The Effect of Public Debt on Growth in Multiple Regimes in the Presence of Long-Memory and Non-Stationary Debt Series

IRINA SYSSOYEVA-MASSON [1]

&

JOÃO DE SOUSA ANDRADE

CeBER and Faculty of Economics of the University of Coimbra

CeBER Working Papers

No. 7 2017

[1] PhD in Economics, Department of Finance and Economics, University of Savoy Mont Blanc, France; Association CEMAFI International. E-mail: irina.syssoyeva-masson@univ-smb.fr; irina_masson@hotmail.fr

The Effect of Public Debt on Growth in Multiple Regimes in the Presence of Long-Memory and Non-Stationary Debt Series

Irina Syssoyeva-Masson[1] and João de Sousa Andrade[2]

Abstract

The study of the relationship between public debt and economic growth came again to the spotlight with the financial crisis (2007-2008) and with the sovereign debt crisis that followed in Europe. This literature aims to shed light about the sign, magnitude, mechanisms and threshold regimes relating debt to growth and to make policy recommendations with important consequences in terms of government's policies.

We empirically investigate this relationship for a group of 60 countries for a long-time period (the shorter one from 1970 to 2012) using the historical public debt database (HPDD) built by the International Monetary Fund (IMF) and we defend that the empirical strategy underlying most of the studies on this topic should be revised. We claim that: a) the study of the long-memory property of the public debt GDP ratio and stationarity (using the last generation tests) has to be performed as a first step of the empirical analysis, what has been done using 87 countries; b) In the presence of a non-stationary public debt GDP ratio cointegration analysis was used to estimate the relationship between the public debt GDP ratio and output; c) under the no rejection of the null of no cointegration, the above mentioned relationship was studied between the public debt GDP ratio first difference and GDP growth rate using threshold models and searching for thresholds using a wide variety of variables.

The main conclusions of this study are that the debt series have a long memory and should not be analyzed in a short-term framework; additionally, the non-stationarity of the debt series does not allow researchers to apply stationary econometrics methods to model its behavior. This finding implies that the relationship between economic growth and national debt that has been characterizing the literature on the subject, has disputable econometric foundations. We thus recommend our empirical strategy to overcome the above-mentioned drawbacks of the existent empirical literature. Finally, it should be mentioned that the relationship between the public debt GDP ratio first difference and GDP growth rate is always negative despite the different threshold regimes identified.

JEL Classification: C24, C51, E62, H6 and O11

Keywords: Public Debt, Growth, Long Memory, Stationary, Co-integration and Thresholds.

1 PhD in Economics, Department of Finance and Economics, University of Savoy Mont Blanc, France; *Association CEMAFI International*. E-mail: irina.syssoyeva-masson@univ-smb.fr; irina_masson@hotmail.fr.

2 *Full Professor, Centre for Business and Economics Research* (CEBER), Department of Economics, University of Coimbra, Portugal; *Association CEMAFI International*. E-mail: jasa@fe.uc.pt (corresponding author).

1. Introduction

The relationship between public debt and economic growth has been, and continues to be the subject of much empirical and theoretical attention in the literature as witnessed by the growing volume of studies recently reported. Those performed by Reinhart and Rogoff (2010), Reinhart et al. (2012), Cecchetti et al. (2011), Checherita-Westphal and Rother (2012), Kumar and Woo (2010), Baum et al. (2013), and Panizza and Presbitero (2012), and Bruce Hansen (2015) are among the many studies so far featured in the literature.

Most of the research studying the relationship between debt (hereafter, *d*) and output growth (from now on, *g*) has found a strong negative relationship between debt levels and real economic growth, especially for Debt-to-GDP ratios above a certain threshold level, i.e. the authors implicitly assume that there is some value of *d* that should be understood as a threshold value: *d* positively affects *g* below it; conversely, *d* negatively affects *g* above it.

In an influential series of papers, Reinhart and Rogoff (2010, 2012)[3] argue that (i) a high level of public debt is associated with lower real GDP growth, (ii) the relationship is causal from public debt to growth and, finally, (iii) there is an important threshold around 90% Public Debt-to-GDP ratio above which growth drops substantially. Despite strong criticism from Herndon et al. (2013), the work of Reinhart and Rogoff (hereafter, RR) has raised the interest in the question, and their results were confirmed (or partly confirmed) by many studies which are reviewed in the next section.

This presumption about the existence of a threshold value of *d* (but also the direction of causality between *d* and *g*) is of critical importance given the historically high level of public debt in most advanced economies, and also because of the existence of the intense debate about austerity versus stimulus as the appropriate policy response. In fact, recently this presumption of existence of a threshold (approximately 90% of GDP) has been the basis for the fiscal austerity imposed in many European Union (EU) countries. This presumption also raises the importance of preliminary analysis of the statistical characteristics of *d* and *g* before studying the "Debt-to-GDP – growth" relationship.

In fact, the preliminary analyses of stationarity of data and the detection of the presence of long memory in series are absent from the existing literature on the "Debt-to-GDP – growth" relationship (i.e. literature that studies the relationship between *g*, the growth rate of output,

[3] According to Krugman (2013), the contribution of Reinhardt and Rogoff (2010) may have had more immediate influence on public debate than any previous paper in the history of economics.

$g = \Delta \log(Y)$, and d_t, Debt-to-GDP ratio). For instance, in many studies d_t is considered: (i) as a short run phenomenon, and (ii) as a stationary variable. Though, if a d_t series contains a unit root [i.e. I(1)], that would imply that the results of many previous studies (some of which had been used as a basis for policy recommendations) are spurious. In such a case, we should study instead, the relationship between Y_t and d_t, and if both series are non-stationary and co-integrated, then appropriate methods for non-stationary variables should be used. If, however, there is no cointegration between Y_t and d_t, then the relationship between g and Δd (debt growth rate) should be studied (again, not between g_t and d_t). Further, even if d_t series is stationary, but has a long memory, this fact should not be ignored and d_t series should be considered as a long run phenomenon and properly taken into account while studying the "Debt-to-GDP – growth" relationship. For the above-mentioned reasons, we propose to study the relation between g_t and Δd_t, (and a group of control variables), in the presence of thresholds that could explain the behavior of growth in a non-spurious way.

To the best of our knowledge, there is no study which has yet examined the statistical property of d and taken into consideration its stationarity property in the analysis of the "Debt-to-GDP – growth" relationship. Therefore, the contribution of this study to the existing empirical literature on the relationship between economic growth and debt is twofold:

- First, to fill this gap by reporting our investigation of the statistical properties of d_t (Debt-to-GDP) and g (economic growth) series and revealing the possible existence of long memory in our data[4]. In conducting this investigation, we apply several statistical methods (that are presented and discussed later) to the new data set on Gross Government Debt-to-GDP ratios for 87 countries (listed in Appendix A) over a long period (i.e. countries for which data is available at least from 1970 to 2012) from the historical public debt database (HPDD) built by the IMF[5]. We also apply last generation of Unit Root tests to confirm the presence of unit roots.
- Second, by taking into the account the presence of long memory and non-stationarity of d series, we propose to investigate (i) whether the debt-growth relation varies with the level of indebtedness or not and (ii) by using Hansen (1999) threshold model, we examine whether there is a common threshold for government debt ratios above which long-term growth rates may drop off significantly. To test the presence of different

[4] In this paper we propose to study mostly d because the data generating process of *output* or g have already been extensively studied in, and after the seminal paper of Nelson and Plosser (1982).

[5] The data covering nearly the entire IMF membership (174 countries) and spanning an exceptionally long time period (for some countries, such the United Kingdom, the data is available since 1692). We select countries with data for 1970.

regimes in the relation between growth and debt we consider four different types of the threshold variable and not one as it is usually done.

In order to capture the presence of the long memory or persistency in *d* we use several methods of long-range dependence such as Spectral analysis, the Hurst (1951) test, the tests of Geweke and Porter-Hudak (1983), and Reisen (1994), and finally the variance analysis (VR) in the revised version of Lo and MacKinlay (1989).

Since most of the analyses require prior information about the stationary nature of the series, we also apply several unit rout tests to our data [i.e. Augmented Dickey-Fuller (1979) (ADF) tests and Schmidt and Phillips (1992) Lagrange multiplier test (S-P)], but also a new generation of unit root tests [ADF tests of Chang (2002)] which take account of cross-section dependence (CSD)[6]. Further, in order to analyze the long-run relationship between debt and GDP, we test for co-integration between them by applying Westerlund (2007) cointegration tests.

As we confirmed the unit root characteristic of debt ratio (d_t) and the absence of co-integration between debt and GDP we propose to confirm the presence of different regimes in the relation between the first differences of these variables using the Hansen (1999) LR tests for the rejection of threshold values.

The remainder of the paper is structured as follows. Section 2 provides a selective[7] overview of the related literature. Section 3 discusses the empirical methods employed, presents and analyses the empirical findings, Section 4 confirms the absence of a threshold effect in the relation between the first difference of debt ratio and GDP growth, and Section 5 concludes the paper and suggests some directions for further research.

2. Literature Review

The relationship between public debt and economic growth has been, and continues to be the subject of much empirical and theoretical attention in the literature. Overall, the main finding of papers studying this relationship suggest that a high level of public debt is associated with lower real GDP growth. This result has important policy and political consequences.

---

[6] *The CSD has been ignored for a long time in the economic literature and it is well known nowadays, that* neglecting CSD can lead to biased estimates and spurious inference. Thus, *this fact may invalidate some previous studies' results.*

[7] The extensive empirical literature on the "debt-growth" relationship varies significantly in terms of the methodology employed, as well as the data set and sample periods covered.

Probably, one of the most influential papers on the topic is that by Reinhart and Rogoff (2010), the key findings of which are that there is an important *threshold* around 90% Public Debt-to-GDP above which, growth drops substantially. This work has been heavily challenged by Herndon et al. (2013), who replicated the RR study, demonstrating that this threshold effect seems to vanish after correcting for a coding error and by using a different weighting of the data. However, despite these criticisms, the RR work has raised interest in the question, and the debate about the relationship between debt and growth is still very much open since the number of articles devoted to this relationship keeps growing. Moreover, as mentioned earlier, the results from the RR study have been (at least partially) confirmed by many empirical studies.

2.1. The presence of a threshold

In discussing this issue, we have no intention to sum up all the contributions, and focus instead on a few. Kumar and Woo (2010) have found that a 10-percentage point increase in the initial Debt-to-GDP ratio is associated with a slowdown in annual real per capita GDP growth of 0.15 percentage points per year in advanced economies. They have also uncovered some evidence of non-linearity. According to the authors, only high levels of debt (i.e. above 90% of GDP) have a significant negative effect on growth.

Similarly, Cecchetti et al. (2011) have found that a 10% increase in government debt reduces real per capita GDP growth by 0.17% per year and that public debt becomes a drag on growth beyond 96% of GDP.

Baum et al. (2013) have analyzed the non-linear impact of public debt on GDP growth in the Euro Area. Their results suggest that the short-run impact of debt on GDP growth is positive, but decreases to around zero and loses significance beyond public debt-to-GDP ratios of around 67%. As for high debt-to-GDP ratios (i.e. above 95%), empirical results suggest that additional debt has a negative impact on economic activity.

We can also cite the more recent and, according to us, very interesting paper of Pescatori et al. (2014). In fact, the authors have raised many doubts about the existence of an appellative idea of a simple threshold in the relation between debt and growth. What is really interesting about the study of Pescatori et al. (2014) is that it concentrates on the long-term relationship between debt and GDP growth (i.e. today's stock of debt over GDP and GDP growth in the next h – years)[8], unlike the analysis of RR (2010) which focuses on the short-run "Debt-to-

[8] By taking a longer-term perspective they have tried to mitigate the reverse causality effects that temporary recessions can have on the debt-to-GDP ratio in the short run.

GDP – growth" relationship. More precisely, Pescatori et al. (2014) study different episodes (characterized by the growth of GDP over 15 years) during which public debt rises above a certain threshold, finding "no evidence of any particular debt threshold above which medium-term growth prospects are dramatically compromised". Furthermore, their results suggest that the "debt trajectory can be as important as the debt level in understanding future growth prospects", and according to them, countries with high but declining debt appear to grow equally as fast as countries with lower debt[9]. Moreover, according to them, using a span of 5, 10 or even 20 years does not modify the main conclusions.

According to these authors, the negative impact of a high level of debt on economic growth cannot be limited only to the very short term, and the relationship between debt and GDP growth should, therefore, be studied over long timeframes. However, we have many doubts about the relevance of this type of analysis. In fact, the problem with episodes is that they very often have different characteristics and contexts**,** so in general, we do not have homogeneous episodes, witnessing instead, variables within episodes with heterogeneous units. For instance, the effects of public debt on GDP growth would certainly be different among different episodes since the level of public debt is not the same, and its composition may differ among episodes.

Ash et al. (2015) re-examine the relationship between public debt and GDP growth by critically reviewing the empirical results from the previous papers: Reinhart and Rogoff (2010, 2012), Cecchetti et al. (2011), and Checherita-Westphal and Rother (2012). They find that there are no consistent thresholds in the data (i.e. 5-year forward growth rates are no lower when public debt exceeds 90% of GDP).

2.2. Causality

Theoretical arguments suggest that there is a negative non-linear correlation between public debt and economic growth. However, correlation does not necessarily imply causation. Hence, the issue of *causality* is additional to the threshold effect.

There are two points of view on the subject in the literature. The first is that the relationship between public debt and GDP growth may be causal from public debt to growth. In this case, public debt negatively affects growth and is responsible for poor economic performance in countries with high value of "Debt-to-GDP" ratio. The second is that high debt may be the result of sluggish GDP growth. For instance, low economic growth may push a

[9] Results of this study also suggest that higher debt is associated with a higher degree of output volatility.

country to borrow more in order to finance domestic economy, and therefore, may imply an increase in the debt ratio. Yet another possibility is the existence of a third variable (an omitted variable), that simultaneously affects both public debt and GDP growth (i.e. increases debt and reduces growth).

The direction of causality between public debt and GDP growth may have important policy implications. In fact, if high levels of public debt negatively affect economic growth, then expansionary fiscal policies that increase the debt-to-GDP ratio may reduce long-run growth perspectives, even if effective in the short-run. Therefore, any potential benefits of such policies may be partially (or fully) offset in the long-run. Although Alesina and Ardagna (2009), Cochrane (2011), and Perotti (2012) have argued in favor of the expansionary effects of fiscal adjustments on the basis of similar arguments, other authors have argued that expansionary fiscal policy is highly effective in environments with very low interest rates. In that case, expansionary fiscal policy is self-financing as increased government spending would not be offset by monetary authorities increasing interest rates (DeLong and Summers, 2012).

In their seminal paper, RR (2010) demonstrated that causality runs from public debt to growth, and these results have subsequently been confirmed by other authors. However, some other authors, like Panizza and Presbitero (2012), reject the hypothesis that high debt causes lower growth in a sample of OECD countries, even if their overall empirical results are consistent with the existing literature (i.e. the existence of a negative correlation between debt and growth).

The more recent paper of Donayre and Taivan (2015) also analyzes the direction of causality between public debt and real economic growth in OECD countries, finding evidence of both types of unidirectional causality, of bidirectional causality, and of no causality between real economic growth and public debt levels. These authors thus conclude that the causal link is intrinsic to each country, and hence, it cannot be inferred that higher debt always leads to lower economic growth.

Ash et al. (2015) re-examine the "Debt-to-GDP – growth" relationship by using the datasets from a few previous studies. Focusing on endogeneity and non-linearity[10] issues that emerge in empirical analysis, these authors find that causality is more likely to run from GDP growth to public debt than vice versa.

3. Empirical Study

[10] In fact, a few key issues that emerge in empirical exploration of the relationship between growth and public debt are endogeneity, non-linearity, and heterogeneity.

3.1. Data

In this paper we use the historical public debt database (HPDD) on gross government debt-to-GDP ratios that was constructed by the IMF Fiscal Affairs Department[11,12] and that covers public debt at the general government level[13]. We select 87 countries conditioned by the existence of data from 1970. Data for GDP per capita in PPP, comes from Penn World Tables (PWT, versions 7.1 and 9) has been previously harmonized. Data for human capital indicator comes from PWT (version 9). Other variables used in our study come from the World Development Indicators (WDI) of the World Bank. Data for some years were not-available (NA) and we calculated them by using autoregressive regressions models were the order of the process was chosen by the Schwarz criterion. The new values were obtained by repeated simulations of this process.

3.2. Long-run memory analysis

As mentioned in the introduction, most literature relating to the "Debt-to-GDP – growth" relationship implicitly assumes that there is some value of *d* that should be understood as a threshold value. Below this value *d* does indeed positively affect *g* and the contrary is observed above that value. We approach this question by studying the statistical characteristics of *d* and *g*.

In this paper we propose to study mostly *d* because the data generating process of *y* or *g* have already been extensively studied. More precisely, one of our main objectives is to identify the memory type of *d* series (i.e. short or long memory). We understand "a series with a short memory" a covariance-stationary series, as being stationary. The mean of a stationary series is the same for any time-period and the covariance of any pair of observations depends only on the time between these observations. As we see below, a stationary series can have a "long" memory, rather than a "short" one. Several tests of long-range dependence are available in the literature, most of these being described in Beran's (1994) reference book.

3.2.1. Analyses of ACF (auto-correlation functions)

[11] More precisely by Abbas et al. (2010), IMF Working Paper WP/10/245.

[12] The HPDD was compiled by bringing together a number of other databases of individual researchers or institutional bodies, as well as information from official government publications and publications of the League of Nations and the United Nations.

[13] The general government sector consists of all government units and all non-market non-profit institutions that are controlled and mainly financed by government units, comprising the central, state, and local governments. The general government sector does not include public corporations or quasi-corporations (p. 6).

To achieve our purpose and to capture the presence of the long memory or persistency in *d*, we use several methods, beginning by evaluating the auto-correlation functions (ACF). If a variable is mean-reverting (i.e. stationary), this should be detected by this analysis.

We follow Poterba and Summers (1988) and assume that a process is mean-reverting if at some frequency the auto-correlation is negative. However, this is a necessary but not sufficient condition.

Table 1: ACF (auto-correlation)

Lag k	15	16	17	18	19	20	25
Countries with ACF<0	0	30	33	36	41	43	54

Table 1 resumes the evolution of ACF values: at *k=15* all countries of our sample have a positive ACF; the first non-positive values appear at k=16 (for 34% of countries) and at *k=25* we have 54 countries (62% of total). Almost 40% of our countries continue to exhibit positive values for the auto-correlation for a period of 25 years. The first result is quite impressive for the 87 countries: during 15 years the relationship is positive. Moreover, these results confirm a long memory process for debt.

3.2.2. Spectral analysis

Following the Beran et al. (2013) heuristic definition of linear dependence, we study the spectral representation of *d* by using Frequency-domain (spectral) analysis. We have opted for this analysis since it can provide an intuitive frequency-based description of the time series, and can indicate interesting features such as long memory, presence of high frequency variation, and cyclical behavior.

In a series with long-memory the auto-correlations are not necessarily very large but they persist for a long time and they decay at a slow rate. The values of this series tend to stay during relatively long periods at high values, and similarly during relatively long periods at low level values (Beran, 1994; Beran et al., 2013).

We know that any stationary process has a frequency-domain representation besides the usual time-domain representation (Hamilton, 1994, Chapter 6). Representing by ω a particular frequency, our series *d* is defined in terms of a weighted sum of periodic functions:

$$d_t = \mu + \int_0^{\pi} \alpha(\omega)\cos(\omega t)d\omega + \int_0^{\pi} \delta(\omega)\sin(\omega t)d\omega \qquad (1)$$

Usually this representation is used for the recognition of cycles in the data but it also allows us to detect the presence of long memory in the series under study.

Taking γ_j as the auto-covariance of order *j* the population spectrum of *d* is given by:

$$s_d(\omega) = \frac{1}{2\pi}\sum_{j=-\infty}^{\infty}\gamma_j e^{-i\omega j} \tag{2}$$

where $i = \sqrt{-1}$. Beran et al. (2013)[14]: propose the following heuristic definition of linear dependence[15]:

	diverges to infinity	long-run memory
$\omega \to 0,\ s_d(\omega)$	converges to a finite constant	short-run memory
	converges to zero	anti-persistence

In a calculable way for a sample from 1 to T, the sample periodogram is given by:

$$\hat{s}_d(\omega) = \frac{1}{2\pi}\left[\hat{\gamma}_0 + 2\sum_{j=1}^{T-1}\hat{\gamma}_j \cos(\omega j)\right] \tag{3}$$

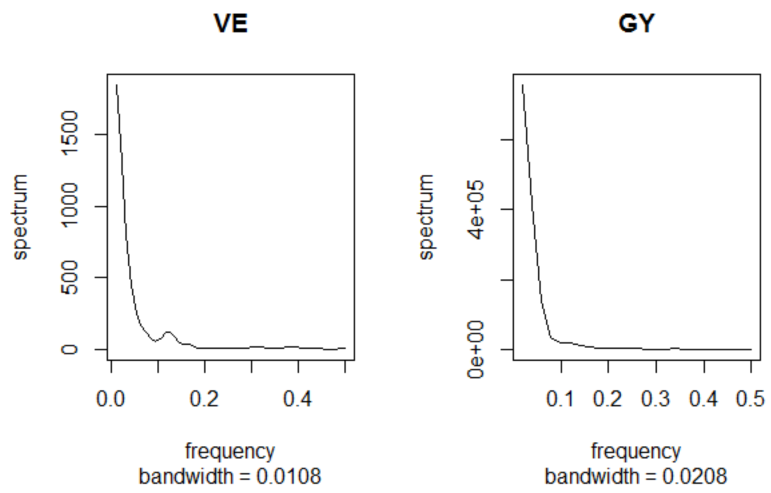
The values of $\hat{s}_d(\omega)$ tell us how much energy is contained within *d* as a function of the frequency. McLeod and Hipel (1978) defined a process of long memory when the limit of the sum of the absolute value of the auto-correlation in not finite. This definition means that the spectral density of a long-memory variable is unbounded at low frequencies.

Spectral Analysis Results

We calculate the spectrum for *d* of the different countries by using the function spectrum[16] of the library stats-package of R Core Team (2015). For all countries of our sample we detect increasing values of the spectrum when frequency $\omega \to 0$ that indicates the presence of long-run memory in *d*. Figure 1 presents two examples of the spectrum evolution of *d* series for Venezuela (VE) that have the lowest value of the median of *d* and for Guyana (GY) that have the highest value of the median. As we see, when the frequency (ω) approaches zero, the spectrum diverges to very high values.

Figure 1. Two examples of the spectral values

[14] Beran et al. (2013), p. 19.
[15] Where we have replaced “long-range” and “short-range” “dependence” by long-run and short-run memory.
[16] We take m=1, and so L=3 for the Daniell kernel.



Note: The spectrum values are not transformed in logs. GY stands for Guyana and VE for Venezuela.

Thus, the results of the spectral analysis applied to all countries of our sample suggest the presence of long memory in *d* series.

3.2.3. Hurst tests: R/S and Whittle

At this point we calculate the rescaled adjusted range statistics (R/S) of Hurst (1951) and Mandelbrot (1972 and 1975) as well as the H estimator of Whittle (1953) for our sample.

In fact, the oldest and best-known method for detecting long memory is the R/S analysis. This method, based on hydrological analysis of Hurst (1951), and lately developed by Mandelbrot (1972 and 1975), allows the calculation of the self-similarity parameter H that measures the *intensity* of long-range dependence in a time series.

If the Hurst exponent is equal to 0.5, it gives an indication of a Brownian motion (random walk), i.e. a random process with no long range memory[17]. Values different from 0.5 and ranging from 0.5 to 1 are indicative of a persistent, trend-reinforcing series (positive long range dependence); positive values that are shorter than 0.5 suggest anti-persistence, (i.e., when a time series reverses itself more often than a random series would). In other words, if the Hurst exponent is $0.5<H<1.0$, the random process will be a long memory process; the $H<0.5$ indicates a short-memory process[18]. This test has been extensively applied to financial data (Hauser, 1997; Weron, 2002; Lillo and Farmer, 2004).

[17] H_0 of the Hurst test is: H = 0.5.
[18] Through Monte Carlo simulation, Hurst noted that if the underlying process is a random draw from a stable distribution, then H = 0.5. If H is greater than 0.5, there is evidence of persistent dependence (large values followed by large values and small values followed by small values) and if H is less than 0.5, an ergodic or mean reverting

The idea behind the rescaled adjusted range statistics (R/S) is to compare the minimum and maximum values of running sums of a normalized series with mean zero[19]. The deviations are larger for long-memory processes than for short-memory processes.

The auto-covariance of a stationary variable *d* with long memory can be represented in the limit, $j \to \infty$, by:

$$\lambda(\mathrm{j}) \sim \mathrm{j}^{-\alpha} L(\mathrm{j}), \quad 0<\alpha<1 \tag{4}$$

$\alpha = 2-2H$ and $L(\mathrm{j})$ is a slowly varying function, $\lim_{j\to\infty} L(tj) \big/ L(j) = 1$. The smaller the value of α the longer the memory in the variable. In the case of a short-memory process $H = 0.5$.

The classical value of the Hurst coefficient, or R/S, is obtained by:

$$R/S = \frac{1}{s_T}\left[\max_{1\le k\le T}\sum_{j=1}^{k}(d_j-\bar{d}) - \min_{1\le k\le T}\sum_{j=1}^{k}(d_j-\bar{d})\right] \tag{5}$$

with s_T the likelihood standard deviation estimator,

$$s_T = \left[\frac{1}{T}\sum_{j=1}^{T}(d_j-\bar{d})^2\right]^{1/2} \tag{6}$$

The first (second) term in brackets in (5) is the maximum (minimum) over *k* of the partial sums of the first *k* deviations of d_j from the sample mean[20]. The difference named “range” is obviously non-negative.

Campbell et al. (1997) refer to several seminal papers of Mandelbrot[21], Taqqu, and Wallis that have demonstrated the superiority of the R/S to other conventional methods for determining long-range dependency such as spectral decomposition, variance ratios or analysis of auto-correlation. However, there are several important shortcomings of this test, i.e. it is sensitive to short-term dependence and heteroscedasticity[22], problems that we might face in our

process is indicated. H = 0.5 – random walk (random process with no long range memory); H ∈ (0, 0.5) – mean reverting; H ∈ (0.5, 1) – mean-averting)

[19] More precisely, the R/S statistic is the range of partial sums of deviations of a time series from its mean, rescaled by its standard deviation.

[20] See Campbell et al. (1997), p.62.

[21] Mandelbrot (1971) was the first to suggest that R/S analysis could be useful in studies of economic data and provided an economic rationale. In Mandelbrot (1972), it was further argued that R/S analysis was superior to auto-correlation and variance analysis since it could consider distributions with infinite variance and was superior to spectral analysis because it could detect non-periodic cycles.

[22] Lo (1991) discusses the lack of robustness of the R/S statistic in the presence of short memory or heteroscedasticity.

empirical analysis. For Campbell et al. (1997), the most important shortcoming of the R/S test is its sensitivity to short-range dependence. Bhattacharya et al. (1983) proved that the R/S test is not robust to departures from stationarity.

Taqqu et al. (1995) propose nine methods to compute the Hurst exponent: (i) the aggregate variance method; (ii) the differenced aggregated variance method; (iii) the aggregate absolute value/moments method; (iv) the Higuchi or fractal dimension method; (v) the Peng or variance of residuals method; (vi) the R/S method (above), (vii) the periodogram method; (vii) the boxed or modified periodogram method; and finally (xix) the Whittle estimator.

Fox and Taqqu (1986), Dalhaus (1989), and Giraitis and Surgailis (1990) proved that the Whittle estimator (Whittle, 1953) is consistent and asymptotically normal for Gaussian long-range dependence. The H statistic by the periodogram method is obtained from the slope of the plots of the logarithm of the spectral density against the logarithm of the frequencies, 1-2H. The minimization of a likelihood function to the above method gives an estimation of H (Whittle) and its confidence intervals[23].

Hurst test results

We calculate *H* values by *R/S* and Whittle for our sample by using packages developed by Pfaff (2008) and Maechler (2015) for *R*. Figures 2 and 3 represent the distribution of H values for the 87 countries. The vertical line is the median value. The second graph of Figure 3 is the distribution of the standard errors associated with the estimated *H* values. For all individual countries, the null of *H=0.5* is rejected at very low p-values of statistical significance.

Figure 2. R/S distribution

[23] A complete presentation is in Beran (2014), pp. 420-31.

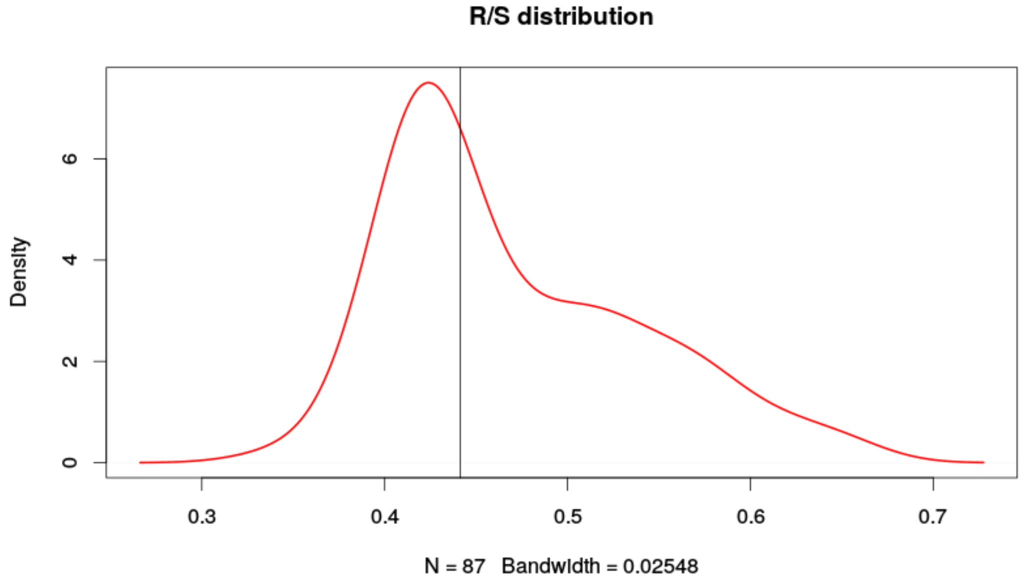


Figure 3. Whittle distribution

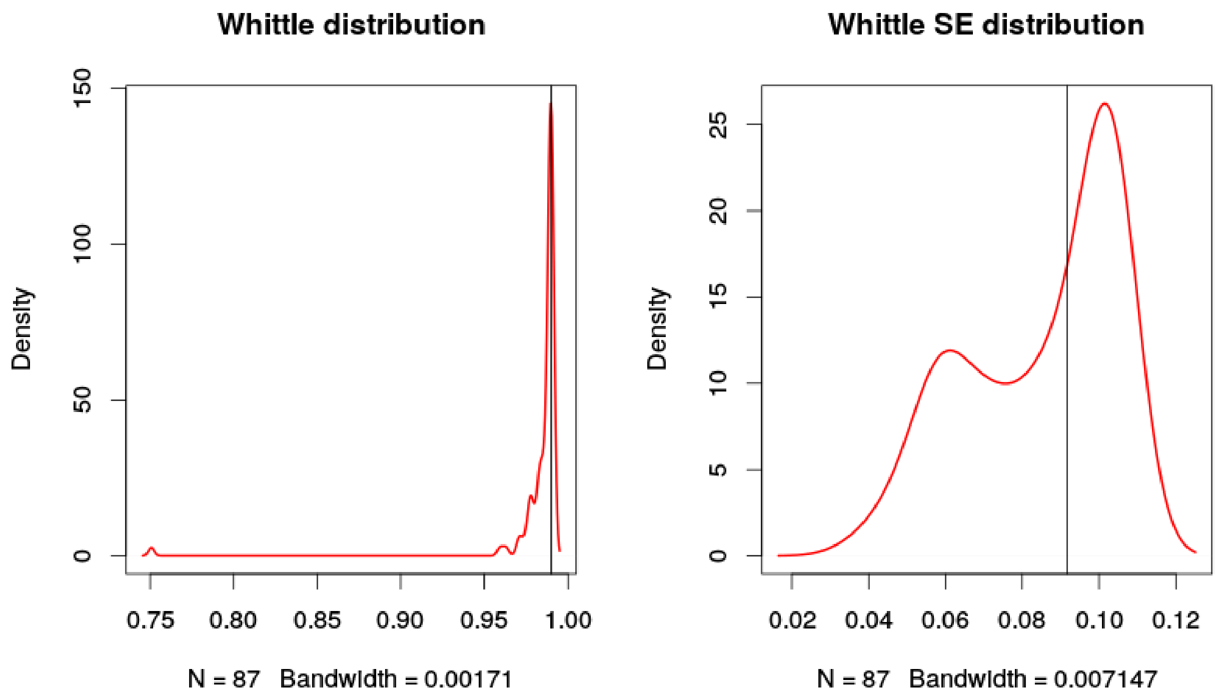


Table 2: H Estimation by the Whittle Method

	Minimum	Median	Mean	Maximum
H (Whittle)	0.7508	0.9899	0.9835	0.9900
Standard-Error	0.038	0.092	0.085	0.104

As we see in Figure 3, and Table 2 that also summarizes our results, the presence of long memory in *d* series is quite obvious (i.e. H exponent is larger than 0.5).

3.2.4. The tests of Geweke and Porter-Hudak and Reisen

In this paragraph, we present the fractionally integrated processes proposed by Granger (1980). We estimate the fractional (or "memory") parameter *d* in the ARFIMA(p,d,q) model, firstly, using the method of Geweke and Porter-Hudak (1983), denoted in the following by (GPH) and, secondly, by the method proposed by Reisen (1994), i.e. smoothed periodogram regression (SPR).

Usually, we consider data-generating processes that are either stationary (i.e. I(0) with the ACF declining exponentially, so that observations separated by a long time span may be considered as totally independent) or integrated of integer order higher than zero (e.g. I(1) with the ACF declining linearly). However, there are some empirically observed times series that share neither of the above characteristics, even if transformed to stationary by appropriate differencing. In addition, these series still exhibit a dependency between distant observations. Granger (1980) theoretically justified these processes and introduced the so-called fractionally integrated processes (i.e. long-memory process).

In order to avoid confusion with the fractional differencing parameter *d*, instead of using d_t for our *d* series (i.e. debt), we use y_t to represent our variable of interest. Suppose we have the following data-generating process (i.e. integrated process of order *d*):

$$(1-L)^d y_t = \mu_t, \quad \mu_t \sim I(0) \tag{7}$$

where *d* is not the integer and represents the fractional order of integration and L is the lag operator.

For $0<d<0.5$, the process y_t is long memory, and its auto-correlations are all positive and exhibit a hyperbolic rate of decay. For -0.5 < d<0 the process has a short memory (i.e. the sum of the absolute values of the auto-correlations tend to a constant, the process is said to be "anti-persistent". In addition, when $-0.5<d<0.5$, we have a covariance-stationary process; $d<1$ we have the characteristic of mean reversion of the process, but if $0.5<d<1$ the process is not covariance stationary; nevertheless, it is still mean reverting (Baillie, 1996).

Geweke and Porter-Hudak (1983) propose calculating *d* by a semi-parametric function known as narrowband least squares (GPH)[24]. The above process can be represented in the frequency domain by:

$$s_y(\omega) = \left|1 - e^{-i\omega}\right|^{-2d} s_\mu(\omega) \tag{8}$$

that in turn after transformation can take the following form:

$$\log[s_y(\omega_j)] = \log[s_\mu(0)] - d\log[4\sin^2(\omega_j/2)] + \log[s_\mu(\omega_j)/s_\mu(0)] \tag{9}$$

The authors suggest estimating *d* by a regression of the ordinates of the log spectral density on a trigonometric function of frequencies:

$$log\left[I_y(\omega_j)\right] = \beta_1 + \beta_2 log\left[4\, sin^2(\omega_j/2)\right] + v_j \tag{10}$$

where $I_y(\omega_j)$ is the periodogram of y_t and $\upsilon_j = \log[s_\mu(\omega_j)/s_\mu(0)]$ ~ $IID(0, \pi^2/6)$. The value of $\hat{d} = -\widehat{\beta_2}$. When μ_t is auto-correlated, the above regression holds approximately for frequencies near zero[25].

Reisen (1994) proposed a modified form of the regression method, based on a smoothed version of the periodogram function (SPR). According to Lopes et al. (2002, 2004), the SPR estimator has better performance than the GPH estimator in the sense of minimizing the mean squared error (mse) values.

GPH and SPR results

In this sub-section, by using the package 'fracdiff' for R developed by Fraley et al. (2015), we propose to estimate *d* by the GPH (Geweke and Porter-Hudak, 1983) and SPR (Reisen, 1994) methods presented above.

The first estimator (GPH) is based on a regression that uses the periodogram function as an estimate of the spectral density. The bandwidth used is $bw = trunc(\mathrm{T}^k)$, where *T* is the number of observations and *k* (*0<k<1*) is a parameter whose default value is 0.5 (Diebold and Rudebusch, 1989).

The second method we use is Reisen's (1994) SPR estimator. This uses the same bandwidth as the first method and we take the value *h*, used in the lag Parzen window, equal to 0.9, $bw2 = trunc(T^h)$.

[24] These authors proposed an estimator of d as the ordinary least squares estimator of the slope parameter in a simple linear regression of the logarithm of the periodogram.

[25] For an exhaustive exposition see Beran (2013), pp. 441-5.

Tables 2 and 3 present the results. We include the 1st and 3rd Quartiles because their values are substantially different from the minimum and maximum respectively.

Table 3: Estimation of *d* by GPH

	Minimum	1stQ	Median	Mean	3rdQ	Maximum
d	0.319	0.836	1.010	1.060	1.220	2.147
C.I.-Inf	-0.298	0.137	0.315	0.334	0.516	1.390
C.I.-Sup	0.855	1.52	1.720	1.780	1.990	2.910
d/Standard-Error	1.23	2.98	4.33	4.44	5.37	10.40

Table 4: Estimation of *d* by SPR

	Minimum	1stQ	Median	Mean	3rdQ	Maximum
D	0.248	0.769	0.937	0.910	1.080	1.338
C.I.-Inf	-0.0081	0.453	0.585	0.575	0.727	0.932
C.I.-Sup	0.488	1.070	1.290	1.250	1.420	1.740
d/Standard-Error	1.42	2.56	3.54	3.90	4.81	11.20

Figure 4: GPH Statistics distribution

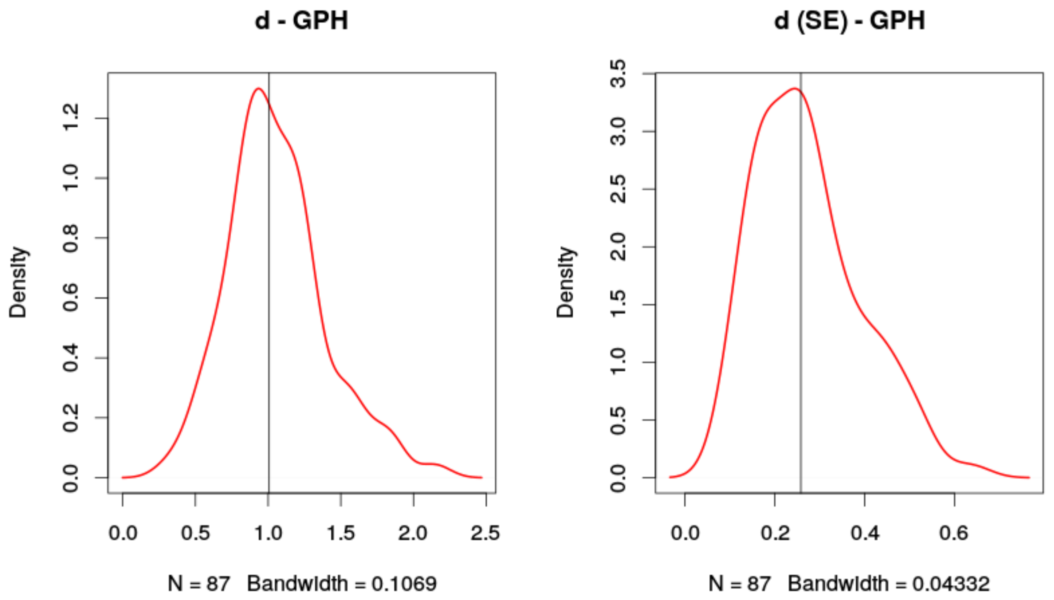
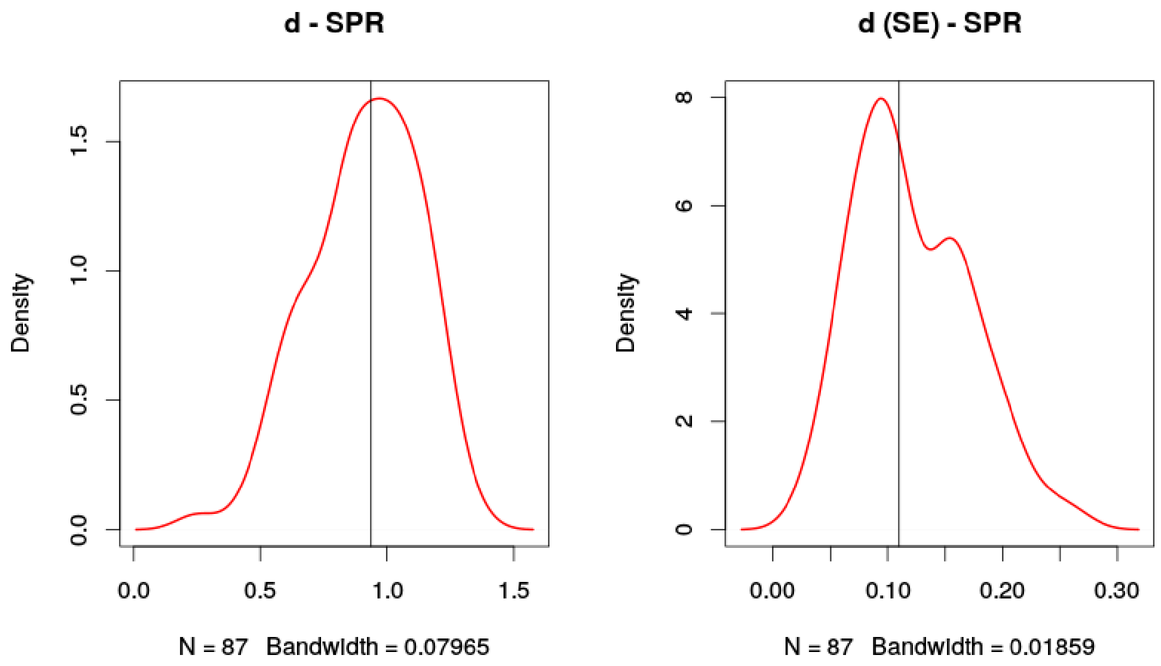


Figure 5: SPR Statistics distribution



Our results (see also Figures 4 and 5) confirm the presence of long memory in y_t (our debt series) since the values of fractional differencing parameter d are positive (d>0) and are outside ($d > 0.5$) the values that characterized a stationary series ($d < 0.5$). As for a few countries that respect the stationary conditions (i.e. when d >0.5), y_t is still characterized by long memory processes (d>0).

3.2.5. The Variance ratio tests

Since the work of Cochrane (1988) and Lo and MacKinlay (1988, 1989), the variance ratio (VR) has been widely used to study the persistence of economic and financial variables. Lo and MacKinlay (1989) proved that in certain circumstances the variance ratio test is more powerful than the Dickey-Fuller or the Box-Pierce tests[26].

The VR methodology consists of testing the Random Walk Hypothesis (RWH) against stationary alternatives, by exploiting the fact that the variance of random walk increments is linear in all sampling intervals.

The idea behind this test is very simple: if the series d_t is stationary the variance for k periods will be k/h times the variance for h periods. Cochrane (1988) and Lo and MacKinlay (1988, 1989) define the VR of order k for a series as:

$$V(k) = 1 + 2\sum_{i=1}^{k-1}\left(\frac{k-1}{k}\right)\rho_i \tag{11}$$

[26] See also Cechetti and Lam (1994).

ρ_i is the i-th lag auto-correlation coefficient of order *i*. In this formulation, the *V(k)* is a particular linear combination of the first *k-1* auto-correlation coefficients (see also Charles and Darné, 2009). A test can be built by considering the statistic based on an estimator of V(k) with the null hypothesis that[27] $\rho_1 = \cdots = \rho_k = 0$, i.e. its values are serially uncorrelated:

$$VR(k) = \frac{\hat{\sigma}^2(k)}{\hat{\sigma}^2(1)} \tag{12}$$

where $\hat{\sigma}^2(1)$ is the unbiased estimator of the one-period return variance using the one-period returns d_t, and is defined as:

$$\hat{\sigma}^2(1) = \frac{1}{(T-1)} \sum_{t=1}^{T} (d_t - \hat{\mu})^2 \tag{13}$$

and $\hat{\sigma}^2(k)$ is:

$$\hat{\sigma}^2(k) = \frac{1}{m} \sum_{t=k}^{T} (d_t + d_{t-1} + \cdots + d_{t-k+1} - k\hat{\mu})^2 \tag{14}$$

with $m = k(T - k + \left(1 - \frac{k}{T}\right)$ and $\hat{\mu}$ - the estimate mean of d_t.

If d_t is a random walk then for all horizons of *k*, the expected values of *VR* $(d_t;k)$ should be equal to unity. A value significantly lower than unity at long horizons of *k* is a characteristic of a mean-reverting series. Conversely, if at long horizons the expected value of *VR* $(d_t;k)$ is significantly higher than unity, the series is said to be “mean averting”, i.e. explosive.

Lo and MacKinlay (1988) propose a different version of the variance ratio test under homoscedastic errors (M_1(k) statistics) and heteroscedastic errors (M_2(k) statistics) with the null hypothesis V(k) =1.

The standard normal test statistics proposed by authors to test the null hypothesis of random walk under the assumption of homoscedasticity is:

$$M_1(k) = \frac{VR(d;k) - 1}{\sqrt{\phi(k)}} \sim N(0,1) \tag{15}$$

the asymptotic variance $\phi(k)$ is equal to:

$$\phi(k) = \frac{2(2k-1)(k-1)}{3kT} \tag{16}$$

The standard normal test statistics under the assumption of heteroscedasticity is:

$$M_2(k) = \frac{VR(d;k) - 1}{\sqrt{\phi^*(k)}} \sim N(0,1), \tag{17}$$

[27] The hypothesis to test random walk against non-random walk is equivalent to testing VR(*k*) = 1 against VR(*k*) ≠ 1.

$$\text{with } \phi^*(k)=\sum_{j=1}^{k-1}\left[\frac{2(k-j)}{k}\right]^2 \frac{\sum_{t=j+1}^{T}(d_t-\hat{\mu})^2(d_{t-j}-\hat{\mu})^2}{\left[\sum_{t=1}^{T}(d_t-\hat{\mu})^2\right]^2} \tag{18}$$

One of the serious shortcomings of the Lo and MacKinlay test is that in a finite sample (i.e. when k is large relative to T) the test statistics are severely biased and right skewed.

Another weakness underlined by Chow and Denning (1993), is that it ignores the joint nature of testing for the RWH. In fact, individual separate tests for k values may be misleading leading to over rejection of the null. To overcome this shortcoming, they proposed Multiple VR tests that allow for the examination of vectors of individual VR tests while controlling for overall test size:

$$MV_h=\sqrt{T}\max_{1\le i\le m}\left|M_h(k_i)\right| \quad \text{for } h=1,2 \tag{19}$$

to test the null of M_1 or M_2 equal to unity. For m values the null is rejected if any one of the VR's is significantly different from unity[28].

The Lo and MacKinlay VR test (1988) results

The first step in the application of the VR(k) tests refers to the non-rejection of heteroscedasticity (i.e. ARCH test)[29]. We apply an ARCH test (Tsay, 2013) where the null of no-ARCH is decided at the critical level of 10% (Table 5). We choose arbitrary $k = 5, 15$ and 25.

Table 5: ARCH Tests (non-rejection)

Number of countries	Lags	SL (median)
4	5	4.5e-07
28	15	0.012
52	25	0.34

For 25 lags, a slight majority of countries does not reject the null of no-ARCH behavior. The presence of the ARCH effect is obvious for almost all countries for 5 lags, and for a third in the case of 15 lags. In the following we use M_1 and M_2 Lo and MacKinlay-type tests by assuming that M_1 is appropriate for small values of k.

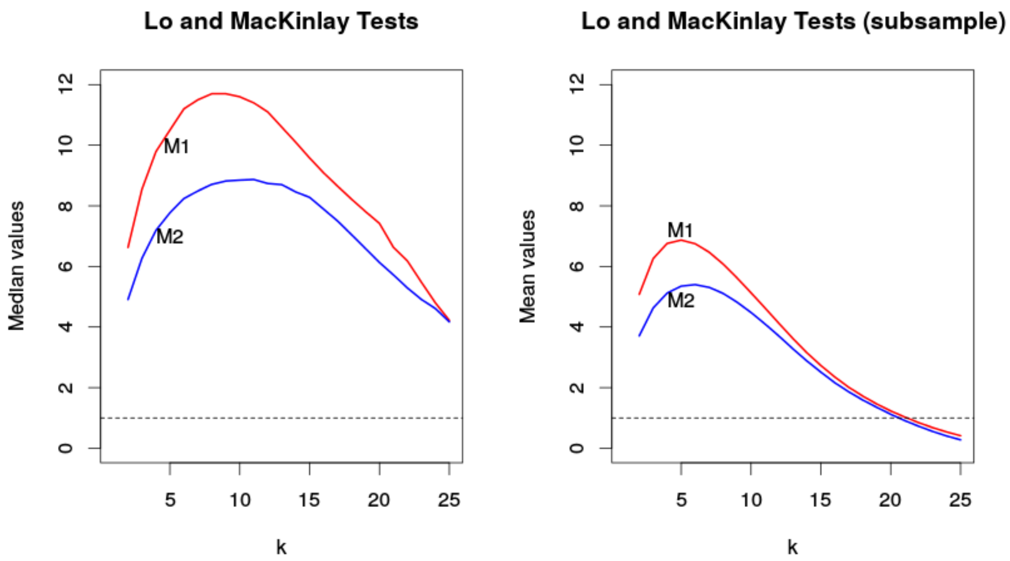
[28] The critical values are in Chow and Denning (1993).

[29] ARCH: the autoregressive conditional heteroscedasticity.

We begin the analysis of the VR by calculating for $M_1(k)$ and $M_2(k)$ statistics for k=2, …, 25. The median of VR for all countries is represented in Figure 6 (left figure). As we can see, for all countries the values of VR are far from unity (that means that the null hypothesis that the variance ratio is unity should be rejected). As expected, the values of M_1 and M_2 converge with an increase in *k*. The variable d_t shows a high level of inertia and, the effects of a positive shock give an expression of being 'explosives' (i.e. 25 years after the initial shock its effects were amplified four times).

The right figure illustrates the mean of $M_1(k)$ and $M_2(k)$ for a group of countries (11 and 10 countries, respectively)[30] with the VR less than unity at $k=25$. For this group the mean equals unity for $k=20$. As for the entire sample, for the groups of 10/11 countries, we obtain similar results with d_t characterized by a high level of inertia; however, it seems that after 5-6 years the effects of a shock are smoothed and will be null after 20-21 years.

Figure 6: Lo and MacKinlay VR Tests



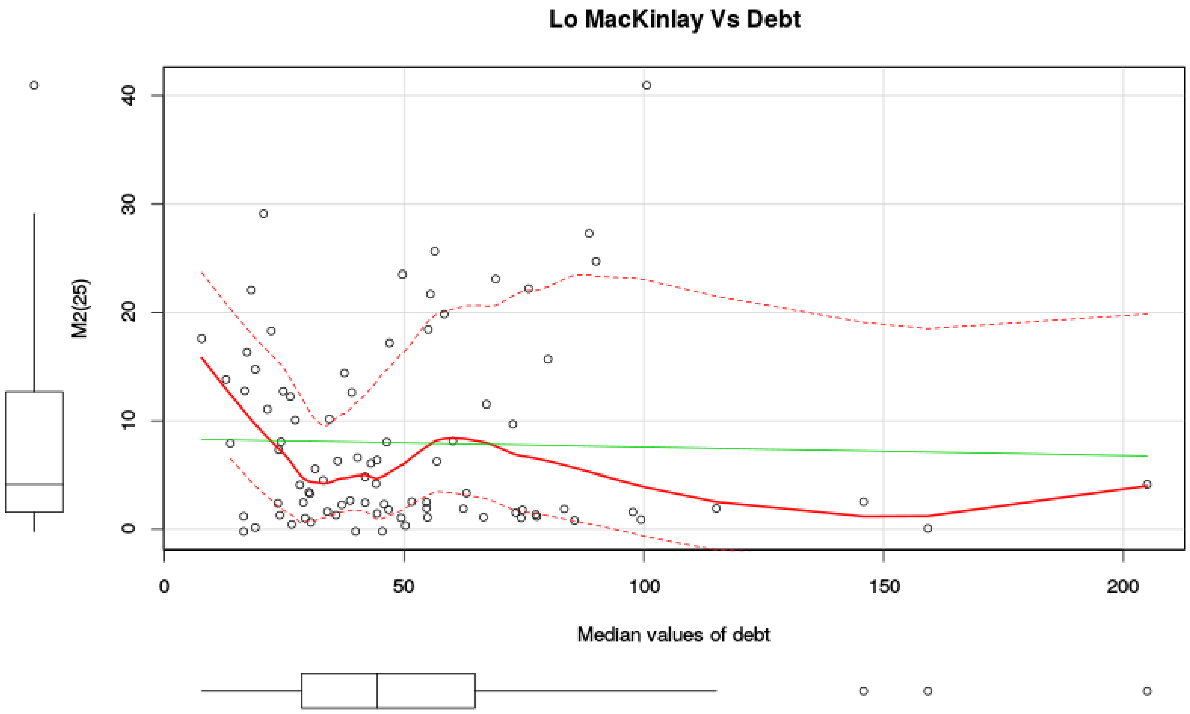
However, it seems to us that these results might hide differences among countries in their responses to shocks depending on the levels of debt. To verify our hypothesis we take the values of the median of debt and the values of $M_2(25)$ statistics (i.e. after 25 years after the initial shock) and represent them together in Figure 7. The box-plot of debt and $M_2(k)$ (left and

[30] We calculate $M_1(k)$ for 11 countries: BB – Belgium; TD – Chile; GA – Germany; GD – Guatemala; IL – Italy; JM – Japan; KR – Malawi; MU – Mexico; NI – Niger; PY – Peru; UY – Venezuela. And $M_2(k)$ for 10 countries: BB – Belgium; TD – Chile; IL – Italy; JM – Japan; KR – Malawi; MU – Mexico; NI – Niger; PY – Peru; SZ – Sweden; UY – Venezuela. These countries were chosen because of the low value of VR (i.e. countries whose d_t series are almost I(0)).

below) show unbalanced sets, and the figure illustrates a very slight decreasing relationship represented by the lowest (solid) line that can be interpreted as follows: the higher the value of the debt, the lower the VR value $M_2(25)$ (i.e. the lower would be the effects of a shock 25 years after). Looking at the overall picture including the 95% confidence interval (represented by dashed lines), the most appropriate conclusion is that there is an absence of any relationship between *d* and VR for the complete group of 87 countries.

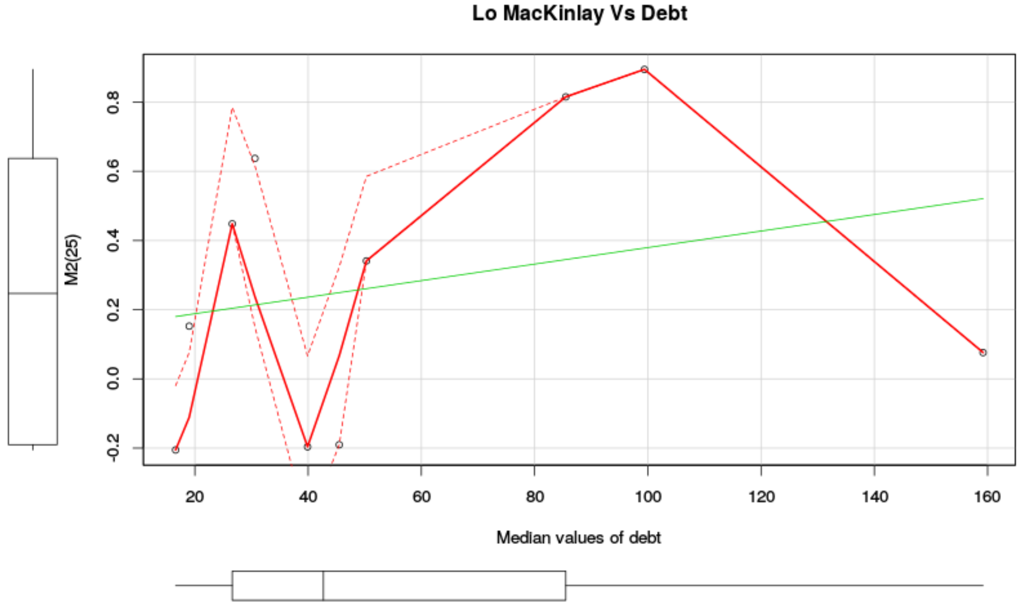
We do the same type of smoothed scatter (Figure 8) for the group of 10 countries that are characterized by a low value of VR (i.e. countries where shock is less persistent), and find a balanced distribution of variance ratios for an unbalanced (median) debt distribution. We obtain a positive relationship between the debt level and the VR.

Figure 7: Lo and MacKinlay Test Results (1)



Note: see Fox and Weisberg (2011) for the scatter plots.

Figure 8: Lo and MacKinlay Test Results (2)

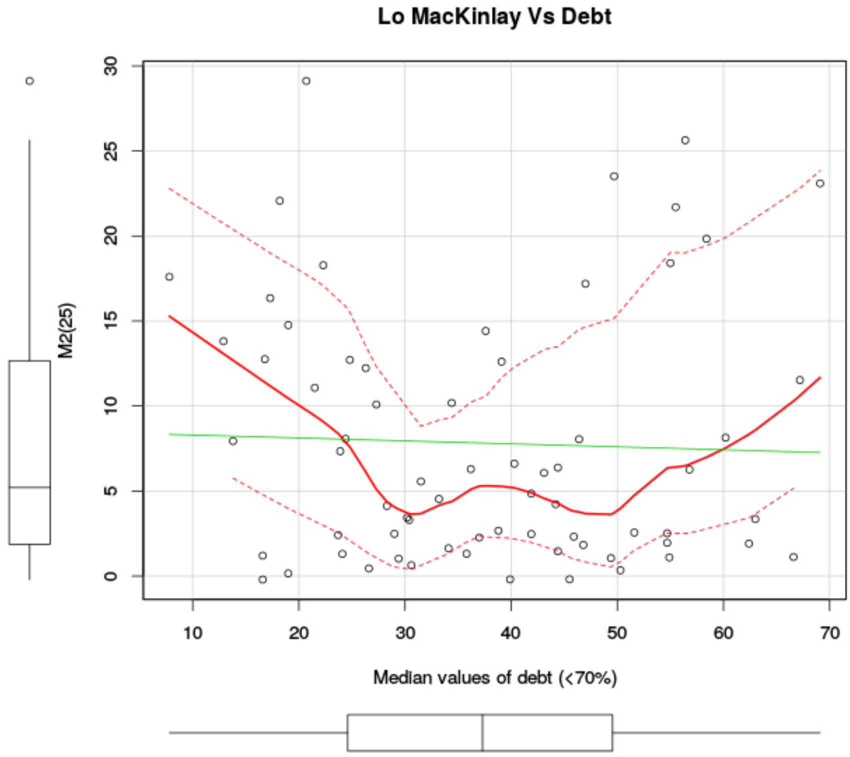


Note: see Figure 7.

Besides this, the standard deviation of this sample is twice the corresponding value for the total countries considered (119.2 and 58.92) for a similar mean (66.3 and 57). A more detailed investigation of Figure 7 tells us that a threshold median value is around 70%.

Therefore, a new scatter is created for these values (Figure 9), from which quite an interesting picture can be observed, e.g. a U-shaped relationship between debt level and VR, that can be interpreted as the absence of a "one-to-one function" of inertia in relation to debt levels.

Figure 9: Lo and MacKinlay Test Results (3)



Note: see Figure 7.

The results for M1(k) and M2(k) are presented in Table 6 and are obtained by using Kim's (2015) method.

Table 6: M_1 and M_2 Lo and MacKinlay Tests

	Minimum	1stQ	Median	Mean	3rdQ	Maximum
$M_1(5)$	2.98	7.67	11.60	13.50	18.40	41.90
$M_1(15)$	-0.12	5.27	9.58	12.80	18.00	48.00
$M_1(25)$	-0.21	1.48	4.22	9.69	13.80	53.10
$M_2(5)$	2.90	6.36	8.85	10.10	13.10	29.40
$M_2(15)$	-0.12	4.90	8.28	10.00	13.70	35.00
$M_2(25)$	-0.20	1.57	4.17	7.96	12.70	41.00

The distribution of the significance levels obtained by bootstrapping (Kim, 2006), consists of near-zero values for the different k values for the M_1(k) and M_2(k) tests. The equivalent significance levels are obtained for the Chow and Denning (1993) joint test, taking again $k = 5$, *15* and *25*. For the wild bootstrap we use the Normal distribution.

Based on the above, we conclude this sub-section by confirming the presence of the long memory in *d*.

3.3. Unit Root (UR) and Stationary Tests

The next step of our study consists of investigating the stationary characteristics of *d* by applying the usual tests of stationarity. If there is some value of *d* for which its values tend to return after a shock, then *d* is stationary[31]. We apply the usual tests of unit root to confirm the existence of a process of mean-reversing or reversing to a determinist trend in *d*. We investigate whether the debt ratios are stationary with a constant, and stationary around a trend. If a country does not reject the hypothesis of the presence of a unit root, this means that a positive shock at some moment is permanent and will not be cancelled on average.

To begin, we apply the usual Augmented Dickey-Fuller (1979) (ADF) tests with drift and with a trend to the *d* series of each country. When we cannot reject the null at 10% of the trend coefficient, we test only the presence of a drift. We know that ADF tests have a low power if the true data-generation process has an auto-regressive coefficient close to 1. The other important problem associated with ADF tests is related to deterministic regressors because they have a different interpretation under the null and alternative hypothesis. So we also apply the Schmidt and Phillips (1992) Lagrange multiplier test (S-P), where the deterministic parameters have the same interpretation under the null or the alternative hypotheses. This test allows the choice of the order of the polynomial trend[32].

However, the first generation of panel unit root tests, through ignoring the presence of CSD, may produce inconsistent results. Therefore, in order to control for the problem of cross-sectional correlation, we use a new generation (second generation[33]) ADF test of Chang (2002) and apply it separately to the individual series and to the entire group. Under the unit root hypothesis, in the context of ADF, $\alpha_i = 1$ we have the well-known equation:

$$y_{i,t} = \alpha_i \cdot y_{i,t-1} + \sum_{k=1}^{p_i} \alpha_{i,k} \cdot \Delta y_{i,t-k} + \epsilon_{i,t} \tag{20}$$

Chang (2002) proposes an instrumental variables (IV) estimation of this equation. More precisely, to deal with CSD he suggests using a non-linear function *F* for the lagged level values of *y*. For the lagged difference, the augmented part of the ADF, he proposes using the variables

[31] However, as we have seen before, a variable may be non-stationary and moreover, mean-reverting.
[32] For ADF and S-P we use the 2015 version of the package 'urca' for R (Pfaff, 2008). We retain the t-values of the tests.
[33] See Hurlin and Mignon (2007).

themselves as instruments. The transformation is named the *instrumental generating function* (IGF). The average IV t-ratio statistic is thus defined as

$$S_n = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} Z_i \tag{21}$$

for the N cross-sectional units, and Z_i is the cross-sectional non-linear IV t-ratio statistic for testing $\alpha_i = 1$ for the i^{th} unit.

There are several advantages of using this test, as noted by Chang (2002) and Breitung and Pesaran (2005): (i) we can apply it for balanced and unbalanced panels; (ii) it is asymptotically Normal; (iii) it is a standardized sum of individual IV t-ratios; and (iv) the non-linear transformations take account of possible contemporaneous dependence among cross-section units[34].

The obtained results are presented in Table 7 (below). For the ADF and Chang tests we used a maximum of 5 lags and the appropriate lag was chosen by the Schwarz Bayesian Information Criterion (BIC). For the S-P test we begin with a polynomial trend of order 4 and reduce the degree until the null hypothesis of its coefficient is not rejected. For the ADF test only four countries reject the null of unit root at 1% significance level; three at 5% and four at 10%, i.e. 11 countries out of 87 reject the unit root at 10%.

For the S-P the picture is not very different, six countries reject the null of unit root at 1% significance level; seven at 5% and four at 10%. This means that in 87 countries only 17 reject the null at 10%.

In respect of the Chang test, no country rejected the unit root at 1% significance level; only three countries out of 87 reject the null of unit root at 5% significance level; and four at 10%. As for the panel, the Chang test does not reject the null of unit root for both hypotheses whether we consider a constant or a constant and trend. Based on the above, we conclude that the variable *d* is not stationary.

Table 7: Results of the Unit Root Tests

	ADF				**S-P**		**Chang**			
Country	***C***		***C,T***				***C***		***C,T***	
AR	-3.97	***			-3.74	***	-1.71	*	-1.74	*
AU	-2.47				-3.02	*	-1.92	*	-1.42	

[34] The package "pdR" of Tsung-wu (2015) was used to compute separate tests for individual countries and for the panel (as a simple average of country values).

AT	-0.72				-1.61		0.91		0.27	
BS			-1.94		-2.52		-0.80		-0.23	
BB			-2.03		-2.34		-0.45		-0.53	
BE			-4.72	***	-3.70	***	-0.63		-0.60	
BJ			-1.63		-0.78		0.93		0.29	
BO			-2.56		-1.87		1.19		0.26	
BI			-2.4		-1.7		-1.43		-0.85	
CM	-1.37				-2.04		0.04		-0.01	
CA	-2.66	*			-3.40	**	0.10		0.15	
CF	-1.22				-1.40		-0.39		0.17	
TD	-1.79				-2.49		0.26		0.43	
CL			-3.69	**	-2.46		0.68		-0.07	
CO			-2.53		-2.66		-1.30		-1.56	
CD	-1.38				-1.46		0.00		0.56	
CG	-2.01				-0.96		0.32		0.35	
CR	-2.49				-2.72		0.19		-0.13	
CY			-2.91		-2.76		0.25		0.40	
DK			-2.84		-3.33	**	-2.01	**	-2.43	**
DO	-1.75				-2.48		-1.42		-1.47	
EC	-1.61				-2.18		0.28		0.28	
EG			-2.35		-2.26		0.69		0.82	
SV	-1.67				-3.54	**	-0.49		-0.70	
FJ			-4.55	***	-2.23		-0.05		0.58	
FI	-2.22				-3.19	**	0.09		-0.17	
FR	-1.70				-2.35		-1.36		-1.31	
GA	-1.82				-1.77		-0.22		0.20	
DE	-0.16				-3.10		1.52		1.03	
GH	-1.88				-2.63		0.02		-0.05	
GR	-1.51				-2.51		0.77		0.67	
GD			-2.54		-2.33		-1.31		-1.54	
GT	-1.92				-3.10	**	-0.08		0.11	
GY	-1.40				-1.77		0.52		-0.52	
HT	-2.42				-2.77	*	-1.20		-1.36	
HN	-1.11				-2.03		0.11		0.33	
IN	-1.14				-2.24		-0.19		-0.28	
IE			-3.13		-3.25		-2.18	**	-2.24	**

IL			-2.98		-1.07		0.96		0.38	
IT	-1.99				-2.15		-1.01		-0.99	
JM	-1.65				-2.76		-0.85		-0.84	
JP			-1.02		-3.41		2.04		1.88	
JO	-2.19				-2.21		0.19		0.05	
KE	-1.40				-0.89		0.79		-0.46	
KR	-1.01				-2.67		-1.06		-1.00	
MW	-2.01				-1.85		-0.91		-0.79	
MY	-2.16				-2.74		-0.53		-0.92	
ML			-1.45		-1.74		0.37		0.09	
MT			-2.6		-0.87		1.60		-0.64	
MU	-3.07	**			-2.43		-1.69	*	-2.05	**
MX	-1.70				-1.89		0.18		0.13	
MA	-1.84				-3.30		0.21		-0.33	
NP			-0.16		-0.12		0.00		-0.51	
NL			-2.43		-2.60		-1.52		-1.69	*
NZ			-2.41		-3.82	*	0.70		-0.77	
NI	-2.50				-6.05	***	1.31		1.08	
NE	-1.57				-2.35		0.32		0.32	
NG	-1.65				-1.84		1.44		1.08	
NO			-4.1	***	-3.97	***	-0.42		-0.39	
PK	-2.03				-2.51		0.60		1.02	
PA	-2.10				-2.34		0.37		0.33	
PY	-2.95	**			-2.36		0.81		0.58	
PE	-2.10				-2.32		0.62		1.12	
PH	-1.43				-2.79		0.72		-1.18	
PT			1.53		-4.91	**	-1.13		-0.67	
RW	-1.72				-1.89		-0.27		-0.33	
WS			-2.64		-2.78		0.08		-0.74	
SN	-1.90				-2.89		0.77		0.14	
SL	-1.22				-2.47		0.03		0.35	
SG			-3.34	*	-1.18		-1.50		-1.23	
ZA			-3.39	*	-3.28		1.23		1.81	
ES	-1.52				-4.00		0.56		0.96	
LK	-1.89				-2.47		-0.55		-0.47	
SZ	-2.47				-1.91		-0.42		-1.29	

SE			-3.29	*	-3.41		0.25		0.21	
CH	-2.02				-3.68		-0.75		-0.68	
TZ	-0.81				-2.68		-0.99		-1.63	
TH			-3.96	**	-3.72	***	-0.72		-0.71	
TT	-1.43				-1.59		-1.24		-0.97	
TN			-0.62		-2.31		-0.01		-0.31	
TR			-2.98		-2.62		-1.50		-1.94	*
UG	-1.70				-1.28		0.32		0.47	
UK	-2.48				-3.02	*	0.06		-0.36	
US			-2.68		-4.07	***	0.52		-0.40	
UY	-2.46				-3.13	**	-1.15		-1.08	
VE			-2.74		-3.15		1.37		-1.51	
ZM	-1.29				-2.12		-1.09		-1.27	
Mean							*-0.13*		*-0.33*	

Note: the stars have the usual meaning, *** for rejection at least of 1% of significance level, ** for 5% and * for 10%.

Therefore, we continue our analysis of the hypothesis that d_t is non-stationary in levels and stationary after differencing, i.e. it is integrated of order one, I(1). Assuming that Y_t has the same order of integration as d_t we should implement a test for panel co-integration between these two variables.

3.4. Co-integration Tests

Once the order of stationarity was defined, the next step of our empirical analysis was to test whether d_t and Y_t are co-integrated. To do so, we apply the Westerlund (2007) tests for co-integration to our data (panel of 87 countries over the period 1970-2012). For this test we use panel data since the additional cross-sectional components incorporated in panel data models provide better properties of panel co-integration tests compared with standard co-integration tests for time series samples. First, we consider d_t and Y_t alone and, secondly, by conditioning them on the following variables: the level of merchandise exports on GDP (*x*), external trade on GDP (*op*), and gross fixed capital formation on GDP (*gfc*).

The four Westerlund (2007) panel tests do not impose any common-factor restriction since the four tests are based on structural rather than residual dynamics; and the bootstrap method takes into account the presence of cross-sectional dependence[35].

[35] In fact, some of the first panel tests required that the long-run parameters for the variables in their levels should be equal to the short-run parameters for the variables in their differences, the so called common-factor restriction.

If the null of the error-correction term (ECM) in a conditional panel error-correction model is not rejected then we should accept the null hypothesis of no co-integration.

The ECM equation is given by:

$$\Delta y_{it} = \delta'_i d_t + \alpha_i y_{i,t-1} + \lambda' x_{i,t-1} + \sum_{j=1}^{p_i} \alpha_{ij} \Delta y_{i,t-j} + \sum_{j=-1}^{-p_i} \gamma_{ij} \Delta x_{i,t-j} + e_{it} \tag{22}$$

where $\lambda'_i = -\alpha_i \beta'_i$. The parameter α_i determines the speed at which the system corrects back to the equilibrium relationship, $y_{i,t-1} - \beta'_i x_{i,t-1}$, after a sudden shock. If $\alpha_i<0$, then there is error correction, which implies that y_{it} and x_{it} are co-integrated; if $\alpha_i=0$, then there is no error correction and, thus, no co-integration (Persyn and Westerlund, 2008, p. 233). The β_i are the beta coefficients of the long-run relationship; indexes *i* refer to the *N* individuals and *t* to the T time periods and *d* refers to the deterministic components. In the ECM equation the lags and leads can vary across individuals. After the estimation of (22) we compute:

$$\hat{u}_{it} = \sum_{j=-q_i}^{p_i} \hat{\gamma}_{ij} \Delta x_{i,t-j} + \hat{e}_{it} \tag{23}$$

The proposed group-mean tests are calculated as:

$$G_\tau = \frac{1}{N} \sum_{i=1}^{N} \frac{\hat{\alpha}_i}{SE(\hat{\alpha}_i)} \text{ and } G_\alpha = \frac{1}{N} \sum_{i=1}^{N} \frac{T\hat{\alpha}_i}{\hat{\alpha}_i(1)}. \tag{24}$$

Where $\hat{\alpha}_i(1) = \hat{\omega}_{ui} / \hat{\omega}_{yi}$ and $\hat{\omega}_{ui}$ and $\hat{\omega}_{yi}$ are the Newey and West (1994) long-run variance estimators based on $\hat{u}_{it}$ and Δy_{it} (Westerlund, 2007).

The null of the test is that $\alpha_i = 0 \ \forall i$ against $\alpha_i < 0$ for at least one *i*. The panel tests proposed by Westerlund take the $\alpha_i = \alpha$ for all the *i* and the alternative $\alpha_i = \alpha < 0, \ \forall$i [36]. Table 8 reports the group-mean co-integration tests results (that are also confirmed by those of the panel tests). These results confirm the absence of any co-integration involving *Y* and *d*. For the first analysis of the relationship between Y_t and d_t we used data for 87 countries; as for the

The failure of this requirement can cause a significant loss of power for residual-based co-integration tests (Kremes et al., 1992).

[36] As we said, the null hypothesis of no co-integration is H_0: $\alpha_i = 0$ for all *i*. The alternative hypothesis depends on what is being assumed about the homogeneity of αi. Two of the tests, called group-mean tests, do not require the αis to be equal, which means that H_0 is tested versus H_{g1}: $\alpha_i<0$ for at least one *i*. The second pair of tests, called panel tests, assume that α_i is equal for all *i* and are, therefore, designed to test H_0 versus H_{p1}: $\alpha_i=\alpha<0$ for all *i*. (Persyn and Westerlund, 2008, p. 233)

analysis of the relationship between *Y*, *d*, and *x* or *op*, we used data for 74 countries; and for *y*, *d* and *gfc* data for 60 countries was (a list of the countries appears in Appendix A).

Table 8: Co-integration Tests Results

		G_{tau}			G_{alpha}				
Variables	Det	Value	Z-Value	P-value	Value	Z-Value	P-value	L	ML
ly, d	C	-1.49	2.96	0.980	-3.26	6.65	1.000	1.61-1.51	5
ly, d	C, T	-2.61	-2.77	0.730	-9.16	3.92	1.000	1.75-1.44	5
ly, d, x	C	1.13	29.53	0.970	-0.16	12.29	0.470	4.84-4.69	5
ly, d, x	C, T	-2.57	-0.46	0.950	-6.09	8.85	1.000	2.55-2.2	4
ly. d, op	C	-1.64	3.67	0.990	-3.11	8.25	1.000	2.3-2.09	4
ly. d, op	C, T	-2.52	0.06	0.930	-5.70	9.30	0.990	2.57-2.41	4
ly. d, gfc	C	-1.23	6.72	1.000	-2.58	8.09	1.000	1.43-1.3	3
ly. d, gfc	C, T	-2.57	-0.37	0.890	-7.95	6.00	1.000	1.52-1.38	3

Note: the column Det refers to the deterministic component, C for constant and T for trend; the P-value is the robust P-value, taking in account the cross-sectional dependence; L represents the mean lag and mean lead applied to the different countries; and ML represents the maximum value of the lags and leads chosen by the Akaike criterion (as the sample was reduced after the second line, we have to also reduce this maximum value).

4. Is there a threshold value in the relation between growth and debt?

As we have already mentioned, the presence of debt thresholds and, more in general, the presence of a non-monotone relationship between debt and growth may have very dramatic policy implications. In fact, it is widely admitted that when the level of debt exceeds a certain threshold, its impact on growth is negative. From this, it is easy to define a quite simple rule to be followed by national governments – "if a country wants to promote future growth and to prosper – its debt level should not exceed a certain percentage of GDP" (around 90% of GDP, according to RR).

Therefore, the second main objective of this paper is, by taking into the account the statistical properties of debt series that has been analyzed in the previous section (i.e. presence of long memory and non-stationarity[37]), to investigate (i) whether the debt-growth relation varies with the level of indebtedness or not and (ii) whether there is a common threshold for government debt ratios above which long-term growth rates may drop off significantly (for 60

[37] In the previous sections, we have demonstrated the presence of long memory in debt series (that should be properly taken into account while studying "growth-debt" relationship); as well its non-stationarity (that do not allow researchers to exploit a relation between economic growth and debt by using econometric methods developed for stationary variables). We have also found that output and debt series are not cointegrated.

countries over 1970-2012 period). We have reduced the number of countries based on data availability for control variables.

To do so, we apply the Hansen (1999) threshold model that aims to capture non-linearities in the "debt-growth" relationship to a sample of 60 countries over 1970-2012 period. Hansen proposed an estimation methodology for the identification of different regimes based on the tests for existence of the thresholds. We have opted for this method since it permits us to identify different regimes of explanatory variables that have been selected according to statistical criteria. In other words, the threshold model allows us to split the sample into different groups that potentially capture a non-linear effect of different explanatory variables on growth and, therefore, to determine specific policy implications for different regimes.

To test the presence of different regimes in the relation between growth and debt we consider four different types of the threshold variable (and not only one as is usually done in the literature). They are:

(a) the relative weight of debt (i.e. the variable that has been preferred so far in the existing literature);
(b) the level of the GDP per capita, hereafter YRPC;
(c) a variable representing the time, (i.e. Year);
(d) and finally, the instability of the relative weight of debt represented by the sum of its two consecutive growth rates ($d\_D_t+d\_D_{t-1}$) as its proxy.

The literature that we have already cited above in this article refers to (a). From a policy point of view the relative weight of debt is the most appealing variable, i.e. it corresponds to the limit above which the debt would have a negative influence on growth. From this perspective, the fact that country maintains its debt level below this limit should be considered as an evidence of a "good governance", and, more generally, as an imperative rule that must be respected by countries' authorities.

The second (b) translates the importance of national wealth considered as a collateral for debt. And because a high level of wealth is usually correlated with good institutional practices it would represent also the proper role of capitalist institutions on debt markets.

Since 1970s, the world has experienced dramatical shifts in financial system and in politics, and therefore, has known different political regimes and various financial structures. That is why, it seems useful to us to test for the existence of "points" of time (c) that could possibly change the relation between debt and growth overtime and, if so, to understand why?

Finally, we propose (d) to shed some light to the fact that some countries have historically high levels of debt without being considered as financially unsustainable and others, whose national debt increased rapidly only recently, have raised suspicions on the international financial market about their solvency. Thus, we want to test if the short-term instability, measured by the sum of only two variations of debt weight, would influence the "growth-debt" relationship.

In order to apply the test of Hansen (1999) for the rejection of threshold values in fixed-effects models, we use the pdR package of Ho Tsung-Wu (2017). A set of control variables is included in the estimations: namely, *x* (the level of merchandise exports on GDP); *gfc* (gross fixed capital formation on GDP), *gc* (government consumption on GDP) and *hc* (human capital indicator). This variables have been previously transformed into stationary variables (i.e. I(1)). The first step is to choose between pooling, fixed-effects (FE) and random-effects (RE) models. This methodology is adequate for several reasons: first, because our sample is quite large (i.e. as we said above, we have data for 60 individuals over 1970-2012 period); second, because the literature confirms a simple threshold imposes a homogeneous econometric estimation strategy. The results obtained by the LM and Hausman tests suggests using the FE model. To reduce the number of individual dummies we have excluded in a second moment those that doesn't reject the nulls of those tests. The coefficients of our models are robust to heteroskedasticity, Kleiber and Zeileis (2008) and Lumley and Zeileis (2015).

Our point of departure is estimation of following models: first, we estimate a basic model (without "regimes" based on thresholds) (9) and a model based on them (10), where the observations are divided into several "regimes" depending on whether the threshold variable (TV) is smaller or larger than the threshold γ . We take the limiting case of three thresholds. The suffix "*d_*" is used to represent first differences. On the left side, we have a polynomial lag of order 3 of the dependent variable.

$$\begin{aligned}\alpha^3(L)g_{i,t} &= \alpha_i + \beta_1 d\_x_{i,t} + \beta_2 d\_gfc_{i,t} + \beta_3 d\_hc_{i,t} + \beta_4 d\_gc_{i,t} + \beta_5 d\_D_{i,t} + \\ &+ \beta_6 d\_D_{i,t-1} + \sum_j^k \mu_j DUM_{i,t}\end{aligned} \tag{9}$$

$$\begin{aligned}\alpha^3(L)g_{i,t} &= \alpha_i + \beta_1 d\_x_{i,t} + \beta_2 d\_gfc_{i,t} + \beta_3 d\_hc_{i,t} + \beta_4 d\_gc_{i,t} + \theta_1\left(d\_D_{i,t} I_1(\cdot)\right) + \\ &+ \theta_2\left(d\_D_{i,t} I_2(\cdot)\right) + \theta_3\left(d\_D_{i,t} I_3(\cdot)\right) + \theta_4\left(d\_D_{i,t} I_4(\cdot)\right) + \theta_1\left(d\_D_{i,t-1} I_1(\cdot)\right) + \\ &+ \theta_2\left(d\_D_{i,t-1} I_2(\cdot)\right) + \theta_3\left(d\_D_{i,t-1} I_3(\cdot)\right) + \theta_4\left(d\_D_{i,t-1} I_4(\cdot)\right) + \sum_j^k \mu_j DUM_{i,t}\end{aligned} \tag{10}$$

The indicator function $I_k(\cdot)$ is defined as follows, where $\gamma_1,\ \gamma_2,\ \gamma_3$ are the values of the selected threshold variable:

$$I_1(\cdot)=\begin{cases}1, & \text{when} \quad TV \le \gamma_1 \\ 0, & \text{otherwise}\end{cases}, \qquad I_2(\cdot)=\begin{cases}1, & \text{when} \quad \gamma_1 < TV \le \gamma_2 \\ 0, & \text{otherwise}\end{cases}$$

$$I_3(\cdot)=\begin{cases}1, & \text{when} \quad \gamma_2 < TV \le \gamma_3 \\ 0, & \text{otherwise}\end{cases}, \quad I_4(\cdot)=\begin{cases}1, & \text{when} \quad \gamma_3 < TV \\ 0, & \text{otherwise}\end{cases}$$

We test three variants of these models. Our model 1 (m1) has the current and one lag of the first difference of debt, model 2 (m2) has only the current value of the first difference of debt and model 3 (m3) is model 2 without the variable "*d_x*" when the null (i.e. H_0: coefficient is statistically different from zero) of it is not rejected.

The estimations of our basic model are presented in Table 9 where we have omitted the individual dummies. The zero of the first lag of the first difference of debt cannot be rejected but by a LM test we can exclude the null of the sum of the current and lag value of the variation of debt. The process of adjustment of GDPpc growth is very slowly (0.110[38]) and debt has a negative effect on growth. For all our estimated models the effect was always negative.

Table 9: Basic model (m1) without thresholds

	Coeff.	T	SL
d_YRPC_1	.177	3.108	.002
d_YRPC 2	-.178	1.805	.071
d_YRPC_3	.111	3.107	.002
d_X	.001	1.881	.060
d_GFC	.002	2.313	.021
d_HC	.716	7.431	.000
d_GC_1	.171	2.138	.033
d_D	-.124	-4.981	.000
d_D_1	-.023	-1.325	.185
SEE	0.07012		

Note: where T is the robust (HC1) T value and SL indicates significance level.

4.1 First choice for the threshold variable: d (debt to GDP ratio)

The Hansen tests (Likelihood-ratio test, LR) are obtained for 100, 200 and 300 bootstrap simulation. The values are presented in Table 10.

Table 10: Threshold LR test for m1 with TV=D

D	LR	SL
115.7	33.5	***

[38]The sum of three coefficients (d_YRPC_1, d_YRPC_2 and d_YRPC_3).

101.7, 115.7	47.2	***
101.7, 113, 115.7	42.7	***

Note: TV – threshold variable; *** indicates significant at 1%.

As we can see, the presence of one, two and three thresholds is not rejected. The three values of debt for the changes of regimes are 102%, 113% and 116%. All of them are greater than the conventional 90% of GDP. We estimate accordingly a model with 4 regimes (r.1 to r.4, for the variable d_D and d_D_1).

Table 11: m1 with four regimes (TV=D)

	Coeff.	T	SL
d_YRPC_1	0.178	3.121	0.002
d_YRPC 2	-0.179	1.812	0.070
d_YRPC_3	0.112	3.111	0.002
d_X	0.001	1.577	0.115
d_GFC	0.002	2.594	0.010
d_HC	0.729	7.472	0.000
d_GC_1	0.166	2.122	0.034
d_D.r.1	-0.159	-4.894	0.000
d_D.r.2	-0.316	-4.171	0.000
d_D.r.3	-0.996	-1.950	0.051
d_D.r.4	-0.033	-1.264	0.206
d_D_1.r.1	-0.002	-0.102	0.919
d_D_1.r.2	0.111	2.023	0.043
d_D_1.r.3	0.751	1.148	0.251
d_D_1.r.4	-0.055	-1.637	0.102
SEE	0.06913		
Sum of coefficients:			
d_D.r.1	d_D_1.r.1	-0.162	***
d_D.r.2	d_D_1.r.2	-0.205	**
d_D.r.3	d_D_1.r.3	-0.245	
d_D.r.4	d_D_1.r.4	-0.087	***

Note: T is the robust (HC1) T value, SL is the significance level; SEE is the standard error of the estimate. d_D.r.1 is the first lag of the variable d_D in regime 1.

All control variables have the expected signs (see Table 11). The dominant literature says that the coefficients of debt should increase with the change of regime, with the growing value of the debt. This does not happen. When debt is between 102% and 113% (-0.205) the negative effect of debt is higher than for the situation when debt is less than 102% (-0.162). This situation is in accordance with the literature but after that, we have an absence of effect when debt is between 113% and 116% and the lowest negative effect of debt on growth for the situations of debt greater than 116% (-0.087).

However, the estimation in Table 11 presents two weaknesses: the non-significance of d_x and of 3 out of 4 of the coefficients of lagged debt variables. Therefore, we estimate now the same model but without these lagged debt variables. While estimating m2 the null of "*d_x*" is rejected when there are no regimes but when we consider the presence of regimes in accordance with the thresholds values it continues not to be rejected. Therefore, we estimate the m3 model.

Table 12: Basic model (m3) without thresholds (TV=D)

	Coeff.	T	SL
d_YRPC_1	0.183	3.202	0.001
d_YRPC_2	-0.178	1.848	0.065
d_YRPC_3	0.111	3.067	0.002
d_GFC	0.002	2.166	0.030
d_HC	0.717	7.593	0.000
D_GC_1	0.160	2.011	0.044
d_D	-0.124	-5.071	0.000
SEE	0.0703		

Note: where T is the robust (HC1) T value; SL indicates significance level and SEE states for Standard error of the estimate.

The two main features of model m1 are also represented here: the low value of lagged growth (0.116) and the negative sign of the debt coefficient lower than above but still negative. The exclusion of exports reduces this influence.

Table 13: Threshold LR test for m3 with TV=D

D	LR	SL
115.7	32.8	***
101.7, 115.7	42.1	***
101.7, 113, 115.7	37.3	**

Note: TV – threshold variable; *** and ** indicate significant at 1% and 5% respectively

The estimation of threshold values allows us to continue to considerer 4 regimes (3 threshold values). The threshold values are the same of those obtained in the previous model (m1).

Table 14: m3 with four regimes (TV=D)

	Coeff.	T	SL
d_YRPC_1	0.182	3.205	0.001
d_YRPC_2	-0.180	1.865	0.062
d_YRPC_3	0.111	3.086	0.002
d_GFC	0.002	2.354	0.019
d_HC	0.728	7.637	0.000
d_GC_1	0.167	2.103	0.036
d_D.r.1	-0.161	-5.075	0.000
d_D.r.2	-0.294	-3.355	0.001
d_D.r.3	-0.819	-1.711	0.087
d_D.r.4	-0.030	-1.180	0.238
SEE	0.06938		

Note: where T is the robust (HC1) T value; SL indicates significance level and SEE states for Standard error of the estimate.

We continue to register the low value of adjustment of the growth variable (0.113). We have now the non-rejection of the null for the regime of debt values over 116% and the negative effect of debt is growing from regime 1 to regime 2 and 3. Our results doesn't confirm a presence of a threshold dividing to the left of that value a positive effect of debt on growth and to the right a negative effect. The first three regimes the negative effect of debt is growing with the level of debt but we register an absence of effect for situations of very high debt. In conclusion, we cannot take as evident that a threshold based on debt values is a good choice to understand the relation between growth and debt.

4.2 Second choice for the threshold variable: YRPC (level of GDP per capita)

As we have mention earlier, another possibility is a difference in behavior of debt resulting from the level of GDPpc. This effect may be an indirect one, i.e. via the quality of institutions and the risk that markets consider for those countries. The threshold tests presented in Table 15 indicate that we can retain the presence of three thresholds and estimate a model with four regimes.

Table 15: Threshold LR test for m1 with TV=YRPC

YRPC	LR	SL
3250.896	26.3	**
3250.896, 24604.22	12.7	
3250.896, 13704.97, 24604.22	37.2	*

Note: TV – threshold variable; * indicates significant at 10%.

As we can see, the low value of the lags of growth is maintained (0.113), Table 16.

Table 16: m1 with four regimes (TV=YRPC)

	Coeff.	T	SL
d_YRPC_1	0.180	3.129	0.002
d_YRPC_2	-0.180	1.828	0.068
d_YRPC_3	0.113	3.143	0.002
d_X	0.001	2.027	0.043
d_GFC	0.002	1.912	0.056
d_HC	0.710	7.391	0.000
d_GC_1	0.175	2.220	0.027
d_D.r.1	-0.066	-1.679	0.093
d_D.r.2	-0.253	-9.890	0.000
d_D.r.3	-0.038	-1.461	0.144
d_D.r.4	-0.435	-7.643	0.000
d_D_1.r.1	-0.041	-1.508	0.132
d_D_1.r.2	0.008	0.283	0.777
d_D_1.r.3	-0.030	-1.487	0.137
d_D_1.r.4	0.204	3.499	0.000
SEE	0.06944		
Sum of coefficients		LM - Test	
d_D.r.1	d_D_1.r.1	-0.107	***
d_D.r.2	d_D_1.r.2	-0.245	***
d_D.r.3	d_D_1.r.3	-0.068	
d_D.r.4	d_D_1.r.4	-0.231	***

Note: the hypothesis of equal values for debt in regimes 2 and 4 is rejected.

We expect that ~~as~~ if the income is higher the negative effect of debt would be lower. The four regimes are well represented in terms of observations, 27%, 37%, 22% and 15% for regime 1, 2, 3 and 4 respectively. The higher incomes group has a coefficient of debt greater than a coefficient of the lowest income group (0.231 against 0.107), and this goes against common sense. The 3rd regime registers no relation between debt and growth. Based on this, it seems not reasonable to choose the income level as a threshold variable defining different regimes for the relation between growth and debt.

The choice of YRPC, as it was done above, leads us to select observations for one country in different regimes; to solve this problem the mean of YRPC from 1990 to 2007 is defined as a threshold variable and so every country will belong to one identified regime and not to more than one as above. The tests of the threshold values are presented in Table 17.

Table 17: Threshold LR test for m1 with TV=YRPC (90-07)

	LR	SL
781.8556	35.0	***
781.8556, 880.3093	15.9	***
781.8556, 880.3093, 3545.663	62.2	***

We don't reject the possibility of three thresholds with the four regimes associated. The distribution of observations from regime 1 to regime 4 is the following: 6.4%, 3.9%, 14.2% and 75.4%, respectively. The most part of observations, and, therefore, also the countries, are in the regimes 3 and 4.

Table 18: m1 with four regimes (TV=YRPC (90-07))

	Coeff.	T	SL
d_YRPC_1	0.177	3.156	0.002
d_YRPC_2	-0.172	1.728	0.084
d_YRPC_3	0.110	3.067	0.002
d_X	0.001	1.792	0.073
d_GFC	0.002	2.161	0.031
d_HC	0.712	7.278	0.000
d_G_1	0.169	2.223	0.026
d_D.r.1	-0.089	-1.748	0.081
d_D.r.2	-0.305	-1.403	0.161
d_D.r.3	0.031	1.041	0.298
d_D.r.4	-0.174	-7.536	0.000
d_D_1.r.1	-0.075	-1.938	0.053
d_D_1.r.2	0.164	1.247	0.212
d_D_1.r.3	-0.020	-0.822	0.411
d_D_1.r.4	0.006	0.309	0.757
SEE	0.06961		
Sum of coefficients:		LM-Test	
d_D.r.1	d_D_1.r.1	-0.164	***
d_D.r.2	d_D_1.r.2	-0.140	*
d_D.r.3	d_D_1.r.3	0.011	
d_D.r.4	d_D_1.r.4	-0.168	***

Considering only regimes 3 and 4 in Table 18 we conclude that in the regime with lower income the national debt has no effect on growth and, conversely, in regime 4 (with higher incomes) the debt negatively affects growth. This result also goes against well-established facts and good sense.

4.3 Third choice for the threshold variable: YEAR

We try now the hypothesis of regime changes caused by time evolution.

Table 19: Threshold LR test for m1 with TV=YEAR

YEAR	LR	SL
1986	6.6	
1986, 1987	17.0	
1986, 1987, 1995	29.3	**

The results obtained for the non-rejection of thresholds in terms of time (Table 19) are not as robust like the other hypotheses that we have so far analyzed. Therefore, we decide to try our last hypothesis: the stability of the evolution of debt measured by the sum of current and lag first differences.

4.4 Fourth choice for the threshold variable: d_D + d_D_1

By using this variable as a threshold, we intend to take into account the fact that the low values of yearly variation of debt ratio would mean~~s~~ for markets "debt sustainability" and, conversely, the high values of yearly variations would mean the "unsustainability of debt". For model m1 the null of the lag of debt for every regime is not rejected. For this reason, we estimate m2 model.

Table 20: Threshold LR test for m2 with TV=d_D+d_D_1

d_D + d_D_1	LR	SL
-17.3	35.0	***
-17.1, 10.2	15.9	***
-17.1, 10.2, 16.2	62.2	***

The LR tests for the rejection of thresholds allows us not to reject the presence of three values for thresholds and, thus, to continue with four regimes. Very low values for the lags of the growth variable (0.110) are also estimated (Table 21).

Table 21: m2 with four regimes (TV=d_D+d_D_1)

	Coeff.	T	SL
d_YRPC_1	0.179	3.117	0.002
d_YRPC_2	-0.176	1.809	0.071
d_YRPC_3	0.107	3.002	0.003
d_X	0.001	1.970	0.049
d_GFC	0.002	2.335	0.020
d_HC	0.710	7.733	0.000
d_G_1	0.171	2.264	0.024
d_D.r.1	-0.011	-0.409	0.682
d_D.r.2	-0.293	-8.023	0.000
d_D.r.3	0.096	1.038	0.299
d_D.r.4	-0.164	-3.728	0.000
SEE	0.06932		

The different regimes are represented very unevenly: 2.8%, 44.4%, 48.1% and 4.6%, for regimes from 1 to 4 respectively. Thus, we should ignore the 1st and 4th regime and concentrate our attention on the other two regimes, the 2nd and 3rd. In the 3rd regime, we cannot reject the absence of influence of debt on growth and in the 2nd we register a very high negative value (-0.293). This is against well-established facts. In fact, one should assume that debt should negatively affect economic growth in countries that register a high value of debt variation in relative terms. However, our results suggest that it does not happen this way: increase in debt variation leads to the “absence of relation between debt and economic growth”.

5. Conclusion

The main purpose of the study reported in this paper was: (i) to investigate the statistical properties of the d_t series (Debt-to-GDP ratio) over a long period of time, since the statistical properties of d_t (more precisely, its stationarity or non-stationarity and the possible presence of the long memory in d_t) have not been taken into consideration by the existing literature on the “Debt-to-GDP – growth” relationship. The absence of such preliminary analysis of data may, therefore, invalidate the results of many previous studies. (ii) to investigate whether the debt-growth relation varies with the level of indebtedness or not; and whether there is a common threshold for government debt ratios above which long-term growth rates may drop off significantly.

For this reason, in this paper we have analyzed the statistical properties of d_t using various statistical methods. The analysis of the ACF values confirms the absence of a general mean-reverting process for debt and the presence of long-run memory for this variable. This last result is also perfectly confirmed by the spectral analysis. The determination of the Hurst coefficient either by the R/S, but mainly by the Whittle estimators, confirms in an obvious way, the presence of long memory for debt. The GPH and SPR methods for the determination of the fractional integration parameter also confirm the presence of long memory process even for a few countries for which d_t is not rejected as stationary. The Lo-MacKinlay VR estimations clearly confirm the absence of short-run memory and the absence of a relation of short-memory in relation with low values of debt.

Concerning stationarity, ADF, S-P and Chang tests confirm the non-rejection of a unit root in the levels of debt and the rejection of its first difference. So d_t is obviously I(1). The hypothesis of a co-integration relationship between output and debt was rejected by Westerlund tests.

The first main conclusion of this study is that d_t series has a long memory and so should not be considered as a short-run phenomenon but rather in a long-term context. Debt is composed of two components: long and medium term, the most important, and short term debt. The decision to issue long and medium term is based on interest rates spreads and ability to pay in future. Once issued it will exist during the next years. Short term debt is usually rollover and replaced for long term debt if interest rates so justify. To these facts, we should add a political element. Optimistic behavior of democratic governments: why to pay what the others have spent to their benefit and lost popularity therefore? We attribute to these factors the long memory characteristic of debt.

Even ignoring our conclusion about stationarity, a study between growth and debt should include debt dynamics to incorporate its long run memory. The second main conclusion is that d_t is I(1) and hence, non-stationary econometric methods should be applied to avoid the emergence of spurious relationships. We have also confirmed that debt and output, two I(1) variables, are not co-integrated.

This finding suggests that the previous studies examining the relationship between economic growth g and national debt (d) suffer from the well-known problems of mis-specification and spurious associations. In the last section, we confirm the presence of a negative effect between growth and the first difference of debt and the influence of this variable evolves very slowly in view of the low value of the growth lag coefficients. We also confirm that the different thresholds that we cannot reject doesn't add relevant information to our understanding of the relation between growth and debt.

In fact, countries have their own past, their history of democratic institutions and debt defaults, current transparency and financial stability is also very important. They must also be distinguished accordingly to their natural resources and other forms of collateral. Consequently, the approach in terms of homogeneous coefficients is not the most adequate and we will continue our research using different econometric methods based on heterogeneous coefficients.

Therefore, in the second part of this paper we study the relationship between *g* and national debt *d* taking the above conclusions into consideration (i.e. presence of long memory and non-stationarity) and examine if the debt-growth relation varies in accordance with different regimes defined by appropriate threshold identification. We have confirmed that the presence of different regimes defined by robust threshold estimations goes against well-established facts and also good-sense. We also confirmed that the relation between growth and debt is clearly negative in a context of very slow adjustment between those two variables.

References:

Abbas S., Belhocine N., El Ganainy A., and Horton M., 2010. A Historical Public Debt Database. IMF Working Paper WP/10/245.

Alesina A. and Ardagna A. 2009. Large Changes in Fiscal Policy: Taxes versus Spending. NBER Working Paper No. 15438. National Bureau of Economic Research, Inc.

Ash M., Basu D., and Dube A. 2015. Public Debt and Growth: Assessing Key Findings on Causality and Thresholds. Mimeo. University of Massachusetts Amherst.

Baillie R. 1996. Long Memory Processes and Fractional Integration in Economics, Journal of Economics, 73, 5, pp. 5-59.

Baum A., Checherita-Westphal C. and Rother P. 2013. Debt and growth: New evidence for the euro area. Journal of International Money and Finance, 32 (C), pp. 809-821.

Beran J., Feng Y., Ghosh S. and Kulik R. 2013. Long-Memory Processes, Probabilistic Properties and Statistical Methods, Springer-Verlag, Berlin, 2013.

Berlan J. 1994. Statistics for Long-Memory Processes, Chapman & Hall, New York.

Berlan J. 2014. Empirical Economic and Financial Research - Theory, Methods and Practice. Springer, Heidelberg, 2014.

Bhattacharya R., Gupta V. and Waymire E. 1983. The Hurst effect under trends, Journal of Applied Probability, 20, pp. 649-62.

Breitung J. and Pesaran M. 2005. Unit Roors and Cointegration in Panels, Deutsche Bundesbank, Discussion Paper, S. 1, N. 42/2005.

Campbell J., Lo A., MacKinlay A. 1997. The Econometrics of Financial Markets, Princeton University Press, Princeton, 1997.

Cecchetti S. and Lam P-S. 1994. Variance-Ratio Tests: Small-Sample Properties with an Application to International Output Data, Journal of Business & Economic Statistics, 12 (2), pp. 177-86.

Cecchetti S., Mohanty M., and Zampolli F., 2011. The real effects of debt. BIS Working Paper 352, Bank for International Settlement (BIS).

Chang Y. Nonlinear IV Unit Root Tests in Panels with Cross-Sectional Dependency. Journal of Econometrics, 110, pp. 261-292.

Charles A. and Darné O. 2009. Variance Ratio Tests of Random Walk: an overview. Journal of Economic Surveys, 23 (3), pp. 503-27.

Checherita-Westphal C. and Rother P. 2012. The Impact of High Government Debt on Economic Growth and its Channels: An empirical investigation for the euro area. European Economic Review, 56 (7), pp. 1392-1405.

Chow V. and Denning K. 1993. A Simple Multiple Variance Ratio Test. Journal of Econometrics, 58, 385-401.

Cochrane J. 1988. How Big Is the Random Walk in GNP? The Journal of Political Economy, 96 (5), pp. 893-920.

Cochrane J. 2011. Discount Rates, NBER Working Papers 16972, National Bureau of Economic Research.

Dahlhaus R. 1989. Efficient Parameter Estimation for Self-Similar Processes. The Annals of Statistics, 17 (4), pp.1749-1766.

DeLong J. and Summers L. 2012. Fiscal Policy in a Depressed Economy. Brookings Papers on Economic Activity. 2012:2.

Diebold F. and Rudebusch G. 1989. Long Memory and Persistence in Aggregate Output, Journal of Monetary Economics, 24, pp. 189-209.

Donayre L. and Taivan A. 2015. Causality between Public Debt and Real Growth in the OECD: A Country-by-Country Analysis. Available at SSRN: http://ssrn.com/abstract= 2668447 or http://dx.doi.org/10.2139/ssrn.2668447.

Fox J. and Weisberg S. 2011. An R Companion to Applied Regression, 2nd ed., Sage, Thousand Oaks, 2011.

Fox R. and M. Taqqu. 1986. Large-sample properties of parameter estimates for strongly dependent stationary Gaussian time series, Annals of Statistics, 14 (1), pp. 527-32.

Fraley C., Leisch F., Maechler M., Reisen V., and Lemonte A.2015. Package 'fracdiff' for R, Version 1.4-2, February 19th, 2015.

Geweke J. and Porter-Hudak S. 1983. The Estimation and Application of Long Memory Time Series Models. Journal of Time Series Analysis 4(4), pp. 221-238.

Giraitis L. and Surgailis D. 1990. A Central Limit Theorem for Quadratic Forms in Strongly Dependent Linear Variables and its Applications to Asymptotic normality of Whittle's Estimate, Probability Theory Related Fields, 86, pp. 87-104.

Granger C. 1980. Long Memory Relationships and the Aggregation of Dynamic Models, Journal of Econometrics, 14, pp. 227-38.

Hamilton J.1994. Time Series Analysis, Princeton University Press, Princeton, 1994.

Hansen B. 1999. Threshold effects in non-dynamic panels: Estimation, testing and inference, Journal of Econometrics, 93, pp. 345-368.

Hansen B. 2015. Regression Kink with Unknown Threshold, Journal of Business and Economic Statistics (forthcoming).

Hauser M. 1997. Semiparametric and Nonparametric Testing for Long Memory: A Monte Carlo Study, Empirical Economics, 22 (2), pp. 247-71.

Herndon T, Ash M. and Pollin R. 2013. Does High Public Debt Consistently Stifle Economic Growth? A Critique of Reinhart and Rogoff. Working paper. Political Economy Research Institute, 2013.

Hurlin C. and Mignon V. 2007. Second Generation Panel Unit Root Tests, HAL, 2007.

Hurst H. 1951. Long Term Storage Capacity of Reservoirs. Transactions of the American Society of Civil Engineers, 116, pp. 770-99.

Kim J. 2006. Wild Bootstrapping Variance Ratio Test, Economic Letters, 92 (1), pp. 38-43.

Kim J. 2015. vrtest: Variance Ratio tests and other tests for Martingale Difference Hypothesis, Version, 0.97, February 20, 2015.

Kleiber C. and Zeileis A. 2008. Applied Econometrics with R, Springer, N. York

Kremes J., Ericsson N., and Dolado J. 1992. The Power of Cointegration Tests. Oxford Bulletin of Statistics, 54 (3).

Kumar M. and Woo J. 2010. Public Debt and Growth. IMF Working Paper WP/10/174.

Lillo F. and Farmer J. 2004. The Long Memory of the Efficient Market. Studies in Nonlinear Dynamics & Econometrics, 8 (3), pp. 1-33.

Lo A. 1991. Long term memory in stock market prices, Econometrica, 59, pp. 451-474.

Lo A. and MacKinlay C. 1988. Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, The Review of Financial Studies, 1, pp. 41-66.

Lo A. and MacKinlay C. 1989. The Size and Power of the Variance Ratio Test in Finite Samples: A Monte Carlo Investigation, Journal of Econometrics, 40, pp. 203-38.

Lopes S., Olbermann B. and Reisen V. 2002. Non-stationary Gaussian ARFIMA processes: Estimation and application, Brazilian Review of Econometrics, 22 (1), pp. 103-26.

Lopes S., Olbermann B. and Reisen V. 2004. A Comparison of Estimation Methods in Non-Stationary ARFIMA processes, Journl of Statistical Computation and Simulation, 74, pp. 339-47.

Lumley T and Zeileis A. 2015. Package ‘sandwich’, Version 2.3-4, R-Cran, September 24.

Maechler M. 2015. Package “longmemo” for R, CRAN, 2015.

Mandelbrot B. 1972. Statistical Methodology for Nonperiodic Cycles: From the Covariance to R/S Analysis, Annals of Economic and Social Measurement, 1 (3), pp. 259-90.

Mandelbrot B. 1975. Limit theorems on the self-normalized range for weakly and strongly dependent processes, Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete, 31 (4), pp. 271-285.

Mandelbrot B. and Taqqu M. 1979. Robust R/S analysis of long-run serial correlation. In Proceedings of the 42nd Session of the International Statistical Institute, Manila (1979) Bulletin of the International Statistical Institute, 48 (2), pp. 69-104.

Mandelbrot B. and Wallis J. 1968. Noah, Joseph, and Operational Hydrology. Water Resources Research, 4 (5), pp. 909-918.

Mandelbrot B. and Wallis J. 1969. Some Long-Run Properties of Geophysical Records. Water Resources Research, 5 (2), pp. 321-340.

McLeod A. and Hipel K. 1978. Preservation of the Rescaled Adjustment Range, 1: A Reassessment of the Hurts Phenomenon. Water Resources Research, 14 (3), pp. 491-508.

Nelson C. and Plosser C. 1982. Trends and random walks in macroeconmic time series: Some evidence and implications, Journal of Monetary Economics, 10 (2), pp. 139-162.

Panizza H. and Presbitero A. 2012. Public Debt and Economic Growth: Is There a Causal Effect? Journal of Macroeconomics, 41 (September), pp. 21-41.

Perotti R. 2012. The Effects of Tax Shocks on Output: Not So Large, but Not Small Either. American Economic Journal: Economic Policy 2012, 4(2), pp. 214-237.

Persyn D. and Westerlund J. and (2008) Error-correction–based Cointegration Tests for Panel Data. The Stata Journal, 8 (2), pp. 232-241.

Pescatori A., Sandri D., and Simon J. 2014. Debt and Growth: Is There a Magic Threshold? IMF Working Papers, WP/14/34.

Pfaff B. 2008. Analysis of Integrated and Cointegrated Time Series with R, Springer, Berlin, 2008.

Poterba J. and Summers L. 1988. Mean Reversion in Stock Prices: Evidence and Implications, Journal of Financial Economics, 22 (1), pp. 27-59.

R Core Team 2015. R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Viena, 2015.

Reinhart C. and Rogoff K. 2010. Growth in a Time of Debt. American Economic Review 100 (2), pp. 573-78.

Reinhart C., Reinhart V., and Rogoff K. 2012. Public Debt Overhangs: Advanced-Economy Episodes since 1800. Journal of Economic Perspectives 26 (3), pp. 69-86.

Reisen V. 1994. Estimation of the Fractional Difference Parameter in the ARFIMA(p,d,q) Model using the Smoothed periodogram. Journal Time Series Analysis, 15(1), 335-350.

Schmidt P. and Phillips P. 1992. Lm Tests of a Unit Root in the Presence of Deterministic Trends. Oxford Bulletin of Economics and Statistics, 54 (3), pp. 257-87.

Taqqu M., Teverovsky V., and Willinger W. 1995. Estimators for Long-range Dependence: an empirical study. Fractals, 3(4), pp. 785-788.

Tsay R. 2013. An Introduction to Analysis of Financial Data with R, Wiley, 2013.

Tsung-Wu. H. 2017. pdR: Panel Data Regression: Threshold Model and Unit Root Tests, R package version 1.5.

Weron R. 2002. Estimating Long Range Dependence: finite sample properties and confidence intervals, Physica A, 312, pp. 285-99.

Westerlund J. 2007. Testing for Error Correction in Panel Data. Oxford Bulletin of Economics and Statistics, 69 (6), pp. 709-748.

Whittle P. 1953. Estimation and information in stationary time series. Arkiv för Matematik, 2 (5), pp. 423-434.

Appendix A. Isocodes of countries

	Country	Isocode
1	Argentina	AR
2	Australia	AU
3	Austria	AT
4	Bahamas, The	BS
5	Barbados	BB
6	Belgium	BE
7	Benin	BJ
8	Bolivia	BO
9	Burundi	BI
10	Cameroon	CM
11	Canada	CA
12	Central African Rep.	CF
13	Chad	TD
14	Chile	CL
15	Colombia	CO
16	Congo, Dem. Rep. of	CD
17	Congo, Republic of	CG
18	Costa Rica	CR
19	Cyprus	CY
20	Denmark	DK
21	Dominican Republic	DO
22	Ecuador	EC
23	Egypt	EG
24	El Salvador	SV
25	Fiji	FJ
26	Finland	FI
27	France	FR
28	Gabon	GA
29	Germany	DE
30	Ghana	GH
31	Greece	GR
32	Grenada	GD
33	Guatemala	GT
34	Guyana	GY
35	Haiti	HT
36	Honduras	HN
37	India	IN
38	Ireland	IE
39	Israel	IL
40	Italy	IT
41	Jamaica	JM
42	Japan	JP
43	Jordan	JO
44	Kenya	KE
45	Korea, Republic of	KR
46	Malawi	MW

47	Malaysia	MY
48	Mali	ML
49	Malta	MT
50	Mauritius	MU
51	Mexico	MX
52	Morocco	MA
53	Nepal	NP
54	Netherlands	NL
55	New Zealand	NZ
56	Nicaragua	NI
57	Niger	NE
58	Nigeria	NG
59	Norway	NO
60	Pakistan	PK
61	Panama	PA
62	Paraguay	PY
63	Peru	PE
64	Phillipines	PH
65	Portugal	PT
66	Rwanda	RW
67	Samoa	WS
68	Senegal	SN
69	Sierra Leone	SL
70	Singapore	SG
71	South Africa	ZA
72	Spain	ES
73	Sri Lanka	LK
74	Swaziland	SZ
75	Sweden	SE
76	Switzerland	CH
77	Tanzania	TZ
78	Thailand	TH
79	Trinidad & Tobago	TT
80	Tunisia	TN
81	Turkey	TR
82	Uganda	UG
83	United Kingdom	UK
84	United States	US
85	Uruguay	UY
86	Venezuela, Rep. Bol.	VE
87	Zambia	ZM

Appendix B. Isocodes of countries in different samples

Table 1:

87, AR, AT, AU, BB, BE, BI, BJ, BO, BS, CA, CD, CF, CG, CH, CL, CM, CO, CR, CY, DE, DK, DO, EC, EG, ES, FI, FJ, FR, GA, UK, GD, GH, GR, GT, GY, HN, HT, IE, IL, IN, IT, JM, JO, JP, KE, KR, LK, MA, ML, MT, MU, MW, MX, MY, NE, NG, NI, NL, NO, NP, NZ, PA, PE, PH, PK, PT, PY, RW, SE, SG, SL, SN, SV, SZ, TD, TH, TN, TR, TT, TZ, UG, US, UY, VE, WS, ZA, ZM
74, AR, AT, AU, BB, BE, BI, BJ, BO, CA, CD, CF, CG, CL, ,CM, CO, CR, DE, DK, DO, EC, EG, ES, FI, FJ, FR, GA, UK, GH, GR, GT, HN, IE, IL, IN, IT, JP, KE, KR, LK, MA, ML, MT, MW, MX, MY, NE, NG, NI, NL, NO, NP, NZ, PE, PH, PK, PT, RW, SE, SG, SL, SN, SV, SZ, TD, TH, TN, TR, TT, UG, US, UY, VE, ZA, ZM
60, AR, AT, AU, BB, BE, BI, BO, CA, CD, CL, CO, CR, DE, DK, DO, EC, EG, ES, FI, FR, GA, UK, GH, GR, GT, HN, IE, IL, IN, IT, JP, KE, KR, LK, MA, ML, MT, MX, MY, NG, NL, NO, NZ, PE, PH, PK, PT, RW, SE, SG, SN, SV, SZ, TH, TN, TR, US, UY, VE, ZA

The number before the codes is the total number of countries in the sample.

CeBER Working Papers

(Available on-line at www.uc.pt/go/ceber)

2017-07	*The Effect of Public Debt on Growth in Multiple Regimes in the Presence of Long-Memory and Non-Stationary Debt Series - Irina Syssoyeva-Masson & João de Sousa Andrade*
2017-06	*The Blank and the Null: An examination of non-conventional voting choices – Rodrigo Martins*
2017-05	*Where is the information on USD/Bitcoins hourly price movements?- Helder Sebastião, António Portugal Duarte & Gabriel Guerreiro*
2017-04	*The response of non-price competitiveness and productivity due to changes in passed income gaps. Evidence from the OECD countries - Pedro André Cerqueira, Micaela Antunes & Elias Soukiazis*
2017-03	*Dutch Disease in Central and Eastern European Countries - João Sousa Andrade & António Portugal Duarte*
2017-02	*On the gains of using high frequency data and higher moments in Portfolio Selection- Rui Pedro Brito, Hélder Sebastião & Pedro Godinho*
2017-01	*Growth adjustments through non-price competitiveness and productivity. A cumulative causation approach- Elias Soukiazis, Micaela Antunes & Pedro André Cerqueira*

A série CeBER Working Papers foi iniciada em 2017.