EVALUATING THE DYNAMICS OF FISCAL POLICY IN SPAIN: PATTERNS OF INTERDEPENDENCE AND CONSISTENCY OF PUBLIC EXPENDITURE AND REVENUES

Francisco de Castro,
José Manuel González-Páramo
and Pablo Hernández de Cos

Banco de España

Banco de España — Servicio de Estudios
Documento de Trabajo n.º 0103

EVALUATING THE DYNAMICS OF FISCAL POLICY IN SPAIN: PATTERNS OF INTERDEPENDENCE AND CONSISTENCY OF PUBLIC EXPENDITURE AND REVENUES[*]

Francisco de Castro
Banco de España
Universidad Complutense de Madrid

José Manuel González-Páramo
Universidad Complutense de Madrid

Pablo Hernández de Cos
Banco de España

January, 2001

[*] We wish to acknowledge J. Galí, J. M. Marqués and J. Vallés for their suggestions and comments. We also thank J. L. Fernández for providing us with the programs needed to carry out some of the tests, and R. Peruga for very helpful discussions.

Abstract

Two issues are addressed in this paper. First, we attempt to ascertain whether the current fiscal policy in Spain satisfies the intertemporal borrowing constraint. For this purpose, we apply the traditional empirical tests of sustainability proposed in the literature, that pay special attention to integration orders of deficit and debt processes and the existence of cointegration relationships between revenues and expenditures. Under this approach, a sustainable fiscal policy would indicate that if fiscal variables follow the pattern of the past in the future, no problems in marketing public debt are expected to arise. Our results show that a structural break seems to have taken place gradually in the Spanish budget performance, which allows to verify the intertemporal borrowing constraint in a "strong sense".

Second, given the limitations of the previous sustainability analysis, the fact that the change to strong sustainability is not fully confirmed for a large sample period and that, even if this were the case, the Stability and Growth Pact for the EMU countries sets explicit ceilings to public finance that might require additional consolidation efforts in Spain, we consider the issue of which is the most efficient strategy to achieve permanent reductions in fiscal deficits. Thus, we analyse empirically the possible interdependence between government expenditure and revenues. We examine this issue by performing the standard Granger causality tests. We find that there is a bias towards deficit in public sector's size and that public expenditure causes revenues in the short and long term, whereas revenues only cause expenditures in the long term. These results imply that fiscal consolidation could be achieved with a reduction in the size of the public sector and that the most adequate strategy of fiscal consolidation should consist on the reduction of structural public expenditure.

1. INTRODUCTION

The most common definition of fiscal sustainability, and that employed in this paper, is based on the need for a fiscal deficit to be financed, i.e. on the concept of inter-temporal budget constraint, which requires that the current market value of debt be equal to the discounted sum of expected future primary surpluses. In this context, fiscal policy is sustainable if the discounted value of debt reaches zero at the limit.

The issue of whether the current fiscal policy can be maintained indefinitely, i.e. whether it is sustainable or not, is an important one, since it will determine the need for future discretionary policy actions. In this sense, since the concept of sustainability relies on the fact that governments need enough resources to ensure their ability to carry out the functions attributed to them, sustainability analysis helps to determine whether a current policy can be maintained in the long run with the ongoing ability to generate financial resources. In addition, fiscal sustainability has clear implications for other macroeconomic variables. In particular, a non-sustainable fiscal policy involves a risk of future interest rate rises leading to a slowdown in economic growth.

This need for fiscal sustainability and, in the case of EMU countries, the commitment to meet the fiscal targets set out in the Stability and Growth Pact, might require the correction of fiscal imbalances. In this sense, knowing the causes of fiscal deficits seems crucial to guarantee permanent reductions in fiscal deficits. In other words, the possible interdependence between government expenditure and revenues, should be analysed in order to select the adequate strategy of fiscal consolidation. Thus, the central issue is whether higher taxes lead to expenditure changes or whether expenditure growth leads budget dynamics, with taxes following suit.

In this paper, we address these two issues, sustainability of public finances and causality between public revenues and expenditures, for the case of Spain. Testing sustainability and causality in the Spanish case is of particular relevance due to the fact that several policy actions have been adopted in Spain since 1975, aimed at implementing a European Welfare State model and a modern tax system. This has led to a sharp increase in public expenditure and revenue. Furthermore, in recent times, Spain has gradually reduced its public deficit mainly as a result of a drop in spending that might have important consequences for fiscal sustainability.

First, as regards sustainability analysis, different tests are proposed in the literature. These pay special attention to integration orders of deficit and debt processes, and to the underlying stochastic structures and the existence of cointegration relationships between revenues and expenditures. Earlier tests in this literature indicated that the condition for fiscal sustainability is the stationarity of the debt (Hamilton and Flavin, 1986) or that the discounted debt process follows an I(0) process without drift (Wilcox, 1989)[1]. Later work developed alternative conditions for fiscal sustainability: provided that total public revenue and expenditure are first-order integrated, sustainability requires both series to be cointegrated (Hakkio and Rush (1991), Haug (1991), Smith and Zin (1991), Trehan and Walsh (1988,1991)). More recently, Quintos (1995) extended this literature by introducing "strong" and "weak" conditions for fiscal sustainability. The "strong" condition corresponds to the stationarity of the deficit process, while the "weak" condition verifies for higher than one orders of integration of the public debt, or even for some mildly explosive processes for this

[1] Wickens and Uctum (1993) develop a test for sustainability when a feedback rule between the deficit and debt is introduced.

variable, implying the intertemporal borrowing constraint is sastified but at a slower rate than in the stronger version.

From an economic point of view, there are important differences between the concepts of strong and weak sustainability[2]. Strong sustainability is understood as a situation in which no future problems in the deficit's behaviour are expected to arise, and there is therefore no need for structural fiscal reforms, in the absence of significant changes in the processes followed by both public expenditures and revenues. However, weak sustainability implies that governments might have future problems in marketing their debt, involving a substantial risk of a rise in interest rates that may have perverse effects on economic growth and the public budget, necessitating fiscal reforms, or at least a consolidation effort. This possibility of future debt-marketing problems comes from the lower speed at which the intertemporal borrowing constraint is satisfied, as a consequence of the higher growth rate of the debt and, consequently the higher level for this variable in future years. Accordingly, the difference between both concepts of sustainability is quite relevant, both from a positive and normative analysis of fiscal policy developments, in that weak sustainability can be taken as an indicator of the need for fiscal consolidation. On the other hand, the relevance of both concepts of sustainability is of great interest for the associated consequences for the macroeconomic variables of interest implied in both definitions.

Camarero et al. (1998) apply the aforementioned tests to the Spanish case, showing that public revenues and expenditures are cointegrated only when the possibility of structural shifts in this relationship is taken into account. According to their analysis, the deficit process is found to be sustainable in the weak sense. However, since over the sample period many fiscal reforms have taken place in Spain, a deeper univariate analysis of the series involved might be of great interest and may provide useful information for deriving sounder conclusions about the sustainability of Spanish fiscal policy in recent years. In this context, the existence of changes in the order of integration, which can be associated with fiscal reforms or with gradual fiscal adjustments, might modify previous results and thus our conclusions. For this reason, in this paper we apply the traditional tests of sustainability, following Quintos´ approach. In addition, we introduce a univariate analysis of the series, consisting of testing whether the series have undergone changes in the order of integration that may, in some way, vary or even invalidate the cointegration analysis.

Second, as regards the interdependence between revenues and expenditure, theoretical arguments give support to any possible direction of causality. For instance, Buchanan and Wagner (1977) argue that a deficit financing facilitates higher spending due to the existence of *fiscal illusion*; Brennan and Buchanan (1980) suggest that in a leviathan-type government higher taxes today lead to more spending tomorrow; Barro (1979) and Peacock and Wiseman (1979) contend that increases in expenditure in the present tend to be followed by tax increases in the future; Other authors support the hypothesis of interdependence between revenues and expenditure, initially posed by Wicksell (1896) and reformulated by Musgrave (1966) and Meltzer and Richard (1981). This theory predicts that decisions on both variables are adopted and suffered by the same groups, so they will be jointly and interdependently adopted. Empirical evidence is also far from conclusive. Some studies have reported results showing that revenues cause expenditure (Manage and Marlow, 1986; Blackley; 1986) while some others support the opposite conclusion (Anderson, Wallace and Warner, 1986; von

[2] In the context of this paper a sustainability analysis makes sense under the existence of positive debt and persistent deficits. Conversely, under an excess of assets over liabilities and persistent surplus sustainability is, by definition, always guarantied.

Furstenberg, Green and Jeong; 1986) and some searchers have found a bidirectional causality (Owoye, 1995)[3].

In this paper, we examine empirically the issue of causality between general government revenues and expenditure in Spain by performing the Granger causality test, allowing or not for cointegration, and by means of the variance decomposition and impulse response functions in the context of a VAR analysis.

The rest of the paper is structured as follows. Section 2 presents a summary of fiscal policy performance during the sample period (1964-1998), which may be of particular interest for better understanding the results. Section 3 and 4 contains the sustainability and causality analysis, respectively. Section 5 draws the conclusions.

2. DEFICIT AND DEBT IN SPAIN[4]

Since 1964, five distinct periods can be identified in relation to public finances in Spain:

1964-75: a period of strong economic expansion, with average real GDP growth of 6.4%, characterised by small budget surpluses, owing to the steady growth of government revenue and expenditure.

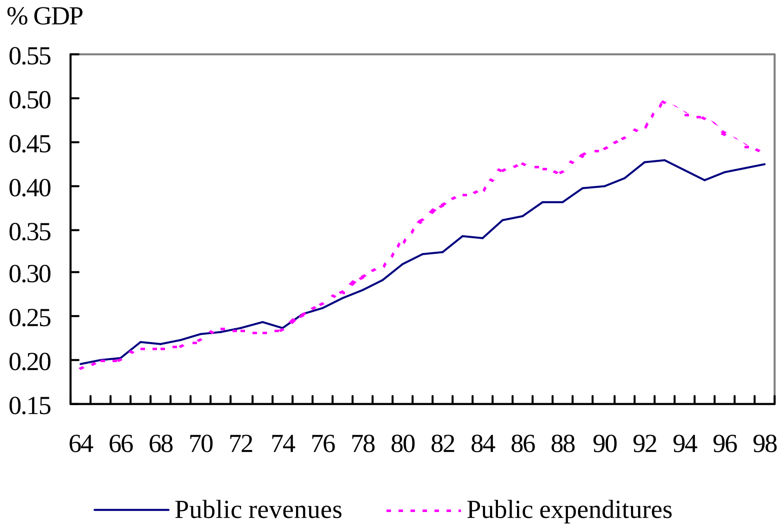
1975-85: against a background of economic crisis and political change, the previous situation changed in 1976, with the appearance of a budget deficit. Although it was small in the first two years, it grew continuously, except in 1979 and 1983, to reach 5.8% of GDP in 1985. On the one hand, public expenditure as a percentage of GDP almost doubled in this period (from 23.5% of GDP in 1974 to 41.6% en 1985, which represents an average annual increase of 1.6% of GDP) due to low economic growth (average real GDP growth of 1.6%) and the building of the Welfare State. On the other hand, public revenues also increased significantly as a consequence of the 1977 and 1978 fiscal reforms, but at lower rates than in the case of expenditure (total revenues moved from 23.6% of GDP in 1974 to 35.8% en 1985, which represents an average annual growth of 1% of GDP).

As a consequence of this budgetary imbalance, public debt also spiralled, from 12.1% of GDP in 1979 to 43.7% in 1985. However, this increase in debt did not lead to a similar rise in the interest burden because, until 1982, around two-thirds of the budget deficit was funded by the Banco de España and financial institutions, primarily through compulsory reserve requirements. In fact, public debt assumed by the private and external sectors under orthodox financing arrangements played a very limited role, covering less than 25% of the state-borrowing requirement. Nonetheless, as from 1983, the deficit was funded in a more orthodox fashion, and the government came to rely more heavily on Treasury bill issuance. This, together with the high interest rates prevailing in the period, caused the interest burden to double as a proportion of GDP between 1982 and 1985 (to 2% in 1985).

1986 to 1988: following Spain's accession to the European Community and the commencement of a new cyclical expansion, there was a change in direction in

[3] Payne (1998) examines the relationship between revenues and expenditures for USA and finds that the argument that tax cause expending is supported for 24 states, while the opposite conclusion is supported for 8 states and a bidirectional causality is valid for 11 states.
[4] See Argimón, Gómez, Hernández de Cos and Martí (1999) for a deeper analysis of fiscal policy in Spain.

Figure 1. Public revenues and expenditures



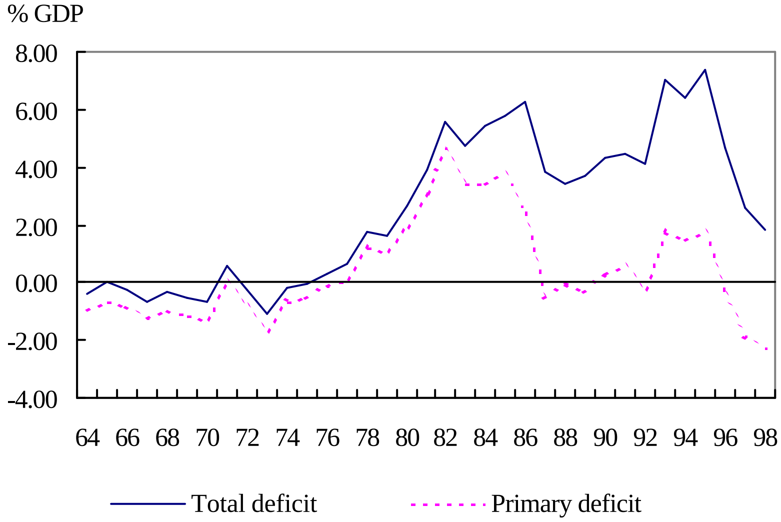
Source : Banco de España

Spanish fiscal policy. The budget deficit was reduced from 5.8% in 1985 to 3.4% in 1988, essentially due to the growth of government revenue. In fact, public revenue as a percentage of GDP increased 2.2 percentage points while public expenditure fell by only 0.2 percentage points. Moreover, there was a significant improvement in the primary balance, which swung from -3.8% in 1985 to a small surplus in 1988, enabling public debt to be whittled down to 41.7% in 1988.

1989 to 1993: the aforementioned period of fiscal restraint came to an end in 1989, when the budget deficit began to grow again, to reach 7% at the height of the economic crisis in 1993. The primary balance followed a similar path to the deficit. After small surpluses between 1987 and 1989, it moved into deficit in 1990, rising to 1.8% of GDP in 1993. As regards public revenues and expenditures, similarly to the period 1975-1985, both increased significantly, reaching 42.8% and 49.8% of GDP, respectively, in 1993. Finally, there was only a slight increase in public debt, to 45.9% of GDP, primarily as a consequence of the strong growth in GDP between 1989 and 1991 (11% in nominal terms), and despite the increase in the cost of debt during this period. Thereafter, however, it rose to exceed 60% of GDP in 1993, as a consequence of the increase in the budget deficits, the fall in nominal GDP growth and the prohibition on monetary financing of the deficit as from 1994, under the Treaty on European Union. At the same time, the interest burden rose, reaching 5.2% of GDP in 1993.

1994 to 1998: fiscal policy was constrained in this period by the commitment to meet the convergence criteria set out in the Treaty on European Union to regulate access to the Third Stage of EMU. In accordance with this commitment, the tendency to imbalance in public finances came to an end in 1994, with a moderate reduction in the deficit. However, this was reversed again in 1995, when the budget deficit reached 7.3%. Thereafter, there was a gradual decline in the deficit, which reached 1.8% in 1998, in a context of economic recovery. The reduction in the public deficit was the result of a drop in spending, which fell by 5.7 points relative to GDP. Meanwhile, the share in GDP of total general government revenue declined slightly.

Figure 2. Total and primary public deficit



Source : Banco de España

Finally, public debt peaked at over 70% of GDP in 1996 and then declined slightly to 67.4% of GDP in 1998. The factors responsible for this decline include the existence of primary surpluses in 1997 and 1998, the fall in interest rates and the revenue obtained from the privatisation of state-owned firms. Finally, interest payments, which peaked in 1995 (at 5.6% of GDP), fell to 4.1% of GDP in 1998. This can be explained by both the reduction in the level of public debt in 1997 and 1998 and the decline in interest rates from 1995 (the average interest rates on Treasury bills fell from 9.1% in 1995 to 3.7% in 1998, while those on government bonds fell from 11.1% in 1994 to 6.4% in 1998).

The empirical results obtained in this paper are based on annual data for Spain of public debt (b_t), public deficit (d_t) and public revenues (t_t) and expenditure (g_t^R) for the period 1964-1998 taken as ratios over GDP[5]. The real GDP growth rate (Δy_t) and the inflation rate (Δp_t) are obtained as the first difference of the natural logs of the original series. Usual unit root tests are carried out in appendix A at the end of the paper. We are aware that any long-run analysis based on such a small number of observations may be somewhat troublesome. Moreover, the well-known lack of power of unit-root tests added to this problem obliges us to treat the results with the greatest care. Without forgetting these difficulties, we will present our main findings in the following sections.

[5] The data employed is based on ESA79 methodology.

3. SUSTAINABILITY ANALYSIS

3.1. Theoretical framework

As stated in the introduction, the definition of sustainability employed in this paper is based on the concept of the fulfilment of the intertemporal borrowing constraint of the government. In period t the budget constraint can be expressed as follows[6]:

$$\Delta B_{t+1} = r_t B_t + G_t - T_t$$

with B_t being the stock of debt at the end of period t-1 in real terms, G_t real public expenditure excluding interest payments, T_t real public revenues and r_t the average real interest rate on the debt in period t-1. Thus, the term G_t - T_t is defined as the primary deficit. Accordingly, total public expenditures are

$$G_t^R = r_t B_t + G_t$$

Therefore, public deficit is defined as $D_t = G_t^R - T_t$. However, the latter variables are not the most accurate ones in a sustainability analysis. In fact, few or no conclusive results can be drawn from variables that show an upward trend if the economy shows a similar pattern. In other words, the relevant variables must be considered by taking into account the size of the economy, and any sustainability analysis should thus be performed using the latter variables as percentages of GDP and focusing on the burden that public debt imposes on the economy. Therefore, the budget constraint in period t and the definition of total public expenditures, both in GDP terms[7], are now

$$\Delta b_{t+1} = l_t b_t + g_t - t_t \qquad [1]$$

$$g_t^R = l_t b_t + g_t$$

where $l_t = \dfrac{r_t - h_t}{1 + h_t}$, which can be understood as the addition to net debt due to the excess of the real interest rate over h_t , the real GDP growth rate. Taking the excess of the real interest rate over the growth rate of the economy as stationary around a mean l[8], [1] can be expressed as

$$\Delta b_{t+1} = l b_t + gx_t - t_t \qquad [2]$$

where $gx_t = g_t + (l_t - l) b_t$. Solving forward [2], we obtain

$$b_t = \sum_{j=0}^{\infty} g^{\,j+1} (t_{t+j} - gx_{t+j}) + \lim_{j \to \infty} g^{\,j+1} b_{t+j+1} \quad ; \quad g^{\,j+1} = (1 + l)^{-(j+1)} \qquad [3]$$

---

[6] In this paper, seigniorage is not considered as a source of public revenues since the current institutional framework in EMU avoids the possibility of deficit financing through monetisation. Nevertheless, we are aware that, in the sample period of analysis, the fiscal policy has been often conducted by extensive use of seigniorage.

[7] The lower case letters indicate the same concepts in terms of GDP.

[8] If l_t is always negative the deficit process is sustainable and such an analysis would lack of relevance.

Equation [2] and its implication [3] cannot be a subject of controversy, for they only summarize definitions of fiscal policy. As Hamilton and Flavin (1986) point out, what is of economic interest and subject to empirical refutation is what the creditors expect about the behaviour of bubble term in [3]. Taking expectations in this equation, the hypothesis that the government is subject to the intertemporal borrowing constraint can be expressed as

$$b_t = E_t \sum_{j=0}^{\infty} g^{j+1}(t_{t+j} - gx_{t+j})$$

which is mathematically equivalent to the transversality condition $E_t \lim_{j \to \infty} g^{j+1} b_{t+j+1} = 0$. Such transversality condition has a very well defined economic sense. It implies that, for a process to be sustainable, the current debt must equal the expected present value of future primary surpluses. Otherwise, stabilisation measures will be required in order to coax the public deficit back to a sustainable path.

Our empirical testing requires using the representation of [3] in terms of Δb_t, which yields to the following expression:

$$g_t^R - t_t = \sum_{j=0}^{\infty} g^{j+1}(\Delta t_{t+j} - \Delta gx_{t+j}) + \lim_{j \to \infty} g^{j+1} \Delta b_{t+j+1}. \qquad [4]$$

where the left-hand side of [4] represents the public deficit. In order to impose a constraint analogous to the intertemporal borrowing constraint faced by an individual the following transversality condition should hold:

$$E_t \lim_{j \to \infty} g^{j+1} \Delta b_{t+j+1} = 0 \qquad [5]$$

Sustainability tests in literature aim to verify whether this transversality condition in the government budget constraint holds. These tests pay special attention to integration orders of deficit and debt processes, and to the underlying stochastic structures and the existence of cointegration relationships between revenues and expenditures. A usual procedure consists of testing the stationarity of Δb_t in various forms, or alternatively the stationarity of $g_t^R - t_t$ if both are I(1), according to the method employed by Trehan and Walsh (1988). This procedure implies testing cointegration between revenues and expenditures when the cointegrating vector (1,-1) is imposed. An alternative procedure would be to test cointegration in

$$t_t = a + b g_t^R + e_t \qquad [6]$$

and afterwards, test the null H_0: b=1. Accordingly, the deficit would be non-sustainable if Δb_t is non-stationary, or if cointegration in [6] does not hold with cointegrating vector (1,-1). In this case, the transversality condition holds because $\Delta b_t = Op(1)$ and, accordingly, the limit term in [5] behaves as

$$E_t \lim_{T \to \infty} \exp(-T\, k) = 0 \qquad [7]$$

where k is a positive constant and $Op(\cdot)$ the rate at which a stochastic sequence converges in probability to a non-stochastic sequence. However, as Quintos (1995) shows, these methods only refer to sufficient conditions for sustainability. In general, it is not necessary for Δb_t to be I(0) for [5] to hold. If Δb_t is I(d), being d a finite order of integration, it verifies $\Delta b_t = Op(T^{d/2})$. In this case, the limit term in [5] behaves as

$$E_t \lim_{T\to\infty} \exp(-T\,k)\,T^{d/2} = 0 \qquad [8]$$

This result determines that if Δb_t is an integrated process of any finite order, the discount factor decreases at a higher rate than Δb_t, making the transversality condition, and thus the intertemporal borrowing constraint, hold, although the limit term in [5] approaches zero at a lower speed than in the case when Δb_t is I(0)[9]. Consequently, using Quintos' terminology we will say that a deficit process is sustainable in its strong form if the limit term in the transversality condition behaves as [7], whereas if this limit behaves as [8] the process will be said to be sustainable in its weak form. Therefore, only when Δb_t contains explosive roots of high enough magnitude to offset the discount factor will the deficit be non-sustainable.

As stated before, strong sustainability means that no future problems, according to the current state of affairs, are likely to arise, whereas a weakly sustainable budget performance might lead in the future to problems in debt-marketing, that would involve a risk of interest rates increases. This risk only arises because of the higher stock of debt that a weakly sustainable fiscal policy would imply. Should debt-marketing problems occur, macroeconomic stability would be endangered and severe fiscal reforms should be adopted.

In this context, Quintos shows that b=1 in [6] is only a sufficient condition for sustainability, in that it implies that the transversality condition behaves as [7]. However, it is not a necessary condition. Therefore, the necessary and sufficient condition is 0<b≤1, whereas cointegration is only a sufficient condition. Substituting [6] in [1] we obtain

$$b_{t+1} = (1 + l_t(1-b))b_t + (1-b)g_t - a - e_t \qquad [9]$$

or equivalently

$$\Delta b_{t+1} = l_t(1-b)b_t + (1-b)g_t - a - e_t = (1-b)g_t^R - a - e_t. \qquad [10]$$

If g_t^R is I(1), 0<b<1 implies, given [10], that Δb_t is I(1), no matter whether e_t is I(0) or I(1). In other words, cointegration in [6] plays no role, and consequently the transversality condition will behave as [8], being the deficit process sustainable only in its weak form. On the contrary, Δb_t will be I(0) and thus the deficit strongly sustainable, when simultaneously b=1 and e_t are I(0), i.e. cointegration between public revenues and expenditures holds. If we reject cointegration in [6] and b equals 1, the deficit will be sustainable in its weak form, because according to [10], Δb_t will be I(1) as well. Finally, if b=0 the deficit is not sustainable because according to [10] Δb_t would grow at a rate greater than l, and thus the discount factor would never offset Δb_t. A summary of all the possibilities is found in table 1.

[9] Quintos shows that when Δb_t contains a trend the bubble term still goes to zero, although at a lower rate than in [8], being the deficit weakly sustainable in this case.

Table 1: Quintos' test

		Cases for g_t^R ~ I(1)				
Values for b	**and**	**Cointegration in [6]**	**Yields**	Δb_t	Þ	**Conclusion for sustainability**
b=1		Yes		I(0)		Strong sustainability
b=1		No		I(1)		Weak sustainability
0<b<1		Plays no role		I(1)		Weak sustainability
b=0		Plays no role		I(1)		No sustainability

According to the process described earlier, Quintos suggests first to analyse the orders of integration of the variables g_t^R and t_t, and provided that they are I(1), to estimate [6] and test the null H_0: b=0 against the alternative Ha: b>0. If H_0 is accepted the deficit is not sustainable, whereas if it is rejected the null H_0: b=1 against Ha: b<1 should be tested. Should H_0 be rejected, the result 0<b<1 is obtained and the transversality condition would behave as [8], and accordingly the deficit would be weakly sustainable. In this case, as [9] shows, the undiscounted debt process contains an explosive root. On the other hand, if one cannot reject H_0: b=1, one should test for a cointegration relationship in [6]. In case cointegration is accepted, the transversality condition will behave as [7], and therefore, the strong sustainability result will hold. If, on the contrary, cointegration is rejected in [6], the transversality condition will behave again as [8], and thus the deficit will be weakly sustainable.

3.2. Empirical results for the sustainability analysis

Given that t_t, and g_t^R are I(1) processes (see appendix A), we perform Quintos' test and estimate [6] for the whole sample. We estimated [6] by OLS, the maximum likelihood procedure suggested by Johansen, and by the non-parametrical procedure proposed by Phillips and Hansen (1990)[10]. Once [6] was estimated, we performed several cointegration tests based on the ADF and Phillips (1987) statistics and the Trace Statistic suggested by Johansen. The estimation and cointegration tests results are summarised in table 2 and none of the tests reject the null of absence of cointegration between both variables. Moreover, the estimated coefficient is between zero and one, which, according to Quintos, would lead us to conclude that deficit is sustainable in the weaker form[11]. However, such a result is not at all informative because the absence of cointegration yields a spurious estimation of b. In order to avoid this problem and to complete Quintos' test, we estimated [6] in first differences (Hamilton, 1994) by OLS, yielding an estimated coefficient of b=0.31. This coefficient was statistically different from 0 and 1. We also estimated the equation in first differences by instrumental variables in order to avoid estimation problems derived from the possible endogeneity of the regressors. In this case the estimated coefficient was b=0.27, also statistically different from 0 and 1. Given that the condition 0<b<1 holds, the transversality condition behaves as [8], and accordingly the deficit process is sustainable in its weaker form. Furthermore, by [9] we know that the debt process should have an explosive root, which is consistent with the positive t-ratios obtained in

[10] The latter may be advisable when the regressors may be endogenous, which leads to a second order asymptotic bias in the OLS estimators. The second-order asymptotic bias arises because the estimators are still consistent when cointegration holds. In order to correct this bias, they suggest estimating by instrumental variables, but the instruments do not fully eliminate the asymptotic bias when the regressors are endogenous. Therefore, they suggest semi-parametric corrections in the long run covariance matrix, which lead to asymptotically unbiased-in-median estimators. These fully-modified estimators form the basis of the so-called fully-modified Wald tests, which can be used for testing general linear hypotheses of the coefficients in cointegrating regressions, and their asymptotic distributions are χ^2. The correction in the long-run covariance matrix is based on the procedure suggested by Andrews and Monahan (1992).

[11] The results derived from our estimations here are not very different from those of Camarero, Esteve and Tamarit (1998).

the unit-root tests on b_t without considering either a constant or a deterministic trend (see appendix A). The estimated explosive root in this case has been 1.01, and its small magnitude can explain why the process followed by this variable can be better approximated by a I(1) rather than by a I(2) one.

The results so far obtained are not conclusive in that the power of the ADF and other cointegration tests diminishes in the presence of structural breaks[12]. Over the course of sample period, on the other hand, many fiscal reforms have taken place in Spain. In particular, since the late seventies, fiscal policy in Spain changed from a system in which the general government budgets were formally balanced, or even showed a small surplus, to another –as from 1976- with public deficits, linked to the expansion of spending as a consequence of moving towards European welfare state models. Also, the tax system was thoroughly overhauled, with the introduction of personal income tax in 1978 and VAT in 1986. Further, the progressive move as from 1983 from monetisation to a more orthodox financing of the deficit, which coincides in time with high budget imbalances, resulted in the emergence of a significant public debt balance and a subsequent increase in the interest burden. Lastly, further to the 1978 Constitution, there has been a gradual fiscal decentralisation process, involving a drift of responsibilities for the management of certain services form the State to the regional governments along with the developments of the arrangements for financing these responsibilities. Therefore, in the following section we will consider the possibility of presence of structural breaks and explore whether the conclusions are substantially affected.

Table 2: Long run relationship between t_t and g_t^R

$t_t = a + b\ g_t^R + u_t$

	OLS (Phillips-Ouliaris)	Phillips-Hansen	Johansen	Critical values		
				10%	5%	1%
b	0.77	0.79	0.74			
ADF	-2.01	-1.92		-3.51	-3.80	-4.36
Z_α	-16.96	-17.01		-23.19	-27.08	-32.19
Z_t	-1.05	-1.79		-3.51	-3.80	-4.36
Trace r=0			19.54		19.96	24.60
r≤1			4.41		9.24	12.97
Fully modified Wald test		b=0 704.09*** b=1 51.97***		2.71	3.84	6.63

Note: Critical values for the statistics Z_α and Z_t have been taken fron Phillips and Ouliaris (1990).

The Gregory and Hansen test

Gregory and Hansen (1996) are concerned with the possibility that the cointegrating vector may change during the sample period at a single unknown point in time. If this is true, the standard ADF and Phillips Z_α and Z_t tests lose power. As a consequence, if the true model is cointegrated with a regime shift, standard analysis consisting of estimating [6] and performing ADF or Phillips Z_α and Z_t tests on the cointegrating residuals does not reject the null of no cointegration. Gregory and Hansen propose a statistic that attempts to test the null hypothesis of no cointegration

[12] See Hansen (1992), Hansen and Johansen (1992), Gregory, Nason and Watt (1996) and Gregory and Hansen (1996).

against the alternative of cointegration with a structural break at an unknown point in time, and accordingly consider three possible models. Thus, they allow only for changes in the intercept, without and with time trend, and shifts in both the intercept and slope coefficients. These are referred to as Level shift (C), Level shift with trend (C/T) and Regime shift (C/S) models, respectively. In this context, the stable cointegration relationship without structural breaks is only a particular case. Their procedure consists of estimating by OLS and computing the cointegration tests for every possible break point and selecting as the most probable break point that associated with the highest absolute value for these tests[13] (InfADF).

Table 3: Gregory-Hansen tests for structural breaks

Model	InfADF	Year	Mean ADF	Z_t	Z_α	Year	Crit.Values[a] InfADF, Z_t		Crit values MeanADF		Crit.values Z_α	
							10%	5%	10%	5%	10%	5%
Level shift (C)	-3.79	1990	-2.62	-3.36	-20.46	1987	-4.34	-4.61	-3.53	-3.80	-36.19	-40.48
Regime Shift (C/S)	-4.56	1988	-2.76	-4.21	-26.31	1987	-4.68	-4.95	-3.52	-3.78	-41.85	-47.04

Notes: The critical values have been taken from Gregory and Hansen (1996). Year columns refer to the most probable break points according to the reported statistics.

Table 3 shows the results from the implementation of these tests to the Spanish case. As can be seen, none of the statistics turned out to be significant. These results, together with those in table 2, support the null hypothesis of no cointegration in [6]. However, the statistic InfADF applied to the (C/S) model shows that, albeit non-significant, between 1987 and 1988 a change in the fiscal policy regime may have taken place.

The Hansen test

Hansen (1992) also considers the possibility of a structural break at an unknown point in time, although the null hypothesis is the existence of cointegration, in contrast with the Gregory and Hansen test. Hansen provides three tests for parameter instability based on the information derived from the fully-modified residuals in the cointegrating equation[14]. The statistics of these tests are complementary to those proposed by Gregory and Hansen, in the sense that Hansen tests the null of cointegration with no regime shift against the alternative that a regime shift has occurred. Following Hansen we will call these statistics SupF, MeanF and L_c. Their values are reported in table 4. Accordingly, the L_c and the SupF statistics show evidence of parameter stability, although the MeanF statistic could suggest that a gradual change in the behaviour of the fiscal variables may have taken place.

Note that the null can be rejected not only because there is a regime shift, but also because cointegration does not hold in [6]. Therefore, the L_c statistic can also be understood as a LM cointegration test. According to its value, it suggests a stable long-term relationship between public revenues and expenditures, in contrast with the results derived from the standard cointegration tests and those proposed by Gregory and Hansen.

[13] See appendix B for details on the implementation of this test.
[14] See appendix C for a detailed explanation of these tests.

Table 4: Hansen test

Statistic	Value	P-value
L_c	0.10	0.20
MeanF	3.05	0.17
SupF	6.88	0.20

Note: P-value column shows the probability of parameter instability. Probability equal to 0.20 means ≥ 0.20. According to Hansen (1992), a P-value over 0.20 can be taken as evidence of parameter stability.

The Hansen and Johansen Test

Hansen and Johansen (1993) do not examine directly the stability of the parameters in the cointegration equation, but the stability of the eigenvalues associated with the Error Correction Model (ECM henceforth) that yield to the estimation of the cointegrating vector. They propose a recursive Likelihood Ratio test with null of cointegration for every subsample[15]. The statistics to be considered will be referred to as SupHJ and MeanHJ, which are the maximum and mean of the sequence of all the HJ(t) statistics for every possible break point. The results drawn from this test are presented in table 5. Only the MeanHJ statistic turned out to be significant at the 10% level which could indicate a gradual regime shift, although the evidence on parameter instability is not conclusive at all.

Table 5: Hansen and Johansen test

Statisic	Value	Critical values		
		10%	5%	1%
SupHJ	3.20	3.69	4.81	7.39
Break point	1966			
MeanHJ	0.78	0.69	0.98	1.65

The results from the tests above do not support the hypothesis of the existence of a structural break in the behaviour of the fiscal variables. Rather, they could be taken as evidence of a gradual shift, although the evidence on this latter issue is far from conclusive. Moreover, we do not have clear evidence of cointegration. Neither the Gregory and Hansen test nor the standard cointegration tests allow us to reject the null of no cointegration, although the L_c test would suggest the contrary. Nevertheless, according to Gregory and Hansen (1996), it would be advisable to test first the null of no cointegration and, if it is rejected in favour of the (C) or the (C/S) models, then perform the Hansen test in order to obtain deeper evidence of a regime shift.

So far, our results do not qualitatively differ from those obtained by Camarero et al. (1998). However, given the above results we consider that a deeper univariate analysis of the series involved is of great interest when addressing the problem of sustainability, and may provide us with useful information in order to derive our conclusions. If we consider that in a given period the order of integration of the series

[15] More detailed information on this test can be found in appendix D. Programs have been provided by J.L. Fernández.

involved in the analysis has changed, then the latter analysis is misleading. In this context, changes in the order of integration can be associated with fiscal reforms or even with the achievement of a given level, after a gradual adjustment, in the variables considered after the implementation of the above-mentioned reforms. This is our justification for considering an alternative way of addressing the problem of fiscal sustainability.

Tests of changes in the order of integration

If the order of integration varies over time the implications for a sustainability analysis may be very important, because it can make us change our initial view derived from the estimation of [6] with or without structural breaks. Concretely, if Δb_t is not stationary in the first part of the sample but becomes stationary in the last part, although a global analysis would lead us to conclude that, according to [9], the deficit process is sustainable in a weak sense, the relevant issue for analysing the future behaviour will be the current process followed by this variable. As a result, we should say that the sustainability seems to be turning to its strong form and no future fiscal problems seem to arise in the horizon.

Earlier work on this issue has been done by Leybourne, McCabe and Tremayne (1996) and Maeso (1997). Here, we follow Fernández (1999) and use a sequential procedure in order to get the following statistics: $Supt_{\delta 1}$, $Meant_{\delta 1}$, $Supt_{\delta 2}$, $Meant_{\delta 2}$, $Supt_{\alpha 1}$, $Meant_{\alpha 1}$, $Supt_{\alpha 2}$ and $Meant_{\alpha 2}$. As before, the Sup statistics have power for a unique break point, whereas the mean ones have power for gradual changes. According to Zivot and Andrews (1992), the break point is associated with the observation that corresponds to the Sup[16].

The results derived from these tests are shown in table 6. The general conclusion that can be extracted is that the Mean statistics for the last part of the sample tend to reject the null of I(1) at the 10% significance level, whereas the null is not rejected for the first part. This indicates that the processes followed by the relevant variables are becoming stationary, and accordingly the debt in GDP percentage points is becoming I(1). The change is taking place in a gradual form and begins between the late 80s and the early 90s. This result has a direct economic interpretation since the first sample period, which covers from 1964 to the early 90s, was characterised by the implementation of a modern fiscal policy in Spain, which implied the building of the Welfare State and a new tax system, moving towards European models, that required a rapid expansion of public revenues and expenditures. Once this target was achieved and with the additional constraint derived from the commitment to meet the convergence criteria set out in the Treaty on European Union, consolidation issues become more relevant in conducting fiscal policy.

Therefore, the general conclusions drawn from the estimation of [6] and the structural-break tests on the cointegrating relationship must be questioned because the variables in [6] are not always I(1), and, consequently, the cointegrating analysis may lose sense at least for the whole sample period. Furthermore, our results show that in recent years the “sustainability in the weak sense” seems to be changing into “sustainability in a strong sense”, according to Quintos’ terminology, and no future problems in marketing public debt are expected to arise as far as this trend is confirmed.

[16] See appendix E for technicalities.

Table 6: Partial integration tests[17]

Statistic	Δb_t	t_t	g_t^R	d_t	Critical values[18] 10%	5%	1%
Supt$_{d\ 1}$	-2.35	-0.59	-1.26	-1.89	-3.44	-3.76	-4.44
Year	1993	1974	1988	1976			
Meant$_{d\ 1}$	-0.97	0.98	0.25	-1.21	-2.37	-2.46	-3.23
Supt$_{d\ 2}$	-3.89*	-3.38	-3.59	-3.33	-3.80	-4.12	-4.76
Year	1987	1991	1988	1980			
Meant$_{d\ 2}$	-2.96**	-2.42*	-2.44*	-2.40*	-2.36	-2.59	-3.07
Supt$_{a\ 1}$	-2.35	-0.68	-1.23	-1.67	-3.18	-3.48	-4.12
Year	1968	1970	1968	1990			
Meant$_{a\ 1}$	-0.73	0.62	0.07	-0.68	-2.09	-2.35	-2.88
Supt$_{a\ 2}$	-3.71*	-3.55	-3.37	-3.22	-3.60	-3.91	-4.52
Year	1987	1991	1988	1980			
Meant$_{a\ 2}$	-2.80***	-2.09*	-2.23*	-2.11*	-2.09	-2.28	-2.65

Note: (*), (**) and (***) mean rejection of the null at the 10%, 5% and 1% significance levels, respectively.

However, any conclusion to be derived from these results should take into account the limitations of the analysis, in particular the fact that this is based on past data. Therefore, under this approach, the sustainability of the Spanish fiscal policy should be interpreted in the sense that no problems in marketing public debt are expected to arise if the variables involved follow the pattern of the past in the future.

4. CAUSALITY ANALYSIS

Given the limitations of the sustainability analysis already mentioned, the fact that the change from weak sustainability to strong sustainability is not fully confirmed for a large sample period and that, even if this is the case, the Stability and Growth Pact for the EMU countries sets explicit ceilings to public finance that might require additional consolidation efforts in Spain, we consider the issue of which is the most efficient strategy to achieve permanent reductions in fiscal deficits. Thus, we analyse the possible interdependence between government expenditure and revenues and the issue of whether higher taxed lead to expenditure changes or whether expenditure growth leads budget dynamics, with taxes following suit.

4.1. Main theoretical hypotheses

The effect of an unanticipated change in public spending or revenue on the size of the budget and the budget deficit depends in a complex way on the characteristics of the tax system and the way in which the political system articulates spending demands and redistribution objectives. Economic theory on the relationship between public spending and revenue offers explanations for all possible directions of causality between public spending and revenue.

Independence of public spending and revenue is consistent with the Ricardian equivalence theorem (Barro, 1974). Given an exogenously determined time path for public spending, there exist an infinite number of distributions of the tax burden over time that satisfy the government's intertemporal budget constraint (GIBC). By postulating exogenous government behaviour with respect to spending decisions, current tax changes merely entail future changes in revenue with the same present

[17] The programs for performing these tests have also been provided by J. L. Fernández.
[18] Taken from Fernandez (1999).

value and the opposite sign. Moreover, an increase in public spending today may be financed by changes in current taxes or by debt (future tax changes). Accordingly, we should not expect to find a stable and significant contemporaneous relationship between public spending and revenue. In a Ricardian economy, any plan to control the deficit is, a priori, equally effective as long as it is consistent with the GIBC.

If some endogeneity is admitted in the behaviour of government, the fulfilment of the GIBC gives support to the hypothesis that spending determines revenue. Applying conventional results of optimal tax theory, Barro's "tax smoothing" theory (1979) predicts that an unanticipated increase in the public spending/output ratio will be followed by a rise in the public revenue/output ratio, that will be achieved by choosing a constant rate of tax[19] that minimises the cost of raising revenue over time, by the amount necessary to balance the GIBC.

In its basic version, the "tax smoothing" hypothesis predicts that higher spending generates higher taxes (and a larger temporary deficit). This pattern of behaviour is also characteristic of other traditional explanations of public economics. Peacock and Wiseman (1979) argue that increases in spending associated with situations of social crisis or war can force a change in attitude on the part of citizens with respect to the "tolerable" tax burden, which gradually rises (shift effect). Following the crisis, part of the initial shift is consolidated as permanent, with the emergence of new spending demands (inspection effect). There are also other more general explanations associated with the institutional features of the budgetary process. Revenue and spending programmes involve different time horizons. If consumers are partly non-Ricardian and politicians discount the future, the political pressures for higher public spending - exerted during the stages of preparation (interest groups), approval (lobbying) and implementation (bureaucracy) of the budget - will tend to dominate budget dynamics. The resulting increase in taxes (the deficit) will be all the smaller (larger) the more lax are the budget rules (von Hagen, 1992) and the more fragmented is political power (Roubini and Sachs, 1989). When the direction of causality goes from spending to revenue, as it does under this behavioural hypothesis, control of the deficit can be achieved through unanticipated increases in the tax burden or strict (legal) limits on the level of public spending.

There are two different public finance traditions, both with great political prestige, according to which, over time, taxes determine spending. First we have the "government leviathan" hypothesis (Brennan and Buchanan, 1980): unless constitutional limits are imposed on the expansion of government, in post-constitutional periods the latter will attempt to maximise the revenue from any source of tax that is constitutionally within its reach. This rule of behaviour enables a minimum level of specific goods and services demanded by citizens to be financed, while maximising the amount of resources that the government can use at its own discretion. The "tax and spend" hypothesis, which can count Friedman (1978) and Gramlich (1989) among its proponents[20], predicts that as technological advances increase the capacity to raise revenue and as economic activity becomes increasingly focussed on the market, the tax burden will increase and, consequently, so will government spending[21].

---

[19] The constancy of the tax burden is not a general result. Barro (1979) obtains it by assuming that the function of revenue-raising costs is homogeneous to the first degree in total taxes and output. Aschauer and Greenwood (1985) derive the same result postulating intertemporally separable preferences for consumption and leisure. Without restrictions on preferences the tax rate must change over time.

[20] This view of government is characteristic of European and North-American political liberalism. In the US, with its well-known terminological peculiarity, the economic programmes of the Republican Party regularly define their position on fiscal and budgetary policy as being the antithesis of that of the "tax and spend liberals" of the Democratic Party.

[21] Ward (1982) and Kau and Rubin (1981) have found empirical support for this prediction.

In this scenario, increases in the tax burden will only reduce the budget deficit in the short term, but may increase it in future if there are no restrictions on the issuance of public debt. The control of the deficit should be based on the establishment of legal restrictions on the deficit or on borrowing, combined with discretionary cuts in spending. However, in accordance with a second behavioural hypothesis, the restrictive effect of taxes may be partly offset by the presence of fiscal illusion. If taxpayers believe that the "price" of public services coincides with the taxes currently paid, increases in the budget deficit will prompt greater spending demands (Buchanan and Wagner, 1977). An increase in the tax burden will be perceived in this case as an increase in the cost of public benefits, which will reduce the social demands for public spending.

Interdependence of spending and taxes is the fourth possibility. When political decisions on spending and taxes are taken and endured by the same group, they will be joint and interdependent. This description of the budgetary process originates in Wicksell (1896) and has been reformulated by Musgrave (1966) and Meltzer and Richard (1981). The result of interdependence is also consistent with "tax smoothing" theory: when the political process which decides the path of spending takes into account the marginal cost of raising and administering taxes, spending and taxes will be determined together (Blanchard and Fisher, 1989). When public spending and revenue are interdependent budget discipline requires action to be taken on both sides of the budget simultaneously.

The empirical literature on dynamic interdependence of public spending and revenue, mostly referring to the US economy, offers a great variety of results. Shibata and Kimura (1986) cannot reject the null hypothesis of absence of causality between spending and revenue. Von Furstemburg et al. (1986) find decisive evidence of causality, although the "spending causes taxes" direction seems dominant. In Anderson et al. (1986) spending causes, in the Granger sense, taxes. Causality in the opposite direction is the main result of Blackley (1986). Ram (1988a) distinguishes between the federal and state levels of government, identifying two opposite patterns of behaviour: spending causes taxes at the state level, while the opposite is the case at the federal level. Owoke (1995), Miller and Russek (1990) and Manage and Marlow (1986) identify a pattern of causality in both directions. Finally, in a recent work, Payne (1998) examines the relationship between revenues and expenditures for USA and finds that the argument that tax cause expending is supported for 24 states, while the opposite conclusion is supported for 8 estates and a bidirectional causality is valid for 11 estates.

The evidence available for other countries is also mixed. Ram (1988b) finds an absence of causality most of the time. Joulfaian and Mookerjee (1991) analyse the interdependence of spending and taxes in 22 OECD countries, using first differences of autoregressive vectors (VAR). In 11 cases no causal relationship is detected, in 8 spending causes taxes, in 2 the "tax and spend" hypothesis is confirmed, and in 1 causality is in both directions. In the case of Spain, this study indicates that spending causes revenue when this pair of variables is considered alone. However, when the cyclical position of the economy and inflation are controlled for, the result is independence. Raymond and González-Páramo (1988) find weak evidence of causality, in the Granger sense, running from taxes to the level of public spending, whereas González-Páramo (1994) finds strong support to the tax and spend hypothesis for the sample period 1955-1991. Finally, Belessiotis (1995) examines this issue for the European Community Member States and finds that there is interdependence between fiscal variables for 12 countries, while only in the case of UK, the data suggests independence. In 2 cases, this study finds that government expenditure causes revenues; in 5 cases, causality is found from revenues to expenditure and, in 6 cases (including Spain), bi-directional causality is found.

4.2. Preliminary causality results

The most generalized hypothesis about the behaviour of the Spanish public sector assigns a prominent role to public expenditure. The oil crises, the change of the political system in the mid seventies and the implementation of the Welfare System have conditioned the evolution of this variable, followed with a certain delay by tax reforms and changes in the fiscal pressure. This hypothesis seems to be supported by the pattern followed by spending and revenues in figure 1 above. In order to provide some empirical support to this statement we look at Granger causality tests (see Granger, 1969), which is based in the notion that lagged realizations of one variable x help to predict current values of another w. More precisely, x G-causes w if the forecast error variance of w is significantly lower when lagged realizations of x are in the information set than when they are not. Therefore, departing from the model

$$w_t = d_0 + \sum_{i=1}^{k} d_i w_{t-i} + \sum_{i=1}^{k} m_i x_{t-i} + e_t$$

where e_t is white noise, the null hypothesis that x does not G-cause w can be contrasted with usual tests of the joint restriction $m_1=\ldots=m_k=0$. In principle, these tests are appropriate when the variables involved are stationary and the process is not misspecified. Omission of relevant variables may lead to incorrectly detect directions of causality or even discover causality when it does not really exists, yielding to spurious results (Granger and Newbold, 1986). Hence, if a shock in variable u generates a reduction in x in period t and a further reduction in w in period t+1, the exclusion of u in the information set could produce the spurious result that x causes w.

Theories of behaviour of expenditure and revenues point to GDP and prices as relevant variables to include in such an analysis. The "Wagner's Law" associates the level of the public expenditure to the degree of economic development, which can be approximated by real income. Moreover, Musgrave relates the tax collecting capacity with technological development and the degree of monetisation of transactions, factors that are also positively correlated with real income[22]. Furthermore, revenues and expenditures respond automatically to the cycle as automatic stabilizers. On the other hand, Baumol states that the relative cost of the public output is increasing in time due to the productivity differential between private and public sector. If this differential is stable, increases in prices could help to explain increases in public expenditure. Finally, taxes and spending also respond to the inflation rate, due to the existence of indexation clauses in many spending programs or the real increase in tax collections for the absence of automatic indexation of tables in personal taxes. This phenomenon is known as *fiscal drag*.

In order to perform a preliminary causality analysis we differenced the variables in order to achieve stationarity (see table A.1 in appendix A) and we considered several possibilities as to the inclusion of the GDP growth rate and the inflation rate. The results are shown in table 7 and offer very weak evidence that, if any, revenues G-cause expenditures. This evidence only appears when using two lags in the analysis and $\Delta^2 y_t$ and $\Delta^2 p_t$ are included in the regressions. On the contrary, the null that

[22] A brief survey of theories explaining public expenditure growth can be found in González-Páramo and Raymond (1988). As regards to the evolution of public revenues, the classical reference is Musgrave (1969).

expenditures G-cause revenues finds no empirical support so far. This finding contradicts the generally accepted explanation for the behaviour of the public sector stated above. However, these results may be biased in that we have not consider so far the possibility of the existence of long-run relationships among the variables involved. We will see in next section that taking into account this possibility changes the conclusions sharply and help to discover relevant directions of causality that constitute the basis of the conclusions of this paper.

Table 7. Granger causality tests

Null hypothesis	Without $\Delta^2 y_t$, $\Delta^2 p_t$		With $\Delta^2 y_t$, $\Delta^2 p_t$	
	k=1	k=2	k=1	k=2
t_t does not G-cause g_t^R	0.72	1.95	0.38	5.33*
g_t^R does not G-cause t_t	0.73	0.27	1.05	0.52

Note: *, ** and *** denote rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

4.3. Long run analysis and causality

Cointegration analysis

The analysis of the existence of cointegration between revenues and expenditures is important for two reasons: First, because the cointegration relationships must be included in causality tests. The Granger-causality analysis performed in last section can be misleading if it contains a specification error. If the variables involved are linked in the long term by cointegration relationships the estimated cointegrating vectors should be included in the specification of the VAR used to test Granger-causality (Granger, 1988). Thus, the inclusion of the so called “equilibrium residuals” may modify the sense of the causality on the one hand, and may give us useful information by distinguishing between long term and short term causality. Second, because the estimated parameters may indicate the existence of a possible long-term relationship between deficit and public sector’s size. Thus, a coefficient affecting revenues in the cointegrating vectors equal to –1 implies that deficit is independent of fiscal pressure or the size of the public sector. Therefore, a long-lasting reduction of deficit would require altering its generating process by reducing its structural component or by altering the elasticity to the cycle or inflation. On the other hand, if this coefficient is greater than one in absolute value, fiscal consolidation could be achieved with a reduction in the public sector’s size. Without altering the generating process, that is, without structural breaks, consolidation should be leaned on expenditures, revenues, or both, depending on the causality structure.

In table 8 we present the estimation of several models in order to test whether there is cointegration[23] among the relevant variables. These models have been

[23] Previous to these results we estimated three models by OLS in which the dependent variable was public expenditure and the regressors were public revenues, growth rate of real output and the inflation rate; public revenues and growth rate of real output and public revenues only. These estimations gave coefficients for the public revenues between 1.23 and 1.27. The ADF test performed on the residuals from the three specifications could, in none of the cases, reject the null hypothesis of absence of cointegration. However, due to the well-known problem of common factor restrictions implied in this way of testing cointegration we considered convenient to proceed in a different way. Nevertheless, the coefficients for the

estimated using the maximum likelihood procedure suggested by Johansen (1988), Johansen and Juselius (1990) and Johansen (1991). The critical values have been taken from Osterwald-Lenum (1992). LR tests were performed in order to contrast several hypothesis that could give us clues to discern about different specifications.

Model [11] is estimated with an unrestricted constant, which implies that the model includes a linear deterministic trend in levels. Although this assumption can be controversial, given that we are using public revenues and expenditures as percentages of GDP and it is not expected that in the long run this trend will be present, it fits very well with the sample period we have. Thus, the specification of the model is

$$\Delta X_t = m_0 + a\,b'(X_{t-1},1) + \sum_{i=1}^{2} \Gamma_i \Delta X_{t-i} + e_t \;\; ; \;\; X_t = (g_t^R, t_t, \Delta y_t, \Delta p_t) \qquad [11]$$

We find two cointegrating vectors according to the Trace and LR_{max} statistics. The null hypothesis of weak exogeneity of the variables is rejected except for the case of Δp_t (χ_2^2=3.15), although the joint hypothesis of weak exogeneity and long run exclusion of this variable is rejected at the 1% significance level (χ_4^2=21.67). It is worth to notice that the public revenues' coefficients in the cointegrating vectors are always greater than one in absolute value. The null hypothesis of this coefficient equal to –1 is rejected, according to the LR test that appears in the table, at the 1% significance level.

Given the weak exogeneity of Δp_t the model was re-specified including $\Delta^2 p_t$ in the VAR. This corresponds to model [12].

$$\Delta Z_t = m_0 + a\,b'(X_{t-1},1) + \Gamma_0 \Delta^2 p_t + \sum_{i=1}^{2} \Gamma_i \Delta Z_{t-i} + e_t \;\; ; \;\; Z_t = (g_t^R, t_t, \Delta y_t) \qquad [12]$$

According to the Trace statistic we find three cointegrating vectors. In all of them the coefficient affecting revenues is greater than one, and in the two first vectors the value of this coefficient is almost equal than in model [11], so the change of the dynamic structure does not affect the long run relationship among the variables significantly. The null of the revenues' coefficient equal to –1 is also rejected at the 1% significance level. In this second model we cannot reject the null that g_t^R is weakly exogenous (χ_4^2=7.16), although the value of this test is very close to significance. For the rest of the variables the null of weak exogeneity is rejected. Furthermore, the long run exclusion of Δp_t was also rejected at the 1% significance level (χ_3^2=14.23). Therefore, we decided not to re-specify the model again.

As stated above, the inclusion of a linear deterministic trend in the data may be a controversial hypothesis from the economic point of view when using variables as percentages of GDP. Accordingly, we tried a similar specification to model [11] but with the constant restricted. This is what we call model [13], and the whole specification is

$$\Delta X_t = a\,b'(X_{t-1},1) + \sum_{i=1}^{2} \Gamma_i \Delta X_{t-i} + e_t \;\; ; \quad X_t = (g_t^R, t_t, \Delta y_t, \Delta p_t) \qquad [13]$$

In all cases but for Δp_t (χ_2^2=2.72) the null of weak exogeneity was clearly rejected. The joint hypothesis of weak exogeneity and long run exclusion of this

public revenues can be informative about the long run relationship between expenditures and revenues. This coefficient seems in all cases to be greater than one.

variable is rejected at the 1% significance level (χ_4^2=17.82). In this model the coefficient affecting revenues was, as in the other cases, greater than one. As the table shows, the hypothesis that this coefficient equals –1 is also rejected at the 1% significance level. For all the models above, tests of lags reduction were performed and were significant in all cases[24].

According to our results, a bias towards deficit seems to exist in the public sector's size. This bias arises because the coefficient for t_t in the cointegrating vectors is greater than one in absolute value. This implies, as stated before, that fiscal consolidation could be achieved with a reduction in the public sector size. As regards the most adequate strategy of fiscal consolidation to be selected to reach this target, the analysis of the direction of causality between the fiscal variables can offer a relevant guideline.

Table 8. Cointegration results (Johansen tests)

	$H_0(r)$	LR_{max}	*Trace*	g^R	Cointegrating vectors t	Δy	Δp	*Intercept*
Model [11]	r=0	24.90*	55.87***	1	-1.20	0.46	0.39	
	r≤1	22.78**	30.97**	1	-1.52	-1.00	-0.45	
	r≤2	5.48	8.19					
	r≤3	2.71*	2.71*			χ_2^2=17.01***		
Model [12]	r=0	22.27**	38.26***	1	-1.50	-0.72	-0.26	
	r≤1	11.69	15.99**	1	-1.21	0.39	0.26	
	r≤2	4.30**	4.30**	1	-1.05	1.06	-0.02	
						χ_3^2=18.67***		
Model [13]	r=0	28.75**	66.69***	1	-1.11	0.87	0.55	-0.05
	r≤1	22.99**	37.39**	1	-1.45	-0.65	-0.24	0.16
	r≤2	10.66	14.95					
	r≤3	4.29	4.29			χ_2^2=12.51***		

Note: *, ** and *** denote rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

Causality analysis with cointegrated variables

As regards to the estimations in table 8, we performed Granger-causality tests including the cointegrating vectors. Weakly exogenous variables such as Δp_t in models [11] and [13] were included as $\Delta^2 p_t$ in the VAR. The high degree of correlation between the residuals from the expenditures' and revenues' equations leaded us to estimate these equations by both OLS and SURE, which is a more efficient method than OLS. Nevertheless, both procedures yield similar results. In this context, short-term causality is understood as a situation in which lagged changes in one variable have predictive power in current changes in another one, whereas long-term causality is detected when the lagged level of one variable (equilibrium residuals) explains current changes in another variable. These chi-squared tests are shown in table 9 and have been performed with one and two lags (the number of lags is *k*).

According to table 9 results, we observe a long-run direction of causality from public expenditures to revenues, which is more evident when we introduce two lags in the equations. On the other hand, the inclusion of the cointegrating vectors from models [11] and [12] also show evidence of long-run causality from revenues to

[24] F(16,46)=2.02**, F(9,41)=2.16** and F(16,46)=2.02** for models [1], [2] and [3], respectively.

expenditures, which is stronger with one, rather than two, lags. Lastly, revenues-to-expenditure long-run causality is not detected in model [13] with two lags, whereas with one lag this causality seems to be present. However, the estimated models of cointegration seemed to support the inclusion of the two lags (see footnote 17). Contrary to the long-run case, short-run causality seems to hold only in the expenditures-to-revenues direction when we specify the model with two lags, and this evidence is only significant at the 10% significance level.

Table 9. Granger causality tests with cointegrated variables

Null hypothesis	OLS estimates				SURE estimates			
	k=1		k=2		k=1		k=2	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
t_t does not G-cause g_t^R								
[11]	0.02	7.34**	2.37	5.17*	0.03	9.40***	3.50	7.63**
[12]	0.00	9.18**	2.57	7.16*	0.00	9.51**	2.59	7.47*
[13]	0.35	6.09**	1.87	3.51	0.35	6.18**	1.88	3.55
g_t^R does not G-cause t_t								
[11]	0.40	5.04*	5.16*	9.41***	0.17	6.46**	7.62**	13.90***
[12]	0.04	6.73*	4.75*	11.76***	0.04	6.80*	4.79*	12.15***
[13]	0.04	15.34***	4.83*	17.48***	0.04	15.93***	4.86*	18.15***

Note: *, ** and *** denote rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

In sum, our tests seem to show clear evidence of public expenditures G-causing public revenues in the long run and also, although less clear, evidence of long-run G-causality from revenues to public expenditure. In other words, we obtain evidence of long-run bidirectional causality. Furthermore, in the short run the direction of causality seems to hold only from expenditures to public revenues. These results, together with the bias to deficit in the size of the public sector that arises from coefficients affecting revenues greater than one in absolute value in the cointegrating vectors, tend to support the idea that efficient fiscal consolidation could be attained by paying special attention to public expenditures. This variable seems to play a crucial role, at least in a first step, given its importance in explaining the process followed by tax collections.

Variance decomposition: VAR in levels

Another tentative way of characterising the interdependence between the most relevant variables in our analysis, public expenditures and revenues, is by means of the variance decomposition and impulse response functions in a context of a VAR analysis. Variance decomposition functions indicate which part of the variance of the forecast error in one variable can be attributed to innovations in another after some periods. Accordingly, this decomposition can be used to approximate the contribution of every variable to variability of the whole system. If the errors in a VAR system are contemporaneously correlated (in our case the correlation is about 0.56), it is not possible to attribute the shocks to a single variable, making it necessary to identify the model. We shall use a simple recursive system based on the Cholesky decomposition. When this identification scheme is used, the order of the equations in the VAR is not a trivial issue, and the impulse response functions and variance decomposition may show different results depending on the order. Thus, we offer the estimations with the order (g^R,t), which means that the common component in the error term is totally attributed to the expenditures, whereas a shock to revenues only affect this variable contemporaneously, and the order (t,g^R), which does the opposite. In order to check whether our conclusions are modified by the introduction of the GDP growth rate, we also present the same results for the three-variable VAR[25] with the identification schemes (Δy,g^R,t) and (Δy,t,g^R).

It is convenient to point out that the Cholesky decomposition used to identify the VAR is not the most convenient one, mostly when the causality analysis shows that revenues and spending show bidirectional causality. Thus, other identification schemes would be more accurate. However, our purpose in this section is not analizing in depth the effects of fiscal shocks. Rather, we are only interested in showing that, even with such an extreme simplification, none of both variables behave independently from each other.

According to table 10, the variance of public expenditure in the long term is explained in a non-negligible share by revenues, although this percentage depends on the identification scheme. The percentage in which public expenditure explains the variance of revenues is very sensitive to the order of the variables in the two-variable VAR. With the order (g^R,t), 55.98% of the variance of the forecast error of revenues is explained by expenditure, whereas this percentage falls down to 7.32% when we use the reverse order. The three-variable VAR leads to similar conclusions, in that expenditures' forecast error variance is explained by revenues in percentages between 30.02% and 48.27%, depending on the order. The percentage explained by the growth rate raises up to 40.87%. In the case of revenues, the variance is explained in a 29.36% by the GDP growth rate, whereas the percentage explained by expenditures goes from 24.05% to 1.60%, depending, as above, on the order. Consequently, we do not get evidence of any of both, revenues and spending, behaving independently from the other.

Table 10: Variance decomposition of the forecast error

Percentage of the forecast error in:	Periods	Explained by shocks in:					
		Order g^R, t			Order t, g^R		
		g^R	t		g^R	t	
g^R	5	90.11	9.89		43.00	56.00	
	10	80.21	19.79		31.91	68.09	
	15	76.00	24.00		27.24	72.76	
t	5	50.24	49.76		5.37	94.63	
	10	54.80	45.20		6.96	93.04	
	15	55.98	44.02		7.32	92.68	
		Order Δy, g^R, t			Order Δy, t, g^R		
		g^R	t	Δy	g^R	t	Δy
g^R	5	37.77	12.80	49.43	19.22	31.35	49.43
	10	30.12	26.24	43.64	13.01	43.35	43.64
	15	29.11	30.02	40.87	10.86	48.27	40.87
t	5	25.89	49.22	24.89	2.10	73.01	24.89
	10	24.61	46.54	28.85	1.80	69.35	28.85
	15	24.05	46.59	29.36	1.60	69.04	29.36
Δy	5	3.14	8.85	88.01	8.19	3.80	88.01
	10	3.98	9.70	86.32	9.62	4.06	86.32
	15	4.45	9.75	85.80	9.75	4.45	85.80

[25] The inflation rate has not been included in the analysis because it is weakly exogenous with respect to the rest of the variables in the analysis, as stated in the estimation of models [11] to [13] above. Current realizations of this variable turned out to be non-significant, so they were excluded form the VAR.

Impulse response functions: VAR in levels

Figures 3 and 4 show impulse response functions of public deficit to innovations in revenues and expenditures. Both specifications of the VAR yield very similar results. A large degree of persistence in the shocks is clearly observed in all cases. A negative shock in g_t^R yields to a long lasting surplus that only disappears after eight years or even more when we include only two variables in the VAR. Including the GDP growth rate in the VAR modifies a little the picture (Figure 4), in that a negative shock to spending yields to a large initial increase in surplus that after five years is totally offset, turning to deficit, when the common component is attributed to revenues in the identification scheme. This deficit is of very small magnitude and non-significant, and also tends to disappear.

However, an asymmetric response is found when a positive shock is simulated on t_t. A large initial reduction of deficit is observed, that after two or three periods reverts to a deterioration of the public budget, as a consequence of the induced response of spending. Therefore, a positive shock in t_t in the long term yields to responses of greater magnitude in g_t^R than in t_t. This result is consistent with our estimates of the revenues' coefficient in the cointegrating vectors greater than one in absolute value, which can be interpreted as a bias towards deficit of the public sector's size. Moreover, the response of spending that we observe after a shock in revenues in the long term is supported by the long-run causality observed from t_t to g_t^R. Consequently, any deficit reduction strategies based on tax increases are only effective in the short term, and will have adverse effects in the long term.

Figure 3. Impulse response function for public deficit. VAR with two varibles

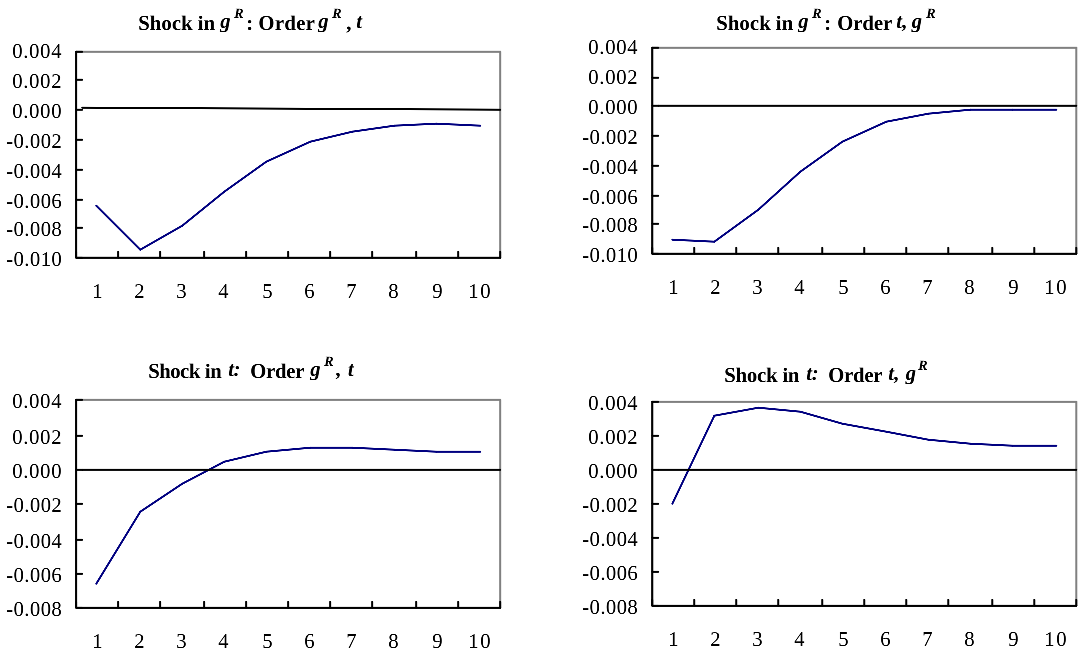
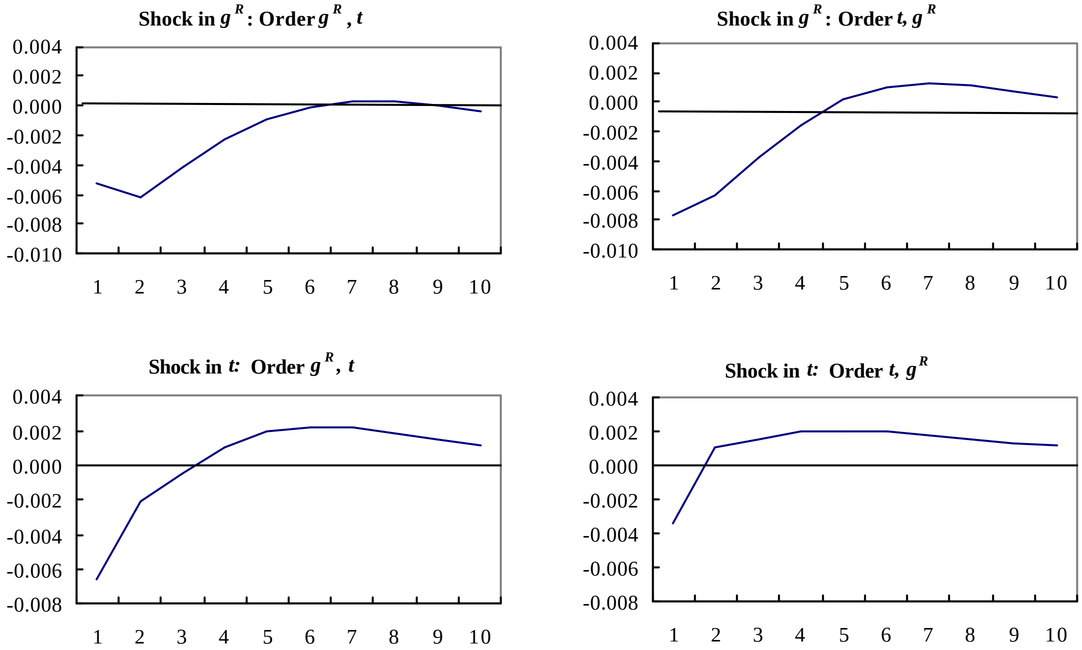


Figure 4. Impulse response function for public deficit. VAR with three varibles



Thus, credible consolidation programs must be based on spending cuts in a first stage, which are efficient both in the short and the long term. The subsequent improvement in deficit realizations should, in a second stage, be supported by tax cuts that would help to maintain the consolidation strategy in the long term by the reduction of the public sector's size. This result provides additional support to the conclusions drawn from the causality with cointegrated variables and from the cointegration analysis itself.

4.4. Restricted sample analysis

According to the results presented in section 3, fiscal sustainability in the "weak sense" seems to be changing into sustainability in the "strong sense" in recent years. Our purpose is now to analyse whether this result is related to changes in the dynamic relationship between revenues and expenditures. More concretely, we check whether the direction of causality obtained above is conditioned to some extent to the more recent realisations of the variables of interest. In order to provide some insights on this issue, we restricted the sample period to finish in 1993, since this year represents the break point for public expenditure in Spain. Accordingly, model [12] was re-estimated yielding the results presented in table 11.

Table 11. Cointegration results (Johansen tests). Variables as % of GDP

Sample 1964-1993	$H_0(r)$	LR_{max}	Trace	Cointegrating vectors g^R	t	Δy	Δp	*Intercept*
Model [12]	r=0	45.97***	71.90***	1	-1.17	1.65	0.17	
	r≤1	18.39**	25.94***	1	-1.64	-1.06	-0.37	
	r≤2	7.55***	7.55***	1	-1.15	0.29	0.17	
						χ_2^2=22.74***		

Note: *, ** and *** denote rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

Table 12. Granger causality tests with cointegrated variables (Sample 1964-1993)

Null hypothesis	OLS estimates				SURE estimates			
	k=1		*k*=2		*k*=1		*k*=2	
	Short run	Long run	Short run	Long run	Short run	Long run	Short run	Long run
t_t does not G-cause g_t^R	5.53**	14.76***	29.38***	41.36***	5.74**	17.06***	34.03***	48.56***
g_t^R does not G-cause t_t	0.01	6.44*	2.76	17.05***	0.01	6.48*	2.76	17.09***

Note: *, ** and *** denote rejection of the null hypothesis at the 10%, 5% and 1% significance levels, respectively.

As in table 8, we find three cointegrating vectors. All the coefficients affecting revenues in the cointegration vector are above 1 in absolute value (the null that these coefficients are –1 is rejected at the 1% significance level), providing additional support to the hypothesis that there is a bias towards deficit in the public sector's size. The null of weak exogeneity is rejected in all cases at the 1% significance level as well.

When we include the cointegration vectors in the Granger-causality analysis we also obtain the *interdependence* result between revenues and expenditures in the long term. However, the short causality tests yield an interesting result. Contrary to our analysis for the whole sample period (table 9), revenues G-cause spending (table 12), confirming previous empirical evidence (González-Páramo, 1994) and supporting the hypothesis that in the last years there has been a change in the dynamic relationship between both variables, as result of the fiscal consolidation strategy implemented, based mainly on a drop in spending (see section 2). Additionally, this fiscal strategy may have contributed to change the direction of causality between revenues and spending and might be behind the change towards strong sustainability[26].

5. CONCLUSIONS

In this paper, we address two issues, the sustainability of public finances and the interdependence between public revenues and expenditures, for the case of Spain. First, we attempt to ascertain whether the current fiscal policy in Spain is sustainable or not, in the sense of the current market value of debt being equal to the discounted sum of expected future surpluses. For this purpose, we apply the traditional tests of sustainability, following Quintos´ approach. In addition, we introduce a deeper univariate analysis of the series. Our findings can be summarised as follows.

As regards the unit root tests applied to the variables used in the analysis, the null hypothesis of the existence of one unit root is accepted, while all the tests reject the null of the existence of two unit roots. Only in the case of the public debt, do the tests not offer conclusive results about the existence of one or two unit roots in the process followed by this variable. Furthermore, cointegration test results show the absence of cointegration between public revenues and expenditures, and given that the condition $0<b<1$ holds, the interpretation has to be that the deficit process is sustainable in its weak form. Moreover, the results from the tests do not support the

[26] The results drawn from the variance decomposition and impulse response functions (not presented here) yield similar conclusions to those obtained for the whole sample.

hypothesis of the existence of a structural break in the long-term relationship between public revenues and expenditures.

The weak sustainability result would imply that, although the transversality condition holds, the government could eventually have problems in marketing its debt, and this involves a risk of rising interest rates in the future. On the hand, that would increase the primary deficit via interest payments. And, on the other, the prospects of future increases in interest rates would cause a slowdown in economic growth and in the government's capacity to generate resources. This second channel also tends to increase future deficits and could ensue in a non-sustainable path, necessitating a fiscal adjustment.

As regards the tests applied to find changes in the order of integration of the series, these indicate that the processes followed by revenues, expenditures and deficit are becoming stationary, and the debt as percentage of GDP is thus turning from I(2) to I(1), making the transversality condition hold in its strong sense, in contrast with the results obtained by Camerero et al. (1998). This change is taking place gradually and starts between the late 80s and early 90s. Therefore, the general conclusion drawn from the cointegration estimation and the structural-break tests on the cointegrating relationship must be questioned because the variables are not first-order integrated for the whole sample period, and the cointegrating analysis may, therefore, be somewhat lacking in meaning for this whole period.

Consequently, our results show that in recent years the "sustainability in the weak sense" seems to be changing into "sustainability in a strong sense", according to Quintos' terminology. This result seems consistent with the evolution of fiscal policy during recent years, characterised by a gradual decline in the deficit as a result of a drop in spending and a slight decline of total revenue over GDP that reversed the previous tendency to imbalance in public finances.

However, any conclusion to be derived from these results should bear in mind the limitations of the analysis, in particular the fact that this is based on past data and, consequently, the future evolution of some structural factors (for example, demographic trends) and their impact on public finances are not taken into account. Therefore, under this approach, the sustainability of the Spanish fiscal policy should be interpreted in the sense that no problems in marketing public debt are expected to arise if the variables involved follow the pattern of the past in the future. Moreover, the well-known lack of power of unit-root tests added to the problem of the small number of observations of the series used in our analysis obliges us to treat the results with the greatest care.

Given the limitations of the sustainability analysis already mentioned, the fact that the change from weak sustainability to strong sustainability is not fully confirmed for a large sample period and that, even if this is the case, the Stability and Growth Pact for the EMU countries sets explicit ceilings to public finance that might require additional consolidation efforts in Spain, we consider the issue of which is the most efficient strategy to achieve permanent reductions in fiscal deficits. Our results are based on causality analysis, variance decomposition and impulse response functions.

Preliminary Granger causality tests, which are based in the notion that lagged realisations of one variable help to predict current values of another variable, offer two main results. First, the deficit is not independent of fiscal pressure or the size of the public sector. Second, we find only very weak evidence, if any, that revenues G-cause expenditures. However, these results may be biased in that they do not consider the possibility of the existence of long-run relationships among the variables involved. Once we consider this possibility and develop the causality analysis with cointegrated variables, we find clear evidence of public expenditures G-causing public revenues in the long run and also, although less clear, evidence of long-run G-causality from revenues to public expenditure. In other words, we obtain evidence of long-run bi-directional causality. Furthermore, in the short run the direction of causality seems to hold only from expenditures to public revenues in the whole sample. These findings for the whole sample period contrast to the conclusions obtained by Joulfaian and Mookerjee (1991), González-Páramo and Raymond (1988) and González-Páramo (1994) for Spain, and provide support to the *interdependence* hypothesis.

The existence of a bias towards deficit in public sector's size indicates that fiscal consolidation could be achieved with a reduction in the size of the public sector. As regards the most adequate strategy to reach this target, the evidence that public expenditure causes revenues in the short and long term, whereas revenues only G-cause expenditures in the long term, implies that a credible and effective fiscal adjustment will require structural public expenditure cuts.

The variance decomposition and impulse response functions analysis in a context of a VAR analysis confirm the latter picture. On the one hand, we get that there is a non-negligible dynamic relationship of dependence between revenues and expenditure. On the other, the results of the impulse response functions confirm the result of the existence of a bias towards deficit in public sector's size and support the idea that efficient fiscal consolidation could be attained by reducing structural public expenditure. This variable seems to play a crucial role, at least in a first step, given its importance in explaining the process followed by tax collections.

Finally, a restricted sample analysis (1964-1993) offers a somewhat different picture, in that we find evidence of causality from revenues to expenditure in the short term. This result confirms previous empirical evidence for similar sample periods (González-Páramo, 1994) and supports the hypothesis that in the last years there has been a change in the dynamic relationship between both variables as result of the fiscal consolidation strategy implemented (based mainly on a drop in spending). Additionally, this fiscal strategy might be behind the observed change towards strong sustainability.

APPENDIX A: UNIT ROOT TESTS

In table A.1 we summarise the unit root tests for the variables used in the analysis. In none of the cases do the tests reject the null hypothesis of the existence of one unit root for the fiscal variables. Since no constant or deterministic trend turned out to be significant for t_t, g_t^R or d_t, the tests reject the null of the existence of two unit roots.

The tests performed on the debt process do not clearly reject the existence of a constant and a deterministic trend. Thus, the ADF test does not offer conclusive results about the existence of one or two unit roots in the process followed by this variable, although it seems to favour the I(2) hypothesis. The Phillips-Perron method offers a different view in that it rejects the null of two unit roots against the alternative of only one unit root. However, given that d_t can be considered as I(1), according to [1], we might expect b_t to be I(2), or accordingly Δb_t to be I(1). The lack of power of unit-root tests, together with the difficulty of distinguishing between a I(2) process and a I(1) process with drift and time trend, prevents us, for the moment, from drawing further conclusions on this issue. Finally, GDP growth rate and inflation rate are, according to the unit root tests, I(1) variables.

Table A.1: Unit root tests

	ADF statistics			Phillips-Perron statistics		
	t_a	t_{a*}	t_{a**}	$Z(t_a)$	$Z(t_{a*})$	$Z(t_{a**})$
I(1) vs. I(0)						
b_t	0.92	-0.26	-2.46	2.07	0.77	-2.54
t_t	3.53	-0.76	-0.99	4.35	-0.78	-1.19
g_t^R	1.38	-1.12	-0.93	2.30	-1.14	-0.39
d_t	-0.84	-1.45	-1.40	-0.82	-1.45	-1.31
Δp_t	-0.76	-0.87	-1.46	-0.92	-1.11	-1.49
Δy_t	-1.70*	-2.61	-2.57	-1.41	-2.52	-2.79
I(2) vs. I(1)						
b_t	-2.26**	-2.59	-2.44	-2.83***	-3.19**	-3.25*
t_t	-1.86*	-3.34**	-3.38*	-3.94***	-6.47***	-6.48***
g_t^R	-1.98**	-2.46	-2.63	-2.97***	-3.73***	-3.92**
d_t	-3.16***	-3.10**	-3.19	-5.25***	-5.18***	-5.27***
Δp_t	-3.44***	-3.42**	-3.55*	-5.96***	-5.90***	-6.08***
Δy_t	-5.95***	-5.84***	-5.76***	-7.24***	-7.17***	-7.26***

Note: The symbols *, ** and *** denote rejection of the null at the 10%, 5%, and 1% significance levels, respectively. The number of lags used has been set to 1.

APPENDIX B: THE GREGORY AND HANSEN TEST (1996)

Gregory and Hansen (1996) propose a statistic that attempts to test the null hypothesis of no cointegration against the alternative of cointegration with a structural break at an unknown point in time, and accordingly consider three possible models. The first one is referred to as “level shift” (C), which is expressed as:

$$y_{1t} = m_1 + m_2 D_{tt} + a y_{2t} + e_t \quad \text{[B.1]}$$

where

$$D_{tt} = \begin{cases} 0 & t \le [tT] \\ 1 & t > [tT] \end{cases} \quad , \quad t \in (0,1)$$

with [ ] meaning “integer part” of the argument inside. Thus, [B.1] aims to test whether there is a change in the intercept term at the time of the shift. The second possibility is called “level shift with trend” (C/T) with the form:

$$y_{1t} = m_1 + m_2 D_{tt} + bt + a y_{2t} + e_t \qquad \text{[B.2]}$$

The last model considered is known as “regime shift” (C/S) and takes the form:

$$y_{1t} = m_1 + m_2 D_{tt} + a_1 y_{2t} + a_2 y_{2t} D_{tt} + e_t \qquad \text{[B.3]}$$

These models are estimated recursively by OLS for all possible break points in the interval $t \in$ [0.15,0.85]. A sequence of ADF and Phillips Z_α and Z_t residual-based tests is computed, and they calculate the highest absolute value of the sequence. The observation associated with this statistic is taken as the most probable break point.

More recently Fernández (1999) has tabulated the distribution for the Mean of the ADF test (MeanADF) which was not originally tabulated by Gregory and Hansen. This latter statistic could be used for testing a gradual change in the policy regime and shows that this statistic has an acceptable power in finite samples. He also shows that the test proposed by Gregory and Hansen has more power than the tests proposed by Hansen (1992) and Hansen and Johansen (1992) (which we shall refer to later) in finite samples for detecting parameter instability in cointegrated relationships, although all of them have lower power the lower the sample period.

APPENDIX C: THE HANSEN TEST (1992)

This test also considers the possibility of a structural break at an unknown point in time, although the null hypothesis is the existence of cointegration, in contrast with the Gregory and Hansen test. Thus, the alternative is the existence of a structural break. He considers the following relationship between the variables:

$$y_t = A_t x_t + u_{1t} \qquad \text{[C.1]}$$

with

$$x_t = (x'_{1t}, x'_{2t})'$$
$$x_{1t} = 1$$
$$x_{2t} = x_{2t-1} + u_{2t}$$

He proposes four tests for instability. The first two are called F_t and SupF for the alternative of a single structural break in A_t, which yields:

$$A_i = \begin{cases} A_1 & i \le t \\ A_2 & i > t \end{cases}$$

where 1<*t*<*n*. These tests for parameter instability are based on the *scores* obtained from the fully-modified residuals in the cointegrating equation and a long-run estimation of the covariance matrix as suggested by Andrews and Monahan (1992), which uses a prewhitened kernel estimator with a plug-in bandwidth.

The test F_t assumes that the break point is known and takes the expression:

$$F_t = trace\{S'_{nt} V_{nt}^{-1} S_{nt} \hat{\Omega}_{1\cdot 2}^{-1}\}$$

and S_{nt} and V_{nt} are

$$S_{nt} = \sum_{i=1}^{t} \left(x_t \hat{u}_{1t}^{+}{}' - \begin{pmatrix} 0 \\ \hat{\Lambda}_{21}^{+} \end{pmatrix} \right)$$

$$V_{nt} = M_{nt} - M_{nt} M_{nn}^{-1} M_{nt}$$

and

$$M_{nt} = \sum_{i=1}^{t} x_i x_i'$$

where $\hat{u}_{1t}^{+}{}'$ are the fully-modified residuals from the estimation of [C.1], corrected by the endogeneity bias of the regressors and $\hat{\Omega}_{1\cdot 2}$ is semi-parametric estimation of the long term variance of u_{1t} conditioned to u_{2t} as suggested by Andrews and Monahan.

$$\Omega = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} E(u_j u_t')$$

Under the null of stationarity of A_t this contrast follows a χ^2 with degrees of freedom equal to the number of cointegrating vectors. This test is similar to the Chow test, but can only be used when t can be chosen independently of the sample size, and thus has a low power. Therefore, when the break point is unknown Hansen proposes the statistic

$$SupF = \sup_{t/n \in x} F_{nt}$$

where x is a compact subset of the interval (0,1). Hansen suggests considering the F_{nt} statistics in the interval x = [0.15 , 0.85] in order to avoid distortions induced by break points close to the first and final observations. The observation associated with *SupF*, *NsupF*, can be interpreted as an indicator of the possible break point. The *SupF* statistic has power against a swift shift in regime. On the other hand, when the parameter shifts gradually over time, when A_t follows a martingale process, Hansen suggests the *MeanF* statistic, which takes the form:

$$MeanF = \frac{1}{n^*} \sum_{t/n \in x} F_{nt} \quad \text{, where } n^* = \sum_{t/n \in x} 1.$$

The last test posed by Hansen is a LM one called L_c, which is appropriate when the likelihood of parameter variation is relatively constant throughout the sample, and takes the form:

$$L_c = trace\left\{M_{nn}^{-1}\sum_{t=1}^{n} S_t \hat{\Omega}_{1\cdot 2}^{-1} S_t'\right\}$$

This test does not require specification of an interval for t a and can be taken as a cointegration test with the null of existence of cointegration.

APPENDIX D: THE HANSEN AND JOHANSEN TEST (1993)

The Hansen and Johansen test is a recursive test that can be applied to the maximum-likelihood method proposed by Johansen (1988,1991) for estimating cointegrating vectors. It examines the stability of the eigenvalues associated with the Error Correction Model, which measures the correlation between the vector of variables in levels and in first differences.

A vector with p I(1) variables, whose dynamics are defined by a VAR, has the form:

$$\Delta X_t = a\,b'\,X_{t-1} + \Lambda Z_t + e_t, \qquad \text{with } t=1,\ldots T \qquad \text{[D.1]}$$

where

$$Z_t = (\Delta X_{t-1},\ldots,\Delta X_{t-k+1}, D_t, 1)'$$

$$\Lambda = (\Lambda_1,\ldots,\Lambda_{k-1},\Psi)'$$

D_t is a set of seasonal dummies, b is the cointegrating vector and a a vector of adjustment coefficients for transitory deviations from the long-term relationship. Regressing DX_t and X_{t-1} over Z_t the residuals R_{0t} and R_{1t} are obtained. Using these residuals the matrixes of moments and the eigenvalues are obtained:

$$S_{ij} = \sum R_{it} R_{jt}'$$

$1 > \hat{l}_1 > \ldots > \hat{l}_p > 0$ and the corresponding eigenvectors $\hat{V} = (\hat{v}_1,\ldots,\hat{v}_p)$

by solving the equation

$$|\, l S_{11} - S_{10} S_{00}^{-1} S_{01} \,| = 0 \qquad \text{[D.2]}$$

These eigenvalues and eigenvectors yield the estimation of b and the range of the matrix of cointegrating vectors, r. They propose the following statistic:

$$HJ(t) = t\sum_{i=1}^{r} \ln\left|\frac{1-\hat{r}_i(t)}{1-\hat{l}_i(t)}\right|$$

where $\hat{l}_i(t)$ are the eigenvalues without restrictions obtained from [D.2] for the subsample 1,…,t, while $\hat{r}_i(t)$ are the eigenvalues obtained for the same subsample according to:

$$| rb'S_{11}(t)b - b'S_{10(t)}S_{00}^{-1}(t)S_{01}(t)b | = 0$$

or in other words, imposing the restriction that the cointegrating-vectors matrix in the subsample 1,…t equals *b*, the cointegrating-vectors matrix for the whole sample. For every possible break point, the *HJ* statistic is a LR test that compares the eigenvalues obtained with and without restrictions, and follows a χ^2 distribution with (*p*-*r*)*r* degrees of freedom. As t approaches the end of the sample, the statistic converges to 0, so it is expected that its asymptotic power is greater for structural breaks in the beginning of the sample.

In their paper Hansen and Johansen do not tabulate the empirical distribution associated with these statistics, but they have recently been obtained by Fernández (1999).

APPENDIX E: TESTS ON CHANGES IN THE ORDER OF INTEGRATION

Earlier work on this issue has been done by Leybourne, McCabe and Tremayne (1996) and Maeso (1997). The former paper tests the null of I(1) with invariant coefficient against the alternative of random coefficient. The latter tests the same null against the alternative of a constant coefficient with a different level since a given date, using rolling regressions. Here we follow Fernández (1999) and use a sequential procedure consisting of estimating the following set of equations:

$$\Delta y_t = m + d_1 D_{tt} y_{t-1} + d_2(1 - D_{tt}) y_{t-1} + \sum_{i=1}^{p} g_i \Delta y_{t-i} + e_t \qquad [E.1]$$

$$\Delta y_t = m + a_1 D_{tt} y_{t-1} + \sum_{i=1}^{p} g_i \Delta y_{t-i} + e_t \qquad [E.2]$$

$$\Delta y_t = m + a_2(1 - D_{tt}) y_{t-1} + \sum_{i=1}^{p} g_i \Delta y_{t-i} + e_t \qquad [E.3]$$

where D_{tt} has been defined above. [E.1] aims to test the existence of one unit root in both subsamples simultaneously. On the contrary, [E.2] and [E.3] impose I(1) in one part of the sample. Accordingly, a sequence for the pseudo-t-ratios t_{d1}, t_{d2}, t_{a1} and t_{a2} associated with the coefficients for every possible break point is obtained. From every sequence two summary statistics are calculated: the mean and the lowest one. Following Fernandez these will be referred to as $\mathrm{Supt}_{\delta 1}$, $\mathrm{Meant}_{\delta 1}$, $\mathrm{Supt}_{\delta 2}$, $\mathrm{Meant}_{\delta 2}$, $\mathrm{Supt}_{\alpha 1}$, $\mathrm{Meant}_{\alpha 1}$, $\mathrm{Supt}_{\alpha 2}$ and $\mathrm{Meant}_{\alpha 2}$. As before, the Sup statistics have power for a unique break point, whereas the mean ones have power for gradual changes. According to Zivot and Andrews (1992), the break point is associated with the observation that corresponds to the Sup.

Fernández (1999) shows that the standard ADF behaves badly with changes in the order of integration, and his results are thus summarised:

- When the non-stationarity appears in the second part of the sample and there is no stochastic trend, the statistics that show most power are $Supt_{\delta 1}$ and $Meant_{\delta 1}$, followed by $Supt_{\alpha 1}$ and $Meant_{\alpha 1.}$
- When there is a stochastic trend and the first part of the sample is I(0), the $Supt_{\delta 1}$ and $Supt_{\alpha 1}$ become significant.
- If the second part is I(0) and we cannot reject a stochastic trend then the $Supt_{\delta 2}$, $Meant_{\delta 2}$, $Supt_{\alpha 2}$ and $Meant_{\alpha 2}$ are expected to be significant.

REFERENCES

Anderson, W., M. Wallace and J. Warner, 1986, "Government Spending and Taxation: What Causes What?", *Southern Economic Journal*, January, 630-639.

Andrews, D.W.K. and J.C. Monahan, 1992, "An Improved Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimator" *Econometrica* 60, 953-966.

Argimón, I., A.L. Gómez, P. Hernández de Cos and F. Martí, 1999, "El Sector de las Administraciones Públicas en España", Estudios Económicos nº 68, Servicio de Estudios, Banco de España.

Aschauer, A.J. and J. Greenwood, 1985, "Macroeconomic Effects of Fiscal Policy", in *Carnegie-Rochester Series on Public Policy* 23, K. Brunner and A. Meltzer (eds.), 138-154.

Banerjee, A., R. Lumsdaine and J. Stock, 1992, "Recursive and Sequential Tests of the Unit-Root and Trend-Break Hypothesis: Theory and International Evidence", *Journal of Business & Economic Statistics* 10, 271-287.

Barro, J.R., 1974, "Are Government Bonds Net Wealth?", *Journal of Political Economy* 82, 1095-1117.

Barro, J.R., 1979, "On the Determination of Public Debt", *Journal of Political Economy* 87, 940-971.

Belessiotis, T., 1995, "Fiscal Revenues and Expenditure in the Community. Granger-Causality Among Fiscal Variables in Thirteen Member States and Implications for Fiscal Adjustment", *Economic Papers* 114, European Commission.

Blackley, P., 1986, "Causality Between Revenues and Expenditures and the Size of the Federal Budget", *Public Finance Quarterly* 14, 139-156.

Blanchard, O.J. and S. Fischer, 1989, *Lectures on Macroeconomics*. Cambridge, MA: MIT Press.

Brennan, G. and J. Buchanan, 1980, "The Power to Tax: Analytical Foundations of the Fiscal Constitution", *Cambridge University Press*, Cambridge, Massachusetts.

Buchanan, J. and R. Wagner, 1977, 'Democracy in Deficit", Academic Press, New York.

Camarero, M., V. Esteve and R. Tamarit, 1998, "Cambio de Régimen y Sostenibilidad a Largo Plazo de la Política Fiscal: El Caso de España", *Instituto Valenciano de Investigaciones Económicas, Working Paper Series*, WP-EC 98-15.

De Castro, F., and R. Peruga, 1999, "Weak Conditions for Deficit and Debt Sustainability", *mimeo*.

De Castro, F., and P. Hernandez de Cos, 2000, "On the Sustainability of the Spanish Public Budget Performance", Public Finance Workshop on Fiscal Sustainability, Research Department, Banca D'Italia.

Dickey, D. and W. Fuller, 1981, "Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root", *Econometrica* 49, 1057-1072.

Dolado, J. and J. Viñals, "Macroeconomic Policy, External Targets and Constraints: The Case of Spain", in *External constraints on macroeconomic policy: The european experience*, Alogoskufis, G, Papademos, L and Porter, R (editors), Cambridge University Press, 1991.

Engle, R. and C. Granger, 1987, "Cointegration and Error Correction: Representation, Estimation and Testing", *Econometrica* 55, 251-276.

Fernández, J.L., 1999, "Efecto de los Cambios Estructurales en el Análisis de Series Económicas no Estacionarias", Ph. Dissertation, Universidad Complutense de Madrid.

Friedman, M., 1978, "The Limitations of Tax Limitations", *Policy Review*, summer, 7-14.

González-Páramo, J.M. and J.L. Raymond, 1988, "Déficit, Impuestos y Crecimiento del Gasto Público", *Papeles de Economía Española* 33.

González-Páramo, J.M., 1994, “Tax and Spend or Spend and Tax? Causality Between Revenues and Expenditure and the Size of the Public Budget in Spain, 1955-91”, Towards Greater Fiscal Discipline, *European Economy* 3, European Commission.

Gramlich, F., 1989, “Budget Deficits and National Savings: Are Politicians Endogenous?”, *Journal of Economic Perspectives* 3, 23-35.

Granger, C.W., 1969, “Investigating Casual Relations by Econometric Models and Cross-Spectral Methods”, *Econometrica* 37, 424-438.

Granger, C.W. and P. Newbold, 1986, *Forecasting Economic Time-Series.* Academic Press, Orlando, Florida.

Gregory, A.W. and B.A. Hansen, 1996, “Residual-Based Tests for Cointegration in Models with Regime Shifts”, *Journal of Econometrics* 70, 99-126.

Gregory, A.W., J.M. Nason and D.G. Watt, 1996, “Testing for Structural Breaks in Cointegrated Relationships”, *Journal of Econometrics* 71, 321-341.

Hakkio, C.S. and M. Rush, 1991, “Is the Budget Defcit ‘Too Large’?”, *Economic Inquiry* 24, 429-445.

Hamilton, J., 1994, *Time Series Analysis*. Princeton, NJ: Princeton University Press.

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

Hansen, B.E., 1992, “Tests for Parameter Instability in Regressions with I(1) Processes”, *Journal of Businesss & Economic Statistics* 10, 321-335.

Hansen, H. and S. Johansen, 1993, “Recursive Estimation in Cointegrated VAR Models”, Institute of Mathematical Statistics, University of Copenhagen.

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

Haug, A., 1992, “ Has the Federal Budget Deficit Policy Changed in Recent Years?”, *unpublished manuscript*,University of Saskatckewan, Dept. of Economics.

Hendry, D.F., and A. Neale, 1990, “The Impact of Structural Breaks on Unit Root Tests”, in *Economic Structural Change*, P. Hackl and A. Westlung (eds.), IIASA and Springer-Verlag, Berlin.

Johansen, S., 1988, “Statistical Analysis of Cointegrating Vectors”, *Journal of Economic Dynamics and Control* 12, 231-254.

Johansen, S., 1991, “Estimation and Hypothesis Testing of Cointegrating Vectors in Gaussian Vector Autorregresive Models”, *Econometrica* 60, 1551-1580.

Johansen, S. and K. Juselius, 1990, “Maximum Likelihood Estimation and Inference on Cointegration with Applications to the Demand for Money”, *Oxford Bulletin of Economics and Statistics*, 52, 169-210.

Joulfaian, D. and R. Mookerjee, 1991, “Dynamics of Government Revenues and Expenditures in Industrial Economies”, *Applied Economics* 23, 1839-1844.

Kau, J. and P. Rubin, 1981, “The Size of Government”, *Public Choice* 37, 261-274.

Leybourne, S.J., B.P.M. McCabe and A.R. Tremayne, 1996, “Can Economic Time Series be Differenced to Stationarity”, *Journal of Business & Economic Statistics* 14, 435-446.

Maeso, F., 1997, “Testing Unit Root in Processes with Changing Coefficient: An Application to Real Exchange Rate”, *mimeo*, U.A.B.

Manage, N. and M. Marlow, 1986, “The Casual Relation Between Federal Expenditures and Receipts”, *Southern Economic Journal*, January, 617-629.

Meltzer, A. and S. Richard, 1981, “A Rational Theory of the Size of the Government”, *Journal of Political Economy* 89, 914-927.

Miller, S. and F. Russek, 1990, “Co-Integration and Error-Correction Models: The Temporal Causality Between Government Taxes and Spending”, *Southern Economic Journal*, august, 221-229.

Musgrave, R., 1966, “Principles of Budget Determination”, in *Public Finance: Selected Readings*, H. Camson and W. Henderson (eds.), Random House, New York, 15-27.

Osterwald-Lenum, M., 1992, “A Note with Fractiles of the Asymptotic Distribution of the Maximum-Likelihood Cointegration Rank Test Statistic”, *Oxford Bulletin of Economics and Statistics*, 54, 461-472.

Owoye, O., 1995, “The Casual Relationship Between Taxes and Expenditures in the G7 Countries: Cointegration and Error-Correction Models”, *Applied Economics Letters* 2, 19-22.

Payne, J.E., 1998, “The Tax-Spend Debate: Time Series Evidence from State Budgets”, *Public Choice* 95, 307-320.

Peacock, A. and J. Wiseman, 1979, “Approches to the Analysis of Government Expenditure Growth”, *Public Finance Quarterly* 7, 3-23.

Perron, P., 1989, “The Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”, *Econometrica* 57, 1361-1401.

Perron, P., 1990, “Testing for Unit Root in a Time Series with a Changing Mean”, *Journal of Business & Economic Statistics* 8, 153-162.

Perron, P. and T.J. Vogelsang, 1992, “Non-stationarity and Level Shifts with an Application to Purchasing Power Parity”, *Journal of Business & Economic Statistics* 10, 301-320.

Phillips, P.C.B., 1987, “Time Series Regression with a Unit Root”, *Econometrica* 55, 277-301.

Phillips, P.C.B. and B. E. Hansen, 1990, “Statistical Inference in Instrumental Variables Regression with I(1) Processes”, *Review of Economic Studies* 57, 99-125.

Phillips, P.C.B. and S. Ouliaris, 1990, “Asymptotic Properties of Residual Based Tests for Cointegration”, *Econometrica* 58, 165-193.

Phillips, P.C.B. and P. Perron, 1988, "Testing for a Unit Root in Time Series Regression", *Biometrica* 75, 335-346.

Quintos, C.E., 1995, "Sustainability of the Deficit Process with Structural Shifts", *Journal of Business & Economic Statistics* 13, 409-417.

Ram, R., 1988a, “Additional Evidence on Causality Between Revenues and Government Expenditure”, *Southern Economic Journal*, january, 763-769.

Ram, R., 1988b, “A Multicountry Perspective on Causality Between Government Revenue and Government Expenditure”, *Public Finance* 43, 261-270.

Roubini, N. and J. Sachs, 1989, “Political and Economic Determinants of Budget Deficits in Industrial Democracies”, *European Economic Review* 33, 903-938.

Shibata, H. and Y. Kimura, 1986, “Are the Budget Deficits the Cause of Growth in Government Expenditures?”, in Various: *Public Finance and Public Debt*, Wayne State University Press, Detroit.

Smith, G. and S.E. Zin, 1991, “Persistent Deficits and the Market Value of Government Debt”, *Journal of Applied Econometrics* 6, 31-44.

Trehan, B. and C.E. Walsh, 1988, “Common Trends, the Government Budget Constraint and Revenue Smoothing”, *Journal of Economic Dynamics and Control* 12, 425-444.

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

Von Furstemberg, G., R.J. Green and J.H. Jeong, 1986, “Tax and Spend, or Spend and Tax”, *Review of Economics and Statistics* 68, 179-188.

Von Hagen, J., 1992, “Budgeting Procedures and Fiscal Performance in theEuropean Communities”, Commission of the European Communities, *Economic Papers* 96, october.

Walsh, C.E., 1998, *Monetary Theory and Policy*. Cambridge, MA: MIT Press.

Ward, B., 1982, “Taxes and the Size of Government”, *American Economic Review* 72, 346-350.

Wickens, M.R. and M. Uctum, 1993, "The Sustainability of Current Account Deficits: A Test of the US Intertemporal Budget Constraint", *Journal of Economic Dynamics and Control* 17, 423-441.

Wicksell, K., 1896, “A New Principle of Just Taxation”, in *Classics in the Theory of Public Finance*, R. Musgrave and R. Peacock (eds.), St. Martins Press, New York, 72-118.

Wilcox, D.W., 1989, "The Sustainability of Government Deficits: Implications of the Present-Value Borrowing Constraint", *Journal of Money, Credit and Banking* 21, 291-306.

Zivot, E. and D. Andrews, 1992, “Further Evidence on the Great Crash, the Oil Price Shock, and the Unit Root Hypothesis”, *Journal of Business & Economic Statistics* 10, 251-270.