Applied Economics

ISSN: 0003-6846 (Print) 1466-4283 (Online) Journal homepage: https://www.tandfonline.com/loi/raec20

The elusive character of fiscal sustainability

António Afonso & João Tovar Jalles

APPLIED ECONOMICS, 2016
VOL. 48, NO. 28, 2651–2664
http://dx.doi.org/10.1080/00036846.2015.1128074

The elusive character of fiscal sustainability

António Afonso[a,b] and João Tovar Jalles[c]

[a]Department of Economics, ISEG/ULisbon – University of Lisbon, Lisbon, Portugal; [b]UECE – Research Unit on Complexity and Economics, Lisbon, Portugal; [c]Center for Globalization and Governance, Nova School of Business and Economics, Lisboa, Portugal

ABSTRACT
We assess the sustainability of public finances in OECD countries using unit root and cointegration analysis, controlling for endogenous breaks. Results deem fiscal sustainability as rather elusive since we find lack of cointegration – absence of sustainability – between government revenues and expenditures (except for Austria, Canada, France, Germany, Japan, Netherlands, Sweden and UK); improvements of the primary balance after worsening debt ratios only for Australia, Belgium, Germany, Ireland, Netherlands and the UK; and Granger causality from government debt to primary balances for 12 countries (suggesting Ricardian regimes).

KEYWORDS
Debt; primary balance; fiscal regimes; stationarity; breaks; causality; DOLS

JEL CLASSIFICATION
C33; E62; H62; H63

I. Introduction

The importance of sustainable public finances has received increasing attention particularly in the context and following the 2008–2009 economic and financial crises. 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 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. Given the detrimental impact of persistent deficits, practices on debt sustainability and appropriate fiscal policies are extremely important.

The existing work on fiscal sustainability has looked at the OECD cases, notably Feve and Henin (2000); Martin (2000); Bravo and Silvestre (2002); Hatemi-J (2002); Afonso (2005); Mendoza and Ostry (2008); Arghyrou and Luintel (2007); Afonso, (2008); Afonso and Rault (2010); Camarero, Carrion-i-Silvestre, and Tamarit (2014); Bajo-Rubio, Díaz-Roldán, and Esteve (2014). 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.

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.

CONTACT António Afonso aafonso@iseg.utl.pt

[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).

Additionally, another strand of the literature, as for example, Davig, Leeper, and Walker (2011), alternatively 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, Leeper, and Walker 2011). As in other related studies, this is usually done for the United States. 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 of those studies can indirectly be read as providing also information on the sustainability of public finances. In fact, the assessment via fiscal reaction functions is more a view on the activism of fiscal policy. But in practice, even in the case of passive fiscal behaviour, sustainability can be at risk, depending on the magnitude of the fiscal response, as can be uncovered via a cointegration evaluation.

The main purpose of this article 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 add to the literature by 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 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, using annual data over the period 1970–2010.

Therefore, our contribution is twofold: (i) we take a longer time span and make use of uniform and comparable data for 18 OECD countries; (ii) we explore three different channels to evaluate fiscal sustainability as put forward in theoretical terms, by looking at the first-differenced debt ratios, the relationship between government revenues, and expenditures and the relationship between (lagged) public debt and primary balances.

Our results show that the first-differenced debt series for most countries is nonstationarity, suggesting that the solvency condition would not be satisfied. Moreover, evidence suggests the existence of one cointegrating relationship between revenues and expenditures in only six countries. 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, explaining the steady building up of fiscal imbalances. 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, again pointing to elusive fiscal sustainability, at least until the 2009–2010 economic and financial crisis. 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.

The structure of the article is as follows. Section II discusses the underlying theoretical framework which serves as the basis for the empirical strategy. Section III looks at some data stylized facts. Section IV presents the methodology and discusses our main findings. The last section concludes.

II. 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

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

where G is the government expenditures, excluding interest payments; R is the government revenues; B is the public debt; $r$is the real interest rate.

Rewriting Equation 1 for the subsequent periods and recursively solving leads to the intertemporal budget constraint:

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

When the second term on the right hand side of Equation 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

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

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

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

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 Equation 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.

From the PVBC, Equation 4, it is possible to present analytically two complementary definitions of sustainability that set the background for empirical testing[2]:

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

$$B_{t-1} = \sum_{s=0}^{\infty} \frac{1}{(1 + r)^{s+1}} (R_{t+s} - E_{t+s}) \tag{5}$$

2. The present value of public debt must approach zero in infinity:

$$\lim_{s\to\infty} \frac{B_{t+s}}{(1 + r)^{s+1}} = 0 \tag{6}$$

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) 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 mentioned in the context of the Fiscal Theory of the Price Level (see, e.g. 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).

It is also possible to assess fiscal policy sustainability through cointegration between government revenues and expenditures. The implicit hypothesis concerning the real interest rate, with mean r, is also stationarity. Using $E_t = G_t + (r_t - r)B_{t-1}$ and $GG_t = G_t + r_t B_{t-1}$, with the no-Ponzi game condition, GG_t and R_t must be cointegrated 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 cointegration is rejected, the residual u_t must be stationary and should not display a unit root. Hakkio and Rush (1991) demonstrate that if GG

[2]Hamilton and Flavin (1986) first used these procedures.

and R are nonstationary variables in levels (by levels we mean in terms of percentage in GDP), the condition $0 < \beta < 1$ 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 $\beta = 1$ in order for the trajectory of the debt to GDP not to diverge in an infinite horizon.

III. Data overview and stylized facts

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 United States, primary balance (% of GDP) data comes from the OECD database. Government Debt (% of GDP) series are retrieved from the IMF's International Financial Statistics.

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.

IV. Econometric methodology and empirical results

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

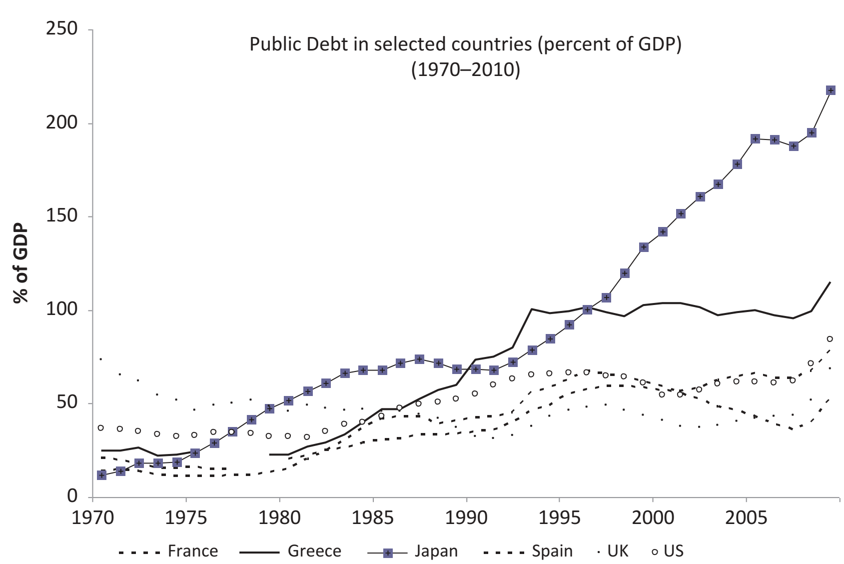

Figure 1. Public debt series: 1970–2010 (selected countries).

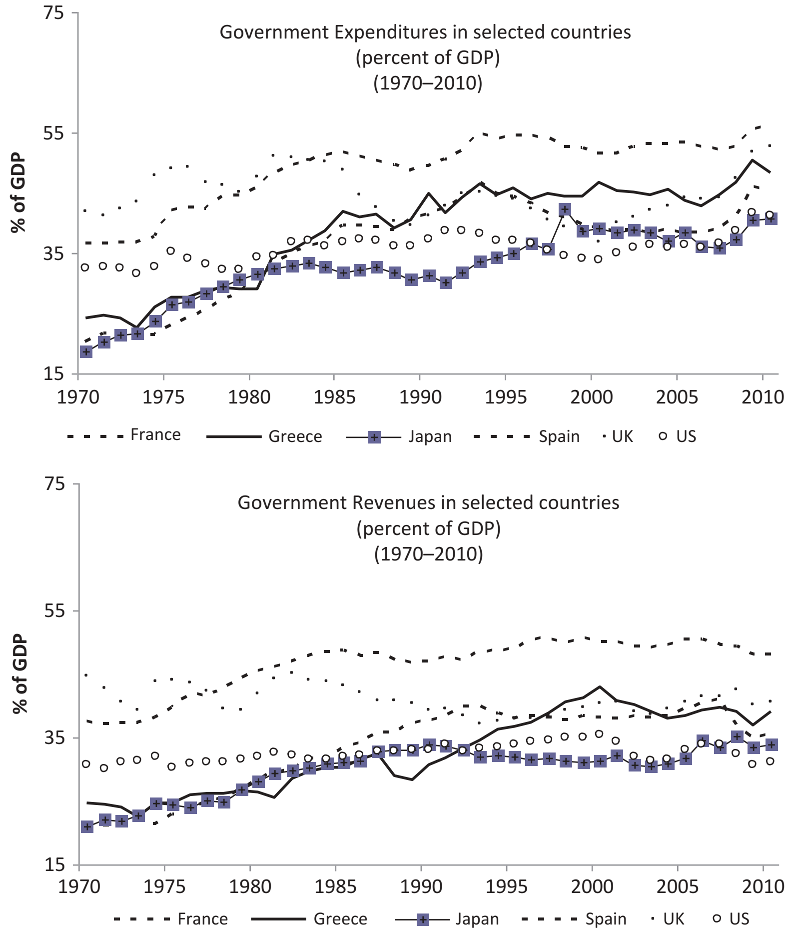

Figure 2. Total government expenditures and revenues: 1970–2010 (selected countries).

Dickey Fuller (ADF) and Phillips-Perron (PP) unit root tests, we also conduct the four tests (M-tests) proposed by Ng and Perron (2001) (NP) based on modified information criteria (MIC): the modified PP test MZ_{α}; the modified Sargan-Bhargava test (MSB); the modified point optimal test MP_T; and the modified PP 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 and 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). We complement with the modified ADF test proposed by Vogelsang and Perron (1998) (VP) also allowing for one endogenously determined break. These test the null of unit root against the break stationary alternative hypothesis.

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. 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. This newer test has better properties in terms of size and power.[3]

The fact that our series are in ratio to GDP does not rule them out being integrated processes. 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 shows the stationarity test results for the first difference of the debt ratio.

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 nonstationarity can be rejected. The Ng and Perron (2001) tests add Finland, Netherlands and Portugal to the 'rejection-of-the-null' set of countries.[4]

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 towards the nonrejection of a unit root. Therefore, in Table 1, 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 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

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

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

[3]We thank Pierre Perron and Tomoyoshi Yabu for providing their GAUSS code.
[4]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 (2014) make use of the longer time series debt data going back to the late nineteenth century to study fiscal sustainability, and such considerations do not apply.

1974, 1982–83 and 2003, corresponding to the 'Carnation Revolution', the IMF program intervention and one severe recession, respectively. 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 2 – we find similar results, with the nonrejection of the null of unit

Table 2. Unit Root Tests and Structural Breaks 1970-2010.

	ADF	PP	NP				ZA	VP(AO)	VP(IO)	PY2009
			MZa	MZt	MSB	MPT				
Countries/Series	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Revenues										
Australia	−1.77	−1.77	−5.11	−1.33	0.26	16.7	2004	1981	1980	2003***
Austria	−1.48	−1.39	−1.07	−0.50	0.47	48.18	1976	1977	1974**	1982***
Belgium	−2.08	−2.07	−3.14	−1.13	0.35	26.22	1977	1978	1973	1980
Canada	−0.87	−0.64	−5.74	−1.46	0.25	15.43	1990	1983	1978	1997***
Denmark	−0.97	−1.40	−4.43	−1.24	0.28	18.61	1983	1980	1981	1984
Finland	−1.61	−1.22	−2.88	−1.01	0.35	26.76	1990	1987	1973**	1990
France	−0.51	−0.63	−0.75	−0.34	0.46	48.77	1979	1980	1973**	1985***
Germany	−2.77	−2.80	−4.32	−1.30	0.30	19.64	1992	1988	1989	2000
Greece	−1.77	−1.92	−6.17	−167	0.27	14.70	1997	1991	1988	1996
Ireland	−1.47	−1.43	−3.17	−1.19	0.37	27.25	1985	1977	1978	1984*
Italy	−0.90	−0.95	−3.03	−1.07	0.35	26.23	1992	1984	1974	1991
Japan	−1.80	−1.80	−2.95	−1.16	0.39	29.41	1979	1981	1977	1991
Netherlands	−2.69	−2.69	−1.96	−0.95	0.48	44.07	1994	1994	1992**	1981**
Portugal	−0.66	−1.82	−3.43	1.07	0.31	22.46	1995	1993	1973	1991***
Spain	0.07	0.28	−0.53	−0.23	0.44	46.63	1994	1982	1977	1993***
Sweden	−1.73	−1.68	−2.40	−0.88	0.36	29.62	1976	1978	1974	1986**
United Kingdom	−3.09	−2.53	−18.85**	−3.04**	0.16**	4.98**	1992	1988	1988	1989
United States	−2.70	−1.40	−19.03**	−2.94**	0.15**	5.63*	2002**	1988**	1985	2001***
Expenditures										
Australia	−4.50***	−3.13	−8.14	−1.98	0.24	11.27	1987	1977	1973	1985***
Austria	−1.96	−1.96	−2.15	−0.96	0.44	38.50	1976	1977	1973**	1982
Belgium	−2.31	−2.31	−2.87	−1.19	0.41	31.59	1984	1979	1992	1983***
Canada	−1.90	−1.79	−5.49	−1.65	0.30	16.55	1982	2001	1997	1995
Denmark	−2.03	−1.93	−6.51	−1.80	0.27	13.99	1980	1978	1972	1980
Finland	−2.42	−1.74	−11.41	−2.37	0.20	8.03	1990	1987	1988**	1997*
France	−2.17	−1.73	−5.19	−1.58	0.30	17.42	1981	1982	1973	1985
Germany	−3.07	−3.07	−5.62	−1.67	0.29	16.20	1991	1976	1972**	1980
Greece	−1.76	−1.64	−5.70	−1.64	0.28	15.86	1981	1982	1979	1983
Ireland	−2.10	−1.65	−7.11	−1.84	0.25	12.86	1980	1983	1984	1988
Italy	−2.10	−1.65	−7.11	−1.84	0.25	12.86	1980	1983	1984	1988
Japan	−2.74	−2.66	6.31	−1.76	0.27	14.42	1984	1995	1991	1983***
Netherlands	−2.01	−1.97	−4.24	−1.45	0.34	21.39	1978	1988**	1992	1983
Portugal	−2.62	−1.87	−4.37	−1.45	0.33	20.65	1978	1982	1972	1981
Spain	−1.41	−1.40	−3.50	−1.31	0.37	25.80	1980	1982	1973	1995**
Sweden	−2.17	−1.87	−6.70	−1.78	0.26	13.62	1981	1978	1972	1981**
United Kingdom	−2.49	−1.48	−16.09*	−2.70*	0.16*	6.42*	2002	1989	1983	1985
United States	−2.56	−1.93	−15.83*	−2.74*	0.17*	6.18*	1997	1983	1978	1996
Primary Balance										
Australia	−3.87**	−2.46	−26.94***	−3.60***	0.13***	3.75***	1986	1989	1984**	1985
Austria	−3.70**	−2.54	−12.57	−2.43	0.19	7.63	1979	1972**	1994	1996
Belgium	−1.66	−1.64	−5.93	−1.58	0.26	15.16	2003	1979	1980	1984
Canada	−1.92	−1.75	−9.53	−2.06	0.21	10.02	1996	1990	1992	1995***
Denmark	−3.52*	−1.74	−27.40***	−3.63***	0.13***	3.69***	1983	1984	1981**	1983
Finland	−3.03	−2.48	−20.70**	−3.17**	0.15**	4.68***	2000	1986**	1987	1997
France	−2.46	−2.63	−10.84	−2.16	0.19	9.16	1997	2006	1992	1997
Germany	−4.38***	−4.15	−17.36**	−2.93**	0.16**	5.32*	1977	1993	1994	1981***
Greece	−2.06	−2.00	−7.30	−1.87	0.25	12.52	1994	2005	1989	1993
Ireland	3.82	2.60	−15.50***	−27.30***	0.01***	0.15***	2004	2007	1985	2003***
Italy	−1.97	−2..04	−6.37	−1.69	0.26	14.27	1992	1989	1989**	1991
Japan	−2.24	−2.15	−8.32	−2.02	0.24	10.99	1993	1996	1990	1993***
Netherlands	−3.30*	−3.40*	−15.38*	−2.59	0.16**	6.94	1996	2006**	1994**	1995***
Portugal	−2.45	−2.42	−9.06	−2.04	0.22	10.04	1982	2006	1980	1982
Spain	−2.88	−1.65	−22.36**	−3..18**	0.14**	4.98**	1986	1997	1984	2003
Sweden	−3.94**	−2.47	−24.49***	−3.49***	0.14**	3.74***	1991	1987**	1988**	1991
United Kingdom	−3.15	−2.43	−18.67**	−2.98**	0.16**	5.28**	1998	2005	1999	1997
United States	−2.56	−1.24	−22.19**	−3.14**	0.14**	5.18**	1997**	2004**	1999	1996***

Note: see Table 1.a for details. In this table, instead of using first-differenced variables, these are tested in levels.

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 United States for the ZA and VP tests in primary balance case).

It is important to realize that fiscal solutions require a change to current fiscal policies, so a structural break in the time series must occur before policy becomes unsustainable. In the present context, the interpretation of these type of tests changes since the time series test is then more a test of when a structural adjustment occurs (perhaps more rightly so that a 'test for sustainability'). Moreover, fiscal policy in many countries may still be sustainable if other economic factors would be taken into account (which is not the case in our analysis). For example, a government may start a costly unsustainable reform that raises potential economic growth, and such a deficit would seem to be unsustainably in a simple unit root test, ceteris paribus, but it may be self-financing due to future tax revenues. We do not have a general framework that could allow for such possibility.

Cointegration and stability

Consider the following cointegrating relationship regression, as from Section II:

$$R_t = \alpha + \beta GG_t + u_t \quad (7)$$

R_t denotes government revenues and GG_t government expenditures. ε_t and u_t are iid disturbance terms satisfying standard assumptions. In addition, an alternative can be tested via a fiscal reaction function (Bohn 1998),

$$s_t = \alpha + \gamma B_{t-1} + \varepsilon_t \quad (8)$$

where B_t is the government debt and s_t is the government primary balance.

Given the nonstationarity of each individual time series, 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, 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 the equations above: 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.[5] 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 cointegration test. 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.

Stationarity aside, we now address the cointegration issue, given by Equations 7 and 8, by analysing the relationship between revenues and expenditures and between the primary balance and (lagged) debt. Table 3 left-hand-side presents the results for the Johansen 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 3 right-hand-side (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.[6] Moreover, in these cases, results from the Hansen-stability test did not reject

[5]In a MD regime, we would observe that an increase in previous period debt would lead to a larger primary balancel 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.

[6]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.

Table 3. Johansen Cointegration Tests:

Test/null hypothesis vs. alternative	Revenues and Expenditures λ_{trace}: $r = 0_{vs.}$ $r \geq 1$	Revenues and Expenditures λ_{max}: $r = 0_{vs.}$ $r \geq 1$	(lagged) Debt and Primary Balance λ_{trace}: $r = 0_{vs.}$ $r \geq 1$	(lagged) Debt and Primary Balance λ_{max}: $r = 0_{vs.}$ $r \geq 1$
Australia	20.33*	11.66	12.19	10.17
Austria	28.84**	19.05**	21.16**	15.55
Belgium	14.41	12.55	18.14	12.8
Canada	14.53	8.62	28.08**	26.45**
Denmark	23.40**	18.40**	19.73	11.99
Finland	19.12	11.86	12.68	11.21
France	18.39	12.51	21.45**	20.59**
Germany	30.89**	21.23**	25.83**	24.69**
Greece	13.31	8.88	4.84	4.09
Ireland	14.35	9.33	15.07	12.02
Italy	13.79	8.73	9.36	7.97
Japan	22.95**	19.96**	16.98**	16.96**
Netherlands	21.04**	18.09**	27.51**	23.86**
Portugal	18.56	11.46	9.41	8.84
Spain	19.55	10.7	18.42**	17.50**
Sweden	19.85	14.3	9.63	6.46
UK	19.2	13.39	21.91**	19.31**
US	18.78	14.64	12.8	12.47

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

the null hypothesis that the series are cointegrated at conventional levels (with p values larger than 20%).[7] Overall, test results allow the rejection of the cointegration hypothesis for the majority of the countries in both relationships under scrutiny.

However, and as in the case of unit roots, a test for cointegration 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 cointegration if there is a cointegration 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 cointegrated 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 cointegration relationships is then studied.[8]

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 4 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)[9]; 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.

In order to estimate the parameters γ and β in Equations 7 and 8, we resort to the method of Dynamic Ordinary Least Squares (DOLS) of Stock and Watson (1993). This method has the advantage of providing a robust 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.

It is important 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 estimations. 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. We apply this approach to test the stability of the cointegrating relation.

[7]Results are available from the authors upon request.
[8]We thank Bruce Hansen for making the GAUSS routine available.
[9]Our results, similarly to those reported in the literature, do not consider additional sources of government revenues: for instance, seignorage and privatization revenues. Additionally, government assets could be taken into account.

Table 4. Testing for regime shifts in cointegration: Gregory-Hansen.

relation	Revenues and Expenditures				(lagged) Debt and Primary Balance			
	ADF test		Phillips Test		ADF test		Phillips Test	
Country	ADF^*stat	Estimated break date	$Z_a{}^*$ stat	Estimated break date	ADF^*stat	Estimated break date	$Z_a{}^*$ 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 $Z_a{}^*$refer to the Augmented Dickey-Fuller (ADF) and to the Phillips $Z_a{}^*$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.

We are now in position to estimate the parameters γ and β in Equations 7 and 8. 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 coefficients and the statistic C_μ, an LM statistic from the DOLS residuals, which tests for deterministic cointegration (i.e. when no trend is present in the regression), appear in Table 5. Two main results can be obtained from Table 5. 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 γ in 8 out of 18 cases. Therefore, a preliminary conclusion can emerge: we cannot say that fiscal policy has been sustainable for half the countries in our sample. One should note, however, that in some cases more than one break could be identified (see, e.g. Afonso and Rault 2007). That said, there is no single universally accepted method in the literature to test for cointegration while allowing for endogenously determined structural breaks in the underlying long-run relationship.[10]

Causality

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. 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 have also employed Toda and Yamamoto (1995) approach for Granger causality.[11]

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

[10]Some examples include Westerlund (2006) and Hatemi-J (2008), for multiple and two-breaks in the cointegrating relationship, respectively.

[11]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(l) model (where l is the lag length in the system).

Table 5. Estimation of long-run relationships: Stock–Watson–Shin cointegration.

	Revenues and Expenditures			(lagged) Debt and Primary Balance		
Country\relation	β	$\bar{R}^2$	C_μ	γ	$\bar{R}^2$	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.

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).[12]
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.

An additional exercise is to explore the causality direction between total government revenues and expenditures. Table 6 presents our results for the standard Granger causality test. In general, 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 six 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.[13]

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 5 can be found in both of them. Hence, Table 7 presents the results from the standard Granger causality tests together with the Toda-Yamamoto version, similarly to Table 6. 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

[12]See also Payne (2003) for a survey on the international evidence concerning the tax and spend debate.
[13]As far as the Toda and Yamamoto (1995) results are concerned, they are not necessarily better, notably due to the limited time sample and are, therefore, not reported here for reasons of parsimony

Table 6. Causality tests – Revenues and Expenditures.

	Standard Granger causality			
Country\relation	$GG \rightarrow R$	Yes/No	$R \rightarrow GG$	Yes/No
Australia	3.58	No	5.21*	Yes
Austria	6.01**	Yes	0.88	No
Belgium	1.74	No	0.22	No
Canada	6.02**	Yes	15.48***	Yes
Denmark	1.03	No	13.08***	Yes
Finland	1.15	No	5.07*	Yes
France	1.66	No	4.85*	Yes
Germany	4.46	No	7.56**	Yes
Greece	0.09	No	5.16*	Yes
Ireland	3.15	No	6.21**	Yes
Italy	1.85	No	2.58	No
Japan	0.14	No	1.34	No
Netherlands	7.95**	Yes	4.52	No
Portugal	6.95**	Yes	3.00	No
Spain	4.84*	Yes	3.48	No
Sweden	3.02	No	5.09*	Yes
United Kingdom	1.87	No	0.50	No
United States	9.16**	Yes	2.48	No

Note: The null is of non-Granger causality. This test is based on a VAR with lag equal to 2 has identified using different lag-length criteria (available from the authors upon request). Chi-squared statistics are displayed. *, ** and *** denote significance at 10, 5 and 1% levels, respectively.

Table 7. Causality tests – Primary Balance and (lagged) Debt.

	Standard Granger causality			
Country\relation	$B_{t-1} \rightarrow s$	Yes/No	$s_{t-1} \rightarrow B$	Yes/No
Australia	1.06	No	3.71	No
Austria	2.09	No	0.33	No
Belgium	14.99***	Yes	1.78	No
Canada	13.35***	Yes	2.53	No
Denmark	5.29*	Yes	1.69	No
Finland	4.31	No	9.14**	Yes
France	1.09	No	3.50	No
Germany	2.68	No	1.53	No
Greece	12.98***	Yes	3.05	No
Ireland	6.37**	Yes	2.27	No
Italy	17.04***	Yes	4.58*	Yes
Japan	1.15	No	6.79**	Yes
Netherlands	0.65	No	2.10	No
Portugal	1.60	No	1.87	No
Spain	5.90*	Yes	1.77	No
Sweden	4.70*	Yes	2.53	No
United Kingdom	3.77	No	12.56***	Yes
United States	1.24	No	2.02	No

Note: see Table 5.a.

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. 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 FD (or non-Ricardian) regime.

V. Conclusion

In this article, 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 time-series techniques. Given the volume of empirical findings, which in some instances could present less clear cut evidence, our analysis is a good example of how difficult is to reach conclusions using alternative econometric techniques when working with data of limited sample size.

Our main results point to the nonstationarity 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, hinting strongly towards the elusive character of fiscal sustainability up to the 2009–2010 economic and financial crisis. 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, our evidence suggests the existence of one cointegrating relationship between revenues and expenditures in only six countries. 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 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. Additionally, in only six 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. 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.

Acknowledgments

The authors are grateful for comments from an anonymous referee, and from participaants at ECB, and ISEG seminars, and at the INFER Annual Conference (Orléans), 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 author's employers.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

UECE is supported by the Portuguese Foundation for Science and Technology through the project PEst-OE/EGE/UI0436/2011.

References

Afonso, A. 2005. "Fiscal Sustainability: The Unpleasant European Case." *FinanzArchiv* 61 (1): 19–44. doi:10.1628/0015221053722532.

Afonso, A. 2008. "Ricardian Fiscal Regimes in the European Union." *Empirica* 35 (3): 313–334. doi:10.1007/s10663-008-9066-3.

Afonso, A., and J. Jalles. 2014. "A Longer-Run Perspective on Fiscal Sustainability." *Empirica* 41: 821–847. doi:10.1007/s10663-013-9240-0.

Afonso, A., and C. Rault. 2007. "Should We Care for Structural Breaks When Assessing Fiscal Sustainability?" *Economics Bulletin* 3 (63): 1–19.

Afonso, A., and C. Rault. 2010. "What Do We Really Know about Fiscal Sustainability in the EU? A Panel Data Diagnostic." *Review of World Economics* 145 (4): 731–755. doi:10.1007/s10290-009-0034-1.

Arghyrou, M. G., and K. B. Luintel. 2007. "Government Solvency: Revisiting Some EMU Countries." *Journal of Macroeconomics* 29: 387–410. doi:10.1016/j.jmacro.2005.09.001.

Bajo-Rubio, O., C. Díaz-Roldán, and V. Esteve. 2014. "Deficit Sustainability, and Monetary versus Fiscal Dominance: The Case of Spain, 1850–2000." *Journal of Policy Modeling* 36 (5): 924–937. doi:10.1016/j.jpolmod.2014.07.004.

Barro, R. 1979. "On the Determination of the Public Debt." *Journal of Political Economy* 87: 940–971. doi:10.1086/jpe.1979.87.issue-5.

Bohn, H. 1998. "The Behavior of U.S. Public Debt and Deficits." *The Quarterly Journal of Economics* 113: 949–963. doi:10.1162/003355398555793.

Bohn, H. 2007. "Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint?" *Journal of Monetary Economics* 54 (7): 1837–1847. doi:10.1016/j.jmoneco.2006.12.012.

Bravo, A. B. S., and A. L. Silvestre. 2002. "Intertemporal Sustainability of Fiscal Policies: Some Tests for European Countries." *European Journal of Political Economy* 18: 517–528. doi:10.1016/S0176-2680(02)00103-9.

Camarero, M., J. Carrion-i-Silvestre, and C. Tamarit. 2014. "The Relationship between Debt Level and Fiscal Sustainability in OECD Countries." *Economic Inquiry* 53 (1): 129–149.

Davig, T., and E. Leeper. 2011. "Monetary–Fiscal Policy Interactions and Fiscal Stimulus." *European Economic Review* 55 (2): 211–227. doi:10.1016/j.euroecorev.2010.04.004.

Davig, T., E. Leeper, and T. Walker. 2011. "Inflation and the Fiscal Limit." *European Economic Review* 55 (1): 31–47. doi:10.1016/j.euroecorev.2010.11.005.

Feve, P., and P.-Y. Henin. 2000. "Assessing Effective Sustainability of Fiscal Policy within the G-7." *Oxford Bulletin of Economics and Statistics* 62: 175–195. doi:10.1111/obes.2000.62.issue-2.

Friedman, M. 1978. "The Limitations of Tax Limitation." *Policy Review* 5: 7–14.

Gregory, A., and B. Hansen. 1996. "Residual-Based Tests for Cointegration in Models with Regime Shifts." *Journal of Econometrics* 70: 99–126. doi:10.1016/0304-4076(69)41685-7.

Hakkio, G., and M. Rush. 1991. "Is the Budget Deficit 'Too Large?'" *Economic Inquiry* 29 (3): 429–445.

Hamilton, J. D., and M. A. Flavin. 1986. "On the Limitations of Government Borrowing: A Framework for Empirical Testing." *American Economic Review* 76: 808–819.

Hansen, H., and S. Johansen. 1993. "Recursive Estimation in Cointegrated VAR-Models." *Mimeo*, Institute of Mathematical Statistics, University of Copenhagen.

Hatemi-J, A. 2002. "Fiscal Policy in Sweden: Effects of EMU Criteria Convergence." *Economic Modelling* 19: 121–136. doi:10.1016/S0264-9993(00)00066-3.

Hatemi-J, A. 2008. "Tests for Cointegration with Two Unknown Regime Shifts with an Application to Financial Market Integration." *Empirical Economics* 35 (3): 497–505. doi:10.1007/s00181-007-0175-9.

Hoover, K. D., and S. M. Sheffrin. 1992. "Causation, Spending and Taxes: Sand in the Sandbox or Tax Collector for the Welfare State." *American Economic Review* 82 (1): 225–248.

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.

Martin, G. M. 2000. "U.S. Deficit Sustainability: A New Approach Based on Multiple Endogenous Breaks." *Journal of Applied Econometrics* 15 (1): 83–105. doi:10.1002/(ISSN)1099-1255.

McCallum, B. T., and E. Nelson. 2005. "Targeting vs. Instrument Rules for Monetary Policy." *Federal Reserve Bank of St. Louis Review* 87 (5): 597–611.

Mendoza, E. G., and J. D. Ostry. 2008. "International Evidence on Fiscal Solvency: Is Fiscal Policy 'Responsible'?" *Journal of Monetary Economics* 55 (6): 1081–1093. doi:10.1016/j.jmoneco.2008.06.003.

Musgrave, R. 1966. "Prinicples of Budget Determination." In *Public Finance: Selected Readings*, edited by H. Cameron and W. Henderson. New York: Random House.

Newey, W., and K. West. 1987. "A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix." *Econometrica* 55 (3): 703–708.

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. doi:10.1111/ecta.2001.69.issue-6.

Payne, J. 2003. "A Survey of the International Empirical Evidence on the Tax-Spend Debate." *Public Finance Review* 31: 302–324. doi:10.1177/1091142103031003005.

Perron, P., and T. J. Vogelsang. 1992. "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity." *Journal of Business and Economic Statistics* 10: 301–320.

Perron, P., and T. Yabu. 2009. "Testing for Shifts in Trend with an Integrated or Stationary Noise Component." *Journal of Business & Economic Statistics* 27: 369–396. doi:10.1198/jbes.2009.07268.

Shin, Y. 1994. "A Residual-based Test of the Null of Cointegration Against the Alternative of No Cointegration." *Econometric Theory* 10: 91–115.

Stock, J., and M. Watson. 1993. "A Simple Estimator of Cointegrating Vectors in Higher Order Integrated Systems." *Econometrica* 61 (4): 783–820. doi:10.2307/2951763.

Toda, H., and P. Phillips. 1994. "Vector Autoregression and Causality: A Theoretical Overview and Simulation Study." *Econometric Reviews* 13: 259–285. doi:10.1080/07474939408800286.

Toda, H., and T. Yamamoto. 1995. "Statistical Inference in Vector Autoregressions with Possibly Integrated Processes." *Journal of Econometrics* 66: 225–250. doi:10.1016/0304-4076(94)01616-8.

Trehan, B., and C. Walsh. 1991. "Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits." *Journal of Money, Credit, and Banking* 23 (2): 206–223.

Vogelsang, T. (2001), "Testing for a Shift in Trend When Serial Correlation is of Unknown Form." Unpublished Manuscript, Department of Economics, Cornell University.

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. doi:10.2307/2527353.

Westerlund, J. 2006. "Testing for Panel Cointegration with Multiple Structural Breaks." *Oxford Bulletin of Economics and Statistics* 68: 101–132. doi:10.1111/obes.2006.68.issue-1.

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

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