.2 Application to the conformity of the BOPS data set

We included all the criteria discussed above in Table III. We see that all three criteria C1–C3 indicate a good adequation of the first digits of the balance of payments data to Benford's law. There is no evidence that Benford's distribution does not describe well either the first digits of balance of payments entries either for the whole sample or our subsample of countries for which we have at least one data point for each year after 1995 (the "stable" sample, defined below). It does not appear either important whether we restrict the samples to include only OECD members (as of 1989) or not. This is very important, as we do not find that non-OECD countries have non-Benford distributions of the first digits which is very different from the results of Nye and

[20]More precisely, to ensure that for each chi-square goodness-of-fit that need to be performed at least 110 observations, as mentioned in (9), are available, at least $1,100$ and $2,200$ are respectively needed to compute the 1-in-10 and 1-in-20 criteria.

Moul (2007) on the Penn World Table dataset and we do not believe that these countries provide data of low quality.

We turn now to our main object of interest, detecting deviations from Benford's law for different groups of countries.

5 Results for country-quarter groups based on economic conditions

As discussed at the end of Sections 3 and 4, in our investigation we offer results for two samples – the full sample of balance of payments data for 1989–2007 and a "stable" subsample of countries which provided data at least once each year. We perform chi-square goodness-of-fit tests on the subsamples grouping country-quarter pairs by economically meaningful characteristics to detect any irregularities in the first-digit behavior. The null hypothesis is that the first digits of the data are drawn from Benford's distribution. Based on our model in Section 2, we believe that some groups of countries have higher incentives to tweak their balance of payments statistics than others. Fixed exchange rate regime countries may be wary to provide truthful information about the developments in the balance of payments because of a fear of an attack on their currency that unfavorable information may trigger. Countries with high net foreign asset positions in terms of GDP or those having negative current account balances may fear that too much of negative information in the balance of payments data may spur a nervous reaction of investors and capital outflows, which they would like to avoid. This would be also especially true in times of a global crisis.

While investigating the specificity of countries with fixed exchange rate regimes or negative current account balances, we also look at subgroups of countries that have relatively higher negative foreign assets or higher foreign liabilities. We look at the following criteria, for which we indicate in parentheses the shortcut used in the tables to report their values:

– net foreign assets as a ratio of GDP (NFA_GDP);

– net foreign assets excluding foreign direct investment to GDP (NFA_EXCL_FDI_GDP);

– liabilities to GDP ratio (LIAB_GDP);

– liabilities excluding foreign direct investment to GDP (LIAB_EXCL_FDI_GDP);

– foreign equity liabilities to GDP ratio (EQ_LIAB_GDP);

– the sign of the current account in a given quarter (CA).

In grouping countries we investigate datasets that are the unions of various conditions. For the international investment positions, we took the top 75 % of countries ranked in terms of negative net foreign assets or foreign liabilities.[21] In all cases, we use previous year's figures for grouping countries. So, for example to test for a group of countries with fixed exchange rate regimes that also belong to countries with high total foreign liabilities in the whole sample, we took all the countries which had the total liabilities to GDP ratio higher than 0.5427 in the preceding year that also had a fixed regime.[22] As a measure of capital openness we used the foreign equity liabilities/GDP ratio. A relatively higher value of this indicator shows that a country is de facto more open to private capital flows than others, and hence may be also more prone to capital outflows.[23]

[21] In the robustness checks, we also check what happens when we use the top 90 %, 80 %, 66.66 % or 50 % of countries (not reported).

[22] The quantiles of a given economic quantity (e.g., net foreign assets to GDP) were computed by considering all the available values of the quantity as countries and quarters vary; that is, at most one value of the quantity per country-quarter pair of the sample (full or stable) was considered: none when the quantity was unavailable and one when the data contained information about it. For this reason the number of available observations in, for instance, the top 10 % or lower 10 % of the country–quarter pairs for this criterion may differ; the difference is due to the grouping according to the criterion by itself.

[23] We did not use the total portfolio liabilities (equity and debt) as many debt liabilities for a country consist of foreign denominated debt which is traded outside of the country borders. There are also fewer datapoints for this measure.

The amount of data that we were able to gather allows for grouping the countries into many different subsamples; however with many restricting conditions very quickly these subsamples may become small, containing few countries but with many quarters of data which would lead to easy rejections of Benford's distribution due solely to persistence, as discussed in Section 4.3. In such cases, we want to rely on the 1-in-10 condition but then we need about 2,000 observations to be able to use these results, as discussed in Section 4.3.3. Hence we are unable to slice the data very finely; this is against our story as we cannot condition very well in each category for countries that may have higher incentives to cheat.

We offer several robustness checks of our main goodness-of-fit tests. First of all, we typically test whether for the complement of the set for a given restricting condition (or an alternative regime in the case of an exchange rate regime) we reject Benford's distribution for the first digits as well or not. Such a rejection would make our claim of the relevance of a potential category for grouping countries and detecting data manipulation vacuous. Next, we want to know whether it is not only one country which drives the result for a particular category. This may indicate that out of the whole group only one country is either providing false information or, for example, suffers from a dire problem of persistence in its quarterly data. The fact that such a country would drive the results would invalidate the generality of the claim, as single individual countries might report data with non-Benford distributed first digits for different reasons. This number of countries driving the results (which we label as the "stability index") is quantified by the number of countries which, when excluded from the subset, lead to acceptance of the null hypothesis at a level of 10 %; that is, we count the number of countries such that when they are excluded one at a time the P–value associated with the goodness-of-fit test to Benford's law increases from less than 5 % to more than 10 %. We would like our stability index to be zero when we expect to reject the null hypothesis. We also test for the rejection of Benford's distribution for the less persistent items in the balance of payments and offer results for different criteria based on random sub-subsets (1-in-5 and 1-in-20).

5.1 Fixed exchange rate regimes

In Tables VI–VII we present the tests for the conformity of the first digits of the balance of payments data with Benford's distribution for country–quarter pairs groups in terms of their exchange regimes (fixed or floating, as defined in Section 3) at the end of a quarter. The left hand column presents the characteristics of the groups considered.

First of all, we observe that the null hypothesis of conformity to Benford's law is rejected for fixed exchange rate countries both for the stable and full samples at a 1 % significance level. The subsets of countries that have fixed exchange rate regimes contains many countries (84 countries with 32.6 quarters on average for the whole sample and 49 countries with 26.6 quarters on average for the stable sample) so that we believe based on Section 4.3.1 that the persistence is not driving our results. The 1-in-10 criterion confirms this. The important thing to notice is that any irregular behavior detected here may not be country specific, as many countries in our samples (27 and 40 respectively for the stable and full samples) change the fixed exchange rate regime to floating (or the other way round) at least once.

Moreover, in all further considered scenarios (after imposing more restrictive conditions) groups of countries with fixed exchange rate regimes have distributions of the first digits for which the hypothesis of Benford's distribution is rejected at a 1 % level for all series. The 1-in-10 criterion confirms our findings as these statistics are always greater than 10. More precisely, the level of significance of the rejection increases by a large amount when we further refine the picture and consider countries which belong to the upper 75 % in terms of the size of their total foreign liabilities, the (negative) net foreign assets (with or without foreign direct investment), and equity liabilities. This is also true for fixed exchange rate regime countries that have a negative current account balance in a particular quarter. This supports our hypothesis that countries with fixed exchange rate regimes that would be more sensitive to capital flows have an interest in tinkering with the signals that they send out to investors. In particular, the rejection rate of Benford's distribution for countries with fixed exchange rate regimes and relatively liquid equity markets within our sample (which means they need to be de facto open to individual financial flows) is 0.00004 % in the stable sample and 0.0007 % in the full sample! This occurs for a large number of countries (38 and 58 countries respectively) in this category with an average dataspan of 17.8 and 24.2 quarters.

Our findings square with the model that we presented in Section 2: it is the countries with fixed exchange rate regimes that wish to misreport data; and among these countries it is those that have higher negative net asset positions or higher openness that have more incentives to do so.

We believe that the results discussed above survive our stress tests (shown in Tables VIII–IX). For the full sample, the hypotheses of Benford's distribution of the first digits is still rejected even when we exclude from the sample each country one-by-one (as the values "0" for the stability indexes indicate); which means that there may be several countries in a given group that are responsible for the result. Even when we run the tests on less persistent series of the balance of payments, we still get rejections of the hypothesis of the first-digit distribution being Benford at a 5 % level–except for the full sample for the fixed exchange rate regimes that have more liquid equity markets or the entire unconditioned group. This is also a strong indication that the rejections may not be due to the persistence of the data. We also run the tests excluding the entries on the reserve assets, which in fixed exchange rate regimes may experience large movements due to the readiness of the central bank to buy and sell currency at a prespecified rate. It does not seem thus that the unusual distribution of the first digits comes from the activity in this entry. The 1-in-5 and 1-in-20 criteria tell the same story as the 1-in-10 criterion. Taking different quantiles for the conditions preserves all of the results for the top 90, 80, and 66.66 percents of the conditions and most for the 50 percent (however, we start having few observations here for some conditions). The picture does not change when we condition on contemporaneous year values of net foreign asset/GDP or liabilities/GDP ratios.

These results contrast strongly with similar tests for the floating exchange rates regimes, for which the null hypothesis cannot be rejected even for one category, also when we use the same conditions on the investment positions of a country. Notice also that the number of countries with floating regimes included in each sample is always lower than for the fixed regimes in the same category; again this may be an indication that the persistence issue does not matter here.

It seems odd that our results would occur only due to the specificity of the economic processes under a fixed exchange rate regime. The fact that some items of the balance of payments (like changes in reserves) may be more variable in a fixed exchange rate regime should not matter as long as the processes generating the data are well behaved and properties F1–F3 are preserved. In a fixed exchange regime, there is still inflation and growth that would lead the balance of payments data to evolve according to a process that leads to Benford's law; and we still aggregate across many countries. In fact, some fixed exchange regimes country groups do exhibit first digit distributions for which the hypothesis that these are drawn from Benford's distribution cannot be rejected. These are OECD countries (as of 1989) or countries that have high Institutional Investor ratings (above 68), which would be one of the indicators according to Reinhart and Rogoff (2009) placing them in the group of advanced economies (results not shown). Another question thus appears: can this be driven by the fact that fixed exchange rate regimes are more often adopted by less developed countries that for example score badly in terms of the quality of institutions? This again seems unlikely, as we obtain changes in the strength of rejections (lower P–values, higher 1-in-10 criteria) when adding further conditions. In particular, we obtain strongest rejections for the category of countries with fixed regimes and most internationally open capital markets, which is typically the feature of more developed countries. Consistent rejections of the null hypothesis across categories also for the less persistent series for countries with fixed exchange rate regimes also indicate that this is a non-issue, as for most categories of countries with low quality of institutions the hypothesis of the first digits conforming to Benford's distribution cannot be rejected for these series, as is discussed in Section 5.5.

It thus appears that among countries with fixed exchange rate regimes there are some which strategically misinform investors as the distribution of the first digits of the balance of payments data they report is unusual. This is true especially for countries which may face higher outflows due to the stock of liabilities they owe to the outside world. It becomes even more so for countries which are de facto open in terms of capital flows (and have higher equity/GDP liabilities than other countries). There is economic rationale why this may be the case: in a fixed exchange rate regime the aggregate information contained in the behavior of the exchange rate is missing, and investors need to rely more on the information provided by the country, which may want to tweak it in order to put itself in a more favorable light.

5.2 Countries with high negative foreign asset liabilities

We discuss here tests on whether countries with higher net foreign asset positions misreport their economic data.

In Tables X–XI we show the P–values of tests for different centiles of the data when countries are ranked according to the ratios of net foreign assets excluding foreign direct investment (FDI) stocks to GDP for the full and the stable samples. This is the most apt measure of the capital stock that could rapidly move out of a country; FDI flows, which are largely immobile in the short run, are excluded, and we take the net value of assets. We use the previous year's value for the current year quarters to determine in which group a country should be placed in a particular quarter.

We see that for the first digits of the balance of payments data for the 10 %, 20 %, and 25 % of countries with the worst net foreign asset position (excluding FDI) in the stable sample the null hypothesis is rejected at the 5 % significance level, with the 1-in-10 criterion confirming this. We also obtain strong evidence on the full sample but only for the top 10 % most indebted countries with liquid liabilities.

The results described above survive several robustness checks while some other conditions that lead to rejections and are not meaningful from the point of view of economics do not (these checks are reported in Tables XII–XIII). First, for the complement of the sets defined by conditions we cannot reject the null hypothesis of conformity to Benford's law. With the exclusion of one or more countries we still reject Benford's distribution for the previously flagged categories as indicated by the stability indexes. This is not the case for other categories (the 75 % *least* indebted countries with liquid liabilities on the full sample or the 75 % most indebted countries for the stable sample). In addition, we obtain a rejection of Benford's distribution for the first digits of the data for the 10 % most indebted countries on the stable sample also for the less persistent series. The 1-in-20 and 1-in-5 indicators give the same indication as the 1-in-10 criterion.

Let us comment further on the rejection of the null hypothesis for the 25 % of countries with the *highest* net foreign assets for the full sample, which is not confirmed further for the highest 10 % of such countries and which does not pass the stability check. This may be driven by one country that is providing suspect statistics (for example, at the upper 20th centile); our further investigations revealed that the exclusion of CIS countries (former Soviet Union) – some of which have small negative or even positive NFA positions – eliminates this feature of the data completely.

We conclude that countries that have a high ratio of liquid indebtedness/GDP provided, in contrast to others, balance of payments data which had an unusual distribution.

We do not get consistent (i.e., in terms of the observed patterns) and strong results for groups of countries created using other measures of investment positions. In particular, for the full sample we obtain rejections of Benford's distribution of the first digits also for countries with the highest net foreign asset position and the lowest liabilities/GDP, although we get similar results as for countries with net foreign assets (excluding FDI) for the stable sample.[24] These series may be less pertinent for our tests. For example, the fact that a country has a high total liability/GDP ratio, like Switzerland, may not per se be an indicator that a country is vulnerable if it is has high assets as well. It may also mean that we cannot condition the data well enough, for example we do not control for many features of countries (like differences in technology) which may affect the payoffs from misreporting. The data on liabilities may not be free of errors as well; and it is data that we have on a yearly basis only. Therefore, if the dataset heterogeneity is large, the picture may be blurred with these less precise measures.

5.3 Countries with negative current account balances

We report in Tables XIV–XV the results of tests for countries with negative current account positions in a given quarter.

The evidence here is weaker than in the case of countries with fixed exchange regimes. By taking into account the P–values and the 1-in-10 criterion we observe that in both the full (P–value of 5.6 % and 1-in-10 criterion indicating rejection at less than a 2 % significance level) and the stable sample (P–value of 0.2 % and 1-in-10 criterion at a 1 % significance level) we obtain a rejection of the conformity of the first digits of the balance of payments data to Benford's distribution for

[24] Available upon request.

countries that have negative current account positions in a given quarter. 99 out of 103 countries at some point in time have a negative current account figure (63 out of 66 for the stable sample); this shows that the rejection of this condition is not due to some country-specific data dissemination practices but rather to what is reported when the country in question is borrowing from the rest of the world in a quarter. If we include only the non-OECD countries, this picture is even stronger.

The same is true when we additionally select countries with the highest level of net indebtedness or highest liabilities that are running negative current accounts. The only exception are the countries with the highest equity/GDP liabilities for the full sample, where we cannot reject the null hypothesis at conventional values.

The robustness tests are presented in Tables XVI–XVII. They raise a couple of question marks about the robustness of some of our findings. For the main condition – negative current account position in a quarter – after excluding countries one-by-one from the sample we cannot reject Benford's distribution for the first digits of the data for the full sample, but we can do so for the stable sample. For the complement of the above conditions, we can reject the null hypothesis at a 5 % significance level for the full sample, but this rejection is mostly driven by data from one country only. It seems that some additional conditioning (for example, choosing countries with high total foreign liabilities/GDP or liabilities/GDP ratios) have in fact little importance in characterizing countries with unusual first-digit distributions, as we reject the null hypothesis for the complement of the set. The less persistent series for the full sample for all categories seem to conform to Benford's distribution, which is not the case for the stable sample, for which some rejections of the null hypothesis on all series are confirmed. What is also important, we reject the null hypothesis unequivocally for countries with negative current account balances that also have lower liquid net foreign assets (i.e., NFA excluding FDI) for either of the samples.

Overall, the rejections we find seem to indicate that countries requiring the financing of their borrowing (as their current account balances are negative) may be willing to misinform investors, especially among the countries in the stable sample.

We obtained no consistent results when the scope of the current account position was taken under scrutiny (akin to the exercise in Section 5.2). The difference is, however, that we could obtain current account/GDP figures on a yearly basis only whereas above we could investigate quarterly evidence based just on the sign of the current account balance.

5.4 Crisis time evidence

Our model tells us that the urge to misinform investors would be higher in the time when many countries have bad states of productivity or, in other words, during a world crisis. To investigate this, we use the stable sample and yearly data, because we want to have globally roughly the same number of observations for each studied time period. We are more constrained here in the number of available observations and the conditions that we impose on the data than in the case of other tests. We believe on the other hand, however, that persistence is not of a problem in these tests as the number of consecutive quarters is small but many countries are involved (cf. Section 4.3.1).

An important issue is to determine time periods when there would be a time of a crisis touching a wide group of countries so that our statistical procedure may have a chance to detect some deviations. There is no easy way to determine and define "crisis" years, although recent attempts were made for example in Reinhart and Rogoff (2009) and Barro and Ursúa (2009). In the time period 1995–2007 Reinhart and Rogoff (2009) claim that there was one global financial crisis in the years 1997–1998 that originated in Asia and then spread across the globe. We present thus the evidence for 1997 in Table XVIII. No results were found for 1998. An interpretation of this may be that in 1998 it was already clear that the crisis is going to affect many countries; the possibilities of misinforming investors were thus lower. We note that the last balance of payments figures for the last quarter of 1997 were published in early 1998, so any deliberate alterations might be in part driven by the events in early 1998. Therefore the data for 1997 in fact covers well the first part of the global crisis of 1997–1998.

The evidence for 1997 shows that countries open to capital flows (also those with lower net foreign assets, higher total liabilities and with fixed exchange rates) have rejections of Benford's distribution for their first-digit data at levels lower than a 5 % level. This is what we would expect for our model in Section 2: the benefit of embellishing statistics would be higher when many countries would be in a bad state of the world, and for those that would be open to capital flows. We consider these results encouraging, but similar patterns appear for many categories for 1996 and 2006; without comprehensive further research it is difficult to determine whether these were years with large shocks to many countries across the world that would cause such rejections. [25, 26]

5.5 Quality of institutions and data provision

We also study whether institutional quality affects the truthfulness of data provided. For this purpose, we use various World Governance Indicators from the World Bank and report the results in Table XIX (main results for all indicators) and Table XX (main results and robustness checks only for the corruption indicator). Detailed tables with robustness checks computed for all indicators are omitted from the main text but are available upon request.

Both the P–values from the goodness of fit tests and the 1-in-10 criterion indicate that countries ranked below the 50th percentile (and below the 25th as well) in the WGI dataset in terms of corruption, political stability and the rule of law and below the 25th percentile in terms of government effectiveness (in terms of the global WGI dataset) exhibit balance of payments data with first digits that have non-Benford behavior. It appears thus that countries with poor institutions provide data of low quality. This is a finding similar to Nye and Moul (2007) who show that in samples of economic data from the Penn World Tables for non-OECD countries the hypothesis of Benford's distribution of the first digits is rejected (but remember that for our BOPS data this is not the case for the non-OECD countries, as witnessed by Table III). For rankings in terms of voice and accountability and regulatory quality we obtained no consistent and interesting results.

For all the categories flagged above the rejection of Benford's distribution for the complement of the set (i.e., for countries with better institutional rankings) of the first-digits fails. The exclusion of one of the countries from the samples does not drive the results mentioned above, except for the lower 25th percentile of government effectiveness on the stable sample. For tests ran on less persistent series, we can reject Benford's distribution at the 10 % significance level for the countries ranking lowest 25 % in terms of government effectiveness and political stability for both samples, and for those ranking most corrupt for the stable sample; the evidence based on these is weak. The 1-in-20 and 1-in-5 criteria give rejections for the same categories, albeit sometimes at different conventionally accepted significance levels, as the 1-in-10 criterion.

Our results lead us to believe that countries with poor institutions produce balance of payments data of questionable quality. However, we do not have a story why such countries would manipulate the data. The answer may be that is not due to bad data collection procedures or methodologies. Firstly, even high measurement errors should in practice cause the data to have first digits obeying Benford's law if the errors are well behaved, in view of the theoretical foundation F1. Secondly, when we look at the tests based on the less persistent series, there is hardly any evidence that countries with bad institutions do have weird first-digit distributions. If the rejection of Benford's distribution would be caused by bad data collection, the first-digit distribution of the less persistent series should be also affected. A better story may be that the institutions in these countries have lower scruples or fewer control mechanisms (so that rightly they are ranked as having bad institutional quality) that prevent data falsification.

5.6 Geographical groups

The goodness-of-fit tests of conformity of the first-digit distributions with Benford's distribution broken down geographically are shown in Table XXI (main results and robustness checks). Here we need to rely heavily on the 1-in-10 criterion as the persistence issue may be severe. We find that countries in Africa and the Middle East (grouped together in order to have a sufficient large

[25] Inspecting for example global stock market indexes for emerging markets one can observe substantial falls (of over 20 % over some periods) in 2006. But this does not indicate how many countries were implicated and whether this would in fact constitute a global and substantial shock.

[26] We do not investigate quarterly data; the reason for this is threefold. First of all, we have relatively few observations for each quarter. Secondly, any way of creating additional groups is going to yield even fewer observations. Not less importantly, it is difficult to know when the results for a given quarter were published. It may well be that the same underlying data would be reported differently if the following quarter (or quarters, depending on the time it takes for a country to publish data) were characterized for example by turbulence in the financial markets or not.

number of observations) have suspect distributions of the first digits both in the full and the stable samples. Latin American countries in the stable sample, i.e., that provided data at least every year between 1995 and 2007, also have a distribution of first digits for which we can reject the hypothesis that it was drawn from a Benford's distribution. The 1-in-10 criterion flags also the rejection of the null hypothesis at 5 % for the stable sample and at 10 % for the full sample for East European countries. These findings should be taken with caution, however, as there are relatively few countries in each group with many quarters of data each, which may make rejecting the null hypothesis easier (see Section 4.3 for discussion). Tests on the least persistent series do not confirm any of the findings. This may mean that it so happens that African, Middle Eastern and Latin American (and potentially East European) countries have some economic series whose first digits are very persistent (i.e., evolve slowly) when termed in U.S. dollars, and these drive the rejections of Benford's distribution. It may not be a primary feature of these countries that they provide bad quality data on purpose.[27]

5.7 Other conditions for defining the subsets of the data

We have tried various other conditioning of the data in our search of unusual reporting patterns.

First of all, we studied (akin to Section 5.2) whether countries with more capital openness (defined as high foreign equity liabilities/GDP ratios) exhibit non-Benford distribution of their first digits of the balance of payments data. Scrutinizing the tails of data sorted according to this measure, we find no results that this is the case.

Next, for countries that joined the Euro zone we inspected a period in which they had to maintain a fixed exchange rate regime and satisfy certain economic performance criteria for admission. We do not find any evidence that the balance of payments data is unusual for up to 5 years prior to joining the zone.

We also investigated countries with episodes of sovereign default and banking crises (data on their timing obtained from Reinhart and Rogoff [2009]). For countries that defaulted on their sovereign debt within our sample period, we do not find rejections of the null hypothesis either for one, two or three year brackets preceding a default. However, we find evidence that within one year after enduring banking crises countries report suspicious distributions of the balance of payments data. For these, we can reject the hypothesis that the first digits conform to Benford's distribution at 0.2 % with 5,094 observations on the full sample and 5.6 % for the stable sample (2,487 observations).

There seems not to be any evidence that countries suffering substantial falls in their Institutional Investor country credit ratings, either contemporaneous, lead or lagged, offer suspect statistics.

5.8 Summary of empirical results

We find results that support the hypothesis that countries strategically provide manipulated data to economic agents. We observe rejections of Benford's distribution for the first digits of data issued by groups of countries that on average are more vulnerable to high capital outflows. These rejections are rather category than country specific, as data from many countries for different quarters enters different categories. To our surprise, we obtain the strongest (i.e., most robust) results for countries with fixed exchange rate regimes though we also find that countries with highest levels of net indebtedness and those that were running current account deficits have first digits that have unusual, non-Benford distributions. These findings confirm the viability of the intuition developed by the simple model in Section 2. However, we do not find general results for countries that were de facto more open to capital flows. This may be due to the fact that we are unable to capture the extent of their vulnerability very finely; but for some subgroups of countries (for example those with fixed exchange rate regimes and relatively high de facto openness) we still get very strong rejections of the null hypothesis. We also find weak crisis time evidence, pointing that countries may engage in data falsification in stress times. The interesting finding on the fixed exchange rate regime countries may show that the public is more readily misinformed

[27] This is also to be put in perspective with the results of Nye and Moul (2007) who indicate that imprecise transformations of the data alter its quality while we suspect it might be only due to the more severe persistence of data when termed in U.S. dollars.

by governments when there is a larger scope for misinformation – for example when the market prices aggregating private information are missing. We find also evidence that countries with weak institutions provide data that are non-Benford distributed.

6 Conclusions

In this paper we took a glimpse at the dark side of the moon of government statistics. We conclude that country–quarter pairs that correspond to economic situations in which countries would have higher incentives to misinform investors lead indeed to different distributions of the corresponding first digits of the balance of payments data than Benford's distribution, while for other country–quarter pairs the data conforms to this distribution. This may of course be because there is a different underlying process that rules the generation of balance of payments distributions in country–quarter pairs for vulnerable countries, which we think is unlikely. It may be, however, due to a simple fact: such countries will falsify their balance of payments data in these quarters. This partial evidence on the strategic character of misreporting points out that models in which governments emit public signals (for example, the discussion on central bank transparency) should seriously consider the possibility that this signal may be at times intentionally misleading. On the policy side, this paper calls for the need to establish independent statistical agencies akin to that of independent central bankers. Some countries that failed in having such an independent agency, like Argentina or Greece, were caught red handed in altering economic data that they disseminated to the public. The possible welfare implications of misinformation should also be investigated; the actions of Argentina or Greece show that economic gains to governments, even if short-run, may exist and be substantial.

References

Angeletos, George-Marios and Alessandro Pavan, "Efficient Use of Information and Social Value of Information," *Econometrica*, 2007, *75* (4), 1103–1142.

_ , Christian Hellwig, and Alessandro Pavan, "Signaling in a Global Game: Coordination and Policy Traps," *Journal of Political Economy*, 2006, *114* (3), 452–484.

Aumann, Robert J., Michael B. Marschler, and Richard E. Stearns, *Repeated Games With Incomplete Information*, MIT Press, 1995.

Barro, Robert J. and José F. Ursúa, "Stock-Market Crashes and Depressions," Working Paper 14760, National Bureau of Economic Research 2009.

Benford, Frank, "The law of anomalous numbers," *Proceedings of the American Philosophy Society*, 1938, *78*, 551–572.

Crawford, Vincent and Joel Sobel, "Strategic Information Transmission," *Econometrica*, 1982, *50*, 1431–1451.

Diaconis, Persi, "The distribution of leading digits and uniform distribution mod 1," *The Annals of Probability*, 1977, *5*, 359–364.

Heinemann, Frank and Gerhard Illing, "Speculative attacks: unique equilibrium and transparency.," *Journal of International Economics*, 2002, *58*, 429–450.

Hill, Theodore P., "Base-invariance implies Benford's law," *Proceedings of the American Mathematical Society*, 1995, *123*, 887–895.

_ , "A statistical derivation of the significant-digit law," *Statistical Science*, 1995, *10*, 354–363.

IMF, International Monetary Fund, "Annual report on exchange arrangements and exchange restrictions," Technical Report 1989–2008.

_ , "Balance of Payments Manual," Technical Report 2005.

Janvresse, Elise and Thierry de la Rue, "From uniform distributions to Benford's law," *Journal of Applied Probability*, 2004, *41*, 1203–1210.

Kaufmann, Daniel, Aart Kraay, and Massimo Mastruzzi, "Govenance Matters VIII: Aggregate and Individual Governance Indicators, 1996–2008," Research Working Paper 4978, World Bank Policy 2009.

Kydland, Finn E. and Edward C. Prescott, "Rules rather than discretion: The inconsistency of optimal plans," *Journal of Political Economy*, 1977, *85*, 473–491.

Lane, Phillip and Gian-Maria Milesi-Ferretti, "The External Wealth of Nations Mark II: Revised and Extended Estimates of Foreign Assets and Liabilities, 1970–2004," *Journal of International Economics*, 2007, *73*, 223–250.

Morris, Stephen and H.S. Shin, "The Social Value of Public Information," *American Economic Review*, 2002, *92*, 1521–1534.

Newcomb, Simon, "Note on the frequency of use of the different digits in natural numbers," *American Journal of Mathematics*, 1881, *4*, 39–40.

Nigrini, Mark, "Taxpayer compliance application of Benford's law," *Journal of the American Taxation Association*, 1996, *18*, 72–92.

_ , "I've got your number: How a mathematical phenomenon can help CPAs uncover fraud and other irregularities," *Journal of Accountancy*, May 1999.

Nye, John and Charles Moul, "The political economy of numbers: on the application of Benford's law to international macroeconomic statistics," *The B.E. Journal of Macroeconomics*, 2007, *7* (1), article 17.

Obstfeld, Maurice, "Rational and self-fulfilling balance of payments crises.," *American Economic Review*, 1986, *76*, 72–81.

— , "Models of Currency Crises with Self-Fulfilling Features.," *European Economic Review*, 1996, *40*, 1037–1047.

Pinkham, Roger S., "On the distribution of first significant digits," *The Annals of Mathematical Statistics*, 1961, *32*, 1223–1230.

Raimi, Ralph A., "The peculiar distribution of first digits," *Scientific American*, December 1969, pp. 109–119.

Reinhart, Carmen and Kenneth Rogoff, *This Time is Different*, Princeton University Press, 2009.

Sandleris, Guido, "Sovereign defaults: Information, investment and credit," *Journal of International Economics*, 2008, *76*, 267–275.

Sbracia, Massimo and Andrea Zaghini, "Expectations and information in second-generation currency crises models," *Economic Modelling*, 2001, *18*, 203–222.

Stein, Jeremy C., "Cheap Talk and the Fed: A Theory of Imprecise Policy Announcements," *American Economic Review*, 1989, *79* (1), 32–42.

Tarashev, Nikola A., "Speculative Attacks and the Information Role of the Interest Rate," *Journal of the European Economic Association*, 2007, *5*, 1–36.

Varian, Hal R., "Benford's Law," *The American Statistician*, 1972, *26*, 65–66.

Newspaper articles

Financial Times (2001, via FT.Com). R. Bream, J. Blitz and P. Norman. "Doubts over way Italy qualified for euro." 4 November 2001.

Financial Times (2006). Christopher Condon. "The Speech that sparked rioting." 20 September 2006.

Financial Times (2007). Stefan Wagstyl. "Lies haunt a reformer's grip on power." 25 October 2007.

Financial Times (2008). J. Webber. "As prices surge, Argentines cry foul over the official figures." 9 June 2008.

Financial Times (2009). Bernie McSherry. "Investors have to be sure that statistics do not lie." 16 July 2009.

Financial Times (2010). T. Barber and K. Hope. "Brussels attacks Greece over false data." 13 January 2010.

The Wall Street Journal (2004). J. W. Miller. "Obscure EU statistics bureau may be getting a higher profile." 10 November 2004.

The Wall Street Journal (2008, online version). M. O'Grady. "Argentina Impoverishes Itself Again; What happens when government meddles with private wealth." 3 November 2008.

A Countries included in the study

Albania, Argentina, Armenia, Australia, Austria, Azerbaijan, Bangladesh, Belarus, Belize, Bolivia, Brazil, Bulgaria, Cambodia, Canada, Cape Verde, Chile, Colombia, Costa Rica, Croatia, Cyprus, Czech Republic, Denmark, Ecuador, El Salvador, Eritrea, Estonia, Ethiopia, Finland, France, Georgia, Germany, Greece, Guatemala, Guinea, Honduras, Hungary, Iceland, India, Indonesia, Ireland, Israel, Italy, Japan, Jordan, Kazachstan, Korea, Kirghiz Republic, Laos, Latvia, Lebanon, Lesotho, Lithuania, Macedonia (former Yugoslav Republic of), Madagascar, Malaysia, Malta, Mauritius, Mexico, Moldova, Mongolia, Morocco, Mozambique, Myanmar, Namibia, Nepal, Netherlands, New Zealand, Nicaragua, Nigeria, Norway, Pakistan, Panama, Papua New Guinea, Paraguay, Peru, Philippines, Poland, Portugal, Romania, Russia, Serbia, Singapore, Slovakia, Slovenia, South Africa, Spain, Sri Lanka, Sudan, Suriname, Sweden, Switzerland, Tajikistan, Thailand, Turkey, Turkmenistan, Uganda, Ukraine, United Kingdom, United States, Uruguay, Venezuela, Vietnam, Yemen (Republic of), Zimbabwe

B Proof of Proposition 1

First, note that (5) indicates that $K\big(h, n'(h)+1\big) > K\big(\ell, n'(h)\big)$, no matter the value of $n(h)$; indeed,

$$K\big(\ell, n'(h)\big) \leqslant \frac{\overline{K} + NC_1 + C_{N+1}}{N+1} \leqslant K\big(h, n'(h)+1\big)\,,$$

where the first inequality is an equality if and only if $n'(h) = 0$, whereas the second one is an equality if and only if $n'(h) = N$. We denote in the sequel $\gamma = K\big(\ell, n'(h)\big)/K\big(h, n'(h)+1\big) < 1$ (and omit the indexation of γ in $n'(h)$ for the sake of simplicity).

The term (7) after substitutions equals

$$\begin{aligned}
& A(\ell)\Big((1-\alpha)\Big(k\big(h,n'(h)+1\big)\Big)^{\alpha} + \alpha\Big(k\big(h,n'(h)+1\big)\Big)^{\alpha-1} C_m\Big) \\
- \;& A(\ell)\Big((1-\alpha)\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha} + \alpha\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha-1} C_m\Big) \\
= \;& A(\ell)\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha}\left((1-\alpha)(1-\gamma^{\alpha}) + \alpha\frac{C_{N+1}}{k\big(\ell,n'(h)\big)}(1-\gamma^{\alpha-1})\right) \qquad (12) \\
\geqslant \;& A(\ell)\Big(k\big(\ell,n'(h)\big)\Big)^{\alpha}\big((1-\alpha)(1-\gamma^{\alpha}) + \alpha\gamma(1-\gamma^{\alpha-1})\big) \qquad (13) \\
> \;& 0\,,
\end{aligned}$$

where the inequality from (12) to (13) is because $C_{N+1} < C_1 \leqslant K\big(\ell, n'(h)\big)$ and $L_{N+1} = 1$ while $\alpha \in (0,1)$. The final inequality indicating that the difference of payoffs is positive comes by function study. To show that the function $t \in (0,1) \mapsto (1-\alpha)(1-t^{\alpha}) + \alpha t(1-t^{\alpha-1}) = 1 - \alpha - t^{\alpha} + \alpha t$ is positive for all $\alpha \in (0,1)$ it suffices to note that it tends to 0 as $t \to 1$ and that its derivative with respect to t equals $\alpha(1-t^{\alpha-1}) < 0$.

C Tables and figures

They are reported in the following pages (one table per page).

Variables	Full sample					Stable sample				
	Observations	Mean	Std. Dev.	Min	Max	Observations	Mean	Std. Dev.	Min	Max
First digit of BOP figure	69,287	3.43	2.46	1	9	41,245	3.40	2.45	1	9
Stable sample = 1	69,287	0.60		0	1					
Fixed exchange rate = 1	69,287	0.46		0	1	41,245	0.38		0	1
Floating exchange rate = 1	69,287	0.28		0	1	41,245	0.36		0	1
Debt liabilities in year (-1) / GDP; see Lane and Milesi-Feretti	67,881	0.73	0.67	0.04	26.88	40,583	0.76	0.67	0.06	6.94
Equity liabilities in year (-1) / GDP; see Lane and Milesi-Feretti	69,287	0.10	0.29	0	4.27	41,245	0.13	0.34	0	4.27
Liabilities excl. FDI in year (-1) / GDP; see Lane and Milesi-Feretti	67,837	0.83	0.87	0.06	26.88	40,583	0.90	0.94	0.06	11.21
Net foreign asset position in year (-1) / GDP (-1); see Lane and Milesi-Feretti	68,285	1.14	1.09	0.11	27.21	40,677	1.27	1.21	0.20	12.95
Net foreign asset position excl. FDI in year (-1) / GDP (-1); see Lane and Milesi-Feretti	67,837	-0.39	0.55	-25.66	1.96	40,583	-0.37	0.51	-3.94	1.96
Liabilities in year (-1) / GDP; see Lane and Milesi-Feretti	68,285	-0.22	0.52	-25.34	2.51	40,677	-0.18	0.47	-2.99	2.51
Current Account balance positive in the quarter = 1	69,204	0.35		0	1	41,209	0.37		0	1
WGI control of corruption (rank)	50,372	56.39	29.24	0.49	100	38,088	59.94	29.91	0.49	100
WGI government effectiveness (rank)	50,460	60.04	27.23	1.90	100	38,176	63.57	27.55	1.90	100
WGI political stability (rank)	50,460	51.18	29.20	0.96	100	38,176	53.34	30.14	0.96	100
WGI rule of law (rank)	50,460	56.22	29.01	1.43	100	38,176	60.12	29.39	2.86	100
WGI regulatory quality (rank)	50,460	60.14	27.03	0	100	38,176	64.09	26.92	0	100
WGI voice and accountability (rank)	50,460	58.61	27.31	0	100	38,176	62.20	26.85	0	100
Institutional Investor country ranking	64,999	51.27	25.78	4.5	96.4	39,334	55.18	26.30	6	96
OECD member prior to 1989 = 1	69,287	0.29		0	1	41,245	0.32		0	1
Africa and Middle East = 1	69,287	0.14		0	1	41,245	0.10		0	1
CIS member = 1	69,287	0.12		0	1	41,245	0.14		0	1
Eastern Europe = 1	69,287	0.15		0	1	41,245	0.17		0	1
Developed Europe + Western offshoots + Japan = 1	69,287	0.29		0	1	41,245	0.32		0	1
Latin America = 1	69,287	0.16		0	1	41,245	0.15		0	1
Asia excluding Middle East and CIS = 1	69,287	0.16		0	1	41,245	0.17		0	1

The indications of = 1 flag variables with values in {0,1}; the value 1 indicated that the given condition is fulfilled. The reported means then correspond to proportions and no standard deviation is reported.

Table II: Data description for the full and stable samples.

	Number of countries and quarters involved		All series		Less persistent series		Rejection rates (in %)		
	Number of countries	Avg. quarters	P-value (in %)	Number of observations	P-value (in %)	Number of observations	1-in-20 criterion	1-in-10 criterion	1-in-5 criterion
Full sample	103	57.2	64.4	69,287	66.4	30,075	6.5	6.1	5.4
Stable sample (countries having observations each year after 1995)	66	51.9	14.6	41,245	14.8	18,066	5.5	6.9	8.5
OECD countries	22	74.0	92.9	20,105	94.2	8,970	3.7	3.7	2.9
Non-OECD countries	81	52.6	25.0	49,182	76.1	21,105	6.1	5.6	5.8

Table III: Assessment of the global conformity of the dataset to Benford's law.

	Number C of countries chosen at random								
	C = 1	C = 10	C = 20	C = 30	C = 40	C = 50	C = 60	C = 70	C = 80
Rejection rates	43.1%	39.8%	35.4%	32.4%	29.6%	24.5%	16.1%	14.0%	7.8%

All series / Less persistent series	C = 20 countries	C = 40 countries	C = 60 countries	C = 80 countries
M = 10 quarters	11.2% /	9.6% /	9.2% /	6.2% /
	5.9% /	5.8% /	5.6% /	4.3% /
M = 20 quarters	16.2% /	14.3% /	9.6% /	7.3% /
	7.0% /	6.7% /	5.0% /	3.8% /
M = 30 quarters	25.0% /	16.9% /	11.9% /	8.5% /
	8.2% /	6.3% /	4.0% /	2.1% /
M = 40 quarters	27.7% /	22.0% /	12.9% /	7.0% /
	8.5% /	8.3% /	5.3% /	1.8% /
All quarters	35.4% /	29.6% /	16.1% /	7.8% /
	11.9% /	5.9% /	2.4% /	0.3% /

Table IV: *Top table*: Rejection rates at the 5 % level of the hypothesis of conformity to Benford's law on data subsets corresponding to given numbers C of countries chosen at random; estimated by performing 1,000 random draws for each value of C. *Bottom table*: Rejection rates at the 5 % level of the hypothesis of conformity to Benford's law on data subsets corresponding to given numbers C of countries and quarters M thereof, all chosen at random, for two different sets of series; estimated by performing 1,000 random draws for each value. The last but one line corresponds to entries in the top table.

Proportion f = 5%	Estimated quantiles (values given in %)			
	90%	95%	98%	99%
N = 2,500	5.3	5.7	6.2	6.5
N = 5,000	6.2	6.7	7.5	7.8
N = 10,000	6.3	6.7	7.4	7.9
N = 30,000	6.3	6.7	7.2	7.6
N = 50,000	6.2	6.6	7.2	7.6
	Conclusion: upper bounds			
N ≥ 5,000	6.3	6.7	7.5	7.9

Proportion f = 10%	Estimated quantiles (values given in %)			
	90%	95%	98%	99%
N = 1,200	5.5	6.3	7.2	8.0
N = 2,000	7.0	7.9	9.1	9.5
N = 5,000	6.9	7.8	9.0	9.8
N = 10,000	6.9	7.5	8.7	9.9
N = 30,000	6.8	7.5	8.7	9.5
N = 50,000	6.8	7.7	9.0	9.7
	Conclusion: upper bounds			
N ≥ 2,000	7.0	8.0	9.0	10.0

Proportion f = 20%	Estimated quantiles (values given in %)			
	90%	95%	98%	99%
N = 1,000	8.7	10.4	12.8	14.2
N = 2,000	8.2	9.3	11.6	13.2
N = 5,000	8.2	10.5	13.4	14.5
N = 10,000	8.5	10.2	12.6	14.3
N = 30,000	8.6	10.9	13.0	14.1
N = 50,000	8.4	9.7	11.8	13.3
	Conclusion: upper bounds			
N ≥ 1,000	8.7	10.9	13.4	14.5

Table V: Estimates of the quantiles of the distribution of the random variable $\overline{R}_{D,1000}$ when D is chosen at random according to Benford's law and f equals 5 % (top table), 10 % (middle table), and 20 % (bottom table); estimates were obtained by drawing D at random 1,000 times and computing the associated values of $\overline{R}_{D,1000}$.

Economic conditions		Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1- in-10 criterion
	Stable sample					
Floating exchange		39	30.9	15,003	37.6	5.8
Floating exchange	& OECD	19	40.6	9,742	17.3	5.7
Floating exchange	& Non-OECD	20	21.6	5,261	57.7	6.4
Fixed exchange		49	26.6	15,488	0.5***	13.3***
Fixed exchange	& OECD	12	19.2	2,881	31.1	5.5
Fixed exchange	& Non-OECD	37	29.0	12,607	2.5*	11.2***
	NFA_EXCL_FDI_GDP ≤ - 0.0201	59	43.0	30,480	3.2*	8.4*
Floating exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0201	34	27.8	11,702	35.4	5.6
Fixed exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0201	44	21.0	11,099	0.9***	10.5***
	NFA_GDP ≤ - 0.1233	58	43.9	30,427	3.3*	9.5**
Floating exchange	& NFA_GDP ≤ - 0.1233	35	24.6	10,619	71.3	5.2
Fixed exchange	& NFA_GDP ≤ - 0.1233	42	23.1	11,569	0.2***	11.9***
	EQ_LIAB_GDP > 0.0098	51	43.5	27,339	72.0	4.1
Floating exchange	& EQ_LIAB_GDP > 0.0098	31	35.1	13,631	77.9	3.6
Fixed exchange	& EQ_LIAB_GDP > 0.0098	38	17.8	8,316	0.00004***	21.7***
	LIAB_EXCL_FDI_GDP > 0.4349	62	40.5	30,503	3.2*	7.5
Floating exchange	& LIAB_EXCL_FDI_GDP > 0.4349	37	26.6	12,303	65.1	5.7
Fixed exchange	& LIAB_EXCL_FDI_GDP > 0.4349	44	21.1	11,219	0.3***	10.6***
	LIAB_GDP > 0.6102	59	42.5	30,527	10.3	6.5
Floating exchange	& LIAB_GDP > 0.6102	37	26.0	12,050	89.9	5.1
Fixed exchange	& LIAB_GDP > 0.6102	40	22.5	10,950	0.6***	10.6***
	CA < 0	63	34.3	25,849	0.2***	10.9***
Floating exchange	& CA < 0	35	19.6	8,413	5.2	7.5
Fixed exchange	& CA < 0	45	20.4	10,958	0.007***	15.6***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table VI: Results for fixed and floating exchange rate regime countries (main results, stable sample).

Economic conditions		Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1- in-10 criterion
	Full sample					
Floating exchange		55	28.9	19,545	81.9	4.6
Floating exchange	& OECD	21	46.6	12,260	31.8	5.3
Floating exchange	& Non-OECD	34	17.9	7,285	40.8	6.4
Fixed exchange		84	32.6	31,736	0.7***	11.0***
Fixed exchange	& OECD	16	35.9	6,958	13.6	4.8
Fixed exchange	& Non-OECD	68	31.9	24,778	0.6***	13.2***
	NFA_EXCL_FDI_GDP ≤ - 0.0347	93	46.3	50,281	18.1	5.5
Floating exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0347	46	26.2	14,766	68.9	5.3
Fixed exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0347	73	27.2	22,845	0.04***	17.1***
	NFA_GDP ≤ - 0.1353	94	45.8	50,171	15.0	5.8
Floating exchange	& NFA_GDP ≤ - 0.1353	47	23.2	13,293	91.5	5.2
Fixed exchange	& NFA_GDP ≤ - 0.1353	72	28.4	23,505	0.007***	17.0***
	EQ_LIAB_GDP > 0.0061	77	45.8	43,043	73.8	4.4
Floating exchange	& EQ_LIAB_GDP > 0.0061	40	34.5	17,162	90.9	4.6
Fixed exchange	& EQ_LIAB_GDP > 0.0061	58	24.2	17,000	0.0007***	17.4***
	LIAB_EXCL_FDI_GDP > 0.4008	96	44.7	50,758	76.6	4.1
Floating exchange	& LIAB_EXCL_FDI_GDP > 0.4008	52	25.5	16,358	94.7	4.3
Fixed exchange	& LIAB_EXCL_FDI_GDP > 0.4008	73	26.3	22,380	0.5***	13.0***
	LIAB _GDP > 0.5427	98	44.1	51,359	50.2	6.7
Floating exchange	& LIAB _GDP > 0.5427	51	26.2	16,476	96.5	5.0
Fixed exchange	& LIAB _GDP > 0.5427	74	25.4	22,103	0.03***	13.8***
	CA < 0	99	38.7	45,022	5.6	9.2**
Floating exchange	& CA < 0	51	18.1	11,232	9.2	7.9
Fixed exchange	& CA < 0	81	23.6	22,216	0.1***	14.9***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table VII: Results for fixed and floating exchange rate regime countries (main results, full sample).

Economic conditions		Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	Excluding entry "Reserve Assets: net": P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
	Stable sample							
Floating exchange		37.6	---	58.1 (6,709)	35.0 (13,798)	4.9	5.8	5.1
Floating exchange	& OECD	17.3	---	69.7 (4,374)	16.0 (8,970)	5.5	5.7	5.7
Floating exchange	& Non-OECD	57.7	---	47.9 (2,335)	70.3 (4,828)	5.9	6.4	4.4
Fixed exchange		0.5***	0	3.6* (6,727)	0.6*** (14,185)	7.6**	13.3***	18.8***
Fixed exchange	& OECD	31.1	---	28.7 (1,311)	43.8 (2,651)	4.0	5.5	4.1
Fixed exchange	& Non-OECD	2.5*	4	13.9 (5,416)	2.6* (11,534)	8.0***	11.2***	14.3**
	NFA_EXCL_FDI_GDP ≤ - 0.0201	3.2*	1	37.6 (13,344)	7.4 (27,943)	7.2*	8.4*	10.3
Floating exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0201	35.4	---	91.9 (5,214)	40.6 (10,758)	6.0	5.6	7.1
Fixed exchange	& NFA_EXCL_FDI_GDP ≤ - 0.0201	0.9***	0	4.3* (4,823)	1.7** (10,174)	9.1***	10.5***	19.4***
	NFA_GDP ≤ - 0.1233	3.3*	2	41.1 (13,276)	7.7 (27,882)	6.4	9.5**	11.6*
Floating exchange	& NFA_GDP ≤ - 0.1233	71.3	---	77.9 (4,726)	66.2 (9,757)	4.9	5.2	4.1
Fixed exchange	& NFA_GDP ≤ - 0.1233	0.2***	0	2.5* (5,001)	0.3*** (10,600)	7.8**	11.9***	22.4***
	EQ_LIAB_GDP > 0.0098	72.0	---	33.9 (12,236)	76.3 (25,123)	3.8	4.1	2.9
Floating exchange	& EQ_LIAB_GDP > 0.0098	77.9	---	68.2 (6,148)	84.3 (12,544)	5.1	3.6	3.0
Fixed exchange	& EQ_LIAB_GDP > 0.0098	0.00004***	0	1.2** (3,724)	0.00007*** (7,639)	11.9***	21.7***	41.1***
	LIAB_EXCL_FDI_GDP > 0.4349	3.2*	1	9.8 (13,409)	8.4 (27,990)	7.1*	7.5	13.2*
Floating exchange	& LIAB_EXCL_FDI_GDP > 0.4349	65.1	---	93.7 (5,518)	61.6 (11,320)	5.4	5.7	4.5
Fixed exchange	& LIAB_EXCL_FDI_GDP > 0.4349	0.3***	0	0.6*** (4,889)	1.5** (10,290)	8.8***	10.6***	21.1***
	LIAB_GDP > 0.6102	10.3	---	12.7 (13,445)	26.0 (28,022)	5.7	6.5	7.0
Floating exchange	& LIAB_GDP > 0.6102	89.9	---	88.9 (5,402)	92.5 (11,088)	5.8	5.1	3.3
Fixed exchange	& LIAB_GDP > 0.6102	0.6***	0	1.6** (4,797)	2.9* (10,050)	9.1***	10.6***	16.5***
	CA < 0	0.2***	0	6.1 (11,219)	0.8*** (23,689)	8.0***	10.9***	20.1***
Floating exchange	& CA < 0	5.2	---	15.0 (3,711)	3.2* (7,727)	5.3	7.5	10.3
Fixed exchange	& CA < 0	0.007***	0	2.2* (4,731)	0.03*** (10,041)	10.4***	15.6***	29.2***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table VIII: Results for fixed and floating exchange rate regime countries (robustness tests, stable sample).

		Economic conditions	Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	Excluding entry "Reserve Assets: net": P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
		Full sample							
Floating exchange			81.9	---	39.3 (8,677)	83.1 (17,958)	4.8	4.6	3.5
Floating exchange	&	OECD	31.8	---	58.1 (5,460)	39.7 (11,281)	5.9	5.3	6.3
Floating exchange	&	Non-OECD	40.8	---	54.8 (3,217)	40.1 (6,677)	5.8	6.4	5.9
Fixed exchange			0.7***	0	24.6 (13,710)	0.4*** (29,003)	9.9***	11.0***	17.6***
Fixed exchange	&	OECD	13.6	---	43.3 (3,119)	10.8 (6,384)	5.2	4.8	6.9
Fixed exchange	&	Non-OECD	0.6***	0	37.4 (10,591)	0.8*** (22,619)	10.2***	13.2***	20.0***
		NFA_EXCL_FDI_GDP ≤ - 0.0347	18.1	---	96.2 (21,799)	14.8 (45,986)	6.4	5.5	7.0
Floating exchange	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	68.9	---	67.5 (6,541)	79.9 (13,562)	5.6	5.3	2.7
Fixed exchange	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	0.04***	0	2.4* (9,855)	0.03*** (20,868)	11.5***	17.1***	26.7***
		NFA_GDP ≤ - 0.1353	15.0	---	97.8 (21,682)	10.6 (45,874)	6.5	5.8	8.6
Floating exchange	&	NFA_GDP ≤ - 0.1353	91.5	---	66.1 (5,880)	77.9 (12,203)	6.3	5.2	3.6
Fixed exchange	&	NFA_GDP ≤ - 0.1353	0.007***	0	6.8 (10,097)	0.004*** (21,467)	14.1***	17.0***	33.5***
		EQ_LIAB_GDP > 0.0061	73.8	---	60.8 (19,145)	72.1 (39,515)	5.2	4.4	3.2
Floating exchange	&	EQ_LIAB_GDP > 0.0061	90.9	---	49.3 (7,691)	94.1 (15,784)	5.7	4.6	3.4
Fixed exchange	&	EQ_LIAB_GDP > 0.0061	0.0007***	0	18.2 (7,566)	0.0004*** (15,595)	11.5***	17.4***	36.8***
		LIAB_EXCL_FDI_GDP > 0.4008	76.6	---	38.9 (22,108)	90.3 (46,473)	5.1	4.1	3.5
Floating exchange	&	LIAB_EXCL_FDI_GDP > 0.4008	94.7	---	66.1 (7,272)	92.1 (15,033)	5.2	4.3	4.0
Fixed exchange	&	LIAB_EXCL_FDI_GDP > 0.4008	0.5***	0	1.9** (9,700)	1.4** (20,465)	9.3***	13.0***	17.3***
		LIAB _GDP > 0.5427	50.2	---	54.5 (22,364)	52.6 (47,044)	5.3	6.7	6.3
Floating exchange	&	LIAB _GDP > 0.5427	96.5	---	92.2 (7,323)	94.8 (15,140)	5.1	5.0	2.7
Fixed exchange	&	LIAB _GDP > 0.5427	0.03***	0	2.0* (9,578)	0.1*** (20,229)	9.8***	13.8***	25.7***
		CA < 0	5.6	3	46.4 (19,367)	11.0 (41,186)	6.6	9.2**	12.4*
Floating exchange	&	CA < 0	9.2	---	1.8** (4,918)	7.4 (10,310)	6.5	7.9	10.1
Fixed exchange	&	CA < 0	0.1***	0	2.8* (9,530)	0.3*** (20,306)	10.9***	14.9***	23.3***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table IX: Results for fixed and floating exchange rate regime countries (robustness tests, full sample).

Economic conditions NFA_EXCL_FDI_GDP		Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Stable sample						
≤ 10% quantile	= - 0.581	18	19.3	4,003	0.003***	14.4***
≥ 10% quantile	= - 0.581	64	47.1	36,580	31.8	5.6
≤ 20% quantile	= - 0.432	24	28.5	8,008	0.1***	10.0**
≥ 20% quantile	= - 0.432	63	42.5	32,575	39.3	6.0
≤ 25% quantile	= - 0.379	30	28.5	10,086	0.1***	9.9**
≥ 25% quantile	= - 0.379	62	40.4	30,497	40.7	7.0
≤ 50% quantile	= - 0.193	48	35.4	20,234	7.3	7.5
≥ 50% quantile	= - 0.193	46	36.1	20,349	35.3	5.2
≤ 75% quantile	= - 0.020	59	43.0	30,480	3.2*	8.4*
≥ 75% quantile	= - 0.020	30	27.5	10,103	37.8	5.2
≤ 80% quantile	= 0.015	60	45.0	32,494	6.7	7.7
≥ 80% quantile	= 0.015	29	22.8	8,089	52.1	5.5
≤ 90% quantile	= 0.144	63	48.1	36,527	8.5	6.9
≥ 90% quantile	= 0.144	15	21.9	4,056	15.8	5.7

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table X: Results for countries grouped by net foreign asset liabilities, excluding foreign direct investment (main results, stable sample).

Economic conditions NFA_EXCL_FDI_GDP		Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Full sample						
≤ 10% quantile	= - 0.678	24	20.1	5,277	0.5***	11.5***
≥ 10% quantile	= - 0.678	101	52.1	62,560	70.3	5.1
≤ 20% quantile	= - 0.471	39	27.6	12,089	45.5	6.3
≥ 20% quantile	= - 0.471	97	48.2	55,748	49.2	6.6
≤ 25% quantile	= - 0.413	44	31.1	15,504	62.6	5.7
≥ 25% quantile	= - 0.413	95	46.1	52,333	43.2	5.9
≤ 50% quantile	= - 0.199	71	40.2	32,894	14.7	8.0
≥ 50% quantile	= - 0.199	73	39.6	34,943	7.5	7.7
≤ 75% quantile	= - 0.035	93	46.3	50,281	18.1	5.5
≥ 75% quantile	= - 0.035	54	26.8	17,556	0.7***	9.7**
≤ 80% quantile	= 0.004	95	48.4	53,907	27.0	6.7
≥ 80% quantile	= 0.004	48	23.9	13,930	0.5***	10.6***
≤ 90% quantile	= 0.130	98	53.0	61,066	44.1	5.8
≥ 90% quantile	= 0.130	26	21.2	6,771	6.2	7.7

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XI: Results for countries grouped by net foreign asset liabilities, excluding foreign direct investment (main results, full sample).

Economic conditions NFA_EXCL_FDI_GDP		Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Stable sample							
≤ 10% quantile	= - 0.581	0.003***	0	1.3** (1,700)	9.0***	14.4***	31.3***
≥ 10% quantile	= - 0.581	31.8	---	13.9 (16,084)	6.1	5.6	5.8
≤ 20% quantile	= - 0.432	0.1***	0	42.9 (3,456)	6.9*	10.0**	21.3***
≥ 20% quantile	= - 0.432	39.3	---	10.0 (14,328)	5.9	6.0	5.3
≤ 25% quantile	= - 0.379	0.1***	0	79.0 (4,368)	8.5***	9.9**	19.9***
≥ 25% quantile	= - 0.379	40.7	---	13.9 (13,416)	6.3	7.0	6.3
≤ 50% quantile	= - 0.193	7.3	---	24.1 (8,814)	5.2	7.5	8.2
≥ 50% quantile	= - 0.193	35.3	---	1.0*** (8,970)	5.7	5.2	5.7
≤ 75% quantile	= - 0.020	3.2*	1	37.6 (13,344)	7.2*	8.4*	10.3
≥ 75% quantile	= - 0.020	37.8	---	8.8 (4,440)	5.0	5.2	5.7
≤ 80% quantile	= 0.015	6.7	---	39.1 (14,236)	7.5*	7.7	8.3
≥ 80% quantile	= 0.015	52.1	---	8.9 (3,548)	6.0	5.5	6.1
≤ 90% quantile	= 0.144	8.5	---	25.3 (15,999)	6.7	6.9	8.5
≥ 90% quantile	= 0.144	15.8	---	3.1* (1,785)	6.3	5.7	6.5

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XII: Results for countries grouped by net foreign asset liabilities, excluding foreign direct investment (robustness tests, stable sample).

Economic conditions NFA_EXCL_FDI_GDP		Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Full sample							
≤ 10% quantile	= - 0.678	0.5***	0	6.2 (2,196)	7.7**	11.5***	18.2***
≥ 10% quantile	= - 0.678	70.3	---	82.5 (27,280)	5.6	5.1	4.5
≤ 20% quantile	= - 0.471	45.5	---	84.6 (5,113)	4.6	6.3	6.6
≥ 20% quantile	= - 0.471	49.2	---	65.8 (24,363)	5.3	6.6	4.8
≤ 25% quantile	= - 0.413	62.6	---	79.6 (6,598)	6.1	5.7	4.6
≥ 25% quantile	= - 0.413	43.2	---	43.5 (22,878)	6.3	5.9	5.0
≤ 50% quantile	= - 0.199	14.7	---	32.4 (14,128)	6.7	8.0	8.1
≥ 50% quantile	= - 0.199	7.5	---	1.5** (15,348)	5.7	7.7	10.2
≤ 75% quantile	= - 0.035	18.1	---	96.2 (21,799)	6.4	5.5	7.0
≥ 75% quantile	= - 0.035	0.7***	1	14.8 (7,677)	9.1***	9.7**	19.5***
≤ 80% quantile	= 0.004	27.0	---	99.4 (23,390)	6.2	6.7	6.8
≥ 80% quantile	= 0.004	0.5***	1	6.3 (6,086)	8.4***	10.6***	16.2***
≤ 90% quantile	= 0.130	44.1	---	87.0 (26,520)	5.9	5.8	5.9
≥ 90% quantile	= 0.130	6.2	0	14.1 (2,956)	6.9*	7.7	11.3*

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XIII: Results for countries grouped by net foreign asset liabilities, excluding foreign direct investment (robustness tests, full sample).

Economic conditions			Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
		Stable sample					
CA ≥ 0			61	20.6	15,360	15.6	5.2
CA ≥ 0	&	OECD	18	28.3	6,499	25.8	7.4
CA ≥ 0	&	Non-OECD	43	17.4	8,861	7.2	7.8
CA < 0			63	34.3	25,849	0.2***	10.9***
CA < 0	&	OECD	18	29.4	6,600	62.4	5.0
CA < 0	&	Non-OECD	45	36.2	19,249	0.6***	9.7**
		NFA_EXCL_FDI_GDP ≤ - 0.0201	59	43.0	30,480	3.2*	8.4*
CA ≥ 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0201	48	17.5	10,230	46.7	4.1
CA < 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0201	57	29.8	20,250	0.1***	11.5***
		NFA_GDP ≤ - 0.1233	58	43.9	30,427	3.3*	9.5**
CA ≥ 0	&	NFA_GDP ≤ - 0.1233	49	15.6	9,224	9.2	6.0
CA < 0	&	NFA_GDP ≤ - 0.1233	57	31.2	21,203	0.2***	10.5***
		EQ_LIAB_GDP > 0.0098	51	43.5	27,339	72.0	4.1
CA ≥ 0	&	EQ_LIAB_GDP > 0.0098	43	22.9	12,228	40.5	4.9
CA < 0	&	EQ_LIAB_GDP > 0.0098	48	25.6	15,111	2.0*	9.2**
		LIAB_EXCL_FDI_GDP > 0.4349	62	40.5	30,503	3.2*	7.5
CA ≥ 0	&	LIAB_EXCL_FDI_GDP > 0.4349	55	18.3	12,378	1.2**	10.8***
CA < 0	&	LIAB_EXCL_FDI_GDP > 0.4349	59	25.5	18,125	0.1***	12.6***
		LIAB _GDP > 0.6102	59	42.5	30,527	10.3	6.5
CA ≥ 0	&	LIAB _GDP > 0.6102	52	18.8	11,991	4.6*	6.4
CA < 0	&	LIAB _GDP > 0.6102	55	27.8	18,536	0.02***	12.8***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XIV: Results for countries grouped by the sign of their current account balances (main results, stable sample).

Economic conditions			Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
		Full sample					
CA ≥ 0			97	20.8	24,182	3.5*	9.8**
CA ≥ 0	&	OECD	21	34.6	9,144	77.4	5.0
CA ≥ 0	&	Non-OECD	76	17.0	15,038	0.4***	11.0***
CA < 0			99	38.7	45,022	5.6	9.2**
CA < 0	&	OECD	20	44.6	10,937	90.0	3.7
CA < 0	&	Non-OECD	79	37.3	34,085	1.2**	11.5***
		NFA_EXCL_FDI_GDP ≤ - 0.0347	93	46.3	50,281	18.1	5.5
CA ≥ 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	79	16.1	15,050	20.2	6.1
CA < 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	91	33.1	35,184	3.2*	9.7**
		NFA_GDP ≤ - 0.1353	94	45.8	50,171	15.0	5.8
CA ≥ 0	&	NFA_GDP ≤ - 0.1353	81	14.7	13,958	2.1*	10.5***
CA < 0	&	NFA_GDP ≤ - 0.1353	91	34.1	36,190	2.7*	11.4***
		EQ_LIAB_GDP > 0.0061	77	45.8	43,043	73.8	4.4
CA ≥ 0	&	EQ_LIAB_GDP > 0.0061	64	22.5	17,713	83.5	4.2
CA < 0	&	EQ_LIAB_GDP > 0.0061	73	28.5	25,306	17.5	6.5
		LIAB_EXCL_FDI_GDP > 0.4008	96	44.7	50,758	76.6	4.1
CA ≥ 0	&	LIAB_EXCL_FDI_GDP > 0.4008	86	18.0	18,635	19.2	7.0
CA < 0	&	LIAB_EXCL_FDI_GDP > 0.4008	89	30.6	32,076	3.3*	9.6**
		LIAB_GDP > 0.5427	98	44.1	51,359	50.2	6.7
CA ≥ 0	&	LIAB_GDP > 0.5427	87	17.8	18,626	56.7	5.6
CA < 0	&	LIAB_GDP > 0.5427	92	30.0	32,698	0.3***	12.4***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XV: Results for countries grouped by the sign of their current account balances (main results, full sample).

Economic conditions			Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
		Stable sample						
CA ≥ 0			15.6	---	17.8 (6,811)	6.3	5.2	7.1
CA ≥ 0	&	OECD	25.8	---	54.7 (2,956)	5.4	7.4	7.3
CA ≥ 0	&	Non-OECD	7.2	---	20.1 (3,855)	5.6	7.8	9.9
CA < 0			0.2***	0	6.1 (11,219)	8.0***	10.9***	20.1***
CA < 0	&	OECD	62.4	---	92.9 (2,939)	3.6	5.0	4.1
CA < 0	&	Non-OECD	0.6***	0	3.6* (8,280)	7.7**	9.7**	15.5***
		NFA_EXCL_FDI_GDP ≤ - 0.0201	3.2*	1	37.6 (13,344)	7.2*	8.4*	10.3
CA ≥ 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0201	46.7	---	19.8 (4,538)	4.8	4.1	4.0
CA < 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0201	0.1***	0	10.8 (8,806)	9.9***	11.5***	20.9***
		NFA_GDP ≤ - 0.1233	3.3*	2	41.1 (13,276)	6.4	9.5**	11.6*
CA ≥ 0	&	NFA_GDP ≤ - 0.1233	9.2	---	24.6 (4,071)	5.7	6.0	10.1
CA < 0	&	NFA_GDP ≤ - 0.1233	0.2***	0	4.9* (9,205)	7.9**	10.5***	18.6***
		EQ_LIAB_GDP > 0.0098	72.0	---	33.9 (12,236)	3.8	4.1	2.9
CA ≥ 0	&	EQ_LIAB_GDP > 0.0098	40.5	---	64.4 (5,513)	5.3	4.9	6.7
CA < 0	&	EQ_LIAB_GDP > 0.0098	2.0*	2	24.2 (6,723)	7.5*	9.2**	13.4*
		LIAB_EXCL_FDI_GDP > 0.4349	3.2*	1	9.8 (13,409)	7.1*	7.5	13.2*
CA ≥ 0	&	LIAB_EXCL_FDI_GDP > 0.4349	1.2**	0	19.1 (5,501)	7.7**	10.8***	15.4***
CA < 0	&	LIAB_EXCL_FDI_GDP > 0.4349	0.1***	0	6.4 (7,908)	8.2***	12.6***	23.1***
		LIAB _GDP > 0.6102	10.3	---	12.7 (13,445)	5.7	6.5	7.0
CA ≥ 0	&	LIAB _GDP > 0.6102	4.6*	3	30.3 (5,338)	6.0	6.4	10.2
CA < 0	&	LIAB _GDP > 0.6102	0.02***	0	4.1* (8,107)	9.6***	12.8***	25.5***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XVI: Results for countries grouped by the sign of their current account balances (robustness tests, stable sample).

Economic conditions			Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
		Full sample						
CA ≥ 0			3.5*	1	10.5 (10,625)	7.3*	9.8**	12.8*
CA ≥ 0	&	OECD	77.4	---	28.6 (4,124)	5.3	5.0	2.3
CA ≥ 0	&	Non-OECD	0.4***	0	17.7 (6,501)	8.3***	11.0***	18.6***
CA < 0			5.6	3	46.4 (19,367)	6.6	9.2**	12.4*
CA < 0	&	OECD	90.0	---	79.9 (4,822)	3.5	3.7	2.4
CA < 0	&	Non-OECD	1.2**	0	30.6 (14,545)	7.7**	11.5***	14.8***
		NFA_EXCL_FDI_GDP ≤ - 0.0347	18.1	---	96.2 (21,799)	6.4	5.5	7.0
CA ≥ 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	20.2	---	14.7 (6,620)	5.7	6.1	8.2
CA < 0	&	NFA_EXCL_FDI_GDP ≤ - 0.0347	3.2*	1	79.8 (15,132)	6.9*	9.7**	13.5**
		NFA_GDP ≤ - 0.1353	15.0	---	97.8 (21,682)	6.5	5.8	8.6
CA ≥ 0	&	NFA_GDP ≤ - 0.1353	2.1*	0	11.4 (6,117)	8.3***	10.5***	14.3**
CA < 0	&	NFA_GDP ≤ - 0.1353	2.7*	0	59.1 (15,542)	8.4***	11.4***	12.2*
		EQ_LIAB_GDP > 0.0061	73.8	---	60.8 (19,145)	5.2	4.4	3.2
CA ≥ 0	&	EQ_LIAB_GDP > 0.0061	83.5	---	38.2 (7,929)	4.9	4.2	2.9
CA < 0	&	EQ_LIAB_GDP > 0.0061	17.5	---	73.2 (11,192)	6.0	6.5	8.8
		LIAB_EXCL_FDI_GDP > 0.4008	76.6	---	38.9 (22,108)	5.1	4.1	3.5
CA ≥ 0	&	LIAB_EXCL_FDI_GDP > 0.4008	19.2	---	8.3 (8,244)	5.9	7.0	6.0
CA < 0	&	LIAB_EXCL_FDI_GDP > 0.4008	3.3*	2	43.5 (13,817)	8.2***	9.6**	14.0**
		LIAB_GDP > 0.5427	50.2	---	54.5 (22,364)	5.3	6.7	6.3
CA ≥ 0	&	LIAB_GDP > 0.5427	56.7	---	52.4 (8,228)	6.4	5.6	4.5
CA < 0	&	LIAB_GDP > 0.5427	0.3***	0	35.1 (14,101)	7.7**	12.4***	19.0***

Note: The threshold values are given by quartiles (first or third) of the given quantities within the sample (full or stable).
Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XVII: Results for countries grouped by the sign of their current account balances (robustness tests, full sample).

Full sample	P-value (in %)	No. of observations	No. of countries	Other years for which there is a rejection at at least 5%
No condition	30.6	3,544	75	1996, 2003, 2006
EQ_LIAB_GDP > 0.0061	0.7***	2,127	44	1996, 2006
Fixed exchange	13.2	2,161	49	2003, 2004, 2007
Fixed exchange & EQ_LIAB_GDP > 0.0061	3.8*	1,337	30	2004
LIAB_EXCL_FDI_GDP > 0.4008	10.3	2,259	48	2006
LIAB_EXCL_FDI_GDP > 0.4008 & EQ_LIAB_GDP > 0.0061	.08***	1,458	29	1996, 2006
NFA_GDP ≤ -0.1353	33.6	2,316	50	1995
NFA_GDP ≤ -0.1353 & EQ_LIAB_GDP > 0.0061	0.2***	1,195	25	1996, 2006
NFA_EXCL_FDI_GDP ≤ -0.0347	31.4	2,532	54	2006
NFA_EXCL_FDI_GDP ≤ -0.0347 & EQ_LIAB_GDP > 0.0061	1.4**	1,455	30	2006
LIAB_GDP > 0.5427	32.4	2,137	45	2006
LIAB_GDP > 0.5427 & EQ_LIAB_GDP > 0.0061	1.4**	1,422	30	1996
CA < 0	59	2,677	65	2003
CA < 0 & EQ_LIAB_GDP > 0.0061	2.8*	1,443	35	2004

Stable sample	P-value (in %)	No. of observations	No. of countries	Other years for which there is a rejection at at least 5%
No condition	17.5	3,102	66	1996, 2006
EQ_LIAB_GDP > 0.0098	2.9*	1,901	39	1996, 2006
Fixed exchange	16.8	1,842	41	2003, 2006
Fixed exchange & EQ_LIAB_GDP > 0.0098	8.1	1,111	25	1996, 1999, 2006
LIAB_EXCL_FDI_GDP > 0.4349	4.7*	2,001	42	2006
LIAB_EXCL_FDI_GDP > 0.4349 & EQ_LIAB_GDP > 0.0098	0.5***	1,325	27	1996, 2006
NFA_GDP ≤ -0.1233	10.0	2,153	46	1995, 1996
NFA_GDP ≤ -0.1233 & EQ_LIAB_GDP > 0.0098	0.2***	1,201	25	1995, 1996, 2003, 2004, 2006, 2007
NFA_EXCL_FDI_GDP ≤ -0.0201	14.2	2,362	60	1996
NFA_EXCL_FDI_GDP ≤ -0.0201 & EQ_LIAB_GDP > 0.0098	3.7*	1,457	30	2006
LIAB_GDP > 0.6102	25.0	1,914	40	---
LIAB_GDP > 0.6102 & EQ_LIAB_GDP > 0.0098	7.1	1,278	26	1996, 2006
CA < 0	51.6	2,257	55	1996, 2003
CA < 0 & EQ_LIAB_GDP > 0.0098	10.4	1,241	30	2003

Table XVIII: Results for year 1997 and other years of the period 1995–2007 (top table: full sample; bottom table: stable sample).

Economic conditions	Corruption		Government effectiveness		Political stability		Rule of law	
	Global P-value (in %)	1-in-10 criterion	Global P-value (in %)	1-in-10 criterion	Global P-value (in %)	1-in-10 criterion	Global P-value (in %)	1-in-10 criterion
Full sample								
Top 25%	99.4	3.5	68.2	5.4	13.5	7	78.6	5.7
Bottom 75%	19	7.7	6.4	7.4	4.4*	8.1*	4.7*	10.4***
Top 50%	88.8	4.7	79.8	4.5	77.8	5.3	66.9	4.8
Bottom 50%	0.7***	10.8***	28.1	7.3	0.5***	10.4***	1.0**	11.7***
Top 75%	77	5	77	6.1	74.5	4.7	83	4.5
Bottom 25%	2.1*	8.9*	0.02***	17.8***	0.003***	18.4***	0.2***	11.0***
Stable sample								
Top 25%	92.5	5.6	86.7	4	22.6	6.8	77	3.5
Bottom 75%	14.4	5.7	11.9	6.6	2.0*	9.4**	7.6	6.8
Top 50%	88.7	4.4	79.1	3.6	92.2	4.3	51.5	5.7
Bottom 50%	0.6***	10.0**	11.5	7	0.6***	9.9**	1.2**	10.7***
Top 75%	66.3	4.7	64.4	5.4	64.8	4.9	82.2	4.1
Bottom 25%	5.6	9.1**	2.6*	8.9*	0.02***	14.8***	3.1*	8.1*

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XIX: Results for countries grouped by WGI indicators: all indicators (main results, full and stable samples).

Economic conditions Corruption	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-20 criterion
Full sample									
25% least corrupt	34	37.9	16,207	99.4	3.5	---	61.9 (7,252)	4.0	1.5
75% most corrupt	76	38.5	34,217	19.0	7.7	---	44.9 (14,700)	6.1	6.4
50% least corrupt	65	36.1	28,978	88.8	4.7	---	20.0 (12,821)	5.1	3.3
50% most corrupt	57	33.0	21,575	0.7***	10.8***	0	24.3 (9,184)	8.8***	17.6***
75% least corrupt	92	36.6	40,686	77.0	5.0	---	22.9 (17,869)	4.8	3.1
25% most corrupt	34	25.1	9,774	2.1*	8.9*	0	16.5 (4,099)	8.1***	13.9**
Stable sample									
25% least corrupt	28	41.4	14,658	92.5	5.6	---	91.0 (6,584)	5.6	2.2
75% most corrupt	44	45.5	23,482	14.4	5.7	---	47.0 (10,154)	6.5	8.4
50% least corrupt	47	40.3	23,581	88.7	4.4	---	18.9 (10,480)	4.5	3.2
50% most corrupt	34	37.6	14,688	0.6***	10.0**	0	21.5 (6,311)	7.7**	16.2***
75% least corrupt	62	41.2	31,138	66.3	4.7	---	28.5 (13,758)	4.9	3.3
25% most corrupt	21	28.8	6,994	5.6	9.1**	---	6.4 (2,976)	6.4	10.9

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XX: Results for countries grouped by WGI indicators: Corruption (main results and robustness checks, full and stable samples).

	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-20 criterion
Full sample									
Africa and Middle East	19	44.7	9,562	0.000000007***	36.3***	0	59.7 (3,988)	18.3***	67.3***
CIS (former Soviet Union)	15	46.6	8,381	67.9	4.1	---	8.3 (3,574)	5.4	2.8
Eastern Europe	15	56.7	10,204	46.2	7.2	---	66.7 (4,452)	6.6	6.9
Developed Europe / Western offshoots / Japan	23	71.0	20,267	75.0	4.8	---	88.7 (9,031)	5.5	3.5
Latin America	18	53.9	11,266	26.3	6.7	---	21.3 (4,896)	5.7	7.3
South and East Asia	15	66.5	11,208	37.4	5.7	---	37.0 (4,856)	6.3	6.1
Stable sample									
Africa and Middle East	7	52.0	4,100	0.03***	15.2***	1	70.2 (1,725)	9.0***	26.7***
CIS (former Soviet Union)	9	52.0	5,761	95.8	4.3	---	38.1 (2,509)	5.3	2.6
Eastern Europe	11	52.0	7,006	16.3	8.2*	---	88.6 (3,068)	6.9*	8.9
Developed Europe / Western offshoots / Japan	20	52.0	13,168	55.5	5.4	---	88.5 (5,912)	3.4	4.1
Latin America	10	52.0	6,150	1.0***	11.5***	2	23.0 (2,687)	6.9*	14.3**
South and East Asia	12	51.8	7,142	33.1	5.4	---	7.9 (3,089)	4.8	4.7

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXI: Results for countries grouped by geographic regions (main results and robustness tests, full and stable samples).

D Additional material for referees

BOP series	Independent series	Less persistent series	All series
Current account	yes	no	yes
Goods	yes	no	yes
Services	yes	no	yes
Income	yes	no	yes
Current transfers	no	no	yes
Capital account	no	no	yes
Financial account	yes	yes	yes
Direct investment	yes	no	yes
Portfolio investment	yes	yes	yes
Other investment	no	yes	yes
Financial derivatives	yes	yes	yes
Reserve assets	yes	yes	yes
Net errors and omissions	yes	yes	yes

Table XXII: Summary of the three sets of balance of payments series considered in this study.

Rejection rates	C = 10 countries	C = 20 countries	C = 30 countries	C = 40 countries	C = 50 countries	C = 60 countries	C = 70 countries	C = 80 countries
M = 10 quarters	12.5%	11.2%	11.1%	9.6%	9.1%	9.2%	6.9%	6.2%
M = 15 quarters	16.6%	14.7%	14.6%	10.5%	11.5%	8.0%	7.2%	6.0%
M = 20 quarters	20.8%	16.2%	17.3%	14.3%	10.9%	9.6%	6.8%	7.3%
M = 25 quarters	22.3%	20.1%	17.9%	15.7%	13.0%	11.0%	7.5%	7.1%
M = 30 quarters	25.8%	25.0%	20.7%	16.9%	15.1%	11.9%	8.6%	8.5%
M = 35 quarters	29.4%	26.2%	24.2%	19.3%	17.4%	15.1%	10.4%	7.0%
M = 40 quarters	34.4%	27.7%	24.3%	22.0%	20.1%	12.9%	11.8%	7.0%
All quarters	39.8%	35.4%	32.4%	29.6%	24.5%	16.1%	14.0%	7.8%

Table XXIII: Rejection rates at the $5\,\%$ level of the hypothesis of conformity to Benford's law on data subsets corresponding to given numbers C of countries and quarters M thereof, all chosen at random; estimated by performing $1,000$ random draws for each value. The last line is identical to the top table of Table IV.

All series / Independent series / Less persistent series	C = 10 countries	C = 20 countries	C = 30 countries	C = 40 countries	C = 50 countries	C = 60 countries	C = 70 countries	C = 80 countries
M = 10 quarters	12.5% /	11.2% /	11.1% /	9.6% /	9.1% /	9.2% /	6.9% /	6.2% /
	12.3% /	10.9% /	10.0% /	8.6% /	9.0% /	7.8% /	7.2% /	5.4% /
	7.6% /	5.9% /	5.6% /	5.8% /	5.0% /	5.6% /	4.5% /	4.3% /
M = 15 quarters	16.6% /	14.7% /	14.6% /	10.5% /	11.5% /	8.0% /	7.2% /	6.0% /
	17.2% /	12.1% /	12.1% /	10.0% /	8.6% /	6.4% /	6.3% /	5.9% /
	7.2% /	6.6% /	8.6% /	5.7% /	6.2% /	6.1% /	4.1% /	5.2% /
M = 20 quarters	20.8% /	16.2% /	17.3% /	14.3% /	10.9% /	9.6% /	6.8% /	7.3% /
	20.6% /	16.6% /	14.8% /	13.5% /	10.3% /	9.5% /	6.9% /	5.0% /
	7.8% /	7.0% /	6.7% /	6.7% /	5.8% /	5.0% /	4.9% /	3.8% /
M = 25 quarters	22.3% /	20.1% /	17.9% /	15.7% /	13.0% /	11.0% /	7.5% /	7.1% /
	22.2% /	19.9% /	17.6% /	15.4% /	13.0% /	11.0% /	6.9% /	4.8% /
	11.0% /	7.4% /	8.0% /	6.4% /	5.6% /	4.2% /	3.3% /	2.7% /
M = 30 quarters	25.8% /	25.0% /	20.7% /	16.9% /	15.1% /	11.9% /	8.6% /	8.5% /
	25.8% /	23.0% /	19.1% /	13.9% /	13.8% /	8.0% /	6.1% /	4.9% /
	9.6% /	8.2% /	8.7% /	6.3% /	5.2% /	4.0% /	3.3% /	2.1% /
M = 35 quarters	29.4% /	26.2% /	24.2% /	19.3% /	17.4% /	15.1% /	10.4% /	7.0% /
	27.0% /	23.4% /	20.8% /	16.6% /	14.7% /	10.3% /	7.4% /	4.7% /
	10.7% /	8.3% /	8.4% /	7.6% /	6.5% /	4.7% /	3.1% /	2.8% /
M = 40 quarters	34.4% /	27.7% /	24.3% /	22.0% /	20.1% /	12.9% /	11.8% /	7.0% /
	33.7% /	28.3% /	22.2% /	19.3% /	16.1% /	11.7% /	7.3% /	2.9% /
	12.6% /	8.5% /	7.6% /	8.3% /	6.5% /	5.3% /	3.5% /	1.8% /
All quarters	39.8% /	35.4% /	32.4% /	29.6% /	24.5% /	16.1% /	14.0% /	7.8% /
	38.5% /	36.2% /	31.9% /	24.8% /	19.2% /	13.5% /	7.0% /	2.0% /
	13.4% /	11.9% /	9.4% /	5.9% /	6.0% /	2.4% /	1.5% /	0.3% /

Table XXIV: Rejection rates at the 5 % level of the hypothesis of conformity to Benford's law on data subsets corresponding to given numbers C of countries and quarters M thereof, all chosen at random, for different sets of series; estimated by performing $1,000$ random draws for each value. The first element of each cell is identical to the values reported in Table XXIII. Note that this table is a more detailed version of the bottom table of Table IV.

Economic conditions Corruption	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Full sample					
25% least corrupt	34	37.9	16,207	99.4	3.5
75% most corrupt	76	38.5	34,217	19.0	7.7
50% least corrupt	65	36.1	28,978	88.8	4.7
50% most corrupt	57	33.0	21,575	0.7***	10.8***
75% least corrupt	92	36.6	40,686	77.0	5.0
25% most corrupt	34	25.1	9,774	2.1*	8.9*
Stable sample					
25% least corrupt	28	41.4	14,658	92.5	5.6
75% most corrupt	44	45.5	23,482	14.4	5.7
50% least corrupt	47	40.3	23,581	88.7	4.4
50% most corrupt	34	37.6	14,688	0.6***	10.0**
75% least corrupt	62	41.2	31,138	66.3	4.7
25% most corrupt	21	28.8	6,994	5.6	9.1**

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXV: Results for countries grouped by WGI indicators: Corruption (main results, full and stable samples).

Economic conditions Government effectiveness (GE)	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Full sample					
Top 25% GE	34	43.8	18,790	68.2	5.4
Bottom 75% GE	74	36.9	31,670	6.4	7.4
Top 50% GE	70	36.7	31,642	79.8	4.5
Bottom 50% GE	54	30.6	18,818	28.1	7.3
Top 75% GE	93	39.0	43,940	77.0	6.1
Bottom 25% GE	27	21.9	6,520	0.02***	17.8***
Stable sample					
Top 25% GE	30	45.3	17,210	86.7	4.0
Bottom 75% GE	42	43.0	20,966	11.9	6.6
Top 50% GE	51	40.8	25,765	79.1	3.6
Bottom 50% GE	31	35.0	12,411	11.5	7.0
Top 75% GE	63	44.1	33,963	64.4	5.4
Bottom 25% GE	17	22.6	4,213	2.6*	8.9*

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXVI: Results for countries grouped by WGI indicators: Government Effectiveness (main results, full and stable samples).

Economic conditions Political stability (PS)	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Full sample					
Top 25% PS	34	32.0	13,629	13.5	7.0
Bottom 75% PS	84	37.4	36,980	4.4*	8.1*
Top 50% PS	60	35.5	26,261	77.8	5.3
Bottom 50% PS	63	33.5	24,535	0.5***	10.4***
Top 75% PS	84	36.2	36,956	74.5	4.7
Bottom 25% PS	41	29.1	13,743	0.003***	18.4***
Stable sample					
Top 25% PS	27	36.4	12,410	22.6	6.8
Bottom 75% PS	51	42.9	25,867	2.0*	9.4**
Top 50% PS	41	39.6	20,241	92.2	4.3
Bottom 50% PS	41	38.2	18,223	0.6***	9.9**
Top 75% PS	57	40.3	28,202	64.8	4.9
Bottom 25% PS	28	31.6	10,166	0.02***	14.8***

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXVII: Results for countries grouped by WGI indicators: Political Stability (main results, full and stable samples).

Economic conditions Rule of law (ROL)	Number of countries	Avg. number of quarters	Number of observations	Global P-value (in %)	1-in-10 criterion
Full sample					
Top 25% ROL	33	39.0	16,218	78.6	5.7
Bottom 75% ROL	78	37.6	34,294	4.7*	10.4***
Top 50% ROL	61	37.9	28,551	66.9	4.8
Bottom 50% ROL	53	36.3	22,155	1.0**	11.7***
Top 75% ROL	91	37.1	40,809	83.0	4.5
Bottom 25% ROL	32	26.7	9,770	0.2***	11.0***
Stable sample					
Top 25% ROL	28	42.0	14,850	77.0	3.5
Bottom 75% ROL	46	43.3	23,378	7.6	6.8
Top 50% ROL	46	41.0	23,370	51.5	5.7
Bottom 50% ROL	30	43.0	14,906	1.2**	10.7***
Top 75% ROL	63	41.8	32,075	82.2	4.1
Bottom 25% ROL	19	28.5	6,220	3.1*	8.1*

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXVIII: Results for countries grouped by WGI indicators: Rule of Law (main results, full and stable samples).

Economic conditions Corruption	Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Full sample						
25% least corrupt	99.4	---	61.9 (7,252)	4.0	3.5	1.5
75% most corrupt	19.0	---	44.9 (14,700)	6.1	7.7	6.4
50% least corrupt	88.8	---	20.0 (12,821)	5.1	4.7	3.3
50% most corrupt	0.7***	0	24.3 (9,184)	8.8***	10.8***	17.6***
75% least corrupt	77.0	---	22.9 (17,869)	4.8	5.0	3.1
25% most corrupt	2.1*	0	16.5 (4,099)	8.1***	8.9*	13.9**
Stable sample						
25% least corrupt	92.5	---	91.0 (6,584)	5.6	5.6	2.2
75% most corrupt	14.4	---	47.0 (10,154)	6.5	5.7	8.4
50% least corrupt	88.7	---	18.9 (10,480)	4.5	4.4	3.2
50% most corrupt	0.6***	0	21.5 (6,311)	7.7**	10.0**	16.2***
75% least corrupt	66.3	---	28.5 (13,758)	4.9	4.7	3.3
25% most corrupt	5.6	---	6.4 (2,976)	6.4	9.1**	10.9

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXIX: Results for countries grouped by WGI indicators: Corruption (robustness tests, full and stable samples).

Economic conditions Government effectiveness (GE)	Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Full sample						
Top 25% GE	68.2	---	25.0 (8,456)	5.7	5.4	4.3
Bottom 75% GE	6.4	---	24.8 (13,504)	6.4	7.4	10.6
Top 50% GE	79.8	---	19.1 (14,046)	4.7	4.5	3.2
Bottom 50% GE	28.1	---	41.0 (7,914)	6.7	7.3	6.9
Top 75% GE	77.0	---	53.3 (19,253)	5.3	6.1	4.4
Bottom 25% GE	0.02***	0	0.7*** (2,707)	11.1***	17.8***	32.7***
Stable sample						
Top 25% GE	86.7	---	33.0 (7,736)	6.6	4.0	3.0
Bottom 75% GE	11.9	---	37.0 (9,010)	6.3	6.6	9.0
Top 50% GE	79.1	---	33.3 (11,477)	5.2	3.6	3.8
Bottom 50% GE	11.5	---	62.2 (5,269)	5.5	7.0	8.3
Top 75% GE	64.4	---	43.5 (14,972)	4.6	5.4	4.3
Bottom 25% GE	2.6*	2	6.3 (1,774)	6.8*	8.9*	13.8**

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXX: Results for countries grouped by WGI indicators: Government Effectiveness (robustness tests, full and stable samples).

Economic conditions Political stability (PS)	Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Full sample						
Top 25% PS	13.5	---	78.0 (6,045)	6.5	7.0	9.4
Bottom 75% PS	4.4*	4	39.5 (15,983)	7.3*	8.1*	14.2**
Top 50% PS	77.8	---	63.1 (11,647)	6.4	5.3	4.7
Bottom 50% PS	0.5***	0	31.5 (10,453)	7.4*	10.4***	18.0***
Top 75% PS	74.5	---	35.3 (16,178)	4.7	4.7	2.9
Bottom 25% PS	0.003***	0	6.9 (5,881)	11.1***	18.4***	34.5***
Stable sample						
Top 25% PS	22.6	---	93.1 (5,544)	5.4	6.8	7.3
Bottom 75% PS	2.0*	2	39.5 (11,250)	6.8*	9.4**	13.6**
Top 50% PS	92.2	---	64.0 (9,053)	5.1	4.3	2.8
Bottom 50% PS	0.6***	0	29.8 (7,813)	7.2*	9.9**	14.9***
Top 75% PS	64.8	---	37.5 (12,460)	5.1	4.9	3.9
Bottom 25% PS	0.02***	0	6.2 (4,366)	7.1*	14.8***	26.8***

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXXI: Results for countries grouped by WGI indicators: Political Stability (robustness tests, full and stable samples).

Economic conditions Rule of law (ROL)	Global P-value (in %)	Stability index	Less persistent series: P-value in % (number of obs.)	1-in-5 criterion	1-in-10 criterion	1-in-20 criterion
Full sample						
Top 25% ROL	78.6	---	57.4 (7,256)	5.9	5.7	3.3
Bottom 75% ROL	4.7*	3	12.9 (14,728)	7.4*	10.4***	12.2*
Top 50% ROL	66.9	---	21.7 (12,666)	6.6	4.8	3.3
Bottom 50% ROL	1.0**	0	25.1 (9,402)	8.5***	11.7***	15.4***
Top 75% ROL	83.0	---	26.0 (17,915)	4.4	4.5	3.2
Bottom 25% ROL	0.2***	0	12.3 (4,097)	9.2***	11.0***	24.5***
Stable sample						
Top 25% ROL	77.0	---	70.7 (6,664)	5.9	3.5	4.2
Bottom 75% ROL	7.6	---	34.7 (10,106)	4.7	6.8	9.4
Top 50% ROL	51.5	---	35.2 (10,413)	5.8	5.7	4.4
Bottom 50% ROL	1.2**	0	54.6 (6,377)	7.9**	10.7***	14.9***
Top 75% ROL	82.2	---	29.0 (14,167)	5.8	4.1	3.6
Bottom 25% ROL	3.1*	1	31.0 (2,631)	6.7	8.1*	9.4

Legend: * if significant at the 5% level, ** at the 2% level, *** at the 1% level.

Table XXXII: Results for countries grouped by WGI indicators: Rule of Law (robustness tests, full and stable samples).