

Afonso, António; Jalles, João Tovar

Working Paper

Revisiting fiscal sustainability: panel cointegration and structural breaks in OECD countries

ECB Working Paper, No. 1465

Provided in Cooperation with:

European Central Bank (ECB)

Suggested Citation: Afonso, António; Jalles, João Tovar (2012) : Revisiting fiscal sustainability: panel cointegration and structural breaks in OECD countries, ECB Working Paper, No. 1465, European Central Bank (ECB), Frankfurt a. M.

This Version is available at:

<https://hdl.handle.net/10419/153898>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.



EUROPEAN CENTRAL BANK

EUROSYSTEM



WORKING PAPER SERIES

NO 1465 / AUGUST 2012

REVISITING FISCAL SUSTAINABILITY

PANEL COINTEGRATION AND STRUCTURAL BREAKS IN OECD COUNTRIES

António Afonso and João Tovar Jalles



In 2012 all ECB publications feature a motif taken from the €50 banknote.



NOTE: This Working Paper should not be reported as representing the views of the European Central Bank (ECB). The views expressed are those of the authors and do not necessarily reflect those of the ECB.

Acknowledgements

The authors are grateful for comments from Artur Lopes and from an anonymous referee of the ECB WP series, from participants at ECB and at ISEG/UTL seminars, and to Roberta De Stefani for research assistance. The opinions expressed herein are those of the authors and do not necessarily reflect those of the ECB or the Eurosystem.

António Afonso

at European Central Bank, Directorate General Economics, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany, ISEG/UTL - Technical University of Lisbon, Department of Economics and UECE – Research Unit on Complexity and Economics;
e-mail: antonio.afonso@ecb.europa.eu, aafonso@iseg.utl.pt

João Tovar Jalles

at European Central Bank, Directorate General Economics, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany and University of Aberdeen; e-mail: joao.jalles@ecb.europa.eu, j.jalles@abdn.ac.uk

© European Central Bank, 2012

Address

Kaiserstrasse 29, 60311 Frankfurt am Main, Germany

Postal address

Postfach 16 03 19, 60066 Frankfurt am Main, Germany

Telephone

+49 69 1344 0

Internet

<http://www.ecb.europa.eu>

Fax

+49 69 1344 6000

All rights reserved.

ISSN 1725-2806 (online)

Any reproduction, publication and reprint in the form of a different publication, whether printed or produced electronically, in whole or in part, is permitted only with the explicit written authorisation of the ECB or the authors.

This paper can be downloaded without charge from <http://www.ecb.europa.eu> or from the Social Science Research Network electronic library at http://ssrn.com/abstract_id=2128484.

Information on all of the papers published in the ECB Working Paper Series can be found on the ECB's website, <http://www.ecb.europa.eu/pub/scientific/wps/date/html/index.en.html>

Abstract

We assess the sustainability of public finances in OECD countries, over the period 1970-2010, using unit root and cointegration analysis, both country and panel based, controlling for endogenous breaks. Results notably show: lack of cointegration – absence of sustainability – between government revenues and expenditures for most countries (except for Austria, Canada, France, Germany, Japan, Netherlands, Sweden, and UK); improvements of the primary balance after past worsening in debt ratios for Australia, Belgium, Germany, Ireland, Netherlands and the UK; Granger causality from government debt to the primary balance for 12 countries (suggesting the existence of Ricardian regimes). Overall, fiscal policy has been less sustainable for several countries, and panel data results corroborate the time-series findings.

JEL: C32, E62, H62, H63

Keywords: debt, primary balance, fiscal regimes, stationarity, breaks, causality, panel cointegration, FMOLS

Non-technical summary

The importance of sustainable public finances has received increasing attention particularly in the context and following the 2008-2009 economic and financial crisis. Sustainable fiscal policies can be continued indefinitely without any change in the policy stance, and when the intertemporal government budget constraint holds in present value terms. Conversely, if budgetary imbalances prevail, economic policies at both macro and microeconomic levels will quickly become unsupportable and changes would be required. If such a phenomenon occurs, then fiscal imbalances would imply a need for larger and more painful adjustments for the economy.

The main purpose of this paper is to investigate and draw some policy lessons on the sustainability of fiscal policy in a set of 18 OECD countries, using annual data over the period 1970-2010. Besides answering this policy question, we are also interested, among other things, in ascertaining the causal direction between government expenditures and revenues. The causal direction between the two budgetary variables may provide useful insights into how policy makers can manage budget deficits in the future. In our empirical approach we perform a systematic analysis of the stationarity properties of the first-differenced stock of government debt as well as, on the one hand, the relation between government revenues and expenditures and, on the other hand, the relation between primary balances and debt. These approaches provide us with an indirect test on the solvency of public finances in these countries. We conduct this analysis on a country-by-country basis, by means of several time series techniques, for robustness purposes, as well as for the country panel as a whole.

Our contributions are as follows: i) we combine both individual-country analysis by means of (recent) time-series techniques with panel data approaches for completeness and robustness purposes; ii) we take a longer time span and make use of uniform and comparable data for 18 OECD countries; iii) and we explore three different channels to evaluate fiscal sustainability as put forward in theoretical terms, that is, by looking at the first-differenced debt ratios, the relationship between government revenues and expenditures and, finally, the relationship between (lagged) public debt and primary balances.

Our results show that the first-differenced debt series for most countries is non-stationarity suggesting that the solvency condition would not be satisfied. Moreover, evidence suggests the existence of one cointegrating relationship in only 6 countries between revenues and expenditures. However, the overall test results allow the rejection of the cointegration hypothesis in both relationships under scrutiny. In other words, government expenditures, in half of the countries, exhibited a higher growth rate than government revenues, challenging therefore the hypothesis of fiscal sustainability.

On the other hand, the cointegrating coefficients for the revenues-expenditures relationship are positive (but less than one) and statistically significant, meaning that for each percentage point of GDP increase in public expenditures, revenues increase by less than one percentage point of GDP. In terms of causality, our evidence suggests stronger effects running from revenues to expenditures and most countries are not able to generate the revenues required to finance the planned expenditures. We find Granger-causality from government debt to the primary balance, which can be seen as evidence of the existence of a Ricardian regime.

Finally, panel data results corroborate time-series findings, and even though we find that long-run causality seems to run from lagged debt to the primary balance, on average the marginal long-run impact is zero. All in all, we cannot say that fiscal policy has been sustainable for most countries in our sample.

1. Introduction

The importance of sustainable public finances has received increasing attention particularly in the context and following the 2008-2009 economic and financial crisis. From a fiscal perspective, maintaining a stable long-term relationship between expenditures and revenues is one of the key requirements for a stable macroeconomic environment and a sustainable economy. Therefore, our purpose is to find out whether fiscal imbalances in a number of OECD countries need to be curtailed before they become economically unsustainable, leading to insolvency situations.

Sustainable fiscal policies can be continued indefinitely without any change in the policy stance, and when the intertemporal government budget constraint holds in present value terms.¹ Conversely, if budgetary imbalances prevail, economic policies at both macro and microeconomic levels will quickly become unworkable and changes would be required. If such a phenomenon occurs, then fiscal imbalances would imply a need for larger and more painful adjustments for the economy. Given the detrimental impact of persistent deficits, practices on debt sustainability and appropriate fiscal policies are extremely important. For instance, Blanchard et al. (1990) present as a definition of a sustainable fiscal policy one that allows, in the short-term, that the debt-to-GDP ratio returns to its original level after some excessive variation.

There has been a substantial volume of empirical studies focusing on fiscal sustainability, tackling explicit government liabilities, looking notably at the US and European cases (see, Hamilton and Flavin, 1986; Hakkio and Rush, 1991; Trehan and Walsh, 1991; MacDonald, 1992; Ahmed and Rogers, 1995; Quintos, 1995; Makrydakis et al., 1999; Feve and Henin, 2000; Martin, 2000; Bravo and Silvestre, 2002; Hatemi-J, 2002; Afonso, 2005, Mendoza and Ostry, 2007; Arghyrou and Luintel, 2007; Afonso, 2008; Afonso and Rault, 2010; Afonso and Jalles, 2011a, to name a few). In particular, Trehan and Walsh (1991) and Afonso (2008) are of interest in what follows since they emphasize the relationship between primary balances and government debt. In addition, Fincke and Greiner (2011), and Legrenzi and Milas (2011) focus on the possibility of finding structural breaks or possible non-linearities notably vis-à-vis debt thresholds, also in a fiscal reaction approach.²

¹ Analysis on fiscal sustainability has focused on both the univariate properties of government debt (e.g. Hamilton and Flavin, 1986) and on the long-run relationship between government revenues and expenditures (e.g. Hakkio and Rush, 1991).

² Regarding the specification of fiscal reaction functions, not the main purpose of the current paper, Afonso and Jalles (2011b) have tackled the issue also in a multi-country framework.

On the other hand, Bohn (2007) provides a substantial challenge to the time series literature on fiscal policy. Specifically, Bohn suggested that rejections of stationarity-based sustainability tests are invalid because in an infinite sample, any order of integration of debt is consistent with the transversality condition which implies that the intertemporal budget constraint may be satisfied even if these particular time series tests are not. Instead, Bohn (2007) emphasizes whether a country's primary balance responds positively to debt as an indicator of sustainability. As in Bohn (1998), this depends upon the assumption that the series are stationary, or when they are nonstationary, for them to be related in a statistical sense, they must be of the same order of integration and the primary balance and government debt must be cointegrated.

Additionally, another strand of the literature, as for example, Davig et al.(2011), estimate fiscal reaction functions, for the assessment of whether governments behave in a Ricardian. In the same vein, other papers study theoretical thresholds or limits to assess when fiscal policies need to be changed (see, Davig et al., 2011, and Richter, 2011). As in other related studies, this is usually done for the US. However, this is not the main objective of our study, since we derive our cointegration relationships directly from the intertemporal government budget constraint. Still, some of the results those studies can indirectly be read as providing also information on the sustainability of public finances.

The main purpose of this paper is to investigate and draw some policy lessons on the sustainability of fiscal policy in a set of 18 OECD countries. Besides answering this policy question, we are also interested, among other things, in ascertaining the causal direction between government expenditures and revenues. The causal direction between the two budgetary variables may provide useful insights into how policy makers can manage budget deficits in the future. In our empirical approach we perform a systematic analysis of the stationarity properties of the first-differenced stock of government debt as well as, on the one hand, the relation between government revenues and expenditures and, on the other hand, the relation between primary balances and debt, in line with Bohn (1998). These approaches provide us with an indirect test on the solvency of public finances in these countries. We conduct this analysis on a country-by-country basis, by means of several time series techniques, for robustness purposes, as well as for the panel as a whole, using annual data over the period 1970-2010.

Thus, our contributions are as follows: i) we combine both individual-country analysis by means of (recent) time-series techniques with panel data approaches for completeness and robustness purposes; ii) we take a longer time span and make use of uniform and comparable data for 18 OECD countries; iii) we identified structural breaks in the series and in the cointegration relations; iv) and we explore three different channels to evaluate fiscal

sustainability as put forward in theoretical terms, that is, by looking at the first-differenced debt ratios, the relationship between government revenues and expenditures and, finally, the relationship between (lagged) public debt and primary balances.

Briefly, our results show that the first-differenced debt series for most countries is non-stationarity suggesting that the solvency condition would not be satisfied. Moreover, evidence suggests the existence of one cointegrating relationship in only 6 countries between revenues and expenditures. However, the overall test results allow the rejection of the cointegration hypothesis in both relationships under scrutiny. In other words, for the period 1970-2010, government expenditures, in half of the countries, exhibited a higher growth rate than government revenues, challenging therefore the hypothesis of fiscal sustainability.

On the other hand, the cointegrating coefficients for the revenues-expenditures relationship are positive (but less than one) and statistically significant, meaning that for each percentage point of GDP increase in public expenditures, revenues increase by less than one percentage point of GDP. In terms of causality, our evidence suggests stronger effects running from revenues to expenditures and most countries are not able to generate the revenues required to finance the planned expenditures. We find Granger-causality from government debt to the primary balance, which can be seen as evidence of the existence of a Ricardian regime. Finally, panel data analysis corroborates time-series findings and even though we find that long-run causality seems to run from lagged debt to the primary balance, on average the marginal long-run impact is zero. All in all, we cannot say that fiscal policy has been sustainable for most countries in our sample.

The structure of the paper is as follows. Section 2 discusses the underlying theoretical framework which serves as the basis for the empirical strategy that follows. Section 3 presents the time-series and panel data econometric methodology. Section 4 presents and discusses our main results and findings. The last section concludes.

2. Theoretical Framework

Regarding the sustainability of fiscal policy the empirical literature usually tests for the possibility of both public expenditures and government revenues continuing their historical growth patterns. In principle, any value for the budget deficit would be possible if the government could raise its liabilities without limit, which is naturally impossible since the government is faced with the present value of its own budget constraint.

The government budget constraint can be used to derive the present value of the budget constraint. The flow budget constraint is written as:

$$G_t + (1 + r_t)B_{t-1} = R_t + B_t \quad (1)$$

where: G – government expenditures, excluding interest payments; R – government revenues; B – public debt; r – real interest rate.

Rewriting equation (1) for the subsequent periods, and recursively solving leads to the intertemporal budget constraint:

$$B_t = \sum_{s=1}^{\infty} \frac{R_{t+s} - G_{t+s}}{\prod_{j=1}^s (1 + r_{t+j})} + \lim_{s \rightarrow \infty} \prod_{j=1}^s \frac{B_{t+s}}{(1 + r_{t+j})}. \quad (2)$$

When the second term in the right-hand side of (2) is zero, the present value of the existing stock of public debt will be identical to the present value of future primary surpluses. However, it is more useful to make several algebraic modifications to equation (1) to obtain an appropriate specification for the empirical tests. Assuming that the real interest rate is stationary, with mean r , and defining:

$$E_t = G_t + (r_t - r)B_{t-1}. \quad (3)$$

we obtain the following so-called Present Value Borrowing Constraint (PVBC):

$$B_{t-1} = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s+1}} (R_{t+s} - E_{t+s}) + \lim_{s \rightarrow \infty} \frac{B_{t+s}}{(1+r)^{s+1}}. \quad (4)$$

A sustainable fiscal policy should ensure that the present value of the stock of public debt, the second term of the right-hand side of (4), goes to zero in infinity, constraining the debt to grow no faster than the real interest rate, imposing the absence of Ponzi games. Therefore, the government needs to achieve future primary surpluses whose present value adds up to the current value of the stock of public debt.

It is also possible to derive the solvency condition, with all the variables defined as a percentage of GDP. Thus, the PVBC, with the variables expressed as ratios of GDP, with y being the GDP real growth rate, is written as:

$$\frac{B_t}{Y_t} = \frac{(1 + r_t)}{(1 + y_t)} \frac{B_{t-1}}{Y_{t-1}} + \frac{G_t}{Y_t} - \frac{R_t}{Y_t}. \quad (5)$$

Assuming the real interest rate to be stationary, with mean r , and considering also constant real growth, the budget constraint is then given by:

$$b_{t-1} = \sum_{s=0}^{\infty} \left(\frac{1+y}{1+r} \right)^{s+1} [\rho_{t+s} - e_{t+s}] + \lim_{s \rightarrow \infty} b_{t+s} \left(\frac{1+y}{1+r} \right)^{s+1}. \quad (6)$$

with $b_t = B_t / Y_t$, $e_t = E_t / Y_t$ and $\rho_t = R_t / Y_t$. When $r > y$ it is necessary to introduce a solvency

condition, given by $\lim_{s \rightarrow \infty} b_{t+s} \left(\frac{1+y}{1+r} \right)^{s+1} = 0$ in order to bound public debt growth. This yields the

familiar result that fiscal policy will be sustainable if the present value of the future stream of primary surpluses, as a percentage of GDP, matches the “inherited” stock of government debt.

Recalling the PVBC, equation (4), it is possible to present analytically two complementary definitions of sustainability that set the background for empirical testing:³

i) The value of public current debt must be equal to the sum of future primary surpluses:

$$B_{t-1} = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s+1}} (R_{t+s} - E_{t+s}). \quad (7)$$

ii) The present value of public debt must approach zero in infinity:

$$\lim_{s \rightarrow \infty} \frac{B_{t+s}}{(1+r)^{s+1}} = 0 \quad (8)$$

In order to test empirically the absence of Ponzi games, one can test the stationarity of the first difference of the stock of public debt (ΔB_t) and the cointegration between primary balance (s_t) and the (lagged) stock of the public debt, in line with Bohn (2007), using the following cointegration regression: $s_t = \alpha + \beta B_{t-1} + u_t$. This, so called, “backward-looking” approach implies that an increase in the previous level of debt would result in a larger primary balance today.

Such relationship has been explored notably by Bajo-Rubio et al. (2009) through the lenses of the Fiscal Theory of the Price Level (see, e.g., Leeper 1991, Kocherlakota and Phelan, 1999, survey and McCallum and Nelson, 2005, critical appraisal) and the distinction between what is referred in the literature as a Ricardian or Monetary-dominant regime (hereafter MD) (“active” monetary policy, being the determination of prices its nominal anchor; “passive” fiscal policy with the budget balance path being endogenous) and a non-Ricardian or Fiscal-dominant regime (hereafter FD) (which allows fiscal policy to set primary balances – “active” – and to follow an arbitrary process, not necessarily compatible with solvency). These concepts will be further addressed when discussing our empirical results in Section 4.

It is also possible to assess fiscal policy sustainability through cointegration between government revenues and expenditures. The implicit hypothesis concerning the real interest rate,

³ Hamilton and Flavin (1986) first used these procedures. See also Trehan and Walsh (1991) and Hakkio and Rush (1991).

with mean r , is also stationarity. Using $E_t = G_t + (r_t - r)B_{t-1}$, and $GG_t = G_t + r_t B_{t-1}$, the intertemporal budget constraint becomes

$$GG_t - R_t = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s-1}} (\Delta R_{t+s} - \Delta E_{t+s}) + \lim_{s \rightarrow \infty} \frac{B_{t+s}}{(1+r)^{s+1}} \quad (9)$$

and with the no-Ponzi game condition, GG_t and R_t must be co-integrated variables of order one for their first differences to be stationary.

Therefore the procedure to assess the sustainability of the intertemporal government budget constraint involves testing the following cointegration regression: $R_t = \alpha + \beta GG_t + u_t$. If the null of no co-integration is rejected, the residual u_t must be stationary, and should not display a unit root. Hakkio and Rush (1991) demonstrate that if GG and R are non-stationary variables in levels, the condition $0 < b < 1$ is a sufficient condition for the budget constraint to be obeyed. However, when revenues and expenditures are expressed as a percentage of GDP or in per capita terms, it is necessary to have $b = 1$ in order for the trajectory of the debt to GDP not to diverge in an infinite horizon.

3. Econometric Methodology

3.1. Time series

3.1.1. Unit Roots and Structural Breaks

Stationarity-wise, unit root tests can provide a valuable insight into the presence of either a deterministic or stochastic secular component in the series. In this context, in addition to standard Augmented Dickey Fuller (ADF) and Phillips-Perron (PP) unit root tests – for purposes of robustness and completeness⁴ – we also conduct the four tests (M-tests) proposed by Ng and Perron (2001) (NP) based on modified information criteria (MIC): the modified Phillips-Perron test MZ_{α} ; the modified Sargan-Bhargava test (MSB); the modified point optimal test MP_T ; and the modified Phillips-Perron MZ_T . These improve the PP-tests both with regard to size distortions and power.

We then resort to unit root tests allowing for breaks and we begin with the Zivot-Andrews (1992) (ZA) one. This endogenous structural break test is a sequential test which utilizes the full sample and uses a different dummy variable for each possible break date. The break date is selected where the t-statistic from the ADF test of unit root is at a minimum (most negative). Consequently a break date will be chosen where the evidence is least favourable for the unit root

⁴ Moreover, these tests are especially appropriate under certain dynamic data structure, and when their random components are not white noise.

null.⁵ We complement with the modified ADF test proposed by Vogelsang and Perron (1998) (VP) also allowing for one endogenously determined break. Finally, we take the two-break unit root test described by Clemente, Montanes and Reyes (1998) (CMR).⁶ These test the null of unit root against the break-stationary alternative hypothesis and provide us supplementary insights vis-à-vis conventional unit root tests that do not account for any break in the data.

For the unit root tests that allow for one or two endogenously determined breaks it is assumed that the shift can be modelled by a dummy variable $DU_t = 0$ for $t \leq TB$ and for $t > TB$, where TB is the shift date (time break). In the time series literature, two generating mechanisms of shifts are distinguished, additive outlier (AO) and innovational outlier (IO) models. The former results in an abrupt shift in the level, whereas the latter allows for a smooth shift from the initial level to a new level. Although both results are reported, we will mainly discuss tests constructed for AO models. As discussed in Vogelsang and Perron (1998), who consider an unknown shift date situation, the AO framework may be preferable to the IO statistics, even if the Data Generating Process (DGP) is an IO process.

However, it is important to recognize some important drawbacks in both previous unit root tests, particularly, the ZA and VP tests. In particular, with relation to the VP test, it has been shown that the critical values are substantially smaller in the $I(0)$ case than in the $I(1)$ case (therefore, suggesting that the test is conservative in the $I(0)$ case). The solution was then to devise a procedure that would have the same limit distribution in both cases. This was first attempted by Vogelsang (2001) but simulations provided support for the lack of power in the $I(1)$ case. Perron and Yabu (2009) (PY) were more successful on this endeavour by proposing a new test for structural changes in the trend function of the time series without any prior knowledge of whether the noise component was stationary or integrated and making use of Andrews and Ploberger's (1994) exponential functional and Roy and Fuller's (2001) finite sample correction procedure. This newer test has better properties in terms of size and power.⁷

3.1.2 Cointegration, Stability and Causality

Consider the following two cointegrating-relationship regressions, as identified in Section 2:

$$s_t = \alpha + \beta B_{t-1} + u_t. \quad (10)$$

⁵ The critical values in Zivot and Andrews (1992) are different to the critical values in Perron (1989). The difference is due to the fact that the selection of the time of the break is treated as the outcome of an estimation procedure, rather than predetermined exogenously.

⁶ For more detailed discussion of these tests that allow for endogenously determined breaks, the reader should refer to the original references.

⁷ We thank Pierre Perron and Tomoyoshi Yabu for providing their GAUSS code.

$$R_t = \alpha + \beta GG_t + u_t. \quad (11)$$

where B_t is the government debt and s_t is the government primary balance; R_t are government revenues and GG_t government expenditures. u_t is an iid disturbance term satisfying standard assumptions.

Given the nonstationarity of each individual time series (to be tested and confirmed in Section 5), the relevant question becomes whether a linear combination of these two pairs of variables is stationary. If such a combination exists, government revenues and expenditures (government debt and primary balance) become cointegrated,⁸ which implies that the variables are attracted to a stable long-run (equilibrium) relation and any deviation from this relation reflected short-run (temporary) disequilibria. A remark is worth making with respect to equation (10): if a positive and significant coefficient is to be found that would be a sufficient condition for solvency, indicating that the government satisfies its present-value budget constraint. A problem with such finding is that it is compatible with both the MD and FD regimes.⁹ Hence, we will combine cointegration with Granger-causality analysis.

We test for cointegrating (long-run) relations between government revenues and expenditures (primary balance and government debt) using the Johansen and Juselius (1990) methodology. This approach estimates the long-run attracting set in a VAR context that incorporates both the short- and long-run dynamics of the various models.

However, and as in the case of unit roots, a test for co-integration that does not take into account possible breaks in the long-run relationship will have lower power. The test will tend to under-reject the null of no co-integration if there is a co-integration relationship that has changed at some time during the sample period. Therefore, to further evaluate the previous results, one should also entertain the possibility that the series are co-integrated but that the linear combination has shifted at an unknown point in the data sample, in other words, that there might be a relevant break date. Following Gregory and Hansen (1996), the hypothesis of a structural shift in the co-integration relationships is then studied.¹⁰

In order to estimate the parameter β in both equations (10) and (11) we resort to the method of Dynamic Ordinary Least Squares (DOLS) of Stock and Watson (1993), following the methodology proposed by Shin (1994). This method has the advantage of providing a robust

⁸ To test for cointegration all variables must be integrated of the same order.

⁹ In a MD regime we would observe that an increase in previous' period debt would lead to a larger primary balance ex-post. Equivalently, in a FD regime, a decrease in the expected primary balance would lead to a decrease in the current debt ratio, through a price increase.

¹⁰ The authors have extended the Engle-Granger model to allow for a single break in the co-integration relationship. We thank Bruce Hansen for making the GAUSS routine available.

correction to the possible presence of endogeneity in the explanatory variable, as well as of serial correlation in the error terms of the OLS estimation.

As emphasized by Bruggemann et al. (2003), it is of some importance to formally investigate the stability of the cointegrating vectors further, once a long-run relationship has been identified. The temporal stability of estimated relations is also indicative of the usefulness of these relations for policy (forecasting) purposes. Hansen and Johansen (1993) outline a procedure that formally tests the constancy of cointegrating vectors in the context of Full Information Maximum Likelihood (FIML) estimations. Holding the short-run dynamics of the model constant, the procedure then treats these estimates as the null hypothesis in consecutive recursive tests. In this way, any rejection of the null of cointegration stability (constancy) should emanate from a breakdown in the long-run relation, rather than from any positive shift in the underlying short-run dynamics (Hoffmann et al., 1995). We apply this approach to test the stability of the cointegrating relation.

Moreover, by taking a VAR approach we can use one further important tool: Granger-causality tests. Many tests of Granger-type causality have been derived and implemented to test the direction of causality – Granger (1969), Sims (1972) and Gweke et al. (1983). These tests are based on null hypotheses formulated as zero restrictions on the coefficients of the lags of a subset of the variables. Thus, the tests are grounded in asymptotic theory. Other shortcomings of these tests have been discussed in Toda and Phillips (1994). Also, it is well documented that the exclusion of relevant variables induces spurious significance and inefficient estimates. In dealing with these problems and for robustness purposes, we employ Toda and Yamamoto (1995) and Dolado and Lutkepohl (1996) approach for Granger causality. They suggest a technique that is applicable irrespective of the integration and cointegration properties of the system. The method involves using a Modified Wald statistic for testing the significance of the parameters of a VAR(s) model (where s is the lag length in the system).¹¹

We follow Rambaldi and Doran (1996) in formulating these tests. Define d_{\max} as the maximum order of integration in the system, a VAR($k + d_{\max}$) has to be estimated to use the Modified Wald test for linear restrictions on the parameters of a VAR(k) which has an asymptotic χ^2 distribution.¹² In our case, we will run a 2 variables' VAR (for i) government

¹¹ As demonstrated by Toda and Yamamoto (1995), if variables are integrated of order d , the usual selection procedure is valid whenever $k \geq d$. Thus, if $d = 1$, the lag selection is always consistent.

¹² The traditional F tests and its Wald test counterpart to determine whether some parameter of a stable VAR model are jointly zero are not valid for non-stationary processes, as the test statistics do not have a standard distribution (Toda and Phillips, 1993).

revenues and expenditures; and ii) primary balance and (lagged) debt), with $k=2$ (AIC-based) and $d_{\max}=1$ but for the sake of notation simplicity, we denote them as $y_i, i=1,2$. For our VAR(3) we estimate the following system of equations:

$$\begin{bmatrix} y_{1t} \\ y_{2t} \end{bmatrix} = A_0 + A_1 \begin{bmatrix} y_{1t-1} \\ y_{2t-1} \end{bmatrix} + A_2 \begin{bmatrix} y_{1t-2} \\ y_{2t-2} \end{bmatrix} + A_3 \begin{bmatrix} y_{1t-3} \\ y_{2t-3} \end{bmatrix} + \begin{bmatrix} e_{y_1} \\ e_{y_2} \end{bmatrix}. \quad (12)$$

The above system of equations is estimated via the seemingly unrelated regression (SUR) method. The test consists of taking the first k VAR coefficient matrix (but not all lagged coefficients) to make Granger causal inference. If, e.g., we want to test that y_{2t} does not Granger-cause y_{1t} , the null hypothesis will be $H_0 : a^{(1)}_{12} = a^{(2)}_{12} = 0$, where $a^{(i)}_{12}$ are the coefficients of $y_{2t-i}, i=1,2$.

There are four main hypotheses with regard to the causal nexus of government revenues and expenditures:

1. One way causation from expenditures to revenues. This suggests that the government adjusts revenues to the level of the planned expenditures (see Barro, 1979).
2. One way causation from revenues to expenditures. Following this hypothesis, the authorities adjust their expenditures to the level of the revenue so that control over revenues leads to limited growth in the public sector (Friedman, 1978).
3. Bidirectional causality (fiscal synchronization). This hypothesis is based on the equivalence of marginal cost and marginal revenue that the utility-maximizing suppliers and demanders of the public services make. That is, the fiscal authorities made simultaneous decisions on expenditures and revenues. Hence, the two variables mutually reinforce each other. This is the classical view of public finance (Musgrave, 1966).
4. No causality. The authorities can set the level of expenditures and revenues by rule of thumb. This phenomenon reflects the institutional separation of allocation and taxation functions of the government (Hoover and Sheffrin, 1992). This view is also consistent with no cointegration and a potential sustainability problem.

It is worth noting that causality per se has implications only for the dynamics of the fiscal adjustments process and not for the sustainability condition. The latter depends on the existence of the condition outlined earlier. Since causality tests do give estimates of the authorities' reactions to the past fiscal imbalances, it would provide a useful indicator of how the authorities may respond to the imbalances in the future.

3.2 Panel Data

3.2.1 Panel Unit Roots

Given the notoriously low power of individual country-by-country tests for unit roots and cointegration, it may be preferable to pool the time series of interest and conduct panel analysis. We implement three different types of panel unit root tests: two first generation tests, namely the Im et al. (2003) test (IPS); the Maddala and Wu (1999) test (MW) and one second generation test – the Pesaran (2007) CIPS test. The latter is associated with the fact that previous tests do not account for cross-sectional dependence of the contemporaneous error terms and failure to consider it may cause substantial size distortions in panel unit root tests (O’Connell, 1998 and Pesaran, 2007).

3.2.2 Panel Cointegration and Panel Causality

We implement the panel cointegration tests proposed by Pedroni (2004). This is a residual-based test for the null of no cointegration in heterogeneous panels. Two classes of statistics are considered in the context of the Pedroni test. The first type is based on pooling the residuals of the regression along the within-dimension of the panel, whereas the second type is based on pooling the residuals of the regression along the between-dimension of the panel. For the first type, the test statistics are the panel ν -statistic, the panel ρ -statistic, the panel PP-statistic, and the panel ADF-statistic. These statistics are constructed by taking the ratio of the sum of the numerators and the sum of the denominators of the analogous conventional time-series statistics across the individual members of the panel. The tests for the second type include the group ρ -statistic, the group PP-statistic, and the group ADF-statistic. They are simply the group mean statistics of the conventional individual time series statistics. All statistics have been standardised by the means and variances so that they are asymptotically distributed $N(0,1)$ under the null of no cointegration. As one-sided tests, large positive values of the panel ρ -statistic reject the null hypothesis of no cointegration. For the remaining statistics, large negative values reject the null.¹³ See Pedroni (2004) for a detailed discussion.

Assuming that government revenues and expenditures (government debt and primary balance) are cointegrated (to be confirmed in Section 5), one thus need to estimate the cointegrating coefficients to investigate the long-run relationship between them. We apply the between-dimension panel fully modified OLS (FMOLS) (Pedroni, 2000, 2001). The rationale

¹³ Pedroni (1999) shows that the panel ADF and group ADF tests have better small-sample properties than the other tests, hence they are more reliable (Lee et al., 2008).

for using FMOLS is that in the presence of unit root variables, the effect of superconsistency may not dominate the endogeneity effect of the regressors if OLS is employed (Lee et al., 2008). Pedroni (2000) showed that the FMOLS approach can be used to draw an inference about cointegration with heterogeneous dynamics. FMOLS takes care of endogeneity problem and provides unbiased estimates of the coefficients, which can be interpreted as long-run elasticities. Individual estimates and standard errors for $H_0 : b_i = 0$ in (10) and (11) are reported as well as the panel estimation results.¹⁴

Therefore, if in each country the series R_t and GG_t are individually non-stationary but together, generally speaking, are cointegrated, we know from the Granger representation theorem that these series can be represented in the form of a dynamic error correction model (ECM). In line with Canning and Pedroni (2008) we estimate the following ECM:

$$\begin{aligned} \Delta R_{it} &= c_{it} + \lambda_{1i} \hat{e}_{it-1} + \sum_{j=1}^K \varphi_{11ij} \Delta R_{it-j} + \sum_{j=1}^K \varphi_{12ij} \Delta GG_{it-j} + \varepsilon_{1it} \\ \Delta GG_{it} &= c_{it} + \lambda_{2i} \hat{e}_{it-1} + \sum_{j=1}^K \varphi_{21ij} \Delta GG_{it-j} + \sum_{j=1}^K \varphi_{22ij} \Delta R_{it-j} + \varepsilon_{2it} \end{aligned}, \quad (13)$$

where $\hat{e}_{it} = R_{it} - \hat{\alpha}_i - \hat{b}_i - \hat{\beta}_i GG_{it}$ is the disequilibrium term and it represents how far our variables are from the equilibrium relationship, and the error correction mechanism estimates how this disequilibrium causes the variables to adjust towards equilibrium in order to keep the long-run relationship intact. The Granger representation theorem implies that at least one of the adjustment coefficients λ_{1i} or λ_{2i} must be non-zero if a long-run relationship between the variables is to hold. According to Canning and Pedroni (2008) one can test hypotheses about long-run effects by testing restrictions on the estimated coefficients in the dynamic ECM.

Hence, a test for the significance of λ_{1i} (λ_{2i}) for any one country can be interpreted as a test of whether shocks or innovations in government expenditures (revenues) have a long-run effect on government revenues (expenditures) and a test for the sign of the ratio $-\lambda_{1i} / \lambda_{2i}$ can be interpreted as a test of the sign of the long-run effect of shocks or innovations to government expenditures on revenues. In the next Section we present and discuss our main findings.

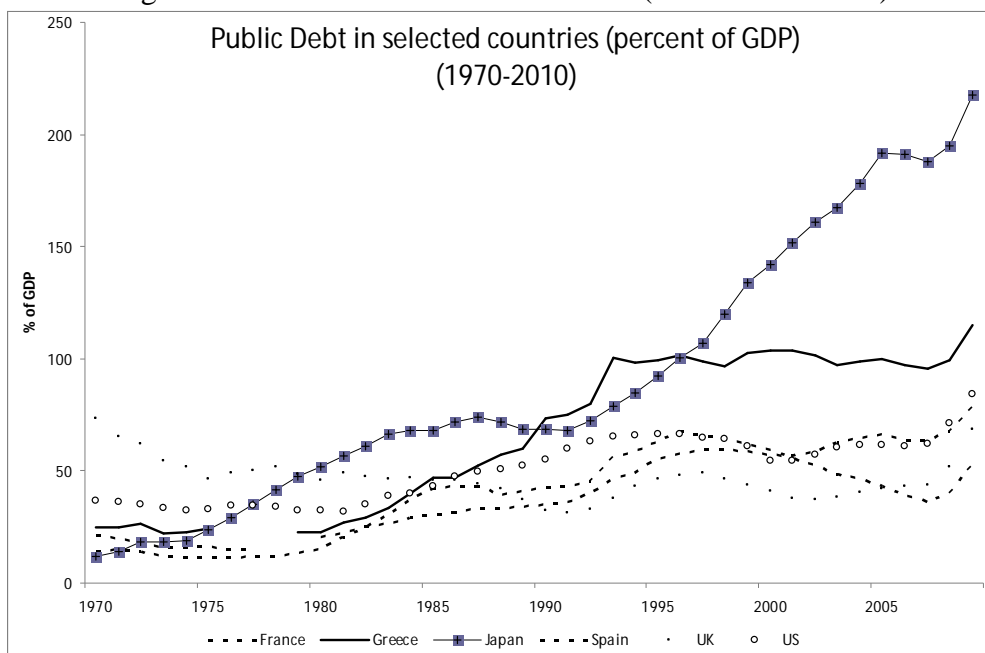
¹⁴ We thank Peter Pedroni for providing his RATS code.

4. Empirical Analysis

4.1 Stylised Facts and Data Overview

Most data are taken from the European Commission AMECO (Annual MacroEconomic Data) database, covering the period 1970-2010 for the 18 OECD countries considered in our sample. For Australia, Canada, Japan and the USA, primary balance (% of GDP) data comes from the OECD database. Government Debt (% of GDP) series are retrieved from the recently compiled Abbas et al. (2010) dataset. Appendix A provides additional information on variables, definitions and their sources.

Figure 1: Public Debt Series: 1970-2010 (selected countries)



Source: Abbas et al. (2010).

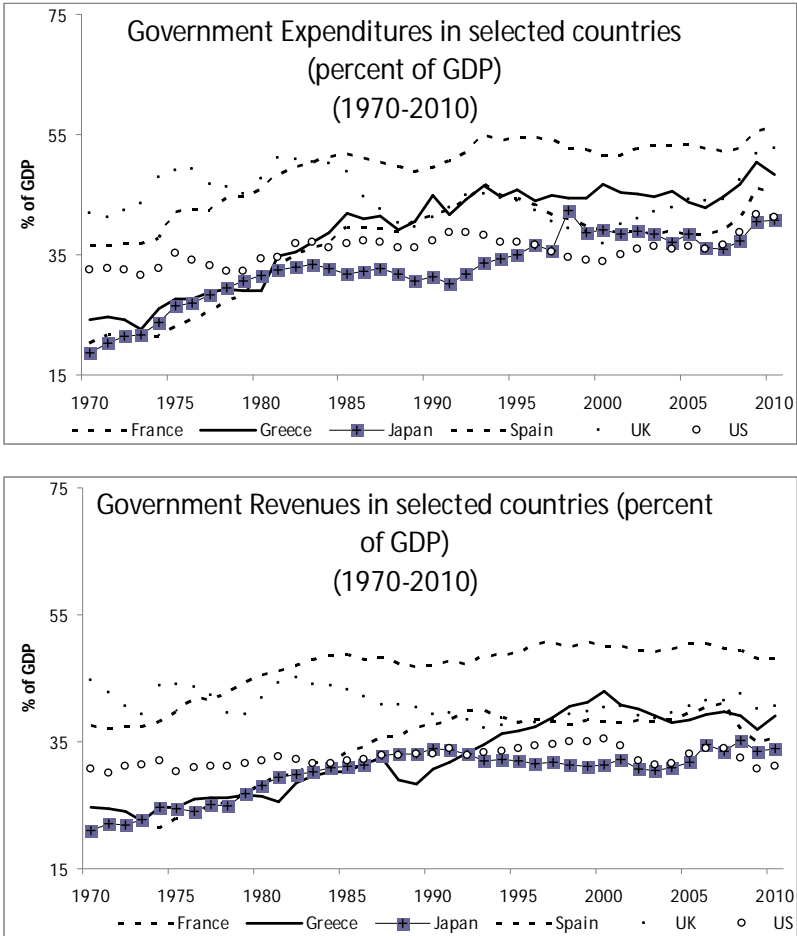
A brief characterization of the data seems appropriate to begin with. The consequences of choosing different fiscal policies may be exemplified by looking at the public debt paths in these countries, as depicted in Figure 1. It is clear from this chart that government debt-to-GDP ratios peaked after the 1970s till the end of the 1990s – with the exception of Japan where debt kept on rising. Government debt restarted an increasing trend with the 2008 economic crisis and the continuous worsening state of public finances in most advanced economies.

For instance, government debt increased in Italy from an average of 51.8% of GDP in the 1970s to an average of 112.3% in the 2000s. In the case of Greece, Italy and Japan government debt has surpassed 100% of GDP, an average value that was kept during the 2000s. In the cases of Belgium and Italy, their high debt service payments induced substantial budget deficits despite primary surpluses. A reversal of that general trend is noticeable only at the end of the

1990s, as several “more indebted” European countries tried to fulfil or at least come closer to the Maastricht criteria (much of that effort was reversed in the most recent crisis). All in all, the main conclusion is that the burden of government debt has increased over time in almost every country under scrutiny.

Plotting an equivalent graph for total government expenditures and revenues (% GDP) yields Figure 2 (for the same selected countries). This visual inspection may help to assess sustainability issues in individual cases.

Figure 2: Total Government Expenditures and Revenues: 1970-2010 (selected countries)



Source: authors' computations

We continue with some descriptive statistics and in order to evaluate the possibility of the presence of a cyclical component in the series, objects such as the correlogram and the power spectrum can provide useful information. If sometimes the correlogram shows only small individual autocorrelations not providing enough strong evidence of the presence of cyclical movement in the series (despite some cycle evidence may be buried with noise), then a much

clearer message emerges from the examination of the spectrum (not shown). Figure 3 shows the extracted (Structural Time Series Model-based)¹⁵ annual trend growth for total government debt, total government expenditures and total government revenues (all in % GDP), respectively in panels (a), (b) and (c) for France, Japan, the UK and the US (for reasons of parsimony).¹⁶

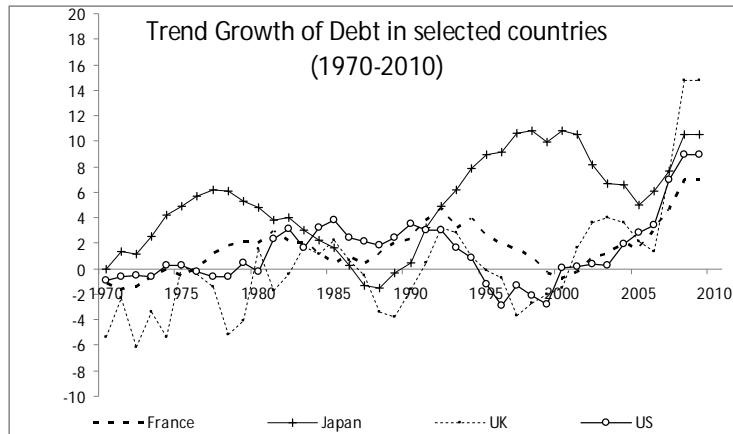
A few facts are worth noticing. Looking at Figure 3 a) we see that following Japan's crisis in the early 1990s we observe a rise in the country's debt trend growth till around 2000. It then decreased slightly until regaining impetus following the latest 2007-08 economic and financial crisis. For France, UK and US the biggest jump in the debt trend growth took place after the latest 2007-08 crisis. In panel b) the highest expenditures trend growth is attributed to the US which has been rising since the early 2000s and the events following the 9/11 of 2001 and wars against terror. Finally, the latest crisis has naturally affected the inflows of government revenues (panel c)) with the most dramatic falls taking place once again in the US following the 9/11 and the latest crisis.

¹⁵ See Harvey (1989, 1991) for details. This class of models is preferred to other trend-cycle filters, as the HP. This one, which has been employed widely in the recent RBC literature, is not considered appropriate as it explicitly neglects low frequency swings by assumption and Harvey and Jaeger (1993) show that the HP filter may create spurious cycles and other distortions.

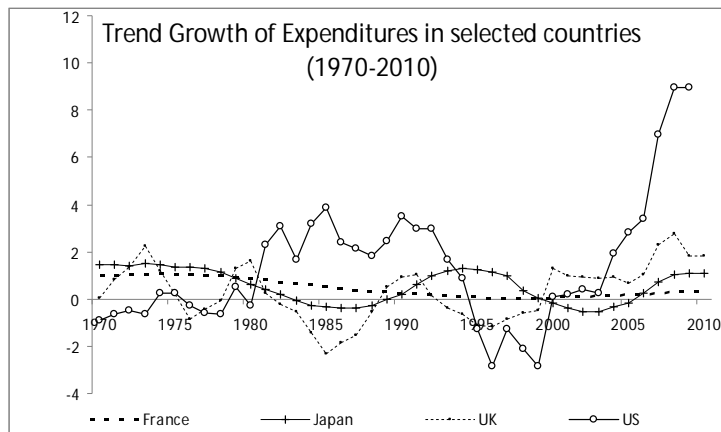
¹⁶ Upon careful examination of several descriptive statistics, a likely specification for the trend and the cycle has been estimated. For reasons of parsimony, basic diagnostic, goodness-of-fit statistics and estimated variances for the hyperparameters for different structural time series models are available upon request. All these models assumed the presence of a trend, one cycle and an irregular component. Three STMs were considered: the first statistical specification assumed that the trend component follows a random walk with drift, with a deterministic slope; the second introduced a somewhat smoother trend with a deterministic level and a stochastic slope; finally, we had a local linear trend model, which stipulates the level and slope to be stochastic. The best model was selected based, primarily, on both AIC and BIC Information Criteria (amongst others) and this corresponds to the version(s) plotted in Figure 3.

Figure 3: Trend Growth of Debt, Expenditures and Revenues, 1970-2010

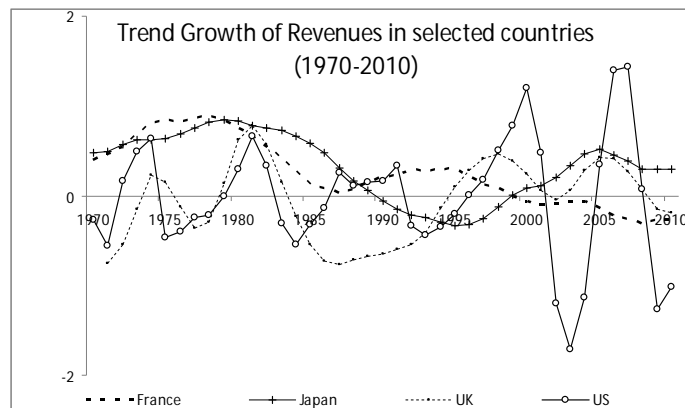
(a)



(b)



(c)



Source: authors' computations

4.2 Country Analysis

The fact that our series are in ratio to GDP does not rule them out being integrated processes (see, Ahmed and Yoo, 1989). Hence, we now focus on fiscal policy sustainability for

each of the 18 countries by means of several unit root tests in an attempt to validate the sufficient sustainability condition using the stock of government debt. Table 1.a shows the stationarity tests results for the first difference of the debt ratio for the period 1970-2010.

The results for the ADF and PP test (considering both a constant and a time trend) allow the rejection of the null of a unit root only in Australia, Germany, Greece and the UK. Therefore the series of the first difference of government debt might be $I(0)$ and the solvency condition would be satisfied in those cases since non-stationarity can be rejected. The Ng and Perron (2001) tests add Finland, Netherlands and Portugal to the “rejection-of-the-null” set of countries.¹⁷

[Table 1.a]

The previous set of results assumes that there is no structural break in the government debt series. However, this might not be the case in some countries – for example, in periods of war or important economic downturns. In the presence of structural changes in the trend function, ADF and PP tests that do not take into account the break in the series have low power, and are biased toward the non-rejection of a unit root. Therefore, in Table 1.a we also report the identified structural breaks. Depending on the precise test, we may get different results for break dates, with the overwhelmingly conclusion that most series are $I(1)$, apart from Australia, Canada, Denmark and Finland for the ZA, VP and CMR tests. For instance, we get reported breaks for Finland in 1990-1993, time when the country was experiencing a severe recession. For Portugal we get dates around 1974, 1982-83 and 2003, corresponding to the “Carnation Revolution”, the IMF program intervention and one severe recession, respectively.¹⁸ One can also note the different power attributed to the PY2009 test (particularly as the ZA and VP tests are conservative in the $I(0)$ case, and have lack of power in the $I(1)$ case) where in all but 4 cases we reject the null of unit root.

Turning to total government expenditures, total government revenues and the primary balance series (all in % GDP) - Table 1.b - we find similar results, with the non-rejection of the null of unit root in levels for most countries (apart from Australia in the case of expenditures and primary balance, and Germany and Sweden in the case of the primary balance). We observe fewer rejections of the null of unit root in the break-type tests (in particular, Austria, Finland, France and the Netherlands for the VP test in the revenues’ case; Austria, Finland and Germany for the VP test in the expenditures’ case; and Australia, Austria, Denmark, Finland, Netherlands, Sweden and the US for the ZA and VP tests in primary balance case).

¹⁷ One should also note that the number of observations used is only 41 at most, and the accuracy problems of unit-root tests with small samples are well known. Afonso and Jalles (2011a) make use of the longer time series debt data going back to the late 19th century to study fiscal sustainability, and such considerations do not apply.

¹⁸ Several studies detect in such periods expansionary fiscal consolidation episodes (see, Afonso, 2010).

[Table 1.b]

Stationarity aside, we now address the cointegration issue, given by (10) and (11), by analysing the relationship between revenues and expenditures and between the primary balance and (lagged) debt. Table 2.a presents the results for the Johansen-Juselius cointegration test for the former case. We find evidence of one cointegrating relationship in only 6 countries (Australia, Austria, Denmark, Germany, Japan and Netherlands); whereas in Table 2.b (for the latter case), we find evidence of cointegration in 8 countries (Austria, Canada, France, Germany, Japan, Netherlands, Sweden and the UK). Therefore, one would not reject the idea that public finances have been less unsustainable in those countries.¹⁹ Moreover, in these cases results from the Hansen-stability test did not reject the null hypothesis that the series are cointegrated at conventional levels (with p-values larger than 20%). Overall, test results allow the rejection of the cointegration hypothesis for the majority of the countries in both relationships under scrutiny.

[Table 2.a, 2.b]

As previously discussed in Section 3.1.1 we further test the hypothesis of a structural shift in the cointegration relationship for all countries in our sample, by using the Gregory and Hansen (1996) procedure. Table 3 presents our results. After taking into account the possibility of breaks in the series, we get for the revenues-expenditure relationship, rejections of the null of no cointegration in 9 countries for the ADF* statistic (relatively in line with previous findings);²⁰ similarly for the balance-(lagged) debt relationship, we reject the null in only 4 countries. In other words, for the period 1970-2010, government expenditures, in half of the countries, exhibited a higher growth rate than public revenues, challenging therefore the hypothesis of fiscal policy sustainability.

[Table 3]

We are now in position to estimate the parameter β (10) and (11). The estimation is made using the DOLS of Stock and Watson (1993) as previously described. The results of the estimation of this equation for each country, in terms of the coefficient β and the statistic C_μ , a LM statistic from the DOLS residuals, which tests for deterministic cointegration (i.e., when no

¹⁹ For instance Davig and Leeper (2011), estimated fiscal reaction rules for the US, and concluded for an active fiscal stance in the early to mid-1980s, and through the 2000s, which indirectly hints at the lack of fiscal sustainability. This is rather in line with our results in Table 2.b, where for the case of the US we cannot reject the hypothesis of the absence of cointegration between government spending and revenues.

²⁰ Our results, as most of the results reported in the literature, were obtained without considering additional sources of government revenues: for instance, seignorage and privatization revenues. Additionally, government assets (wealth) should be taken into account to make judgements about the sustainability of public finances (even though data are mostly lacking).

trend is present in the regression), appear in Table 4. Two main results can be obtained from Table 4.a. First, since most of the cointegration statistics are highly significant at usual levels, the null of deterministic cointegration is rejected (less so in the case of the (lagged) debt-primary balance relationship). And, second, the estimates of β are in 15 out of 18 cases positive and statistically significant for the revenues-expenditures relationship. Moreover, they are always less than one, that is, for each percentage point of GDP increase in public expenditures, for instance in Denmark and in Canada, public revenues only increase by respectively 0.70 and 0.33 percentage points of GDP. In the case of the primary balance-debt relationship we obtain positive and statistically significant estimates of β in 8 out of 18 cases (our results are in line with Bajo-Rubio et al. (2009) findings).

[Table 4.a]

As two robustness checks, we have redone the exercise by: i) allowing for more lags in (10), that is, this time testing cointegration between the primary balance and the second lagged debt ratio; and ii) testing cointegration between the primary balance and the lagged debt deviation to each country's mean over the entire period. Results (see Table 4.b) yield positive and statistically significant estimates of β in 9 (7) out of 18 cases for the former (latter) exercise and roughly corresponding to the same set of countries.²¹ Therefore, the conclusion that emerges is that we cannot say that fiscal policy has been sustainable for half of the countries in our sample.

[Table 4.b]

An additional exercise is to explore the causality direction between total government revenues and expenditures. Table 5.a present our results for the standard Granger causality test but also the Toda-Yamamoto one. In general, focusing on the former test first, evidence suggests stronger effects running from revenues to expenditures. In Canada, however, two-way causality is found, that is, we have “fiscal synchronization” in line with the assumption of equivalency between the marginal costs and marginal revenues that the utility-maximization suppliers and demanders of the public services make. It seems that in only 6 cases we have causality running from expenditures to revenues (the “spend and tax” hypothesis), meaning that the majority of fiscal authorities are not able to generate the revenues required to finance the planned expenditures; that is, the authorities have not kept fiscal budgets under control. Similar conclusions result from inspecting the Toda-Yamamoto test with Germany and Netherlands

²¹ We thank an anonymous referee for helping us explore additional possible channels through which cointegration could be realised.

emerging as the two cases where two-way causality is present, that is, where we have “fiscal synchronization”.

[Table 5a]

As remarked before, in equilibrium the fiscal solvency condition holds in both the MD and FD regimes and the positive estimates of β found in Table 4 can be found in both of them. Hence, Table 5.b presents the results from the standard Granger-causality tests together with the Toda-Yamamoto version, similarly to Table 5.a. Two-way causality was found in Italy in the former test, and in Japan and the Netherlands in the latter test. For these countries results from causality tests do not allow us to conclude whether fiscal solvency would have followed a MD or FD regime between 1970 and 2010. Granger-causality just from primary balance to debt appears for Finland, Ireland, Spain and the UK. On the other hand, Granger causality from government debt to the primary balance is found for 12 countries, which can be seen as evidence of the existence of a Ricardian regime.

[Table 5b]

4.3 Panel Analysis

Tables B1.a and B1.b in Appendix B report the outcome for the full sample of three panel unit root tests: two first generation type, namely IPS and MW and one second generation type, namely CIPS. They show that the null hypothesis of unit roots for the panel data for debt, total government expenditures, revenues and the primary balance cannot be rejected when variables are taken in levels.²² These results strongly indicate that the variables are non-stationary in level and stationary in first differences (with the exception of the primary balance with first generation panel unit root tests – Table B1, last column).

Table 6 shows the outcomes of Pedroni’s (1999) cointegration tests between total government revenues and expenditures and the primary balance and (lagged) debt. We use four within-group tests and three between-group tests to check whether the panel data are cointegrated. The columns labelled within-dimension contain the computed value of the statistics based on estimators that pool the autoregressive coefficient across different countries for the unit root tests on the estimated residuals. The columns labelled between-dimension report the computed value of the statistics based on estimators that average individually calculated coefficients for each country. Results of the within-group tests and the between-group tests show that the null hypothesis of no cointegration can be rejected. Therefore, the

²² Panel stationarity was also conducted for the debt deviation to the mean for each country in our dataset over the entire period. Results are in line with the ones for the original first difference debt variable (see Appendix B).

relationships identified in equations (10) and (11) are cointegrated for the panel of all countries in our sample.²³

[Table 6]

We then estimate the cointegrating vector using the FMOLS estimator. Table 7 shows the coefficients obtained with this estimator. The estimated coefficient for the pool of all countries is 0.51 and 0.03 (statistically significant at the 1% level) for the revenues-expenditures and primary balance-debt relationships, respectively. Focusing on the former, as before (Table 4), it seems that in general the greatest share of results point to a positive long-run co-movement between the levels of total government revenues and expenditures. On the second relationship, the average result points to solvency, even though, a country-by-country inspection shows that only Australia, Belgium, Germany, Ireland, Netherlands and the UK present statistically significant positive coefficient estimates for the improvement of the primary balance when there is a past worsening of the debt ratio. This result improves slightly if one considers the (lagged) debt-mean deviation variable instead. That is, a country-by-country inspection shows that now Austria, Belgium, Canada, France, Germany, Greece, Ireland, Italy and Netherlands present statistically significant positive coefficient estimates.

[Table 7]

Turning to the Pedroni causality tests, one should note first that despite the fact that these tests can be implemented on a country-by-country basis²⁴, in practice the reliability of these various point estimates and associated tests for any one country is likely to be poor given the relatively short time sample over which the data are observed. Therefore, our tests will be panel based. In particular, we want to know more about the pervasiveness of a long-run causal effect in the panel rather than simply finding that there is at least some long-run causality present in at least one specific country. To this end, we use both a group mean based test²⁵ and a lambda-Pearson based test²⁶. The combination of the group mean and the lambda-Pearson can be particularly informative when the underlying parameters of interest are heterogeneous. For

²³ The same is true if instead of the debt-ratio if one computes the same test statistics for the debt-mean deviation.

²⁴ Results are available upon request.

²⁵ The group mean test is based on the sample average of the individual country λ_{1i} tests and will allow us to ask whether the long-run causal effect is zero on average for the panel. The group mean panel estimate is computed as $\bar{\lambda}_1 = N^{-1} \sum_{i=1}^N \hat{\lambda}_{1i}$ and the group mean panel test for the null of no long-run causal effect from unemployment to productivity is computed as $\bar{t}_{\lambda_1} = N^{-1} \sum_{i=1}^N t_{\lambda_{1i}}$, where $t_{\lambda_{1i}}$ is the individual country test for the null that $\lambda_{1i} = 0$.

²⁶ The lambda-Pearson panel test uses the p-values associated with each of the individual country t -tests to compute the accumulated marginal significance associated with these. It takes the form $P_{\lambda_1} = -2 \sum_{i=1}^N \ln p_{\lambda_{1i}}$, where $\ln p_{\lambda_{1i}}$ is the log of the p-value associated with individual country i 's t -test for the null that $\lambda_{1i} = 0$.

instance, when \bar{t}_{λ_1} fails to reject the null while P_{λ_1} succeeds in rejecting the null, this can be interpreted as a situation in which we do not reject that the average value for λ_{1i} is zero, even though we reject that it is pervasively zero in the panel.²⁷ The results for each of these panel tests for the direction of long-run causality and the sign of the long-run causal effect are reported in Table 8.

[Table 8]

In examining the details of Table 8, panel A, the first note goes to the λ_{1i} parameters as reported in columns 2 through 4, which indicate that long-run causality that does not run from expenditures to revenues (p-values above 10%). This confirms our earlier “time-series” findings, that is, the non-validity of the “Spend and Tax” hypothesis, meaning that most fiscal authorities are not able to generate the revenues required to finance the planned expenditures. Turning to λ_{2i} , we reject the hypothesis that revenues have a zero average long-run effect globally (group mean tests) on spending. The results hold pervasively among individual countries and on average for the entire panel (based on the group-mean and Lamba-Pearson tests). The implication of these results is that changes in revenues appear to induce permanent changes in long-run expenditures. On average the marginal long-run impact is zero.

Turning to Table 8, panel B, looking at the λ_{1i} parameters we conclude that long-run causality seems to run from lagged debt to the primary balance (p-values below 10%). This confirms our previous “time-series” results, that is, we cannot reject that government debt Granger causes the primary balance. The results hold pervasively among individual countries and on average for the entire panel (based on the group-mean and Lamba-Pearson tests). Turning to λ_{2i} , we cannot reject the hypothesis that primary balances have a zero average long-run effect globally (group mean tests). In other words, there is no long-run causality in this case running from the primary balance to government debt. At the same time, the sign of the effect is mixed, so that the average is still zero. As before, on average the marginal long-run impact is zero.

In order to control additionally for other effects on the a fiscal reaction specification as (10), we have made the estimation using information on financial crisis. The rationale is that when such crisis occur, the response of primary balances to increases in government indebtedness will tend to be lower, given the additional effort asked to public finances under those circumstances. Therefore, we use financial crises dummies derived from Leaven and Valencia’s (2010) publicly

²⁷ This can occur when the value for λ_{1i} is significantly positive for some fraction of the panel and significantly negative for another fraction of the panel. In this case, we can say that a long-run causal effect is present, even if for some members of the panel it is positive while for others it is negative.

available database, and use them both in an additive and multiplicative (with the debt variable) in a pooled estimation of (1).

[Table 9]

The results in Table 9 show that financial crisis negatively impinge on the primary balance, as expected, and this is a consistently statistically significant result. For instance, in the case of the multiplicative dummy, the overall reaction of the primary balance to government debt developments may even be reversed, undermining fiscal sustainability.²⁸

5. Conclusion

In this paper we have revisited the issue of fiscal policy sustainability in a sample of 18 OECD countries, with annual data between 1970 and 2010, by means of both time-series and panel data techniques. Our main results point to the non-stationarity of the first-differenced debt series for most countries (with the exception of Australia, Germany, Greece and the UK with the ADF and PP tests and adding Finland, Netherlands and Portugal with the Ng and Perron tests) suggesting that the solvency condition would not be satisfied. We find similar results in the cases of total government expenditures, total government revenues and the primary balance series, with the non-rejection of the null of unit root (in levels) for most countries.

Moreover, evidence suggests the existence of one cointegrating relationship in only 6 countries between revenues and expenditures. However, the overall test results allow the rejection of the cointegration hypothesis in both relationships under scrutiny. In other words, for the period 1970-2010, government expenditures, in half of the countries, exhibited a higher growth rate than public revenues, challenging therefore the hypothesis of fiscal policy sustainability.

Estimating the cointegrating coefficient we get 15 out of 18 cases positive and statistically significant estimates for the revenues-expenditures relationship and these are always less than one, that is, for each percentage point of GDP increase in public expenditures, revenues increase by less than one percentage point of GDP.

In terms of individual-country causality, evidence suggests stronger effects running from revenues to expenditures. Moreover in only 6 cases we have causality running from expenditures to revenues (the “spend and tax” hypothesis), meaning that the majority of fiscal authorities are not able to generate the revenues required to finance the planned expenditures.

²⁸ For instance, in column (3) in Table 9, we have: $\partial s / \partial B = 0.014 - 0.049FC$, implying an effect of -0.035 when a financial crisis takes place.

On the other hand, Granger causality from government debt to the primary balance is found for 12 countries, which can be seen as evidence of the existence of a Ricardian regime.

Our panel data results corroborate time-series findings with the greatest share of empirical evidence pointing to a positive long-run co-movement between the levels of total government revenues and expenditures, that is, changes in revenues appear to induce permanent changes in long-run expenditures.

Even though we find that long-run causality seems to run from lagged debt to the primary balance, on average the marginal long-run impact is zero. All in all, we cannot say that fiscal policy has been sustainable for many countries in our sample.

References

1. Abbas , A., Belhocine, N., El Ganainy, A., Horton, M. (2010), “A Historical Public Debt Database”, IMF Working Paper No. 10/245.
2. Afonso, A. (2005), “Fiscal Sustainability: the Unpleasant European Case”, FinanzArchiv, 61 (1), 19-44.
3. Afonso, A. (2008), “Ricardian Fiscal Regimes in the European Union”, Empirica, 35 (3), 313–334.
4. Afonso, A. (2010), “Expansionary fiscal consolidations in Europe: new evidence”, Applied Economics Letters, 17(2), 105-109.
5. Afonso, A., Rault, C. (2010). “What do we really know about fiscal sustainability in the EU? A panel data diagnostic”, Review of World Economics, 145 (4), 731-755.
6. Afonso, A., Jalles, J. (2011a). “A longer-run perspective on fiscal sustainability”, ISEG-UTL, Department of Economics, Working Paper n° 17/2011/DE/UECE.
7. Afonso, A, Jalles, J. (2011b). “Appraising fiscal reaction functions”, Economics Bulletin, 31 (4), 3320-3330.
8. Ahmed, S. and Rogers J. H. (1995), “Government budget deficits and trade deficits Are present value constraints satisfied in long-term data”, Journal of Monetary Economics 36(2), 351-374.
9. Ahmed, S. and Yoo, B. (1989), “Fiscal trends and real business cycles”, Working paper, Pennsylvania State University, University Park, PA.
10. Andrews, D.W. K., and Ploberger, W. (1994), “Optimal Tests When a Nuisance Parameter Is Present Only Under the Alternative”, Econometrica, 62, 1383-1414.
11. Arellano, M. and Bover, O. (1995), "Another Look at the Instrumental Variable Estimation of Error Component Models", Journal of Econometrics, 68: 29-51.
12. Arghyrou, Michael G. and Kul B. Luintel (2007), “Government Solvency: Revisiting Some EMU Countries”, Journal of Macroeconomics 29, 387-410.

13. Bajo-Rubio, Óscar, Díaz-Roldán, Carmen and Esteve, Vicente (2009), "Deficit sustainability and inflation in EMU: An analysis from the Fiscal Theory of the Price Level", European Journal of Political Economy, 25(4), 525-539.
14. Barro, R., (1979), "On the Determination of the Public Debt," Journal of Political Economy, 87, 940-971.
15. Bergman, M. (2001), "Testing Government Solvency and the No Ponzi Game Condition", Applied Economics Letters, 8(1), 27-29.
16. Blanchard, O., J.C. Chouraqui, R.P. Hagemann and N. Sartor (1990), "The sustainability of Fiscal Policy: New Answers to an Old question", OECD Economic Studies, 15, 7-36.
17. Bohn, H. (1991), "The Sustainability of Budget Deficits with Lump-Sum and with Income-Based Taxation", Journal of Money, Credit, and Banking 23 (3), Part 2: 581-604.
18. Bohn, H. (1998), "The Behavior of U.S. Public Debt and Deficits", Quarterly Journal of Economics 113, 949-963.
19. Bohn, H. (2007), "Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint?" Journal of Monetary Economics, 54(7), 1837-1847.
20. Bravo, A. B. S., Silvestre, A. L. (2002), "Intertemporal Sustainability of Fiscal Policies: Some Tests for European Countries", European Journal of Political Economy, 18, 517-528.
21. Bruggeman, A., P. Donati, and A. Warne (2003), "Is the Demand for Euro Area M3 Stable?", European Central Bank - Working Paper Series, No. 255.
22. Canning, C. and Pedroni, P. (2008), "Infrastructure, Long Run Economic Growth and Causality Tests for Cointegrated Panels", Manchester School, 76, 504-527.
23. Canzoneri, M., Cumby, R., Diba, B. (2001), "Fiscal discipline and exchange rate regimes", Economic Journal, October, 667-690.
24. Chalk, N. and Hemming, R. (2000), "Assessing fiscal sustainability in theory and practice", (IMF Working Paper No. 00/81). International Monetary Fund.
25. Clemente, J., Montañés, A., and Reyes, M. (1998), "Testing for a unit root in variables with a double change in the mean", Economics Letters, 59, 175-182.
26. Davig, T., Leeper, E. (2011). "Monetary-fiscal policy interactions and fiscal stimulus", European Economic Review, 55 (2), 211-227
27. Davig, T., Leeper, E., Walker, T. (2011). "Inflation and the fiscal limit," European Economic Review, 55(1), 31-47.
28. Dolato, J., Lutkepohl, H. (1996), "Making Wald test work for cointegrated VAR systems", Econometrics Review 15, 369-386.
29. Feve, P., and Henin, P.Y. (2000), "Assessing Effective Sustainability of Fiscal Policy within the G-7", Oxford Bulletin of Economics and Statistics, 62, 175.

30. Fincke, B., Greiner, A. (2011). "Debt Sustainability in Selected Euro Area Countries: Empirical Evidence Estimating Time-Varying Parameters", *Studies in Nonlinear Dynamics & Econometrics*. 15 (3), article 2.
31. Friedman, M. (1978), "The Limitations of Tax Limitation," Policy Review, 7-14.
32. Geweke, J., Meese, R. and Dent, W. T. (1983), "Comparing alternative tests of causality in temporal systems: Analytic results and experimental evidence", Journal of Econometrics, 21, 161- 94.
33. Granger, C. (1969), "Investigating causal relations by econometric models and cross-spectral methods", Econometrica 37, 424-38.
34. Gregory, A. and Hansen, B. (1996), "Residual-based tests for cointegration in models with regime shifts", Journal of Econometrics, 70, 99-126.
35. Hakkio, C.S., and Rush, M. (1991), "Is the Budget Deficit "Too Large"?" *Economic Inquiry* 29, 429-445.
36. Hamilton, J.D., and Flavin, M.A. (1986), "On the Limitations of Government Borrowing: A Framework for Empirical Testing", American Economic Review 76, 808-819.
37. Hansen, H. and S. Johansen (1993), "Recursive Estimation in Cointegrated VAR-Models", Mimeo. Institute of Mathematical Statistics, University of Copenhagen.
38. Harvey, A.C. (1989), *Forecasting, Structural Time Series Models and the Kalman Filter*, Cambridge University Press.
39. Harvey, A.C. (1991), *An Econometric Analysis of Time Series*, 2nd Edition, Cambridge University Press.
40. Harvey, A.C. and A. Jaeger, 1993, Detrending, stylized facts and the business cycle, Journal of Applied Econometrics 8, 231-247.
41. Hatemi-J A. (2002), "Fiscal Policy in Sweden: Effects of EMU Criteria Convergence", Economic Modelling, 19 (1).
42. Hoffman, D. L., R. H. Rasche, and M. A. Tieslau (1995), "The Stability of Long-Run Money Demand in Five Industrial Countries", Journal of Monetary Economics 35, 317-339.
43. Hoover K. D. and Sheffrin S. M. (1992), "Causation, spending and taxes: Sand in the sandbox or tax collector for the welfare state", American Economic Review, 82(1), 225-248.
44. Im, K. S., Pesaran, M. H. and Shin, Y. (2003), "Testing for unit roots in heterogeneous panels", Journal of Econometrics 115, 53-74.
45. Johansen, S. and Juselius, K. (1990), "Maximum Likelihood Estimation and Inference on Cointegration – with Applications to the Demand for Money", Oxford Bulletin of Economics and Statistics 52, 169-210.
46. Kocherlakota, N., and C. Phelan (1999) "Explaining the Fiscal Theory of the Price Level," Federal Reserve Bank of Minnesota Quarterly Review 23(4), 14-23.
47. Lee, C., Chang, C., and Chen, P. (2008), "Energy-income causality in OECD countries revisited: The key role of capital stock", Energy Economics, 30, 2359-2373.

48. Leeper, E.M. (1991) "Equilibria Under 'Active' and 'Passive' Monetary and Fiscal Policies," Journal of Monetary Economics 27, 129-147.
49. Laeven, L., Valencia, F. (2010), "Resolution of Banking Crises: The Good, the Bad, and the Ugly," IMF Working Paper 10/146.
50. Legrenzi, G., Milas, C. (2011). "Debt Sustainability and Financial Crises: Evidence from the GIIPS", CESifo WP 3594.
51. MacDonald, R. (1992), "Some Tests of the Government Intertemporal Budget Constraint Using U.S. Data", Applied Economics 24(12): 1287-92.
52. Maddala, G. S., and Wu, S. (1999). A Comparative Study of Unit Root Tests with Panel Data and New Simple Test", Oxford Bulletin of Economics and Statistics, 61, 631-652
53. Makrydakis S., Tzavalis E. and Balfoussias A. (1999), "Policy regime changes and the long-run sustainability of fiscal policy: an application to Greece", Economic Modeling 16, 71-86.
54. Martin, G.M. (2000), "U.S. Deficit Sustainability: A New Approach Based on Multiple Endogenous Breaks", Journal of Applied Econometrics, 15(1), pp. 83–105.
55. McCallum, Bennett T. and Nelson, Edward. "Targeting vs. Instrument Rules for Monetary Policy." Federal Reserve Bank of St. Louis Review, September/October 2005, 87(5), 597-611.
56. Mendoza, E.G., and J.D. Ostry (2008), "International Evidence on Fiscal Solvency: Is Fiscal Policy 'Responsible'?" Journal of Monetary Economics, 55(6) 1081–93.
57. Musgrave, R. (1996), "Principles of Budget Determination," in H. Cameron and W. Henderson, *Public Finance: Selected Readings*, New York: Random House.
58. Ng, S., and P. Perron (2001), "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," Econometrica, 69, 1519-1554.
59. O'Connell, P. (1998), "The Overvaluation of Purchasing Power Parity", Journal of International Economics, 44, 1-19.
60. Pedroni P. (2004), "Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time series Tests, With an Application to the PPP Hypothesis," Econometric Theory, 20, 597-625.
61. Pedroni, P. (1999), "Critical Values for Cointegration Tests in Heterogeneous Panels with Multiple Regressors", Oxford Bulletin of Economics and Statistics, 61, 653-670.
62. Pedroni, P. (2000), "Fully Modified OLS for Heterogeneous Cointegrated Panels", in Baltagi, B., C.D. Kao (Eds.), *Advances in Econometrics, Nonstationary Panels, Panel Cointegration and Dynamic Panels*, Elsevier Science, New York, 93-130.
63. Pedroni, P. (2001), "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the PPP Hypothesis", Mimeo, Department of Economics, Indiana University.
64. Perron, P. (1989), "The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis." Econometrica 57, 1361-1401.

65. Perron, P. and Yabu, T. (2009), "Testing for Shifts in Trend with an Integrated or Stationary Noise Component", Journal of Business and Economics Statistics, 27, 369-396.
66. Pesaran, M.H., (2007), "A simple panel unit root test in the presence of cross section dependence", Journal of Applied Econometrics, 22, 265-312.
67. Quintos, C.E. (1995), "Sustainability of the Deficit Process with Structural Shifts", Journal of Business and Economic Statistics 13, 409-417.
68. Rambaldi, A.N. and Doran, H. E. (1996), "Testing for Granger non-causality in Cointegrated System made easy" Working Papers in Econometrics and Applied Statistics No. 88, Department of Econometrics, University of New England.
69. Richter, A. (2011). "The Fiscal Limit and Non-Ricardian Consumers", mimeo, Indiana University.
70. Shin, Y., "A Residual-based Test of the Null of Cointegration against the Alternative of no Cointegration," Econometric Theory, 1994, (10), 91–115.
71. Sims, C. A. (1972) Money, income and causality, American Economic Review, 62, 540-52.
72. Stock, J. and Watson, M., (1993), "A simple estimator of cointegrating vectors in higher order integrated systems", Econometrica , 61(4), 783-820.
73. Toda, H and P. Phillips (1994), "Vector autorgression and causality: A theoretical overview and simulation study", Econometric Review 13, 259-85.
74. Toda, H. and Yamamoto, T. (1995), "Statistical inference in vector autoregressions with possibly integrated processes", Journal of Econometrics 66, 225--250.
75. Trehan, B., and Walsh, C. (1991), "Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits", Journal of Money, Credit and Banking 23, 210-223.
76. Vogelsang, T. (2001), "Testing for a Shift in Trend When Serial Correlation is of Unknown Form," Unpublished Manuscript, Department of Economics, Cornell University.
77. Vogelsang, T. and P. Perron (1998), "Additional Tests for a Unit Root Allowing for a Break in the Trend Function at an Unknown Time", International Economic Review 39(4), 1073-1100.
78. Zivot, E. and D. W. K. Andrews (1992), "Further Evidence on the Great Crash, the Oil-Price Shock and the Unit Root Hypothesis", Journal of Business and Economic Statistics 10, 251-270.

Appendix A

Data sources

| Original series | Ameco codes |
|---|--|
| Total expenditure: general government, Excessive deficit procedure (% of GDP at market prices) | 1.0.319.0.UUTGF, 1.0.319.0.UUTGE |
| Total revenue: general government, Excessive deficit procedure (% of GDP at market prices) | 1.0.319.0.URTGF, 1.0.319.0.URTGE |
| General government consolidated gross debt, excessive deficit procedure (based on ESA 1995) and former definition (linked series) (% of GDP at market prices) | 1.0.319.0.UDGGF, 1.0.319.0.UDGGL |
| Primary Balance (% GDP at market prices) | 1.0.319.0.UBLGI, for EU countries; OECD database for Australia, Canada, Japan, US and |

Appendix B

Table B1: First Generation Panel Unit Root Tests

Im, Pesaran and Shin (2003) Panel Unit Root Test (IPS) (a)

| Full | debt | | debt-mean deviation | | revenues | | expenditures | | primary balance | |
|------------------|-------------|----------------|---------------------|----------------|-------------|----------------|--------------|----------------|-----------------|----------------|
| <i>in levels</i> | <i>lags</i> | <i>[t-bar]</i> | <i>lags</i> | <i>[t-bar]</i> | <i>lags</i> | <i>[t-bar]</i> | <i>lags</i> | <i>[t-bar]</i> | <i>lags</i> | <i>[t-bar]</i> |
| | 1.78 | -0.63 | 1.78 | -0.64 | 0.39 | -3.56** | 1.17 | -3.67** | 0.78 | -4.95** |

Maddala and Wu (1999) Panel Unit Root Test (MW) (b)

| Full | debt | | debt-mean deviation | | revenues | | expenditures | | primary balance | |
|-----------------------------|-------------|------|---------------------|------|-------------|------|--------------|------|-----------------|------|
| <i>lags</i> | P_λ | (p) | P_λ | (p) | P_λ | (p) | P_λ | (p) | P_λ | (p) |
| <i>in levels</i> | | | | | | | | | | |
| 0 | 8.06 | 1.00 | 8.07 | 1.00 | 15.29 | 0.99 | 10.96 | 1.00 | 26.44 | 0.87 |
| 1 | 26.27 | 0.88 | 26.27 | 0.88 | 20.65 | 0.98 | 20.60 | 0.98 | 54.40 | 0.03 |
| 2 | 23.20 | 0.95 | 23.21 | 0.95 | 19.37 | 0.98 | 22.56 | 0.96 | 45.73 | 0.13 |
| <i>in first differences</i> | | | | | | | | | | |
| 0 | 98.35 | 0.00 | 98.36 | 0.00 | 345.88 | 0.00 | 276.25 | 0.00 | 0.18 | 0.57 |
| 1 | 54.39 | 0.02 | 54.39 | 0.03 | 206.00 | 0.00 | 142.29 | 0.00 | -0.47 | 0.32 |
| 2 | 34.08 | 0.56 | 34.09 | 0.56 | 161.21 | 0.00 | 84.46 | 0.00 | 0.45 | 0.67 |

Notes: (a) We report the average of the country-specific "ideal" lag-augmentation (via AIC). We report the t-bar statistic, constructed as $t - bar = (1/N) \sum_i t_i$ (t_i are country ADF t-statistics). Under the null of all country series containing a nonstationary process this statistic has a non-standard distribution: the critical values are -1.73 for 5%, -1.69 for 10% significance level – distribution is approximately t . We indicate the cases where the null is rejected with **. (b) We report the MW statistic constructed as $p_\lambda = -2 \sum_i \log(p_i)$ (p_i are country ADF statistic p-values) for different lag-augmentations. Under the null of all country series containing a nonstationary process this statistic is distributed $\chi^2(2N)$. We further report the p-values for each of the MW tests.

Table B2: Second Generation Panel Unit Root Tests

Pesaran (2007) Panel Unit Root Test (CIPS)

| Full | debt | | debt-mean deviation | | revenues | | expenditures | | primary balance | |
|-----------------------------|-------------|------|---------------------|------|-------------|------|--------------|------|-----------------|------|
| <i>lags</i> | P_λ | (p) | P_λ | (p) | P_λ | (p) | P_λ | (p) | P_λ | (p) |
| <i>in levels</i> | | | | | | | | | | |
| 0 | 4.22 | 1.00 | 4.94 | 1.00 | 0.32 | 0.62 | 1.41 | 0.92 | 405.13 | 0.00 |
| 1 | 2.29 | 0.98 | 2.84 | 0.99 | -0.07 | 0.47 | 0.44 | 0.67 | 208.96 | 0.00 |
| 2 | 3.52 | 1.00 | 3.40 | 1.00 | 0.62 | 0.73 | 1.67 | 0.95 | 134.98 | 0.00 |
| <i>in first differences</i> | | | | | | | | | | |
| 0 | -5.63 | 0.00 | -5.64 | 0.00 | -13.76 | 0.00 | -12.12 | 0.00 | -15.97 | 0.00 |
| 1 | -2.53 | 0.00 | -2.54 | 0.01 | -7.65 | 0.00 | -5.74 | 0.00 | -11.07 | 0.00 |
| 2 | 0.57 | 0.71 | 0.57 | 0.72 | -3.82 | 0.00 | -0.82 | 0.20 | -7.47 | 0.00 |

Notes: Null hypothesis of non-stationarity. We further report the p-values for each of the CIPS tests.

TABLE 1.a Unit Root Tests and Structural Breaks: First-Differenced Public Debt 1970-2010

| Countries | ADF | PP | NP | | | | ZA | VP(AO) | VP(IO) | CMR(AO) | CMR(IO) | PY2009 |
|----------------|----------|----------|------------|------------|------------|------------|--------|--------|--------|------------|--------------|---------|
| | (1) | (2) | MZa (3) | MZi (4) | MSB (5) | MPT (6) | (7) | (8) | (9) | (10) | (11) | (12) |
| Australia | -4.98*** | -4.98*** | -3.66 | -1.30 | 0.35 | 24.02 | 2002 | 2005 | 2004** | 1998, 2005 | 2004, 2008** | 1995*** |
| Austria | -2.15 | -2.42 | -8.34 | -1.99 | 0.23 | 11.07 | 1992 | 2003 | 1999 | 1988, 2002 | 1990, 1999** | 2002 |
| Belgium | -1.89 | -1.78 | -6.78 | -1.69 | 0.24 | 13.53 | 1990 | 1989 | 1979 | 1985, 2002 | 1980, 1996 | 1993*** |
| Canada | -3.25* | -3.39* | -13.81 | -2.57 | 0.186 | 6.92 | 1992 | 1988** | 1981 | 1988, 2002 | 1981, 2002 | 1991*** |
| Denmark | -2.18 | -2.20 | -9.55 | -1.91 | 0.20 | 10.56 | 1983 | 1987** | 2004 | 1984, 2003 | 1980, 2004 | 1980*** |
| Finland | -4.05** | -2.09 | -31.86*** | -3.99*** | 0.12*** | 2.86*** | 1993** | 1995 | 1990** | 1985, 1995 | 1980, 1990** | 1990*** |
| France | -3.04 | -3.07 | -11.45 | -2.37 | 0.20 | 8.05 | 1994 | 1998 | 1991 | 1984, 1995 | 1981, 1991 | 2003*** |
| Germany | -3.96** | -3.96** | -16.12 | -2.83 | 0.17 | 5.65 | 1996 | 1998 | 1992 | 1984, 1997 | 1974, 1992 | 2003*** |
| Greece | -4.95*** | -5.03*** | -16.11* | -2.83* | 0.17* | 5.67* | 1991 | 1990 | 1983 | 1984, 1991 | 1983, 1989 | 2000 |
| Ireland | -3.25* | -3.19 | -14.80* | -2.40 | 0.16** | 7.90 | 1987 | 1989 | 1996 | 1984, 1998 | 1981, 1996 | 1995*** |
| Italy | -3.34* | -3.41* | -13.99 | -2.61 | 0.18 | 6.70 | 1993 | 1987 | 1980 | 1987, 1995 | 1981, 1990 | 1999 |
| Japan | -3.04 | -3.28* | -4.92 | -1.56 | 0.31 | 18.51 | 1989 | 2007 | 1994 | 1983, 2001 | 1992, 1997 | 2003*** |
| Netherlands | -3.04 | -3.12 | -15.4* | -2.50 | 0.16** | 7.41 | 1982 | 1994 | 1995 | 1985, 2003 | 1981, 1998** | 1998*** |
| Portugal | -1.56 | -4.28** | -21.06** | -3.21** | 0.15** | 4.51** | 1982 | 1983 | 1980 | 1973, 1983 | 1974, 1980 | 2003*** |
| Spain | -2.73 | -2.79 | -11.39 | -2.32 | 0.20 | 8.32 | 1994 | 1987 | 1980 | 1989, 1993 | 1980, 1989** | 1995*** |
| Sweden | -2.80 | -2.93 | -10.66 | -2.30 | 0.21 | 8.54 | 1997 | 1984 | 1979 | 1984, 2006 | 1979, 2002 | 1998*** |
| United Kingdom | -3.25* | -3.34* | -12.03 | -2.35 | 0.19 | 8.08 | 1990 | 1977 | 1978 | 1977, 1989 | 1988, 1992 | 2003 |
| United States | -2.24 | -2.23 | -11.19 | -2.08 | 0.18 | 9.48 | 1986 | 1988 | 1981 | 1988, 1994 | 1984, 1989 | 1990*** |

Note: ADF critical values: -4.028, -3.445, -3.145 for 1, 5 and 10% levels respectively. For the Ng-Perron test (NP), none of the test statistics are significant at the usual levels. The critical values are taken from Ng and Perron (2001), table 1 and the autoregressive truncation lag (zero) has been selected using the modified AIC. The ZA test statistic reported is the minimum Dickey-Fuller statistic calculated across all possible breaks in the series, when both a break in the intercept and the time trend is allowed for. The year in parenthesis denotes the year when this minimum DF statistic is obtained. The 1% critical value is -5.57 and the 5% critical value is -5.08. As for the VP test, "AO" means addictive outlier and "IO" means innovational outlier and critical values are taken from Perron and Vogelsang (1992), in particular, -3.56 (AO) and -4.27 (IO) for 5% level. As for CMR the 5% critical value is -5.49 (both AO and IO), also taken from Perron and Vogelsang (1992). In column 12 we run the Perron-Yabu (PY) unit root test. For the structural-break type tests only dates are presented and when applicable, a statistically significant symbol is added. The null in the non-break type tests is of unit root. The null in the break-type tests is of unit root against the break stationary alternative hypothesis.

TABLE 1.b Unit Root Tests and Structural Breaks 1970-2010

| Countries/Series | ADF | | PP | | NP | | | | ZA | VP(AO) | VP(IO) | CMR(AO) | CMR(IO) | PY2009 |
|------------------|----------|----------|--------|-----------|-----------|-----------|---------|---------|--------|--------|--------|--------------|--------------|---------|
| | Levels | FD | Levels | FD | MZa | MZi | MSB | MPT | | | | | | |
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | | | | | | |
| Revenues | | | | | | | | | | | | | | |
| Australia | -1.77 | -7.60*** | -1.77 | -7.60*** | -5.11 | -1.33 | 0.26 | 16.7 | 2004 | 1981 | 1980 | 1981, 1996 | 1980, 1993 | 2003*** |
| Austria | -1.48 | -4.69*** | -1.39 | -4.69*** | -1.07 | -0.50 | 0.47 | 48.18 | 1976 | 1977 | 1974** | 1977, 1999 | 1974, 2000** | 1982*** |
| Belgium | -2.08 | -6.50*** | -2.07 | -6.50*** | -3.14 | -1.13 | 0.35 | 26.22 | 1977 | 1978 | 1973 | 1976, 1995 | 1973, 1991 | 1980 |
| Canada | -0.87 | -3.85** | -0.64 | -4.51*** | -5.74 | -1.46 | 0.25 | 15.43 | 1990 | 1983 | 1978 | 1987, 2003 | 1978, 1988 | 1997*** |
| Denmark | -0.97 | -4.58*** | -1.40 | -6.82*** | -4.43 | -1.24 | 0.28 | 18.61 | 1983 | 1980 | 1981 | 1973, 1982 | 1977, 1982 | 1984 |
| Finland | -1.61 | -5.47*** | -1.22 | -10.13*** | -2.88 | -1.01 | 0.35 | 26.76 | 1990 | 1987 | 1973** | 1972, 1987 | 1973, 1983 | 1990 |
| France | -0.51 | -5.85*** | -0.63 | -5.85*** | -0.75 | -0.34 | 0.46 | 48.77 | 1979 | 1980 | 1973** | 1979, 1994** | 1973, 1991 | 1985*** |
| Germany | -2.77 | -6.02*** | -2.80 | -6.01*** | -4.32 | -1.30 | 0.30 | 19.64 | 1992 | 1988 | 1989 | 1993, 2002 | 1989, 1999** | 2000 |
| Greece | -1.77 | -5.46*** | -1.92 | -5.42*** | -6.17 | -1.67 | 0.27 | 14.70 | 1997 | 1991 | 1988 | 1983, 1993 | 1980, 1988 | 1996 |
| Ireland | -1.47 | -6.67*** | -1.43 | -6.67*** | -3.17 | -1.19 | 0.37 | 27.25 | 1985 | 1977 | 1978 | 1982, 1996 | 1978, 1993 | 1984* |
| Italy | -0.90 | -5.83*** | -0.95 | -5.82*** | -3.03 | -1.07 | 0.35 | 26.23 | 1992 | 1984 | 1974 | 1979, 1989 | 1978, 1988 | 1991 |
| Japan | -1.80 | -7.24*** | -1.80 | -7.24*** | -2.95 | -1.16 | 0.39 | 29.41 | 1979 | 1981 | 1977 | 1981, 2007 | 1977, 2004 | 1991 |
| Netherlands | -2.69 | -5.62*** | -2.69 | -5.62*** | -1.96 | -0.95 | 0.48 | 44.07 | 1994 | 1994 | 1992** | 1976, 1996 | 1973, 1992** | 1981** |
| Portugal | -0.66 | -7.53*** | -1.82 | -9.13*** | -3.43 | -1.07 | 0.31 | 22.46 | 1995 | 1993 | 1973 | 1979, 1990 | 1983, 1993 | 1991*** |
| Spain | 0.07 | -5.47*** | 0.28 | -5.46*** | -0.53 | -0.23 | 0.44 | 46.63 | 1994 | 1982 | 1977 | 1978, 1986 | 1973, 1978 | 1993*** |
| Sweden | -1.73 | -6.05*** | -1.68 | -11.91*** | -2.40 | -0.88 | 0.36 | 29.62 | 1976 | 1978 | 1974 | 1978, 2002 | 1974, 1999 | 1986** |
| United Kingdom | -3.09 | -4.97*** | -2.53 | -5.26*** | -18.85** | -3.04** | 0.16** | 4.98** | 1992 | 1988 | 1988 | 1988, 2006 | 1988, 2002 | 1989 |
| United States | -2.70 | -5.95*** | -1.40 | -8.76*** | -19.03*** | -2.94** | 0.15** | 5.63* | 2002** | 1988** | 1985 | 1988, 1998 | 1985, 1999 | 2001*** |
| Expenditures | | | | | | | | | | | | | | |
| Australia | -4.50*** | -4.06*** | -3.13 | -3.37* | -8.14 | -1.98 | 0.24 | 11.27 | 1987 | 1977 | 1973 | 1979, 1999 | 1973, 1979 | 1985*** |
| Austria | -1.96 | -5.69*** | -1.96 | -5.69*** | -2.15 | -0.96 | 0.44 | 38.50 | 1976 | 1977 | 1973** | 1977, 2004 | 1973, 2003 | 1982 |
| Belgium | -2.31 | -6.91*** | -2.31 | -6.87*** | -2.87 | -1.19 | 0.41 | 31.59 | 1984 | 1979 | 1992 | 1978, 1986 | 1973, 1986 | 1983*** |
| Canada | -1.90 | -4.75*** | -1.79 | -4.75*** | -5.49 | -1.65 | 0.30 | 16.55 | 1982 | 2001 | 1997 | 1983, 1998 | 1978, 1997 | 1995 |
| Denmark | -2.03 | -4.61*** | -1.93 | -4.58*** | -6.51 | -1.80 | 0.27 | 13.99 | 1980 | 1978 | 1972 | 1980, 2001 | 1977, 2003 | 1980 |
| Finland | -2.42 | -3.99** | -1.74 | -3.27* | -11.41 | -2.37 | 0.20 | 8.03 | 1990 | 1987 | 1988** | 1987, 1999 | 1988, 1995 | 1997* |
| France | -2.17 | -4.95** | -1.73 | -4.95** | -5.19 | -1.58 | 0.30 | 17.42 | 1981 | 1982 | 1973 | 1979, 1993 | 1973, 1979 | 1985 |
| Germany | -3.07 | -4.94** | -3.07 | -4.80** | -5.62 | -1.67 | 0.29 | 16.20 | 1991 | 1976 | 1972** | 1994, 1999 | 1972, 1988 | 1980 |
| Greece | -1.76 | -7.39*** | -1.64 | -7.55*** | -5.70 | -1.64 | 0.28 | 15.86 | 1981 | 1982 | 1979 | 1982, 1991 | 1972, 1979 | 1983 |
| Ireland | -2.10 | -4.26*** | -1.65 | -4.21*** | -7.11 | -1.84 | 0.25 | 12.86 | 1980 | 1983 | 1984 | 1983, 1996 | 1972, 1993 | 1988 |
| Italy | -2.10 | -4.26*** | -1.65 | -4.21*** | -7.11 | -1.84 | 0.25 | 12.86 | 1980 | 1983 | 1984 | 1983, 1996 | 1972, 1993 | 1988 |
| Japan | -2.74 | -8.12*** | -2.66 | -8.01*** | 6.31 | -1.76 | 0.27 | 14.42 | 1984 | 1995 | 1991 | 1976, 1995 | 1973, 1991 | 1983*** |
| Netherlands | -2.01 | -4.26*** | -1.97 | -4.28*** | -4.24 | -1.45 | 0.34 | 21.39 | 1978 | 1988** | 1992 | 1977, 1995 | 1972, 1992 | 1983 |
| Portugal | -2.62 | -4.94*** | -1.87 | -6.65*** | -4.37 | -1.45 | 0.33 | 20.65 | 1978 | 1982 | 1972 | 1977, 1992 | 1973, 1988 | 1981 |
| Spain | -1.41 | -4.21*** | -1.40 | -4.19** | -3.50 | -1.31 | 0.37 | 25.80 | 1980 | 1982 | 1973 | 1982, 1991 | 1973, 1978 | 1995** |
| Sweden | -2.17 | -4.08** | -1.87 | -4.06** | -6.70 | -1.78 | 0.26 | 13.62 | 1981 | 1978 | 1972 | 1978, 1999 | 1974, 1992 | 1981** |
| United Kingdom | -2.49 | -3.69** | -1.48 | -3.76** | -16.09* | -2.70* | 0.16* | 6.42* | 2002 | 1989 | 1983 | 1987, 2006 | 1983, 2007 | 1985 |
| United States | -2.56 | -4.65*** | -1.93 | -4.43*** | -15.83* | -2.74* | 0.17* | 6.18* | 1997 | 1983 | 1978 | 1983, 1995 | 1980, 1994 | 1996 |
| Primary Balance | | | | | | | | | | | | | | |
| Australia | -3.87** | -4.52*** | -2.46 | -3.74** | -26.94*** | -3.60*** | 0.13*** | 3.75*** | 1986 | 1989 | 1984** | 1986, 1996 | 1972, 1984 | 1985 |
| Austria | -3.70** | -4.92*** | -2.54 | -9.71*** | -12.57 | -2.43 | 0.19 | 7.63 | 1979 | 1972** | 1994 | 1972, 1997 | 1973, 1994 | 1996 |
| Belgium | -1.66 | -7.82*** | -1.64 | -7.75*** | -5.93 | -1.58 | 0.26 | 15.16 | 2003 | 1979 | 1980 | 1979, 1986 | 1980, 1982 | 1984 |
| Canada | -1.92 | -4.94*** | -1.75 | -4.95*** | -9.53 | -2.06 | 0.21 | 10.02 | 1996 | 1990 | 1992 | 1991, 1997 | 1973, 1992 | 1995*** |
| Denmark | -3.52* | -4.63*** | -1.74 | -4.13*** | -27.40*** | -3.63*** | 0.13*** | 3.69*** | 1983 | 1984 | 1981** | 1985, 1989 | 1981, 1990 | 1983 |
| Finland | -3.03 | -5.09*** | -2.48 | -5.02*** | -20.70** | -3.17** | 0.15** | 4.68*** | 2000 | 1986** | 1987 | 1991, 1997 | 1989, 1996 | 1997 |
| France | -2.46 | -6.36*** | -2.63 | -6.37*** | -10.84 | -2.16 | 0.19 | 9.16 | 1997 | 2006 | 1992 | 1989, 1997 | 1990 | 1997 |
| Germany | -4.38*** | -7.96*** | -4.15 | -11.92*** | -17.36** | -2.93** | 0.16** | 5.32* | 1977 | 1993 | 1994 | 1983, 1993** | 1982, 1994** | 1981*** |
| Greece | -2.06 | -6.99** | -2.00 | -7.00 | -7.30 | -1.87 | 0.25 | 12.52 | 1994 | 2005 | 1989 | 1991, 2002 | 1989, 2001 | 1993 |
| Ireland | 3.82 | -0.98 | 2.60 | -0.54 | -15.50*** | -27.30*** | 0.01*** | 0.15*** | 2004 | 2007 | 1985 | 1985, 2004 | 1985 | 2003*** |
| Italy | -1.97 | -6.83*** | -2.04 | -6.94*** | -6.37 | -1.69 | 0.26 | 14.27 | 1992 | 1989 | 1989** | 1991, 2002 | 1990, 1999 | 1991 |
| Japan | -2.24 | -7.56*** | -2.15 | -7.56*** | -8.32 | -2.02 | 0.24 | 10.99 | 1993 | 1996 | 1990 | 1984, 1994 | 1982, 1991 | 1993*** |
| Netherlands | -3.30* | -7.40*** | -3.40* | -9.08*** | -15.38* | -2.59 | 0.16** | 6.94 | 1996 | 2006** | 1994** | 1993, 2001 | 1994, 1999 | 1995*** |
| Portugal | -2.45 | -4.08** | -2.42 | -7.52*** | -9.06 | -2.04 | 0.22 | 10.04 | 1982 | 2006 | 1980 | 1985, 1990 | 1973, 1981 | 1982 |
| Spain | -2.88 | -5.34*** | -1.65 | -4.48*** | -22.36** | -3.18** | 0.14** | 4.98** | 1986 | 1997 | 1984 | 1979, 1988 | 1978, 1984 | 2003 |
| Sweden | -3.94** | -4.20** | -2.47 | -4.23** | -24.49*** | -3.49*** | 0.14** | 3.74*** | 1991 | 1987** | 1988** | 1987, 1993 | 1988, 1999 | 1991 |
| United Kingdom | -3.15 | -4.68*** | -2.43 | -4.66*** | -18.67** | -2.98** | 0.16** | 5.28** | 1998 | 2005 | 1999 | 1991, 1998 | 1987, 1999 | 1997 |
| United States | -2.56 | -4.83*** | -1.24 | -4.66*** | -22.19** | -3.14** | 0.14** | 5.18** | 1997** | 2004** | 1999 | 1995, 2002 | 1995, 2000 | 1996*** |

Note: see Table 1.a for details. In this table, instead of using first-differenced variables, these are tested in levels and, in the cases of the ADF and PP tests, also in first-differences, as identified in columns 3 and 5.

TABLE 2.a Johansen-Juselius Cointegration Tests: Revenues and Expenditures

| Null | Alternative | Aus | Aut | Bel | Can | Den | Fin | Fra | Ger | Gre | Ire | Ita | Jap | Net | Por | Spa | Swe | UK | US |
|------------------------------|-------------|--------|---------|-------|-------|---------|-------|-------|---------|-------|-------|-------|---------|---------|-------|-------|-------|-------|-------|
| λ_{trace} $r = 0$ | $r \geq 1$ | 20.33* | 28.84** | 14.41 | 14.53 | 23.40** | 19.12 | 18.39 | 30.89** | 13.31 | 14.35 | 13.79 | 22.95** | 21.04** | 18.56 | 19.55 | 19.85 | 19.20 | 18.78 |
| | $r \geq 2$ | 8.66 | 9.79 | 1.86 | 5.91 | 4.99 | 7.26 | 5.88 | 9.65 | 4.43 | 5.01 | 5.05 | 2.99 | 3.30 | 7.10 | 8.84 | 5.55 | 5.80 | 4.14 |
| λ_{max} $r = 0$ | $r = 1$ | 11.66 | 19.05** | 12.55 | 8.62 | 18.40** | 11.86 | 12.51 | 21.23** | 8.88 | 9.33 | 8.73 | 19.96** | 18.09** | 11.46 | 10.70 | 14.30 | 13.39 | 14.64 |
| | $r = 2$ | 8.66 | 9.79 | 1.86 | 5.91 | 4.99 | 7.26 | 5.88 | 9.65 | 4.43 | 5.01 | 5.05 | 2.99 | 3.30 | 7.10 | 8.84 | 5.55 | 5.80 | 4.14 |
| Cointegration* | | Yes | Yes | No | No | Yes | No | No | Yes | No | No | No | Yes | Yes | No | No | No | No | No |

Note: * denotes rejection of the null hypothesis of no cointegration at the 5% level (based on MacKinnon-Haug-Michelis p-values).

TABLE 2.b Johansen-Juselius Cointegration Tests: (lagged) Debt and Primary Balance

| Null | Alternative | Aus | Aut | Bel | Can | Den | Fin | Fra | Ger | Gre | Ire | Ita | Jap | Net | Por | Spa | Swe | UK | US |
|------------------------------|-------------|-------|---------|-------|---------|-------|-------|---------|---------|------|-------|------|---------|---------|------|---------|------|---------|-------|
| λ_{trace} $r = 0$ | $r \geq 1$ | 12.19 | 21.16** | 18.14 | 28.08** | 19.73 | 12.68 | 21.45** | 25.83** | 4.84 | 15.07 | 9.36 | 16.98** | 27.51** | 9.41 | 18.42** | 9.63 | 21.91** | 12.80 |
| | $r \geq 2$ | 2.01 | 5.60 | 5.33 | 1.63 | 7.73 | 1.46 | 0.86 | 1.13 | 0.75 | 3.04 | 1.41 | 0.02 | 3.64 | 0.57 | 0.91 | 3.17 | 2.59 | 0.32 |
| λ_{max} $r = 0$ | $r = 1$ | 10.17 | 15.55 | 12.80 | 26.45** | 11.99 | 11.21 | 20.59** | 24.69** | 4.09 | 12.02 | 7.97 | 16.96** | 23.86** | 8.84 | 17.50** | 6.46 | 19.31** | 12.47 |
| | $r = 2$ | 2.01 | 0.22 | 5.33 | 1.63 | 7.73 | 1.46 | 0.86 | 1.13 | 0.75 | 3.04 | 1.41 | 0.02 | 3.64 | 0.57 | 0.91 | 3.17 | 2.59 | 0.32 |
| Cointegration* | | No | Yes | No | Yes | No | No | Yes | Yes | No | No | No | Yes | Yes | No | Yes | No | Yes | No |

Note: * denotes rejection of the null hypothesis of no cointegration at the 5% level (based on MacKinnon-Haug-Michelis p-values).

TABLE 3. Testing for regime shifts in cointegration: Gregory-Hansen

| Country | Revenues and Expenditures | | | | (lagged) Debt and Primary Balance | | | |
|-------------|---------------------------|----------------------|-------------------|----------------------|-----------------------------------|----------------------|-------------------|----------------------|
| | ADF test | | Phillips Test | | ADF test | | Phillips Test | |
| | ADF^* stat | Estimated break date | Z_α^* stat | Estimated break date | ADF^* stat | Estimated break date | Z_α^* stat | Estimated break date |
| Australia | -3.02 | 1994 | -15.22 | 2002 | -3.52 | 1993 | -17.20 | 2003 |
| Austria | -4.86** | 2003 | -29.11 | 1996 | -3.85 | 1987 | -19.92 | 1989 |
| Belgium | -3.13 | 1976 | -16.63 | 1994 | -3.82 | 1988 | -20.38 | 1979 |
| Canada | -3.26 | 2001 | -21.95 | 2001 | -3.60 | 2000 | -18.17 | 1997 |
| Denmark | -5.36** | 1983 | -27.05 | 1983 | -5.12** | 1982 | -25.74 | 1983 |
| Finland | -4.63* | 1980 | -20.38 | 1976 | -3.13 | 1989 | -18.36 | 1994 |
| France | -3.27 | 1996 | -20.06 | 1996 | -3.97 | 1993 | -24.27 | 1992 |
| Germany | -4.92** | 2002 | -31.81 | 2001 | -3.71 | 1993 | -16.68 | 1987 |
| Greece | -4.34 | 1993 | -27.26 | 1994 | -4.60* | 1996 | -28.16 | 1986 |
| Ireland | -3.86 | 1983 | -24.11 | 1983 | -3.60 | 1997 | -23.78 | 1997 |
| Italy | -3.56 | 2000 | -20.95 | 2000 | -4.74* | 1984 | -24.27 | 1984 |
| Japan | -3.65 | 1979 | -20.85 | 1995 | -3.52 | 1976 | -15.61 | 2000 |
| Netherlands | -6.54*** | 1975 | -30.61 | 1974 | -3.04 | 1996 | -19.30 | 1999 |
| Portugal | -5.10** | 2002 | -30.67 | 2003 | -26.19*** | 1987 | -22.98 | 1996 |
| Spain | -4.39 | 1982 | -15.46 | 2003 | -3.71 | 1996 | -17.77 | 2001 |
| Sweden | -4.83* | 1981 | -19.69 | 1983 | -4.12 | 1999 | -13.52 | 2000 |
| UK | -5.12** | 1989 | -25.25 | 1988 | -3.50 | 1996 | -21.86 | 1995 |
| US | -5.81** | 2000 | -32.07 | 2000 | -3.82 | 1992 | -17.30 | 1987 |

Note: ADF^* and Z_α^* refer to the Augmented Dickey-Fuller (ADF) and to the Phillips Z_α^* tests statistics; null of no cointegration. *, ** and *** denote significance at 10, 5 and 1% levels, respectively, using the critical values from Gregory and Hansen (1996), table 1.

TABLE 4.a Estimation of long-run relationships: Stock-Watson-Shin cointegration

| Country\relation | Revenues and Expenditures | | | (lagged) Debt and Primary Balance | | |
|------------------|---------------------------|-------------|------------------|-----------------------------------|-------------|------------------|
| | β | \bar{R}^2 | C_μ | β | \bar{R}^2 | C_μ |
| Australia | -0.02 (0.25) | 0.77 | 34.25 (9.16)*** | -0.03 (0.17) | 0.30 | 3.43 (5.71) |
| Austria | 0.67 (0.08)*** | 0.92 | 14.27 (4.13)*** | -0.07 (0.08) | 0.45 | 1.36 (1.93) |
| Belgium | -0.11 (0.19) | 0.41 | 53.25 (10.33)*** | 0.06 (0.02)** | 0.92 | -3.21 (2.10) |
| Canada | 0.33 (0.10)*** | 0.76 | 26.44 (4.50)*** | 0.14 (0.02)*** | 0.95 | -10.81 (1.26)*** |
| Denmark | 0.70 (0.10)*** | 0.88 | 16.05 (6.01)** | 0.10 (0.06) | 0.87 | -0.90 (3.44) |
| Finland | 0.63 (0.05)*** | 0.90 | 20.37 (2.53)*** | -0.05 (0.02)*** | 0.93 | 6.69 (0.60)*** |
| France | 0.69 (0.07)*** | 0.95 | 13.03 (4.01)*** | 0.01 (0.01) | 0.84 | -0.34 (0.46) |
| Germany | 0.84 (0.16)*** | 0.72 | 4.94 (7.68) | -0.02 (0.02) | 0.67 | 1.78 (0.91)* |
| Greece | 0.61 (0.16)*** | 0.82 | 10.42 (7.16) | 0.12 (0.04)*** | 0.73 | -11.22 (3.97)** |
| Ireland | 0.38 (0.15)** | 0.64 | 21.52 (6.34)*** | 0.13 (0.05)** | 0.66 | -7.33 (3.44)** |
| Italy | -0.30 (0.30) | 0.50 | 52.92 (12.76)*** | 0.12 (0.02)*** | 0.89 | -10.64 (1.78)*** |
| Japan | 0.37 (0.16)** | 0.69 | 18.75 (5.68)*** | 0.02 (0.01)*** | 0.90 | 1.15 (0.45)** |
| Netherlands | 0.67 (0.05)*** | 0.93 | 14.36 (2.56)*** | 0.01 (0.03) | 0.52 | 0.38 (2.29) |
| Portugal | 0.98 (0.12)*** | 0.92 | 3.94 (5.47) | -0.13 (0.12) | 0.57 | 7.17 (6.78) |
| Spain | 0.89 (0.06)*** | 0.96 | 1.32 (2.78) | 0.04 (0.01)*** | 0.88 | -1.23 (0.57)** |
| Sweden | 0.61 (0.07)*** | 0.79 | 21.83 (4.43)*** | -0.04 (0.04) | 0.82 | 6.93 (2.63)** |
| United Kingdom | 0.56 (0.16)*** | 0.58 | 15.52 (7.51)** | 0.04 (0.04) | 0.89 | -0.99 (-0.60) |
| United States | 0.46 (0.14)*** | 0.66 | 16.22 (5.12)*** | 0.03 (0.02) | 0.81 | -1.60 (0.99) |

Note: The C_μ is the Shin (1994) LM statistic which tests for deterministic cointegration. The critical values are taken from Shin (1994), Table 1, for m=1. Standard errors in parentheses, adjusted for long-run variance. The long-run variance of the cointegrating regression residuals was estimated using the Barlett window with $l = 6 \approx INT(T^{1/2})$ as proposed by Newey and West (1987). The number of leads and lags selected was $q = 3 \approx INT(T^{1/3})$ as proposed in Stock and Watson (1993). *, ** and *** denote significance at 10, 5 and 1% levels, respectively.

TABLE 4.b Estimation of long-run relationships (robustness): Stock-Watson-Shin cointegration

| Country\relation | (lagged) Debt-mean deviation and Primary Balance | | | (twice lagged) Debt and Primary Balance | | |
|------------------|--|-------------|----------------|---|-------------|------------------|
| | β | \bar{R}^2 | C_μ | β | \bar{R}^2 | C_μ |
| Australia | 0.04 (0.05) | 0.09 | 2.18 (0.29)*** | 0.25 (0.09)** | 0.52 | -6.02 (3.18)* |
| Austria | -0.05 (0.05) | 0.21 | 0.32 (0.29) | -0.03 (0.05) | 0.26 | 0.71 (1.14) |
| Belgium | 0.07 (0.01)*** | 0.90 | 2.00 (0.25)*** | 0.08 (0.01)*** | 0.87 | -5.19 (0.94)*** |
| Canada | 0.14 (0.01)*** | 0.93 | -0.51 (0.22)** | 0.16 (0.02)*** | 0.86 | -11.65 (1.09)*** |
| Denmark | 0.04 (0.05) | 0.57 | 4.86 (0.67)*** | 0.08 (0.05)* | 0.63 | 0.17 (2.54) |
| Finland | -0.06 (0.01)*** | 0.87 | 5.07 (0.26)*** | -0.03 (0.021)* | 0.62 | 5.69 (0.72)*** |
| France | 0.01 (0.01) | 0.75 | 0.38 (0.15)** | -0.01 (0.01) | 0.55 | 0.07 (0.55) |
| Germany | 0.02 (0.01) | 0.63 | 1.34 (0.24)*** | 0.04 (0.02)* | 0.21 | -1.56 (1.20) |
| Greece | 0.06 (0.03)* | 0.39 | -0.94 (1.05) | 0.05 (0.03) | 0.20 | -5.41 (2.88)* |
| Ireland | 0.08 (0.03)** | 0.62 | 0.48 (0.73) | 0.12 (0.03)*** | 0.63 | -6.72 (2.21)*** |
| Italy | 0.12 (0.01)*** | 0.87 | -0.06 (0.40) | 0.13 (0.02)*** | 0.84 | -11.26 (1.45)*** |
| Japan | 0.01 (0.01) | 0.85 | 1.92 (0.43)*** | 0.01 (0.01) | 0.62 | 0.53 (0.85) |
| Netherlands | 0.01 (0.01) | 0.36 | 1.70 (0.27)*** | 0.04 (0.02)* | 0.26 | -0.83 (1.49) |
| Portugal | -0.01 (0.04) | 0.32 | 0.03 (0.36) | 0.02 (0.06) | 0.16 | -1.79 (2.98) |
| Spain | 0.03 (0.01)*** | 0.81 | 0.28 (0.20) | 0.04 (0.01)** | 0.53 | -1.74 (0.84)** |
| Sweden | -0.02 (0.03) | 0.69 | 3.95 (0.61)*** | 0.001 (0.04) | 0.54 | 3.37 (2.41) |
| United Kingdom | 0.03 (0.07) | 0.52 | 0.68 (0.43) | 0.11 (0.08) | 0.47 | -4.45 (3.80) |
| United States | 0.04 (0.02)** | 0.76 | -0.05 (0.26) | 0.03 (0.03) | 0.59 | -2.13 (1.46) |

Note: The C_μ is the Shin (1994) LM statistic which tests for deterministic cointegration. The critical values are taken from Shin (1994), Table 1, for m=1. Standard errors in parentheses, adjusted for long-run variance. The long-run variance of the cointegrating regression residuals was estimated using the Barlett window with $l = 6 \approx INT(T^{1/2})$ as proposed by Newey and West (1987). The number of leads and lags selected was $q = 3 \approx INT(T^{1/3})$ as proposed in Stock and Watson (1993). *, ** and *** denote significance at 10, 5 and 1% levels, respectively.

TABLE 5.a Causality tests – Revenues and Expenditures

| Country\relation | Standard Granger causality | | | | Toda-Yamamoto | | | |
|------------------|----------------------------|--------|--------------------|--------|--------------------|--------|--------------------|--------|
| | $GG \rightarrow R$ | Yes/No | $R \rightarrow GG$ | Yes/No | $GG \rightarrow R$ | Yes/No | $R \rightarrow GG$ | Yes/No |
| Australia | 3.58 | No | 5.21* | Yes | 2.88 | No | 6.86** | Yes |
| Austria | 6.01** | Yes | 0.88 | No | 4.21 | No | 2.56 | No |
| Belgium | 1.74 | No | 0.22 | No | 0.30 | No | 1.30 | No |
| Canada | 6.02** | Yes | 15.48*** | Yes | 1.50 | No | 10.25*** | Yes |
| Denmark | 1.03 | No | 13.08*** | Yes | 0.56 | No | 1.51 | No |
| Finland | 1.15 | No | 5.07* | Yes | 7.40** | Yes | 2.20 | No |
| France | 1.66 | No | 4.85* | Yes | 1.37 | No | 6.29** | Yes |
| Germany | 4.46 | No | 7.56** | Yes | 5.09* | Yes | 9.73*** | Yes |
| Greece | 0.09 | No | 5.16* | Yes | 0.16 | No | 2.58 | No |
| Ireland | 3.15 | No | 6.21** | Yes | 5.08* | Yes | 3.06 | No |
| Italy | 1.85 | No | 2.58 | No | 9.60*** | Yes | 0.43 | No |
| Japan | 0.14 | No | 1.34 | No | 0.20 | No | 2.72 | No |
| Netherlands | 7.95** | Yes | 4.52 | No | 8.78** | Yes | 8.37** | Yes |
| Portugal | 6.95** | Yes | 3.00 | No | 11.82*** | Yes | 2.11 | No |
| Spain | 4.84* | Yes | 3.48 | No | 1.23 | No | 17.22*** | Yes |
| Sweden | 3.02 | No | 5.09* | Yes | 1.69 | No | 1.13 | No |
| United Kingdom | 1.87 | No | 0.50 | No | 1.75 | No | 2.85 | No |
| United States | 9.16** | Yes | 2.48 | No | 3.49 | No | 1.22 | No |

Note: In the case of standard Granger-causality tests the null is of non-Granger causality. These tests are based on a VAR with lag equal to 2 has identified using different lag-length criteria. In the case of the Toda-Yamamoto tests the null is of non-Granger causality. These tests are based on a VAR(3) – see main text for details. Chi-squared statistics are displayed. *, ** and *** denote significance at 10, 5 and 1% levels, respectively.

TABLE 5.b Causality tests – Primary Balance and (lagged) Debt

| Country\relation | Standard Granger causality | | | | Toda-Yamamoto | | | |
|------------------|----------------------------|--------|-------------------------|--------|-------------------------|--------|-------------------------|--------|
| | $B_{t-1} \rightarrow s$ | Yes/No | $s_{t-1} \rightarrow B$ | Yes/No | $B_{t-1} \rightarrow s$ | Yes/No | $s_{t-1} \rightarrow B$ | Yes/No |
| Australia | 1.06 | No | 3.71 | No | 0.33 | No | 0.52 | No |
| Austria | 2.09 | No | 0.33 | No | 4.16* | Yes | 0.19 | No |
| Belgium | 14.99*** | Yes | 1.78 | No | 15.48*** | Yes | 0.45 | No |
| Canada | 13.35*** | Yes | 2.53 | No | 6.42** | Yes | 0.96 | No |
| Denmark | 5.29* | Yes | 1.69 | No | 8.81** | Yes | 2.83 | No |
| Finland | 4.31 | No | 9.14** | Yes | 4.84* | Yes | 3.95 | No |
| France | 1.09 | No | 3.50 | No | 9.77*** | Yes | 2.86 | No |
| Germany | 2.68 | No | 1.53 | No | 2.12 | No | 1.01 | No |
| Greece | 12.98*** | Yes | 3.05 | No | 9.94*** | Yes | 0.16 | No |
| Ireland | 6.37** | Yes | 2.27 | No | 0.62 | No | 9.25*** | Yes |
| Italy | 17.04*** | Yes | 4.58* | Yes | 5.73* | Yes | 3.91 | No |
| Japan | 1.15 | No | 6.79** | Yes | 6.04** | Yes | 12.59*** | Yes |
| Netherlands | 0.65 | No | 2.10 | No | 13.23*** | Yes | 5.35* | Yes |
| Portugal | 1.60 | No | 1.87 | No | 1.95 | No | 2.06 | No |
| Spain | 5.90* | Yes | 1.77 | No | 1.31 | No | 7.01** | Yes |
| Sweden | 4.70* | Yes | 2.53 | No | 12.2*** | Yes | 3.32 | No |
| United Kingdom | 3.77 | No | 12.56*** | Yes | 2.09 | No | 16.09*** | Yes |
| United States | 1.24 | No | 2.02 | No | 10.44*** | Yes | 2.93 | No |

Note: see Table 5.a.

Table 6: Pedroni (2004) panel cointegration tests

| | relation | Revenues and Expenditures | | (lagged) Debt and Primary Balance | | (lagged) Debt-mean deviation and Primary Balance | |
|-------------------|--------------|---------------------------|-----------|-----------------------------------|--------|--|--------|
| | | No trend | Trend | No trend | Trend | No trend | Trend |
| | | Within-dimension | Panel V | 3.42 | 1.15 | 5.29 | 2.85 |
| | Panel ρ | -2.93* | -1.82* | -3.98* | -3.48* | -3.78* | -3.4* |
| | Panel PP | -2.61* | -2.87* | -3.07* | -3.83* | -2.92* | -3.74* |
| | Panel ADF | -2.51* | -2.75* | -1.72 | -1.91 | -1.72 | -1.92 |
| Between-dimension | Group ρ | -1.80 | -0.55 | -2.91* | -1.58 | -2.89* | -1.57 |
| | Group PP | -2.50* | -2.48* | -2.86* | -3.15* | -2.81* | -3.12* |
| | Group ADF | -2.51* | -3.17* | -1.81 | -1.80 | -1.93 | -1.83 |

Notes: The null hypothesis is that there is no cointegration. Under the null all the statistics are distributed as standard Normal distributions. An asterisk (*) indicates rejection at the 10% level or better.

Table 7: Panel estimates of the cointegrating relationship: FMOLS in Pedroni (2000, 2001)

| Country\relation | Revenues and Expenditures | | (lagged) Debt and Primary Balance | | (lagged) Debt-mean deviation and Primary Balance | |
|------------------|---------------------------|---------------|-----------------------------------|---------------|--|---------------|
| | FMOLS | | FMOLS | | FMOLS | |
| | β | <i>t-stat</i> | β | <i>t-stat</i> | β | <i>t-stat</i> |
| Australia | 0.60** | 2.47 | 0.17*** | 4.59 | 0.05 | 0.91 |
| Austria | 0.69*** | 14.98 | 0.03* | 1.98 | 0.08*** | 5.14 |
| Belgium | 0.30* | 1.80 | 0.11*** | 5.16 | 0.15** | 3.58 |
| Canada | 0.35** | 2.34 | 0.02 | 0.55 | 0.09** | 2.22 |
| Denmark | 0.63*** | 4.78 | -0.06 | -1.51 | -0.01 | -0.24 |
| Finland | 0.68*** | 10.01 | 0.05 | 1.25 | 0.01 | 0.95 |
| France | 0.69*** | 14.25 | 0.00 | 0.12 | 0.06** | 2.72 |
| Germany | 0.53*** | 6.50 | 0.08** | 3.66 | 0.11** | 3.68 |
| Greece | 0.67*** | 6.54 | 0.04 | 0.69 | 0.11** | 3.10 |
| Ireland | 0.40** | 2.65 | 0.07** | 2.57 | 0.07*** | 4.94 |
| Italy | -0.14 | -0.45 | 0.01 | 0.37 | -0.04* | -2.19 |
| Japan | 0.52*** | 4.28 | -0.03 | -3.15 | 0.06 | 1.45 |
| Netherlands | 0.65*** | 16.06 | 0.05* | 1.59 | 0.04* | 1.63 |
| Portugal | 0.84** | 11.87 | 0.01 | 0.48 | -0.00 | -0.07 |
| Spain | 0.76*** | 9.85 | 0.01 | 0.17 | 0.09 | 1.23 |
| Sweden | 0.50*** | 5.52 | -0.01 | -0.08 | 0.02 | 0.56 |
| United Kingdom | 0.34** | 3.36 | 0.07* | 1.83 | -0.01 | -0.17 |
| United States | 0.24 | 1.52 | -0.00 | -0.07 | 0.04 | 1.23 |
| Panel | 0.51*** | 27.93 | 0.03*** | 4.76 | 0.05*** | 7.22 |

Notes: The regression equations correspond to (10) and (11) as discussed in the main text. *t-stat.* stands for t-statistics. *, **, *** denote significance at 10, 5 and 1% levels.

Table 8: Panel Long-run Causality (Pedroni)

| Panel A: | $\lambda_1 : GG_{it} \rightarrow R_{it}$ | | | $\lambda_2 : R_{it} \rightarrow GG_{it}$ | | | $-\lambda_1/\lambda_2$ median |
|---------------|---|--------|---------|---|--------|---------|----------------------------------|
| | Estimate | Test | p-value | Estimate | Test | p-value | |
| Group mean | 0.14 | 0.24 | 0.60 | 0.34 | 2.31 | 0.01 | -0.55 |
| Lamba-Pearson | | 47.35 | 0.10 | | 145.39 | 0.00 | 0.42 |
| Panel B: | $\lambda_1 : B_{it-1} \rightarrow s_{it}$ | | | $\lambda_2 : s_{it} \rightarrow B_{it+1}$ | | | $-\lambda_1/\lambda_2$ median |
| | Estimate | Test | p-value | Estimate | Test | p-value | |
| Group mean | 0.66 | 1.56 | 0.06 | -0.42 | -1.18 | 0.12 | 0.97 |
| Lamba-Pearson | | 104.58 | 0.00 | | 90.64 | 0.00 | 0.41 |

Note: Panel A explores the causality relation between government revenues and expenditures. Panel B explores the causality relation between (lagged) public debt and primary balance. For the full sample considered above there are two rows, one for the group mean based test, and one for the lambda-Pearson based test. Columns 2-4 report these for tests based on the parameter λ_1 , which reflected the presence or absence of long-run causality running from expenditures to revenues. The second column reports the panel point estimate, which exists only for the group mean, not for the lambda-Pearson. The third column reports the corresponding panel test statistics and the fourth column reports the p-value for outcome of the panel test statistic. The next three columns repeat this same pattern for analogous tests based on the parameter λ_2 , which reflects the presence or absence of long-run causality running from revenues to expenditures. Finally, the last column reports the group median point estimate of the sign ratio in the first row, with the simulated standard error reported in parenthesis in the second row.

Table 9: Pooled OLS estimation of (10) with dummy for financial crises

| Equation (10) | (1) | (2) | (3) | (4) |
|-----------------------------|----------------------|----------------------|----------------------|----------------------|
| Lagged debt | 0.011** (0.005) | | 0.014*** (0.005) | |
| Lagged debt mean dev. | | 0.031*** (0.007) | | 0.038*** (0.007) |
| FC dummy (additive) | -3.145*** (0.621) | -3.277*** (0.653) | | |
| FC interaction (times debt) | | | -0.049*** (0.007) | -0.145*** (0.023) |
| Constant | -0.463 (0.299) | 0.120 (0.128) | -0.711** (0.307) | -0.021 (0.129) |
| Observations | 552 | 552 | 543 | 543 |
| R-squared | 0.071 | 0.116 | 0.087 | 0.131 |

Note: Pooled OLS estimation of Eq.(10) using as regressor either lagged debt or lagged debt-mean deviation together with additive or multiplicative dummy for financial crises (FC). Heteroskedastic-robust standard errors are in parenthesis. . *, **, *** denote significance at 10, 5 and 1% levels.