

## Panel cointegration tests of the sustainability hypothesis in rich OECD countries

Westerlund, Joakim; Prohl, Silika

Postprint / Postprint

Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:

[www.peerproject.eu](http://www.peerproject.eu)

### Empfohlene Zitierung / Suggested Citation:

Westerlund, J., & Prohl, S. (2010). Panel cointegration tests of the sustainability hypothesis in rich OECD countries. *Applied Economics*, 42(11), 1355-1364. <https://doi.org/10.1080/00036840701721323>

### Nutzungsbedingungen:

Dieser Text wird unter dem "PEER Licence Agreement zur Verfügung" gestellt. Nähere Auskünfte zum PEER-Projekt finden Sie hier: <http://www.peerproject.eu> Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

**gesis**  
Leibniz-Institut  
für Sozialwissenschaften

### Terms of use:

This document is made available under the "PEER Licence Agreement". For more information regarding the PEER-project see: <http://www.peerproject.eu> This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Mitglied der  
  
Leibniz-Gemeinschaft

Diese Version ist zitierbar unter / This version is citable under:

<https://nbn-resolving.org/urn:nbn:de:0168-ssoar-250140>



### Panel Cointegration Tests of the Sustainability Hypothesis in Rich OECD Countries

Journal:	<i>Applied Economics</i>
Manuscript ID:	APE-06-0624.R1
Journal Selection:	Applied Economics
Date Submitted by the Author:	27-Jan-2007
Complete List of Authors:	Westerlund, Joakim; Lund University, Department of Economics Prohl, Silika; University of St. Gallen
JEL Code:	C22 - Time-Series Models < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods, C23 - Models with Panel Data < C2 - Econometric Methods: Single Equation Models < C - Mathematical and Quantitative Methods, H60 - General < H6 - National Budget, Deficit, and Debt < H - Public Economics
Keywords:	Fiscal Sustainability, Panel Cointegration, Structural Change



# Panel Cointegration Tests of the Sustainability Hypothesis in Rich OECD Countries\*

Joakim Westerlund<sup>†</sup> and Silika Prohl<sup>‡</sup>

January 27, 2007

## Abstract

This study reexamines the sustainability hypothesis by testing whether government revenues and expenditures for eight rich OECD countries between 1977Q1 and 2005Q4 are cointegrated. For this purpose, a non-stationary panel data approach is adopted, which is general enough to permit for cross-country dependence as well as structural breaks representing major shifts in fiscal policy. In contrast to many earlier studies, the results reported in this study suggest that the sustainability hypothesis cannot be rejected.

**JEL Classification:** C22; C23; H60.

**Keywords:** Fiscal sustainability; Panel Cointegration; Structural Change.

## 1 Introduction

The increased budget deficit and public debt experienced by many OECD countries raises the issue of sustainability of the government finance in the long run, and its potential effects on the economy as a whole. Fiscal policy is constrained by the need to finance the government deficit, which over the longer term implies that the market value of public debt must be offset by the present value of all discounted future budget surpluses. In other words, the fiscal policy is sustainable if the discounted value of debt is zero in the limit. This is the essence of the sustainability hypothesis. If the hypothesis holds, then the ongoing fiscal policy can, at least in principle, be maintained indefinitely, whereas if it fails, then there is a need for discretionary policy actions. In particular, a failure

---

\*Westerlund gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation, research grant number W2006-0068:1.

<sup>†</sup>Corresponding author: Department of Economics, Lund University, P. O. Box 7082, S-220 07 Lund, Sweden. E-mail address: joakim.westerlund@nek.lu.se.

<sup>‡</sup>SIAW - University of St. Gallen, Bodanstraße 8, CH-9000 St. Gallen, Switzerland. E-mail address: silika.prohl@unisg.ch.

1  
2  
3  
4  
5  
6  
7  
8 implies that the government will inevitably run into problems in managing its  
9 debt, thus requiring compensation in terms of higher interest rates leading to a  
10 slowdown in the economic growth. Thus, from a policy point of view, the issue  
11 of whether the sustainability hypothesis holds or not is indeed an important  
12 one. Because of this, and because of the growing concern regarding the debt  
13 situation in many OECD countries, sustainability analysis has received much  
14 interest in recent empirical research.

15  
16 Most of the earlier work in this field focused on the stationarity of public  
17 deficit and debt as a way to empirically test the sustainability hypothesis.  
18 Prominent examples include, among others, Wilcox (1989), MacDonald (1992),  
19 Hakkio and Rush (1991) and Quintos (1995) for the United States, Smith and  
20 Zin (1991) for Canada, Baglioni and Cherubini (1993) for Italy, Olekalns (2000)  
21 for Australia, and perhaps most recently Kirchgassner and Prohl (2006) for  
22 Switzerland. The overall picture being that the sustainability hypothesis does  
23 not hold.

24  
25 As a response to this, more recent work has moved away from examining the  
26 stationarity of debt, and towards more flexible testing strategies based on cointe-  
27 gration. In particular, as shown by Quintos (1995), given that public revenues  
28 and expenditures are non-stationary, sustainability requires these variables to  
29 be cointegrated with a unit slope on expenditures. In the terminology of Quin-  
30 tos (1995), this is known as strong sustainability. It means that no problems,  
31 according to the ongoing fiscal policy, are likely to arise in the future. By con-  
32 trast, weak sustainability refers to the case when the slope lies between zero and  
33 one, regardless of whether revenues and expenditures are cointegrated or not.  
34 In this case, although the deficit is sustainable, the government is expected to  
35 have difficulty in marketing its debt, thus requiring fiscal reforms. Finally, if  
36 the slope on expenditures is equal to zero, then the deficit is not sustainable, in  
37 which case the sustainability hypothesis is refuted.

38  
39 Payne (1997) applies this approach to the G7 countries between 1950 and  
40 1990, and rejects the sustainability hypothesis. Similarly, using the same ap-  
41 proach to a sample of five countries within the European Union between 1961  
42 and 1995, Papadopoulos and Sidiropoulos (1999) only find evidence of cointe-  
43 gration for Greece and Spain. However, the cointegrating slope on government  
44 expenditures is estimated significantly lower than one, suggesting that only the  
45 weak form of the sustainability hypothesis is supported. Apparently, despite  
46 the strong theoretical appeal of the sustainability hypothesis, most empirical  
47 studies tend to reject it.

48  
49 Although there are many explanations for why the sustainability hypothesis  
50 does not seem to hold, this paper focuses on two that might potentially go a  
51 long way towards explaining the weak empirical evidence on the cointegration  
52 between revenues and expenditures. The first explanation is that conventional  
53 time series cointegration tests may have low power against persistent alterna-  
54 tives because of the short sample periods usually employed, see Afonso (2005).  
55  
56

1  
2  
3  
4  
5  
6  
7  
8 The second explanation is that the empirical relationship between public revenues and expenditures that is being estimated need not be invariant to policy regime changes, which may lead to structural shifts in both revenues and expenditures, and hence in the relationship between them, see for example Quintos (1995) and Martin (2000).

9  
10  
11  
12  
13 However, while very reasonable and potentially appealing, when analyzed separately, these explanations have been found to be empirically inadequate and far from convincing. This paper therefore offers an alternative route. The idea is that to be able to provide any robust evidence on the cointegration between revenues and expenditures, one needs to consider not the low power and the presence of structural change separately but simultaneously. The intuition is simple. On the one hand, while increasing the length of the sample may be justified from a power point of view, it also increases the probability of a break. On the other hand, modifying existing tests so as to accommodate for structural change is typically very costly in terms of power. Therefore, it is the joint consideration of both these aspects that it is likely to be key when testing the sustainability hypothesis.

14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26 One way to accomplish this is to resort to panel data, and the recent econometrical advances within this field that make it possible to construct powerful cointegration tools while simultaneously entertaining the possibility of structural change. Applying this approach to a panel of eight rich OECD countries between 1977Q1 and 2005Q4, in contrast to much of the earlier evidence, we are unable to reject the strong sustainability hypothesis. We also provide evidence suggesting that all countries have been subject to several structural breaks representing major changes in fiscal policy.

27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40 The remainder of this paper is organized as follows. In Section 2, we give the brief description of the theoretical model of fiscal sustainability, and of the empirical panel data framework that will be used to test it. Section 3 then reports the empirical results, while Section 4 concludes.

## 41 2 Testing the sustainability hypothesis

42  
43  
44  
45  
46  
47  
48  
49 In this section, we begin with a brief account of the economic theory underlying the sustainability hypothesis, and then we go on to discuss the empirical method that we will use to test it.

### 50 2.1 Theoretical underpinnings

51  
52  
53  
54  
55  
56  
57  
58  
59  
60 The theoretical model of sustainability of the fiscal policy starts with the budget constraint of the government at time  $t$ , which is given in nominal terms by

$$g_t + (1 + i_t)b_{t-1} = r_t + b_t, \quad (1)$$

where  $b_t$  is the stock of public debt,  $i_t$  is the nominal interest rate payable on the public debt,  $r_t$  is the government revenue including revenue from seignorage, and  $g_t$  is the government expenditure excluding interest payments. Of course, these variables are relatively uninteresting as they do not take into account the size of the economy. We therefore follow the conventional practice of rewriting (1) in terms of GDP ratios as

$$\frac{g_t}{y_t} + \frac{(1+i_t)b_{t-1}}{(1+\sigma_t)y_{t-1}} = \frac{r_t}{y_t} + \frac{b_t}{y_t},$$

where  $y_t$  is the nominal GDP, while  $\sigma_t$  represents the nominal GDP growth rate. Using capital letters to denote the ratios of the corresponding upper-case variables to nominal GDP, we get

$$C_t + (1+i_t^*)B_{t-1} = R_t + B_t, \quad (2)$$

where  $C_t = G_t + (i_t^* - i^*)B_{t-1}$  and  $i^*$  is the mean of  $i_t^* = (i_t - \sigma_t)/(1 + \sigma_t)$ , the growth adjusted interest rate. If we assume that the budget constraint in (2) holds continuously at each  $t$ , we can use forward substitution to obtain

$$B_{t-1} = \sum_{j=t+1}^{\infty} \left( \frac{1}{1+i^*} \right)^{j-t} (R_{j-1} - C_{j-1}) + \lim_{j \rightarrow \infty} \left( \frac{1}{1+i^*} \right)^{j-t} B_{j-1}. \quad (3)$$

Now, consider taking expectations conditional on all the information available at time  $t$ . Since  $B_{t-1}$  is known at  $t$ , the hypothesis that the government obeys its intertemporal budget constraint can be expressed as

$$\lim_{j \rightarrow \infty} E_t \left( \frac{1}{1+i^*} \right)^{j-t} B_{j-1} = 0. \quad (4)$$

This transversality condition, also known as the no Ponzi game rule, says that the public debt-GDP ratio should grow no faster on average than the mean rate of interest, meaning that the principal repayments and interests cannot be financed indefinitely by issuing new debt. It then follows that the current stock of public debt is offset by all current and expected discounted future primary surpluses.<sup>1</sup>

## 2.2 A time series cointegration test

The empirical literature has developed several tests to investigate whether the transversality condition (3) holds. The particular test opted for in this paper is that of Quintos (1995), who suggests testing the sustainability hypothesis by examining whether government revenues and expenditures are cointegrated or not. The intuition behind this test lies in first taking differences, and then

<sup>1</sup>Note that  $1/(1+i^*) = (i+\sigma)/(1+i)$ , where  $i$  and  $\sigma$  are the mean values of  $i_t$  and  $\sigma_t$ , respectively. Thus, if  $i < \sigma$ , then the transversality condition is satisfied automatically.

rewriting (2) in terms of  $\tilde{G}_t = G_t + i_t^* B_{t-1}$ , the total expenditure including interest payments, as

$$\begin{aligned} \tilde{G}_{t-1} - R_{t-1} &= \sum_{j=t+1}^{\infty} \left( \frac{1}{1+i^*} \right)^{j-t} (\Delta R_{j-1} - \Delta C_{j-1}) \\ &+ \lim_{j \rightarrow \infty} \left( \frac{1}{1+i^*} \right)^{j-t} \Delta B_{j-1}. \end{aligned} \quad (5)$$

As before, the sustainability hypothesis holds if the expected value of the second term on the right-hand side is zero. One way to test whether this is in fact the case, given that  $\tilde{G}_t$  and  $R_t$  possess unit roots, involves first estimating and then testing for cointegration in the following regression

$$R_t = \alpha + \beta \tilde{G}_t + e_t, \quad (6)$$

where  $e_t$  is a mean zero disturbance term. To see what this means for the test of the sustainability hypothesis, note that  $\Delta B_t = \tilde{G}_t - R_t$ , which together with (5) implies that

$$\Delta B_t = (1 - \beta) \tilde{G}_t - \alpha - e_t. \quad (7)$$

As shown by Quintos (1995), although the sustainability hypothesis holds regardless of the integratedness of  $\Delta B_t$ , the speed with which this happens does not. If  $\Delta B_t$  is stationary, then the second term in (5) vanishes at a much faster rate than if  $\Delta B_t$  is non-stationary. In the terminology of Quintos (1995), the sustainability of the public deficit is said to be strong if  $\Delta B_t$  is stationary, whereas it is said to be weak if  $\Delta B_t$  is non-stationary.<sup>2</sup>

In terms of the regression in (6), by looking at (7), we see that cointegration with  $\beta = 1$  is enough for strong sustainability, because only then will  $\Delta B_t$  be stationary. But since  $\tilde{G}_t$  is non-stationary, the sustainability is weak if either  $\beta = 1$  and there is no cointegration or  $0 < \beta < 1$ . It follows that  $0 < \beta \leq 1$  is both necessary and sufficient for sustainability, while cointegration is only sufficient. The distinction between strong and weak sustainability is made because although  $0 < \beta < 1$  is enough for the sustainability hypothesis to hold, the government is now spending more than it earns, thus making the public debt difficult to market in the long run. Finally, if  $\beta \leq 0$  the deficit is no longer sustainable because this would make  $\Delta B_t$  grow faster than the growth rate of the economy, as approximated by the mean interest rate.<sup>3</sup>

The above discussion suggests that testing for cointegration is key in this kind of sustainability analysis, not only because of its role in determining strong sustainability but also because we cannot do inference regarding  $\beta$  unless revenues and expenditures are cointegrated.

<sup>2</sup>Bergman (2001) extends the results of Quintos (1995) by showing that sufficient condition for government solvency is that government debt is integrated of any finite order.

<sup>3</sup>Note that the case with  $\beta > 1$  is not consistent with a deficit, since revenues are growing faster than expenditures.



1  
2  
3  
4  
5  
6  
7  
8 The conventional way in which earlier studies have been trying to test the  
9 sustainability hypothesis within this framework involves first estimating (6) by  
10 ordinary least squares (OLS) and then testing whether the residuals from that  
11 equation can be treated as stationary or not by using any conventional cointe-  
12 gration test. This test is then repeated for every country in the sample, each  
13 time using only the sample information for that particular country. Studies  
14 based on this approach are generally unable to reject the null hypothesis of no  
15 cointegration, which is seen to imply that the strong form of the sustainability  
16 hypothesis should be rejected. In this paper, we argue that this result should  
17 not be taken at face value as a failure to reject the null may well be due to low  
18 power.

19  
20 In fact, it is well known that tests that take no cointegration as the null  
21 hypothesis can have very little power against nearly cointegrated alternatives,  
22 especially if there are structural breaks present. In a situation like this, it is  
23 therefore essential to devise tests with increased power. A natural approach to  
24 do this would be to combine the sample information obtained from the time  
25 series dimension with that obtained from the cross-sectional. This not only in-  
26 creases the power by taking the total number of observations and their variation  
27 into account, but also increases the precision of the test by effectively reducing  
28 the noise coming from the individual time series regressions. Therefore, one way  
29 to augment the power of univariate tests would be to subject the residuals from  
30 (6) to a panel data cointegration test.

31 Unfortunately, most existing tests of this kind rely critically on assuming  
32 that the countries are independent of each other, which is unlikely to hold  
33 in the present application because of strong intra-economy linkages. Another  
34 limitation of most panel cointegration tests is that they ignore the possibility  
35 of structural change, which is likely to be the rule rather than the exception in  
36 this kind of fiscal data. Thus, a first crucial step in testing the sustainability  
37 hypothesis using panel data is to employ tests that allow for structural breaks,  
38 and that do not rely to such a large extent on the countries being independent.  
39  
40

### 41 2.3 A panel cointegration test with breaks

42 The previous section suggests that testing for cointegration in cross-country  
43 dependent data with structural change is key in inferring the sustainability  
44 hypothesis. In this section, we outline a test that fits this description, and that  
45 is general enough to allow for both cross-country dependence and an unknown  
46 number of breaks that may be located at different dates for different countries.  
47 In so doing, we will use the index  $i = 1, \dots, N$  to denote countries, while  $t$  again  
48 denotes time.  
49

50 Consider the following panel version of the regression in (6)

$$51 R_{it} = \alpha_{ij} + \beta_i \tilde{G}_{it} + e_{it}, \quad j = 1, \dots, M_i + 1, \quad (8)$$

52 where  $\beta_i$  is a country specific slope that is assumed to be constant over time,  
53  
54  
55  
56  
57  
58  
59  
60



while  $\alpha_{ij}$  is a country specific intercept that is subject to  $M_i$  structural breaks. In other words, there are  $M_i + 1$  regimes for each country  $i$  with the  $j^{th}$  regime running from  $T_{ij-1}$  to  $T_{ij}$  time series observations.

Sustainability requires (8) being a cointegrated relationship. If there is no structural change, then this hypothesis can be readily tested by using the existing tests for cointegration in panel data. If there are breaks, however, then this test procedure is no longer valid since the relationship in (8) is now non-linear. This poses a serious problem for inference since conventional tests cannot be used to discriminate between cointegration with structural change and the absence of cointegration. This issue was recently addressed by Westerlund (2006), who develops a panel Lagrange multiplier test for cointegration that allows for multiple structural breaks.

The null hypothesis is formulated as that all the countries in the panel are cointegrated, while the alternative is formulated as that there is at least some country for which cointegration does not hold. The test statistic for this particular hypothesis can be written as

$$Z(M) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M_i+1} \sum_{t=T_{ij-1}+1}^{T_{ij}} S_{it}^2 / (T_{ij} - T_{ij-1})^2 \hat{\sigma}_i^2,$$

where  $S_{it} = \sum_{s=T_{ij-1}+1}^t \hat{e}_{is}$  and  $\hat{e}_{it}$  is the regression residual obtained by using any efficient estimator of the cointegration vector such as the conventional fully modified OLS (FMOLS) estimator. Given that the countries are independent, and some technical conditions regarding the persistency of the individual time series, Westerlund (2006) shows that  $Z(M)$  reaches the following sequential limit as  $T \rightarrow \infty$  and then  $N \rightarrow \infty$  under the null hypothesis

$$\frac{\sqrt{N}(Z(M) - E(Z(M)))}{\sqrt{\text{var}(Z(M))}} \Rightarrow N(0, 1),$$

where the mean and variance adjustment terms  $E(Z(M))$  and  $\text{var}(Z(M))$  are defined in Westerlund (2006). Thus, by standardizing  $Z(M)$  by its mean and standard deviation, we obtain a new test statistic that has an asymptotic standard normal distribution under the null hypothesis. Under the alternative hypothesis, the statistic diverges to positive infinity suggesting that the right tail of the normal distribution should be used to reject the null.

For the estimation of the number of breaks and their locations, Westerlund (2006) suggests using the Bai and Perron (2003) procedure, which can be implemented in two steps. In the first, the breakpoints are estimated by minimizing the sum of squared residuals for all permissible values of  $M_i$ . In the second, the estimated breakpoints are used together with the associated sum of squares to estimate the number of breaks using an information criterion. These steps are then repeated  $N$  times to obtain the estimated number of breaks and their locations for each country.

1  
2  
3  
4  
5  
6  
7  
8 The fact that both the number of breaks and their locations are treated  
9 as unknown here is a clear advantage in comparison to other approaches. For  
10 example, most studies that allow for structural change, see Wilcox (1989) and  
11 Hakkio and Rush (1991), often maintain that the breaks are known, which  
12 means that they are likely to suffer from pre-test bias. Other studies such as  
13 Haug (1995) and Quintos (1995) allow for unknown breakpoints but restrict the  
14 number of breaks to one, which is not likely to be the case.

15 To handle the impact of cross-country dependence Westerlund (2006) sug-  
16 gests using the bootstrap approach. The particular bootstrap scheme opted for  
17 in this paper uses the sieve approach of Psaradakis (2003), who proposes a boot-  
18 strapped stationarity test for pure time series. The advantage with this scheme  
19 is that it can be modified to preserve not only the cross-country correlations  
20 but also the serial correlations.  
21

## 22 **3 Empirical results**

23  
24 In this section, we first briefly describe the data, and then we present the em-  
25 pirical unit root and cointegration test results. Finally, we present some results  
26 of the estimated cointegration vectors.  
27

### 28 **3.1 Data**

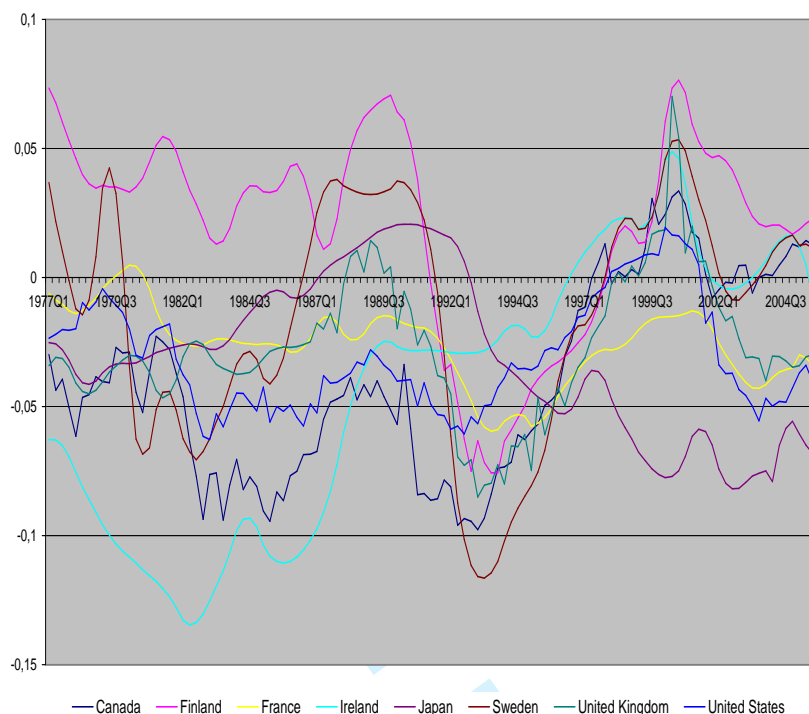
29  
30 The sustainability hypothesis is tested by using quarterly data that covers eight  
31 rich OECD countries over the period 1977Q1 to 2005Q4. The countries included  
32 in the panel are Canada, Finland, France, Ireland, Japan, Sweden, the United  
33 Kingdom, and the United States. The data are taken from the OECD Economic  
34 Outlook Database Inventory, updated in February 2006, and include general  
35 government receipts and disbursement, and GDP at market prices. As in Section  
36 2, the variables used in the analysis were converted into GDP ratios.  
37

### 38 **3.2 Graphical analysis**

39  
40 In order to get a feeling of the behavior of the fiscal debt situation in the OECD  
41 area, we begin with a graphical inspection of the data. To foreshadow the more  
42 formal treatment in the next two sections, Figure 1 plots the deficit-GDP ratios  
43 of the eight countries in the sample. The figure clearly illustrates that while  
44 individually quite dispersed, the series tend to revert back towards a common  
45 mean value, which seem to have become more positive in recent years. Thus, if  
46 sustainability is to be interpreted as the mean reversion, or stationarity, of the  
47 public deficit, it appears that there is some indication of sustainability among  
48 the countries.  
49

50  
51 It is also interesting to consider Figure 1 in view of the major public reforms  
52 that has taken place during the length of the sample. The first thing to notice  
53  
54  
55  
56

Figure 1: Deficit-GDP ratios.



is increased deficit-GDP ratios in the first half of the sample, which originated with public reforms based on tight monetary and loose fiscal policy (OECD, 2002). These deficits made it impossible for governments to compensate the spiraling debt interest payments. Hence, in order to stabilize and restore public finances, and to enhance output growth in the long-run, new fiscal rules was put in place in the United States, and the European Union member countries introduced stricter budget deficit and fiscal debt criteria. These reforms have in recent years made it possible for governments to move away from a strategy of controlling unstable debt dynamics, and towards policy actions designed to strengthen the fiscal stance. This improvement in the overall debt-GDP stance in the latter half of the sample is clearly visible in the figure.

### 3.3 Stationarity tests

We begin by considering the integratedness of the government revenue and expenditure series using the panel stationarity test of Carrion-i-Silvestre *et al.*

(2005), which is very similar to the Westerlund (2006) cointegration test in the sense that it permits for both cross-country dependence and structural breaks. This is therefore an appropriate test for our purposes. Another similarity with the cointegration test is that the breaks are estimated using the Bai and Perron (2003) procedure and that the bootstrap is of sieve type.

In applying the Carrion-i-Silvestre *et al.* (2005) test, we follow the recommendation of Newey and West (1994) and use the Bartlett kernel, where the bandwidth parameter is permitted to grow with  $T$  at the rate  $T^{1/3}$ . In estimating the number of breaks, we allow for a maximum of five shifts for each country, which seem sufficient to capture most of the breaks in the data. The exact number of breaks for each country is estimated using the Schwartz information criterion. Also, to ensure that the break date estimator work properly, we set the minimum length of each regime equal to  $0.1T$ . The sieve bootstrap is implemented using a lag length of five and we make 1,000 replications.

As for the deterministic specification, since all variables are expressed in GDP ratios, no time trends are required. Two models are considered. The first assumes a fixed mean while the second allows for up to five breaks in the mean of each series. Among these models, since most of the literature indicates the presence of breaks, the latter stands out as the most natural one.

Table 1 shows the results of panel stationary test for each variable. For each model, the first row represents the test value, the second row contains the asymptotic  $p$ -values, and the third row contains the bootstrapped  $p$ -values. The results from the asymptotic  $p$ -values suggest that the null hypothesis of stationary can be safely rejected. However, these values assume that the countries are independent, which is unlikely to hold. In order to account for this dependence, we use the bootstrapped  $p$ -values instead. Except for revenues in the structural break model, the conclusions are not altered by taking the cross-country correlations into account. Thus, since the overall evidence in favor of a rejection is quite overwhelming, we choose to proceed as if the variables are in fact non-stationary.

### 3.4 Cointegration tests

The second step in our analysis is to test whether the variables are cointegrated using the  $Z(M)$  test. The results are presented in the rightmost column of Table 1. It is seen that the  $p$ -values for the model with a fixed constant provide very little support of cointegration. However, these results do not take into account the possibility of structural breaks, and are therefore prone to erroneous conclusions. Indeed, if we allow for structural shifts and cross-country correlation, the null hypothesis of cointegration cannot be rejected at the 5 percent level. We therefore conclude that the variables appear to be cointegrated around a broken intercept.

Table 2 reports the estimated breakpoints obtained from the Bai and Perron (2003) procedure. It is seen that there are at least three breaks for each coun-

try, which is indicative of structural instability.<sup>4</sup> This evidence seems consistent with the gradual shift towards increased business cycles synchronization among the OECD countries, which has substantially affected the fiscal position of many governments, see Fatas and Mihov (2003). The first break period in the 1980's reflects the effect of the Great Depression, during which the synchronized recession coincided with an increase in real interest rate. This caused a substantial deterioration of the fiscal stance of governments of rich OECD countries, as the deficits created by the spiralling interest payments could not be offset, see OECD (2002).

Later, following the Maastricht Treaty and the Stability and Growth Pact in the early 1990's, the fiscal stance of many OECD countries were driven primarily by an increased coordination of both fiscal and monetary policy within the European Monetary Union, see Darvas *et al.* (2005). This period is also clearly visible in the table, as indicated by the concentration of breaks in the early 1990's, and seems consistent with the results reported in many earlier studies, see for example Afonso (2005). The many breaks in the late 1990's are most likely due to the formation of the European Monetary Union in January 1999.

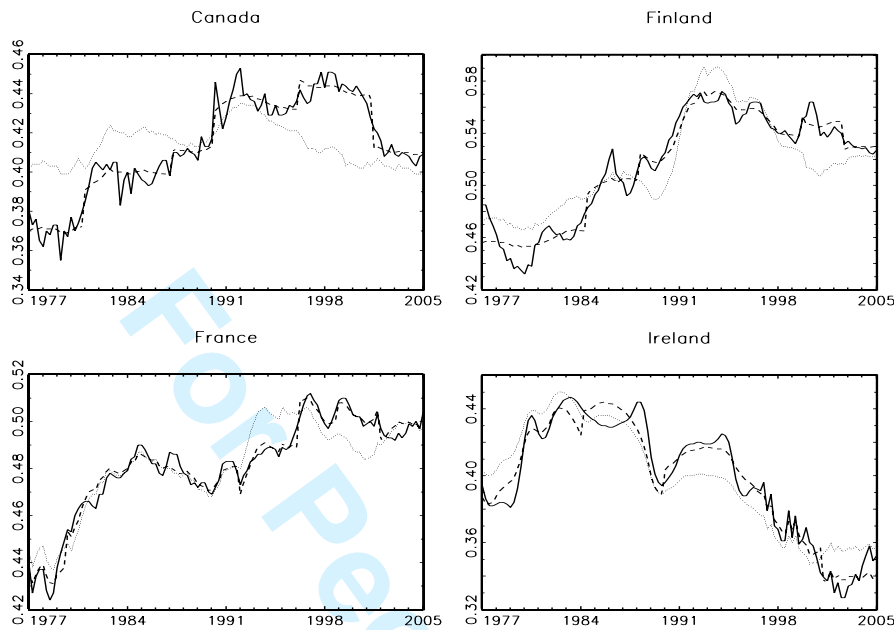
To further illustrate the importance of accounting for structural change, Figures 2 and 3 plot the revenues-GDP ratio, the dependent variable, together with the fitted trend functions for the models with and without breaks. In the figures, the solid line represents revenues measured as a fraction of GDP, the dashed line represents the model with breaks and the dotted line represents the model without any breaks. It can easily be seen that nearly all revenue series exhibit several level breaks, most of which seem to reflect the effect of economic fluctuations, and changes in the monetary and fiscal policy objectives in the individual countries. Interestingly, it is seen that the revenues-GDP ratio of most countries have been rising during the first half of the sample and falling during the second half, which reflects a more flexible response of the recent fiscal policy of most OECD governments to the economic downturn.

More importantly, we see that the trend lines for the model with breaks seem to provide a very good fit to the revenues for all eight countries. By contrast, the trend lines for the model without breaks are generally incapable of accounting for the all variation in revenues, thus leading to a poor fit. This effect is particularly striking for the United States.

---

<sup>4</sup>Note that for five of the eight countries we end up estimating the maximum number of breaks. This is in agreement with the results of Ahamada *et al.* (2004) and Perron (1997), who point out that the type of information criteria used here cannot directly take into account the effect of different distributions of the data and possible serial correlation in the regression errors. In fact, as Perron (1997) showed in simulations, although most criteria perform reasonably well when the errors are not correlated, they have a strong tendency to overestimate the number of breaks when serial correlation is present. Of course, the purpose in this paper is not the correct estimation of the number of breaks *per se* but rather the accounting of the effects of these breaks. Thus, overestimation is not a serious problem in the sense that the model is still free to estimate the associated break parameters to zero.

Figure 2: Revenues-GDP ratio with fitted trend functions.



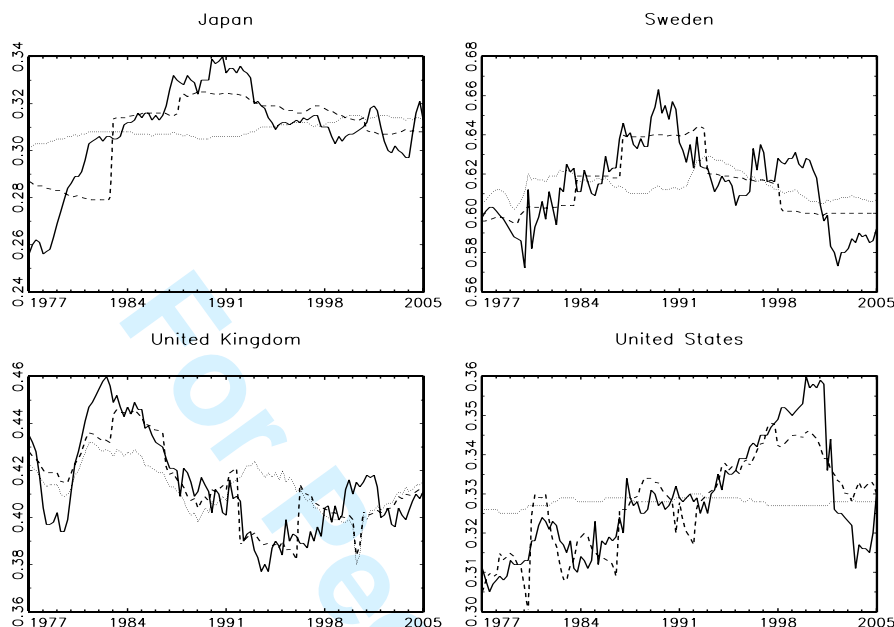
### 3.5 Cointegration estimation

Since the variables appear to be cointegrated, it is possible to estimate and test whether the cointegrating slope is indeed equal to one as required by sustainability hypothesis.

It is well known that the OLS estimator is consistent under fairly general conditions when applied to the cointegrated regression in (6). Unfortunately, if the regressor is endogenous, then this estimator suffers from nuisance parameter dependencies even asymptotically, which makes it a poor candidate for inference. To account for this, in addition to using OLS, we employ both FMOLS and dynamic OLS (DOLS) panel estimation techniques, see Kao and Chiang (2000). All three estimators are based on pooling along the cross-sectional dimension, and are appropriate for testing the null hypothesis that  $\beta_i = 1$  for all  $i$  against the alternative hypothesis that  $\beta_i$  is equal to some common slope value, which is different from unity.

Moreover, as when testing for stationarity and cointegration in panel data, the presence of cross-country dependence makes inference based on the asymptotic normal distribution inappropriate, in which case bootstrap inference might be better. The particular bootstrap opted for this section is taken from Chang *et al.* (2006), who propose a sieve resampling scheme that preserves the serial correlation properties of the errors, and that can be generalized along the lines

Figure 3: Revenues-GDP ratio with fitted trend functions.



of Westerlund (2006) to also accommodate cross-country dependence. As an indication of the severity of this problem, we applied the Pesaran (2004) test of cross-section correlation to the OLS residuals of the estimated cointegrated regression.<sup>5</sup> The computed test value was 4.98, which, when compared to the right tail of the asymptotic normal distribution, lead to a clear rejection of the no correlation null at all conventional significance levels. Thus, cross-country correlation is indeed a problem that needs to be addressed.

The results from the OLS, FMOLS and DOLS estimators are reported in Table 3 together with the double-sided  $p$ -values for the null hypothesis of a unit slope. The first thing to notice is that the estimated slopes lie very close to their hypothesized values of one. Indeed, the range of the estimated slopes is 0.909 to 0.934. The closeness of these estimates to their expected value based on the sustainability hypothesis is supported by the  $p$ -values.

Indeed, by looking at the  $p$ -values based on the asymptotic normal distribution, we see that the null hypothesis of a unit slope cannot be rejected even at the most generous 10 percent level. The results from the bootstrapped  $p$ -values, which allow for cross-country dependence, are even more supportive of the unit slope null, and are all well above 0.6. Thus, based on this evidence we cannot

<sup>5</sup>The Pesaran (2004) test is basically the average pair-wise correlation coefficient of the regression residuals, which has an asymptotic normal distribution under the null hypothesis of no cross-correlation.



1  
2  
3  
4  
5  
6  
7  
8 reject the null hypothesis of a unit slope in the panel. Thus, in contrast to much  
9 of the earlier evidence regarding the sustainability of fiscal policy, we are unable  
10 to reject the sustainability hypothesis for the panel as a whole.  
11

## 12 4 Concluding remarks

13  
14  
15 Earlier empirical research based on testing for stationarity of public debt and  
16 budget deficit has tended to reject the hypothesis of fiscal sustainability. As a  
17 response to this, recent work within this field has turned towards more flexible  
18 testing approaches based on cointegration. The idea is that if fiscal policy is  
19 sustainable, then government revenues and expenditures should be cointegrated  
20 with a unit slope on expenditures. Unfortunately, the results of the earlier  
21 studies have not been very convincing.  
22

23 In this paper, we provide some insights suggesting that the weak empirical  
24 support of the sustainability hypothesis may be due to inadequate econometric  
25 methods. In particular, we argue that a failure of accounting for structural  
26 breaks and the poor precision of commonly applied time series tests may well  
27 result in a rejection of the sustainability hypothesis. To circumvent these prob-  
28 lems, we resort to some recent advances in the area of non-stationary panel  
29 data. These methods are more powerful than conventional time series methods,  
30 and are general enough to allow for cross-country dependence and an unknown  
31 number of structural breaks in the cointegrating relationship.  
32

33 Based on data covering eight rich OECD countries over the period from  
34 1977Q1 to 2005Q4, we show that revenues and expenditures are non-stationary  
35 and cointegrated. We also provide evidence suggesting that the null hypothesis  
36 of a common unit slope on expenditures cannot be rejected, suggesting that  
37 the sustainability hypothesis cannot be rejected. The results on the estimated  
38 structural breaks are suggestive of strong structural instability, and are com-  
39 patible with the findings of many earlier studies such as Quintos (1995), Martin  
40 (2000) and Afonso (2005). Some of the estimated breaks are related to the in-  
41 creased business cycle synchronization among the OECD countries, while others  
42 are related to the policy interventions resulting from the increased coordination  
43 of the monetary and fiscal policy within the European community.  
44

45 Thus, although there is a growing concern about increased budget deficits  
46 within the OECD area, our results indicate that after taking into account struc-  
47 tural change and economic growth, the budget of the eight rich countries con-  
48 sidered in this study has in fact been balanced, at least intertemporally over the  
49 sample considered. In other words, for these countries the sequence of primary  
50 surpluses has been sufficient to cover the marked value of public debt. Thus,  
51 from an historical point of view, there is no reason to be concerned about the  
52 public debt situation.  
53

54 This finding is in agreement with the results reported by Quintos (1995) for  
55 the United States, and is particularly interesting in view of the recent discus-  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6  
7  
8 sion regarding the effect of the projected increase in future public liabilities to  
9 age-related public spending and health care on public finances during the de-  
10 mographic transition of OECD countries to older societies. Our results imply  
11 that if governments are able to raise revenues at the same rate as the expected  
12 increase in spending, so that revenues and expenditures are one-to-one in the  
13 long run, then one need not worry about the future public debt situation.  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

For Peer Review

Table 1: Panel stationarity and cointegration tests.

Specification	Test	Stationarity tests		
		$R_{it}$	$\tilde{G}_{it}$	$Z(M)$
Fixed constant	Value	20.147	18.684	16.661
	$p$ -value <sup>a</sup>	0.000	0.000	0.000
	$p$ -value <sup>b</sup>	0.000	0.002	0.028
Constant with breaks	Value	2.192	5.114	1.416
	$p$ -value <sup>a</sup>	0.014	0.000	0.078
	$p$ -value <sup>b</sup>	0.242	0.008	0.678

*Notes:* All test results allow for up to five structural breaks for each state. To handle the serial correlation, the Bartlett kernel is used with a bandwidth parameter of  $T^{1/3}$ .

<sup>a</sup>The  $p$ -value is based on the asymptotic normal distribution.

<sup>b</sup>The  $p$ -value is based on the bootstrapped distribution. The number of lags in the sieve approximation is five and we make 1,000 bootstrap replications.

Table 2: Estimated breaks.

Country	No.	Breakpoint				
		1	2	3	4	5
Canada	5	1981Q1	1987Q2	1990Q2	1996Q2	2001Q3
Finland	4	1984Q3	1988Q2	1999Q4	2002Q4	–
France	5	1979Q4	1992Q1	1996Q2	1999Q1	2002Q1
Ireland	3	1983Q1	1987Q4	2001Q1	–	–
Japan	5	1979Q4	1983Q4	1987Q1	1993Q1	1998Q2
Sweden	5	1983Q1	1986Q4	1992Q1	1996Q2	2001Q2
United Kingdom	4	1980Q3	1986Q4	1992Q2	1998Q1	–
United States	5	1980Q3	1987Q2	1994Q2	1997Q3	2002Q1

*Notes:* The breaks are estimated using the Bai and Perron (2003) procedure with a maximum number of five breaks for each state. The minimum length of each break regime is set to  $0.1T$ .

Table 3: Panel cointegration slope estimates.

Estimator	Value	$p$ -value <sup>a</sup>	$p$ -value <sup>b</sup>
OLS	0.925	0.594	0.708
DOLS	0.934	0.689	0.718
FMOLS	0.909	0.444	0.602

*Notes:* All results allow for up to five structural breaks for each state. The  $p$ -values for the between and within estimates are for the null of a unit slope against the alternative that the slope is different from unity. <sup>a</sup>The  $p$ -value is based on the asymptotic normal distribution. <sup>b</sup>The  $p$ -value is based on the bootstrapped distribution. The number of lags in the sieve approximation is five and we make 1,000 bootstrap replications.

## References

- Afonso, A. (2005). Fiscal Sustainability: the unpleasant European case. *FinanzArchiv* 61, 19-44.
- Ahamada, I., Jouini, J., and Boutahar, M. (2004). Detecting multiple breaks in time series covariance structure: a non-parametric approach based on the evolutionary spectral density. *Applied Economics* 36, 1095-1101.
- Ahmed, S., and Rogers, J. (1995). Government budget deficits and trade deficits. Are present value constraint satisfied in long-term data? *Journal of Monetary Economics* 36, 351-374.
- Baglioni, A., and Cherubini, U. (1993). Intertemporal budget constraint and public debt sustainability: the case of Italy. *Applied Economics* 25, 275-283.
- Bai, J., and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18, 1-22.
- Bergman, M. (2001). Testing government solvency and the No Ponzi Game condition. *Applied Economics Letters* 8, 27-29.
- Bohn, H. (1991). The sustainability of budget deficit with lump-sum and with income-based taxation. *Journal of Money, Credit and Banking* 23, 581-604.
- Bordo, M.D., and Heibling, T. (2003). Have national business cycles become more synchronized? NBER Working Paper No. 10130.
- Breitung, J., and Pesaran, M. H. (2005). Unit roots and cointegration in panels. In Matyas, L., and Sevestre, P. (Eds.), *The Econometrics of Panel Data*. Klüver Academic Publishers.
- Carrion-i-Silvestre, J. L., Del Barrio-Castro, T., and López-Bazo, E. (2005). Breaking the panels: an application to the GDP per capita. *Econometrics Journal* 8, 159-175.
- Chang, Y., Park, J. Y., and Song, K. (2006). Bootstrapping cointegrating regressions. *Journal of Econometrics* 133, 703-739.
- Darvas, Z., Rose, A. K., and Szapary, G. (2005). Fiscal divergence and business cycle synchronization: irresponsibility is idiosyncratic. NBER Working Paper No. 11580.
- European Commission. (2002). Public finances in EMU. *European Economy* 3.

- 1  
2  
3  
4  
5  
6  
7  
8 Fatas, A., and Mihov, I. (2003). The case of restricting fiscal policy discretion.  
9 *Quarterly Journal of Economics* 118-4, 1419-1447.
- 10  
11 Hakkio, G., and Rush, M. (1991). Is the budget deficit too large? *Economic*  
12 *Inquiry* 29, 429-445.
- 13  
14 Haug, A. (1995). Has federal budget deficit policy changed in recent years?  
15 *Economic Inquiry* 33, 104-118.
- 16  
17 IMF (2001). Business cycle linkages among major advanced economies. *World*  
18 *Economic Outlook*, October 2001.
- 19  
20 Kao, C., and Chiang, M. H. (2000). On the estimation and inference of a  
21 cointegrated regression in panel data. *Advances in Econometrics* 15, 179-  
22 222.
- 23  
24 Kirchgassner, G., and Prohl, S. (2006). Sustainability of Swiss fiscal policy.  
25 Forthcoming in *Swiss Journal of Economics and Statistics*.
- 26  
27 MacDonald, R. (1992). Some tests of the government's intertemporal budget  
28 constraint using U.S. data. *Applied Economics* 24, 1287-1292.
- 29  
30 Martin, G. (2000). U.S. deficit sustainability: a new approach based on mul-  
31 tiple endogenous breaks. *Journal of Applied Econometrics* 15, 83-105.
- 32  
33 Newey, W., and West, K. (1994). Autocovariance lag selection in covariance  
34 matrix estimation. *Review of Economic Studies* 61, 631-653.
- 35  
36 OECD (2002). Fiscal sustainability: the contribution of fiscal rules. *OECD*  
37 *Economic Outlook* 72.
- 38  
39 Olekalns, N. (2000). Sustainability and stability? Australian fiscal policy in  
40 the 20th century. *Australian Economic Papers* 39, 138-151.
- 41  
42 Papadopoulos, A., and Sidiropoulos, M. (1999). The sustainability of fiscal  
43 policies in the European Union. *International Advances in Economic Re-*  
44 *search* 5, 289-307.
- 45  
46 Payne, J. (1997). International evidence on the sustainability of budget deficits.  
47 *Applied Economic Letters* 12, 775-779.
- 48  
49 Perron, P. (1997). L'estimation de modèles avec changements structurels mul-  
50 tiples. *Actualit conomique* 73, 457-505.
- 51  
52 Pesaran, H. (2004). General diagnostic tests for cross section dependence in  
53 panels. Cambridge Working Paper in Economics 435.
- 54  
55 Psaradakis, Z. (2003). A sieve bootstrap test for stationary. *Statistics and*  
56 *Probability Letters* 62, 263-274.

- 1  
2  
3  
4  
5  
6  
7  
8 Quintos, C. (1995). Sustainability of the deficit process with structural shifts.  
9 *Journal of Business and Economic Statistics* 13, 409-417.  
10  
11 Smith, G., and Zin, S. (1991). Persistent deficits and the market value of  
12 government debt. *Journal of Applied Econometrics* 6, 31-44.  
13  
14 Westerlund, J. (2006). Testing for panel cointegration with multiple structural  
15 breaks. *Oxford Bulletin of Economics and Statistics* 68, 101-132.  
16  
17 Wilcox, D. (1989). The sustainability of government deficits: implications of  
18 the present-value borrowing constraint. *Journal of Money, Credit, and*  
19 *Banking* 21, 291-306.  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60



# Panel Cointegration Tests of the Sustainability Hypothesis in Rich OECD Countries\*

Joakim Westerlund<sup>†</sup> and Silika Prohl<sup>‡</sup>

October 7, 2006

## Abstract

This study reexamines the sustainability hypothesis by testing whether government revenues and expenditures for eight rich OECD countries between 1977Q1 and 2005Q4 are cointegrated. For this purpose, a non-stationary panel data approach is adopted, which is general enough to permit for cross-country dependence as well as structural breaks representing major shifts in fiscal policy. In contrast to many earlier studies, the results reported in this study suggest that the sustainability hypothesis cannot be rejected.

**JEL Classification:** C22; C23; H60.

**Keywords:** Fiscal Sustainability; Panel Cointegration; Structural Change.

## 1 Introduction

The increased budget deficit and public debt experienced by many OECD countries raises the issue of sustainability of the government finance in the

---

\*Westerlund gratefully acknowledges financial support from the Jan Wallander and Tom Hedelius Foundation, research grant number W2006-0068:1.

<sup>†</sup>Corresponding author: Department of Economics, Lund University, P. O. Box 7082, S-220 07 Lund, Sweden. E-mail address: joakim.westerlund@nek.lu.se.

<sup>‡</sup>SIAW - University of St. Gallen, Bodanstraße 8, CH-9000 St. Gallen, Switzerland. E-mail address: silika.prohl@unisg.ch.

1  
2  
3  
4  
5  
6 long run, and its potential effects on the economy as a whole. Fiscal policy  
7 is constrained by the need to finance the government deficit, which over the  
8 longer term implies that the market value of public debt must be offset by  
9 the present value of all discounted future budget surpluses. In other words,  
10 the fiscal policy is sustainable if the discounted value of debt is zero in the  
11 limit. This is the essence of the sustainability hypothesis. If the hypothesis  
12 holds, then the ongoing fiscal policy can, at least in principle, be maintained  
13 indefinitely, whereas if it fails, then there is a need for discretionary policy  
14 actions. In particular, a failure implies that the government will inevitably  
15 run into problems in managing its debt, thus requiring compensation in terms  
16 of higher interest rates leading to a slowdown in the economic growth.  
17  
18

19 Thus, from a policy point of view, the issue of whether the sustainabil-  
20 ity hypothesis holds or not is indeed an important one. Because of this, and  
21 because of the growing concern regarding the debt situation in many OECD  
22 countries, sustainability analysis has received much interest in recent empirical  
23 research. Most of the earlier work in this field focused on the stationarity of  
24 public deficit and debt as a way to empirically test the sustainability hypoth-  
25 esis. Prominent examples include, among others, Wilcox (1989), MacDonald  
26 (1992), Hakkio and Rush (1991) and Quintos (1995) for the United States,  
27 Smith and Zin (1991) for Canada, Baglioni and Cherubini (1993) for Italy,  
28 Olekalns (2000) for Australia, and perhaps most recently Kirchgaessner and  
29 Prohl (2006) for Switzerland. The overall picture being that the sustainability  
30 hypothesis does not hold.  
31  
32

33 As a response to this, more recent work has moved away from examining  
34 the stationarity of debt, and towards more flexible testing strategies based  
35 on cointegration. In particular, as shown by Quintos (1995), given that pub-  
36 lic revenues and expenditures are non-stationary, sustainability requires these  
37 variables to be cointegrated with a unit slope on expenditures. In the ter-  
38 minology of Quintos (1995), this is known as strong sustainability. It means  
39 that no problems, according to the ongoing fiscal policy, are likely to arise in  
40 the future. By contrast, weak sustainability refers to the case when the slope  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6 lies between zero and one, regardless of whether revenues and expenditures  
7 are cointegrated or not. In this case, although the deficit is sustainable, the  
8 government is expected to have difficulty in marketing its debt, thus requiring  
9 fiscal reforms. Finally, if the slope on expenditures is equal to zero, then the  
10 deficit is not sustainable, in which case the sustainability hypothesis is refuted.  
11  
12

13  
14 Payne (1997) applies this approach to the G7 countries between 1950 and  
15 1990, and rejects the sustainability hypothesis. Similarly, using the same ap-  
16 proach to a sample of five EU countries between 1961 and 1995, Papadopou-  
17 los and Sidiropoulos (1999) only find evidence of cointegration for Greece and  
18 Spain. However, the cointegrating slope on government expenditures is esti-  
19 mated significantly lower than one, suggesting that only the weak form of the  
20 sustainability hypothesis is supported. Apparently, despite the strong theo-  
21 retical appeal of the sustainability hypothesis, most empirical studies tend to  
22 reject it.  
23  
24

25  
26 Although there are many explanations for why the sustainability hypoth-  
27 esis does not hold, this paper focuses on two that might potentially go a long  
28 way towards explaining the weak empirical evidence on the cointegration be-  
29 tween revenues and expenditures. The first explanation is that conventional  
30 time series cointegration tests may have low power against persistent alterna-  
31 tives because of the short sample periods usually employed, see Afonso (2005).  
32 The second explanation is that the empirical relationship between public rev-  
33 enues and expenditures that is being estimated need not be invariant to pol-  
34 icy regime changes, which may lead to structural shifts in both revenues and  
35 expenditures, and hence in the relationship between them, see for example  
36 Quintos (1995) and Martin (2000).  
37  
38

39  
40 However, while very reasonable and potentially appealing, when analyzed  
41 separately, these explanations have been found to be empirically inadequate  
42 and far from convincing. This paper therefore offers an alternative route. The  
43 idea is that, to be able to provide any robust evidence on the cointegration  
44 between revenues and expenditures, one needs to consider not the low power  
45 and the presence of structural change separately but simultaneously. The  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

intuition is simple. First, while increasing the length of the sample may be justified from a power point of view, it also increases the probability of a break. On the other hand, modifying existing tests so as to accommodate for structural change is typically very costly in terms of power. Therefore, it is the joint consideration of both these aspects that it is likely to be key approach when testing the sustainability hypothesis.

One way to accomplish this is to resort to panel data, and the recent econometrical advances within this field that make it possible to construct powerful cointegration methods while simultaneously entertaining the possibility of structural change. Applying this approach to panel of eight rich OECD countries between 1977Q1 and 2005Q4, in contrast to much of the earlier evidence, we are unable to reject the strong sustainability hypothesis. We also provide evidence suggesting that all countries have been subject to several structural breaks representing major changes in fiscal policy.

The remainder of this paper is organized as follows. In Section 2, we give the brief description of the theoretical model of fiscal sustainability and the empirical panel data framework that will be used to test it. Section 3 then reports the empirical results, while Section 4 concludes.

## 2 Testing the sustainability hypothesis

In this section, we begin with a brief account of the economic theory underlying the sustainability hypothesis, and then we go on to discuss the empirical method that we will use to test it.

### 2.1 Theoretical underpinnings

The theoretical model of sustainability of the fiscal policy starts with the budget constraint of the government at time  $t$ , which is given in nominal terms by

$$g_t + (1 + i_t)b_{t-1} = r_t + b_t, \quad (1)$$

where  $b_t$  is the stock of public debt,  $i_t$  is the nominal interest rate payable on the public debt,  $r_t$  is the government revenue including revenue from seigniorage, and  $g_t$  is the government expenditure excluding interest payments. Of course, these variables are relatively uninteresting as they do not take into account the size of the economy. We therefore follow the conventional practice of rewriting (1) in terms of GDP ratios as

$$\frac{g_t}{y_t} + \frac{(1 + i_t) b_{t-1}}{(1 + \sigma_t) y_{t-1}} = \frac{r_t}{y_t} + \frac{b_t}{y_t},$$

where  $y_t$  is the nominal GDP, while  $\sigma_t$  represents the nominal GDP growth rate. Using capital letters to denote the ratios of the corresponding upper-case variables to nominal GDP, we get

$$C_t + (1 + i_t^*) B_{t-1} = R_t + B_t, \quad (2)$$

where  $C_t = G_t + (i_t^* - i^*) B_{t-1}$  and  $i^*$  is the mean of  $i_t^* = (i_t - \sigma_t)/(1 + \sigma_t)$ , the growth adjusted interest rate. If we assume that the budget constraint in (2) holds continuously at each  $t$ , we can use forward substitution to obtain

$$B_{t-1} = \sum_{j=t+1}^{\infty} \left( \frac{1}{1 + i^*} \right)^{j-t} (R_{j-1} - C_{j-1}) + \lim_{j \rightarrow \infty} \left( \frac{1}{1 + i^*} \right)^{j-t} B_{j-1}. \quad (3)$$

Now, consider taking expectations conditional on all the information available at time  $t$ . Since  $B_{t-1}$  is known at  $t$ , the hypothesis that the government obeys its intertemporal budget constraint can be expressed as

$$\lim_{j \rightarrow \infty} E_t \left( \frac{1}{1 + i^*} \right)^{j-t} B_{j-1} = 0. \quad (4)$$

This transversality condition, also known as the no Ponzi game rule, says that the public debt-GDP ratio should grow no faster on average than the mean rate of interest, meaning that the principal repayments and interests cannot be financed indefinitely by issuing new debt. It then follows that the current stock of public debt is offset by all current and expected discounted future primary surpluses.<sup>1</sup>

<sup>1</sup>Note that  $1/(1 + i^*) = (i + \sigma)/(1 + i)$ , where  $i$  and  $\sigma$  are the mean values of  $i_t$  and  $\sigma_t$ , respectively. Thus, if  $i < \sigma$ , then the transversality condition is satisfied automatically.

## 2.2 A time series cointegration test

The empirical literature has developed several tests to investigate whether the transversality condition (3) holds. The particular test opted for in this paper is that of Quintos (1995), who suggests testing the sustainability hypothesis by examining whether government revenues and expenditures are cointegrated or not. The intuition behind this test lies in first taking differences, and then rewriting (2) in terms of  $\tilde{G}_t = G_t + i_t^* B_{t-1}$ , the total expenditure including interest payments, as

$$\begin{aligned} \tilde{G}_{t-1} - R_{t-1} &= \sum_{j=t+1}^{\infty} \left( \frac{1}{1+i^*} \right)^{j-t} (\Delta R_{j-1} - \Delta C_{j-1}) \\ &+ \lim_{j \rightarrow \infty} \left( \frac{1}{1+i^*} \right)^{j-t} \Delta B_{j-1}. \end{aligned} \quad (5)$$

As before, the sustainability hypothesis holds if the expected value of the second term on the right-hand side is zero. One way to test whether this is in fact the case, given that  $\tilde{G}_t$  and  $R_t$  possess unit roots, involves first estimating and then testing for cointegration in the following regression

$$R_t = \alpha + \beta \tilde{G}_t + e_t, \quad (6)$$

where  $e_t$  is a mean zero disturbance term. To see what this means for the test of the sustainability hypothesis, note that  $\Delta B_t = \tilde{G}_t - R_t$ , which together with (5) implies that

$$\Delta B_t = (1 - \beta) \tilde{G}_t - \alpha - e_t. \quad (7)$$

As shown by Quintos (1995), although the sustainability hypothesis holds regardless of the integratedness of  $\Delta B_t$ , the speed with which this happens does not. If  $\Delta B_t$  is stationary, then the second term in (5) vanishes at a much faster rate than if  $\Delta B_t$  is non-stationary. In the terminology of Quintos (1995), the sustainability of the public deficit is said to be strong if  $\Delta B_t$  is stationary, whereas it is said to be weak if  $\Delta B_t$  is non-stationary.

In terms of the regression in (6), by looking at (7), we see that cointegration with  $\beta = 1$  is enough for strong sustainability, because only then will  $\Delta B_t$  be

1  
2  
3  
4  
5  
6 stationary. But since  $\tilde{G}_t$  is non-stationary, the sustainability is weak if either  
7  $\beta = 1$  and there is no cointegration or  $0 < \beta < 1$ . It follows that  $0 < \beta \leq 1$   
8 is both necessary and sufficient for sustainability, while cointegration is only  
9 sufficient. The distinction between strong and weak sustainability is made  
10 because although  $0 < \beta < 1$  is enough for the sustainability hypothesis to  
11 hold, the government is now spending more than it earns, thus making the  
12 public debt difficult to market in the long run. Finally, if  $\beta \leq 0$ , then the  
13 deficit is no longer sustainable, because this would make  $\Delta B_t$  grow faster  
14 than the growth rate of the economy, as approximated by the mean interest  
15 rate.<sup>2</sup>  
16  
17  
18  
19  
20  
21

22 The above discussion suggests that testing for cointegration is key in this  
23 kind of sustainability analysis, not only because of its role in determining  
24 strong sustainability but also because we cannot do inference regarding  $\beta$   
25 unless revenues and expenditures are cointegrated.  
26  
27

28 The conventional way in which earlier studies have been trying to test the  
29 sustainability hypothesis within this framework involves first estimating (6) by  
30 ordinary least squares (OLS) and then testing whether the residuals from that  
31 equation can be treated as stationary or not by using any conventional cointe-  
32 gration test. This test is then repeated for every country in the sample, each  
33 time using only the sample information for that particular country. Studies  
34 based on this approach are generally unable to reject the null hypothesis of no  
35 cointegration, which is seen to imply that the strong form of the sustainability  
36 hypothesis should be rejected. In this paper, we argue that this result should  
37 not be taken at face value as a failure to reject the null may well be due to  
38 low power.  
39  
40  
41  
42  
43  
44  
45

46 In fact, it is well known that tests that take no cointegration as the null  
47 hypothesis can have very little power against nearly cointegrated alternatives,  
48 especially if there are structural breaks present. In a situation like this, it is  
49 therefore essential to device tests with increased power. A natural approach  
50  
51  
52

---

53 <sup>2</sup>Note that the case with  $\beta > 0$  is not consistent with a deficit, since revenues are growing  
54 faster than expenditures.  
55  
56  
57  
58  
59  
60



to do this would be to combine the sample information obtained from the time series dimension with that obtained from the cross-sectional. This not only increases the power by taking the total number of observations and their variation into account, but also increases the precision of the test by effectively reducing the noise coming from the individual time series regressions. Therefore, one way to augment the power of univariate tests would be to subject the residuals from (6) to a panel data cointegration test.

Unfortunately, most existing tests of this kind rely critically on assuming that the countries are independent of each other, which is unlikely to hold in the present application because of strong intra-economy linkages. Another limitation of most panel cointegration tests is that they ignore the possibility of structural change, which is likely to be the rule rather than the exception in this kind of fiscal data. Thus, a first crucial step in testing the sustainability hypothesis using panel data is to employ tests that allow for structural breaks, and that do not rely to such a large extent on the countries being independent.

### 2.3 A panel cointegration test with breaks

The previous section suggests that testing for cointegration in cross-country dependent data with structural change is key in inferring the sustainability hypothesis. In this section, we outline a test that fits this description, and that is general enough to allow for both cross-country dependence and an unknown number of breaks that may be located at different dates for different countries. In so doing, we will use the index  $i = 1, \dots, N$  to denote countries, while  $t$  again denotes time.

Consider the following panel version of the regression in (6)

$$R_{it} = \alpha_{ij} + \beta_i \tilde{G}_{it} + e_{it}, \quad j = 1, \dots, M_i + 1, \quad (8)$$

where  $\beta_i$  is a country specific slope that is assumed to be constant over time, while  $\alpha_{ij}$  is a country specific intercept that is subject to  $M_i$  structural breaks. In other words, there are  $M_i + 1$  regimes for each country  $i$  with the  $j^{th}$  regime running from  $T_{ij-1}$  to  $T_{ij}$  time series observations.

Sustainability requires (8) being a cointegrated relationship. If there is no structural change, then this hypothesis can be readily tested by using the existing tests for cointegration in panel data. If there are breaks, however, then this test procedure is no longer valid since the relationship in (8) is now non-linear. This poses a serious problem for inference since conventional tests cannot be used to discriminate between cointegration with structural change and the absence of cointegration. This issue was recently addressed by Westerlund (2005), who develops a panel Lagrange multiplier test for cointegration that allows for multiple structural breaks.

The null hypothesis is formulated as that all the countries in the panel are cointegrated, while the alternative is formulated as that there is at least some country for which cointegration does not hold. The test statistic for this particular hypothesis can be written as

$$Z(M) = \frac{1}{N} \sum_{i=1}^N \sum_{j=1}^{M_i+1} \sum_{t=T_{i,j-1}+1}^{T_{ij}} S_{it}^2 / (T_{ij} - T_{i,j-1})^2 \hat{\sigma}_i^2,$$

where  $S_{it} = \sum_{s=T_{i,j-1}+1}^t \hat{e}_{is}$  and  $\hat{e}_{it}$  is the regression residual obtained by using any efficient estimator of the cointegration vector such as the conventional fully modified OLS (FMOLS) estimator. Given that the countries are independent, and some technical conditions regarding the persistency of the individual time series, Westerlund (2005) shows that  $Z(M)$  reaches the following sequential limit as  $T \rightarrow \infty$  and then  $N \rightarrow \infty$  under the null hypothesis

$$\frac{\sqrt{N}(Z(M) - E(Z(M)))}{\sqrt{\text{var}(Z(M))}} \Rightarrow N(0, 1),$$

where the mean and variance adjustment terms  $E(Z(M))$  and  $\text{var}(Z(M))$  are defined in Westerlund (2005). Thus, by standardizing  $Z(M)$  by its mean and standard deviation, we obtain a new test statistic that has an asymptotic standard normal distribution under the null hypothesis. Under the alternative hypothesis, the statistic diverges to positive infinity suggesting that the right tail of the normal distribution should be used to reject the null.

For the estimation of the number of breaks and their locations, Westerlund (2005) suggest using the Bai and Perron (2003) procedure, which can

1  
2  
3  
4  
5  
6 be implemented in two steps. In the first, the breakpoints are estimated by  
7 minimizing the sum of squared residuals for all permissible values of  $M_i$ . In  
8 the second, the estimated breakpoints are used together with the associated  
9 sum of squares to estimate the number of breaks using an information crite-  
10 rion. These steps are then repeated  $N$  times to obtain the estimated number  
11 of breaks and their locations for each country.  
12  
13  
14

15  
16 The fact that both, the number of breaks and their locations are treated  
17 as unknown here is a clear advantage in comparison to other approaches. For  
18 example, most studies that allow for structural change, for example Wilcox  
19 (1989) and Hakkio and Rush (1991), often maintain that the breaks are known,  
20 which means that they are likely to suffer from pre-test bias. Other studies  
21 such as Haug (1995) and Quintos (1995) allow for unknown breakpoints but  
22 restrict the number of breaks to one, which is not likely to be the case.  
23  
24  
25  
26

27 To handle the impact of cross-country dependence Westerlund (2005) sug-  
28 gests using the bootstrap approach. The particular bootstrap scheme opted  
29 for in this paper uses the sieve approach of Psaradakis (2003), who proposes  
30 a bootstrapped stationarity test for pure time series. The advantage with  
31 this scheme is that it can be modified to preserve not only the cross-country  
32 correlations but also the serial correlations.  
33  
34  
35  
36  
37

### 38 **3 Empirical results**

39  
40 In this section, we first briefly describe the data, and then we present the  
41 empirical unit root and cointegration test results. Finally, we present some  
42 results on the estimated cointegration vectors.  
43  
44  
45  
46

#### 47 **3.1 Data**

48  
49 We test the sustainability hypothesis by using the quarterly data that covers  
50 eight rich OECD countries over the period 1977Q1 to 2005Q4. The countries  
51 included in the panel are Canada, Finland, France, Ireland, Japan, Sweden,  
52 the United Kingdom, and the United States. The data are taken from the  
53  
54  
55  
56  
57

1  
2  
3  
4  
5  
6 OECD Economic Outlook Database Inventory, updated in February 2006, and  
7 include general government receipts and disbursement, and GDP at market  
8 prices. As in Section 2, the variables used in the analysis are defined in terms  
9 of GDP ratios.  
10  
11

### 12 3.2 Stationarity tests

13  
14  
15  
16 We begin by considering the integratedness of the government revenue and  
17 expenditure series using the panel stationarity test of Carrion-i-Silvestre *et al.*  
18 (2005), which is very similar to the Westerlund (2005) cointegration test in the  
19 sense that it permits for both cross-country dependence and structural breaks.  
20 This is therefore an appropriate test for our purposes. Another similarity with  
21 the cointegration test is that the breaks are estimated using the Bai and Perron  
22 (2003) procedure and that the bootstrap of the sieve type.  
23  
24  
25  
26

27 In applying the Carrion-i-Silvestre *et al.* (2005) test, we follow the recom-  
28 mendation of Newey and West (1994) and use the Bartlett kernel, where the  
29 bandwidth parameter is permitted to grow with  $T$  at the rate  $T^{1/3}$ . In esti-  
30 mating the number of breaks, we allow for a maximum of five shifts for each  
31 country, which seem sufficient to capture most of the breaks in the data. The  
32 exact number of breaks for each country is estimated using the Schwartz infor-  
33 mation criterion. Also, to ensure that the break date estimator work properly,  
34 we set the minimum length of each regime equal to  $0.1T$ . The sieve bootstrap  
35 is implemented using a lag length of five and we make 1,000 replications.  
36  
37  
38  
39  
40  
41

42 As for the deterministic specification, since all variables are expressed in  
43 GDP ratios, no time trends are required. Two models are considered. The  
44 first assumes a fixed mean while the second allows for up to five breaks in  
45 the mean of each series. Among these models, since most of the literature  
46 indicates the presence of breaks, the latter stands out as the most natural  
47 one.  
48  
49  
50

51 Table 1 shows the results of panel stationary test for each variable. For  
52 each model, the first row represents the test value, the second row contains the  
53 asymptotic  $p$ -values, and the third row contains the bootstrapped  $p$ -values.  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6 The results from the asymptotic  $p$ -values suggest that the null hypothesis  
7 of stationary can be safely rejected. However, these values assume that the  
8 countries are independent, which is unlikely to hold. In order to account  
9 for this dependence, we use the bootstrapped  $p$ -values instead. Except for  
10 revenues in the structural break model, the conclusions are not altered by  
11 taking the cross-country correlations into account. Thus, since the overall  
12 evidence in favor of a rejection is quite overwhelming, we choose to proceed  
13 as if the variables are in fact non-stationary.  
14  
15  
16  
17  
18

### 19 20 **3.3 Cointegration tests**

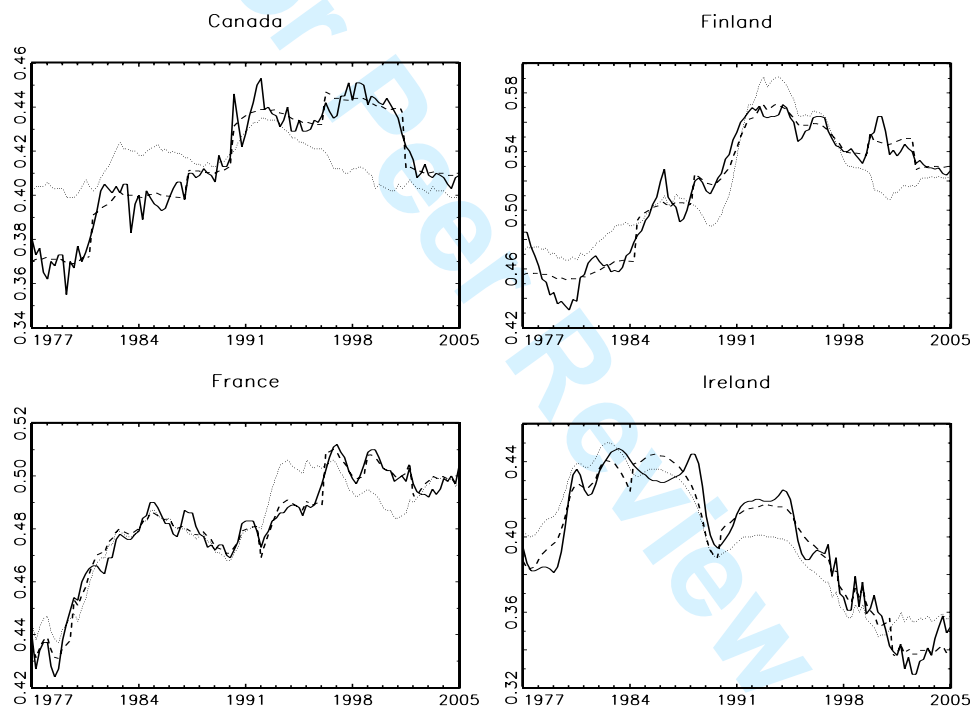
21  
22 The second step in our analysis is to test whether the variables are cointegrated  
23 using the  $Z(M)$  test. The results are presented in the rightmost column of  
24 Table 1. It is seen that the  $p$ -values for the model with a fixed constant  
25 provide very little support of cointegration. However, these results do not  
26 take into account the possibility of structural breaks, and are therefore prone  
27 to erroneous conclusions. Indeed, if we allow for structural shifts and cross-  
28 country correlation, the null hypothesis of cointegration cannot be rejected at  
29 the 5 percent level. We therefore conclude that the variables appear to be  
30 cointegrated around a broken intercept.  
31  
32  
33  
34  
35  
36

37 Table 2 reports the estimated breakpoints obtained from the Bai and Per-  
38 ron (2003) procedure. It is seen that there are at least three breaks for each  
39 country, which is indicative of structural instability. This evidence seems con-  
40 sistent with the gradual shift towards increased business cycles synchroniza-  
41 tion among the OECD countries, which has substantially affected the fiscal  
42 position of many governments, see Fatas and Mihov (2003). The first break  
43 period in the 1980's reflects the effect of the Great Depression, during which  
44 the synchronized recession coincided with an increase in real interest rate. This  
45 caused a substantial deterioration of the fiscal stance of of governments of rich  
46 OECD countries, as the deficits created by the spiralling interest payments  
47 could not be offset, see OECD (2002).  
48  
49  
50  
51  
52  
53  
54

55 Later, following the Maastricht Treaty and the Stability and Growth Pact  
56  
57

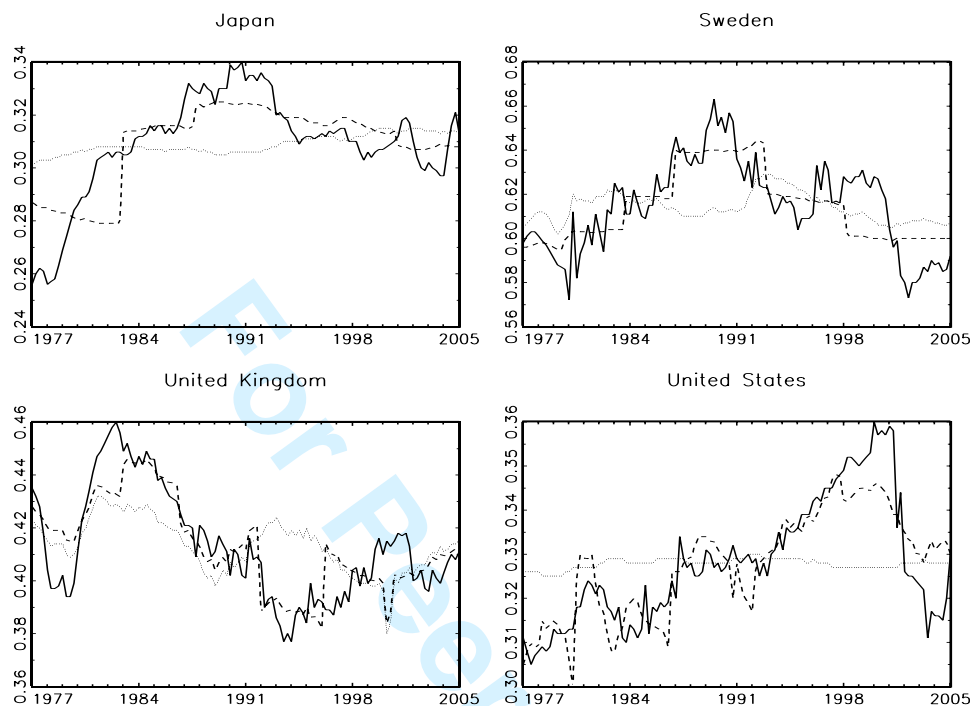
in the early 1990's, the fiscal stance of many OECD countries were driven primarily by an increased coordination of both fiscal and monetary policy within the European Monetary Union, see Darvas *et al.* (2005). This period is also clearly visible in the table, as indicated by the concentration of breaks in the early 1990's, and seems consistent with the results reported in many earlier studies, see for example Afonso (2005). The many breaks in the late 1990's are most likely due to the formation of the European Monetary Union in January 1999.

Figure 1: Revenues-GDP ratio with fitted trend functions.



To further illustrate the importance of accounting for structural change, Figures 1 and 2 plot revenues-GDP ratio, the dependent variable, together with the fitted trend functions for the models with and without breaks. In the figures, the solid line represents revenues measured as GDP ratio, the dashed line represents the model with breaks and the dotted line represents the

Figure 2: Revenues-GDP ratio with fitted trend functions.



model without any breaks. It can easily be seen that nearly all revenue series exhibit several level breaks, most of which seem to reflect the effect of economic fluctuations, and changes in the monetary and fiscal policy objectives in the individual countries. Interestingly, it is seen that the revenues measured as GDP ratio of most countries have been rising during the first half of the sample and falling during the second half, which reflects a more flexible response of the recent fiscal policy of most OECD governments to the economic downturn.

More importantly, we see that the trend lines for the model with breaks seem to provide a very good fit to the revenues for all eight countries. By contrast, the trend lines for the model without breaks are generally incapable of accounting for the all variation in revenues, thus leading to a poor fit. This effect is particularly striking for the United States.



### 3.4 Cointegration estimation

Since the variables appear to be cointegrated, it is possible to estimate and test whether the cointegrating slope is indeed equal to one as required by sustainability hypothesis.

It is well known that the OLS estimator is consistent under fairly general conditions when applied to the cointegrated regression in (7). Unfortunately, if the regressor is endogenous, then this estimator suffers from nuisance parameter dependencies even asymptotically, which makes it a poor candidate for inference. To account for this, in addition to using OLS, we employ both FMOLS and dynamic OLS (DOLS) panel estimation techniques, see Kao and Chiang (2000). All three estimators are based on pooling along the cross-sectional dimension, and are appropriate for testing the null hypothesis that  $\beta_i = 1$  for all  $i$  against the alternative hypothesis that  $\beta_i$  is equal to some common slope value, which is different from unity.

Moreover, as when testing for stationarity and cointegration in panel data, the presence of cross-country dependence makes inference based on the asymptotic normal distribution inappropriate, in which case bootstrap inference might be better. The particular bootstrap opted for this section is taken from Chang *et al.* (2006), who propose a sieve resampling scheme that preserves the serial correlation properties of the errors, and that can be generalized along the lines of Westerlund (2006) to also accommodate cross-country dependence. As an indication of the severity of this problem, we applied the Pesaran (2004) test of cross-section correlation to the OLS residuals of the estimated cointegrated regression. The computed test value was 4.98, which, when compared to the right tail of the asymptotic normal distribution, lead to a clear rejection of the no correlation null at all conventional significance levels. Thus, cross-country correlation is indeed a problem that needs to be addressed.

The results from the OLS, FMOLS and DOLS estimators are reported in Table 3 together with the double-sided  $p$ -values for the null hypothesis of a unit slope. The first thing to notice is that the estimated slopes lie very close

1  
2  
3  
4  
5  
6 to their hypothesized values of one. Indeed, the range of the estimated slopes  
7 is 0.909 to 0.934. The closeness of these estimates to their expected value  
8 based on the sustainability hypothesis is supported by the  $p$ -values.  
9

10  
11 Indeed, by looking at the  $p$ -values based on the asymptotic normal dis-  
12 tribution, we see that the null hypothesis of a unit slope cannot be rejected  
13 even at the most generous 10 percent level. The results from the bootstrapped  
14  $p$ -values, which allow for cross-country dependence, are even more supportive  
15 of the unit slope null, and are all well above 0.6. Thus, based on this evidence  
16 we cannot reject the null hypothesis of a unit slope in the panel. Thus, in  
17 contrast to much of the earlier evidence regarding the sustainability of fiscal  
18 policy, we are unable to reject the sustainability hypothesis for the panel as a  
19 whole.  
20  
21  
22  
23  
24  
25  
26

## 27 4 Concluding remarks

28  
29 Earlier empirical research based on testing for stationarity of public debt and  
30 budget deficit has tended to reject the hypothesis of fiscal sustainability. As  
31 a response to this, recent work within this field has turned towards more  
32 flexible testing approaches based on cointegration. The idea is that if fiscal  
33 policy is sustainable, then government revenues and expenditures should be  
34 cointegrated with a unit slope on expenditures. Unfortunately, the results of  
35 the earlier studies have not been very convincing.  
36  
37  
38  
39  
40

41 In this paper, we provide some insights suggesting that the weak empirical  
42 support of the sustainability hypothesis may be due to inadequate econometric  
43 methods. In particular, we argue that a failure of accounting for structural  
44 breaks and the poor precision of commonly applied time series tests may well  
45 result in a rejection of the sustainability hypothesis. To circumvent these  
46 problems, we resort to some recent advances in the area of non-stationary  
47 panel data. These methods are more powerful than conventional time series  
48 methods, and are general enough to allow for cross-country dependence and  
49 an unknown number of structural breaks in the cointegrating relationship.  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

1  
2  
3  
4  
5  
6 Based on data covering eight rich OECD countries over the period from  
7 1977Q1 to 2005Q4, we show that revenues and expenditures are non-stationary  
8 and cointegrated. We also provide evidence suggesting that the null hypoth-  
9 esis of a common unit slope on expenditures cannot be rejected, suggesting  
10 that the sustainability hypothesis cannot be rejected. The results on the esti-  
11 mated structural breaks are suggestive of strong structural instability, and are  
12 compatible with the findings of many earlier studies such as Quintos (1995),  
13 Martin (2000) and Afonso (2005). Some of the estimated breaks are related  
14 to the increased business cycle synchronization among the OECD countries,  
15 and the policy interventions resulting from the increased coordination of the  
16 monetary and fiscal policy within the European community.  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Table 1: Panel stationarity and cointegration tests.

Specification	Test	Stationarity tests		
		$R_{it}$	$\tilde{G}_{it}$	$Z(M)$
Fixed constant	Value	20.147	18.684	16.661
	$p$ -value <sup>a</sup>	0.000	0.000	0.000
	$p$ -value <sup>b</sup>	0.000	0.002	0.028
Constant with breaks	Value	2.192	5.114	1.416
	$p$ -value <sup>a</sup>	0.014	0.000	0.078
	$p$ -value <sup>b</sup>	0.242	0.008	0.678

*Notes:* All test results allow for up to five structural breaks for each state. To handle the serial correlation, the Bartlett kernel is used with a bandwidth parameter of  $T^{1/3}$ .

<sup>a</sup>The  $p$ -value is based on the asymptotic normal distribution.

<sup>b</sup>The  $p$ -value is based on the bootstrapped distribution. The number of lags in the sieve approximation is five and we make 1,000 bootstrap replications.

Table 2: Estimated breaks.

Country	No.	Breakpoint				
		1	2	3	4	5
Canada	5	1981Q1	1987Q2	1990Q2	1996Q2	2001Q3
Finland	4	1984Q3	1988Q2	1999Q4	2002Q4	–
France	5	1979Q4	1992Q1	1996Q2	1999Q1	2002Q1
Ireland	3	1983Q1	1987Q4	2001Q1	–	–
Japan	5	1979Q4	1983Q4	1987Q1	1993Q1	1998Q2
Sweden	5	1983Q1	1986Q4	1992Q1	1996Q2	2001Q2
United Kingdom	4	1980Q3	1986Q4	1992Q2	1998Q1	–
United States	5	1980Q3	1987Q2	1994Q2	1997Q3	2002Q1

*Notes:* The breaks are estimated using the Bai and Perron (2003) procedure with a maximum number of five breaks for each state. The minimum length of each break regime is set to  $0.1T$ .

Table 3: Panel cointegration slope estimates.

Estimator	Value	$p$ -value <sup>a</sup>	$p$ -value <sup>b</sup>
OLS	0.925	0.594	0.708
DOLS	0.934	0.689	0.718
FMOLS	0.909	0.444	0.602

*Notes:* All results allow for up to five structural breaks for each state. The  $p$ -values for the between and within estimates are for the null of a unit slope against the alternative that the slope is different from unity. <sup>a</sup>The  $p$ -value is based on the asymptotic normal distribution. <sup>b</sup>The  $p$ -value is based on the bootstrapped distribution. The number of lags in the sieve approximation is five and we make 1,000 bootstrap replications.

## References

- Afonso, A. (2005). Fiscal Sustainability: the unpleasant European case. *FinanzArchive* 61, 19-44.
- Ahmed, S., and Rogers, J. (1995). Government budget deficits and trade deficits. Are present value constraint satisfied in long-term data? *Journal of Monetary Economics* 36, 351-374.
- Baglioni, A., and Cherubini, U. (1993). Intertemporal budget constraint and public debt sustainability: the case of Italy. *Applied Economics* 25, 275-283.
- Bai, J., and Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics* 18, 1-22.
- Bohn, H. (1991). The sustainability of budget deficit with lump-sum and with income-based taxation. *Journal of Money, Credit and Banking* 23, 581-604.
- Bordo, M.D., and Heibling, T. (2003). Have national business cycles become more synchronized? NBER Working Paper No. 10130.
- Breitung, J., and Pesaran, M. H. (2005). Unit roots and cointegration in panels. Forthcoming in Matyas, L., Sevestre, P. (Eds.), *The Econometrics of Panel Data*, Klüver Academic Press.
- Carrion-i-Silvestre, J. L., Del Barrio-Castro, T., and López-Bazo, E. (2005). Breaking the panels: an application to the GDP per capita. *Econometrics Journal* 8, 159-175.
- Chang, Y., Park, J. Y., and Song, K. (2006). Bootstrapping cointegrating regressions. *Journal of Econometrics*, forthcoming.
- Darvas, Z., Rose, A.K., and Szapary, G. (2005). Fiscal divergence and business cycle synchronization: irresponsibility is idiosyncratic. NBER Working Paper No. 11580.

- 1  
2  
3  
4  
5  
6 European Commission. (2002). Public finances in EMU. *European Economy*  
7 3.  
8  
9  
10 Fatas, A., and Mihov, I. (2003). The case of restricting fiscal policy discretion.  
11 *Quarterly Journal of Economics* 118-4, 1419-1447.  
12  
13  
14 Hakkio, G., and Rush, M. (1991). Is the budget deficit too large? *Economic*  
15 *Inquiry* 29, 429-445.  
16  
17  
18 Haug, A. (1995). Has federal budget deficit policy changed in recent years?  
19 *Economic Inquiry* 33, 104-118.  
20  
21  
22 IMF (2001). Business cycle linkages among major advanced economies.  
23 *World Economic Outlook*, October 2001.  
24  
25  
26 Kao, C., and Chiang, M. H. (2000). On the estimation and inference of a  
27 cointegrated regression in panel data. *Advances in Econometrics* 15,  
28 179-222.  
29  
30  
31 Kirchgaessner, G., and Prohl, S. (2006). Sustainability of Swiss fiscal policy.  
32 forthcoming in *Swiss Journal of Economics and Statistics*.  
33  
34  
35 MacDonald, R. (1992). Some tests of the government's intertemporal budget  
36 constraint using U.S. data. *Applied Economics* 24, 1287-1292.  
37  
38  
39 Martin, G. (2000). U.S. deficit sustainability: a new approach based on  
40 multiple endogenous breaks. *Journal of Applied Econometrics* 15, 83-  
41 105.  
42  
43  
44  
45 Newey, W., and West, K. (1994). Autocovariance lag selection in covariance  
46 matrix estimation. *Review of Economic Studies* 61, 631-653.  
47  
48  
49 OECD (2002). Fiscal sustainability: the contribution of fiscal rules. *OECD*  
50 *Economic Outlook* 72.  
51  
52  
53 Olekalns, N. (2000). Sustainability and stability? Australian fiscal policy in  
54 the 20th century. *Australian Economic Papers* 39, 138-151.  
55  
56  
57  
58  
59  
60

- 1  
2  
3  
4  
5  
6 Papadopoulos, A., and Sidiropoulos, M. (1999). The sustainability of fiscal  
7 policies in the European Union. *International Advances in Economic*  
8 *Research* 5, 289-307.  
9  
10  
11 Payne, J. (1997). International evidence on the sustainability of budget  
12 deficits. *Applied Economic Letters* 12, 775-779.  
13  
14  
15 Pesaran, H. (2004). General diagnostic tests for cross section dependence in  
16 panels. Cambridge Working Paper in Economics 435.  
17  
18  
19 Psaradakis, Z. (2003). A sieve bootstrap test for stationary. *Statistics and*  
20 *Probability Letters* 62, 263-274.  
21  
22  
23 item Quintos, C. (1995). Sustainability of the deficit process with struc-  
24 tural shifts. *Journal of Business and Economic Statistics* 13, 409-417.  
25  
26  
27 Smith, G., and Zin, S. (1991). Persistent deficits and the market value of  
28 government debt. *Journal of Applied Econometrics* 6, 31-44.  
29  
30  
31 Westerlund, J. (2005). Testing for panel cointegration with multiple struc-  
32 tural breaks. *Oxford Bulletin of Economics and Statistics* 68, 101-132.  
33  
34  
35 Wilcox, D. (1989). The sustainability of government deficits: implications of  
36 the present-value borrowing constraint. *Journal of Money, Credit, and*  
37 *Banking* 21, 291-306.  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60