"The relationship between debt level and fiscal sustainability in OECD countries"

Mariam Camarero, Josep Lluís Carrion-i-Silvestre and Cecilio Tamarit

IREA
Research Institute of Applied Economics

AQR
Regional Quantitative Analysis Research Group

I|R|E|A|

Institut de Recerca en Economia Aplicada Regional i Públic
Research Institute of Applied Economics

WEBSITE: www.ub-irea.com • CONTACT: irea@ub.edu

AQR

Grup de Recerca Anàlisi Quantitativa Regional
Regional Quantitative Analysis Research Group

WEBSITE: www.ub.edu/aqr/ • CONTACT: aqr@ub.edu

Universitat de Barcelona
Av. Diagonal, 690 • 08034 Barcelona

The Research Institute of Applied Economics (IREA) in Barcelona was founded in 2005, as a research institute in applied economics. Three consolidated research groups make up the institute: AQR, RISK and GiM, and a large number of members are involved in the Institute. IREA focuses on four priority lines of investigation: (i) the quantitative study of regional and urban economic activity and analysis of regional and local economic policies, (ii) study of public economic activity in markets, particularly in the fields of empirical evaluation of privatization, the regulation and competition in the markets of public services using state of industrial economy, (iii) risk analysis in finance and insurance, and (iv) the development of micro and macro econometrics applied for the analysis of economic activity, particularly for quantitative evaluation of public policies.

The relationship between debt level and fiscal sustainability in OECD countries[*]

Mariam Camarero[†] Josep Lluís Carrion-i-Silvestre[‡] Cecilio Tamarit[§]

Jaume I University   University of Barcelona   University of Valencia

September 5, 2013

Abstract

In this paper we unify the traditional approaches to testing for fiscal sustainability considering the stock-flow system that fiscal variables configure. Our approach encompasses previous ways of testing for sustainability. The results obtained for a group of 17 OECD countries point to weak fiscal sustainability, as well as to the existence of cointegration between deficit and debt, confirming the relevance of the stock-flow approach. Allowing for structural breaks and multicointegration turns out to be of critical importance to assess whether the fiscal authorities apply their policies looking for sustainability and whether, simultaneously, they try to stabilize real debt target levels.

Keywords: fiscal sustainability, cointegration, unit roots, structural breaks

JEL codes: H62, E62, C22

[†]Department of Applied Economics. Jaume I University. Campus del Riu Sec, 12071 Castelló, Spain. Tel: +34 964728590. E-mail: camarero@uji.es

[‡]Corresponding author: AQR-IREA research group. Department of Econometrics, Statistics and Spanish Economy. University of Barcelona. Av. Diagonal, 690, 08034 - Barcelona. E-mail: carrion@ub.edu.

[§]Department of Applied Economics II. University of Valencia. Avinguda dels tarongers, s/n, 46071 Valencia, Spain. Tel: +34 963828349. E-mail: Cecilio.Tamarit@uv.es

1   Introduction

The recent financial and economic crisis of 2008-09 has raised serious concerns about the long-term fiscal sustainability of the euro area countries. Although economic analysis agrees on the need to maintain sustainability of debt/GDP levels, the policy debate still devotes most of the attention to deficit/GDP ratios, which is still the main focus of the Stability and Growth Pact (SGP).[1] Therefore, the practical implementation of the SGP has neglected to some extent the importance of the debt stock level and its behavioural relations with the deficit focusing on the fulfilment of the deficit bounds established by the agreement.[2] This fact led to a failure to adopt sufficiently prudent spending policies in good times and has meant one of the major sources of the fiscal vulnerabilities in euro area countries. Fiscal sustainability requires a government to be solvent, which means that it has to be able to repay its debt at some point in the future. The primary budget balance (budget balance net of interests payments) is a key determinant of government debt dynamics, but there are other factors that have been neglected or, at least, underestimated so far by mainstream economic analysis. In fact, gross debt accumulation is driven by three main factors: first, the above-mentioned government primary balance; second, the so-called "snowball effect", which captures the joint impact of interest payments on the outstanding stock of debt and of real GDP growth and inflation rates over the debt-to-GDP ratio; and third, the deficit-debt relationship, also called "the stock-flow adjustment", which relates to all other factors that affect the outstanding stock of debt but are not recorded as part of the primary balance (see European Central Bank (2011))[3].

However, the empirical literature has rarely tackled the problem of the relationship between debt and deficit until very recently. The no-Ponzi scheme restriction, which is regarded as synonymous with the fulfillment of the Intertemporal Budget Constraint (IBC), imposes testable restrictions on the time series of key fiscal measures such as the stock of public debt, the budget deficit, and the long-run relationship between government expenditures and revenues. In the last two decades, fiscal sustainability has been tested through the use of non-stationary time series analysis. Two different approximations can be found in the literature: first, a univariate time series approach that has focused on the stochastic properties of the deficit inclusive of interest payments – Hamilton and Flavin (1986) – or the stock of debt – Wilcox (1989) – and second, a multivariate one focused on the long-run properties of the flows of expenditures and revenues – Trehan and Walsh (1988), Hakkio and Rush (1991), Haug (1991) or Quintos (1995), among others.

Trying to reconcile the two approaches, Trehan and Walsh (1991) derived that the sufficient and necessary conditions for the IBC to be satisfied are the existence of a cointegration relationship between primary deficit and debt, as well as the I(0) stationarity of the quasi-difference of the primary deficit. More recently, Bohn (1998) has suggested that the analysis of the fiscal

---

[1]For an analysis of the impact of the excessive deficit rule over public finances in the euro area see Gali and Periotti (2003).

[2]The 2005 reform, while maintaining a safety margin with respect to the 3 per cent limit, introduces country-specific medium term objectives, which should take into account the economic features of each member state, such as potential growth, population ageing and public debt /GDP ratio.

[3]The above analysis implies that a full assessment of fiscal sustainability requires a comprehensive approach where debt dynamics should capture the feedback effects between fiscal policies, the macroeconomy and the financial sector.

policy soundness should not be limited to the evaluation of the stationarity of the debt-to-GDP ratio. Fiscal policy reaction functions should be used for the assessment of fiscal deficit sustainability. These should focus on the investigation of whether the governments are reacting to the evolution of debt by adjusting primary balances in the following periods. Bohn (1998) considers that univariate analysis alone could be misleading, as he later proves in Bohn (2007).[4] However, Quintos (1995) shows that it is important to know the order of integration of the deficit since it will define how fast do we tend to fulfil the IBC condition.[5] Following a similar argument, Afonso and Sousa (2011) analyze the effects of fiscal policy on the financial, public and production sectors of four developed economies accounting for the government's intertemporal budget constraint.

In this paper and following Berenguer-Rico and Carrion-i-Silvestre (2011), we unify these approaches considering the stock-flow system that links the fiscal variables. Previous analyses have characterized the flow of expenditures inclusive of interest payments (G_t) and the flow of revenues (R_t) as I(1) non-stationary stochastic processes. The assessment of the order of integration of the variables involved in fiscal sustainability studies – i.e., expenditures, revenues, deficit and debt – is important provided that this step is required in order to estimate potential models that link them. Thus, if debt is I(1) and the deficit is I(0) sustainability is given through the deficit but not through the debt. This means that the government controls the flows of revenues and expenditures, and hence, the deficit, but paying little attention to the stock of debt. In this case we can test whether revenues and expenditures are cointegrated to ensure this type of sustainability. Besides, if both the stock of debt and the deficit are I(0) stationary processes, sustainability is ensured through both the deficit and debt. In this case, the government takes corrective measures on the flows such that the (I(0) stationary) debt can be also controlled. The analysis of these situations requires to work within an I(2) stochastic processes framework and our contribution is twofold. First, we follow the procedure applied by Berenguer-Rico and Carrion-i-Silvestre (2011) to test the null hypothesis of non-cointegration in an I(2) framework, and extend it to allow for the presence of up to two structural breaks. The generalization for the two structural breaks is also conducted for the cointegration test statistics in Gregory and Hansen (1996) and Carrion-i-Silvestre and Sansó (2006). Second, we apply this analysis, for the first time, to a large group of developed OECD countries focusing on the euro area.

The paper is organized as follows. Section 2 summarizes the arithmetic of fiscal sustainability, as well as the main empirical approaches used for testing it. Section 3 presents the econometric model and the statistics applied to test for the presence of cointegration in an I(2) framework allowing for up to two structural breaks. Section 4 reports the results of testing for the OECD countries' fiscal deficit sustainability using the approach that has been proposed in this paper. Finally, Section 5 concludes.

---

[4] However, Bohn (2007) points out that the order of integration of debt or deficit is still informative about the government's management of its public finances. A high order of integration is associated with macroeconomic risks in the long-run.

[5] There is also other econometric technical reasons to know the order of integration of the time series involved in the assessment of fiscal deficit sustainability given the potential spurious regression problem that can appear.

2 The arithmetic of deficit and debt sustainability

In this section, we derive the algebra for an "ad hoc" version of the IBC and the implied stationarity restrictions. The one-period government budget constraint can be written as follows:

$$\Delta B_t = G_t - R_t = DEF_t, \tag{1}$$

where B_t is the real market value of government debt, G_t is real government expenditure inclusive of interest payments, R_t represents real tax revenues and $\Delta = (1 - L)$ is the first difference operator. The deficit (DEF_t) is the one-period difference between outlays and revenues and it also equals the change in public debt. However, as claimed by Bohn (2005), while (1) holds in nominal terms, changes in real debt differ from the real value of deficit by an inflation term. Therefore, it is important in this context in order to separate the stock of debt from the outflows of outlays and revenues, to use a scale-invariant definition of debt dynamics.

Denoting i_t as the real interest rate[6] and assuming i_t to be I(0) stationary around a mean i, as in Hakkio and Rush (1991), we can define:

$$G_t = GE_t + i_t B_{t-1}, \tag{2}$$

where GE_t is the real expenditure exclusive of interest payments, and the second term on the right hand side of (2) represents interest payments on the level of debt accumulated at the end of the previous period. Further, we can express the debt as:

$$B_t = (1 + i)B_{t-1} + EXP_t - R_t, \tag{3}$$

where $EXP_t = GE_t + (i_t - i)B_{t-1}$, or, alternatively, $B_t = \left(\frac{1}{1+i}\right)(R_{t+1} - EXP_{t+1}) + \left(\frac{1}{1+i}\right)B_{t+1}$. As the government is subject to the same restriction for periods $t+1$, $t+2$, ..., we can aggregate intertemporaly the different budgetary restrictions for each individual period and obtain:

$$B_t = \sum_{j=0}^{\infty}\left(\frac{1}{1+i}\right)^{j+1}(R_{t+j+1} - EXP_{t+j+1}) + \lim_{j\to\infty}\left(\frac{1}{1+i}\right)^{j+1}B_{t+j+1}. \tag{4}$$

The representation of (4) in terms of the first difference of B_t is the standard specification that is used in the empirical literature to test for fiscal deficit sustainability – see Quintos (1995). If we take first differences on (4) the sustainability of the fiscal deficit is associated with the condition:

$$\lim_{j\to\infty} E_t \left(\frac{1}{1+i}\right)^{j+1}\Delta B_{t+j+1} = 0. \tag{5}$$

In Table 1 we present a summary of the different hypotheses and conditions that have been tested in the empirical literature on fiscal policy sustainability – see also Afonso (1995) for a comprehensive overview. As mentioned in the introduction, the empirical literature has followed

---

[6]Note that the variables could be expressed in nominal terms, real terms, or as a ratio to GDP as long as i_t is adjusted accordingly (i.e., if the variables are in nominal terms, i_t is the nominal interest rate; if the variables are in real terms, as it is our case, i_t is the real interest rate; if the variables are ratios to GDP, $1 + i_t$ is the growth-adjusted real interest rate that follows from dividing the gross real interest rate by the gross rate of output growth).

two different routes when assessing fiscal sustainability. The first group of studies – hereafter, univariate-based approach – has concentrated on the analysis of the univariate properties of B_t – see Hamilton and Flavin (1986) and Wilcox (1989). The second group – henceforth, cointegration-based approach – assesses whether R_t and G_t are cointegrated assuming either that the cointegrating vector is known and equal to $(1, -1)$ or estimating it – see Trehan and Walsh (1988), Trehan and Walsh (1991), Hakkio and Rush (1991), Haug (1991), Quintos (1995) and Martin (2000).

Trehan and Walsh (1991) can be considered as the first contribution that unifies the two previous approaches. In particular, they are the first to explicitly derive the fiscal deficit sustainability conditions in terms of a relationship between primary deficit and debt. In addition, Hakkio and Rush (1991) establish an implicit relationship between the deficit and debt, although they focus on the cointegration relationship among the components of the primary deficit. Hakkio and Rush (1991) derive that provided that total revenue and expenditure define a cointegrated relationship:

$$R_t = \mu + \beta_{1,0} G_t + u_t, \tag{6}$$

with an estimated $\hat{\beta}_{1,0}$ that lies within the interval $0 < \hat{\beta}_{1,0} < 1$, the no-Ponzi scheme condition is still satisfied. In economic terms, this corresponds to a situation when the government reacts to the increase in the public debt, but this correction is not equal to the growth in the public expenditure. In this case, growing budget deficit and non-stationary change in public debt could also be observed.

The above findings were later developed by Quintos (1995), who distinguishes between two degrees of deficit sustainability, i.e., strong and weak sustainability. *Strong sustainability* exists when revenues and expenditures are cointegrated with cointegrating parameter $\beta_{1,0} = 1$ in (6). *Weak sustainability* holds when revenues and expenditures are cointegrated and $\beta_{1,0}$ lies in $(0, 1)$. Quintos (1995) argues that although the condition $0 < \beta_{1,0} < 1$ is sufficient for the fiscal deficit to be sustainable, it is inconsistent with the government's ability to market its debt in the long-run. She also mentions that the condition $0 < \beta_{1,0} < 1$ has serious policy implications because a government that continues to spend more than it earns has a high risk of default and would have to offer higher interest rates to service its debt.[7] Regardless of the degree of sustainability, the cointegration based approach states that the deficit has to be $\Delta B_t \sim I(0)$ to be sustainable, which in turn implies that the debt B_t is $I(1)$. This means that the government is equilibrating the flows but with little attention to the stocks, i.e., they can grow infinitely.

Note that even in the case where the cointegrating vector is set to $(1, -1)$, the univariate-based and cointegration-based approaches are focusing on different variables, as the former concentrates on the debt while the latter on the deficit. In this paper we argue that, in fact, it is possible to reconcile these approaches. It is also worth noticing that this type of analysis only focuses on the cointegration relationship to test for fiscal sustainability, and no attention is paid to the deterministic component in (6). It is difficult to give an interpretation of a particular estimate of the constant in this model, although there is a particular case that deserves some attention. Thus, if there is strong sustainability we can rewrite (6) as $R_t - G_t = \mu + u_t = DEF_t$, so that the constant term will give a measure of the mean around which the deficit is evolving in

---

[7]See Quintos (1995), pp. 410.

the long-run. If the constant term is not statistically significant, the deficit will evolve around zero in the long-run, i.e., we end up with balanced public accounts in the long-run. Positive values of the constant will indicate that the government is increasing his savings (wealth), whereas negative values indicate that the government is increasing the debt. This sort of interpretation is more difficult if we are under weak sustainability, although we can still conclude that positive values of the constant will lead to an increase of government wealth – that can be used to repay the debt – and negative values imply an increase of the debt.

Bohn (1998, 2007) can be considered two path-breaking contributions to the unification of the approach to testing for fiscal deficit sustainability. Bohn (2007) summarizes all the previous research techniques developed to testing whether the no-Ponzi scheme restriction is satisfied. He also shows that the cointegration relationship between total revenue and expenditure is not a necessary condition for the no-Ponzi scheme condition to hold. Thus any finite order of integration of these series separately leads also to the fulfilment of the IBC. He also adopts the concepts of strong and weak sustainability proposed by Quintos (1995) and relate them to the order of integration of the debt: strong sustainability implies that the debt is difference stationary, whereas an I(2) debt is associated with weak sustainability, or even "absurdly" weak when the order of integration is higher – see Table 1.

These findings suggest that unit roots and cointegration techniques are not appropriate to analyze fiscal policy sustainability, as this would always point towards the conclusion that the debt series are integrated of any finite order, which implies that the no-Ponzi scheme restriction holds. Bohn (2007) therefore proposes to pay more attention to the fiscal reaction functions, as in Bohn (1998) for the case of the United States. A positive reaction of the primary surplus to the increase in public debt is referred to as a sustainable fiscal policy. However and even if Bohn' criticism is correct, the analysis of the order of integration of the variables involved in the different approaches of fiscal sustainability is interesting from both an economic and econometric point of views. Thus, note that if relevant economic conclusions are to be obtained, the estimation of the models that are used to test for fiscal deficit sustainability, regardless of the approach that is followed, should be based on a consistent approach. This requires to assess the order of integration of the variables in order to know whether the classical regression approach is valid – when the variables are I(0) – or the cointegration analysis approach should be taken instead – if the variables are I(1). Finally and as will be evidenced below, our analysis involves similar variables to those considered in Bohn (1998) – revenues, expenditures and debt – so that the approximation in this paper can be seen as a unified approach to the analysis of the fiscal deficit sustainability.[8]

3   The econometric model and test statistics

In this paper, we propose to analyze if there is a relationship between the stocks of revenues and expenditures using a model specification that is based upon the cumulated equation (6).

---

[8]Bohn (1998) estimates fiscal policy rules where revenues or expenditures respond to government debt. Note that, in fact, this implies the existence of a (long-run) relationship among revenues, expenditures and debt, which is the goal of our analysis. However, the assessment of whether this long-run relationship exist and, hence, the inference techniques that can be applied, crucially depend on the order of integration and cointegration analysis using these time series, something that is not carried out in Bohn (1998) and other related studies.

Let us define $\sum_{j=1}^{t} R_j$ and $\sum_{j=1}^{t} G_j$ the cumulated government revenues and expenditures, respectively. We can specify a model such as:

$$\sum_{j=1}^{t} R_j = \alpha + \mu t + \beta_{1,0} \sum_{j=1}^{t} G_j + \delta_{1,0} G_t + u_t. \tag{7}$$

Equation (7) includes the level of expenditures among the stochastic regressors in order to cover the feature that policymakers pay attention to both the stock of debt and the level of deficit when designing fiscal policies. This interpretation can be made if we note that setting $\beta_{1,0} = 1$ in (7) leads to:

$$\sum_{j=1}^{t} R_j - \sum_{j=1}^{t} G_j = \alpha + \mu t + \delta_{1,0} G_t + u_t = -B_t, \tag{8}$$

where $B_t = \sum_{j=1}^{t} G_j - \sum_{j=1}^{t} R_j$ is the stock of debt. It can be seen that (8) defines a potential long-run relationship between the (negative value of the) debt and the expenditures once the constraint $\beta_{1,0} = 1$ is imposed – notice that this approach defines the multicointegration testing concept proposed in Granger and Lee (1989), where the cointegrating vector among the levels of the variables is assumed to be known. The specification given in (7) implies working within a stock-flow setup, which allows us to consider whether governments are taking corrective measures on flows – in our case, revenues and expenditures – in such a way that they are also controlling, to some extent, the stock of debt. Specifically, we can investigate if the growth of the debt is under control, maybe basing it on the growth of the economy. This would imply that governments implement fiscal policies that affect the flows of revenues and expenditures that aim to control the stock of debt.

In this multicointegration setup and in order to test for fiscal deficit sustainability, we have to analyze the conditions in terms of the expressions above, and in particular, of the single equation model given by (7). This expression integrates the flow I(1) variable G_t and the I(2) stock variables $\sum_{j=1}^{t} R_j$ and $\sum_{j=1}^{t} G_j$. The parameters $\beta_{1,0}$ and $\delta_{1,0}$ define the first and the second cointegration layers, respectively. The first layer refers to the cointegration relationship between the flow variables, whereas the second relates flow and stock variables. The assessment of the degree of sustainability of the fiscal policy will depend on the values of these parameters. If $\beta_{1,0} > 1$ then surpluses have been, on average, predominant, whereas if $\beta_{1,0} < 1$, deficits have outpaced surpluses. We can combine this information with the one provided by $\delta_{1,0}$, which relates flow and stock variables. This parameter indicates how fiscal policy reacts to the accumulation of debt or wealth, that is, whether this reaction favours budgetary sustainability. Concerning $\delta_{1,0}$, sustainability will depend also on the value of $\beta_{1,0}$ – see Escario, Gadea, and Sabaté (2012):

1. If $\beta_{1,0} > 1$, we have a majority of surpluses, so that sustainability will require $\delta_{1,0} > 0$, that is, expenditure should increase to accommodate increasing levels of cumulated revenues.

2. In contrast, if $\beta_{1,0} < 1$, deficits have been predominant, so that expenditure should decrease to compensate the accumulation of debt (that is, $\delta_{1,0} < 0$).

So far, the discussion that summarizes the different approximations in the literature on

fiscal deficit sustainability leads us to distinguish two broad concepts of sustainability. First, we can test whether the fiscal policy is sustainable looking at the value of the $\beta_{1,0}$ parameter in (6) – with the associated interpretation in terms of weak $(0 < \beta_{1,0} < 1)$ or strong $(\beta_{1,0} = 1)$ sustainability. We can think of this approximation as a way of testing for a *first layer* (weak or strong) *of fiscal sustainability*. In this case, the flow fiscal variables are the ones in which we focus the analysis. Second, we can assess whether the fiscal policy is sustainable not only paying attention to the flow variables but also to the stock of debt. In this case, the $\beta_{1,0}$ parameter in (6) and, equivalently, in (7), can take values smaller or larger than one provided that the inequality that affects the $\delta_{1,0}$ parameter goes on the same direction. This means that the practice of the governments that affect the flow fiscal variables is also influenced by the stock of debt. We can think of this second approximation as a way of testing for a *second layer of fiscal sustainability*.

As noted above, most of the empirical studies have highlighted the existence of regimes in the sustainability process and some of them have introduced structural breaks in the analysis of (6). If such regime changes are present, as the literature concludes, the test statistic developed by Haldrup (1994) and Engsted, Gonzalo, and Haldrup (1997) should not be applied since it does not account for regime shifts. One of the contributions of our paper is to implement a model and test statistic developed by Berenguer-Rico and Carrion-i-Silvestre (2011) to test for I(2) cointegration with regime shifts defining the model:[9]

$$
y_t = \alpha + \mu t + \sum_{i=1}^{l} \theta_i DU_{i,t} + \sum_{i=1}^{l} \gamma_i DT_{i,t} + \beta_{1,0} x_{2,t} + \sum_{i=1}^{l} \beta_{1,i} DU_{i,t} x_{2,t} + \delta_{1,0} x_{1,t} + \sum_{i=1}^{l} \delta_{1,i} DU_{i,t} x_{1,t} + u_t,
\tag{9}
$$

where $y_t = \sum_{j=1}^{t} R_j$, $x_{2,t} = \sum_{j=1}^{t} G_j$ and $x_{1,t} = G_t$, l denotes the number of structural breaks, $DU_{i,t} = 1$ and $DT_{i,t} = (t - T_i)$ for $t > T_i$, and 0 otherwise, where $T_i = [\lambda_i T]$ denotes the i-th break point, $i = 1, 2, \ldots, l - \lambda_i \in \Lambda$ is the break fraction parameter, where Λ is a closed subset of $(0, 1)$, and $[\cdot]$ being the integer part.[10] In order to assess the integration order for $\hat{u}_t$ the augmented Dickey-Fuller (ADF) test statistic $(t_{\varphi_0}(\lambda))$ can be computed as the t-ratio for testing the null hypothesis that $\varphi_0 = 0$ – i.e., $u_t \sim I(1)$ – against the alternative hypothesis that $\varphi_0 < 0$ – i.e., $u_t \sim I(0)$ – in the model

$$
\Delta \hat{u}_t = \varphi_0 \hat{u}_{t-1} + \sum_{j=1}^{p} \varphi_j \Delta \hat{u}_{t-j} + \eta_t,
\tag{10}
$$

for each possible vector of break points, $\lambda \in \Lambda^l$, which defines a sequence of statistics. Then, the minimum of the sequence of ADF statistics is taken, which defines the statistic $t_{\varphi_0}^* = \min_{\lambda \in \Lambda^l} t_{\varphi_0}(\lambda)$. Berenguer-Rico and Carrion-i-Silvestre (2011) consider the case of $l = 1$ structural break providing critical values to test the null hypothesis of $u_t \sim I(1)$. In this paper we extend their analysis allowing for up to two structural breaks for the model specification given in (9).

---

[9] See Model 8 in Berenguer-Rico and Carrion-i-Silvestre (2011).

[10] For instance, a usual definition in the literature, which is also used in this paper, is $\Lambda = [0.15, 0.85]$.

4 Empirical results

In this section we present the results of the analysis of fiscal sustainability and debt solvency for a group of 17 developed OECD countries, paying special attention to the eurozone economies. The database has been obtained from AMECO, based on consistent time series across countries according to the European Commission definitions. The three variables are real revenues, real expenditures and, for completeness, real debt. The countries analyzed are 17, the majority EU members and also EMU participants. The exceptions are Canada, Denmark, Japan, Sweden, the UK and the US. The sample extends from 1970 until 2012 for most of them.[11],[12]

4.1 Order of integration analysis

Given previous results in the literature and the expected effects of the different economic crises that might have affected the variables that we are dealing with, we start the analysis of the order of integration of the time series involved in our study investigating the presence of structural breaks. This is an important feature provided that unit root tests can lead to misleading conclusions if the presence of structural breaks is not accounted for when testing the order of integration. Therefore, the first stage of our analysis has focused on a pre-testing step that aims to assess whether the time series are affected by the presence of structural breaks regardless of their order of integration. This pre-testing stage is a desirable feature, as it provides an indication of whether we should then apply unit root tests with or without structural breaks depending on the outcome of the pre-test.

This testing problem has recently been addressed by Perron and Yabu (2009), who define a test statistic that is based on a quasi-GLS approach using an autoregression for the noise component, with a truncation to 1 when the sum of the autoregressive coefficients is in some neighborhood of 1, along with a bias correction. For given break dates, one constructs the F-test $(Exp - W_{FS})$ for the null hypothesis of no structural change in the deterministic components. The final statistic uses the Exp functional of Andrews and Ploberger (1994). Perron and Yabu (2009) specify three different models depending on whether the structural break only affects the level (Model I), the slope of the trend (Model II) or the level and the slope of the time trend (Model III). The computation of these statistics, which are available in the companion appendix, show that we find more evidence against the null hypothesis of no structural break with Model III. Further, the variable that presents more evidence of structural breaks is real expenditures – the null hypothesis of no structural break is rejected for eleven countries – followed by real debt – nine countries – and real revenues – the null hypothesis of no structural break is rejected

---

[11]Real revenues, expenditures and deficit time series are available for Japan for the 1981-2012 period, and for Canada for the 1970-2008 period. As for the real debt, there is lack of information at the beginning of the time period for France (the time series starts at 1977), the Netherlands (starts at 1975), Portugal (starts at 1973) and Canada (starts at 1980).

[12]Although it would be possible to work with quarterly time series – see Afonso, Agnello, Furceri, and Sousa (2011) and Afonso and Sousa (2011) – we have decided to use annual data provided that it comes from the same source of information and is homogeneous and comparable across countries. Further, what is important for unit root and cointegration analysis is the length of the time period that is covered not the frequency of the data provided that these are long-run properties of the stochastic processes. Finally, the use of high frequency data sometimes imply to work with seasonal adjusted data, which can lead to a spurious detection of the order of integration of the variables. In these regards, see Ghysels and Perron (1996).

for six countries.

Taking into account these results, Table 2 indicates that the null hypothesis of a unit root is not rejected by the modified unit root test statistics MZ_α, MSB, MZ_t and MP_T (M-type test statistics) in Ng and Perron (2001) for the real revenues, expenditures and debt for almost all the countries for which the $Exp - W_{FS}$ indicates that there are no structural breaks – the exception is the debt of Greece, for which the null hypothesis of a unit root is clearly rejected.[13] We also report the Z_α unit root test statistic in Ng and Perron (2001), and the ADF and P_T unit root test statistics in Elliot, Rothenberg, and Stock (1996) for completeness, although Ng and Perron (2001) show that the M-type test statistics have better finite sample performance in terms of empirical size and power.

As for the time series for which the presence of parameter instabilities has been found, we have computed the unit root test statistics in Zivot and Andrews (1992), Lumsdaine and Papell (1997) and Carrion-i-Silvestre, Kim, and Perron (2009). It should be mentioned that the test statistics in Zivot and Andrews (1992) and Lumsdaine and Papell (1997) allow for one or two structural breaks, respectively, only under the alternative hypothesis of I(0) but not under the null hypothesis of I(1). This implies an asymmetric treatment of the structural breaks, since they appear or disappear depending on whether we are under the null or alternative hypothesis. On the contrary, the unit root tests in Carrion-i-Silvestre, Kim, and Perron (2009) allow for multiple structural breaks under both the null and alternative hypotheses. This difference on the treatment of the structural break is important provided that the rejection of the null hypothesis of a unit root when using the test statistics in Zivot and Andrews (1992) and Lumsdaine and Papell (1997) might be simply due to the misspecified null hypothesis. Therefore, more weight should be given to the conclusions drawn from the application of the statistics in Carrion-i-Silvestre, Kim, and Perron (2009), since we have obtained evidence in favor of the presence of structural breaks regardless of their order of integration. However, we also report the ADF test statistics in Zivot and Andrews (1992) and Lumsdaine and Papell (1997) to get a complete picture.

The results of the unit root test statistics in Zivot and Andrews (1992), Lumsdaine and Papell (1997) and Carrion-i-Silvestre, Kim, and Perron (2009) are reported in Table 3, along with the estimated break points $(\hat{T}_i, i = 1, 2, 3)$ that arise from each procedure. As can be seen, the M-type unit root tests proposed by Carrion-i-Silvestre, Kim, and Perron (2009) led to the non-rejection of the null hypothesis of a unit root in all cases at the 5% level of significance. There is some cases where the null hypothesis of unit root is rejected by the ADF and Z_α test statistics, although, as mentioned above, the M-type tests have better finite sample performance.[14] Some contradictory results are obtained when computing the ADF test statistics in Zivot and Andrews (1992) and Lumsdaine and Papell (1997), especially when we allow for two structural breaks. Thus, we can see that the null hypothesis of I(1) is rejected at the 5% level of significance by

---

[13]It should be mentioned that the ADF test statistic rejects the null hypothesis of unit root in the case of the revenues for Finland and Netherlands, and in the case of the expenditures for Austria. However, we base our analysis on the M unit root tests as they show better performance in finite sample than the ADF test statistic.

[14]It should be mentioned that the ADF test statistic rejects the null hypothesis of unit root for four time series, although at the 10% level of significance. Further, the null hypothesis of unit root is rejected by one of the seven test statistics that have been computed in the case of the expenditure of Ireland. However, we base our analysis on the M unit root tests as they show better performance in finite sample than the ADF test statistic.

the ADF test statistic in Zivot and Andrews (1992) in four cases – six cases if we set the level of significance at the 10% – whereas it is rejected by the ADF test statistic in Lumsdaine and Papell (1997) in seven cases – ten cases when working at the 10% level of significance. However, it should be born in mind that these contradictions might be pointing out the need to consider the presence of the structural breaks also under the null hypothesis of a unit root, something that is not allowed when computing the unit root test statistics in Zivot and Andrews (1992) and Lumsdaine and Papell (1997).

To sum up, the order of integration analysis has revealed that, in general, the null hypothesis of unit root cannot be rejected for the time series for which no evidence of structural breaks has been found – the exception is Greece. For the time series for which structural breaks have been considered, the M-type test statistics in Carrion-i-Silvestre, Kim, and Perron (2009) do not reject the null hypothesis of unit root at the 5% level of significance. Taking into account all these results, the main conclusion that we draw is that the real revenues, expenditures and debt can be considered I(1) stochastic processes.[15]

It is worth mentioning that our analysis is conditional on the pre-testing stage that determines the presence of structural breaks using the test statistic in Perron and Yabu (2009). As shown in Harvey, Leybourne, and Taylor (2013), the use of this type of pre-testing statistics can produce poor performance if the magnitude of the change in the slope of the time trend is small. Instead, Harvey, Leybourne, and Taylor (2013) propose the use of a unit root test statistic that generalizes the proposal in Perron and Rodríguez (2003) to test the null hypothesis of unit root with local multiple trend breaks. The idea is to compute a unit root test statistic that models the magnitude of the trend breaks as local-to-zero, capturing situations where the magnitude of the structural breaks are of moderate magnitude. Based on their simulations, Harvey, Leybourne, and Taylor (2013) show their minimum Dickey-Fuller (MDF) test statistic has good power when the local break magnitudes in absolute values are between 1 and 6 – this range defines what is called the "power valley" for the unit root tests in Carrion-i-Silvestre, Kim, and Perron (2009). In order to get robust conclusions of the order of integration analysis that we have conducted, we have also computed the unit root test statistic in Harvey, Leybourne, and Taylor (2013) for up to three structural breaks.

Table 4 presents the results of Harvey, Leybourne, and Taylor (2013) methodology. Looking at the results for the MDF test statistic together with the estimated values of the local break magnitudes ($\kappa_i, i = 1, 2, 3$) that end up in the range between 1 and 6 (in absolute values), we can see that there are few situations where the conclusion about the order of integration that have been obtained above is reversed with the use of the MDF test statistic. To be specific, this only happens for Netherlands (revenues and expenditures), Sweden (revenues), Belgium (expenditures) and Ireland (expenditures and debt), which represents 11.8% of the cases that we have analyzed.[16] Further, it should be noticed that not all the local break magnitudes lie inside the [1,6] range for the cases where more than one structural break has been selected –

---

[15] Results reported in the companion appendix indicate that the null hypothesis of a unit root can be rejected for the variables in first difference, a conclusion that is reached after the unit root test statistics with structural breaks in Perron and Vogelsang (1992), Clemente, Montañés, and Reyes (1998), and Carrion-i-Silvestre, Sansó, and Artís (2004) have been applied.

[16] The number of structural breaks has been selected using the BIC information criterion of the local GLS regression in Harvey, Leybourne, and Taylor (2013) allowing for a maximum of three structural breaks.

i.e., for Belgium and Ireland (debt) – and that the rejection of the unit root hypothesis is done at the 10% level of significance in some cases. Consequently, we can see that the qualitative conclusion pointing out the I(1) non-stationarity of the time series that we have analyzed is not significantly reversed when computing the MDF test statistic.

Although it is not the primary focus of this paper, these results allow us to take a first look at the issue of fiscal sustainability. Notice that assessing that $B_t \sim I(1)$ can be taken as evidence pointing towards fiscal deficit sustainability by a strand of the literature. According to this testing approach, if debt is a I(1) non-stationary process, it implies that the fiscal deficit is I(0). More specifically, this is the approach that was followed in the papers that impose a (1, -1) cointegrating vector among the revenues and expenditures and find that the fiscal deficit is I(0). Note that this sort of analysis imposes the cointegrating vector and, hence, rules out the possibility of getting a richer interpretation in terms of weak or strong fiscal sustainability as discussed above.

Once we allow for multiple structural changes and based on the results of the Carrion-i-Silvestre, Kim, and Perron (2009) approach, the first striking feature is the presence of up to three changes in many cases, especially for the variable debt: 13 out of 17 countries present three breaks, and the variables are still I(1) non-stationary. Two breaks is the most common pattern of revenues and expenditures. Concerning the placement of the breaks, only in a few cases have the structural breaks occurred in the seventies or first years of the eighties: two times (Austria and Belgium) for the revenues; four for the expenditures (Belgium, Denmark, France and Italy), and six for the debt (the latter four plus Sweden and Spain). They are related to the response to the second oil crisis in some European countries. A larger group of structural breaks can be placed between 1984 and 1992. Broadly speaking, the years 1989-91 concentrate the majority of the first structural breaks for many variables and are related to the strengthening of economic integration in Europe, the collapse of the Central and Eastern Europe regimes and the German reunification. They generally correspond to improvements in fiscal variables (with the exception of Finland and Germany) and the beginning of the consolidation process towards monetary union in Europe. The cases of Canada, Japan or the US do not follow this pattern. The next set of structural changes is mainly related with the creation of EMU and can be found around 1999-2001. This is the case, for example, of Germany in revenues, together with Ireland, Finland or Spain, but notably in expenditure or debt: Finland, Ireland, Spain, Italy or France. Earlier in the nineties the Asian crisis hit Japan, the US or Canada (between 1995-1997). Finally, the rest of the breaks mainly correspond to the recent world financial crisis and the recession that followed (breaks in the period 2007-2009).

4.2   Revenues and expenditures relationship: First cointegration layer

In order to analyze the existence of a cointegration relationship between real revenues and expenditures we might have estimated the model:

$$R_t = \mu + \beta_{1,0} G_t + u_t,$$

and compute the Engle and Granger (1987) cointegration test statistic using the estimated residuals of this equation, although the presence of structural breaks affecting the time series in the model would invalidate the statistical inference. Instead, in order to deal with structural breaks, we have proceeded to compute the Gregory and Hansen (1996) cointegration test statistic, which allows for the presence of one structural break under the alternative hypothesis of cointegration. It should be mentioned that we have also extended the test statistics in Gregory and Hansen (1996) to allow for two structural breaks. Table 5 presents the result of the cointegration test statistics for the model that accounts for up to two structural breaks affecting only the level of the relationship (Model C using the notation in Gregory and Hansen (1996)):

$$R_t = \mu + \sum_{i=1}^{l} \theta_i DU_{i,t} + \beta_{1,0} G_t + u_t,$$

$l = 1, 2$, and also the results for the model that considers up to two structural breaks affecting both the level and the cointegrating vector (Model C/S using the notation in Gregory and Hansen (1996)):

$$R_t = \mu + \sum_{i=1}^{l} \theta_i DU_{i,t} + \beta_{1,0} G_t + \sum_{i=1}^{l} \beta_{1,i} DU_{i,t} G_t + u_t,$$

$l = 1, 2$. In this case, the use of the ADF statistic indicates that the evidence against the null of no cointegration increases when considering up to two structural breaks – using Model C/S, the null of no cointegration cannot be rejected for Belgium, Ireland and Japan with one structural break, although it is rejected if we consider two structural breaks.

We have also computed the cointegration test statistics proposed by Carrion-i-Silvestre and Sansó (2006), which accommodate the presence of one structural break affecting the parameters of the model under both the null hypothesis of cointegration and the alternative hypothesis of spurious regression – i.e., this approach reverses the null and alternative hypotheses in Gregory and Hansen (1996). Carrion-i-Silvestre and Sansó (2006) test follows the suggestions made by Engle and Granger (1987), Phillips and Ouliaris (1990) and Engle and Yoo (1991), who argued that the natural specification to test should be the null hypothesis of cointegration rather than the null hypothesis of absence of cointegration provided that the null hypothesis will be only rejected when there is strong evidence against what the economic theory is proposing. Taking into account these concerns, we have estimated the model given by:

$$R_t = \mu + \sum_{i=1}^{2} \theta_i DU_{i,t} + \beta_{1,0} G_t + \sum_{i=1}^{2} \beta_{1,i} DU_{i,t} G_t + u_t, \tag{11}$$

allowing for up to two structural breaks.[17] Two different specifications are considered, i.e., the one given by Model An – which imposes $\beta_{1,1} = \beta_{1,2} = 0$ in (11) – and the one given by Model D – which does not impose any constraint in (11). The application of the Carrion-i-Silvestre and Sansó (2006) test statistics points to the non-rejection of the null hypothesis of cointegration in all cases once the presence of structural breaks has been considered in the model – in order

---

[17]Although we have allowed up to three structural breaks in the order of integration analysis, we consider a maximum of two structural breaks in the cointegration analysis due to the number of time observations that we have and to avoid the criticism of data mining.

to save space, we do not report the tables with these results, although they are available in the companion appendix. This conclusion is obtained regardless of whether we allow for one or two structural breaks.

The dynamic ordinary least squares (DOLS) estimates for each country are reported in Table 6. In order to reduce the presentation of results, we only report the estimates for each country considering one or two structural breaks and the type of the structural breaks effect – i.e., Model An or Model C – according to the selection pointed out by the BIC information criterion – the full set of results is available in the companion appendix. As can be seen, almost all parameters are statistically significant at the 5% level – the exceptions are Greece (the first and second change in level and the first change in the slope), Portugal (the intercept) and Spain (first change in the slope is statistically significant at the 10% level); in these cases, an hybrid specification could be estimated, although this does not affect the consistency of the other estimated parameters.

Using these estimates, we have proceed to test whether the coefficient of the expenditure is equal to one (strong sustainability) or not (weak sustainability). In general, the Wald test statistic rejects the null hypothesis that the coefficient is equal to one in all cases except for the following situations: (i) we find evidence of strong sustainability for Canada after the second structural break that took place in 2000, (ii) for Denmark, strong sustainability is found up to 2003, (iii) for Greece strong sustainability is found after 1997, (iv) for Ireland strong sustainability is only given between 1970-1978, (v) for the Netherlands, between 1970-1988, (vi) for Portugal, between 1991-2004, (vii) for Spain, for the period 1989-2006, and (viii) for US, only between 1998-2001. Therefore, for most cases the evidence points to the presence of weak sustainability, a situation that indicates that these countries will face difficulties to market their debt.

From an economic point of view, and bearing in mind that this analysis only looks at the relation between revenues and expenditures, the results can be related with fiscal reforms or fiscal consolidation processes. In the case of Canada in the mid-nineties the authorities adopted a series of measures to restore fiscal balance and decrease debt ratios, achieving a surplus around 1998 – see Traclet (2004). According to our results, strong sustainability was achieved after 2000. Denmark showed strong sustainability up to 2003, a reflection of sound public finances, also maintained after that date (weak sustainability). The reason for the (surprising) finding of strong sustainability in Greece after 1997 is related to the real surplus that was maintained up to 2005[18] that becomes apparent in Figure 1 in the companion appendix. Concerning Ireland, the rejection of the null hypothesis of $\hat{\beta} = 1$ has been due, in contrast to other countries, to the fact that its value was larger, not smaller, than one. Real surpluses have been present up to the financial crisis of 2007. Dutch fiscal policy has been stable, with no significant disequilibria and sustainability, either strong or weak, of public finances during the whole sample. Portugal presents strong sustainability during the nineties, probably thanks to different consolidations episodes (at the end of the eighties and in 1992). However, Portugal was the first country in the EU to breach the SGP deficit limit in 2001, becoming subject to the Excessive Deficit Procedure (EDP) in 2002 and again in 2005. In the Spanish case, with some cyclical disequilibria, real

---

[18]See the next section for a full discussion comparing the first layer and second layer cointegration results.

budget deficit was kept in line with the Maastricht criteria after the fiscal consolidation in the mid-nineties. This has been captured by the 1989-2006 period of strong sustainability. Finally, from 1998 to 2001 the US had budget surpluses followed by a tax cut,[19] and this is precisely the only period of strong sustainability in the US fiscal policy according to our results.

As for the changing constant term in (11), the estimated coefficients in Table 6 indicate that, in general, they are statistical significant at least at the 10% level of significance. An interesting feature that we can observe is that the magnitude of the level of the model is predominantly positive in the last regime that has been estimated, indicating that the countries in our sample have been forced to increase their revenues above their expenditure levels. For instance, for the US we have that the constant is 150.56 up to 1997, it changed to -1285.8 ($= 150.56 - 1436.36$) in the period 1998-2001, and becomes positive after 2001 (919.97). This is a common behaviour in most cases, with the exceptions of Canada and Sweden, whose estimated level values in the third regime are negative.

To sum up, the evidence found in this section leads to the presence of a long-run relationship between the real revenues and the real expenditures, although such a relationship has suffered the effect of structural breaks. Moreover and except for Canada, all the countries analyzed present fiscal sustainability problems in the last decade as they are only sustainable in the weak sense. This points to the need of a deeper analysis involving the dynamic relationship among the stock of debt and the revenue and expenditure flows.

4.3   Debt, revenues and expenditures: Second cointegration layer

As mentioned above, a stronger definition (or control) of deficit sustainability would not only focus on the long-run relationship between the revenues and the expenditures, but also would take into account the level of debt of the economy. This relationship would correspond to a specification within a stock-flow model. This approach has not attracted too much attention from researchers when analyzing fiscal deficit sustainability, maybe because the level of debt of the developed economies was not really high (compared to their GDP) or because of the absence of constraints in the access to credit in the financial markets. However, the presence of such relationship will give a stronger definition to the deficit sustainability and solvency of the governments. Testing for the presence of this stronger relationship – or second cointegration layer – can be done using the multicointegration framework. There is two different approaches to test for the presence of multicointegration.

The first one is the approach by Granger and Lee (1989). In our case, to implement it we test for the presence of a long-run relationship between debt and real expenditures, which is equivalent to impose that $\beta_{1,0} = 1$ in (7). The second approach to multicointegration bases on the proposal in Engsted, Gonzalo, and Haldrup (1997) and consists of estimating equation (7) without imposing the restriction $\beta_{1,0} = 1$. Therefore, in this paper we follow the latter approximation provided that it does not require to impose a constraint on one of the parameters. Further, in our case, we have proceed to test for the presence of multicointegration using the proposal in Berenguer-Rico and Carrion-i-Silvestre (2011), who extend Engsted, Gonzalo, and Haldrup (1997) approach allowing for one structural break – here, we also consider the case of

---

[19]We thank an anonymous referee for pointing this out.

two structural breaks.

As can be seen in Table 5, the null hypothesis of no multicointegration is rejected in just two cases when one structural break is considered in the model, whereas it is rejected in six cases when two structural breaks are allowed for. For these countries, we can conclude that there is a deeper cointegration relationship that not only links the revenues and the expenditures (first cointegration layer), but also the cumulation of the residuals of this relationship (second cointegration layer) that, in principle, does not need to be equal to the debt – i.e., we do not require to impose $\beta_{1,0} = 1$ in (7) to find this deeper cointegration relationship.

Next, we present the estimated model for those countries with a multicointegration relationship, allowing for either one or two structural breaks. Table 7 reports the DOLS estimates of the parameters corresponding to the model allowing for one or two structural breaks for those countries for which the ADF test statistic has detected the presence of multicointegration. The equation that has been estimated is given by:

$$y_t = \alpha + \mu t + \sum_{i=1}^{l} \theta_i DU_{i,t} + \sum_{i=1}^{l} \gamma_i DT_{i,t} + \beta_{1,0} x_{2,t} + \sum_{i=1}^{l} \beta_{1,i} DU_{i,t} x_{2,t} + \delta_{1,0} x_{1,t} + \sum_{i=1}^{l} \delta_{1,i} DU_{i,t} x_{1,t} + u_t,$$

(12)

with $l = 1$ or $l = 2$ depending on the case. First, for the cases of Greece and the US the BIC information criterion selects the specification that includes two structural breaks, so we will focus on these estimates. As described above, to assess the existence of fiscal policy sustainability we have to analyze the value of the estimated parameters $\beta_{1,i}$ and $\delta_{1,i}$, $i = 0, 1, 2$. When surpluses have prevailed during the i-th regime analyzed, $i = 0, \ldots, l + 1$, $\sum_{j=0}^{i} \beta_{1,j}$ is larger than one, whereas when deficits have been abundant, $\sum_{j=0}^{i} \beta_{1,j}$ is smaller than 1. For the assessment of fiscal sustainability, in the first case, $\sum_{j=0}^{i} \delta_{1,j}$ should be positive to guarantee sustainability, whereas in the second case $\sum_{j=0}^{i} \delta_{1,j}$ should be negative, $i = 0, \ldots, l + 1$.

For Canada the parameter of the second layer of cointegration is only statistically significant after the second structural break – i.e., the $\delta_{1,2}$ parameter – and has a positive sign. Concerning the parameters of the first layer of cointegration, we can see that $\hat{\beta}_{1,0}$ and $\sum_{j=0}^{1} \hat{\beta}_{1,j}$ are smaller than one – these parameters define the cointegrating vector in the first two sub-periods, respectively – a situation that satisfies the weak sustainability condition. For the third regime we have that $\sum_{j=0}^{2} \hat{\beta}_{1,j}$ is larger than one and $\hat{\delta}_{1,2}$ is positive, a situation that is also compatible with the sustainability of the Canadian fiscal policy after 2005. From the comparison with the first layer results, in that case we found strong sustainability after 2000 and weaker evidence before that date. Moreover, the value of the β_i parameters in the different regimes remains similar (in both cases with a final aggregate value of 1 or slightly above).

In the case of Greece, $\hat{\delta}_{1,0}$ is not statistical significant, but $\hat{\delta}_{1,1}$ and $\hat{\delta}_{1,2}$ are positive and statistical significant. As for the first layer of cointegration parameter values in Table 7, we observe that all $\hat{\beta}_{1,i}$, $i = 0, 1, 2$, are statistical significant. The estimated value of $\hat{\beta}_{1,1} = 0.717$ indicates that the fiscal policy of Greece was weakly sustainable up to 1983. However the negative values of $\sum_{j=0}^{1} \hat{\beta}_{1,j} = -1.911$ and $\sum_{j=0}^{2} \hat{\beta}_{1,j} = -0.076$ together with the positive values of $\hat{\delta}_{1,1} = 11.271$ and $\sum_{j=1}^{2} \hat{\delta}_{1,j} = 0.185$ indicate that the fiscal policy was unsustainable

16

in Greece after 1983. If we compare these results with those obtained in the previous analysis – see Table 6 – there seems to exist a contradiction. In the first cointegration layer and after 1997 fiscal policy in Greece was sustainable. This particular case can serve as an example of the importance of the multicointegration analysis: looking for sustainability between the flows would lead to the conclusion that the Greek finances were stable, as real revenues had been above expenditures for some time – see Figure 1 in the companion appendix. However, the real debt was growing so quickly that, in reality, even after maintaining the surpluses, the debt path was unsustainable. We should note that we have found one structural break in the mid-nineties (in 1994 in the model with one break and in 1995 in the two-break model), possibly related to the fiscal efforts to converge and become an EMU member, only achieved in 2000.

As for Italy, $\hat{\delta}_{1,0}$ is again non-significant, although $\hat{\delta}_{1,1}$ and $\hat{\delta}_{1,2}$ are statistical significant. The parameters of the first layer cointegrating vector are statistical significant in the three regimes. In the first regime the value of $\hat{\beta}_{1,1} = 0.763$ indicates a weakly sustainable fiscal policy. For the other regimes we have $\sum_{j=0}^{1} \hat{\beta}_{1,j} = 0.97$ and $\sum_{j=0}^{2} \hat{\beta}_{1,j} = 0.676$ together with $\hat{\delta}_{1,1} = -0.879$ and $\sum_{j=1}^{2} \hat{\delta}_{1,j} = -0.087$, situations that are compatible with fiscal sustainability. Looking at the cointegration results in the DOLS estimation of Table 6, the only significant $\beta_{1,i}$ parameter is $\hat{\beta}_{1,0} = 0.75$, that is, almost the same value as the one obtained in the multicointegration analysis. The conclusion is the same: weak sustainability, as the debt has been under surveillance by the Italian fiscal authorities since the eighties. In fact, looking at Figure 1 in the companion appendix, we can attribute the 1996 structural break to the stabilization of the Italian debt in an effort to fulfill the Maastricht convergence criteria.

Concerning Japan, $\hat{\delta}_{1,0}$ and $\hat{\delta}_{1,1}$ are not statistical significant, so that no multicointegration relation exists until 2000. As for the first layer cointegrating vector, we can see that the $\hat{\beta}_{1,i}$, $i = 0, 1, 2$, are statistical significant. In the first regime we have $\hat{\beta}_{1,0} > 1$ indicating that sustainability is fulfilled – in fact, the coefficient is larger than one, which indicates that the Japanese government was increasing his wealth in this period. In the second regime we have $\sum_{j=0}^{1} \hat{\beta}_{1,j} = 0.272$ with $\hat{\delta}_{1,1} = -0.083$, a situation that indicates fiscal sustainability. Finally, in the third regime we have $\sum_{j=0}^{2} \hat{\beta}_{1,j} = 0.041$ with $\sum_{j=1}^{2} \hat{\delta}_{1,j} = -0.609$, which also satisfy the fiscal sustainability conditions. As in the previous case, there are no contradictions between the results obtained in the cointegration (first layer) and multicointegration analysis: during the first regime $\hat{\beta}_{1,0}$ is larger than one and afterwards the values of the estimated parameters are compatible with the sustainability.

The case of the UK is similar to the Japanese one, as $\hat{\delta}_{1,0}$ and $\hat{\delta}_{1,1}$ are not statistical significant, so that no multicointegration relation exists until 1993. In the first regime we have that $\hat{\beta}_{1,0} > 1$ indicating that sustainability is met – as mentioned above, a coefficient larger than one would indicate that the government is increasing its wealth. After 1980 the parameter of interest is $\sum_{j=0}^{1} \hat{\beta}_{1,j} = 0.433$, which satisfies the condition of weak sustainability. Finally, for the third regime we have $\sum_{j=0}^{2} \hat{\beta}_{1,j} = 0.581$ and $\hat{\delta}_{1,2} > 0$, a situation that indicates that weak fiscal sustainability is met, although there is not a second layer of sustainability. Therefore, the country is able to maintain (weak) sustainability, but the monetary authorities do not really take their decisions looking at the evolution of the stock of debt. The most plausible explanation for this behaviour is related to the role of the UK in the international financial markets and the

ability to finance their macroeconomic imbalances through foreign capital. Even if the stock of debt over GDP has attained relatively high levels, larger than the Spanish one in 2011, the UK has not suffered from sovereign debt problems, as De Grauwe (2012) points out. De Grauwe (2012) argues that not being part of a monetary union allows British authorities to control their currency and therefore avoiding the difficulties that other countries with similar macroeconomic disequilibria have to face in a monetary union. Also in this case, the evidence is similar in the two analyses – see Tables 6 and 7 – as the parameters corresponding to the second layer are not significant in part of the sample. The conclusion of weak sustainability is confirmed using both approaches.

For the US we can see that all parameters involved in the first and second layers of cointegration are statistically significant. In the first regime the values of $\hat{\beta}_{1,0} < 1$ and $\hat{\delta}_{1,0} < 0$ indicate that the fiscal policy was sustainable. The second regime covers from 1987 till 1995, with the parameters $\sum_{j=0}^{1} \hat{\beta}_{1,j} = -0.363$ and $\sum_{j=1}^{1} \hat{\delta}_{1,j} = 5.077$ that clearly violates the sustainability of the fiscal policy. Finally, in the third regime we have $\sum_{j=0}^{2} \hat{\beta}_{1,j} = 0.533$ and $\sum_{j=1}^{2} \hat{\delta}_{1,j} = -2.297$, values that indicate that fiscal sustainability is met from 1996 on. From the comparison with the first cointegration results presented in Table 6, strong sustainability was then obtained between 1998 and 2001. We also find sustainability during that period. The main difference is, as in the case of Greece, that the period 1987-95 was weakly sustainable then, whereas once we introduce the debt into the analysis, the multicointegration analysis detects an unsustainable period, later corrected after a structural change (detected in the mid-nineties with both methodologies).

The estimation of the parameters of the deterministic component is also relevant in this context. Similar to the analysis above, the slope of the time trend indicates whether the cumulated revenues are above or below the cumulated expenditures. For Canada, the slope of the time trend increased from 43.254 to 133.64 ($= 43.254 + 90.386$), which indicates a clear predominance of a net inflow of money, and then reduced to -9.943 ($= 43.254 + 90.386 - 143.583$). For Greece we observe a positive slope in all regimes, although a deceleration is observed in the last one. Italy has a positive slope in the first regime, which is counteracted in the second regime resulting in a strong negative value, but changing to a positive value in the third regime. For Japan, we have a negative slope in the first regime, which is changed to a positive value for the second and third regimes. A similar situation is found in the UK, with an initial negative slope, turned into a positive value in the second and third regimes – although there is a marginal reduction of the slope in the third regime, the magnitude of the slope is still positive. Finally, for the US we have a strong positive magnitude of the slope for the first two regimes, which is slightly reduced on the third one. The general picture that can be drawn shows a clear tendency for cumulated revenues to be above the cumulated expenditures, something that adds to the fiscal sustainability.

5   Conclusions

In this paper we have applied a unified framework to test for fiscal sustainability to a group of 17 OECD countries that includes the larger EMU members as well as some non-EMU economies,

such as the US, the UK, Japan and Canada, for the period 1970-2012. Within this framework we not only test for the long-run relationship between revenues and expenditures, but also analyze the stock-flow mechanism relating deficit and debt.

From an econometric point of view, we contribute to previous literature in several respects. First, our approach also considers the presence of structural breaks in the variables as well as in the relationships linking them to provide an improved specification of the institutional changes that this type of variables are prone to suffer. Second, we use a testing framework based on I(2) stochastic processes and implement several types of cointegration and multicointegration test statistics. Moreover, we introduce an extension of the test statistics in Gregory and Hansen (1996), Carrion-i-Silvestre and Sansó (2006) and Berenguer-Rico and Carrion-i-Silvestre (2011) to allow for up to two regime shifts.

The results confirm the existence of a cointegration relationship between revenues and expenditures. However, these relationships have been specified with structural changes in the majority of cases. These changes not only affect the level but also the parameters of the long-run equilibrium relationship. This means that we have found evidence in favour of the weak or ex-post version of sustainability and, therefore, that governments were compelled to adjust their fiscal policies in order to fulfill their long-run budgetary constraints. This is the case, for example, of the US in the eighties and the EU countries during the nineties. Moreover, the value of the estimated cointegration vector points as well to a weak fiscal sustainability, signaling possible problems to market the debt by all the countries involved in the study (excepting Canada). This is particularly true for the last sample period in all cases studied. As a consequence, our results show that not only the deficit but the debt level has to be monitored by public authorities, which implies the study of the relationship between both variables. Therefore, in addition, we study the stock-flow mechanism linking deficit and debt accounting for the structural changes detected in the pre-testing stages of the analysis.

The results obtained using the more complex stock-flow mechanism point to the presence of a deeper long-run relationship for six out of seventeen cases that we analyze. Moreover, in four of these cases the estimated parameters lead to conclude that fiscal policy is sustainable – i.e., what we have called second layer fiscal sustainability. The exceptions is Greece where even considering the structural breaks (and therefore the fiscal policy changes implemented), does not have a sustainable fiscal policy at the end of the sample.

We have also been able to compare the results of the two approaches (cointegration versus multicointegration) in the case of the six countries where the second-layer relationship was significant. Four countries (Canada, Italy, Japan and the UK) display similar results, whereas in two cases, unsustainability was detected – in Greece, from 1983 onwards, and in the US, between 1987 and 1995.

Therefore, this study shows the importance of analyzing the interaction between fiscal deficit and debt stock accumulation, an issue that has been normally neglected when studying fiscal sustainability. The current eurozone debt crisis exemplifies the critical role of this servo control mechanism in the context of highly integrated and globalized financial markets.

References

AFONSO, A. (1995): "Fiscal sustainability: the unpleasant European case," *FinanzArchiv*, 61, 19–44.

AFONSO, A., L. AGNELLO, D. FURCERI, AND R. M. SOUSA (2011): "Assessing long-term fiscal developments: a new approach," *Journal of International Money and Finance*, 30, 130–146.

AFONSO, A., AND R. M. SOUSA (2011): "What are the effects of fiscal policy on asset markets?," *Economic Modelling*, 28, 1871–1890.

ANDREWS, D. W. K., AND W. PLOBERGER (1994): "Optimal Tests When a Nuisance Parameter is Present Only Under the Alternative," *Econometrica*, 62, 1383–1414.

BERENGUER-RICO, V., AND J. CARRION-I-SILVESTRE (2011): "Regime Shifts in Stock-Flow I(2)-I(1) Systems: The Case of US Fiscal Sustainability," *Journal of Applied Econometrics*, 26, 298–321.

BOHN, H. (1998): "The Behaviour of US Public Debt and Deficits," *The Quarterly Journal of Economics*, 113, 949–963.

——— (2005): "The Sustainability of Fiscal Policy in the United States," Discussion Paper 1446, CESifo, Munchen.

——— (2007): "Are Stationary and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint?," *Journal of Monetary Economics*, 54, 1837–1847.

CARRION-I-SILVESTRE, J. L., D. KIM, AND P. PERRON (2009): "GLS-Based Unit Root Tests with Multiple Structural Breaks Both under the Null and the Alternative Hypotheses," *Econometric Theory*, 25, 1754–1792.

CARRION-I-SILVESTRE, J. L., AND A. SANSÓ (2006): "Testing the Null of Cointegration with Structural Break," *Oxford Bulletin of Economics and Statistics*, 68, 623–646.

CARRION-I-SILVESTRE, J. L., A. SANSÓ, AND M. ARTÍS (2004): "Raíces Unitarias Y Cambios Estructurales En Las Macromagnitudes Españolas," *Revista de Economía Aplicada*, 35, 5–27.

CLEMENTE, J., A. MONTAÑÉS, AND M. REYES (1998): "Testing for a Unit Root in Variables with a Double Change in the Mean," *Economics Letters*, 59, 175–182.

DE GRAUWE, P. (2012): "The Governance of a Fragile Eurozone," *The Australian Economic Review*, 45, 255–268.

ELLIOT, G., T. J. ROTHENBERG, AND J. H. STOCK (1996): "Efficient Tests for an Autoregressive Unit Root," *Econometrica*, 64, 813–836.

ENGLE, R. F., AND C. W. J. GRANGER (1987): "Cointegration and Error Correction: Representation, Estimation, and Testing," *Econometrica*, 55, 251–276.

ENGLE, R. F., AND S. YOO (1991): "Cointegrated Economic Time Series: An Overview with New Results," in *Long-Run Economic Relationships*, ed. by R. F. Engle, and C. W. J. Granger. Oxford University Press.

ENGSTED, T., J. GONZALO, AND N. HALDRUP (1997): "Testing for multicointegration," *Economics Letters*, 56, 259–266.

ESCARIO, R., M. GADEA, AND M. SABATÉ (2012): "Multicointegration, seignorage and fiscal sustainability. Spain 1857-2000," *Journal of Policy Modeling*, 34(2), 270–283.

EUROPEAN CENTRAL BANK, E. (2011): "Ensuring Fiscal Sustainability in the euro area," *Monthly Bulletin*, April, 61–77.

GALI, J., AND R. PERIOTTI (2003): "Fiscal Policy and Monetary Integration in Europe," Discussion Paper 9773, NBER.

GHYSELS, E., AND P. PERRON (1996): "The effect of linear filters on dynamic time series with structural change," *Journal of Econometrics*, 70, 69–97.

GRANGER, C. W. J., AND T. H. LEE (1989): "Investigation of Production, Sales and Inventory Relations Using Multicointegration and Non-Symmetric Error Correction Models," *Journal of Applied Econometrics*, 4 (supplement), S145–S159.

GREGORY, A. W., AND B. E. HANSEN (1996): "Residual-based Tests for Cointegration in Models with Regime Shifts," *Journal of Econometrics*, 70, 99–126.

HAKKIO, C., AND M. RUSH (1991): "Is the Budget Deficit Too Large?," *Economic Inquiry*, pp. 429–445.

HALDRUP, N. (1994): "The Asymptotics of Single-Equation Cointegration Regressions with I(1) and I(2) Variables," *Journal of Econometrics*, 63, 151–181.

HAMILTON, J., AND M. FLAVIN (1986): "On the Limitations of Government Borrowing: A Framework for Empirical Testing," *American Economic Review*, 76, 808–816.

HARVEY, D., S. LEYBOURNE, AND A. TAYLOR (2013): "Testing for Unit Root in the Possible Presence of Multiple Trend Breaks Using Minimum Dickey-Fuller Statistics," *Journal of Econometrics*, forthcoming.

HAUG, A. (1991): "Cointegration and Government Borrowing Constraints: Evidence for the U.S.," *Journal of Business & Economic Statistics*, 9, 97–101.

LUMSDAINE, R. L., AND D. H. PAPELL (1997): "Multiple Trend Breaks and the Unit Root Hypothesis," *Review of Economics and Statistics*, 79, 212–218.

MARTIN, G. (2000): "U.S. Deficit Sustainability: A New Approach Based on Multiple Endogeneous Breaks," *Journal of Applied Econometrics*, 15, 83–105.

NG, S., AND P. PERRON (2001): "Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power," *Econometrica*, 69(6), 1519–1554.

PERRON, P., AND G. RODRÍGUEZ (2003): "GLS Detrending, Efficient Unit Root Tests and Structural Change," *Journal of Econometrics*, 115, 1–27.

PERRON, P., AND T. VOGELSANG (1992): "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business & Economic Statistics*, 10(3), 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.

PHILLIPS, P. C. B., AND S. OULIARIS (1990): "Asymptotic Properties of Residual Based Tests for Cointegration," *Econometrica*, 58, 165–193.

QUINTOS, C. E. (1995): "Sustainability of the Deficit Process with Structural Shifts," *Journal of Business & Economic Statistics*, 13, 409–418.

TRACLET, V. (2004): "Monetary and Fiscal Policies in Canada: Some Interesting Principles for EMU?," Discussion Paper 28, Bank of Canada.

TREHAN, B., AND C. WALSH (1988): "Common Trends, the Government Budget Constraint, and Revenue Smoothing," *Journal of Economic Dynamics and Control*, 12, 425–444.

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

WILCOX, D. (1989): "The Sustainability of Government Deficits: Implications of the Present Value Borrowing Constraint," *Journal of Money, Credit and Banking*, 54, 1837–1847.

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 & Economic Statistics*, 10(3), 251–270.

Table 1: Sustainability conditions

	Conditions	
Hamilton & Flavin (1986)	$R_t - GE_t \sim I(0)$	$B_t \sim I(0)$
Wilcox (1989)	$R_t - GE_t \sim I(0)$	$B_t \sim I(0)$
		$B_t \sim I(1)$
Hakkio & Rush (1991)	$R_t \sim I(1), G_t \sim I(1)$	
	$R_t - G_t \sim I(0), CI(1,1)$	$B_t \sim I(1)$
Trehan & Walsh (1988)	$(R_t - GE_t) - \delta B_{t-1} \sim I(0)$	$B_t \sim I(1)$
Trehan & Walsh (1991)	$(R_t - GE_t) - \delta B_{t-1} \sim I(0)$	$(R_t - GE_t) \sim I(1), B_{t-1} \sim I(1)$
	or either $\rightarrow$	$(R_t - GE_t) \sim I(0), B_{t-1} \sim I(0)$
Quintos (1995) weak	$R_t \sim I(1), G_t \sim I(1)$	$0 < \beta_{1,0} < 1$
	$R_t - \beta_{1,0} G_t \sim I(1)$	no cointegration, $B_t \sim I(2)$
	$R_t \sim I(1), G_t \sim I(1)$	$0 < \beta_{1,0} < 1$
	$R_t - \beta_{1,0} G_t \sim I(0)$	cointegration, $B_t \sim I(1)$
strong	$R_t \sim I(1), G_t \sim I(1)$	$\beta_{1,0} = 1$
	$R_t - \beta_{1,0} G_t \sim I(0)$	cointegration, $B_t \sim I(1)$
Bohn (1998, 2007)	$R_t \sim I(m_R), G_t \sim I(m_G)$	$m_R \geq 0, \ m_G \geq 0, \ \text{may } m_R \neq m_G$
"absurdly" weak		$B_t \sim I(m)$
weak		$B_t \sim I(2)$
strong		$B_t \sim I(1)$
Berenguer-Rico and	$R_t \sim I(1), G_t \sim I(1)$	$\sum_{j=1}^{t} R_j \sim I(2), \ \sum_{j=1}^{t} G_j \sim I(2)$
Carrion-i-Silvestre(2011)	$\sum_{j=1}^{t} R_j - \sum_{j=1}^{t} G_j = -B_t \sim I(1)$	$\Delta B_t \sim I(0)$
	$R_t - G_t \sim I(0)$	

Table 2: Unit root tests of Ng and Perron (2001)

		Z_α	MZ_α	MSB	ADF	P_T	MP_T	MZ_t
AUT	Rev	-7.26	-6.20	0.28	-2.06	16.03	14.69	-1.76
	Exp	-17.07	-13.43	0.19	-3.28**	6.88	6.83	-2.58
	Debt	-6.97	-5.99	0.28	-1.93	16.22	15.14	-1.66
BEL	Rev	-6.14	-5.50	0.29	-1.77	16.53	16.37	-1.59
	Debt	-6.11	-6.04	0.28	-1.77	14.62	15.05	-1.71
CAN	Rev	-9.00	-7.75	0.25	-2.26	11.85	11.81	-1.94
	Debt	-3.14	-2.62	0.43	-1.35	34.50	33.97	-1.12
DEN	Exp	-3.45	-3.12	0.36	-1.24	26.40	26.25	-1.12
	Debt	-9.14	-8.94	0.23	-2.22	10.29	10.25	-2.10
FIN	Rev	-17.82**	-13.98	0.19	-3.36**	6.37	6.55	-2.64
FRA	Rev	-8.15	-7.27	0.25	-2.02	12.45	12.71	-1.81
	Debt	-7.74	-6.67	0.22	-1.07	15.28	13.81	-1.48
GRE	Exp	-12.93	-12.62	0.19	-2.32	7.58	7.54	-2.45
	Debt	-39.91**	-39.49**	0.11**	-3.63**	2.80**	2.77**	-4.36**
IRE	Rev	-3.98	-3.56	0.37	-1.05	25.50	25.20	-1.31
ITA	Debt	-7.34	-7.31	0.25	-1.58	12.06	12.57	-1.85
JAP	Rev	-6.14	-5.00	0.32	-1.94	19.02	18.20	-1.58
NLD	Rev	-14.92	-11.75	0.21	-3.08**	8.69	7.76	-2.42
	Debt	-3.00	-2.77	0.42	-1.25	31.43	32.11	-1.15
POR	Rev	-8.32	-7.34	0.24	-2.00	12.73	12.67	-1.77
	Exp	-9.00	-8.03	0.22	-1.97	11.53	11.99	-1.76
SPA	Rev	-7.20	-6.28	0.27	-1.93	15.24	14.48	-1.68
	Exp	-8.57	-6.95	0.27	-2.28	15.67	13.12	-1.85

Rev indicates revenues and Exp expenditures. ** denotes rejection of the null hypothesis of unit root at the 5 % level of significance

Table 3: Zivot and Andrews (1992), Lumsdaine and Papell (1997), and Carrion-i-Silvestre et al. (2009) unit root test with structural breaks

		Zivot-Andrews One structural break		Lumsdaine-Papell Two structural breaks			CKP Up to three structural breaks									
		ADF	$\hat{T}_1$	ADF	$\hat{T}_1$	$\hat{T}_2$	P_T	MP_T	ADF	Z_α	MZ_α	MSB	MZ_t	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
BEL	Exp	-2.57	2005	-6.41	1980	2000	14.11	13.62	-3.34	-17.65	-13.94	0.19	-2.64	1980	2004	
CAN	Exp	-5.59*	1989	-6.90	1989	2004	15.71	14.34	-3.34	-17.29	-13.19	0.19	-2.55	1989	1997	
DEN	Rev	-6.53**	2003	-7.13*	1982	2004	18.62	16.94	-3.56	-19.56	-14.82	0.18	-2.69	1984	1991	2007
FIN	Exp	-4.49	1989	-7.57**	1990	2007	17.96	15.85	-3.04	-15.10	-12.37	0.20	-2.49	1989	2000	
	Debt	-4.88	1990	-8.06**	1991	2006	23.83	20.92	-2.06	-8.03	-7.15	0.26	-1.83	1990		
FRA	Exp	-6.25**	1980	-6.99	1981	2009	18.62	17.42	-3.72	-20.76	-15.58	0.18	-2.79	1976	1985	1992
GER	Rev	-5.73*	1990	-7.69**	1990	2000	17.88	16.30	-3.19	-15.65	-11.87	0.20	-2.42	1990	1999	
	Exp	-7.20**	1990	-7.19*	1988	1994	16.16	13.83	-2.87	-15.86	-12.89	0.20	-2.53	1990	1994	
	Debt	-4.77	1994	-6.96	1988	1994	16.25	15.44	-3.79	-20.74	-14.73	0.18	-2.69	1988	1994	2008
GRE	Rev	-3.27	2005	-5.49	1980	2008	18.16	15.23	-2.94	-14.14	-11.53	0.21	-2.40	1983	2008	
IRE	Exp	-3.61	2001	-4.90	1985	1995	10.60	10.49	-2.50	-34.62**	-13.70	0.19	-2.60	2000	2008	
	Debt	-3.31	2009	-10.66**	1992	2005	19.91	18.24	-2.26	-15.23	-9.71	0.22	-2.14	1988	2007	
ITA	Rev	-3.47	1994	-5.23	1981	1996	15.67	14.22	-3.15	-15.58	-12.04	0.20	-2.43	1993	2006	
	Exp	-5.48	1990	-6.03	1979	1991	25.88	22.94	-3.18	-16.32	-12.51	0.19	-2.44	1977	1992	2000
JAP	Exp	-4.45	1997	-9.73**	1997	2009	18.96	18.28	-2.74	-17.42	-10.11	0.22	-2.19	1997	2004	
	Debt	-4.73	1990	-7.53**	1993	2003	20.62	18.98	-2.69	-11.86	-9.43	0.23	-2.14	1992	2004	
NLD	Exp	-5.41	1995	-6.02	1977	1995	16.38	15.34	-2.55	-14.76	-10.60	0.21	-2.27	1995	2007	
POR	Debt	-2.79	2006	-4.75	1996	2006	16.97	14.77	-3.88*	-21.20	-14.83	0.18	-2.72	1991	1996	2008
SPA	Debt	-2.53	2002	-5.13	1981	2006	16.66	16.53	-3.98	-23.23	-16.80	0.17	-2.88	1980	1992	2005
SWE	Rev	-4.71	2008	-5.47	1986	2006	16.09	15.00	-4.05*	-23.43	-16.56	0.17	-2.86	1985	1992	2007
	Exp	-8.80**	1992	-10.64**	1981	1992	13.37	12.97	-2.91	-13.85	-11.31	0.21	-2.38	1992	1992	
	Debt	-4.58	1999	-5.59	1986	1999	29.25	26.67	-2.80	-13.05	-10.84	0.21	-2.33	1980	1990	1999
UK	Rev	-4.26	1991	-6.02	1991	2008	16.13	14.22	-2.98	-14.56	-11.92	0.20	-2.44	1990	2008	
	Exp	-4.57	1995	-5.48	1999	2007	19.44	18.46	-3.20	-17.00	-13.38	0.19	-2.52	1984	1999	2006
	Debt	-3.46	1993	-7.42*	1994	2004	21.85	21.12	-2.25	-9.16	-8.10	0.25	-1.99	1991	2007	
US	Rev	-4.75	1994	-6.60	1981	1997	12.25	12.13	-3.95*	-22.87	-16.50	0.17	-2.85	1993	2000	2007
	Exp	-4.21	1996	-4.97	1984	2002	17.21	14.93	-4.22*	-25.92	-17.91	0.16	-2.91	1981	1995	2007
	Debt	-4.02	2005	-5.86	1987	2003	25.28	20.43	-3.01	-15.16	-12.40	0.20	-2.46	1982	1997	2007

Notes: The null hypothesis of these statistics is that the variable is $I(1)$, whereas the alternative hypothesis is that it is $I(0)$. The critical values at the 5 and 10% level of significances for Model C of Zivot and Andrews (1992) with one structural break are -5.87 and -5.49, respectively. For the Model CC of Lumsdaine and Papell (1997) with two structural breaks they are -7.45 and -7.04, respectively. These critical values are obtained by simulation using $T = 43$ with 10,000 replications. ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 4: MDF unit root test statistic and local break magnitudes in Harvey et al. (2013) with up to three structural breaks, where the number of breaks is estimated using the BIC information criterion

| | Revenues | | | | Expenditures | | | | Debt | | | |
| | MDF | Local break magnitudes | | | MDF | Local break magnitudes | | | MDF | Local break magnitudes | | |
		$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$		$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$		$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
AUT	-2.825	2.390			-5.841**	6.870	-9.641	7.265	-4.252	10.066	-18.001	33.407
BEL	-1.998	-0.708			-7.107**	-14.030	5.638	6.592	-3.760	2.378	-6.126	6.907
CAN	-2.707	-0.137			-3.942	-13.589	9.424		-7.710**	10.527	-36.213	33.588
DEN	-2.724	-6.635			-2.535	-2.115			-3.971	3.204	-4.175	
FIN	-3.419	-1.630			-4.700	5.203	-13.113	14.662	-3.545	6.949	-9.097	7.972
FRA	-3.001	-3.507			-3.405	3.986	-4.356		-3.778	9.077	-6.857	18.590
GER	-2.028	-1.085			-4.000	4.447	-6.252		-4.491	16.696	-15.644	7.482
GRE	-2.096	1.153			-2.916	6.739	-8.126	3.384	-3.731	4.012	-7.105	17.808
IRE	-2.478	-0.842			-5.054**	2.607			-4.819*	-2.690	-1.976	35.141
ITA	-3.153	-4.233			-3.260	5.023	-5.558	-4.404	-3.949	19.750	-40.836	22.011
JAP	-2.705	-2.087			-3.387	-5.619	6.258		-3.627	4.453	16.207	-9.474
NLD	-3.658*	-1.332			-4.910*	-4.436	2.833	4.307	-4.804	-8.141	8.232	3.082
POR	-2.455	-1.882			-3.189	-4.992			-3.875	-2.223	14.114	
SPA	-2.495	-2.725			-2.797	2.837			-4.136	4.648	8.943	
SWE	-3.694*	-3.681			-3.756	4.875	-7.379		-3.451	-3.446		
UK	-2.843	2.198			-2.775	6.828	-5.462		-3.510	2.754	16.270	
US	-2.682	-0.949			-2.775	3.671			-4.204	3.304	-2.435	17.765

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 5: Gregory and Hansen (1996) cointegration tests statistics and Carrion-i-Silvestre and Berenguer-Rico (2011) multicointegration test statistics

	Gregory and Hansen														Berenguer-Rico and Carrion-i-Silvestre						
	Model C							Model C/S													
	One break			Two breaks				One break			Two breaks				One break			Two breaks			
	ADF	k	$\hat{T}_1$	ADF	k	$\hat{T}_1$	$\hat{T}_2$	ADF	k	$\hat{T}_1$	ADF	k	$\hat{T}_1$	$\hat{T}_2$	ADF	k	$\hat{T}_1$	ADF	k	$\hat{T}_1$	$\hat{T}_2$
AUT	-5.13**	2	1996	-5.134*	2	1996	2002	-5.40**	1	1998	-6.621*	1	1978	1998	-6.046	1	1997	-7.656	1	1986	2002
BEL	-3.60	0	1995	-3.599	0	1995	2001	-3.72	0	1995	-6.625*	5	1987	1999	-4.588	4	1991	-7.055	4	1994	2008
CAN	-4.69**	1	1996	-4.691	1	1996	2001	-4.76*	1	1996	-6.137	3	1987	1995	-5.127	1	1996	-8.631**	1	1991	2001
DEN	-4.12	1	2001	-4.062	1	1991	1997	-4.86*	2	1999	-5.801	2	1982	1999	-5.343	2	1987	-7.260	1	1988	2006
FIN	-4.37*	1	1999	-4.373	1	1999	2005	-4.72*	1	1999	-6.139	1	1987	1997	-5.297	1	1993	-7.370	1	1989	2001
FRA	-4.30	1	2006	-3.831	1	1995	2001	-5.64**	3	1999	-6.494	3	1991	2002	-5.937	3	1999	-7.793	3	1984	2004
GER	-5.22**	0	1996	-5.215**	0	1996	2002	-5.24**	0	1997	-6.514	0	1997	2006	-6.390	1	1991	-6.991	1	1983	1997
GRE	-5.66**	1	1996	-5.656**	1	1996	2002	-5.68**	1	1996	-6.695*	1	1993	2000	-7.293**	1	1995	-8.114*	1	1998	2005
IRE	-3.14	0	1994	-3.138	0	1994	2000	-4.54	1	1999	-6.970**	2	1987	2002	-4.174	2	1993	-6.722	4	1995	2008
ITA	-5.32**	1	1994	-5.324**	1	1994	2000	-5.33**	1	1994	-6.634*	1	1992	2000	-5.536	1	1993	-8.393**	1	1993	2005
JAP	-3.49	0	1993	-3.485	0	1993	1997	-3.39	0	1993	-6.753*	2	1988	2004	-4.954	1	1983	-8.203*	1	1983	1997
NLD	-6.10**	3	1995	-6.097**	3	1995	2001	-5.50**	3	1995	-7.643**	1	1997	2005	-4.957	1	1994	-6.691	1	1984	1999
POR	-5.07**	0	1995	-5.074*	0	1995	2001	-5.21**	0	1995	-7.054**	1	1978	1996	-4.494	5	1993	-5.839	3	1988	2003
SPA	-4.58*	1	2005	-4.073	1	1993	1999	-4.95**	1	2003	-6.302	1	1993	2004	-5.155	5	1996	-7.251	5	1982	2002
SWE	-4.99**	1	1999	-4.994*	1	1999	2005	-4.97**	1	1999	-6.377	3	1984	1997	-5.281	3	2003	-7.227	4	1990	1998
UK	-4.22	1	2006	-3.891	1	1990	1996	-4.79*	1	1999	-6.420	2	1984	1996	-5.896	2	2005	-8.427**	1	1988	2001
US	-5.43**	1	2005	-4.800	1	1990	1996	-6.76**	1	1996	-7.433**	2	1996	2003	-9.682**	1	1995	-10.283**	1	1990	2000

Notes: For the test statistics in Gregory and Hansen (1996), with one structural break the critical values at the 5 and 10% levels of significance from Table 1 in Gregory and Hansen (1996) are, respectively, -4.61 and -4.34 for Model C and -4.95 and -4.68 for Model C/S for the ADF and $Z_t(\tau)$ statistics, and -40.48 and -36.19 for Model C and -47.04 and -41.85 for Model C/S for the $Z_\alpha(\tau)$ statistic. For two structural breaks, the critical values are obtained by simulation using $T = 43$ observations and 10,000 replications. The estimated critical values in this case are -5.169 (5%) and -4.856 (10%) for Model C, and -6.918 (5%) and -6.55 (10%) for Model C/S. For the test statistics in Berenguer-Rico and Carrion-i-Silvestre (2011) the critical values at the 5 and 10% levels of significance from Table II in Berenguer-Rico and Carrion-i-Silvestre (2011) are, respectively, -6.97 and -6.65 for the one structural break case, whereas they are -8.33 and -7.95 for the two structural breaks case (obtained by simulation in this paper). ** and * denote rejection of the null hypothesis of no cointegration at the 5 and 10% levels of significance, respectively

Table 6: Revenue and expenditure cointegration relationship with up to two structural breaks. DOLS estimates

	$R_t = \mu + \sum_{i=1}^{2} \theta_i DU_{i,t} + \beta_{1,0} G_t + \sum_{i=1}^{2} \beta_{1,i} DU_{i,t} G_t + u_t$						
	μ	θ_1	θ_2	$\beta_{1,0}$	$\beta_{1,1}$	$\beta_{1,2}$	$\hat{T}_1$	$\hat{T}_2$
AUT	5.35	2.44		0.92			1990	
	(2.67)	(1.91)		(33.65)				
BEL	12.86	-125.63	162.68	0.70	1.20	-1.29	1988	2000
	(2.78)	(-11.74)	(11.66)	(16.04)	(12.40)	(-11.53)		
CAN	34.64	-1739.15	1630.40	0.82	5.57	-5.12	1997	2000
	(6.04)	(-4.04)	(3.75)	(44.92)	(4.14)	(-3.77)		
DEN	7.29	111.75		0.96	-1.07		2003	
	(3.59)	(6.60)		(42.53)	(-6.49)			
FIN	7.46	5.60		0.89			2003	
	(3.66)	(3.69)		(31.14)				
FRA	63.07	268.29		0.85	-0.30		1998	
	(5.98)	(5.96)		(48.37)	(-5.46)			
GER	53.92	-67.39	110.25	0.85			1987	1990
	(1.31)	(-3.84)	(6.39)	(12.67)				
GRE	1.88	8.90	-9.49	0.65	0.00	0.36	1985	1997
	(2.30)	(1.22)	(-1.05)	(6.55)	(0.02)	(2.07)		
IRE	-1.58	-10.57	33.24	0.81	0.57	-0.95	1978	2000
	(-1.48)	(-6.85)	(12.48)	(7.97)	(5.10)	(-12.99)		
ITA	11.44	37.25	58.78	0.75			1991	1996
	(1.87)	(5.38)	(11.52)	(51.09)				
JAP	-520.10	985.93	4015.54	1.34	-0.78	-1.87	1992	2004
	(-5.18)	(3.43)	(4.26)	(15.62)	(-4.42)	(-4.03)		
NLD	-26.87	-94.37	210.31	1.01	0.54	-1.10	1988	1993
	(-4.81)	(-3.65)	(7.70)	(30.29)	(3.52)	(-7.07)		
POR	0.01	-10.95	85.41	0.78	0.32	-1.45	1990	2004
	(0.02)	(-4.71)	(4.26)	(31.01)	(6.16)	(-4.40)		
SPA	6.98	-32.93	696.90	0.91	0.07	-1.86	1988	2006
	(2.48)	(-4.23)	(7.28)	(20.93)	(1.68)	(-7.51)		
SWE	-19.42	-12.19	-22.24	1.35			1978	1989
	(-3.36)	(-3.48)	(-4.97)	(18.59)				
UK	29.98	-55.50	49.03	0.87			1977	1981
	(1.95)	(-4.00)	(4.39)	(20.86)				
US	150.56	-1436.36	2205.77	0.94	0.48	-0.65	1997	2001
	(2.29)	(-1.82)	(2.72)	(42.94)	(2.17)	(-2.94)		

Note: t-ratio test statistics between parentheses

Table 7: DOLS estimation of the multicointegration relationship

	α	μ	θ_1	γ_1	θ_2	γ_2	$\delta_{1,0}$	$\beta_{1,0}$	$\delta_{1,1}$	$\beta_{1,1}$	$\delta_{1,2}$	$\beta_{1,2}$	$\hat{T}_1$	$\hat{T}_2$	BIC
CAN	-0.174	43.254	806.062	90.386	-3781.662	-143.583	-0.122	0.664	-0.469	-0.238	0.957	0.620	1994	2005	4.714
	(-0.005)	(16.220)	(2.223)	(2.779)	(-5.476)	(-4.089)	(-0.344)	(39.539)	(-0.998)	(-2.286)	(1.807)	(5.633)			
GRE	-1.118	1.301	849.654	145.477	-1314.914	-69.830	-0.059	0.717	11.271	-2.628	-11.086	1.835	1983	1995	1.939
	(-1.438)	(10.368)	(4.737)	(5.995)	(-2.988)	(-2.705)	(-0.840)	(116.365)	(6.845)	(-5.488)	(-6.725)	(3.725)			
	0.545	0.937	-19.120	14.878			-0.440	0.774	0.599	-0.026			1994		2.191
	(0.368)	(4.050)	(-1.150)	(7.964)			(-5.051)	(93.342)	(2.765)	(-0.895)					
ITA	-1.064	5.830	-302.944	-87.025	2262.252	234.070	0.158	0.763	-0.879	0.207	0.792	-0.294	1975	1996	5.475
	(-0.072)	(1.629)	(-2.004)	(-4.834)	(5.107)	(9.589)	(0.739)	(47.693)	(-3.244)	(6.035)	(3.375)	(-7.074)			
JAP	-51.046	-711.128	17449.983	1763.497	9590.372	751.257	0.024	1.558	-0.083	-1.286	-0.526	-0.231	1988	2000	8.050
	(-0.173)	(-24.882)	(44.577)	(36.372)	(4.375)	(5.870)	(0.081)	(51.085)	(-0.288)	(-34.400)	(-4.188)	(-3.664)			
UK	267.545	-255.468	6924.578	512.059	-2580.982	-27.102	0.731	1.408	-0.394	-0.975	1.110	0.148	1980	1993	6.974
	(1.042)	(-6.029)	(10.498)	(10.027)	(-3.147)	(-0.820)	(1.115)	(14.613)	(-0.760)	(-9.578)	(2.960)	(2.127)			
US	1976.392	205.934	48197.655	4063.949	-47217.729	-2283.157	-1.691	0.867	6.768	-1.230	-7.374	0.896	1986	1995	9.667
	(6.096)	(10.986)	(6.608)	(7.304)	(-2.903)	(-2.616)	(-6.241)	(99.687)	(5.895)	(-6.642)	(-5.455)	(3.655)			
	2875.505	176.665	17843.934	1426.492			-2.457	0.905	0.057	-0.303			1995		9.803
	(7.337)	(7.104)	(12.934)	(12.999)			(-7.193)	(88.696)	(0.082)	(-6.562)					

29

Companion appendix of the paper "The Relationship between Debt Level and Fiscal Sustainability in OECD countries: A Regime-shift Multicointegration Approach"[*]

Mariam Camarero
Jaume I University

Josep Lluís Carrion-i-Silvestre[†]
University of Barcelona

Cecilio Tamarit
University of Valencia

September 5, 2013

Figure 1: Real revenue, expenditure and debt

Figure 1 (cont.): Real revenue, expenditure and debt

Figure 1 (cont.): Real revenue, expenditure and debt

Table 1: Perron and Vogelsang unit root test for the first difference of time series

	Revenues		Expenditures		Debt	
	$t_{\hat{\alpha}}$	$\hat{T}_1$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$t_{\hat{\alpha}}$	$\hat{T}_1$
AUT	-5.19**	1991	-8.66**	2003	-5.67**	2006
BEL	-7.42**	2007	-9.72**	1980	-3.18	1992
CAN	-4.88**	1999	-5.02**	1991	-6.74**	1994
DEN	-5.84**	2004	-5.69**	1993	-3.50	1992
FIN	-7.08**	2008	-5.21**	1990	-4.11	1991
FRA	-7.42**	2008	-5.79**	1983	-4.77**	2006
GER	-7.16**	1990	-5.37**	1994	-5.40**	1994
GRE	-5.13**	2008	-4.28	2007	-5.78**	2006
IRE	-4.65*	1994	-4.85**	2006	-6.54**	2004
ITA	-6.06**	1991	-5.58**	1993	-3.80	1993
JAP	-5.11**	1989	-10.19**	1995	-3.48	1992
NLD	-6.10**	2007	-7.43**	2008	-5.91**	2005
POR	-8.07**	2008	-4.24	2007	-6.37**	2007
SPA	-6.13**	2008	-7.27**	2008	-5.38**	2007
SWE	-6.76**	1990	-9.21**	1992	-4.48*	1997
UK	-6.33**	2008	-4.94**	1999	-4.90**	2006
US	-5.79**	2006	-2.92	2000	-6.29**	2006

Notes: Perron and Vogelsang (1992) test statistic, which accounts for one structural break under the null and alternative hypotheses affecting the level the time series. Critical values at the 5 and 10% level of significances are -4.76 and -4.42, respectively, which are obtained from Table 2 in Perron and Vogelsang (1992). ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 2: Clemente, Montanes and Reyes unit root test for the first difference of time series

	Revenues			Expenditures			Debt		
	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$
AUT	-5.26	1979	1998	-8.50**	1985	2003	-5.66*	1985	2006
BEL	-7.81**	1988	2007	-11.00**	1980	2008	-3.59	1987	2005
CAN	-5.00	1980	1999	-4.40	1978	2000	-2.85	1992	2010
DEN	-6.48**	1982	2004	-5.68*	1981	2008	-4.12	1992	2010
FIN	-7.06**	1990	2008	-5.28	1990	2008	-4.05	1991	2010
FRA	-7.56**	1978	2008	-6.53**	1983	2009	-5.25	1989	2007
GER	-7.61**	1990	2008	-3.31	1990	2010	-4.76	1988	2006
GRE	-6.72**	1985	2008	-4.53	1985	2007	-5.85*	1983	2006
IRE	-4.37	1979	1997	-5.25	1987	2005	-9.87**	1988	2006
ITA	-6.36**	1979	2006	-4.53	1987	2005	-4.00	1978	1996
JAP	-5.61*	1989	2008	-9.65**	1992	2010	-3.34	1992	2010
NLD	-7.15**	1979	2007	-7.55**	1979	2008	-6.47**	1987	2005
POR	-8.22**	1984	2008	-4.66	1989	2010	-6.52**	1984	2007
SPA	-9.92**	1981	2006	-7.98**	1980	2008	-6.55**	1981	2007
SWE	-7.47**	1990	2008	-8.97**	1992	2010	-4.40	1978	1997
UK	-6.86**	1990	2008	-4.91	1985	2007	-6.99**	1988	2006
US	-6.19**	1981	2006	-5.32	1981	2002	-6.93**	1980	2006

Notes: Test statistic in Clemente, Montañés, and Reyes (1998), which accounts for two structural breaks under the null and alternative hypotheses affecting the level of the time series. Critical values at the 5 and 10% level of significances are -5.88 and -5.52, respectively, which are obtained from Table 1 in Clemente, Montañés, and Reyes (1998). ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 3: Lumsdaine and Papell unit root test (Model AAn) for the first difference of time series

	Revenues			Expenditures			Debt		
	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$
AUT	-5.96**	1989	2002	-8.68**	2004	2006	-6.38**	2001	2007
BEL	-7.94**	1978	2008	-10.22**	1981	2007	-5.07	1993	2007
CAN	-5.96**	1996	2000	-6.35**	1995	2000	-8.46**	1996	2007
DEN	-6.40**	1981	2005	-5.96**	1992	1994	-4.60	1993	2007
FIN	-6.82**	1990	1993	-5.85**	1993	2001	-6.86**	1990	1997
FRA	-6.91**	2007	2009	-6.14**	1984	1989	-5.98**	1991	2008
GER	-8.21**	1990	1992	-7.54**	1990	1995	-5.52*	1989	1997
GRE	-6.11**	1981	2007	-5.65**	1982	2009	-7.03**	1999	2007
IRE	-6.76**	2002	2006	-5.99**	1999	2006	-11.60**	1991	2007
ITA	-6.70**	1992	2007	-6.05**	1993	2000	-5.39*	1983	1994
JAP	-5.67**	1991	2003	-13.31**	1998	2007	-5.84**	1994	2006
NLD	-6.77**	1979	2005	-9.21**	1995	1997	-8.41**	1993	2007
POR	-7.62**	2007	2009	-4.57	1990	1993	-7.04**	2000	2008
SPA	-9.30**	1982	2007	-8.65**	2005	2009	-6.69**	1982	2008
SWE	-7.15**	2004	2006	-7.48**	1991	1995	-4.83	1990	1998
UK	-7.80**	1996	2008	-6.12**	2000	2008	-8.16**	1991	2007
US	-6.85**	1996	2000	-5.26	1982	2002	-7.03**	1981	2007

Notes: Model AA of Lumsdaine and Papell (1997) without a time trend is used, which accounts for two structural breaks under the alternative hypothesis of I(0) affecting the level of the time series. Critical values at the 5 and 10% level of significances are -5.63 and -5.29, respectively, which are obtained from Table A.2 in Carrion-i-Silvestre, Sansó, and Artís (2004). ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 4: Perron-Yabu test statistics to test the null hypothesis of no structural breaks against the alternative hypothesis of one structural break

	Revenues			Expenditures			Debt		
	I	II	III	I	II	III	I	II	III
AUT	-0.19	-0.32	0.05	-0.15	-0.53	-0.08	0.78	-0.49	0.83
BEL	-0.21	-0.21	0.16	0.79	1.60**	5.04**	0.24	-0.20	0.73
CAN	0.18	-0.52	0.27	0.92*	1.51**	15.72**	-0.15	0.69	1.37
DEN	-0.22	2.54**	6.67**	-0.04	0.32	1.05	0.15	-0.17	0.77
FIN	-0.04	-0.56	0.03	1.09*	-0.45	1.34	3.42**	-0.23	4.09**
FRA	-0.40	-0.31	-0.05	0.11	10.83**	10.78**	-0.19	0.50	1.08
GER	18.39**	0.36	18.77**	6.60**	0.14	7.29**	8.73**	0.95*	22.84**
GRE	0.96*	0.18	1.88	0.41	-0.41	0.65	0.25	-0.53	0.66
IRE	-0.36	-0.54	-0.26	-0.60	6.04**	5.01**	-0.40	5.77**	7.99**
ITA	-0.06	3.72**	4.96**	0.34	76.12**	44.68**	0.51	0.51	2.08
JAP	-0.10	0.29	0.83	6.14**	-0.23	9.24**	-0.38	25.13**	10.02**
NLD	0.15	-0.47	0.61	12.18**	-0.32	12.24**	-0.04	-0.13	0.38
POR	0.12	-0.35	0.47	-0.06	-0.38	0.20	0.33	3.81**	5.05**
SPA	-0.31	-0.53	-0.20	0.07	-0.22	0.33	0.21	5.87**	6.98**
SWE	1.92**	-0.25	2.61*	24.29**	0.00	30.34**	3.37**	4.17**	5.02**
UK	-0.33	0.36	2.44*	2.10**	0.46	7.40**	-0.25	4.46**	6.74**
US	2.11**	-0.52	2.18*	-0.51	0.80*	10.36**	-0.05	2.26**	4.83**

Notes: The columns labelled as I, II and III present the results for the specifications defined by Models I, II and III, respectively, in Perron and Yabu (2009). The critical values at the 5 and 10% levels of significance are, respectively, 1.33 and 0.91 (Model I), 1.28 and 0.74 (Model II) and 2.79 and 2.15 (Model III). ** and * denote rejection of the null hypothesis of no structural break at the 5 and 10% levels of significance, respectively

Table 5: Zivot and Andrews unit root test (Model C) for the level of time series

	Revenues		Expenditures		Debt	
	$t_{\hat{\alpha}}$	$\hat{T}_1$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$t_{\hat{\alpha}}$	$\hat{T}_1$
AUT	-4.13	1989	-4.55	1990	-3.81	2002
BEL	-3.91	1996	-2.57	2005	-4.62	1989
CAN	-3.95	1996	-5.59*	1989	-3.75	1991
DEN	-6.53**	2003	-3.73	1993	-3.78	1980
FIN	-4.26	1991	-4.49	1989	-4.88	1990
FRA	-3.79	2003	-6.25**	1980	-3.85	2005
GER	-5.73*	1990	-7.20**	1990	-4.77	1994
GRE	-3.27	2005	-4.48	1991	-3.99	1988
IRE	-3.83	1992	-3.61	2001	-3.31	2009
ITA	-3.47	1994	-5.48	1990	-5.33	1991
JAP	-3.69	2008	-4.45	1997	-4.73	1990
NLD	-5.00	1994	-5.41	1995	-3.73	2008
POR	-3.82	1998	-5.35	1997	-2.79	2006
SPA	-4.41	2005	-3.72	1985	-2.53	2002
SWE	-4.71	2008	-8.80**	1992	-4.58	1999
UK	-4.26	1991	-4.57	1995	-3.46	1993
US	-4.75	1994	-4.21	1996	-4.02	2005

Notes: Model C of Zivot and Andrews (1992) is used, which accounts for one structural break under the alternative hypothesis of I(0) affecting the level and the slope of the time series. Critical values at the 5 and 10% level of significances are -5.87 and -5.49, respectively, which are obtained by simulation using $T = 43$ with 10,000 replications. ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 6: Lumsdaine and Papell unit root test (Model CC) for the level of time series

	Revenues			Expenditures			Debt		
	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$	$t_{\hat{\alpha}}$	$\hat{T}_1$	$\hat{T}_2$
AUT	-5.08	1977	1995	-5.54	1990	2004	-5.62	1998	2006
BEL	-5.12	1985	2008	-6.41	1980	2000	-5.33	1992	2002
CAN	-5.71	1981	2002	-6.90	1989	2004	-6.07	1994	2008
DEN	-7.13*	1982	2004	-5.22	1988	1997	-5.09	1993	2004
FIN	-5.80	1987	1999	-7.57**	1990	2007	-8.06**	1991	2006
FRA	-5.08	1985	2003	-6.99	1981	2009	-5.94	1994	2005
GER	-7.69**	1990	2000	-7.19*	1988	1994	-6.96	1988	1994
GRE	-5.49	1980	2008	-7.31*	1984	2008	-5.75	1992	2004
IRE	-9.65**	1996	2005	-4.90	1985	1995	-10.66**	1992	2005
ITA	-5.23	1981	1996	-6.03	1979	1991	-5.95	1978	1991
JAP	-7.68**	1990	2005	-9.73**	1997	2009	-7.53**	1993	2003
NLD	-7.04	1978	2001	-6.02	1977	1995	-5.12	1994	2005
POR	-5.72	1990	2005	-6.04	1995	2008	-4.75	1996	2006
SPA	-7.25*	1982	2004	-4.79	1978	1995	-5.13	1981	2006
SWE	-5.47	1986	2006	-10.64**	1981	1992	-5.59	1986	1999
UK	-6.02	1991	2008	-5.48	1999	2007	-7.42*	1994	2004
US	-6.60	1981	1997	-4.97	1984	2002	-5.86	1987	2003

Notes: Model CC of Lumsdaine and Papell (1997) is used, which accounts for two structural breaks under the alternative hypothesis of I(0) affecting the level and the slope of the time series. Critical values at the 5 and 10% level of significances are -7.45 and -7.04, respectively, which are obtained by simulation using $T = 43$ with 10,000 replications. ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 7: Unit root test with multiple structural breaks. Revenues

	P_T	MP_T	ADF	Z_α	MZ_α	MSB	MZ_t	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
AUT	15.90	15.79	-3.22	-16.57	-13.10	0.19	-2.55	1975	1989	
BEL	21.14	17.52	-3.40	-18.11	-14.16	0.19	-2.66	1978	1995	2008
CAN	11.16	10.83	-3.08	-15.65	-12.34	0.20	-2.43	1996		
DEN	18.62	16.94	-3.56	-19.56	-14.82	0.18	-2.69	1984	1991	2007
FIN	16.98	17.05	-3.09	-15.60	-12.70	0.20	-2.52	1990	1999	2008
FRA	10.36	10.47	-2.58	-11.52	-9.92	0.22	-2.22	2008		
GER	17.88	16.30	-3.19	-15.65	-11.87	0.20	-2.42	1990	1999	
GRE	18.16	15.23	-2.94	-14.14	-11.53	0.21	-2.40	1983	2008	
IRE	11.04	11.00	-4.45**	-27.44*	-18.47	0.16	-3.00	1993	2000	2007
ITA	15.67	14.22	-3.15	-15.58	-12.04	0.20	-2.43	1993	2006	
JAP	13.93	14.13	-4.28**	-23.33	-14.19	0.18	-2.60	1991	2005	
NLD	10.74	9.48	-3.02*	-14.49	-11.38	0.21	-2.37	2007		
POR	13.69	12.66	-3.36	-17.38	-13.30	0.19	-2.57	1991	2008	
SPA	12.42	12.33	-4.82**	-30.33*	-19.15	0.16	-3.05	1982	2001	2007
SWE	16.09	15.00	-4.05*	-23.43	-16.56	0.17	-2.86	1985	1992	2007
UK	16.13	14.22	-2.98	-14.56	-11.92	0.20	-2.44	1990	2008	
US	12.25	12.13	-3.95*	-22.87	-16.50	0.17	-2.85	1993	2000	2007

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 8: Unit root test with multiple structural breaks. Expenditures

	P_T	MP_T	ADF	Z_α	MZ_α	MSB	MZ_t	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
AUT	15.08	13.93	-4.02*	-23.55	-16.90	0.17	-2.89	1989	1996	2003
BEL	14.11	13.62	-3.34	-17.65	-13.94	0.19	-2.64	1980	2004	
CAN	15.71	14.34	-3.34	-17.29	-13.19	0.19	-2.55	1989	1997	
DEN	17.74	17.14	-3.04	-15.10	-12.36	0.20	-2.49	1978	1991	
FIN	17.96	15.85	-3.04	-15.10	-12.37	0.20	-2.49	1989	2000	
FRA	18.62	17.42	-3.72	-20.76	-15.58	0.18	-2.79	1976	1985	1992
GER	16.16	13.83	-2.87	-15.86	-12.89	0.20	-2.53	1990	1994	
GRE	24.23	21.03	-2.06	-8.41	-7.40	0.24	-1.81	1984		
IRE	10.60	10.49	-2.50	-34.62**	-13.70	0.19	-2.60	2000	2008	
ITA	25.88	22.94	-3.18	-16.32	-12.51	0.19	-2.44	1977	1992	2000
JAP	18.96	18.28	-2.74	-17.42	-10.11	0.22	-2.19	1997	2004	
NLD	16.38	15.34	-2.55	-14.76	-10.60	0.21	-2.27	1995	2007	
POR	12.53	12.82	-3.31	-17.48	-13.76	0.19	-2.60	1989	2007	
SPA	22.59	18.03	-3.40	-17.40	-12.94	0.20	-2.53	1986	1999	2008
SWE	13.37	12.97	-2.91	-13.85	-11.31	0.21	-2.38	1992		
UK	19.44	18.46	-3.20	-17.00	-13.38	0.19	-2.52	1984	1999	2006
US	17.21	14.93	-4.22*	-25.92	-17.91	0.16	-2.91	1981	1995	2007

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 9: Unit root test with multiple structural breaks. Debt

	P_T	MP_T	ADF	Z_α	MZ_α	MSB	MZ_t	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
AUT	22.47	17.04	-3.60	-17.23	-11.45	0.21	-2.39	1993	1997	2007
BEL	30.14	27.61	-2.54	-11.06	-9.49	0.23	-2.18	1979	1991	2007
CAN	16.47	14.26	-4.35**	-24.09	-14.80	0.18	-2.67	1993	2005	2008
DEN	29.48	25.32	-2.91	-13.45	-10.67	0.22	-2.31	1980	1991	2006
FIN	23.83	20.92	-2.06	-8.03	-7.15	0.26	-1.83	1990		
FRA	21.29	18.72	-3.31	-16.40	-12.05	0.20	-2.43	1992	2002	2007
GER	16.25	15.44	-3.79	-20.74	-14.73	0.18	-2.69	1988	1994	2008
GRE	28.88	27.33	-2.49	-11.06	-9.53	0.23	-2.15	1981	1992	2007
IRE	19.91	18.24	-2.26	-15.23	-9.71	0.22	-2.14	1988	2007	
ITA	24.52	19.96	-3.42	-18.94	-14.60	0.18	-2.63	1979	1992	2000
JAP	20.62	18.98	-2.69	-11.86	-9.43	0.23	-2.14	1992	2004	
NLD	28.09	24.06	-3.33	-15.00	-10.13	0.22	-2.25	1990	1998	2007
POR	16.97	14.77	-3.88*	-21.20	-14.83	0.18	-2.72	1991	1996	2008
SPA	16.66	16.53	-3.98	-23.23	-16.80	0.17	-2.88	1980	1992	2005
SWE	29.25	26.67	-2.80	-13.05	-10.84	0.21	-2.33	1980	1990	1999
UK	21.85	21.12	-2.25	-9.16	-8.10	0.25	-1.99	1991	2007	
US	25.28	20.43	-3.01	-15.16	-12.40	0.20	-2.46	1982	1997	2007

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 10: Unit root test statistic in Harvey, Leybourne and Taylor (2012) with one structural break

	Revenues		Expenditures		Debt	
	MDF	$\hat{T}_1$	MDF	$\hat{T}_1$	MDF	$\hat{T}_1$
AUT	-2.825	1977	-3.612*	1975	-2.169	1996
BEL	-1.998	2006	-2.050	2004	-2.517	1990
CAN	-2.707	1981	-2.630	1990	-2.428	1995
DEN	-2.724	2006	-2.535	1995	-3.834*	1993
FIN	-3.419	2006	-1.743	2005	-2.575	1980
FRA	-3.001	2004	-2.983	1982	-2.571	1981
GER	-2.028	2001	-2.586	2000	-2.623	1994
GRE	-2.096	2006	-2.653	1995	-3.688*	1996
IRE	-2.478	1983	-5.054**	2000	-1.487	1975
ITA	-3.153	1996	-2.397	1990	-2.827	2006
JAP	-2.705	1989	-2.153	1996	-2.415	2003
NLD	-3.658*	1975	-2.507	1978	-1.330	2005
POR	-2.455	1996	-3.189	2004	-2.619	2005
SPA	-2.495	2002	-2.797	1976	-2.170	1994
SWE	-3.694*	2006	-3.186	1996	-3.451	1990
UK	-2.843	1994	-2.337	1998	-3.337	1998
US	-2.682	1981	-2.775	1999	-2.556	1996

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 11: Unit root test statistic in Harvey, Leybourne and Taylor (2012) with two structural breaks

	Revenues			Expenditures			Debt		
	MDF	$\hat{T}_1$	$\hat{T}_2$	MDF	$\hat{T}_1$	$\hat{T}_2$	MDF	$\hat{T}_1$	$\hat{T}_2$
AUT	-3.434	1977	2000	-4.492*	1988	1994	-3.631	2000	2006
BEL	-2.635	1997	2003	-3.826	1981	1999	-3.489	1982	1988
CAN	-3.235	1994	1999	-3.942	1993	1998	-5.176**	1997	2007
DEN	-2.933	1976	2006	-2.818	1983	1997	-3.971	1980	1993
FIN	-3.663	2000	2006	-2.718	1995	2001	-3.157	1984	1993
FRA	-3.307	1980	2006	-3.405	1975	1981	-2.401	1987	2007
GER	-3.726	1988	1995	-4.000	1987	1996	-3.164	1990	1996
GRE	-3.256	1980	2006	-2.729	1987	1996	-3.821	1994	2002
IRE	-3.635	1997	2006	-5.486**	1981	1999	-4.451*	1993	2006
ITA	-3.513	1982	1998	-2.714	1983	1991	-3.575	1983	1989
JAP	-3.691	1991	2001	-3.387	2000	2006	-3.493	1994	2004
NLD	-4.044	1977	2002	-4.119	1982	2004	-4.305*	1993	2005
POR	-4.897**	1986	2003	-3.243	1993	2006	-3.875	1989	2005
SPA	-4.128	1978	2006	-3.227	1980	1989	-4.136	1981	2006
SWE	-4.152	1975	2006	-3.756	1988	1994	-3.662	1981	1989
UK	-3.920	1995	2006	-2.775	1999	2006	-3.510	1975	2002
US	-3.530	1993	1999	-3.186	1978	2000	-3.566	1995	2001

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 12: Unit root test statistic in Harvey, Leybourne and Taylor (2012) with three structural breaks

	Revenues				Expenditures				Debt			
	MDF	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	MDF	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$	MDF	$\hat{T}_1$	$\hat{T}_2$	$\hat{T}_3$
AUT	-4.345	1976	1988	1999	-5.841**	1989	1995	2006	-4.252	1975	2000	2006
BEL	-6.263**	1978	1992	2002	-7.107**	1982	1988	2001	-3.760	1982	1988	2003
CAN	-3.526	1981	1994	1999	-4.240	1980	1992	1998	-7.710**	1991	1996	2008
DEN	-3.099	1976	1985	2006	-3.135	1975	1981	1997	-4.051	1976	1993	2000
FIN	-4.231	1990	1996	2006	-4.700	1988	1994	2000	-3.545	1989	1995	2006
FRA	-3.885	1982	1994	2003	-3.674	1975	1981	1988	-3.778	1990	1996	2007
GER	-4.372	1978	1987	1995	-7.188**	1977	1988	1994	-4.491	1991	1997	2003
GRE	-3.806	1982	1988	2006	-2.916	1981	1991	1998	-3.731	1975	1982	2001
IRE	-3.938	1994	2000	2006	-5.679**	1984	1990	2000	-4.819*	1988	1995	2006
ITA	-3.936	1975	1981	1998	-3.260	1982	1992	1998	-3.949	1982	1995	2006
JAP	-4.734	1991	2003	2007	-11.035**	1989	1998	2007	-3.627	1975	1994	2004
NLD	-4.083	1978	1984	2003	-4.910*	1979	1994	2000	-4.804	1979	1993	2005
POR	-5.018*	1986	1993	2003	-3.467	1980	1986	2003	-4.726	1994	2000	2006
SPA	-4.498	1978	2000	2006	-3.539	1980	1991	1997	-5.094**	1983	1994	2006
SWE	-4.607	1980	1988	2006	-4.007	1989	1995	2001	-3.322	1975	1981	1988
UK	-4.493	1988	1994	2006	-3.165	1975	1999	2006	-3.552	1981	1994	2001
US	-3.739	1981	1993	1999	-3.651	1980	1991	1999	-4.204	1979	1990	2001

Notes: ** and * denote rejection of the null hypothesis of unit root at the 5 and 10% levels of significance, respectively

Table 13: Local break magnitudes in Harvey et al. (2012)

| | Revenues | | | | | | Expenditures | | | | | | Debt | | | | | |
| | $m=1$ | $m=2$ | | $m=3$ | | | $m=1$ | $m=2$ | | $m=3$ | | | $m=1$ | $m=2$ | | $m=3$ | | |
	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
AUT	2.390	4.297	-1.365	5.456	1.752	-3.729	1.111	2.808	-3.027	6.870	-9.641	7.265	0.962	-3.596	7.675	10.066	-18.001	33.407
BEL	-0.708	1.191	-2.669	-12.970	11.775	-13.813	-1.327	-6.967	3.462	-14.030	5.638	6.592	-1.770	-3.763	4.760	2.378	-6.126	6.907
CAN	-0.137	4.002	-5.706	2.839	5.051	-9.985	-2.584	-13.589	9.424	2.954	-24.016	16.477	-3.354	-8.985	10.115	10.527	-36.213	33.588
DEN	-6.635	2.351	-9.797	5.819	-1.709	-12.404	-2.115	2.503	-3.255	2.944	-3.417	-1.868	-1.715	3.204	-4.175	3.173	-4.222	1.626
FIN	-1.630	-2.210	2.237	-4.698	6.679	-7.933	0.316	-2.892	4.139	5.203	-13.113	14.662	0.407	2.973	-2.929	6.949	-9.097	7.972
FRA	-3.507	-0.681	-4.033	1.265	-1.385	-3.710	-1.068	3.986	-4.356	6.221	-8.383	2.161	4.753	1.580	5.983	9.077	-6.857	18.590
GER	-1.085	0.872	-2.439	7.571	-8.622	-1.293	-0.760	4.447	-6.252	-2.695	7.051	-8.444	1.695	8.313	-5.231	16.696	-15.644	7.482
GRE	1.153	3.389	-6.120	10.427	-4.593	-9.195	0.311	2.842	-2.621	6.739	-8.126	3.384	2.021	-2.118	7.167	4.012	-7.105	17.808
IRE	-0.842	11.409	-27.797	-0.667	12.753	-30.963	2.607	5.312	-4.078	-3.019	2.969	4.299	5.721	-2.429	18.660	-2.690	-1.976	35.141
ITA	-4.233	-0.095	-4.376	4.563	-4.218	-7.802	-2.864	-0.412	-2.957	5.023	-5.558	-4.404	-1.312	3.931	-6.056	19.750	-40.836	22.011
JAP	-2.087	-4.833	3.148	-14.015	17.773	-17.070	-2.095	-5.619	6.258	2.436	-10.546	9.697	5.230	10.709	-6.188	4.453	16.207	-9.474
NLD	-1.332	-0.842	1.553	-3.394	0.572	1.878	0.686	-0.957	2.621	-4.436	2.833	4.307	0.553	-2.013	5.105	-8.141	8.232	3.082
POR	-1.882	2.662	-2.659	6.788	-3.759	-10.360	-4.992	0.682	-5.380	2.098	-1.113	-4.845	4.233	-2.223	14.114	0.060	-2.539	15.286
SPA	-2.725	2.889	-10.143	8.188	2.113	-21.380	2.837	3.079	-1.712	5.883	-5.838	2.964	2.953	4.648	8.943	14.315	-5.037	26.346
SWE	-3.681	0.611	-2.422	7.386	-5.360	-9.564	-1.521	4.875	-7.379	1.547	4.269	-7.619	-3.446	1.890	-5.470	-2.728	5.291	-7.406
UK	2.198	4.911	-6.478	-3.426	10.947	-11.491	2.881	6.828	-5.462	-2.321	11.356	-9.167	7.003	2.754	16.270	4.696	-2.337	21.663
US	-0.949	3.652	-4.677	3.931	4.349	-8.987	3.671	1.502	4.192	5.389	-4.865	13.868	7.617	0.805	7.844	3.304	-2.435	17.765

14

Table 14: Local break magnitudes in Harvey et al. (2012) that end up in a valley

	Revenues						Expenditures						Debt					
	$m=1$	$m=2$		$m=3$			$m=1$	$m=2$		$m=3$			$m=1$	$m=2$		$m=3$		
	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
AUT	2.390	4.297	-1.365	5.456	1.752	-3.729	1.111	2.808	-3.027					-3.596				
BEL		1.191	-2.669				-1.327		3.462		5.638		-1.770	-3.763	4.760	2.378		
CAN		4.002	-5.706	2.839	5.051		-2.584			2.954			-3.354					
DEN		2.351		5.819	-1.709		-2.115	2.503	-3.255	2.944	-3.417	-1.868	-1.715	3.204	-4.175	3.173	-4.222	1.626
FIN	-1.630	-2.210	2.237	-4.698				-2.892	4.139	5.203				2.973	-2.929			
FRA	-3.507	1.265	-4.033		-1.385	-3.710	-1.068	3.986	-4.356			2.161	4.753	1.580	5.983			
GER	-1.085		-2.439			-1.293		4.447		-2.695			1.695		-5.231			
GRE	1.153	3.389			-4.593			2.842	-2.621			3.384	2.021	-2.118		4.012		
IRE							2.607	5.312	-4.078	-3.019	2.969	4.299	5.721	-2.429		-2.690	-1.976	
ITA	-4.233		-4.376	4.563	-4.218		-2.864		-2.957	5.023	-5.558	-4.404	-1.312	3.931				
JAP	-2.087	-4.833	3.148				-2.095	-5.619		2.436			5.230	0.000	0.000	4.453		3.082
NLD	-1.332		1.553	-3.394		1.878			2.621	-4.436	2.833	4.307		-2.013	5.105			
POR	-1.882	2.662	-2.659		-3.759		-4.992	3.079	-5.380	2.098	-1.113	-4.845	4.233	-2.223			-2.539	
SPA	-2.725	2.889			2.113		2.837	4.875	-1.712	5.883	-5.838	2.964	2.953	4.648			-5.037	
SWE	-3.681		-2.422		-5.360		-1.521			1.547	4.269		-3.446	1.890	-5.470	-2.728	5.291	
UK	2.198	4.911		-3.426			2.881		-5.462	-2.321				2.754		4.696	-2.337	
US		3.652	-4.677	3.931	4.349		3.671	1.502	4.192	5.389	-4.865					3.304	-2.435	

Table 15: Local break magnitudes in Harvey et al. (2012), where the number of breaks is estimated using the BIC information criterion

	Revenues			Expenditures			Debt		
	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
AUT	2.390			6.870	-9.641	7.265	10.066	-18.001	33.407
BEL	-0.708			-14.030	5.638	6.592	2.378	-6.126	6.907
CAN	-0.137			-13.589	9.424		10.527	-36.213	33.588
DEN	-6.635			-2.115			3.204	-4.175	
FIN	-1.630			5.203	-13.113	14.662	6.949	-9.097	7.972
FRA	-3.507			3.986	-4.356		9.077	-6.857	18.590
GER	-1.085			4.447	-6.252		16.696	-15.644	7.482
GRE	1.153			6.739	-8.126	3.384	4.012	-7.105	17.808
IRE	-0.842			2.607			-2.690	-1.976	35.141
ITA	-4.233			5.023	-5.558	-4.404	19.750	-40.836	22.011
JAP	-2.087			-5.619	6.258		4.453	16.207	-9.474
NLD	-1.332			-4.436	2.833	4.307	-8.141	8.232	3.082
POR	-1.882			-4.992			-2.223	14.114	
SPA	-2.725			2.837			4.648	8.943	
SWE	-3.681			4.875	-7.379		-3.446		
UK	2.198			6.828	-5.462		2.754	16.270	
US	-0.949			3.671			3.304	-2.435	17.765

16

Table 16: Local break magnitudes in Harvey et al. (2012), where the number of breaks is estimated using the BIC information criterion, with coefficients that end up in a valley

	Revenues			Expenditures			Debt		
	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$	$\hat{\kappa}_1$	$\hat{\kappa}_2$	$\hat{\kappa}_3$
AUT	2.390								
BEL					5.638		2.378		
CAN									
DEN	-1.630			-2.115			3.204	-4.175	
FIN	-3.507			5.203					
FRA	-1.085			3.986	-4.356				
GER	1.153			4.447					
GRE						3.384	4.012		
IRE	-4.233			2.607			-2.690	-1.976	
ITA	-2.087			5.023	-5.558	-4.404			
JAP	-1.332			-5.619			4.453		
NLD	-1.882			-4.436	2.833	4.307			3.082
POR	-2.725			-4.992			-2.223		
SPA	-3.681			2.837			4.648		
SWE				4.875			-3.446		
UK	2.198				-5.462		2.754		
US				3.671			3.304	-2.435	

Table 17: Carrion-i-Silvestre and Sansó SC^+ cointegration test statistic, allowing for one structural break

| | Panel A | | | | | | | | Panel B | | | | | | | |
| | Model An | | | | Model D | | | | Model An | | | | Model D | | | |
	SC_{An}^+	$\hat{T}_1$	Crit. val. 10%	5%	SC_D^+	$\hat{T}_1$	Crit. val. 10%	5%	SC_{An}^+	$\hat{T}_1$	Crit. val. 10%	5%	SC_D^+	$\hat{T}_1$	Crit. val. 10%	5%
AUT	0.05	1990	0.13	0.16	0.03	1998	0.13	0.17	0.08	2003	0.16	0.21	0.07	1991	0.10	0.13
BEL	0.09	1994	0.13	0.16	0.10	1994	0.11	0.14	0.07	1988	0.13	0.16	0.05	1989	0.10	0.13
CAN	0.07	1998	0.14	0.18	0.07	1997	0.13	0.17	0.05	1991	0.14	0.19	0.04	1998	0.11	0.14
DEN	0.05	2003	0.16	0.21	0.05	2003	0.15	0.20	0.12	1982	0.14	0.19	0.05	1993	0.10	0.13
FIN	0.03	2003	0.16	0.21	0.03	2003	0.15	0.20	0.05	1992	0.13	0.16	0.06	1992	0.10	0.13
FRA	0.08	1997	0.14	0.18	0.05	1998	0.13	0.17	0.04	1983	0.16	0.21	0.02	1986	0.15	0.21
GER	0.04	1990	0.13	0.16	0.04	1990	0.10	0.13	0.05	2003	0.16	0.21	0.05	2003	0.15	0.20
GRE	0.05	1998	0.14	0.18	0.06	1997	0.13	0.17	0.04	2002	0.16	0.21	0.03	1988	0.11	0.14
IRE	0.08	1989	0.13	0.16	0.09	2001	0.13	0.17	0.07	1999	0.14	0.18	0.07	2000	0.13	0.17
ITA	0.08	1996	0.13	0.16	0.09	1996	0.11	0.14	0.10	1980	0.14	0.19	0.04	1982	0.13	0.17
JAP	0.05	1992	0.14	0.19	0.07	2003	0.10	0.13	0.08	2003	0.13	0.16	0.06	2003	0.10	0.13
NLD	0.05	1990	0.13	0.16	0.04	1993	0.11	0.14	0.06	1993	0.13	0.16	0.06	1999	0.11	0.14
POR	0.08	1995	0.13	0.16	0.08	1995	0.11	0.14	0.07	1994	0.13	0.16	0.05	1996	0.11	0.14
SPA	0.09	1999	0.14	0.18	0.03	2000	0.13	0.17	0.06	1999	0.14	0.18	0.06	1992	0.10	0.13
SWE	0.05	1990	0.13	0.16	0.03	1989	0.10	0.13	0.10	1979	0.16	0.21	0.05	1992	0.10	0.13
UK	0.05	1981	0.14	0.19	0.04	1998	0.13	0.17	0.06	1980	0.14	0.19	0.06	1991	0.10	0.13
US	0.09	1997	0.14	0.18	0.03	1997	0.13	0.17	0.07	1999	0.14	0.18	0.04	1999	0.13	0.17

Notes: Cointegration test of Carrion-i-Silvestre and Sansó (2006) allowing for one structural break both under the null hypothesis of cointegration and the alternative hypothesis of no cointegration. In Model An the structural break only affects the h level of the relationship, whereas in Model D it also affects the cointegrating vector. ** and * denote rejection of the null hypothesis of cointegration at the 5 and 10% levels of significance, respectively

18

Table 18: Cointegration relationship with one structural break. DOLS estimates

| | Panel A | | | | | | | Panel B | | | | | | |
| | Model An | | | Model D | | | | Model An | | | Model D | | | |
	μ	θ	β_0	μ	θ	β_0	β_1	μ	θ	β_0	μ	θ	β_0	β_1
AUT	5.35	2.44	0.92	2.94	19.09	0.92	-0.14	-90.87	-8.54	2.13	-106.24	30.54	2.40	-0.43
	(2.67)	(1.91)	(33.65)	(1.71)	(2.28)	(36.72)	(-1.96)	(-33.09)	(-4.36)	(64.54)	(-20.91)	(3.25)	(32.12)	(-4.16)
BEL	-5.28	14.27	0.87	-3.40	7.97	0.86	0.05	135.04	71.39	0.90	-7.10	300.20	2.29	-2.22
	(-1.21)	(7.29)	(21.79)	(-0.58)	(0.61)	(16.21)	(0.49)	(7.13)	(8.79)	(4.86)	(-0.20)	(6.01)	(6.87)	(-4.72)
CAN	37.81	47.61	0.83	32.00	-77.02	0.82	0.37	81.93	71.85	1.71	-192.08	1248.19	2.78	-3.73
	(6.00)	(14.56)	(38.29)	(4.56)	(-1.89)	(36.23)	(3.09)	(1.06)	(2.14)	(5.29)	(-3.04)	(7.18)	(14.13)	(-7.16)
DEN	10.41	5.20	0.94	7.29	111.75	0.96	-1.07	0.63	64.84	0.02	-77.02	451.53	2.42	-5.37
	(3.14)	(3.79)	(29.17)	(3.59)	(6.60)	(42.53)	(-6.49)	(0.06)	(8.14)	(0.11)	(-9.78)	(11.25)	(21.59)	(-12.63)
FIN	7.46	5.60	0.89	8.43	41.30	0.89	-0.45	-7.84	26.92	0.62	-9.45	49.27	0.61	-0.31
	(3.66)	(3.69)	(31.14)	(4.42)	(3.07)	(31.23)	(-2.66)	(-3.79)	(9.54)	(8.90)	(-3.99)	(4.86)	(9.20)	(-2.27)
FRA	39.53	24.00	0.85	63.07	268.29	0.85	-0.30	-609.37	-114.88	2.41	789.10	-1472.41	-0.43	2.84
	(2.46)	(3.15)	(43.27)	(5.98)	(5.96)	(48.37)	(-5.46)	(-15.49)	(-7.36)	(74.98)	(3.96)	(-8.12)	(-1.19)	(7.86)
GER	176.29	124.40	0.65	181.42	111.66	0.64	0.02	-664.35	170.24	1.98	-649.71	-3534.53	1.97	3.71
	(5.49)	(5.97)	(12.38)	(3.92)	(1.33)	(7.93)	(0.16)	(-22.02)	(7.97)	(51.40)	(-25.70)	(-7.45)	(58.85)	(7.93)
GRE	-0.14	9.83	0.88	0.30	8.87	0.81	0.06	4.33	-18.47	1.38	2.53	36.48	2.24	-1.44
	(-0.16)	(8.36)	(35.34)	(0.40)	(1.98)	(38.66)	(0.83)	(1.79)	(-6.18)	(26.22)	(1.30)	(5.14)	(13.81)	(-8.84)
IRE	-1.40	4.46	0.89	-6.33	33.18	1.13	-0.86	9.42	-19.58	1.82	15.29	-51.94	1.70	0.93
	(-2.10)	(5.40)	(20.96)	(-7.75)	(7.26)	(34.77)	(-7.79)	(4.17)	(-5.98)	(17.12)	(4.73)	(-4.63)	(14.29)	(2.99)
ITA	9.44	72.96	0.80	8.08	135.59	0.80	-0.11	98.06	-243.00	2.57	301.83	-704.05	1.49	1.52
	(1.17)	(12.67)	(49.34)	(0.98)	(1.80)	(48.60)	(-0.83)	(2.61)	(-7.77)	(26.82)	(9.17)	(-10.89)	(11.67)	(8.59)
JAP	-289.83	-272.91	1.09	279.02	1649.19	0.65	-0.75	-3713.78	2332.68	6.07	-3186.88	-17892.8	5.82	9.99
	(-2.08)	(-3.74)	(10.72)	(2.34)	(1.51)	(9.85)	(-1.41)	(-4.98)	(7.89)	(15.53)	(-7.25)	(-6.83)	(24.18)	(7.95)
NLD	1.92	12.34	0.85	-24.32	115.15	0.99	-0.55	-178.61	-76.13	2.87	-217.90	70.46	2.65	-0.75
	(0.53)	(5.91)	(31.98)	(-4.59)	(6.70)	(35.00)	(-6.10)	(-5.16)	(-5.57)	(11.24)	(-10.72)	(1.57)	(20.68)	(-3.34)
POR	0.04	3.56	0.84	0.01	4.78	0.84	-0.02	-2.83	-12.44	1.74	-1.95	-54.50	1.57	0.85
	(0.06)	(4.66)	(40.78)	(0.02)	(1.44)	(38.39)	(-0.38)	(-0.83)	(-4.99)	(21.30)	(-0.71)	(-5.99)	(24.30)	(4.99)
SPA	7.50	9.54	0.80	5.75	90.29	0.82	-0.25	0.74	67.21	2.44	-2.22	-109.70	1.85	0.88
	(1.78)	(2.08)	(39.83)	(1.65)	(4.36)	(50.19)	(-4.08)	(0.05)	(4.85)	(40.35)	(-0.28)	(-4.50)	(27.74)	(9.55)
SWE	1.18	-11.75	1.10	-7.70	-41.67	1.10	0.22	25.62	39.34	0.57	-5.66	320.63	1.31	-2.34
	(0.29)	(-3.00)	(23.78)	(-1.37)	(-2.95)	(18.08)	(2.11)	(3.21)	(4.50)	(6.48)	(-0.92)	(8.89)	(15.06)	(-9.62)
UK	18.12	28.80	0.82	3.23	179.47	0.90	-0.21	-119.14	-160.07	1.69	1094.22	-1297.69	-1.08	2.70
	(1.04)	(2.50)	(18.23)	(0.09)	(2.81)	(12.94)	(-2.16)	(-2.51)	(-5.39)	(13.25)	(7.62)	(-7.84)	(-3.65)	(8.07)
US	391.64	283.66	0.93	191.35	1127.83	0.95	-0.24	-1851.67	-848.32	2.73	-1634.21	-2240.66	2.71	0.37
	(5.68)	(4.62)	(28.22)	(2.56)	(5.23)	(35.18)	(-4.02)	(-15.82)	(-7.79)	(61.88)	(-13.70)	(-5.20)	(71.65)	(3.31)

Notes: t-ratio test statistics between parentheses

19

Table 19: Revenue and expenditure relationship. Carrion-i-Silvestre and Sansó SC^+ cointegration test statistic, allowing for two structural breaks

| | Model An | | | | | Model D | | | | |
| | | | | Crit. val. | | | | | Crit. val. | |
	SC_{An}^+	$\hat{T}_1$	$\hat{T}_2$	10%	5%	SC_D^+	$\hat{T}_1$	$\hat{T}_2$	10%	5%
AUT	0.029	1998	2001	0.127	0.168	0.033	1999	2003	0.119	0.159
BEL	0.056	1990	1996	0.103	0.129	0.032	1988	2000	0.066	0.081
CAN	0.053	1995	1998	0.111	0.143	0.084	1997	2000	0.119	0.159
DEN	0.047	1993	2003	0.104	0.136	0.023	1980	2004	0.077	0.098
FIN	0.049	1984	2003	0.093	0.117	0.032	1978	2003	0.094	0.125
FRA	0.029	1985	2006	0.105	0.132	0.035	1985	2004	0.072	0.089
GER	0.022	1987	1990	0.104	0.130	0.022	1987	1990	0.086	0.108
GRE	0.035	1981	1997	0.084	0.101	0.039	1985	1997	0.066	0.081
IRE	0.041	1992	2006	0.104	0.129	0.043	1978	2000	0.076	0.097
ITA	0.043	1991	1996	0.103	0.129	0.034	1991	1996	0.086	0.108
JAP	0.046	1989	2005	0.088	0.106	0.041	1992	2004	0.066	0.080
NLD	0.033	1990	1999	0.092	0.115	0.044	1988	1993	0.071	0.088
POR	0.054	1983	2006	0.111	0.143	0.046	1990	2004	0.077	0.097
SPA	0.059	1982	1999	0.084	0.101	0.023	1988	2006	0.087	0.109
SWE	0.023	1978	1989	0.091	0.114	0.034	1989	2004	0.077	0.097
UK	0.050	1977	1981	0.127	0.169	0.056	1976	1986	0.093	0.124
US	0.038	1997	2001	0.127	0.168	0.027	1997	2001	0.119	0.159

Notes: Cointegration test of Carrion-i-Silvestre and Sansó (2006) allowing for two structural breaks both under the null hypothesis of cointegration and the alternative hypothesis of no cointegration. In Model An the structural breaks only affect theh level of the relationship, whereas in Model D they also affect the cointegrating vector. ** and * denote rejection of the null hypothesis of cointegration at the 5 and 10% levels of significance, respectively

Table 20: Revenue and expenditure cointegration relationship with two structural breaks. DOLS estimates

	Model An				Model D					
	μ	θ_1	θ_2	β_0	μ	θ_1	θ_2	β_0	β_1	β_2
AUT	2.39	5.22	-2.54	0.91	1.78	336.28	-303.64	0.93	-3.07	2.81
	(1.45)	(3.66)	(-2.15)	(37.58)	(1.18)	(3.15)	(-2.85)	(45.62)	(-3.12)	(2.87)
BEL	6.79	7.75	12.64	0.76	12.86	-125.63	162.68	0.70	1.20	-1.29
	(1.75)	(6.23)	(8.58)	(21.00)	(2.78)	(-11.74)	(11.66)	(16.04)	(12.40)	(-11.53)
CAN	33.57	16.63	36.14	0.80	34.64	-1739.15	1630.40	0.82	5.57	-5.12
	(5.10)	(3.28)	(8.70)	(39.40)	(6.04)	(-4.04)	(3.75)	(44.92)	(4.14)	(-3.77)
DEN	13.32	3.30	4.71	0.88	-2.41	6.13	134.30	1.21	-0.19	-1.30
	(3.95)	(2.18)	(3.62)	(21.13)	(-0.38)	(0.80)	(7.04)	(8.76)	(-1.31)	(-6.92)
FIN	11.30	3.77	7.92	0.78	-10.17	14.23	34.82	1.58	-0.64	-0.39
	(4.31)	(2.14)	(4.43)	(13.50)	(-1.03)	(1.44)	(2.20)	(4.32)	(-1.74)	(-1.91)
FRA	-28.22	-33.17	-41.04	1.01	-36.30	-9.87	1229.75	1.03	-0.04	-1.45
	(-1.88)	(-3.99)	(-4.42)	(40.55)	(-1.69)	(-0.44)	(6.34)	(25.69)	(-1.00)	(-6.42)
GER	53.92	-67.39	110.25	0.85	15.30	474.16	-436.16	0.92	-0.85	0.83
	(1.31)	(-3.84)	(6.39)	(12.67)	(0.30)	(1.90)	(-1.68)	(10.73)	(-2.17)	(2.12)
GRE	-0.49	-3.77	9.23	0.93	1.88	8.90	-9.49	0.65	0.00	0.36
	(-0.73)	(-2.86)	(8.70)	(27.52)	(2.30)	(1.22)	(-1.05)	(6.55)	(0.02)	(2.07)
IRE	-2.64	3.09	-8.82	0.98	-1.58	-10.57	33.24	0.81	0.57	-0.95
	(-4.06)	(5.02)	(-6.75)	(28.95)	(-1.48)	(-6.85)	(12.48)	(7.97)	(5.10)	(-12.99)
ITA	11.44	37.25	58.78	0.75	9.37	365.94	-207.63	0.76	-0.57	0.47
	(1.87)	(5.38)	(11.52)	(51.09)	(1.47)	(0.80)	(-0.45)	(50.44)	(-0.72)	(0.59)
JAP	1016.05	214.85	264.63	0.19	-520.10	985.93	4015.54	1.34	-0.78	-1.87
	(9.27)	(3.75)	(6.40)	(2.42)	(-5.18)	(3.43)	(4.26)	(15.62)	(-4.42)	(-4.03)
NLD	22.46	10.18	8.95	0.77	-26.87	-94.37	210.31	1.01	0.54	-1.10
	(5.60)	(6.05)	(5.07)	(29.14)	(-4.81)	(-3.65)	(7.70)	(30.29)	(3.52)	(-7.07)
POR	-4.41	-2.23	-2.25	1.02	0.01	-10.95	85.41	0.78	0.32	-1.45
	(-4.46)	(-2.57)	(-2.39)	(37.59)	(0.02)	(-4.71)	(4.26)	(31.01)	(6.16)	(-4.40)
SPA	12.07	16.15	18.29	0.74	6.98	-32.93	696.90	0.91	0.07	-1.86
	(2.99)	(2.94)	(3.63)	(25.63)	(2.48)	(-4.23)	(7.28)	(20.93)	(1.68)	(-7.51)
SWE	-19.42	-12.19	-22.24	1.35	-4.92	-21.01	380.54	1.10	0.08	-2.22
	(-3.36)	(-3.48)	(-4.97)	(18.59)	(-0.97)	(-1.26)	(2.48)	(20.58)	(0.62)	(-2.44)
UK	29.98	-55.50	49.03	0.87	869.51	-1315.41	437.21	-1.03	2.85	-0.91
	(1.95)	(-4.00)	(4.39)	(20.86)	(1.72)	(-2.53)	(5.04)	(-0.87)	(2.37)	(-5.00)
US	170.26	307.92	-257.89	0.93	150.56	-1436.36	2205.77	0.94	0.48	-0.65
	(2.45)	(5.90)	(-4.30)	(37.23)	(2.29)	(-1.82)	(2.72)	(42.94)	(2.17)	(-2.94)

Notes: t-ratio test statistics between parentheses

21

Table 21: Berenguer-Rico and Carrion-i-Silvestre ADF multicointegration test statistic with one and two structural breaks

	One break			Two breaks			
	ADF	k	$\hat{T}_1$	ADF	k	$\hat{T}_1$	$\hat{T}_2$
AUT	-6.046	1	1997	-7.656	1	1986	2002
BEL	-4.588	4	1991	-7.055	4	1994	2008
CAN	-5.127	1	1996	-8.631**	1	1991	2001
DEN	-5.343	2	1987	-7.260	1	1988	2006
FIN	-5.297	1	1993	-7.370	1	1989	2001
FRA	-5.937	3	1999	-7.793	3	1984	2004
GER	-6.390	1	1991	-6.991	1	1983	1997
GRE	-7.293**	1	1995	-8.114*	1	1998	2005
IRE	-4.174	2	1993	-6.722	4	1995	2008
ITA	-5.536	1	1993	-8.393**	1	1993	2005
JAP	-4.954	1	1983	-8.203*	1	1983	1997
NLD	-4.957	1	1994	-6.691	1	1984	1999
POR	-4.494	5	1993	-5.839	3	1988	2003
SPA	-5.155	5	1996	-7.251	5	1982	2002
SWE	-5.281	3	2003	-7.227	4	1990	1998
UK	-5.896	2	2005	-8.427**	1	1988	2001
US	-9.682**	1	1995	-10.283**	1	1990	2000

Notes: The critical values at the 5 and 10% levels of significance from Table II in Berenguer-Rico and Carrion-i-Silvestre (2011) are, respectively, -6.97 and -6.65 for the one structural break case, whereas they are -8.33 and -7.95 for the two structural breaks case (obtained by simulation in this paper). ** and * denote rejection of the null hypothesis of no cointegration at the 5 and 10% levels of significance, respectively

References

BERENGUER-RICO, V., AND J. CARRION-I-SILVESTRE (2011): "Regime Shifts in Stock-Flow I(2)-I(1) Systems: The Case of US Fiscal Sustainability," *Journal of Applied Econometrics*, 26, 298–321.

CARRION-I-SILVESTRE, J. L., AND A. SANSÓ (2006): "Testing the Null of Cointegration with Structural Break," *Oxford Bulletin of Economics and Statistics*, 68, 623–646.

CARRION-I-SILVESTRE, J. L., A. SANSÓ, AND M. ARTÍS (2004): "Raíces Unitarias Y Cambios Estructurales En Las Macromagnitudes Españolas," *Revista de Economía Aplicada*, 35, 5–27.

CLEMENTE, J., A. MONTAÑÉS, AND M. REYES (1998): "Testing for a Unit Root in Variables with a Double Change in the Mean," *Economics Letters*, 59, 175–182.

LUMSDAINE, R. L., AND D. H. PAPELL (1997): "Multiple Trend Breaks and the Unit Root Hypothesis," *Review of Economics and Statistics*, 79, 212–218.

PERRON, P., AND T. VOGELSANG (1992): "Nonstationarity and Level Shifts with an Application to Purchasing Power Parity," *Journal of Business & Economic Statistics*, 10(3), 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.

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 & Econom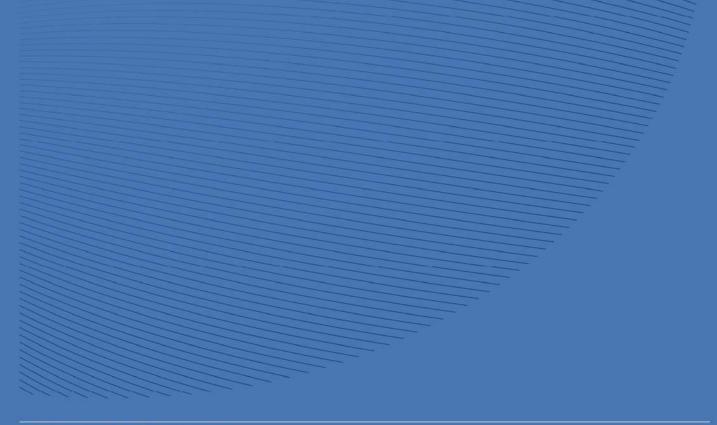ic Statistics*, 10(3), 251–270.