THE WILLIAM DAVIDSON INSTITUTE
AT THE UNIVERSITY OF MICHIGAN

What do we really know about fiscal sustainability in the EU? A panel data diagnostic

By: António Afonso and Christophe Rault

William Davidson Institute Working Paper Number 893
October 2007

What do we really know about fiscal sustainability in the EU? A panel data diagnostic*

António Afonso # *and Christophe Rault* $

2007

* We are grateful to Ad van Riet, Jürgen von Hagen and an anonymous referee for helpful comments and suggestions and to Simone Ruiz for assistance with the data. The opinions expressed herein are those of the authors and do not necessarily reflect those of the European Central Bank or the Eurosystem. Christophe Rault thanks the Fiscal Policies Division of the ECB for its hospitality.
# European Central Bank, Kaiserstraße 29, D-60311 Frankfurt am Main, Germany, email: antonio.afonso@ecb.int.
ISEG/TULisbon – Technical University of Lisbon, Department of Economics; UECE – Research Unit on Complexity and Economics, R. Miguel Lupi 20, 1249-078 Lisbon, Portugal, email: aafonso@iseg.utl.pt. UECE is supported by FCT (Fundação para a Ciência e a Tecnologia, Portugal), financed by ERDF and Portuguese funds.
$ LEO, University of Orleans, Rue de Blois-B.P.6739, 45067 Orléans Cedex 2, e-mail: chrault@hotmail.com, website: http://membres.lycos.fr/chrault/index.html.

Contents

Abstract

We assess the sustainability of public finances in the EU15 over the period 1970-2006 using stationarity and cointegration analysis. Specifically, we use panel unit root tests of the first and second generation allowing in some cases for structural breaks. We also apply modern panel cointegration techniques developed by Pedroni (1999, 2004), generalized by Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2007), to a structural long-run equation between general government expenditures and revenues. While estimations point to fiscal sustainability being an issue in some countries, fiscal policy was sustainable both for the EU15 panel set, and within sub-periods (1970-1991 and 1992-2006).

Keywords: intertemporal budget constraint, fiscal sustainability, EU, panel unit root, panel cointegration.

JEL Classification Numbers: C23, E62, H62, H63.

Non-technical summary

The sustainability of public finances is a key policy issue for the European Union (EU). Within the EU fiscal framework, fiscal discipline is an important support for the implementation of monetary policy, particularly in the case of the European and Monetary Union (EMU) countries. In EMU, the existence of sound fiscal policies is seen as a necessary objective for individual countries to pursue. For example, it is not possible to discard adverse responses from the financial markets when fiscal behaviour is deemed to be unsustainable. Indeed, the accumulation of government debt, following continued budgetary imbalances, may in the end trigger the need for higher long-term interest rates in order to place additional sovereign debt in the markets. Moreover, the Treaties governing the EU also require sustainable public finances. Countries are urged to comply with the budgetary requirements of EMU, by avoiding excessive deficits, keeping government debt levels below the 60 percent of GDP reference value, and respecting the requirements of the Stability and Growth Pact (SGP).

The aim of this paper is to examine the sustainability of public finances for the EU15 countries (covering the EU Member States before the 1 May 2004 enlargement) by applying recent advances in the econometrics of non-stationary panel data methods. The econometric literature on unit roots and cointegration testing has been expanding rapidly, and now distinguishes between the first generation tests developed on the assumption of cross-section independence (except for common time effects), and the second generation tests that allow, in a variety of forms and degrees, the dependence that might prevail across the different units in the panel. In the context of our paper, cross-dependence can mirror possible changes in the behaviour of fiscal authorities related to the signing of the EU Treaty in Maastricht on 7 February 1992, with the setting up of the fiscal convergence criteria that urged the EU countries to consolidate public finances in the run-up to the EMU on 1 January 1999, when most EU legacy currencies were replaced by the euro, and in the context of the SGP since then.

The econometric methods used in the paper to assess the sustainability of public finances in the EU15 rest upon (i) individual unit root tests allowing in some cases for structural breaks; (ii) first generation panel data integration tests that assume cross-sectional independence among panel units (apart from common time effects); (iii) two second generation panel data unit root tests that relax the assumption of cross-sectional independence; (iv) panel data unit root tests that enable to accommodate structural breaks, and (v) the panel data cointegration tests developed by Pedroni (1999, 2004) and generalized by Banerjee and Carrion-i-Silvestre (2006), and the bootstrap panel cointegration test by Westerlund and Edgerton (2007).

The results from these panel unit root tests, allowing for structural breaks, support the results of both the first and second generation panel data unit root tests, leading us to conclude that the first difference of stock of public debt series is integrated of order zero, thus indicating that the solvency condition would be satisfied for EU15 countries, which is a necessary condition for fiscal policy sustainability. Moreover, our results also show that general government expenditure and the revenue series are integrated of order one.

Even if the results of the analysis may question fiscal sustainability in some cases when taken individually, it is nevertheless true that the tests point to the solvency of government public finances when considering the EU15 panel data set. This is an advantage of the panel approach, since the time series dimension of the data is not that long for individual countries. Even if there is no single fiscal policy in the EU, the panel sustainability of public finances indicated by our results is relevant in a context of EU countries seeking to pursue sound fiscal policy behaviour within the Stability and Growth Pact framework.

Interestingly, the panel cointegration results for the entire 1970-2006 period allow us to draw the conclusion that a long-run relationship does exist between general government revenue and expenditure for the set of EU15 countries, at least at the 10 per cent level of significance, both using conventional (asymptotic) critical values given in Pedroni (1999), and bootstrap panel cointegration proposed by Westerlund and Edgerton (2007). Moreover, this conclusion holds for the two sub-periods, 1970-1991 and 1992-2006 (broadly before and after the Maastricht Treaty) for most of the cointegration tests carried out.

1. Introduction

The sustainability of public finances is a key policy issue for the European Union (EU). Within the EU fiscal framework, fiscal discipline is an important support for the implementation of monetary policy, particularly in the case of the EMU member countries. In EMU, the existence of sound fiscal policies is seen as a necessary objective for individual countries to pursue. It is not possible to exclude adverse responses from the financial markets when fiscal behaviour is deemed to be unsustainable. Indeed, the accumulation of government debt, following continued budgetary imbalances, may in the end trigger the need for higher long-term interest rates in order to place additional sovereign debt in the markets.[1] Moreover, the Treaties governing the EU also require sustainable public finances. Countries are urged to comply with the budgetary requirements of EMU, by avoiding excessive deficits, keeping debt levels below the 60 percent of GDP reference value, and respecting the requirements of the Stability and Growth Pact (SGP).[2]

The aim of this paper is to examine the sustainability of public finances for the EU15 countries (covering the EU Member States before the 1 May 2004 enlargement) by applying recent advances in the econometrics of non-stationary panel data methods.[3] The econometric literature on unit roots and cointegration testing has been expanding rapidly, and now distinguishes between the first generation tests developed on the assumption of cross-section independence (except for common time effects), and the second generation tests that allow, in a variety of forms and degrees, the dependence that might prevail across the different units in the panel. This question is crucial and responds to the complex nature of the interactions and dependencies that generally exist over time and across the individual units in the panel. For instance, observations on firms, industries, regions and countries tend to be cross-correlated as well as serially dependent. As pointed out by Breitung and Pesaran (2005), the problem of cross-section dependence is particularly difficult to deal with since it could arise for a

[1] Afonso et al. (2007) report evidence on the relevance of fiscal variables as determinants of sovereign credit ratings, which also points to the need for sound fiscal policies.
[2] See Morris et. al (2006) on the revised framework of the SGP.
[3] The countries are: Belgium, Denmark, Germany, Ireland, Greece, Spain, France, Italy, Luxembourg, the Netherlands, Austria, Portugal, Finland, Sweden and the UK.

variety of reasons, including spatial spillover effects, common unobserved shocks, social interactions, or a combination of these factors.

In the context of our paper, cross-dependence can mirror possible changes in the behaviour of fiscal authorities related to the signing of the EU Treaty in Maastricht on 7 February 1992, with the setting up of the convergence criteria that urged the EU countries to consolidate public finances in the run-up to the EMU on 1 January 1999, when most EU legacy currencies were replaced by the euro, and in the context of the SGP since then.

Generally, fiscal sustainability is considered on a country basis and can usually only be restored by changing national fiscal policies. From a monetary policy point of view, fiscal policy in the current institutional setting of EMU must be considered a largely national competence and responsibility. Although, even if there is no single fiscal policy in the EU, a panel sustainability analysis of public finances has to be seen as relevant in a context of EU countries seeking to pursue common and sound fiscal policy behaviour within the SGP framework. Possible cross-country dependence can be envisaged either in the run-up to EMU or, for example, via integrated financial markets. Indeed, with cross-country spillovers in government bond markets especially after the completion of the single EU15 capital market from 1994 were to be expected, interest rates comovements inside the EU became also more noticeable.

To the best of our knowledge, few comparable studies have taken into account the possible cross-sectional dependence among countries when investigating the sustainability of public finances for the EU15 countries. A few studies provide panel unit root and panel cointegration analysis in this context, notably Prohl and Schneider (2006), for eight OECD countries and Claeys (2007) for the EU (not allowing for cross-section dependence). Indeed, although the main analytical techniques used to analyse the sustainability of public finances have been stationarity tests for the stock of public debt and cointegration tests between government expenditures and government revenues, this has been mostly performed for individual countries, which sometimes poses the problem of relatively short time series.[4] This paper takes these results in the literature regarding the sustainability of public finances, and assesses them to see whether they still hold when more powerful cointegration techniques are employed in a panel framework.

Our econometric methodology uses three approaches for unit root testing: panel data integration tests of "first generation" (Breitung, 2000; Choi, 2006; Hadri, 2000; Im, Pesaran and Shin, 2003; Levin, Lin and Chu, 2002; Maddala and Wu, 1999), which assume cross-sectional independence among panel units (except for common time effects); panel data unit root tests of the "second generation" (Choi, 2006; Moon and Perron, 2004), which allow for more general forms of cross sectional dependency (not only limited to common time effects); and panel unit root tests that allow for structural breaks (Im and Lee, 2001). We also implement panel cointegration techniques developed by Pedroni (1999, 2004), and generalised by Banerjee and Carrion-i-Silvestre (2006) and Westerlund and Edgerton (2007), to a structural long-run equation between general government expenditures and revenues. To make the analysis robust, the results of panel data unit root tests are also compared with those obtained with individual unit root tests.

The advantages of panel data methods within the macro-panel setting include the use of data for which the spans of individual time series data are insufficient for the study of many hypotheses of interest. Other benefits include better properties of the testing procedures when compared to more standard time series methods, and the fact that many of the issues studied, such as convergence, purchasing power parity or the sustainability of public finances, naturally lend themselves to being studied in a panel context.

The remainder of the paper is organised as follows. In section two we briefly review the analytical framework of public finance sustainability. In section three we present a brief overview of our fiscal data. In section four we perform the stationarity analysis

[4] Examples of empirical tests of fiscal sustainability on an individual country basis are provided, for instance, by Hamilton and Flavin (1986), Trehan and Walsh (1991), Kremers (1988), Wilcox (1989), Hakkio and Rush (1991), Tanner and Liu (1994), Quintos (1995), Haug (1991), Ahmed and Rogers (1995), Payne (1997), Bohn (1998), Fève and Hénin (2000), Uctum and Wickens (2000), Bergman (2001), Bravo and Silvestre (2002), and Afonso (2005a).

of the fiscal series. In section five we report the cointegration results for the general government expenditure and revenue series. Finally, section six concludes the paper.

2. The analytical framework of public finance sustainability

In the beginning of the 1920s, when writing about the public debt problem faced by France, Keynes (1923) highlighted the need for the French government to conduct a sustainable fiscal policy in order to satisfy its budget constraint. Keynes stated that the absence of sustainability would be evident when "the State's contractual liabilities (…) have reached an excessive proportion of the national income" (p. 54).

In modern terms, the sustainability of public finances is challenged when the government debt-to-GDP ratio reaches an excessive value. There is a problem of sustainability when the government revenues are not enough to keep on financing the costs associated with the new issuance of public debt or, again in Keynes words, when "it has become clear that the claims of the bond-holders are more than the tax payers can support" (p. 55). At that point the government will have to take measures that restore the sustainability of fiscal policy, meaning that the State "must come in due course to some compromise between increasing taxation, and diminishing expenditure, and reducing what (…) [it] owe[s]" (p. 59).

From an analytical perspective, the issue of fiscal policy sustainability can be presented in a straightforward way with the so-called present value borrowing constraint (PVBC). In order to derive the PVBC of a single country, the flow government budget constraint for a given period *t* can be written as

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

where *G* is the primary government expenditure, *R* is the government revenue, *B* is the government debt, and *r* is the real interest rate.[5] Rewriting (1) for the subsequent periods, and recursively solving that equation leads to the following intertemporal budget constraint:

[5] For the validation of theoretical results, the real interest rate is sometimes assumed in the literature to be stationary, but this is a much more difficult assumption for the nominal interest rate.

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

When the second term from the right-hand side of equation (2) is zero, the present value of the existing stock of public debt will be identical to the present value of future primary surpluses. For empirical purposes it is useful to make several algebraic modifications to equation (1). Assuming that the real interest rate is stationary, with mean r, and defining

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

it is possible to obtain the following PVBC:

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

A sustainable fiscal policy needs to ensure that the present value of the stock of public debt, the second term of the right hand side of (4), goes to zero in infinity, constraining the debt to grow no faster than the real interest rate. In other words, it implies imposing the absence of Ponzi games and the fulfilment of the intertemporal budget constraint. Faced with this transversality condition, the government will have to achieve future primary surpluses whose present value adds up to the current value of the stock of public debt.[6]

It is also worth noting that the hypothesis of fiscal policy sustainability is related to the condition that the trajectory of the main macroeconomic variables is not affected by the choice between the issuance of public debt and the increase in taxation. Under such conditions, it would therefore be irrelevant how the deficits are financed, which also implies the assumption of the Ricardian Equivalence hypothesis.[7]

---

[6] McCallum (1984) discusses whether this is a necessary condition to obtain an optimal growth trajectory for the stock of public debt.

[7] Afonso (2005b) provides evidence of overall Ricardian behaviour on the part of EU15 governments.

In addition, one can also derive the solvency condition, with all the variables defined as a percentage of GDP.[8] The PVBC, with the variables expressed as ratios of GDP, with y being the real GDP growth rate, and neglecting for presentation purposes seigniorage revenues, is then written as

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

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

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

with $b_t = B_t/Y_t$, $e_t = E_t/Y_t$ and $\rho_t = R_t/Y_t$. When $r > y$, it is necessary to introduce a solvency condition, given by $\lim_{s\to\infty} b_{t+s}\left(\frac{1+y}{1+r}\right)^{(s+1)} = 0$, in order to bound public debt growth.[9] This yields the familiar result that fiscal policy will be sustainable if the present value of the future stream of primary surpluses, as a percentage of GDP, matches the "inherited" stock of government debt. In a similar fashion, looking at the US after the end of the Second World War, Domar (1944) pointed out that it would be possible to sustain successive primary budget deficits as long as the real growth rate surpasses the real interest rate ($y > r$).

A common practice in the literature is to investigate past fiscal data to see if government debt follows a stationary process or to establish if there is cointegration between government revenues and government expenditures.[10]

---

[8] For instance, Hakkio and Rush (1991) suggest that an analysis based on ratios (to GDP) is more appropriate for growing economies.

[9] This implies that the growth rate of the debt-to-GDP ratio should be less than the factor $\left((1+y)/(1+r)\right)^{(s+1)}$.

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

Recalling the PVBC in equation (4), it is possible to ascertain empirically the absence of Ponzi games by testing the stationarity of the first difference of the stock of public debt, using unit root tests both at the country level and for a European panel. It is also possible to assess fiscal policy sustainability through cointegration tests. The implicit hypothesis concerning the real interest rate, with mean *r*, is also stationarity. Using again the auxiliary variable $E_t = G_t + (r_t - r)B_{t-1}$, and the additional definition $GG_t = G_t + r_t B_{t-1}$, the intertemporal budget constraint may also be written as

$$GG_t - R_t = \sum_{s=0}^{\infty} \frac{1}{(1+r)^{s-1}} (\Delta R_{t+s} - \Delta E_{t+s}) + \lim_{s \to \infty} \frac{B_{t+s}}{(1+r)^{s+1}}, \tag{7}$$

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

Assuming that *R* and *E* are non-stationary variables, and that the first differences are stationary variables, this implies that the series *R* and *E* in levels are I (1). Then, for equation (7) to hold, its left-hand side will also have to be stationary. If it is possible to conclude that *GG* and *R* are integrated of order 1, these two variables should be cointegrated with cointegration vector (1, -1) for the left-hand side of equation (7) to be stationary.

The procedure to assess the sustainability of the intertemporal government budget constraint therefore involves testing the following cointegration regression: $R_t = a + bGG_t + u_t$. If the null of no cointegration, i.e. the hypothesis that the two I(1) variables are not cointegrated, is rejected (with a high-test statistic), this implies that one should accept the alternative hypothesis of cointegration. For that result to hold true, the series of the residual u_t must be stationary, and should not display a unit root.

Hakkio and Rush (1991) also demonstrate that if *GG* and *R* are non-stationary variables in levels, the condition $0 < b < 1$ is a sufficient condition for the budget constraint to be obeyed. However, when government revenues and expenditures are expressed as a percentage of GDP (or in per capita terms), it is necessary to have $b =$

1 in order for the trajectory of the government debt to GDP ratio not to diverge in an infinite horizon.[11]

In terms of our subsequent empirical analysis, we will assess the stationarity of government debt, a sufficient but not necessary condition for fiscal sustainability, and the existence of cointegration between government revenues and expenditures, a necessary condition for fiscal sustainability.

3. Fiscal data overview

All data are taken from the European Commission AMECO (Annual Macro-Economic Data) database, covering the period 1970-2006 for the EU15 countries. The precise AMECO codes are reported in Appendix A, and Table 1 reports summary statistics for our main fiscal variables.

In the period 1970-2006 the highest government debt-to-GDP ratios were recorded in Belgium, Italy, Greece and Ireland, related to high budget deficits incurred by those countries, and resulted notably in the pushing up of interest payments. The government expenditure-to-GDP ratios ranged overall between some 20 per cent and 70 per cent, with the lower values being recorded in the beginning of the period, while the government revenue-to-GDP ratios were in the interval between 20 and 60 per cent. Additionally, visual inspection of the revenue and expenditure time series as a ratio of GDP, as exemplified in Figure 1 for selected countries, and in advance of the subsequent econometric analysis, may help to assess sustainability issues in individual cases.

[11] Quintos (1995), Ahmed and Rogers (1995) and Bergman (2001) discuss the necessary conditions for sustainability in terms of the order of integration of public debt.

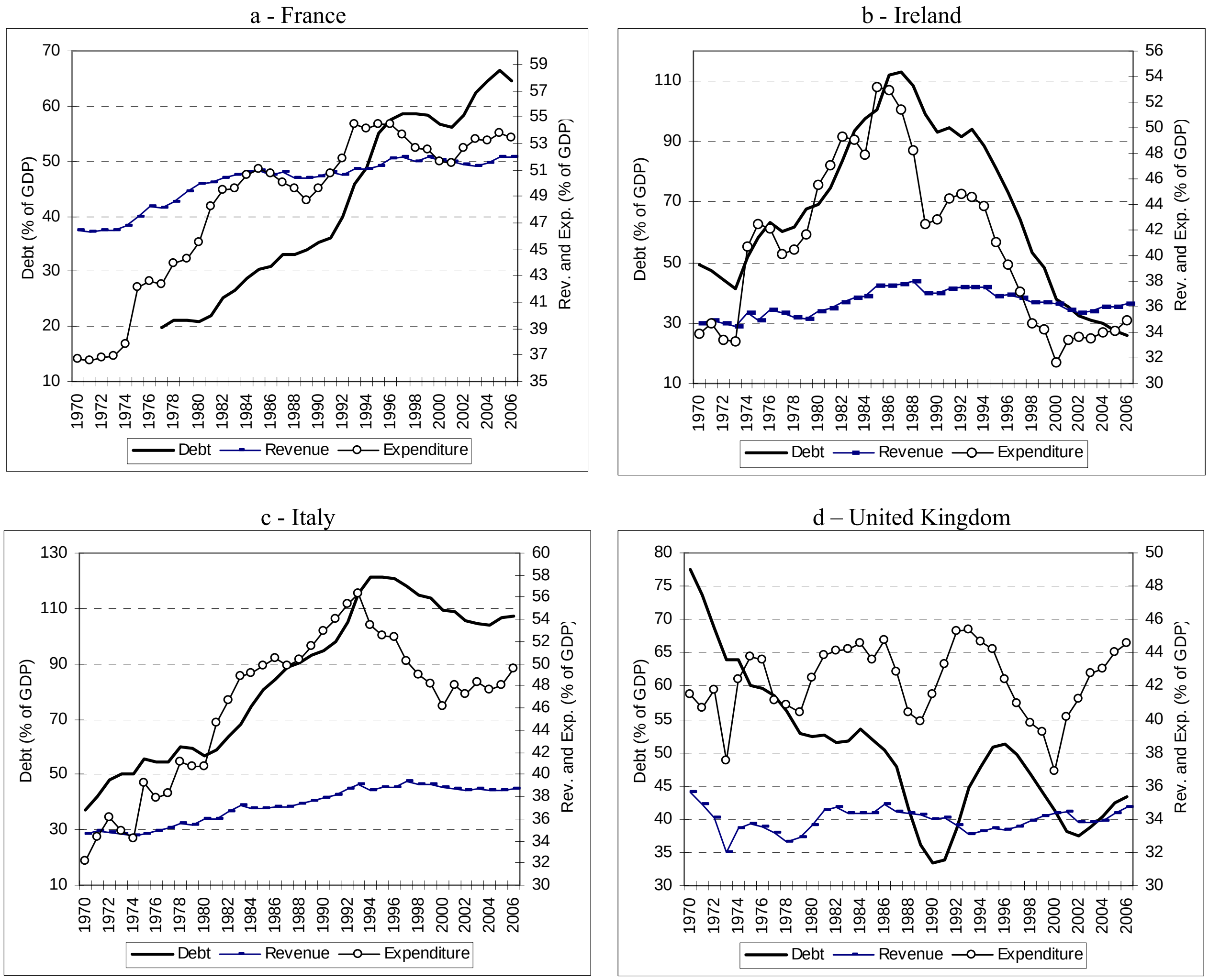
Table 1 – Statistical summary for fiscal variables (% of GDP, 1970-2006)

	Government debt				Primary balance			
Country	Mean	Max	Min	n	Mean	Max	Min	n
Austria	48.0	67.9	16.7	37	0.9	3.5	-2.0	37
Belgium	97.9	133.4	54.3	37	2.0	6.8	-4.8	37
Denmark	48.3	80.1	6.2	36	4.5	11.6	-3.0	37
Finland	26.6	57.8	6.1	36	4.0	9.7	-3.3	37
France	42.3	66.6	19.8	30	0.2	1.9	-2.3	37
Germany	42.5	67.9	18.0	37	0.2	2.8	-4.1	37
Greece	67.2	114.0	17.5	30	-0.7	5.0	-6.7	37
Ireland	67.5	112.9	25.8	37	0.8	6.6	-7.3	37
Italy	84.9	121.5	37.4	37	-0.7	6.6	-6.7	37
Luxembourg	9.3	20.3	4.1	37	2.6	6.4	-1.6	37
Netherlands	60.6	78.5	39.6	32	1.7	5.0	-1.3	37
Portugal	47.7	67.4	14.2	34	-0.4	3.9	-7.4	37
Spain	37.3	66.8	11.8	32	0.0	3.1	-4.4	37
Sweden	49.2	73.2	24.6	34	4.0	10.3	-5.6	37
United Kingdom	49.9	77.4	33.4	37	1.1	6.8	-4.8	37

	Government revenue				Government expenditure			
Country	Mean	Max	Min	n	Mean	Max	Min	n
Austria	48.0	52.5	38.3	37	50.1	56.7	37.1	37
Belgium	46.3	51.1	38.1	37	51.6	62.1	40.2	37
Denmark	52.9	58.1	44.0	37	52.6	60.6	39.5	37
Finland	48.9	57.1	33.6	37	46.5	64.7	29.5	37
France	46.2	50.9	37.1	37	48.4	54.5	36.5	37
Germany	44.3	46.6	39.6	37	46.6	49.9	39.1	37
Greece	34.0	47.0	22.5	37	40.3	52.0	22.6	37
Ireland	36.5	43.6	29.2	37	40.9	53.2	31.6	37
Italy	38.7	47.6	27.9	37	46.2	56.3	32.1	37
Luxembourg	40.4	44.4	27.8	35	38.5	45.2	25.3	35
Netherlands	48.5	53.8	41.2	37	51.0	59.2	42.7	37
Portugal	32.6	43.5	20.6	37	36.9	47.8	18.6	37
Spain	32.8	40.1	20.9	37	35.2	46.6	20.3	37
Sweden	57.4	62.3	46.0	37	57.6	72.4	41.8	37
United Kingdom	39.8	44.1	34.9	37	42.3	45.4	36.9	37

Source: European Commission AMECO database.

Figure 1 – Fiscal variables for selected countries



Source: European Commission AMECO database.

4. Stationarity analysis of fiscal series

In this section we study the stationarity of the fiscal series in our country panel, specifically the stock of government debt in real terms and the ratios to GDP of government revenue and government expenditure, using several unit root tests, which allow notably for cross-country independence and dependence, and for structural breaks.

4.1. Standard (individual) unit root tests

The first step of the analysis is to look at the data univariate properties and to determine the degree of integration of our fiscal series, and also to assess the existence of a unit root in the first difference of the stock of government debt. Theoretically, a process is either I(0), I(1) or I(2). Nevertheless, in practice many variables or variable combinations are borderline cases, so that distinguishing between a strongly autoregressive I(0) or I(1) process (interest rates are a typical example), or between a strongly autoregressive I(1) or I(2) process (nominal prices are a typical example) is far from easy. Therefore, we have applied a sequence of standard unit root tests (Augmented Dickey-Fuller (ADF) test; the Phillips-Perron (PP) test, 1988; the Kwiatkowsky, Phillips and Shin (KPSS) test, 1992; the Elliot, Rothenberg, and Stock Point Optimal (ERS) test, 1996; and the Ng and Perron (NP) tests, 2001), to investigate which of the I(0), I(1), I(2) assumptions is most likely to hold.[12]

All these tests have been implemented for the first difference of the stock of government debt at 1995 constant prices, while taking general government expenditure and the general government revenue as a percentage of GDP. The results for a model with a constant and no trend are reported in Tables B1, B2 and B3 in Appendix B (respectively for the debt, revenue and expenditure variables).[13]

According to the ADF tests (see Table B1 in Appendix B), the first difference of the stock of public debt seems to be non-stationary in most countries, since the null of a unit root is rejected at the five per cent level of significance only for Austria, Finland, France, the UK and Sweden, indicating that the solvency condition would not be satisfied for ten out of the 15 EU15 countries. At the ten per cent level of significance, the non-stationarity hypothesis would also be rejected for Germany and the Netherlands. A similar result is obtained by the PP test, where the unit root assumption is rejected at the five per cent level only for Austria, Ireland, Luxembourg and Sweden. The Ng-Perron test more or less confirms this, suggesting that real

[12] These tests are sufficiently well known, and the reader will find references at the end of the paper.

[13] If the estimated model also includes a linear trend then all tests conclude in favor of a unit root in the debt, revenues and expenditures for all EU15 countries.

government debt is not stationary at the five per cent level for most EU15 countries except for Austria, Finland, Germany, the Netherlands, Sweden and the UK. Additionally, one notices that these results are in line with the ones reported by Afonso (2005a) for the period 1970-2003.[14]

The KPSS test that considers the null hypothesis of stationarity against the alternative unit root hypothesis confirms these results. Finally, the results given by the ERS test are in accordance with the previous ones and favour the non-stationarity of the first difference of the stock of government debt for EU15 countries, with the notable exceptions of Austria, Finland, France, the Netherlands, Sweden and the UK.[15]

Concerning the general government expenditure and revenue-to-GDP ratios, four standard unit root tests out of five (the ADF, PP, Ng-Perron and ERS tests) indicate that the two series would be non-stationary in levels for most countries. Note that the ERS test is similar to an ADF "t" test, as performed by Dickey-Fuller, but has the best overall performance in terms of small-sample size and power, dominating the ordinary Dickey-Fuller test. The problem here is that the test critical values are calculated for 50 observations and may not be accurate for a sample size with a time dimension of 37 observations. Consequently, the results obtained with this test may be questionable. The KPSS test provides relatively similar results for general government revenues (which are found to be non-stationary at the five per cent level of significance except for Germany, Ireland, the Netherlands and UK), but the opposite results for general government expenditures (which appear to be stationary at the five per cent level of significance except for Finland, France, Ireland, Portugal and Spain).

4.2. Individual unit root tests allowing for structural breaks

The results obtained in the previous section are based on the assumption that no structural break exists in the series under consideration. However, it is now well-

[14] According to Afonso (2005a), general government revenues and expenditures are cointegrated only for Austria, Finland, Germany and the Netherlands, but the cointegration coefficient is quite below unity. Furthermore, for Belgium, Denmark, Portugal and the UK, there are significant structural breaks in the cointegration relationship.

[15] Note that as it is a one-sided unit root test, with the alternative H_1:ρ<1, a calculated statistic smaller than the tabulated critical value will lead to the rejection of the null H_1:ρ=1.

established that examining time series for the presence of structural breaks is an important component of any empirical analysis. Indeed, the standard unit root tests have serious power distortions in the presence of structural breaks, which could be the case, for instance, for Germany following reunification in 1990. For this reason, we now investigate this issue, using two endogenous unit root tests for structural breaks, the Zivot and Andrews (1992) and the Lumsdaine and Papell (1997) tests.[16]

We use the C version of the sequential trend break model proposed by Zivot and Andrews (1992) to investigate the presence of a unit root in the first difference of the stock of public debt at 1995 constant prices, and the general government expenditure and the general government revenue-to-GDP ratios. This model combines the one-time change in the level and the slope of the trend function of the series. Hence, to test for a unit root against the alternative of a one-time structural break, Zivot and Andrews use the following regression equation:

$$\Delta Y_t = \alpha + \Phi Y_{t-1} + \beta\, t + \theta DU_t + \gamma DT_t + \sum_{j=1}^{k} d_j \Delta Y_{t-j} + \varepsilon_t \;\; t = 1,...,T \;, \qquad (8)$$

where DU_t is an indicator dummy variable for a mean shift occurring at each possible break-date (*TB*), while DT_t is the corresponding trend shift variable. Formally, $DU_t = 1$ if t $>TB$ and 0 otherwise; and $DT_t = t-TB$ if t $>TB$ and 0 otherwise. The ΔY_{t-j} terms on the right-hand side of the above equation allow for serial correlation and ensure that the disturbance term is white noise.

The null hypothesis of the test is $\Phi=0$, which implies that the series Y_t contains a unit root with a drift that excludes any structural break, while the alternative hypothesis $\Phi<0$ implies that the series is a trend-stationary process with a one-time break occurring at an unknown point in time. The Zivot and Andrews method regards every point as a potential break-date (*TB*), and runs a regression for every possible break-date sequentially.

[16] In doing so, we do not consider structural break-dates as being exogenously determined, but we test endogenously for them, i.e. assuming the break-date to be unknown. We are grateful to J. Lee for providing us with the GAUSS codes (for the Zivot and Andrews and for the Lumsdaine and Papell tests) that we adjusted for our analysis, and which are available upon request.

In applying the Zivot and Andrews (1992) test, some region must be chosen such that the end points of the sample are not included, because in the presence of these end points, the asymptotic distribution of the statistics diverges to infinity. Zivot and Andrews suggest the "trimming region" [0.15, 0.85]. Among all possible break-points (TB), the procedure selects as its choice of break-date (TB^*) the date for which the ADF t-statistic (the absolute value of the t-statistic for Φ) is maximised.

Although asymptotic critical values are available for this test, Zivot and Andrews warn that with small sample sizes (such as the one we are using for our fiscal sustainability analysis) the distribution of the test statistic can deviate substantially from its asymptotic distribution. To circumvent this distortion, 'exact' critical values for the test are computed following the methodology advocated in Zivot and Andrews (1992, p. 262).

The results of the test by Zivot and Andrews (1992) together with exact critical values are reported in Tables C1, C2 and C3 of Appendix C, respectively for the first difference of the stock of government debt, government expenditure and government revenue series. The outcome clearly favours the non-stationary hypothesis. Indeed, at the five per cent level of significance the test allows the rejection of the unit root hypothesis in the stock of government debt for Finland and the UK, in general government expenditures only for Finland, France, Luxembourg, Portugal, and Spain, and cannot reject it in the general government revenues for all EU15 countries.

A possible problem with the Zivot and Andrews (1992) test is the loss of power if there are two structural breaks in the series.[17] Lumsdaine and Papell (1997) explore this possibility by extending Zivot and Andrews's models to allow for two endogenous breaks under the alternative hypothesis and additionally to accommodate for breaks in the level and the trend. Series are generally interpreted as broken trend stationary if the null unit root hypothesis is rejected in favour of the alternative of two breaks. More precisely, model C of Zivot and Andrews (1992), as used previously and

[17] Besides, Lee and Strazicich (2003) suggest that spurious rejection problems may arise akin to that with Zivot and Andrews with a break under the null hypothesis.

extended by Lumsdaine and Papell (1997) (model CC, which allows for two breaks in the intercept and slope of the trend), is written as follows:

$$\Delta Y_t = \alpha + \Phi Y_{t-1} + \beta\, t + \theta_1 DU_{1t} + \gamma_1 DT_{1t} + \theta_2 DU_{2t} + \gamma_2 DT_{2t} + \sum_{j=1}^{k} d_j \Delta Y_{t-j} + \varepsilon_t \;\; t = 1,...,T \; , \quad (9)$$

where DU_{1t} and DU_{2t} are indicator dummy variables for a mean shift occurring respectively at time (TB_1) and (TB_2) with $TB_2 > TB_1 + 2$, while DT_{1t} and DT_{2t} are the corresponding trend shift variables. Formally, DU_{1t} =1 if $t > TB_1$ and 0 otherwise; and DU_{2t} =1 if $t > TB_2$ and 0 otherwise. In the same way, $DT_{1t} = t\text{-}TB_1$ if t $> TB_1$ and 0 and otherwise; and $DT_{2t} = t\text{-}TB_2$ if t $> TB_2$ and 0 and otherwise. As in the previous test, the $\Delta Y_{t\text{-}j}$ terms on the right-hand side of the above equation allow for serial correlation and ensure that the disturbance term is white noise. Note that the break points are chosen using the same approach as in the one break case, and that the critical values are generated as in the case of the Zivot and Andrews (1992) test.

The Lumsdaine and Papell (1997) two-break test results are reported in Tables C1, C2 and C3 of Appendix C, again respectively for the first difference of the stock of public debt, government expenditure, and government revenue. Here, two important results emerge. First, by allowing for two structural breaks it is not possible to reject the unit root null hypothesis in the general government expenditure as well as in the general government revenue series for all EU15 countries, at the five per cent level of significance. A similar result is obtained for the stock of public debt for most EU countries except Finland and Germany.

Second, in the majority of cases the results of the Lumsdaine and Papell (1997) two-break test are in accordance with those previously obtained with the Zivot and Andrews (1992) one-break test, which is a relatively satisfactory result. For instance, no statistically significant structural break can be detected by the two tests in the general government expenditures. The two tests are also able to detect a structural break for Finland and for the UK in the first difference of the stock of government debt, and the date break found is the same for Finland (1991). Of course, some diverging conclusions in the two tests' diagnostics can be noted when one allows for two structural breaks in the series, in particular for the general government revenue series for several EU15 countries. Finally, it may be relevant to note that the years selected as the break points for Finland and Germany in the government debt variable were 1991, a period closely related to an economic downturn in the former country, where the government stepped in to solve the banking crisis, and to reunification in the latter.

4.3. First generation panel unit root tests (cross-country independence)

In addition to the previous unit root tests applied to individual series, we have also carried out a set of panel data unit root test the robustness of the degree of integration of our series. In this sub-section, we implement more particularly the following panel data unit root tests (Breitung, 2000; Hadri, 2000; Im, Pesaran and Shin, 1997, 2003; Levin, Lin and Chu, 2002; Maddala and Wu, 1999; Choi, 2006; and Moon and Perron, 2004). Note that all tests except the last two are "first generation" panel data unit root tests.

First, we used the test proposed by Im, Pesaran and Shin (2003, hereafter IPS), which has been widely implemented in empirical research due to its rather simple methodology and alternative hypothesis of heterogeneity. This test assumes cross-sectional independence among panel units (except for common time effects), but allows for heterogeneity in the form of individual deterministic effects (constant and/or linear time trend), and heterogeneous serial correlation structure of the error terms. Tables 2, 3 and 4 report the results of the IPS test for the government debt, and for the revenue and expenditure ratio series. In order to facilitate comparisons, we also provide the results of five other panel unit root tests: Levin, Lin and Chu (2002), Breitung (2000), and Fisher-type tests using ADF and PP tests (Maddala and Wu, 1999, hereafter MW; and Hadri, 2000).

Concerning the first difference of the stock of government debt, the results given by the panel data unit root tests are more concomitant than those provided by the standard (individual) unit root ones. Indeed, at the five per cent level of significance, five panel data tests out of six (with the exception of the Hadri test) reveal that the null unit root hypothesis can be rejected at the five per cent level for EU15 countries (see Table 2), thus supporting the stationarity of the change in the stock of government debt and hence the non-rejection of the solvency condition for the overall country sample.[18]

Table 2 – Summary of panel data unit root tests for the first difference of the stock of government debt, constant prices (1970-2006)

Method	Statistic	P-value*	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t stat	-1.92991	0.0268	15	494
Breitung t-stat	-2.99756	0.0014	15	479
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	-3.18952	0.0007	15	494
ADF - Fisher Chi-square	58.7550	0.0013	15	494
PP - Fisher Chi-square	77.9679	0.0000	15	509
Null: No unit root (assumes common unit root process)				
Hadri Z-stat	2.57067	0.0051	15	524

* Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
Automatic selection of lags based on SIC. Newey-West bandwidth selection using a Bartlett kernel

As far as the general government revue-to-GDP ratio is concerned, five panel data tests out of six (with the exception of the Breitung test) produce significant evidence in favour of their integration of order one for all EU15 countries at the 5 per cent level of significance (see Table 3). In other words, the non-stationarity of the revenue-to-GDP ratio cannot be rejected.

Table 3 – Summary of panel data unit root tests for general government revenue-to-GDP ratios (1970-2006)

Method	Statistic	P-value*	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu stat	-0.77258	0.2199	15	534
Breitung t-stat	-2.57515	0.0050	15	519
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	2.09943	0.9821	15	534
ADF - Fisher Chi-square	20.8225	0.8934	15	534
PP - Fisher Chi-square	20.1458	0.9127	15	537
Null: No unit root (assumes common unit root process)				
Hadri Z-stat	9.94807	0.0000	15	553

* Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
Automatic selection of lags based on SIC. Newey-West bandwidth selection using a Bartlett kernel

[18] A common feature of the panel tests mentioned above is that they maintained the null hypothesis of a unit root in all panel members (the only exception is the test by Hadri, 2000, whose null hypothesis is stationarity for all panel units). Therefore, their rejection decision actually indicates that at least one panel member is stationary, with no information about how many series or which ones are stationary. This possibility for a mixed panel implies that some of the members may be stationary while others may be non-stationary (see Taylor and Sarno, 1998 and Taylor, 2004 for further details).

Finally, and according to Table 4, the general government expenditure-to-GDP ratio also appears to have a unit root for all countries at the 5 per cent level of significance if one refers to the results of all panel data unit root tests.

Table 4 – Summary of panel data unit root tests for general government expenditure-to-GDP ratios (1970-2006)

Method	Statistic	P-value*	Cross-sections	Obs
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu stat	-0.88260	0.1887	15	450
Breitung t-stat	-1.53137	0.0628	15	435
Null: Unit root (assumes individual unit root process)				
Im, Pesaran and Shin W-stat	2.61169	0.9955	15	450
ADF - Fisher Chi-square	13.3435	0.9963	15	450
PP - Fisher Chi-square	13.1161	0.9968	15	465
Null: No unit root (assumes common unit root process)				
Hadri Z-stat	10.6455	0.0000	15	480

* Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
Automatic selection of lags based on SIC. Newey-West bandwidth selection using a Bartlett kernel

However, as shown by several authors (notably O'Connell, 1998; and Banerjee, Marcellino and Osbat, 2004, 2005), the assumption of cross-sectional dependence limited to the case of common time effects on which the asymptotic results of the IPS's procedure relies (like most panel data unit root tests of "the first generation", including Maddala and Wu, 1999; Levin, Lin and Chu, 2002; and more generally all previous six panel data unit-root tests) is often unrealistic and can be at odds with economic theory and empirical results. Besides, as shown in two simulation studies by Banerjee et al. (2004a, 2004b), if panel members are cross-correlated or even cross-sectionally cointegrated, all these tests experience strong size distortions and limited power This point is analytically confirmed by Lyhagen (2000) and Pedroni and Urbain (2001).

4.4. Second generation panel unit root tests (cross-country dependence)

As Breitung and Pesaran (2005) note, time series are contemporaneously correlated in many macroeconomic applications using country or regional data. Prominent examples of this are the analysis of purchasing power parity and output convergence (see for instance Pesaran, 2004). However, the literature on how to model cross-sectional dependence in large panels is still developing. Cross-sectional dependence can arise due to a variety of factors, such as omitted observed common factors, spatial spillover effects, for example via integrated financial markets, unobserved common factors, or general residual interdependence, all of which could remain even when all observed and unobserved common effects have been taken into account. In the EU context, some possible cross-country dependence can be envisaged in the presence of a similar policy measures (i.e. in the run-up to EMU), coupled with similar fiscal behaviour (e.g. pursuing fiscal consolidation in the run-up to EMU and within the SGP framework), and cross-country spillovers in government bond markets especially after the completion of the single EU15 capital market from 1994 (stage 2 of EMU).

For this reason, various recent studies have proposed panel unit root tests allowing for more general forms of cross-sectional dependency, e.g. Choi (2006), Bai and Ng (2003), Moon and Perron (2004), Pesaran (2007) and Phillips and Sul (2003). We have decided to investigate the presence of a unit root using two second generation tests, namely Choi (2006) and Moon and Perron (2004), to whom we refer the reader for further details.[19] This last test in particular seems to show good size and power for different values of T and N and model specifications, according to the Monte Carlo experiments conducted by Gutierrez (2006).[20]

The results reported in Tables 5 and 6 indicate that the null unit root hypothesis cannot be rejected by the two tests at the 5 per cent level for the government expenditure and revenue ratios, but can be rejected for the government debt for all EU15 countries, which supports the initial results produced by the first generation panel data unit root tests. Furthermore, tests on the series in first differences confirm the hypothesis of stationarity for government expenditure and revenue ratios. Therefore, we may conclude that the general government revenue and expenditure-to-GDP ratios expressed in level are integrated of order 1 for all EU15 countries, independently of the panel unit root tests considered, thereby demonstrating that the non-stationarity property of our revenue and expenditure series is a robust result.

[19] Note that another possibility would be to use a procedure as the one advocated by Breuer et al. (2002) whereby unit root testing is conducted within a seemingly unrelated regression (SUR) framework. An advantage of this procedure is that the SUR framework is another useful way of addressing cross-sectional dependency.

[20] We are grateful to C. Hurlin for making available his Matlab codes to us.

Table 5 – Results of Choi's (2006) test [a] (1970-2006)

	P_m statistic	Z statistic	L* statistic
First difference of the stock of public debt	0.000	0.000	0.000
General government revenue-to-GDP ratios	0.463	0.354	0.354
General government expenditure-to-GDP ratios	0.364	0.382	0.373

Note: All figures reported in the table are P-values.
a - Note that the P_m test is a modification of Fisher's (1932) inverse Chi-square tests, and rejects the null unit root hypothesis for positive large value of the statistics, and that the L^* is a logit test. The tests (Z and L^*) reject the null for large negative values of the statistics. The P, Z and L^* tests converge under the null to a standard normal distribution as ($N, T \to\infty$) (see Choi, 2006 for further details).

Table 6 – Results of Moon and Perron's (2004) test [a] (1970-2006)

	t*a	t*b
First difference of the stock of public debt	0.000	0.000
General government revenue-to-GDP ratios	0.526	0.541
General government expenditure-to-GDP ratios	0.382	0.434

Note: All figures reported in the table are P-values.
a - The null hypothesis of the two tests proposed by Moon and Perron (2004) is the unit root for all panel units. Under the null H_0 , they show that for ($N, T \to\infty$) with $N / T \to 0$, the statistics t*a and t*b have a standard normal distribution.

4.5. Panel unit root tests allowing for structural breaks

The presence of structural breaks in panel series data can induce behaviour similar to that of an integrated process, making it difficult to differentiate between a unit root and a stationary process with a regime shift. For this reason, the panel unit root tests in the previous section, such as the IPS and MW tests, may potentially suffer from a significant loss of power if structural breaks are present in the data.

In this section, we employ the panel data unit root test based on the Lagrangian multiplier (LM) principle developed by Im and Lee (2001), which is very flexible since it can be applied not only when a structural break occurs at a different time period in each time series, but also when the structural break occurs in only some of the time series. The proposed test is not only robust to the presence of structural breaks, but is also more powerful than the popular IPS test in the basic scenario where no structural breaks are involved. Furthermore, as reported by Im and Lee (2001), since the LM test loses little power by controlling for spurious structural breaks when they do not exist, this represents a reasonable strategy to control for breaks even when they are only at a suspicious level. Moreover, this panel LM test does not require the simulation of new critical values that depend on the number and location of breaks.[21]

In order to provide a robust analysis, we compare both univariate and panel LM unit root test results with and without a structural break. We begin with the Schmidt and Phillips (1992) univariate LM unit root test without any structural change. Then, we move to extensions that allow for one break, since our time series covers periods during which structural change may have occurred due to structural and institutional changes in the EU15 countries. In addition to the Schmidt and Phillips (1992) no-break test, we employ the univariate test and the Lee and Strazicich (2003) minimum LM unit root tests with one break to determine the structural break point in each country. After determining the optimal break point, we employ the panel LM unit root test of Im and Lee (2001). For comparison, we also show the panel LM test results with no breaks.

To determine the optimal break point in the panel LM test, we utilize the univariate minimum LM unit root tests of Lee and Strazicich (2003). These tests are comparable to the corresponding Dickey and Fuller-type endogenous break tests of Zivot and Andrews (1992). The performance of the LM test is comparable to or superior to these counterpart tests in terms of size and power. In addition, the LM unit root tests are not subject to spurious rejections under the null. In each test, the break point is determined endogenously from the data via a grid search by selecting the break where the value of the unit root test statistic is at its minimum. Using the minimum LM tests of Lee and Strazicich (2003), the unit root test statistic is estimated at each break point. The procedure is repeated over the time interval [0.1T, 0.9T] in order to eliminate end points, until the break is determined where the unit root t-test statistic is minimized. The optimal number of lags in each country is determined by sequentially examining the t-statistic for the last lag coefficient to see if it is significant at the approximate 5 per cent level in an asymptotic normal distribution. We begin with the one-break LM test. If less than one break is significant, we employ the no-break LM unit root test. The corresponding LM unit root test statistic is then chosen after

[21] It should be noted that these tests assume cross-sectional independence among panel units.

determining the optimal break point. After determining the appropriate unit root test statistic for each country, the panel LM test statistic is then calculated.[22]

The results are reported in Tables 7, 8 and 9, which respectively show the first difference of the stock of government debt at 2000 constant prices, and general government expenditure and revenue taken as a percentage of GDP. For the univariate LM test with no break, the unit root null can be rejected at the 5 per cent level of significance in three countries for government debt (Austria, Greece and Italy), in two countries for government expenditure (Finland and the UK), and in two countries for government revenue (Denmark and Sweden). After allowing for a structural break, the univariate minimum LM test rejects the unit root null in four countries for government debt (Austria, Denmark, France and Italy), in four countries for government expenditure (Finland, France, Italy and the Netherlands), and cannot reject it for government revenue at the 5 per cent level.

Table 7 – Panel LM unit root tests allowing for structural break for the first difference of the stock of government debt (1970-2006)

Country	Individual LM statistic without a break [a]	Lags	Individual LM tatistic with a break [b]	Lags	Optimal break point
Austria	-4.420*	7	-4.707*	7	2003
Belgium	-2.246	8	-1.632	1	1995
Denmark	-2.288	2	-4.126*	7	2000
Finland	-1.945	8	-2.456	3	1993
France	-2.997	3	-3.718*	4	1993
Germany	-2.877	8	-3.075	8	1993
Greece	-3.213*	8	-2.099	8	2002
Ireland	-1.444	2	-2.683	5	1995
Italy	-4.404*	7	-4.905*	7	2003
Luxembourg	-1.449	5	-1.731	4	1997
Netherlands	-0.487	3	-0.868	3	1992
Portugal	-1.874	3	-2.132	3	2002
Spain	-1.599	1	-1.076	1	1993
Sweden	-2.129	1	-3.155	1	2000
United Kingdom	-2.142	4	-2.169	4	2002
Panel LM stat [c]	-3.126*		-5.077*		

Notes: i) As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5 per cent the critical value for the LM test without break is - 3.06. At 5 per cent the critical value for the minimum LM test with one break is - 3.566.
ii) The critical value for the panel LM test (with or without breaks) is -1.645 with an asymptotic standard normal distribution.
iii) * denotes significance at the 5 per cent level.
a) Schmidt and Phillips (1992) test; b) Lee and Strazicich (2003) test; c) Im and Lee (2001) test.

[22] We are grateful to J. Lee for providing us with the GAUSS codes, which we have adapted for our analysis, and that are available upon request.

Without allowing for structural breaks, the panel LM test statistic is -3.126 for the stock of real government debt series clearly indicating that the unit root null can be rejected at the 5 per cent level of significance, due to increased power from panel data (see Table 7). In addition, after allowing for structural breaks, the panel test statistic of -5.95 strongly rejects the unit root null at the 5 per cent level. These results clearly demonstrate the gain in power from combining structural breaks with panel data. Since the panel LM test statistic is calculated using the average test statistic of all countries, it is possible that the panel results are due to a small number of outliers having a relatively large impact.

Examination of the univariate test statistics (with breaks) for each country reveals that Austria, Denmark, France and Italy might qualify as such an outlier, as they are the only four countries that reject the unit root null at the 5 per cent level. In order to see if our panel results are robust to a possible outlier effect, we therefore recalculated the panel LM test statistic (with breaks) omitting these four countries. The resulting panel test statistic of -3.62 continues to reject the unit root null at the 5 per cent level of significance, thus firmly supporting our hypothesis that the panel test results are not due to outliers.

Concerning the general government expenditure and the revenue series taken as a percentage of GDP, it appears that the panel LM test statistics with or without a break cannot reject the null unit root hypothesis at the 5 per cent level of significance, thus providing strong evidence in favour of a unit root in these two EU15 country series.

Table 8 – Panel LM unit root tests allowing for structural break for general government revenue-to-GDP ratios (1970-2006)

Country	Individual LM statistic without a break [a]	Lags	Individual LM statistic with a break [b]	Lags	Optimal break point
Austria	-2.667	2	-2.957	2	1989
Belgium	-1.627	3	-2.313	3	1990
Denmark	-2.128	6	-2.467	6	1989
Finland	-3.901*	8	-3.806*	8	2001
France	-3.063	4	-4.205*	6	1995
Germany	-1.593	7	-2.492	8	1998
Greece	-1.292	0	-1.443	0	1992
Ireland	-0.916	5	-0.346	8	1997
Italy	-2.284	8	-3.950*	8	1991
Luxembourg	0.502	8	0.362	8	1992
Netherlands	-2.070	2	-4.168*	8	1992
Portugal	-1.674	0	0.105	8	1988
Spain	-1.928	0	-1.577	8	1983
Sweden	-0.595	6	-0.811	8	1991
United Kingdom	-3.156*	6	-2.141	5	1987
Panel LM stat [c]	-0.292		-1.62		

Notes: i) As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5 per cent the critical value for the LM test without a break is - 3.06. At 5 per cent the critical value for the minimum LM test with one break is - 3.566.
ii) The critical value for the panel LM test (with or without breaks) is -1.645 with an asymptotic standard normal distribution.
iii) * denotes significance at the 5 per cent level.
a) Schmidt and Phillips (1992) test; b) Lee and Strazicich (2003) test; c) Im and Lee (2001) test.

Table 9 – Panel LM unit root tests allowing for structural break for general government expenditure-to-GDP ratios (1970-2006)

Country	Individual LM statistic without a break [a]	Lags	Individual LM statistic with a break [b]	Lags	Optimal break point
Austria	-1.627	3	-1.253	2	1981
Belgium	-2.128	6	-1.855	6	1991
Denmark	-3.901*	8	-2.055	8	1985
Finland	-3.063	4	-1.935	7	1981
France	-1.593	7	-1.712	2	1993
Germany	-1.292	0	-1.553	6	1993
Greece	-0.916	5	-2.779	7	1994
Ireland	-2.284	8	-1.487	7	1990
Italy	0.502	8	-2.372	7	2000
Luxembourg	-2.070	2	0.234	6	2003
Netherlands	-1.674	0	-1.394	7	1985
Portugal	-1.928	0	-1.966	8	2000
Spain	-0.595	6	-1.898	5	1986
Sweden	-3.156*	6	-1.203	1	1993
United Kingdom	-1.411	2	-1.326	7	2000
Panel LM stat [c]	0.212		0.999		

Notes: i) As all tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null of a unit root. At 5 per cent the critical value for the LM test without a break is - 3.06. At 5 per cent the critical value for the minimum LM test with one break is - 3.566.
ii) The critical value for the panel LM test (with or without breaks) is -1.645 with an asymptotic standard normal distribution.
iii)) * denotes significance at the 5 per cent level.
a) Schmidt and Phillips (1992) test; b) Lee and Strazicich (2003) test; c) Im and Lee (2001) test.

Overall, our findings using panel data unit root tests that allow for structural breaks support the previous results of first and second generation panel data unit root tests, leading us to conclude that the stock of government debt series is integrated of order zero (indicating that the solvency condition would be satisfied for the EU15 countries), and that the general government expenditure and the revenue series are integrated of order one. These findings are summarized in Table 10.

Table 10 – Summary of stationarity tests, 5 per cent level of significance
(H_0: unit root, non-stationarity, for most cases)

Set of results	**1st difference of stock of real government debt** (2000 constant prices)		
1	Individual unit root tests	ADF, no unit root: AT, FI, FR, UK, SW	PP, no unit root: AT, IR, LU, SW
2	Individual unit root tests, with breaks	Zivot-Andrews (1992), no unit root: FI, UK	Lumsdaine-Papell (1997), no unit root: FI, DE, UK
3	Panel unit root 1st generation tests, country independence	Levin-Lin-Chu, Breitung, Im-Pesaran-Shin, ADF, PP: no unit root, stationarity. Hadri, unit root.	
4	Panel unit root 2nd generation tests, country dependence	Choi (2006): no unit root.	Moon-Perron (2004): no unit root.
5	Individual LM unit root tests	Schmidt-Phillips (1992), no breaks, no unit root: AT, GR, IT	Lee-Strazicich (2003), with breaks, no unit root: AT, DK, FR, IT
6	Panel LM unit root tests	Im-Lee (2001), no breaks: no unit root.	Im-Lee (2001), with breaks: no unit root.
		General government revenue (% of GDP)	**General government expenditure (% of GDP)**
7	Individual unit root tests	ADF, no unit root: AT, DE, LU, SW PP, no unit root: AT, FI, DE, LU, SW, UK	ADF, no unit root: DE, UK PP, no unit root: DE, LU, PT, UK
8	Panel unit root 1st generation tests, country independence	Levin-Lin-Chu, Im-Pesaran-Shin, Hadri, ADF, PP: unit root, non-stationarity. Breitung, no unit root.	Levin-Lin-Chu, Breitung, Im-Pesaran-Shin, Hadri, ADF, PP: unit root, non-stationarity.
9	Panel unit root 2nd generation tests, country dependence	Choi (2006) and Moon-Perron (2004): unit root.	Choi (2006) and Moon-Perron (2004): unit root.
10	Individual LM unit root tests	Schmidt-Phillips (1992), no breaks, no unit root: FI, UK. Lee-Strazicich (2003), with breaks, no unit root: FI, FR, IT, NL.	Schmidt-Phillips (1992), no breaks, no unit root: DK, SW. Lee-Strazicich (2003), with breaks, no unit root: reject for all countries.
11	Panel LM unit root tests	Im-Lee (2001), no breaks: unit root.	Im-Lee (2001), with breaks: unit root.

Note: AT – Austria, DE – Germany, DK – Denmark, FI – Finland, FR – France, GR – Greece, IR – Ireland, IT – Italy, LU – Luxembourg, NL – Netherlands, PT – Portugal, SW – Sweden, UK – United Kingdom.

5. Cointegration between government expenditure and revenue ratios

After having confirmed the non-stationarity of our series of government revenue and expenditure for the EU15 as a whole, in particular if one refers to the panel data unit root tests of the previous section, it is natural to test the existence of a structural long-run relationship between both series. This is the procedure we use in this section to assess fiscal sustainability on the basis of the intertemporal budget constraint as given in (7).

Compared to panel unit root tests, the analysis of cointegration in panels is still at an early stage of development. So far, the focus of the panel cointegration literature has been on residual-based approaches, although there have been a number of attempts to develop system approaches as well. As is the case for panel unit root tests, panel cointegration tests are based on homogeneous and heterogeneous alternatives. The residual-based tests were developed to ward against the spurious regression problem that can arise in panels when dealing with I(1) variables. Such tests are appropriate when it is *a priori* known that at most there can be only one within-group cointegration in the panel. Notable contributions to this strand of the literature include Kao (1999), Pedroni (1999, 2000, 2004), and more recently Westerlund and Edgerton (2007) and Westerlund (2005, 2007). System approaches are required in more general settings where more than one within-group cointegrating relation might be present, and/or unobserved common I(1) factors exist. Recent contributions in this area include the work of Larsson, Lyhagen and Lothgren (2001), Groen and Kleibergen (2003) and Breitung (2005), who has generalized the likelihood approach introduced in Pesaran, Shin and Smith (1999).

The computation of the Pedroni test statistics assumes cross-sectional independence across individual units (apart from common time effects), an assumption that, as we have already mentioned, is probably absent for many macroeconomic time series. To take into account the possible cross-sectional dependence when carrying out the cointegration analysis, we decided to compute the bootstrap distribution of Pedroni’s test statistics, thereby generating data-specific critical values. As in Banerjee and Carrion-i-Silvestre (2006), we have of course not used the seven statistics proposed by Pedroni (1999, 2004) to test the null hypothesis of no cointegration using single equation methods based on the estimation of static regressions. These statistics can also be grouped into either parametric or non-parametric statistics, depending on the way that autocorrelation and endogeneity bias are accounted for. In our study, we are only concerned with the parametric version of the statistics, i.e. the normalized bias and the pseudo t-ratio statistics, and with the ADF test statistics in particular. These test statistics are defined by pooling the individual tests, so that they belong to the class of between-dimension test statistics (see Pedroni, 1999, 2004 for further details).

As Banerjee and Carrion-i-Silvestre (2006) stress, some caution is required concerning the method used to bootstrap cointegration relationships, since not all available procedures lead to consistent estimates. In this regard, we have followed Phillips (2001), Park (2002) and Chang, Park and Song (2006) in using a modified version of the sieve bootstrap described in Banerjee et al. (2006).[23]

Table 11 reports the results of the panel data cointegration tests developed by Pedroni (1999, 2004) both using conventional (asymptotic) critical values (as per Pedroni, 1999) and bootstrap critical values. We present the results for the entire sample period, 1970-2006, and for two sub-periods, 1970-1991 and 1992-2006, in order to assess whether different fiscal realities and behaviour can be detected for more recent years in the EU, notably after the signing of the Maastricht Treaty with the setting up of the fiscal convergence criteria.

For the period 1970-2006, using conventional asymptotic critical values (-1.65 at 5 per cent) calculated under the assumption of cross-sectional independence (reported in Pedroni, 1999, and extracted from the standard normal distribution), the null hypothesis of no cointegration between government revenue and expenditure ratios is always rejected by the test statistics, irrespective of whether the model includes a constant or a linear trend. However, if we consider bootstrap critical values (which are valid if there is some dependence among individuals), the conclusions of the test are less straightforward, and instead crucially depend on the level of significance chosen. Indeed, at the 10 per cent level of significance, the null hypothesis of no cointegration is still rejected by the data, but an opposite result is obtained at the 5 per cent level of

[23] We are grateful to A. Banerjee and J. Carrion-i-Silvestre for providing us with their GAUSS codes (for a detailed discussion of the method used, see the end of the paper).

significance for a model including either a constant or a linear trend. Finally, retaining a 10 per cent level of significance, we conclude that a long-run relationship exists between government revenue and expenditure for the set of EU15 countries, whatever the specification of the deterministic component.

Table 11 – Panel cointegration test results between government revenue and expenditure (Pedroni, 1999; 2004)

Period 1970-2006	ADF-stat	P-value	Bootstrap distribution 1%	5%	10%
Model with no deterministic component	-4.38	0.00	-4.88	-4.01	-3.52
Model with a constant term	-3.19	0.00	-4.25	-3.31	-2.82
Model including a time trend	-4.04	0.00	-5.62	-4.70	-4.03
Period 1970-1991			1%	5%	10%
Model with no deterministic component	-5.93	0.00	-7.63	-6.31	-5.63
Model with a constant term	-7.38	0.00	-6.68	-5.40	-4.72
Model including a time trend	-3.50	0.00	-7.56	-6.69	-5.09
Period 1992-2006			1%	5%	10%
Model with no deterministic component	-2.93	0.00	-6.78	-5.53	-4.87
Model with a constant term	-1.79	0.03	-7.78	-6.32	-5.62
Model including a time trend	-5.79	0.00	-9.22	-7.76	-6.98

Notes:
i) – The bootstrap is based on 2000 replications.
ii) – As the tests are one-sided, a calculated statistic smaller than the critical value leads to the rejection of the null hypothesis of no cointegration.

We then investigated the robustness of the previous results, implementing panel data cointegration tests for the two sub-periods 1970-1991 and 1992-2006. The results are easier to interpret and provide econometric elements that justify this split on the basis of economic and institutional grounds, as two different types of behaviour now emerge from the cointegration tests (see Table 11).

First, concerning the 1970-1991 period, if one considers a model with a constant term, a statistical cointegration relationship clearly exists between government revenue and expenditure ratios, irrespective of whether one considers the (asymptotic) p-value or bootstrap critical values at 1, 5 or 10 per cent. The opposite result is however obtained for a model including a time trend, independently of the critical values used (asymptotic or bootstrap). Finally, intermediate results are obtained for a model with no deterministic component, for which a long-run statistical relationship between government revenue and expenditure ratios only exists with the 10% bootstrap critical value.

Second, the results do not seem to confirm the existence of a cointegration relationship for the period 1992-2006 between government revenue and expenditure ratios in the EU15 panel data set. This result is valid for any specification of the deterministic component considered, and is robust to the critical value used (asymptotic or bootstrap) for the conventional levels of significance. In this context, we should recall that after the beginning of the new millennium, the EU faced an economic recession (mirroring the beginning of the 1990s), with several countries entering into an Excessive Deficit Procedure (EDP) situation within the fiscal framework of the SGP. The reason why some countries faced an EDP depended, to some extent, on the difficulties encountered in implementing sound fiscal policies in "good times" and thus the lack of budgetary manoeuvre in the recession period. Such developments may explain the different results regarding fiscal sustainability obtained in our analysis for this more recent period.

In order to assess the robustness of our findings, we also implemented the bootstrap panel cointegration test proposed by Westerlund and Edgerton (2007). Unlike the panel data cointegration tests of Pedroni (1999, 2004), here the null hypothesis is now cointegration. This new test relies on the popular Lagrange multiplier test of McCoskey and Kao (1998), and permits correlation to be accommodated both within and between the individual cross-sectional units. In addition, the bootstrap suggested by Westerlund and Edgerton (2007) is based on the sieve-sampling scheme, and has the appealing advantage of significantly reducing the distortions of the asymptotic test.[24] The results reported in Table 12 for a model including either a constant term or a linear trend clearly indicate the absence of a cointegrating relationship between government revenue and expenditure since with an asymptotic p-value of 0.00, the null hypothesis of cointegration is always rejected. This result is only marginally modified if one refers to the bootstrap critical value, indicating that for a significant level higher than 2 per cent, the null hypothesis is still rejected. Hence at the conventional 5 and 10 per cent levels of significance, we can conclude that there is no

[24] We are grateful to J. Westerlund for making available his GAUSS codes to us.

cointegrating relationship between government revenue and expenditure for the EU15 panel data set.

Interestingly, performing the panel data cointegration tests for the two sub-periods 1970-1991 and 1992-2006 produces strong evidence in favour of the existence of a cointegration relationship between government revenue and expenditure ratios for the model with a constant term, with bootstrap p-values of 44% for the period 1970-1991, and 16% for the period 1992-2006. Hence, the necessary condition for public finance sustainability, i.e. the existence of a cointegration relationship between government revenue and expenditure, seems to be verified for the two sub-periods using this bootstrap panel cointegration test.

Table 12 – Panel cointegration test results between government revenue and expenditure (Westerlund and Edgerton, 2007) [a]

Period 1970-2006	LM-stat	Asymptotic p-value	Bootstrap p-value
Model with a constant term	7.08	0.00	0.02
Model including a time trend	3.90	0.00	0.02
Period 1970-1991			
Model with a constant term	0.63	0.26	0.44
Model including a time trend	2.10	0.01	0.02
Period 1992-2006			
Model with a constant term	1.37	0.08	0.16
Model including a time trend	3.22	0.00	0.19

Note: the bootstrap is based on 2000 replications.
a) - The null hypothesis of the tests is cointegration between government revenue and expenditure.

We further investigated whether public finances were sustainable for the model including a constant term, following the methodology of Pedroni (2004) and using a t-statistic to test whether the panel cointegration coefficient of the general government expenditure-to-GDP ratios is equal to one or not in the cointegrating regression where the government revenue-to-ratio is the dependent variable. For the period 1970-2006, the calculated t-statistic of 5.03 is above the tabulated critical values extracted from the normal distribution (1.96 and 2.33 respectively at the 5 per cent and 1 per cent levels of significance). The confidence interval for this coefficient, at the 5 per cent level of significance, is [1.023; 1.136], which confirms that the value of the coefficient is likely to be higher than one. For the two sub-periods, the 5 per cent confident intervals for the coefficient are respectively [0.868; 1.072] for the period 1970-1991, and [0.678; 0.841] for the period 1992-2006. This therefore indicates that the coefficient in the cointegration relation is likely to be equal to one for the period 1970-1991, which provides evidence of the sustainability of public finances in that period.

Finally, we also tested, along the lines of MacDonald (1992), the possibility of cointegration between the primary balance ratio and the government debt-to-GDP ratio, which represents a possible avenue for assessing the sustainability of public finances, provided that both series are I(1) processes. However, the panel unit root tests for those series, as reported in Appendix D, show that while the government debt-to-GDP ratio is indeed I(1), the primary balance ratio is I(0), which thus excludes the possibility of the existence of a cointegration relationship between these two series.[25]

6. Conclusion

This paper has drawn on recent advances in the econometrics of non-stationary panel data methods to assess the sustainability of public finances for the EU15 countries in the period 1970-2006. Starting from the present value borrowing constraint of governments, we investigate past fiscal data to see if the stock of real government debt follows a stationary process, or if there is cointegration between government revenue and government expenditure as a percentage of GDP.

The econometric methods used in the paper to assess the sustainability of public finances in the EU15 rest upon (i) individual unit root tests allowing in some cases for structural breaks; (ii) first generation panel data integration tests that assume cross-sectional independence among panel units (apart from common time effects); (iii) two second generation panel data unit root tests that relax the assumption of cross-sectional independence; (iv) panel data unit root tests that enable to accommodate structural breaks, and (v) the panel data cointegration tests developed by Pedroni

[25] Similar results, not reported here, are obtained with the implementation of the panel data tests of the second generation by Moon and Perron, 2004 and Choi (2006).

(1999, 2004) and generalized by Banerjee and Carrion-i-Silvestre (2006), and the bootstrap panel cointegration test by Westerlund and Edgerton (2007).

The results from these panel unit root tests, allowing for structural breaks, support the results of both the first and second generation panel data unit root tests, leading us to conclude that the first difference of the stock of real government debt series is integrated of order zero, thus indicating that the solvency condition would be satisfied for EU15 countries, which is a necessary condition for fiscal policy sustainability. Moreover, our results also show that general government expenditure and revenue ratios are integrated of order one.

Even if the results of the analysis may question fiscal sustainability in some cases when taken individually, it is nevertheless true that the tests point to the solvency of government public finances when considering the EU15 panel data set. Naturally, this is an obvious advantage of the panel approach, since the time series dimension of the data is not that long for individual countries. Even if there is no single fiscal policy in the EU, the panel sustainability of public finances indicated by our results is relevant in a context of EU countries seeking to pursue sound fiscal policy behaviour within the Stability and Growth Pact framework. Nevertheless, what we can also conclude from our analysis is that for some particular cases sustainability will not be attained if past fiscal behaviour is to be kept unchanged in the future. For instance, and as we saw, the solvency condition, on the basis of the stationarity tests of government debt, was satisfied for roughly half of the 15 EU countries: Austria, Finland, France, Germany, the Netherlands, the UK and Sweden. This set of countries is even smaller once we take into account the existence of structural breaks in the series.

Interestingly, the panel cointegration results for the entire 1970-2006 period allow us to draw the conclusion that a long-run relationship does exist between general government revenue and expenditure ratios for the set of EU15 countries, at least at the 10 per cent level of significance, both using conventional (asymptotic) critical values given in Pedroni (1999), and bootstrap panel cointegration proposed by Westerlund and Edgerton (2007). Moreover, this conclusion holds for the two sub-periods, 1970-1991 and 1992-2006 (broadly before and after the Maastricht Treaty), for most of the cointegration tests carried out.

Naturally, one has to stress that in this paper we assessed fiscal sustainability taking into account the stock of explicit government debt, and also via the analysis of cointegration relationships between the flows of government expenditures and revenues. Other aspects, outside the scope of analysis of the paper, and which are also relevant for the sustainability of public finances, are on the one hand the existence of implicit government liabilities, and on the other hand population ageing in combination with insufficiently funded public pension schemes that may endanger fiscal sustainability in the future.

References

Afonso, A. (2005a). "Fiscal Sustainability: The Unpleasant European Case", *FinanzArchiv*, 61 (1), 19-44.

Afonso, A. (2005b). "Ricardian Fiscal Regimes in the European Union", European Central Bank, Working Paper 558, November.

Afonso, A., Gomes, P. and Rother, P. (2007). "What 'Hides' behind Sovereign Debt Ratings?" European Central Bank, Working Paper 711, February.

Ahmed, S. and Rogers, J. (1995). "Government Budget Deficits and Trade Deficits. Are Present Value Constraints Satisfied in Long-term Data?", *Journal of Monetary Economics*, 36 (2), 351-374.

Banerjee, A. (1999). "Panel Data Units and Cointegration: An Overview", *Oxford Bulletin of Economics and Statistics*, 61 (3), 607-629.

Banerjee, A., Marcellino, M. and Osbat, C. (2004). "Some Cautions on the Use of Panel Methods for Integrated Series of Macro-economic Data", *Econometrics Journal*, 7 (2), 322–340.

Banerjee, A., Marcellino, M. and Osbat, C. (2005). "Testing for PPP: Should We Use Panel Methods?", *Empirical Economics*, 30 (1), 77-91.

Banerjee, A. and Carrion-i-Silvestre, J. (2006). "Cointegration in Panel Data with Breaks and Cross-section Dependence", European Central Bank, Working Paper 591, February.

Bergman, M. (2001). "Testing Government Solvency and the No Ponzi Game Condition", *Applied Economics Letters*, 8 (1), 27-29.

Bohn, H. (1998). "The Behavior of U.S. Public Debt and Deficits", *Quarterly Journal of Economics*, 113 (3), 949-963.

Bravo, A. and Silvestre, A. (2002). "Intertemporal Sustainability of Fiscal Policies: Some Tests for European Countries", *European Journal of Political Economy*, 18 (3), 517-528.

Breuer, J.; McNown, R. and Wallace, M. (2002). "Series-specific unit root tests with panel data", *Oxford Bulletin of Economics and Statistics* 64, 527–546.

Breitung, J. (2000). "The Local Power of Some Unit Root Tests for Panel Data" in Baltagi, B. (ed.), *Nonstationary Panels, Panel Cointegration, and Dynamic Panels*. *Advances in Econometrics*, Vol. 15. Amsterdam: JAI.

Breitung, J. (2005). "A Parametric Approach to the Estimation of Cointegration Vectors in Panel Data", *Econometric Reviews*, 24 (2), 151-173.

Breitung, J. and Pesaran, M. (2005). "Unit Roots and Cointegration in Panels", IEPR Working Paper No 05.32.

Chang, Y., Park, J. and Song, K. (2006). "Bootstrapping Cointegrating Regressions", Rice University*, Journal of Econometrics*, 133 (2), 703-739.

Choi, I. (2006). "Combination Unit Root Tests for Cross-sectionally Correlated Panels", in Corbae, D., Durlauf, S. and Hansen, B. (eds), *Econometric Theory and Practice: Frontiers of Analysis and Applied Research*, *Essays in Honor of Peter C. B. Phillips*, Cambridge: Cambridge University Press.

Claeys, P. (2007). "Sustainability of EU Fiscal Policies: A Panel Test", Institut de Recerca en Economia Aplicada, University of Barcelona, Documents de Treball 2007/02.

Domar, E. (1944). "The 'Burden of Debt' and the National Income", *American Economic Review*, 34 (4), 798-827.

Elliott, G., Rothenberg, T. and Stock, J. (1996). "Efficient Tests for an Autoregressive Unit Root", *Econometrica*, 64 (4), 813-836.

Fève, P. and Hénin, P. (2000). "Assessing Effective Sustainability of Fiscal Policy within the G-7", *Oxford Bulletin of Economic Research*, 62 (2), 175-195.

Fisher, R. (1932). *Statistical Methods for Research Workers*, London: Oliver and Boyd.

Groen, J. and Kleibergen, F. (2003). "Likelihood-based Cointegration Analysis in Panels of Vector Error-correction Models", *Journal of Business and Economic Statistics*, 21 (2), 295-318.

Gutierrez, L. (2006). "Panel Unit Roots Tests for Cross-sectionally Correlated Panels: A Monte Carlo Comparison", *Oxford Bulletin of Economics and Statistics*, 68 (4), 519-540.

Hadri, K. (2000). "Testing for Stationarity in Heterogeneous Panel Data", *Econometrics Journal*, 3 (2), 148-161.

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

Hamilton, J. and Flavin, M. (1986). "On the Limitations of Government Borrowing: A Framework for Empirical Testing", *American Economic Review*, 76 (4), 808-816.

Haug, A. (1991). "Cointegration and Government Borrowing Constraints: Evidence for the United States", *Journal of Business & Economic Statistics*, 9 (1), 97-101.

Im, K. and Lee, J. (2001). "Panel LM Unit Root Test with Level Shifts", Discussion paper, Department of Economics, University of Central Florida.

Im, K., Pesaran, M. and Shin, Y. (2003). "Testing for Unit Roots in Heterogeneous Panels", *Journal of Econometrics*, 115 (1), 53-74.

Kao, C. (1999). "Spurious Regression and Residual-based Tests for Cointegration in Panel Data", *Journal of Econometrics*, 90 (1), 1-44.

Keynes, J. (1923). "A Tract on Monetary Reform", in *The Collected Writings of John Maynard Keynes*", Vol. IV, London: Macmillan, 1971 edition.

Kremers, J. (1988). "The Long-run Limits of U.S. Federal Debt", *Economics Letters*, 28 (3), 259-262.

Kwiatkowski, D., Phillips, P. and Shin, Y. (1992). "Testing for the Null Hypothesis of Stationarity against the Alternative of a Unit Root", *Journal of Econometrics*, 54 (1-3), 159-178.

Larsson, R., Lyhagen, J. and Löthgren, M. (2001). "Likelihood-based Cointegration Tests in Heterogenous Panels", *Econometrics Journal*, 4 (1), 109-142.

Lee, J. and Strazicich, M. (2003). "Minimum Lagrange Multiplier Unit Root Test with Two Structural Breaks", *Review of Economics and Statistics*, 85 (4), 1082-1089.

Levin, A., Lin, C.-F. and Chu, C.-S. (2002). "Unit Root Tests in Panel Data: Asymptotic and Finite Sample Properties", *Journal of Econometrics*, 108 (1), 1-24.

Lumsdaine, R. and Papell, D. (1997). "Multiple Trend Breaks and the Unit Root Hypothesis", *Review of Economics and Statistics*, 79, 212-218.

Lyhagen, J. (2000). "Why Not Use Standard Panel Unit Root Test for Testing PPP", Stockholm School of Economics Working Paper No 413.

McCallum, B. (1984). "Are Bond-financed Deficits Inflationary? A Ricardian Analysis", *Journal of Political Economy*, 92 (1), 123-135.

McCoskey, S. and Kao, C. (1998). "A Residual-Based Test of the Null of Cointegration in Panel Data", *Econometric Reviews*, 17 (1), 57-84.

MacDonald, R. (1992). "Some Tests of the Government's Intertemporal Budget Constraint using US Data", *Applied Economics*, 24 (12), 1287-1292.

Maddala, G. and Wu, S. (1999). "A Comparative Study of Unit Root Tests and a New Simple Test", *Oxford Bulletin of Economics and Statistics*, 61 (1), 631-652.

Moon, H. and Perron, B. (2004). "Testing for a Unit Root in Panels with Dynamic Factors", *Journal of Econometrics*, 122 (1), 8-126.

Morris, R.; Ongena, H. and Schuknecht, L. (2006). "The reform and implementation of the Stability and Growth Pact", ECB Occasional Paper N. 47.

Ng, S. and Perron, P. (2001). “Lag Length Selection and the Construction of Unit Root Tests with Good Size and Power”, *Econometrica*, 69 (6), 1519-1554.

O’Connell, P. (1998). “The Overvaluation of Purchasing Power Parity”, *Journal of International Economic*s, 44 (1), 1-19.

Park, J. (2002). “An Invariance Principle for Sieve Bootstrap in Time Series”, *Econometric Theory*, 18 (2), 469-490.

Payne, J. (1997). “International Evidence on the Sustainability of Budget Deficits”, *Applied Economics Letters*, 12 (4), 775-779.

Pedroni, P. (1999). “Critical Values for Cointegrating Tests in Heterogeneous Panels with Multiple Regressors”, *Oxford Bulletin of Economics and Statistics*, 61 (1), 653-670.

Pedroni, P. (2000). “Fully Modified OLS for Heterogeneous Cointegrated Panels”, *Advances in Econometrics*, 15, 93-130.

Pedroni, P. (2004). “Panel Cointegration; Asymptotic and Finite Sample Properties of Pooled Time Series Tests with an Application to the Purchasing Power Parity Hypothesis”, *Econometric Theory*, 20 (3), 597-625.

Pedroni, P. and Urbain, J.-P. (2001). “Cross Member Cointegration in Non-stationary Panels”, mimeo, Universtiteit Maastricht.

Pesaran, M. (2004). “General Diagnostic Tests for Cross Section Dependence in Panels, Cambridge Working Papers in Economics, No 435, University of Cambridge.

Pesaran, M. (2007). “A Simple Panel Unit Root Test in the Presence of Cross Section Dependence”, *Journal of Applied Econometrics*, 22 (2), 265-312.

Pesaran, M., Shin, Y. and Smith, R. (1999). “Pooled Mean Group Estimation of Dynamic Heterogeneous Panels”, *Journal of the American Statistical Association*, 94 (446), 621-624.

Phillips, P. (2001). “Bootstrapping Spurious Regressions”, Cowles Foundation Discussion Paper No 1330. Yale.

Phillips, P. and Perron P. (1988). “Testing for a Unit Root in Time Series Regression”, *Biometrika*, 75 (2), 335-346.

Phillips, P. and Sul, D. (2003). “Dynamic Panel Estimation and Homogeneity Testing under Cross-section Dependence”, *Econometrics Journal,* 6 (1), 217-259.

Prohl, S. and Schneider, F. (2006). “Sustainability of Public Debt and Budget Deficit: Panel Cointegration Analysis for the European Union Member Countries”, Department of Economics, Johannes Kepler University Linz, Working Paper No 0610.

Quintos, C. (1995). “Sustainability of the Deficit Process with Structural Shifts,” *Journal of Business & Economic Statistics*, 13 (4), 409-417.

Schmidt, P. and Phillips, P. (1992). “LM Tests for a Unit Root in the Presence of Deterministic Trends”,*Oxford Bulletin of Economics and Statistics*, 54, 257-287.

Tanner, E. and Liu, P. (1994). “Is the Budget Deficit ‘Too Large?’: Some Further Evidence”, *Economic Inquiry*, 32, 511-518.

Taylor, A. and Taylor, M. (2004). “The Purchasing Power Parity Debate”, *Journal of Economic Perspectives*, 18 (4)-Fall.

Taylor, M. and Sarno, L. (1998). “The Behavior of Real Exchange Rates During the Post-Bretton Woods Period”, *Journal of International Economics*, 46 (2), 281-312.

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

Uctum, M. and Wickens, M. (2000). “Debt and Deficit Ceilings, and Sustainability of Fiscal Policies: An Intertemporal Analysis”, *Oxford Bulletin of Economic Research*, 62 (2), 197-222.

Westerlund, J. (2005). “New Simple Tests for Panel Cointegration”, *Econometric Reviews*, 24 (3), 297-316.

Westerlund, J. (2007), “Panel Cointegration Tests of the Fisher Hypothesis”, *Journal of Applied Econometrics*, forthcoming.

Westerlund, J. and Edgerton, D. (2007). “A Panel Bootstrap Cointegration Test”, *Economics Letters*, forthcoming.

Westerlund, J. and Prohl, S. (2006). “Panel Cointegration Tests of the Sustainability Hypothesis in Rich OECD Countries”, mimeo.

Wilcox, D. (1989). “The Sustainability of Government Deficits: Implications of the Present-value Borrowing Constraint”, *Journal of Money, Credit and Banking*, 21 (3), 291-306.

Zivot, E. and Andrews, D. (1992). “Further Evidence of the Great Crash, the Oil-price Shock and the Unit-root Hypothesis”, *Journal of Business and Economic Statistics*, 10, 251-270.

Appendix A. Data sources

Original series	AMECO codes *
Gross domestic product, at current market prices, national currency	.1.0.0.0.UVGD
Gross domestic product, at 2000 market prices, national currency	.1.1.0.0.OVGD
General government consolidated gross debt, Excessive Deficit Procedure (based on ESA 1995) and former definition (linked series) (% of GDP)	.1.0.319.0.UDGGL .1.0.319.0.UDGGF
General government debt (level)	.1.0.0.0.UDGGL .1.0.0.0.UDGGF
General government total expenditure (% of GDP)	.1.0.319.0.UUTGE .1.0.319.0.UUTGF
General government total revenue (% of GDP)	.1.0.319.0.URTG .1.0.319.0.URTGF
General government interest payments (% of GDP)	.1.0.319.0.UYIG .1.0.319.0.UYIGF

Note: * series from the European Commission AMECO database (updated on 04/05/2007).

Appendix B. Standard individual unit root test results

Table B1 – Stationarity tests for the first difference of the stock of government debt with constant (at 2000 prices) [a]

Country	Period	Lags	ADF	PP [b]	KPSS	ERS	NG-PERRON [c]			
			P value	P value for Adj t-Stat	LM-Statistic for level stationarity [d]					
Austria	1970-2006	2	0.0014	0.0104	0.514653	1.485594	-12.3621	-2.48100	0.20069	2.00192
Belgium	1970-2006	1	0.4105	0.4384	0.213485	5.969791	-4.63482	-1.50390	0.32448	5.32258
Denmark	1971-2006	1	0.4656	0.4012	0.370990	4.412965	-6.79899	-1.58747	0.23349	4.43042
Finland	1970-2006	2	0.0104	0.1365	0.171744	1.016047	-24.0726	-3.46576	0.14397	1.02967
France	1977-2006	3	0.0106	0.1082	0.401339	0.886443	-7.93293	-1.96910	0.24822	3.17086
Germany	1970-2006	1	0.0516	0.0582	0.414604	3.104104	-10.0108	-2.22845	0.22260	2.48166
Greece	1970-2006	2	0.8076	0.2683	0.624108	28.58263	-0.86369	-0.52257	0.60504	20.4395
Ireland	1970-2006	2	0.3293	0.0005	0.216461	5.823934	-4.48445	-1.49703	0.33383	5.46400
Italy	1970-2006	1	0.4311	0.4207	0.326493	10.59323	-3.18130	-1.14895	0.36116	7.54889
Luxembourg	1970-2006	2	0.6856	0.0050	0.510071	16.06996	9.36789	1.96072	0.20930	14.5379
Netherlands	1975-2006	1	0.0695	0.0640	0.275236	2.425800	-10.8222	-2.27744	0.21044	2.45088
Portugal	1973-2006	1	0.2117	0.2835	0.621288	5.408992	-7.64496	-1.68737	0.22072	4.11891
Spain	1970-2006	1	0.1552	0.1530	0.257055	3.582032	-7.97205	-1.96187	0.24609	3.20160
Sweden	1970-2006	1	0.0480	0.0433	0.102106	2.282042	-11.6127	-2.37962	0.20492	2.22602
United Kingdom	1970-2006	3	0.0007	0.2068	0.224438	1.988533	-16.7307	-2.78151	0.16625	1.86568

a – Note that the null hypothesis of all tests is that the considered series has a unit root except for the Kwiatkowski-Phillips-Schmidt-Shin test, where it is stationarity around a constant. The lag length in the ADF regression is based on the Schwartz Information Criterion with a maximum lag of 9.

b – Bandwidth: Newey-West using a Bartlett kernel.

c –

		MZa	MZt	MSB	MPT
Critical values:	1%	-13.8	-2.58	0.17	1.78
	5%	-8.10	-1.98	0.23	3.17
	10%	-5.70	-1.62	0.27	4.45

Lag length: Spectral GLS-detrended AR based on the Schwartz Information Criterion with a maximum lag of 9.

d – The critical values provided by Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1) are respectively 0.739 (1 per cent level), 0.463 (5 per cent level) and 0.347 (10 per cent level) for the LM test for level stationarity.

e – The critical values extracted from Elliott-Rothenberg-Stock (1996, Table 1) are respectively 1.87 (1 per cent level), 2.97 (5 per cent level) and 3.91 (10 per cent level). Note that as it is a one-sided unit root test, with the alternative H_1:$\rho<1$, a calculated statistic smaller than the tabulated critical value will lead to the rejection of the null H_1:$\rho=1$.

Table B2 – Stationarity tests for government revenues with constant (percent of GDP) [a]

			ADF	PP [b]	KPSS	ERS	NG-PERRON [c]			
Country	Period	Lags	P value	P value for Adj t-Stat	LM-Statistic for level stationariy [d]					
Austria	1970-2006	0	0.0416	0.0461	0.481611	82.42092	-0.54676	-0.45836	0.83832	36.0130
Belgium	1970-2006	0	0.0871	0.7574	0.688033	49.86760	-0.84923	-0.56183	0.66158	23.1942
Denmark	1971-2006	2	0.5809	0.3288	0.615298	53.90938	-0.62071	-0.41985	0.67640	25.3609
Finland	1970-2006	0	0.1384	0.0320	0.602118	56.88679	-0.56931	-0.40385	0.70936	27.5648
France	1977-2006	0	0.2649	0.2874	0.627758	143.0834	-0.550	0.55602	1.01094	64.5458
Germany	1970-2006	0	0.0104	0.0116	0.250752	29.50308	-2.10129	-1.00725	0.47935	11.4929
Greece	1970-2006	0	0.8580	0.8447	0.671432	109.2311	0.03749	0.02851	0.76040	35.4385
Ireland	1970-2006	0	0.4052	0.3947	0.282870	17.05027	-2.06819	-0.99790	0.48250	11.6596
Italy	1970-2006	1	0.6334	0.6358	0.660573	104.7556	0.34064	0.34222	1.00466	60.9516
Luxembourg	1970-2006	1	0.0326	0.0155	0.468933	49.88170	0.34083	1.21833	3.57456	692.068
Netherlands	1975-2006	1	0.2414	0.2264	0.201091	29.97768	-1.40949	-0.83491	0.59235	17.2672
Portugal	1973-2006	2	0.7474	0.7953	0.716110	279.0827	0.84760	0.78794	0.92961	59.1332
Spain	1970-2006	1	0.2669	0.3457	0.650818	287.1239	-1.15793	-0.55955	0.48323	14.5698
Sweden	1970-2006	1	0.0865	0.0166	0.514078	42.52085	-1.02022	-0.61825	0.60599	19.7016
United Kingdom	1970-2006	1	0.0148	0.0171	0.081815	4.850811	-7.97068	-1.99555	0.25036	3.07668

a – Note that the null hypothesis of all tests is that the considered series has a unit root except for the Kwiatkowski-Phillips-Schmidt-Shin test, where it is stationarity around a constant. The lag length in the ADF regression is based on the Schwartz Information Criterion with a maximum lag of 9.

b – Bandwidth: Newey-West using a Bartlett kernel.

c –

		MZa	MZt	MSB	MPT
Critical values:	1%	-13.8	-2.58	0.17	1.78
	5%	-8.10	-1.98	0.23	3.17
	10%	-5.70	-1.62	0.27	4.45

Lag length: Spectral GLS-detrended AR based on the Schwartz Information Criterion with a maximum lag of 9.

d – The critical values provided by Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1) are respectively 0.739 (1 per cent level), 0.463 (5 per cent level) and 0.347 (10 per cent level) for the LM test for level stationarity.

e – The critical values extracted from Elliott-Rothenberg-Stock (1996, Table 1) are respectively 1.87 (1 per cent level), 2.97 (5 per cent level) and 3.91 (10 per cent level). Note that as it is a one-sided unit root test, with the alternative H_1:ρ<1, a calculated statistic smaller than the tabulated critical value will lead to the rejection of the null H_1:ρ=1.

Table B3 – Stationarity tests for government expenditures with constant (percent of GDP) [a]

			ADF	PP [b]	KPSS	ERS	NG-PERRON [c]			
Country	Period	Lags	P value	P value for Adj t-Stat	LM-Statistic for level stationarity [d]					
Austria	1970-2006	0	0.0590	0.0816	0.412088	74.40395	-0.62648	-0.51475	0.82166	34.1304
Belgium	1970-2006	2	0.2681	0.1980	0.167964	13.33629	-3.74537	-1.36833	0.36534	6.54146
Denmark	1971-2006	1	0.3221	0.2106	0.446637	25.34413	-1.94339	-0.96343	0.49575	12.3474
Finland	1970-2006	1	0.2525	0.3533	0.518290	12.86527	-2.53376	-1.07174	0.42298	9.40257
France	1977-2006	5	0.3261	0.2434	0.628887	250.1467	-0.47147	-0.28933	0.61368	22.8458
Germany	1970-2006	0	0.0155	0.0151	0.240678	26.51387	-2.24493	-1.02856	0.45817	10.6777
Greece	1970-2006	0	0.4561	0.4232	0.240678	75.70858	-0.20703	-0.19007	0.91808	45.7349
Ireland	1970-2006	1	0.5114	0.5985	0.640310	7.403980	-3.76424	-1.35503	0.35997	6.51600
Italy	1970-2006	0	0.1620	0.1862	0.233897	72.26970	-0.29061	-0.23057	0.79340	35.1634
Luxembourg	1970-2006	1	0.1788	0.0080	0.424445	30.89793	-1.11555	-0.55193	0.49476	15.0782
Netherlands	1975-2006	1	0.4848	0.4668	0.200853	11.97715	-2.86158	-1.19209	0.41658	8.54979
Portugal	1973-2006	0	0.2739	0.0104	0.675697	128.6005	0.66476	0.61917	0.93142	57.1760
Spain	1970-2006	1	0.3786	0.2850	0.539324	71.11501	-2.14546	-0.95815	0.44659	10.7549
Sweden	1970-2006	1	0.1732	0.1886	0.381379	14.59689	-3.44981	-1.29807	0.37627	7.09510
United Kingdom	1970-2006	2	0.0184	0.0711	0.080455	0.477187	-62.4567	-5.55697	0.08897	0.46388

a – Note that the null hypothesis of all tests is that the considered series has a unit root except for the Kwiatkowski-Phillips-Schmidt-Shin test, where it is stationarity around a constant. The lag length in the ADF regression is based on the Schwartz Information Criterion with a maximum lag of 9.

b – Bandwidth: Newey-West using a Bartlett kernel.

c –

		MZa	MZt	MSB	MPT
Critical values:	1%	-13.8	-2.58	0.17	1.78
	5%	-8.10	-1.98	0.23	3.17
	10%	-5.70	-1.62	0.27	4.45

Lag length: Spectral GLS-detrended AR based on the Schwartz Information Criterion with a maximum lag of 9.

d – The critical values provided by Kwiatkowski-Phillips-Schmidt-Shin (1992, Table 1) are respectively 0.739 (1 per cent level), 0.463 (5 per cent level) and 0.347 (10 per cent level) for the LM test for level stationarity.

e – The critical values extracted from Elliott-Rothenberg-Stock (1996, Table 1) are respectively 1.87 (1 per cent level), 2.97 (5 per cent level) and 3.91 (10 per cent level). Note that as it is a one-sided unit root test, with the alternative H_1:ρ<1, a calculated statistic smaller than the tabulated critical value will lead to the rejection of the null H_1:ρ=1

Appendix C. Results of individual unit root tests with breaks

Table C1 – Tests for structural change in the first difference of the stock of government debt with constant, at 2000 prices (Innovational Outlier Model)

		ZIVOT and ANDREWS (1992)			LUMSDAINE and PAPELL (1997)			
Country	Period	Lags	Break date TB	ADF[a] break point test	Lags	Break date[b] TB_1	Break date[c] TB_2	ADF break point test[d]
Austria	1970-2006	8	1991	-5.45	8	1992	2002	-6.08
Belgium	1970-2006	7	1992	-5.66	7	1990	1992	-4.66
Denmark	1971-2006	1	1997	-3.50	1	1997	2003	-2.84
Finland	1970-2006	3	1991**	-8.47###	1	1991**	2003**	-12.67###
France	1977-2006	4	1982	-3.09	7	1982	1993	-5.35
Germany	1970-2006	5	1991	-4.52	7	1991**	1994**	-8.57##
Greece	1970-2006	6	1988	-3.76	7	1990	1992	-4.18
Ireland	1970-2006	2	1992	-3.82	1	1982	1987	-4.85
Italy	1970-2006	8	1991	-5.03	8	1991	2002	-5.95
Luxembourg	1970-2006	7	1986	-5.12	3	1987	2002	-2.55
Netherlands	1975-2006	7	1993	-4.13	7	1996	1999	-6.51
Portugal	1973-2006	6	2001	-4.71	6	1989	2001	-5.55
Spain	1970-2006	8	1987	-3.46	8	1987	1992	-7.36
Sweden	1970-2006	2	1990	-5.25	7	1990	2001	-740
United Kingdom	1970-2006	3	1987**	-7.12###	3	1991	2002	-7.05#

a – The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Zivot and Andrews (1992, p. 262) and are respectively -6.68 (1 per cent level), -5.82 (5 per cent level) and -5.37 (10 per cent level).

b – ** denotes statistical significance of the first structural break at the 5 per cent level of significance.

c – ** denotes statistical significance of the second structural break at the 5 per cent level of significance.

d - The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Lumsdaine and Papell (1997), and are respectively -8.78 (1 per cent level), -7.47 (5 per cent level), and -6.98 (10 per cent level).

#, ##, ### denote statistical significance at the 10, 5 and 1 per cent levels of significance respectively.

** denotes statistical significance of the structural break at the 5 per cent level of significance.

Table C2 – Tests for structural change in government revenues with constant (as a percentage of GDP)

		ZIVOT and ANDREWS (1992)			LUMSDAINE and PAPELL (1997)			
Country	Period	Lags	Break date TB	ADF[a] break point test	Lags	Break date TB_1	Break date TB_2	ADF break point test[b]
Austria	1970-2006	0	1983	-2.68	7	1986	1992	-4.96
Belgium	1970-2006	7	1988	-3.70	4	1981	1995	-4.27
Denmark	1971-2006	0	1983	-3.62	1	1983	1989	-3.61
Finland	1970-2006	4	1993**	-7.03###	4	1981	1989	-4.52
France	1977-2006	6	1986**	-6.85###	8	1992	2000	-5.45
Germany	1970-2006	7	1985	-3.59	2	1991	2000	-5.27
Greece	1970-2006	8	1988	-3.69	8	1986	1996	-4.71
Ireland	1970-2006	6	1990	-5.06	6	1982	1986	-4.39
Italy	1970-2006	2	1994	-4.10	7	1989	2001	-3.64
Luxembourg	1970-2006	4	1987**	-5.89##	4	1985	1997	-6.16
Netherlands	1975-2006	2	1981	-2.94	5	1988	1997	-6.47
Portugal	1973-2006	1	1990**	-6.66##	1	1982	1990	-5.81
Spain	1970-2006	8	1991**	-7.35###	1	1994	2000	-4.18
Sweden	1970-2006	1	1984	-4.17	1	1991	2001	-7.09 #
United Kingdom	1970-2006	7	1981	-4.66	1	1981	2001	-5.52

a – The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Zivot and Andrews (1992, p. 262) and are respectively -6.68 (1 per cent level), -5.82 (5 per cent level) and -5.37 (10 per cent level).

b – ** denotes statistical significance of the first structural break at the 5 per cent level of significance.

c – ** denotes statistical significance of the second structural break at the 5 per cent level of significance.

d - The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Lumsdaine and Papell (1997), and are respectively -8.78 (1 per cent level), -7.47 (5 per cent level), and -6.98 (10 per cent level).

#, ##, ### denote statistical significance at the 10, 5 and 1 per cent levels of significance respectively.

** denotes statistical significance of the structural break at the 5 per cent level of significance.

Table C3 – Tests for structural change in government expenditures with constant (as a percentage of GDP)

		ZIVOT and ANDREWS (1992)			LUMSDAINE and PAPELL (1997)			
Country	Period	Lags	Break date TB	ADF[a] break point test	Lags	Break date TB_1	Break date TB_2	ADF break point test[b]
Austria	1970-2006	4	1986	-3.61	8	1986	1999	-4.59
Belgium	1970-2006	1	1981	-3.76	1	1981	1994	-6.12
Denmark	1971-2006	1	1989	-3.03	3	1987	1993	-3.09
Finland	1970-2006	1	1989	-5.65	4	1989	1997	-6.42
France	1977-2006	3	1992	-3.37	8	1981	1992	-4.65
Germany	1970-2006	4	1991	-5.09	4	1991	1994	-5.23
Greece	1970-2006	0	1986	-3.82	1	1985	2002	-2.72
Ireland	1970-2006	5	1984	-4.51	8	1981	1996	-5.24
Italy	1970-2006	4	1989	-4.40	1	1981	1996	-4.46
Luxembourg	1970-2006	8	1988	-5.23	8	1992	1998	-5.22
Netherlands	1975-2006	1	1981	-3.42	6	1981	1996	-4.94
Portugal	1973-2006	1	1982	-3.94	4	1986	1996	-4.87
Spain	1970-2006	4	1991	-4.98	1	1993	1996	-2.42
Sweden	1970-2006	1	1991	-3.87	6	1991	2003	-4.61
United Kingdom	1970-2006	6	1985	-4.56	6	1981	1999	-5.92

a – The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Zivot and Andrews (1992, p. 262) and are respectively -6.68 (1 per cent level), -5.82 (5 per cent level) and -5.37 (10 per cent level).

b – ** denotes statistical significance of the first structural break at the 5 per cent level of significance.

c – ** denotes statistical significance of the second structural break at the 5 per cent level of significance.

d - The 'exact' critical values are calculated based on 7,000 replications of a Monte Carlo simulation as described in Lumsdaine and Papell (1997), and are respectively -8.78 (1 per cent level), -7.47 (5 per cent level), and -6.98 (10 per cent level).

#, ##, ### denote statistical significance at the 10, 5 and 1 per cent levels of significance respectively.

** denotes statistical significance of the structural break at the 5 per cent level of significance.

Appendix D. Panel unit root tests, additional results

Table D1 – Panel data unit root tests for the government debt-to-GDP ratio (1970-2006)

Method	Statistic	P-value*	Cross-sections	Observ.
Null: Unit root (assumes common unit root process)				
Levin, Lin & Chu t*	0.54469	0.7070	15	525
Breitung t-stat	-1.30967	0.0952	15	510
Null: Unit root (assumes an individual unit root process)				
Im, Pesaran and Shin W-stat	1.68879	0.9544	15	525
ADF – Fisher Chi-square	24.5988	0.7443	15	525
PP – Fisher Chi-square	8.90810	0.9999	15	540
Null: No unit root (assumes a common unit root process)				
Hadri Z-stat	9.33368	0.0000	15	555

* Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
Automatic selection of maximum lags. Automatic selection of lags based on SIC: 0 to 2. Newey-West bandwidth selection using a Bartlett kernel

Table D2 – Panel data unit root tests for the primary balance-to-GDP ratio (1970-2006)

Method	Statistic	P-value*	Cross-sections	Observ.
Null: Unit root (assumes a common unit root process)				
Levin, Lin & Chu t*	-2.77057	0.0028	15	525
Breitung t-stat	-3.74204	0.0001	15	510
Null: Unit root (assumes an individual unit root process)				
Im, Pesaran and Shin W-stat	-4.34773	0.0000	15	525
ADF – Fisher Chi-square	76.1011	0.0000	15	525
PP – Fisher Chi-square	49.9668	0.0125	15	525
Null: No unit root (assumes a common unit root process)				
Hadri Z-stat	4.62685	0.0000	15	525

* Probabilities for Fisher tests are computed using an asymptotic Chi-square distribution. All other tests assume asymptotic normality.
Automatic selection of maximum lags. Automatic selection of lags based on SIC: 0 to 2. Newey-West bandwidth selection using a Bartlett kernel

DAVIDSON INSTITUTE WORKING PAPER SERIES - Most Recent Papers

The entire Working Paper Series may be downloaded free of charge at: www.wdi.umich.edu

CURRENT AS OF 10/17/07

Publication	Authors	Date
No. 893: What Do We Really Know About Fiscal Sustainability in the EU? A Panel Data Diagnostic	António Afonso and Christophe Rault	Oct 2007
No. 892: The Political Economy of Corruption and the Role of Financial Institutions	Kira Boerner and Christa Hainz	Oct 2007
No. 891: Growth, Volatility and Political Instability: Non-Linear time Series Evidence for Argentina 1896-2000	Nauro Campos and Menelaos Karanasos	Sept 2007
No. 890: Social Costs of Mass Privatization	David Stuckler and Lawrence King	Sept 2007
No. 889: A Rise By Any Other Name? Sensitivity of Growth Regressions to Data Source	Randall Filer, Dana Hajkova and Jan Hanousek	July 2007
No. 888: Mind the Gap! Social Capital, East and West	Jan Fidrmuc and Klarita Gerxhani	June 2007
No. 887: Ever Closer Union or Babylon Discord?	Jan Fidrmuc, Victor Ginsburgh and Schlomo Weber	July 2007
No. 886: FDI & the Consequences towards more complete capture of spillover effects	Bruno Merlevede and Koen Schoors	Aug 2007
No. 885: Consumption Smoothing and Vulnerability in Russia	Christopher J. Gerry and Carmen A. Li	July 2007
No. 884: National Cultural & Financial Systems	Solomon Tadesse and Chuck Kwok	March 2005
No. 883: Stock Markets Liquidity, Corporate Governance and Small Firms	Solomon Tadesse	June 2005
No. 882: The MNC as an Agent of Change for Host-Country Institutions: FDI & Corruption	Chuck Kwok and Solomon Tadesse	Sept 2006
No. 881: The Allocation and Monitoring Role of Capital Markets: Theory and International Evidence	Solomon Tadesse	March 2003
No. 880: Tunisia: Sources of Real Exchange Rate Fluctuations	Sfia Mohamed Daly	March 2006
No. 879: Financial Development & Technology	Solomon Tadesse	June 2007
No. 878: Consolidation, Scale Economics & Technological Change in Japanese Banking	Solomon Tadesse	Feb 2005
No. 877: Innovation, Information and Financial Architecture	Solomon Tadesse	June 2007
No. 876: Corporate Cash Holdings, National Culture, and Multinationality	Andres Ramirez and Solomon Tadesse	June 2007
No. 875: The Economic Value of Regulated Disclosure: Evidence from the Banking Sector	Solomon Tadesse	Jan 2006
No. 874:Banking Fragility & Disclosure: International Evidence	Solomon Tadesse	Sept 2006
No. 873: The Impact of Outward FDI on Home-Country Employment in a Low-Cost Transition Economy	Jaan Masso, Urmas Varblane and Priit Vahter	May 2007
No. 872: Local Distributional Effects of Government Cash Transfers in Chile	Claudio A. Agostini and Philip Brown	May 2007
No. 871: How do Workers Fare During Transition? Perceptions of Job Insecurity among Russian Workers, 1995-2004	Susan J. Linz and Anastasia Semykina	April 2007
No. 870: Does Reform Work? An Econometric Examination of the Reform-Growth Puzzle	Ian Babetskii and Nauro Campos	April 2007
No. 869: Perceptions and Behavior: Analyzing Wage Arrears in Russia	Susan Linz, Anastasia Semykina and Charles Petrin	June 2006
No. 868: The Endogeneity of Association Agreements and their Impact on Trade for Eastern Countries: Empirical Evidence for Romania	Christophe Rault, Ana Maria Sova and Robert Sova	April 2007
No. 867: Institutions & Entrepreneurship Development in Russia: A Comparative Perspective	Saul Estrin, Ruta Aidis and Tomasz Mickiewicz	Feb 2007