Available online at www.sciencedirect.com
ScienceDirect
Journal of Macroeconomics 28 (2006) 596–608
Journal of MACROECONOMICS
www.elsevier.com/locate/jmacro

Is the budget deficit sustainable when fiscal policy is non-linear? The case of Spain

Oscar Bajo-Rubio [a,b,*], Carmen Díaz-Roldán [a], Vicente Esteve [c]

[a] *Department of Economics, Universidad de Castilla-La Mancha, 13071 Ciudad Real, Spain*
[b] *Instituto de Estudios Fiscales, Avda. Cardenal Herrera Oria 378, 28035 Madrid, Spain*
[c] *Department of Applied Economics II, Universidad de Valencia, 46071 Valencia, Spain*

Received 22 April 2003; accepted 24 August 2004
Available online 7 July 2006

Abstract

In this paper, we re-examine the long-run sustainability of budget deficits, when fiscal policy is conducted as a non-linear process. Our empirical methodology makes use of recent developments on threshold cointegration that consider the possibility of a non-linear relationship between government expenditures and revenues. The analysis is applied to the case of Spain, a country that has recently accomplished an important fiscal consolidation. Overall, our results suggest the presence of significant non-linear effects in Spanish fiscal policy, so that fiscal authorities would cut deficits only if they are 'large', which would assure in turn their long-run sustainability.

JEL classification: E62; H62

Keywords: Fiscal policy; Sustainability; Threshold cointegration; Non-linearity

1. Introduction

The formation of the Economic and Monetary Union (EMU) by 12 member countries of the European Union (EU) means that fiscal policy becomes the main instrument of stabilization policies available to their national authorities. As a consequence, issues related to fiscal

* Corresponding author. Tel.: +34 926 295300x3580; fax: +34 926 295211.
E-mail address: oscar.bajo@uclm.es (O. Bajo-Rubio).

0164-0704/$ - see front matter
doi:10.1016/j.jmacro.2004.08.002

policy have gained a growing interest in recent years, in both academic and policy circles. In particular, the long-run sustainability of budget deficits has become a matter of great concern. On the one hand, emphasizing the sustainability of deficits would mean in itself a limit to the flexibility of fiscal policy as a stabilization tool, insofar as this might lead to a more restrictive fiscal stance. On the other hand, EMU might lead to relax fiscal discipline, so increasing the risks of default and bailout. In turn, all this was reflected in the Treaty of Maastricht, which defined budgetary rules that countries had to satisfy in order to join EMU, later enforced in the clauses of the Pact for Stability and Growth (De Grauwe, 2003).

The traditional approach to the analysis of the sustainability of budget deficits has tested whether the government's intertemporal budget constraint (IBC) holds, that is, whether the current level of debt equals the present discounted value of primary surpluses. Empirical tests on sustainability, however, are still inconclusive due to differences in the econometric methodology, the particular specification of the transversality condition, and the sample period used.

Several procedures to test for the IBC have been proposed in the literature, which focus on the univariate properties of the government deficit and debt (Hamilton and Flavin, 1986; Wilcox, 1989), and on the presence of a long-run, linear, cointegration relationship between government revenues and expenditures (Trehan and Walsh, 1988, 1991; Haug, 1991; Smith and Zin, 1991). Further on, the eventual occurrence of structural breaks in this cointegrating relationship has been examined by Hakkio and Rush (1991a), who assumed the break point as exogenously given; and by Haug (1995), Quintos (1995), Camarero et al. (1998), Makrydakis et al. (1999), and Martin (2000), where the break point was endogenously derived.

However, and unlike the above quoted papers, which assumed a linear relationship between government expenditures and revenues, several recent studies have suggested the possibility that fiscal policy may have non-linear effects, in the sense that both the size and the sign of the response of macroeconomic variables to fiscal policy actions could be different depending on the way and the initial conditions in which such policy actions are implemented; see Giavazzi et al. (2000) for a recent comprehensive empirical analysis.

The objective of this paper is to provide some additional evidence on the sustainability of budget deficits, when fiscal policy is conducted as a non-linear process, which has been hardly treated in the literature. To this end, we make use of recent developments on threshold cointegration that consider the possibility of a non-linear relationship between government expenditures and revenues. Our approach should be relevant for a number of EU countries that have faced problems of fiscal sustainability during the 1980s and 1990s. In particular, we analyze the case of Spain, a country traditionally experiencing high budget deficits, which has accomplished an important fiscal consolidation in last years that has allowed her to be able to participate in EMU since the outset.

The rest of the paper is organized as follows. Some motivation for the analysis in the paper, in terms of the recent literature on non-linear fiscal policy, is provided in Section 2. Next, the empirical methodology is outlined in Section 3. The results of the tests on threshold cointegration between Spanish government expenditures and revenues are presented in Section 4. Finally, the main conclusions are summarized in Section 5.

2. Non-linearities in fiscal policy

A wide attention has been recently given in the literature on stabilization policy, to the possible non-Keynesian effects of fiscal policy and, in general, its non-linear effects. The

pioneering paper was Giavazzi and Pagano (1990), who studied the cases of two strong fiscal contractions performed during the 1980s (namely, Denmark in 1983–86, and Ireland in 1987–89) that were accompanied by a vigorous economic expansion, contrarily to what should be expected according to a standard Keynesian view.

Such a result would be explained through the operation of two different channels, i.e., the wealth channel and the expectations channel. On the one hand, the fiscal contraction can reduce interest rates, which would increase the market value of non-human wealth, and so private consumption and aggregate demand. On the other hand, a large enough fiscal contraction might lead the agents to expect lower taxes in the future, which would increase households' permanent disposable income, and hence their planned and current consumption. Notice that this literature tends to stress the role of the expectations channel, since the wealth channel is deemed to be insufficient to generate the observed effects on private consumption and aggregate demand.

More generally, the possibility of non-Keynesian effects following changes in taxes or government spending would mean that fiscal policy could behave in a non-linear way, so that both the magnitude and the size of its effects on macroeconomic variables would be different according to the particular circumstances in which the fiscal measures are enforced.

So, for instance, the non-Keynesian effects would be more likely if fiscal policy changes are sufficiently large and protracted (Giavazzi and Pagano, 1996). The composition of the fiscal measures also matters, since cutting public sector wages and transfer programs, rather than raising taxes, would tend to produce permanent reductions in the budget deficit, which in turn would favor the possibility of an expansionary effect on aggregate demand (Alesina and Perotti, 1995). And the latter, indeed, would be reinforced when combined with two additional ingredients: some form of wage agreement with the unions that ensures wage moderation, and a devaluation immediately before the fiscal tightening (Alesina and Ardagna, 1998).

Finally, the role of the initial conditions in which the fiscal adjustment is performed, also seems to be of great importance, with an expansionary fiscal contraction being more likely for high initial levels in tax rates or in the ratio of debt to output (Blanchard, 1990). The latter point was later extended by Perotti (1999), who developed a model where fiscal shocks would have Keynesian effects on private consumption (i.e., positive for government expenditure shocks, negative for tax shocks) in 'normal' times, and non-Keynesian effects in 'bad' times, where 'bad' times were defined in terms of a high value of the accumulated government debt.

Until now we have examined under which circumstances a fiscal policy action, and in particular a fiscal adjustment, could lead to effects on macroeconomic variables that not always agree with those expected according to a standard Keynesian view. But, what about the timing of the fiscal actions? As noticed by Alesina and Drazen (1991), although it could be agreed that stabilization requires a change in fiscal policy to eliminate budget deficits, stabilizations can be delayed if there is disagreement about how the burden of the policy change is to be shared. Hence, stabilization would occur only when certain groups give in, and allow the others to decide on how the burden of the fiscal adjustment is distributed.

More specifically, Bertola and Drazen (1993) argue that significant cuts in government spending take place only when the ratio of government spending to output hits a trigger point, which implies that abrupt changes in fiscal policy should not be observed

frequently. In other terms, these authors introduce a non-linearity in the reaction function of fiscal policymakers. And this in turn would imply a non-linear relationship between the consumption-to-output ratio and the government spending-to-output ratio, provided that the current policy generates sufficiently strong expectations of future policy changes in the opposite direction.

The above considerations should be of special relevance for the case of EMU. As is well known, those countries wishing to participate in EMU had to satisfy, according to the Treaty of Maastricht, several 'convergence criteria', two of them related to fiscal policy; namely, the budget deficit and government debt should not exceed 3% and 60%, respectively, of their GDP. Leaving aside the suitability of these fiscal criteria, whose economic rationale was questioned by several authors (see, e.g., Buiter et al., 1993 or De Grauwe, 1996), at that time a wide discussion arose about the potential difficulty of fulfilling them for most of the countries involved. So, for instance, Wyplosz (1997) observed that, given the tight monetary policies performed during the mid-1990s to meet the criterion on inflation, growth slowed down, which reduced tax revenues and made even more difficult to achieve the fiscal targets. However, most countries were able in the end to satisfy the fiscal and other convergence criteria, and so to participate in EMU from the start. What could it happen?

The objective of this paper, then, will be trying to quantify the precise moment in which the authorities would correct a deficit, so that a required fiscal adjustment is finally enforced; i.e., the 'trigger point', according to Bertola and Drazen's (1993) contribution. Notice, on the other hand, that the empirical literature on fiscal sustainability has hardly incorporated the new developments on non-linearities in fiscal policy. We can just quote Cipollini (2001), who introduced a regime shift in the adjustment towards a linear long-run (cointegrating) relationship between total government revenues and expenditures for the UK, using a smooth transition error correction model to test for non-linearities or asymmetries in the adjustment process. Also, Chortareas et al. (2003) use stationarity tests that allow the alternative hypothesis to incorporate non-linearities, for the case of several Latin American countries.

More recently, Bajo-Rubio et al. (2004) have found strong evidence of non-linearities in the evolution of the Spanish budget deficit in terms of a threshold autoregressive model, so that the deficit dynamics would be different depending on whether the change in the deficit was below or above an endogenously estimated threshold; in other words, mean-reverting dynamic behavior in the budget deficit would be expected once such threshold was reached. In particular, significant fiscal stabilizations were found to occur when, in a certain year, the ratio of the budget deficit to GDP showed an increase of more than 1.9% between the previous year and the sixth year before. A similar analysis has been applied to the US case by Arestis et al. (2004).

Unlike our previous paper, which made use of a threshold autoregressive model for the budget deficit, we analyze here the possible presence of threshold cointegration between government expenditures and revenues. In other words, our objective will be analyzing the evolution of government expenditures and revenues, rather than the budget deficit as a whole.

3. Methodology

The concept of threshold cointegration was introduced by Balke and Fomby (1997) as a feasible way to combine non-linearity and cointegration. As is well known, systems in which variables are cointegrated can be characterized by an error correction model

(ECM), which describes how the variables respond to deviations from the equilibrium. In this way, the ECM can be characterized as the adjustment process through which the long-run equilibrium is maintained. The traditional approach, however, assumes that such a tendency to move towards the long-run equilibrium is present every time period.

Balke and Fomby (1997) stressed the possibility that this movement towards the long-run equilibrium might not occur in every time period, due to the presence of some adjustment costs on the side of economic agents. In other words, there could be a discontinuous adjustment to equilibrium so that, only when the deviation from the equilibrium exceeds a critical threshold, the benefits of adjustment are higher than the costs, and economic agents move the system back to equilibrium. Threshold cointegration would characterize this discrete adjustment as follows: the cointegrating relationship does not hold inside a certain range, but holds if the system gets 'too far' from the equilibrium; i.e., cointegration would hold only if the system exceeds a certain threshold.

This type of discrete adjustment could be particularly useful to describe the behavior of fiscal authorities. More specifically, fiscal authorities would intervene by cutting budget deficits only when these are 'too large', in order to meet the IBC. The concept of threshold cointegration would capture the possibility of a non-linear relationship between government expenditures and revenues, so that mean-reverting dynamic behavior in the budget deficit (or a cointegrating relationship between government expenditures and revenues) should be expected only when a certain threshold is reached.

When testing for threshold cointegration, Balke and Fomby (1997) proposed applying several univariate tests previously developed in the literature, to the known cointegrating residual (i.e., the error-correction term). Further contributions include Forbes et al. (1999), who developed a Bayesian estimation procedure; and Lo and Zivot (2001), who extended Balke and Fomby's approach to a multivariate threshold cointegration model with a known cointegrating vector, using Tsay's (1998) and multivariate extensions of Hansen's (1996) tests. More recently, Hansen and Seo (2002) have contributed further to this literature by examining the case of an unknown cointegration vector. In particular, these authors proposed a vector error-correction model (VECM) with one cointegrating vector and a threshold effect based on the error-correction term, and developed a Lagrange multiplier (LM) test for the presence of a threshold effect. This will be the approach followed in this paper.

Hansen and Seo (2002) considered a two-regime threshold cointegration model, or a non-linear VECM of order $l+1$, such as

$$\Delta x_t = \begin{cases} A_1' X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) \leqslant \gamma \\ A_2' X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) > \gamma \end{cases} \tag{1}$$

with

$$X_{t-1}(\beta) = \begin{pmatrix} 1 \\ w_{t-1}(\beta) \\ \Delta x_{t-1} \\ \Delta x_{t-2} \\ \vdots \\ \Delta x_{t-l} \end{pmatrix}$$

where x_t is a p-dimensional $I(1)$ time series which is cointegrated with one $p \times 1$ cointegrating vector β, $w_t(\beta) = \beta' x_t$ is the $I(0)$ error-correction term, u_t is an error term, A_1 and A_2 are coefficient matrices, and γ is the threshold parameter.

As can be seen, the threshold model (1) has two regimes, depending on whether deviations from the equilibrium (defined by the value of the error-correction term) are below or above the threshold, where A_1 and A_2 describe the dynamics in each of the regimes. In one of the regimes there would be no tendency for the variables x_t to revert to an equilibrium (i.e., the variables would not be cointegrated); on the contrary, in the other regime there would be a tendency for the variables x_t to move towards some equilibrium (i.e., the variables would be cointegrated).

Next, Hansen and Seo (2002) proposed two heteroskedastic-consistent LM test statistics for the null hypothesis of linear cointegration (i.e., there is no threshold effect), against the alternative of threshold cointegration (i.e., model (1)). The first test would be used when the true cointegrating vector is known a priori, and is denoted as

$$\sup \mathrm{LM}^0 = \sup_{\gamma_L \leqslant \gamma \leqslant \gamma_U} \mathrm{LM}(\beta_0, \gamma) \tag{2}$$

where β_0 is the known value of β (in the case analyzed below, $\beta_0 = 1$); whereas the second test would be used when the true cointegrating vector is unknown, and is denoted as

$$\sup \mathrm{LM} = \sup_{\gamma_L \leqslant \gamma \leqslant \gamma_U} \mathrm{LM}(\tilde{\beta}, \gamma) \tag{3}$$

where $\tilde{\beta}$ is the null estimate of β. In both tests, $[\gamma_L, \gamma_U]$ is the search region set so that γ_L is the π_0 percentile of $\widetilde{w}_{t-1}$, and γ_U is the $(1 - \pi_0)$ percentile; Andrews (1993) suggested setting π_0 between 0.05 and 0.15. Finally, Hansen and Seo (2002) developed two bootstrap methods to calculate asymptotic critical values and p-values.

4. Results

In this section, we re-examine the issue of the sustainability of the Spanish budget deficit using the new approach of Hansen and Seo (2002), described in the last section. We use data on total (i.e., inclusive of debt interest) government expenditures, and total government revenues, both of them as a ratio to GDP, which are denoted by gr_t and rev_t, respectively. These data are taken at two alternative frequencies: annual, from 1964 to 2003; and quarterly, from 1982:1 to 2004:1. Using the annual series has the advantage of providing a longer span of data, which allows to test more properly for cointegration (Hakkio and Rush, 1991b). However, since the number of observations available from the annual series might be too short for the correct implementation of our empirical methodology, we have made also use of quarterly series, only available from 1982 on, which provide a significantly higher number of observations (i.e., 89 versus 40).

The annual data on government expenditures and revenues come from the Spanish National Institute of Statistics, and the quarterly data from IGAE (National Audit Office); the source for GDP, both at the annual and quarterly frequencies, is the National Institute of Statistics. Notice that the annual series correspond to the general government, unlike the quarterly series that correspond to the central government (i.e., excluding the social security and regional governments), so that both series should not be strictly comparable. However, in a recent paper using data for 9 EMU countries, Pérez (2005) shows that

infra-annual (i.e., monthly/quarterly) data on central government expenditures and revenues provide useful information when anticipating the evolution of the relevant variable for fiscal policy, namely, the annual general government deficit.

As a first step of the analysis, we have tested for the order of integration of the two series. To this end, we have used a modified version of the Dickey–Fuller and Phillips–Perron tests proposed by Ng and Perron (2001), which tries to solve the main problems present in these conventional tests for unit roots.

In general, most of the conventional unit root tests suffer from three problems. First, they have low power when the root of the autoregressive polynomial is close to, but less than unit (DeJong et al., 1992). Second, most of the tests suffer from severe size distortions when the moving-average polynomial of the first differences series has a large negative autoregressive root (Schwert, 1989). Third, implementing the unit root tests often implies the selection of an autoregressive truncation lag, k, which is strongly associated with size distortions and/or the extent of power loss (Ng and Perron, 1995).

Trying to address these critiques, Ng and Perron (2001) have proposed a methodology that would be robust against the three problems quoted above. This consists of a class of modified tests, $\bar{M}Z_{\alpha}^{\mathrm{GLS}}$ and $\bar{M}Z_{t}^{\mathrm{GLS}}$, originally developed in Stock (1999) as M tests, with GLS detrending of the data as proposed in Elliott et al. (1996), and using the modified Akaike information criterion. In addition, Ng and Perron (2001) have proposed a similar procedure that corrects the problems associated with the standard Augmented Dickey–Fuller test, $\mathrm{ADF}^{\mathrm{GLS}}$.

Table 1 shows the results of the three tests, $\bar{M}Z_{\alpha}^{\mathrm{GLS}}$, $\bar{M}Z_{t}^{\mathrm{GLS}}$, and $\mathrm{ADF}^{\mathrm{GLS}}$, for the two data frequencies. As shown in the table, the null hypothesis of non stationarity for the two series in levels cannot be rejected, independently of the test and the frequency of the data; and the presence of two unit roots is clearly rejected at the 1% significance level. Accord-

Table 1
Ng–Perron tests of unit roots

Variable	$\bar{M}Z_{\alpha}^{\mathrm{GLS}}$	$\bar{M}Z_{t}^{\mathrm{GLS}}$	$\mathrm{ADF}^{\mathrm{GLS}}$
(A) Annual data			
$I(2)$ vs. $I(1)$, Case: $p=0,\ \bar{c}=-7.0$			
$\Delta \mathrm{gr}_t$	-16.6[a]	-2.86[a]	-4.24[a]
$\Delta \mathrm{rev}_t$	-18.9[a]	-3.07[a]	-6.25[a]
$I(1)$ vs. $I(0)$, Case: $p=1,\ \bar{c}=-13.5$			
gr_t	-0.43	-0.20	-0.21
rev_t	-2.63	-0.97	-1.01
(B) Quarterly data			
$I(2)$ vs. $I(1)$, Case: $p=0,\ \bar{c}=-7.0$			
$\Delta \mathrm{gr}_t$	-40.2[a]	-4.48[a]	-12.3[a]
$\Delta \mathrm{rev}_t$	-40.9[a]	-4.49[a]	-7.25[a]
$I(1)$ vs. $I(0)$, Case: $p=1,\ \bar{c}=-13.5$			
gr_t	-6.55	-1.58	-1.71
rev_t	-4.22	-1.20	-1.33

Notes: The autoregressive truncation lag, k, has been selected using the modified Akaike information criterion, as proposed by Perron and Ng (1996).

[a] Denotes significance at the 1% level. The critical values are taken from Ng and Perron (2001), Table 1.

ingly, the two series would be concluded to be $I(1)$, both at the annual and quarterly frequencies.

Next, we have applied the tests of threshold cointegration proposed by Hansen and Seo (2002), namely, sup LM^0 (for a given $\beta = 1$) and sup LM (for an estimated β). For the two tests, the p-values are calculated using a parametric bootstrap method (with 5000 simulation replications), as proposed by Hansen and Seo (2002). To select the lag length of the VAR, we have used the Akaike and Bayesian information criteria, both of them leading to $l = 1$ for annual data, and to $l = 2$ for quarterly data. The results of the tests are reported in Table 2.

Beginning with the annual data, threshold cointegration would appear at the 2% significance level for the sup LM^0 test, i.e., when β is fixed at unity, so that the null hypothesis of linear cointegration would be strongly rejected. The estimated threshold is $\hat{\gamma} = 5.30$, with the error-correction term defined as $w_t = \mathrm{gr}_t - \mathrm{rev}_t$ (i.e., the budget deficit). Hence, the first regime would occur when government expenditures are more than 5.30% points above revenues, as a ratio to GDP; in other words, when the government deficit as a ratio to GDP is above 5.30%. This would be the relatively unusual regime, including 13% of the observations (namely, the years 1982, 1984, 1985, 1993, and 1995). In turn, the second or usual regime (with 87% of the observations) would occur when the government deficit as a ratio to GDP is below 5.30%.

The estimated two-regime threshold VAR (heteroskedasticity-consistent standard errors in parentheses) is

$$\Delta \mathrm{gr}_t = \begin{cases} \underset{(1.94)}{20.27} - \underset{(0.31)}{2.55}\, w_{t-1} - \underset{(0.55)}{4.14}\, \Delta \mathrm{gr}_{t-1} + \underset{(0.44)}{2.94}\, \Delta \mathrm{rev}_{t-1} + u_{1t}, & w_{t-1} \geqslant 5.30 \\ \underset{(0.22)}{0.36} + \underset{(0.07)}{0.007}\, w_{t-1} + \underset{(0.17)}{0.65}\, \Delta \mathrm{gr}_{t-1} - \underset{(0.26)}{0.11}\, \Delta \mathrm{rev}_{t-1} + u_{2t}, & w_{t-1} < 5.30 \end{cases}$$

$$\Delta \mathrm{rev}_t = \begin{cases} \underset{(2.97)}{11.38} - \underset{(0.48)}{1.12}\, w_{t-1} - \underset{(0.84)}{3.17}\, \Delta \mathrm{gr}_{t-1} + \underset{(0.67)}{2.17}\, \Delta \mathrm{rev}_{t-1} + u_{1t}, & w_{t-1} \geqslant 5.30 \\ \underset{(0.20)}{0.43} + \underset{(0.09)}{0.03}\, w_{t-1} + \underset{(0.10)}{0.35}\, \Delta \mathrm{gr}_{t-1} - \underset{(0.24)}{0.20}\, \Delta \mathrm{rev}_{t-1} + u_{2t}, & w_{t-1} < 5.30 \end{cases}$$

where significant error-correction effects appear only in the first regime, i.e., when either government expenditures are well above revenues or government deficit is relatively high. On the contrary, in the second regime error-correction effects and dynamics are minimal, both in terms of significance and size of the coefficients, only being significant the coefficient on the lagged change in gr_t.

Fig. 1 plots the error-correction effect, i.e., the estimated response of government expenditures and revenues to the discrepancy between them (i.e., to the size of the government

Table 2
Hansen–Seo tests of threshold cointegration

	Annual data		Quarterly data	
	sup LM^0	sup LM	sup LM^0	sup LM
Test statistic value	16.79	6.51	18.77	18.66
Calculated p-values	0.02	0.95	0.30	0.08
Threshold parameter	5.30	−2.40	2.92	1.41
Estimate of the cointegrating vector	1.00	1.18	1.00	1.02

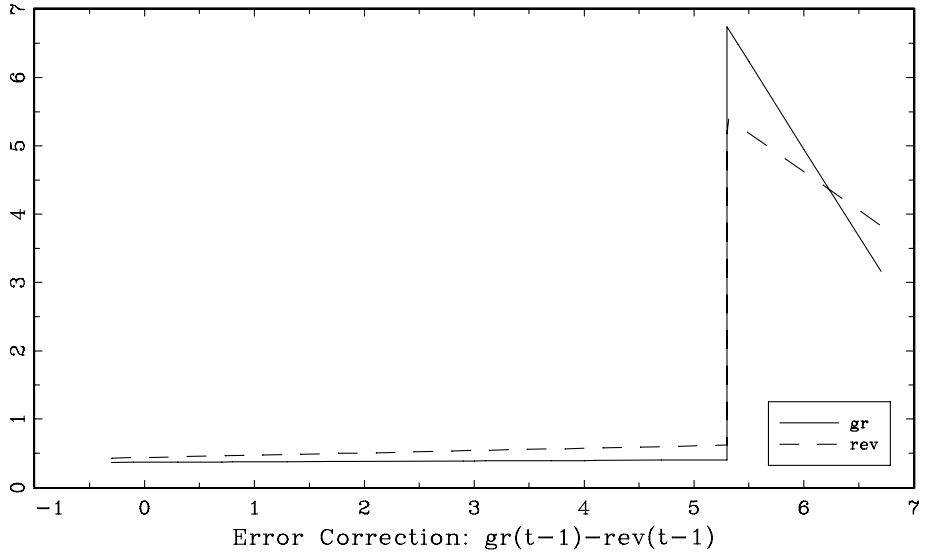

Fig. 1. Response of expenditures and revenues to error correction. Annual data, 1964–2003.

deficit) in the previous period, holding the other variables constant. As can be seen, for a 'small' deficit (i.e., lower than 5.30% of GDP), the response of both expenditures and revenues would be nearly zero. However, if the deficit were 'large' (i.e., greater than 5.30% of GDP), both expenditures and revenues would decrease sharply with the size of the deficit. And, since the response of expenditures would be larger than that of revenues (notice that the estimated error-correction term is twice for expenditures than for revenues), the government deficit would fall accordingly.

When turning to quarterly data, threshold cointegration would now appear at the 8% significance level for the sup LM test, with β estimated at 1.02 (see Table 2). The estimated threshold would be now $\hat{\gamma} = 1.41$, with the error-correction term defined as $w_t = \mathrm{gr}_t - 1.02\mathrm{rev}_t$, or 1.75 in terms of the budget deficit (i.e., 7% of GDP in annual figures), with rev_t computed at its average value over the sample period. The first or usual regime would include 84% of the observations, and the second or unusual regime the remaining 16%; as before, the highest deficits (that would correspond now to the usual regime) would have appeared at the first half of the eighties, and the period 1993–95.

On the other hand, the estimated two-regime threshold VAR (heteroskedasticity-consistent standard errors in parentheses) would be now

$$\Delta \mathrm{gr}_t = \begin{cases} \underset{(0.29)}{-0.64} + \underset{(0.07)}{0.15}\, w_{t-1} - \underset{(0.10)}{0.64}\, \Delta \mathrm{gr}_{t-1} + \underset{(0.10)}{0.32}\, \Delta \mathrm{gr}_{t-2} \\ \quad \underset{(0.09)}{-0.28}\, \Delta \mathrm{rev}_{t-1} + \underset{(0.11)}{0.23}\, \Delta \mathrm{rev}_{t-2} + u_{1t}, \quad w_{t-1} \geqslant 1.41 \\ \\ \underset{(0.98)}{0.91} - \underset{(0.79)}{1.15}\, w_{t-1} + \underset{(1.17)}{1.18}\, \Delta \mathrm{gr}_{t-1} - \underset{(0.41)}{0.03}\, \Delta \mathrm{gr}_{t-2} \\ \quad \underset{(0.51)}{-0.80}\, \Delta \mathrm{rev}_{t-1} + \underset{(0.35)}{0.002}\, \Delta \mathrm{rev}_{t-2} + u_{2t}, \quad w_{t-1} < 1.41 \end{cases}$$

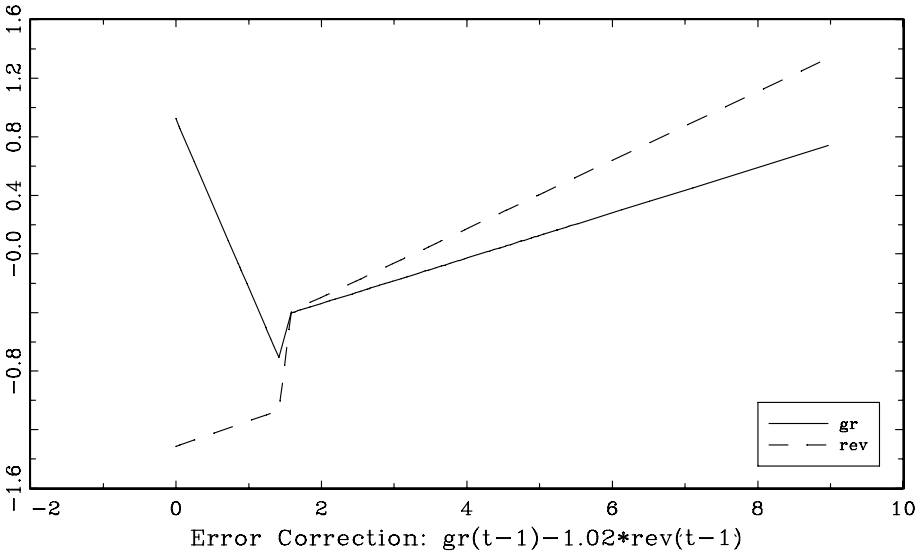

Fig. 2. Response of expenditures and revenues to error correction. Quarterly data, 1982:1–2004:1.

$$\Delta \mathrm{rev}_t = \begin{cases} \underset{(0.23)}{-0.76} + \underset{(0.06)}{0.23}\, w_{t-1} - \underset{(0.07)}{0.32}\, \Delta \mathrm{gr}_{t-1} - \underset{(0.08)}{0.05}\, \Delta \mathrm{gr}_{t-2} \\ \quad \underset{(0.08)}{-0.12}\, \Delta \mathrm{rev}_{t-1} - \underset{(0.09)}{0.19}\, \Delta \mathrm{rev}_{t-2} + u_{1t}, \quad w_{t-1} \geqslant 1.41 \\ \\ \underset{(0.51)}{-1.31} + \underset{(0.41)}{0.17}\, w_{t-1} + \underset{(0.86)}{1.25}\, \Delta \mathrm{gr}_{t-1} + \underset{(0.32)}{0.05}\, \Delta \mathrm{gr}_{t-2} \\ \quad \underset{(0.23)}{-1.29}\, \Delta \mathrm{rev}_{t-1} - \underset{(0.16)}{0.02}\, \Delta \mathrm{rev}_{t-2} + u_{2t}, \quad w_{t-1} < 1.41 \end{cases}$$

where significant error-correction effects appear again just in the first regime. The error-correction effect is plotted in Fig. 2, where government deficits above the threshold would lead now to higher expenditures and revenues, with a stronger response on the side of revenues (since the size of the estimated error-correction term is higher for the latter), so that deficits would fall accordingly.

Overall, the above evidence, taken together with that from a previous paper (Bajo-Rubio et al., 2004), suggests that Spanish fiscal policy would have shown significant non-linear effects. In particular, the outstanding fiscal consolidation experienced in recent years would have occurred only once the budget deficit exceeded a certain threshold and became 'too large', with the main consolidation effort taking place at the end of the first half of the eighties, and the period 1993–95. As a consequence, budget deficits would have shown a mean-reverting dynamic behavior after such threshold was reached, which in turn would have assured their long-run sustainability.

5. Conclusions

This paper has re-examined the long-run sustainability of budget deficits, when fiscal policy is conducted as a non-linear process. The empirical methodology has made use of Hansen and Seo's (2002) recent contribution, based on a threshold cointegration model

that considers the possibility of a non-linear relationship between government expenditures and revenues. Our analysis should be relevant for a number of EU countries that have faced problems of fiscal sustainability in last years, and has been applied to the case of Spain, a country traditionally experiencing high budget deficits, which has recently accomplished an important fiscal consolidation. In the empirical application we have used both annual data for the general government, which allow for a wider time perspective; and quarterly data for the central government, which provide a significantly higher number of observations.

According to our results, the null hypothesis of linear cointegration between government expenditures and revenues would be rejected in favor of a two-regime threshold cointegration model, with the threshold parameter estimated at 5.30% of GDP for the general government deficit (annual data); and at 7% of GDP for the central government deficit (quarterly data). The main consolidation effort would have occurred at the end of the first half of the eighties, and the period 1993–95. In this way, if the budget deficit was 'large' (i.e., greater than the above figures), the immediate adjustment in both expenditures and revenues would lead to a fall in the deficit, which would turn back to more manageable figures.

In the end, these results would suggest the presence of a significant non-linear behavior in Spanish fiscal policy. Specifically, they would support the theoretical insights put forward by Bertola and Drazen (1993) and other authors, who concluded that fiscal authorities would cut deficits only if they were large enough, ensuring their sustainability in the long run; in our case, Bertola and Drazen's 'trigger point' for the Spanish budget deficit has been estimated at 5.30% of GDP for the general government deficit, and at 7% of GDP for the central government deficit. And this in turn would open the door to the possibility of non-linear effects of fiscal policy, provided that the expectations mechanism concerning future policy actions was strong enough.

Acknowledgements

The authors wish to thank two anonymous referees for helpful comments to an earlier version. Financial support from the Spanish Ministry of Science and Technology, through the projects SEC2002-01892 (O. Bajo-Rubio and C. Díaz-Roldán) and SEC2002-03651 (V. Esteve), as well as from the Department of Science and Technology of the regional government of Castilla-La Mancha, through the project PBI-03-003, is also gratefully acknowledged.

References

Alesina, A., Ardagna, S., 1998. Tales of fiscal adjustment. Economic Policy 27, 487–545.

Alesina, A., Drazen, A., 1991. Why are stabilizations delayed? American Economic Review 81, 1170–1188.

Alesina, A., Perotti, R., 1995. Fiscal expansions and adjustments in OECD countries. Economic Policy 21, 205–248.

Andrews, D.W.K., 1993. Tests for parameter instability and structural change with unknown change point. Econometrica 61, 821–856.

Arestis, P., Cipollini, A., Fattouh, B., 2004. Threshold effects in the US budget deficit. Economic Inquiry 42, 214–222.

Bajo-Rubio, O., Díaz-Roldán, C., Esteve, V., 2004. Searching for threshold effects in the evolution of budget deficits: An application to the Spanish case. Economics Letters 82, 239–243.

Balke, N.S., Fomby, T.B., 1997. Threshold cointegration. International Economic Review 38, 627–645.
Bertola, G., Drazen, A., 1993. Trigger points and budget cuts: Explaining the effects of fiscal austerity. American Economic Review 83, 11–26.
Blanchard, O.J., 1990. Comment. NBER Macroeconomics Annual 5, 111–116.
Buiter, W., Corsetti, G., Roubini, N., 1993. Excessive deficits: Sense and nonsense in the Treaty of Maastricht. Economic Policy 16, 57–100.
Camarero, M., Esteve, V., Tamarit, C.R., 1998. Cambio de régimen y sostenibilidad a largo plazo de la política fiscal: El caso de España. Working Paper WP-EC 98-15, Instituto Valenciano de Investigaciones Económicas.
Chortareas, G., Kapetanios, G., Uctum, M., 2003. A non-linear approach to public finance sustainability in Latin America. Working Paper 486, Queen Mary University of London.
Cipollini, A., 2001. Testing for government intertemporal solvency: A smooth transition error correction model approach. The Manchester School 69, 643–655.
De Grauwe, P., 1996. The economics of convergence: Towards monetary union in Europe. Weltwirtschaftliches Archiv 132, 1–27.
De Grauwe, P., 2003. Economics of Monetary Union, 5th ed. Oxford University Press, Oxford.
DeJong, D.N.J., Nankervis, J.C., Savin, N.E., Whiteman, C.H., 1992. Integration versus trend-stationarity in time series. Econometrica 60, 423–433.
Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64, 813–836.
Forbes, C.S., Kalb, G.R.J., Kofman, P., 1999. Bayesian arbitrage threshold analysis. Journal of Business and Economic Statistics 17, 364–372.
Giavazzi, F., Pagano, M., 1990. Can severe fiscal contractions be expansionary? Tales of two small European countries. NBER Macroeconomics Annual 5, 75–111.
Giavazzi, F., Pagano, M., 1996. Non-Keynesian effects of fiscal policy changes: International evidence and the Swedish experience. Swedish Economic Policy Review 3, 67–103.
Giavazzi, F., Jappelli, T., Pagano, M., 2000. Searching for non-linear effects of fiscal policy: Evidence from industrial and developing countries. European Economic Review 44, 1259–1289.
Hakkio, C.S., Rush, M., 1991a. Is the budget deficit 'too large'? Economic Inquiry 29, 429–445.
Hakkio, C.S., Rush, M., 1991b. Cointegration: How short is the long run? Journal of International Money and Finance 10, 571–581.
Hamilton, J.D., Flavin, M.A., 1986. On the limitations of government borrowing: A framework for empirical testing. American Economic Review 76, 808–819.
Hansen, B.E., 1996. Inference when a nuisance parameter is not identified under the null hypothesis. Econometrica 64, 413–430.
Hansen, B.E., Seo, B., 2002. Testing for two-regime threshold cointegration in vector error-correction models. Journal of Econometrics 110, 293–318.
Haug, A.A., 1991. Cointegration and government borrowing constraints: Evidence for the US. Journal of Business and Economic Statistics 9, 97–101.
Haug, A.A., 1995. Has the federal deficit policy changed in recent years? Economic Inquiry 33, 104–118.
Lo, M., Zivot, E., 2001. Threshold cointegration and non-linear adjustment to the law of one price. Macroeconomic Dynamics 5, 533–576.
Makrydakis, S., Tzavalis, E., Balfoussias, A., 1999. Policy regime changes and the long-run sustainability of fiscal policy: An application to Greece. Economic Modelling 16, 71–86.
Martin, G.M., 2000. US deficit sustainability: A new approach based on multiple endogenous breaks. Journal of Applied Econometrics 15, 83–105.
Ng, S., Perron, P., 1995. Unit root tests in ARMA models with data dependent methods for the selection of the truncation lag. Journal of the American Statistical Association 90, 268–281.
Ng, S., Perron, P., 2001. Lag length selection and the construction of unit root tests with good size and power. Econometrica 69, 1529–1554.
Pérez, J.J., 2005. Early-warning tools to forecast general government deficit in the euro area: The role of intra-annual fiscal indicators. Working Paper 497, European Central Bank.
Perotti, R., 1999. Fiscal policy in good times and bad. Quarterly Journal of Economics 114, 1399–1436.
Perron, P., Ng, S., 1996. Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. Review of Economic Studies 63, 435–465.
Quintos, C.E., 1995. Sustainability of the deficit process with structural shifts. Journal of Business and Economic Statistics 13, 409–417.

Schwert, G.W., 1989. Tests for unit roots: A Monte Carlo investigation. Journal of Business and Economic Statistics 7, 147–159.

Smith, G.W., Zin, S.E., 1991. Persistent deficits and the market value of government debt. Journal of Applied Econometrics 6, 31–44.

Stock, J.H., 1999. A class of tests for integration and cointegration. In: Engle, R.F., White, H. (Eds.), Cointegration, Causality and Forecasting. A Festschrift in Honour of Clive W.J. Granger. Oxford University Press, Oxford, pp. 135–167.

Trehan, B., Walsh, C.E., 1988. Common trends, the government's budget balance, and revenue smoothing. Journal of Economic Dynamics and Control 12, 425–444.

Trehan, B., Walsh, C.E., 1991. Testing intertemporal budget constraints: Theory and applications to US federal budget and current account deficits. Journal of Money, Credit, and Banking 23, 206–223.

Tsay, R.S., 1998. Testing and modeling multivariate threshold models. Journal of the American Statistical Association 93, 1188–1198.

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.

Wyplosz, C., 1997. EMU: Why and how it might happen. Journal of Economic Perspectives 11, 3–22.