DOCUMENTOS DE ECONOMÍA Y FINANZAS INTERNACIONALES
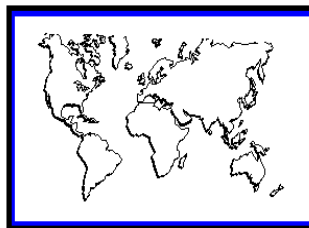
Working Papers on International Economics and Finance

On the sustainability of government deficits: Some long-term evidence for Spain, 1850-2000

Oscar Bajo-Rubio
Carmen Díaz-Roldán
Vicente Esteve

ON THE SUSTAINABILITY OF GOVERNMENT DEFICITS: SOME LONG-TERM EVIDENCE FOR SPAIN, 1850-2000[*]

Oscar Bajo-Rubio
(Universidad de Castilla-La Mancha)

Carmen Díaz-Roldán
(Universidad de Castilla-La Mancha)

Vicente Esteve
(Universidad de Valencia and Universidad de La Laguna)

Abstract

We provide a test of the sustainability of the Spanish government deficit over the period 1850-2000, from the estimation of a cointegration relationship between government expenditures and revenues derived from the intertemporal budget constraint. The longer than usual span of the data allows us to obtain more robust results on the fulfilment of the intertemporal budget constraint than most of the previous analyses. Two additional robustness checks are provided. First, we investigate the possibility of structural changes occurring along the period analyzed, using the new approach of Kejriwal and Perron (2008, 2010) to testing for multiple structural changes in cointegrated regression models. Second, we investigate whether the behaviour of fiscal authorities has been non-linear, by means of the procedure of Hansen and Seo (2002) based on a threshold cointegration model. Our results show that (i) the government deficit has been strongly sustainable in the long run, (ii) no evidence is found on any significant structural break throughout the whole period, and (iii) fiscal sustainability has been attained due to the non-linear behaviour of fiscal authorities, which have only acted on the budget deficit when it has exceeded around 4.5% of GDP.

JEL classification codes: E62, H62.

Key words: fiscal policy, sustainability, structural change, threshold cointegration, nonlinearity

---

I. Introduction

The sustainability of government deficits turns to be a crucial issue for economic policy. In principle, a deficit of the public sector can be sustainable in the short run provided that the government can borrow. However, if the interest rate on the government debt exceeds the growth rate of the economy, the resulting debt dynamics leads to an ever-increasing ratio of debt to GDP. This dynamics of debt accumulation can only be stopped if the ratio of the budget deficit to GDP becomes a surplus; or, alternatively, through a sufficiently large revenue from money creation, i.e., if seigniorage is allowed for.

When assessing the long-run sustainability of fiscal policy, the usual procedure consists of testing the government's intertemporal budget constraint (IBC). In short, for the budget deficit being sustainable, the government must run future budget surpluses equal, in present-value terms, to the current value of its outstanding liabilities. There are several studies available on the sustainability of budget deficits; see, e.g., Hamilton and Flavin (1986), Trehan and Walsh (1988, 1991), Smith and Zin (1991), Haug (1995), Quintos (1995), Martin (2000) or Bajo-Rubio, Díaz-Roldán and Esteve (2009), to name a few. The results, though, seem to be still rather inconclusive due to differences in the econometric methodology, the particular specification of the transversality condition, and the sample period used. However, a common criticism to most of the available literature is that the econometric procedures used require a large number of observations, which is not usually the case in most tests of the IBC.

In this paper, we try to overcome this problem by using an extended data set covering 150 years, for the case of Spain. The Spanish case can be of interest given the permanent difficulties experienced when balancing the government budget across those years. For most of this period, and until the fiscal reform of 1978, public revenues were kept at a very low level in order not to clash with the interests of the upper classes of the society. Accordingly, revenues turned to be insufficient to finance even small amounts of public expenditures, and deficits became chronic, leading the government to a continuous resource to seigniorage.

The aim of this paper will be to provide a test of the sustainability of the Spanish government deficit over the period 1850-2000, from the estimation of a cointegration

1

relationship between government expenditures and revenues derived from the IBC. In this way, the longer than usual span of the data will allow us to obtain more robust results on the fulfilment of the IBC than in most of previous analyses. In addition, we will provide two additional robustness checks of the estimated relationship between government expenditures and revenues. First, given the extended length of the sample, we investigate the possibility of structural changes occurring along the period analyzed. To this end, we will make use of the new approach of Kejriwal and Perron (2008, 2010) to test for multiple structural changes in cointegrated regression models. Second, as an extension of this analysis, we further investigate whether the behaviour of fiscal authorities had been non-linear. We will do this by means of the procedure developed by Hansen and Seo (2002), based on a threshold cointegration model that considers the possibility of a non-linear relationship between government expenditures and revenues. In particular, provided that fiscal sustainability holds, we analyze if fiscal authorities would have adjusted the deficit only when the difference between expenditures and revenues exceeded a certain threshold, thereby ensuring deficit sustainability in the long run.

In Section II, we present and briefly discuss our data set. The results of the sustainability test are presented in Section III; and those of the tests on structural change and non-linearity in Sections IV and V, respectively. Finally, Section VI concludes.

II. The data

We use data on total (i.e., inclusive of debt interest) expenditures, and total revenues, both measured as ratios to GDP, for the Spanish central government (i.e., excluding social security and local and regional governments), over the period 1850-2000. The data sources are Comín and Díaz (2005) for the public sector variables, and Prados de la Escosura (2003) for GDP. Notice that data on local governments are unavailable until 1958, whereas regional governments were just established after the approval of the current Constitution in 1978; also, social security only began to expand after 1966. The evolution of the two series appears in Figure 1.

[Figure 1 here]

As discussed at length in Comín (1995, 1996), the behaviour of the Spanish public sector has been characterized over most of this period by a high degree of protection and regulation, in order to favour some particular groups and sectors, rather than satisfying collective needs (such as infrastructure, or social expenditures). Recall that, at that time and until, say, the first 1960s, Spain was a poor and underdeveloped country, according to Western Europe's standards. The small levels of expenditure were explained by the lack of revenues; and there were insufficient revenues since, in a context of limited democracy, governments representing the wealthy classes of the society were unwilling to affect their interests. Such a situation begins to change in the 1960s, but the process of economic liberalization was insufficient and soon partially reverted, so Spain had to wait until the restoration of democracy after 1977, and especially the integration in the now European Union (EU) in 1986, to enjoy a public sector comparable to that of the rest of Western Europe.

Government expenditure was kept at a minimum over the 19th century, according to the principles of the liberal state. A greater role for the government involvement in the economy, especially regarding public works and social protection, appears between 1900 and 1935. Such a trend, however, finally arrived to a halt due to political instability and the lack of revenues. The increase in public expenditure after the end of the Spanish civil war was mainly due to the higher defence spending, and soon reverted to lower levels than in the pre-war years. Government expenditure as a ratio to GDP only began to increase in the mid-1960s, due to higher spending on education, housing, and social security; however, a modern welfare state was not developed due to the lack of revenues. The process of modernization of the Spanish public sector, coupled with an increase in the functions performed by the state, can be dated at the end of the 1970s, in the aftermath of the economic crisis of that period, the restoration of democracy and, later on, the integration into the EU. Accordingly, the last years have contemplated an intense process of building a welfare state on European standards, and the development of modern and improved infrastructure. Even so, the ratio of government expenditure to GDP is still lower than the EU average.

On the other hand, over most of the period government revenues proved to be insufficient to finance expenditures, despite the fiscal reforms performed in 1845 (the first attempt to build a modern fiscal system), and 1900 (following the loss of the last

Spanish colonies in Cuba and the Philippines). The inability of the governments to hinder the interests of the upper classes, coupled with the intense fiscal fraud, was at the root of this situation. The tax system for most of the period was built on French lines (the so called "Latin tax system"), and was based on the prevalence of indirect taxes, falling on specific consumption goods, and the setting of non-personal or product taxes, proportional and not progressive. The advantage of such a system was the simplicity of tax collection, but at the cost of an easy tax evasion. Only after the fiscal reform of 1978, with the creation of a modern personal income tax, completed in 1986 with the introduction of the value added tax at the time of the integration into the EU, the Spanish fiscal system can be thought to be comparable to that of the rest of Western Europe.

Even if both revenues and expenditures remained at low levels, budget deficits dominated over most of the period, which led the government to a continuous resource to seigniorage until recent years, when a more orthodox financing of the deficits has been set up. In particular, from 1982 on budget deficits, previously financed mostly via credits from the Bank of Spain, were increasingly financed using market mechanisms, through the issuing of public debt. Finally, government deficit financing by the central bank was explicitly forbidden according to the provisions of Article 104a of the Maastricht Treaty, entering into force on November 1, 1993. Despite the fact that surpluses prevailed in only 29 years between 1850 and 2000, in relative terms the magnitude of the government deficit did not tend to increase along the period (Comín and Díaz, 2005, p. 878).

III. The sustainability of Spanish budget deficits, 1850-2000

The customary framework used to test for the sustainability of budget deficits starts from the government's IBC that, in terms of GDP shares, becomes:

$$b_t = \sum_{j=0}^{\infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t s_{t+j+1} + \lim_{j \to \infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t b_{t+j+1} \qquad (1)$$

where b and s denote, respectively, the total government debt and the primary budget surplus (i.e., excluding interest payments), both as ratios to GDP. In addition, E is the

expectations operator; and x and r stand, respectively, for the rate of growth of real GDP and the real interest rate, both assumed to be constant for simplicity.

From here, the condition for fiscal sustainability is:

$$\lim_{j \to \infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t b_{t+j+1} = 0 \qquad (2)$$

i.e., the government must run expected future budget surpluses equal, in present-value terms, to the current value of its outstanding debt.

The cointegration framework to test for the IBC follows once first differences are taken in (1):

$$\Delta b_t = \sum_{j=0}^{\infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t \Delta s_{t+j+1} + \lim_{j \to \infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t \Delta b_{t+j+1} \qquad (3)$$

so that sustainability would require:

$$\lim_{j \to \infty} \left(\frac{1+x}{1+r} \right)^{j+1} E_t \Delta b_{t+j+1} = 0 \qquad (4)$$

Under a no-Ponzi scheme rule, the right-hand side of equation (3) will be stationary as long as the budget surplus and the stock of government debt are all stationary in first differences. In order to test for condition (4), the usual procedure consists of testing for the stationarity of

$$\Delta b_t = exp_t - rev_t$$

where exp_t and rev_t denote the ratios of the government's total expenditures and revenues, respectively, to GDP; that is, $exp_t - rev_t$ would be the total budget deficit (i.e., including interest payments) as a ratio to GDP. Provided that both exp_t and rev_t are I(1) with a cointegration relationship (1, −1), one should then test the linear restriction $\beta = 1$ in a regression model of the form:

$$rev_t = \alpha + \beta exp_t + \varepsilon_t \qquad (5)$$

where ε_t is an error term. In particular, following Quintos (1995), if exp_t and rev_t are cointegrated and $\beta = 1$, the government deficit would be strongly sustainable; whereas, if $0 < \beta < 1$, the deficit would be only weakly sustainable. Finally, if $\beta = 0$ the deficit would be unsustainable.

This framework has been applied in previous analyses of the sustainability of the Spanish government deficit, using data from 1964 to the second half of the 1990s; see Camarero, Esteve and Tamarit (1998), de Castro and Hernández de Cos (2002), or Escario (2005). A common result to these papers is the finding of an estimate of β positive but significantly different from one, indicating weak sustainability of the deficit in the sense of Quintos (1995); together with the finding of cointegration only in the presence of structural changes in the relationship, located around 1986-1988.

On the other hand, Bohn (2007) has recently provided a critique of tests of the IBC based on unit root and cointegration tests, on the grounds that such tests are incapable of rejecting sustainability, suggesting instead the estimation of error-correction-type policy reaction functions. However, the application of the more traditional procedure would give a sufficient condition for sustainability to hold, so, if the deficit were found to be sustainable, this would provide a stronger criterion to assess sustainability. In addition, this exercise will allow us to offer some extra evidence on the sustainability of the Spanish government deficit over a far more extended period[1].

Now, we are in position of testing for the sustainability of the Spanish budget deficit over the period 1850-2000, using the data on government expenditures and revenues examined in the previous section. As a first step of the analysis, we test for the order of integration of the variables exp_t and rev_t using the tests of Ng and Perron (2001). These authors propose using the tests statistics $\overline{MZ}_{\alpha}^{GLS}$ and $\overline{MZ}_{t}^{GLS}$, which are modified versions of the Z_{α} and Z_t Phillips-Perron tests; and ADF^{GLS}, a modified version of the Augmented Dickey-Fuller test. Such modifications improve the tests with regard to both size distortions and power. The results are shown in Table 1, and the null hypothesis of no stationarity cannot be rejected, independently of the test, for the two series in levels; at the same time that the presence of two unit roots is clearly rejected at the 1% significance level. Accordingly, both series are I(1).

[Table 1 here]

---

[1] In a companion paper (Bajo-Rubio, Díaz-Roldán and Esteve, 2010), an alternative test of the sustainability of the Spanish government deficit over the period 1850-2000 is performed along the lines of Bajo-Rubio, Díaz-Roldán and Esteve (2009), addressing Bohn's (2007) critique and emphasizing the role of monetary versus fiscal policy dominance.

Next, we perform a cointegration analysis of equation (5) over the whole sample. The estimation is made using the method of Dynamic Ordinary Least Squares (DOLS) of Stock and Watson (1993), following the methodology proposed by Shin (1994). This method has the advantage of providing a robust correction to the possible presence of endogeneity in the explanatory variables, as well as of serial correlation in the error terms of the OLS estimation. On the other hand, the DOLS estimates are asymptotically equivalent to those obtained from Johansen's (1988, 1991) maximum likelihood procedure, and would be more appropriate in this case since we have just two variables in the theoretical model[2]. Accordingly, we first estimate a long-run dynamic equation including leads and lags of the (first difference of the) explanatory variable in equation (5):

$$rev_t = \alpha + \beta exp_t + \sum_{j=-q}^{q} \varphi_j \, \Delta exp_{t-1-j} + v_t \qquad (6)$$

where v_t is an error term, and then perform Shin's (1994) test from the calculation of C_μ, a LM statistic from the DOLS residuals which tests for deterministic cointegration (i.e., when no trend is present in the regression).

The results of the estimation of equation (6) appear in Table 2. First of all, the null hypothesis of deterministic cointegration between rev_t and exp_t is not rejected at the 1% level of significance. In addition, the estimate of β is 0.88, significantly different from zero at the 1% level; but not significantly different from one at the 1% level, according to a Wald test on the null hypothesis $\hat{\beta} = 1$ against the alternative $\hat{\beta} < 1$, distributed as a χ_1^2 and denoted in the table by W_{DOLS}. Accordingly, the Spanish budget deficit has been strongly sustainable, according to Quintos's (1995) terminology, over the whole period 1850-2000. Thus, when extending the sample period over a 150-year span, the budget deficit turns out to be strongly, rather than weakly, sustainable, unlike previous evidence available from the mid-1960s on.

[Table 2 here]

---

[2]   Notice also that the DOLS estimates are preferable to Johansen's, according to Monte Carlo experiments in Stock and Watson (1993).

IV. Testing for structural change

Issues related to structural change have received a considerable amount of attention in the literature on statistics and econometrics; see Perron (2008) for a recent and far-reaching account of the problem of testing for multiple structural changes in linear regression models. Accounting for parameter shifts is crucial in cointegration analysis as far as long spans of data are involved, as in this paper, since they are more likely to be affected by structural breaks.

Recently, Kejriwal and Perron (2008, 2010) provide a comprehensive treatment of the problem of testing for multiple structural changes in cointegrated systems. Structural changes can manifest themselves through changes in the long-run relationship (6), either in the form of a change in the intercept, or in the slope of the cointegrating vector. In particular, the presence of structural changes affects the size of the residual-based test for the null hypothesis of cointegration by Shin (1994) when the true data generating process exhibits some type of structural break.

Accordingly, we test for the stability of equation (6) using the tests of Kejriwal and Perron (2008, 2010). These authors propose three types of test statistics to test for multiple breaks in cointegrated regression models:

a) First, a sup Wald test of the null hypothesis of no structural break ($m=0$) versus the alternative hypothesis that there are a fixed (arbitrary) number of breaks ($m=k$):

$$\sup F_T^*(k) = \sup_{\lambda \in \Lambda_\varepsilon} \frac{SSR_0 - SSR_k}{\hat{\sigma}^2}$$

where SSR_0 and SSR_k denote, respectively, the sums of squared residuals under the null hypothesis of no breaks, and under the alternative hypothesis of k breaks; $\lambda = \{\lambda_1, ..., \lambda_m\}$ is the vector of breaks fractions defined by $\lambda_i = T_i/T$ for $i=1,..., m$; and T_i are the break dates.

b) Second, a test of the null hypothesis of no structural break ($m=0$) versus the alternative hypothesis that there is an unknown number of breaks given some upper bound M ($1 \leq m \leq M$):

8

$$UD\max F_T^*(M) = \max_{1 \le k \le m} F_T^*(k)$$

c) In addition to the tests above, Kejriwal and Perron also propose a sequential procedure that not only enables detection of parameter instability but also allows a consistent estimation of the number of breaks, i.e., a sequential test of the null hypothesis of k breaks versus the alternative hypothesis of $k+1$ breaks:

$$F_T(k+1|k) = \max_{1 \le j \le k+1} \sup_{\tau \in \Lambda_{j,\varepsilon}} T\left\{SSR_T\left(\hat{T}_1,...,\hat{T}_k\right)\right\} - \left\{SSR_T\left(\hat{T}_1,...,\hat{T}_{j-1},\tau,\hat{T}_j,...,\hat{T}_k\right)\right\}/SSR_{k+1}$$

where $\Lambda_{j,\varepsilon} = \left\{\tau; \hat{T}_{j-1} + \left(\hat{T}_j - \hat{T}_{j-1}\right)\varepsilon \le \tau \le \hat{T}_j - \left(\hat{T}_j - \hat{T}_{j-1}\right)\varepsilon\right\}$, and the model with k breaks is obtained by a global minimization of the sum of squared residuals.

The results of applying the Kejriwal-Perron tests to the relationship between exp_t and rev_t are shown in Table 3. Both the intercept and the slope are allowed to change across regimes, and the level of trimming used is 15%, allowing up to five breaks. As can be seen, none of the tests proves to be significant and the sequential procedure selects no break point, which would point to a stable cointegrating relationship between the two variables over the whole period[3]. A possible explanation to the failure to find any structural change in the long-run relationship between exp_t and rev_t might be that the potential candidate years (i.e., the mid-1980s; see Section II and the previous evidence quoted in Section III) are located at the very end of the sample, leaving an insufficient number of observations available.

[Table 3 here]

V. Was fiscal policy non-linear?

As shown before, Spanish fiscal policy has been strongly sustainable over the long run, along the period 1850-2000. In addition, we find no evidence of any significant structural break over the whole period. In this section, we explore whether sustainability was attained due to the non-linear behaviour of fiscal authorities. In such a case, the adjustment to the long-run equilibrium occurs only when the deviation with respect to

---

[3] The results do not change significantly when using instead levels of trimming of 20% and 10%, allowing up to three and eight breaks, respectively.

this equilibrium exceeds a certain threshold; unlike the traditional approach, where it is assumed that the long-run equilibrium always exists. In particular, regarding the government deficit, if the behaviour of fiscal authorities were non-linear, they would adjust the deficit only when the difference between expenditures and revenues exceeded a certain threshold, thus ensuring the sustainability of the deficit in the long run. A theoretical justification of such behaviour is presented in Bertola and Drazen (1993) where, regarding the timing of fiscal actions, they 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 words, these authors introduce a non-linearity in the reaction function of policymakers.

However, the empirical literature on fiscal sustainability has hardly incorporated the new developments on non-linearities in fiscal policy. A first attempt is Cipollini (2001), who introduces 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. Bajo-Rubio, Díaz-Roldán and Esteve (2004) find 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 are different depending on whether the change in the deficit is below or above an endogenously estimated threshold. A similar analysis is applied to the US case by Arestis, Cipollini and Fattouh (2004). Bajo-Rubio, Díaz-Roldán and Esteve (2006) re-examine the long-run sustainability of the budget deficit in the Spanish case, finding that non-linear effects occur when the deficit is "big enough", which guarantees its long-run sustainability. More recently, a non-linear analysis on fiscal sustainability, using smooth transition autoregressive models, has been also applied to some EMU countries by Arghyrou and Luintel (2007), who find that fiscal disequilibria adjust in a non-linear way to their long-run equilibrium in the countries analysed.

In this section, we will provide some additional evidence on the sustainability of budget deficits, when fiscal policy is conducted as a non-linear process. However, unlike the above papers, we will make use of a much longer sample period, extending over 150 years, which should provide some more robustness to our empirical results. In

particular, we will try to quantify the "trigger point"; i.e., the precise moment in which the authorities correct a deficit, so that a required fiscal adjustment is finally enforced to ensure the deficit sustainability in the long run. To that end, we will follow Bertola and Drazen's (1993) theoretical contribution, making use of recent developments on threshold cointegration that consider the possibility of a non-linear relationship between government expenditures and revenues.

The concept of threshold cointegration is introduced by Balke and Fomby (1997) as a feasible way to combine non-linearity and cointegration. The traditional approach assumes that the error-correction model (ECM), characterizing systems in which variables are cointegrated, is present in every time period. On the contrary, Balke and Fomby (1997) show threshold cointegration as a possibility of a discrete adjustment, due to the presence of some adjustment costs on the side of economic agents. This type of discrete adjustment is particularly useful to describe the behaviour of fiscal authorities, which only cut budget deficits when they were "too large", in order to meet the IBC. In this way, threshold cointegration captures the possible non-linearity of fiscal policy, in the sense that the error correction mechanism should be expected only when a certain threshold is reached.

When testing for threshold cointegration, Balke and Fomby (1997) propose applying several univariate tests previously developed in the literature to the known cointegrated residual (i.e., the error-correction term). More recently, Hansen and Seo (2002) examine the case of an unknown cointegrating vector. In particular, these authors propose a vector ECM (VECM) with one cointegrating vector and a threshold effect based on the error-correction term, and develop an LM test for the presence of a threshold effect. This will be the approach followed in this paper.

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

$$\Delta x_t = \begin{cases} A_1^{'} X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) \leq \gamma \\ A_2^{'} X_{t-1}(\beta) + u_t & \text{if } w_{t-1}(\beta) > \gamma \end{cases} \tag{7}$$

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 that 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 (7) 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.

Next, Hansen and Seo (2002) propose 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 (7)). The first test is used when the true cointegrating vector is known *a priori*, and is denoted as:

$$
\sup\ LM^0 = \sup_{\gamma_L \leq \gamma \leq \gamma_U}\ LM(\beta_0, \gamma)
$$

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

$$
\sup\ LM = \sup_{\gamma_L \leq \gamma \leq \gamma_U}\ LM(\tilde{\beta}, \gamma)
$$

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 $\tilde{w}_{t-1}$ and γ_U is the $(1 - \pi_0)$ percentile.

We have applied the tests $\sup\ LM^0$ (for a given $\beta = 1$) and $\sup\ LM$ (for an estimated β). In both cases, as proposed by Hansen and Seo (2002), the p-values are calculated using a parametric bootstrap method (with 5,000 simulation replications). To select the lag length of the VAR, we have used the Akaike and Bayesian information criteria, both of them leading to $l = 2$. The results are reported in Table 4. Threshold cointegration appears at the 1% significance level when β is fixed at unity and when β is estimated rather than fixed, so that the null hypothesis of linear cointegration are

strongly rejected in both cases. According to the first test, the estimated threshold is $\tilde{\gamma} = 4.38$, with the error-correction term defined as $w_t = exp_t - rev_t$ (i.e., the budget deficit). The results from the second test are quite similar, with β estimated at 1.01 and an estimated threshold of $\tilde{\gamma} = 4.52$, or 4.64 in terms of the budget deficit (with rev_t computed at its average value over the sample period), where the error-correction term is now defined as $w_t = exp_t - 1.01rev_t$.

[Table 4 here]

Hence, the first regime occurs when the government deficit as a ratio to GDP is above 4.38% or 4.64%, which would be the relatively unusual regime including around 11% of the observations (namely, the years 1869, 1870, 1915, 1917, 1919, 1921, 1941, 1943 to 1945, 1982, 1984 to 1986, and 1993 to 1994). In particular, the main consolidation efforts would appear:

-    following the so-called "Glorious Revolution" of 1868
-    at the time of the First World War and immediately after
-    in the first half of the 1940s, in the aftermath of the Spanish Civil War
-    at the end of the first half of the 1980s, immediately before joining the EU
-    and in the years 1993-94, after a strong (and short-lived) recession;

where the last two episodes were also identified in Bajo-Rubio, Díaz-Roldán and Esteve (2006). In turn, the second or usual regime, which applies to the remaining 89% of the observations, would occur when the government deficit as a ratio to GDP is below 4.38% or 4.64%.

The estimated two-regime threshold VAR (heteroskedasticity-consistent standard errors in parentheses) for the first case (i.e., for a given β = 1; the results are very similar for the second case, i.e., for an estimated β = 1.01) is:

$$\Delta exp_t = \begin{cases} \underset{(0.24)}{7.63} - \underset{(0.15)}{0.95}\,w_{t-1} - \underset{(0.20)}{1.00}\,\Delta exp_{t-1} + \underset{(0.22)}{0.93}\,\Delta exp_{t-2} - \underset{(0.19)}{0.34}\,\Delta rev_{t-1} \\ \quad + \underset{(0.21)}{0.72}\,\Delta rev_{t-2} + u_{1t}, \;\; w_{t-1} \geq 4.38 \\ -\underset{(0.11)}{0.04} + \underset{(0.075)}{0.29}\,w_{t-1} + \underset{(0.10)}{0.22}\,\Delta exp_{t-1} - \underset{(0.10)}{0.44}\,\Delta exp_{t-2} + \underset{(0.09)}{0.11}\,\Delta rev_{t-1} \\ \quad - \underset{(0.10)}{0.38}\,\Delta rev_{t-2} + u_{2t}, \;\; w_{t-1} < 4.38 \end{cases}$$

$$\Delta rev_t = \begin{cases} \underset{(0.23)}{4.17} - \underset{(0.14)}{0.61} w_{t-1} - \underset{(0.20)}{0.99} \Delta exp_{t-1} + \underset{(0.21)}{1.23} \Delta exp_{t-2} - \underset{(0.19)}{0.24} \Delta rev_{t-1} \\ \qquad + \underset{(0.20)}{1.13} \Delta rev_{t-2} + u_{1t}, \ w_{t-1} \geq 4.38 \\ \underset{(0.15)}{0.15} + \underset{(0.10)}{0.24} w_{t-1} + \underset{(0.13)}{0.07} \Delta exp_{t-1} - \underset{(0.14)}{0.30} \Delta exp_{t-2} + \underset{(0.12)}{0.45} \Delta rev_{t-1} \\ \qquad - \underset{(0.13)}{0.62} \Delta rev_{t-2} + u_{2t}, \ w_{t-1} < 4.38 \end{cases}$$

Significant error-correction effects appear in both regimes, but in the first regime the error-correction effects are larger in terms of significance and size of the estimated coefficients, i.e., when either government expenditures are well above revenues or government deficit is relatively high.

Figure 2 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 deficit) in the previous period, holding the other variables constant. As can be seen, if the deficit is "large" (i.e., greater than 4.38% of GDP or rightwards of the threshold), both expenditures and revenues decrease sharply with the size of the deficit; and, since the response of expenditures is larger than that of the revenues, the government deficit falls accordingly. However, for a "small" deficit (i.e., lower than 4.38% of GDP or leftwards of the threshold), the response of both expenditures and revenues is smaller.

[Figure 2 here]

Accordingly, the above evidence suggests that the Spanish fiscal policy has shown significant non-linear effects. In particular, budget deficits have shown a mean-reverting dynamic behaviour once the threshold was reached, which in turn has assured their long-run sustainability.

VI. Conclusions

In this paper, we have tried to provide some additional evidence on the long-run sustainability of the Spanish government deficit, using an extended data set covering 150 years. When assessing the long-run sustainability of fiscal policy, the econometric

procedures usually employed require a large number of observations, which is not the case in most tests of the IBC. In order to overcome this problem we have made use of historical series for the period 1850-2000, so that the longer than usual span of the data allows us to obtain more robust results on the fulfilling of the IBC than in most of previous analyses, both for Spain and elsewhere.

After estimating a cointegration relationship between government expenditures and revenues derived from the IBC, we have investigated the possible presence of structural changes following the new approach recently proposed by Kejriwal and Perron (2008, 2010), and we have checked the possible non-linear behaviour of fiscal authorities, using the procedure developed by Hansen and Seo (2002). The results show that the Spanish fiscal policy has been strongly sustainable over the long run, along the period analysed. Moreover, no evidence has been found of any significant structural break throughout the whole period, which would point to a stable cointegrating relationship between government expenditures and revenues. In addition, when exploring the possibility of a non-linear behaviour on the side of the fiscal authorities, we find that the Spanish fiscal policy has shown significant non-linear effects. In particular, budget deficits have shown a mean-reverting dynamic behaviour once their value exceeded around 4.5% of GDP, which in turn has assured their long-run sustainability. Notice, on the other hand, that the prevalence of budget deficits over most of the period (resulting from low levels of both revenues and expenditures) would not be at odds with their long-term sustainability, since their size did not tend to increase in relative terms along the period.

To conclude, our result on the long-run sustainability of the Spanish government deficit says nothing on financing issues. As pointed out elsewhere, "(T)he history of the Spanish public debt until 1978 mirrors the main features of less developed countries, confirming the slow modernization of the Spanish Ministry of Finance" (Comín, 1995, p. 532), with frequent rescheduling of the government debt, and inflationary financing of the deficit. During the 19th century and the first years of the 20th century, deficits were monetized through the sales of government bonds to the Bank of Spain. The system changed after 1917, when government bonds were sold instead to private buyers, mostly private banks, but with the particular feature that these bonds could be automatically pledged at the Bank of Spain, so that its effect on the monetization of the

deficit was not significantly different. Although this specific practice ended in the early 1960s, for most of the period government deficits were ultimately financed through seigniorage, leading to an inflationary trend; accordingly, monetary policy appeared to be subordinated to the needs of financing the budget deficit (Sabaté, Gadea and Escario, 2006; Bajo-Rubio, Díaz-Roldán and Esteve, 2010). Only in recent years, more orthodox practices of deficit financing became prevalent, in particular after the Treaty of Maastricht entered into force in 1993, so that the financing of budget deficits by the central banks was explicitly ruled out. These issues of fiscal sustainability and deficit financing merit further investigation in order to properly characterize the behaviour of the Spanish fiscal policy in a long-term perspective.

References

Arestis, Philip, Andrea Cipollini, and Bassam Fattouh (2004), Threshold effects in the U.S. budget deficit, *Economic Inquiry* **42**: 214-22.

Arghyrou, Michael G., and Kul B. Luintel (2007), Government solvency: Revisiting some EMU countries, *Journal of Macroeconomics* **29**: 387-410.

Bajo-Rubio, Oscar, Carmen Díaz-Roldán, and Vicente Esteve (2004), Searching for threshold effects in the evolution of budget deficits: An application to the Spanish case, *Economics Letters* **82**: 239-43.

Bajo-Rubio, Oscar, Carmen Díaz-Roldán, and Vicente Esteve (2006), Is the budget deficit sustainable when fiscal policy is nonlinear? The case of Spain, *Journal of Macroeconomics* **28**: 596-608.

Bajo-Rubio, Oscar, Carmen Díaz-Roldán, and Vicente Esteve (2009), Deficit sustainability and inflation in EMU: An analysis from the fiscal theory of the price level, *European Journal of Political Economy* **25**: 525-39.

Bajo-Rubio, Oscar, Carmen Díaz-Roldán, and Vicente Esteve (2010), Government deficit sustainability, and monetary versus fiscal dominance: The case of Spain, 1850-2000, Working Paper on International Economics and Finance 10-04, Asociación Española de Economía y Finanzas Internacionales (available at http://www.aeefi.com/RePEc/pdf/defi10-04-final.pdf).

Balke, Nathan S., and Thomas B. Fomby (1997), Threshold cointegration, *International Economic Review* **38**: 627-45.

Bertola, Giuseppe, and Allan Drazen (1993), Trigger points and budget cuts: Explaining the effects of fiscal austerity, *American Economic Review* **83**: 11-26.

Bohn, Henning (2007), Are stationarity and cointegration restrictions really necessary for the intertemporal budget constraint?, *Journal of Monetary Economics* **54**: 1837-47.

Camarero, Mariam, Vicente Esteve, and Cecilio Tamarit (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.

Cipollini, Andrea (2001), Testing for government intertemporal solvency: A smooth transition error correction model approach, *The Manchester School* **69**: 643-55.

Comín, Francisco (1995), Public finance in Spain during the 19th and 20th centuries, in P. Martín-Aceña and J. Simpson, eds., *The economic development of Spain since 1870*, Aldershot, Edward Elgar, 521-60.

Comín, Francisco (1996), *Historia de la Hacienda Pública, II. España (1808-1995)*, Barcelona, Crítica.

Comín, Francisco, and Daniel Díaz (2005), Sector público administrativo y estado del bienestar, in A. Carreras and X. Tafunell, eds., *Estadísticas históricas de España: Siglos XIX y XX* (2nd edition), Bilbao, Fundación BBVA, 873-964.

de Castro, Francisco, and Pablo Hernández de Cos (2002), On the sustainability of the Spanish public budget performance, *Hacienda Pública Española/Revista de Economía Pública* **160**: 9-27.

Escario, Regina (2005), Sostenibilidad del déficit público. Un análisis desagregado 1964-1998, *Cuadernos Aragoneses de Economía* **15**: 117-36.

Hamilton, James D., and Marjorie A. Flavin (1986), On the limitations of government borrowing: A framework for empirical testing, *American Economic Review* **76**: 808-19.

Hansen, Bruce E., and Byeongseon Seo (2002), Testing for two-regime threshold cointegration in vector error-correction models, *Journal of Econometrics* **110**: 293-318.

Haug, Alfred A. (1995), Has federal budget deficit policy changed in recent years?, *Economic Inquiry* **33**: 104-18.

Johansen, Søren (1988), Statistical analysis of cointegration vectors, *Journal of Economic Dynamics and Control* **12**: 231-54.

Johansen, Søren (1991), Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models, *Econometrica* **59**: 1551-80.

Kejriwal, Mohitosh, and Pierre Perron (2008), The limit distribution of the estimates in cointegrated regression models with multiple structural changes, *Journal of Econometrics* **146**: 59-73.

Kejriwal, Mohitosh, and Pierre Perron (2010), Testing for multiple structural changes in cointegrated regression models, forthcoming in *Journal of Business and Economic Statistics* (doi:10.1198/jbes.2009.07220).

Martin, Gael M. (2000), US deficit sustainability: A new approach based on multiple endogenous breaks, *Journal of Applied Econometrics* **15**: 83-105.

Newey, Whitney K., and Kenneth D. West (1987), A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix, *Econometrica* **55**: 703-08.

Ng, Serena, and Pierre Perron (2001), Lag length selection and the construction of unit root tests with good size and power, *Econometrica* **69**: 1519-54.

Perron, Pierre (2008), Structural change, Econometrics of, in S. N. Durlauf and L. E. Blume, eds., *The New Palgrave Dictionary of Economics* (2nd edition), Vol. 8, Basingstoke, Palgrave Macmillan, 55-64.

Perron, Pierre, and Serena Ng (1996), Useful modifications to some unit root tests with dependent errors and their local asymptotic properties, *Review of Economic Studies* **63**: 435-63.

Prados de la Escosura, Leandro (2003), *El progreso económico de España (1850-2000)*, Bilbao, Fundación BBVA.

Quintos, Carmela E. (1995), Sustainability of the deficit process with structural shifts, *Journal of Business and Economic Statistics* **13**: 409-17.

Sabaté, Marcela, María Dolores Gadea, and Regina Escario (2006), Does fiscal policy influence monetary policy? The case of Spain, 1874-1935, *Explorations in Economic History* **43**: 309-31.

Shin, Yongcheol (1994), A residual-based test of the null of cointegration against the alternative of no cointegration, *Econometric Theory* **10**: 91-115.

Smith, Gregor W., and Stanley E. Zin (1991), Persistent deficits and the market value of government debt, *Journal of Applied Econometrics* **6**: 31-44.

Stock, James H., and Mark W. Watson (1993), A simple estimator of cointegrating vectors in higher order integrated systems, *Econometrica* **61**: 783-820.

Trehan, Bharat, and Carl E. Walsh (1988), Common trends, the government's budget constraint, and revenue smoothing, *Journal of Economic Dynamics and Control* **12**: 425-44.

Trehan, Bharat, and Carl 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**: 206-23.

Figure 1
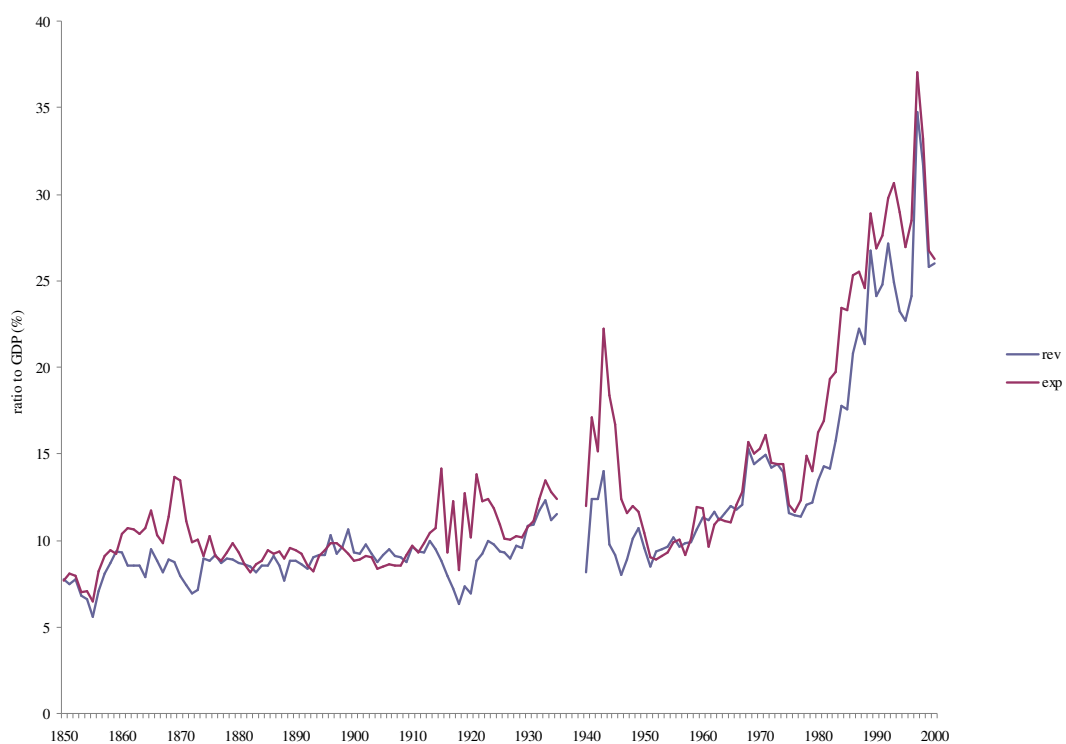## Spanish central government expenditures and revenues, 1850-2000

Figure 2
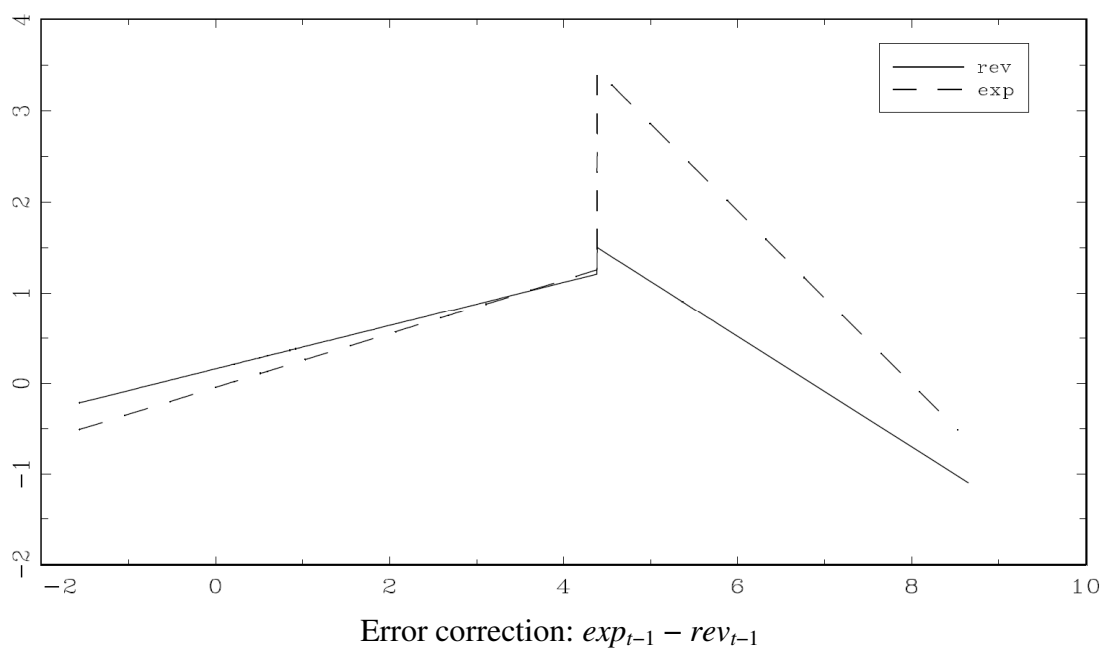## Response of expenditures and revenues to error correction

Error correction: $exp_{t-1} - rev_{t-1}$

Table 1
Ng-Perron tests for unit roots

I(2) vs. I(1)

	$\overline{MZ}_{\alpha}^{GLS}$	$\overline{MZ}_{t}^{GLS}$	ADF^{GLS}
Δrev_t	-72.03^*	-6.00^*	-13.03^*
Δexp_t	-62.46^*	-5.58^*	-8.54^*

I(1) vs. I(0)

	$\overline{MZ}_{\alpha}^{GLS}$	$\overline{MZ}_{t}^{GLS}$	ADF^{GLS}
rev_t	-8.18	-1.82	-1.83
exp_t	-7.59	-1.85	-1.87

Notes:
(i) * denotes significance at the 1% level. The critical values are taken from Ng and Perron (2001), Table 1.
(ii)   The autoregressive truncation lag has been selected using the modified Akaike information criterion, as proposed by Perron and Ng (1996).

Table 2
Estimation of the long-run relationship:
Stock-Watson-Shin cointegration tests

α	β	C_{μ}	W_{DOLS}
0.19	0.88	0.144	1.20
(0.67)	(7.57)		

Notes:
(i) t-statistics in parentheses.
(ii)   The C_{μ} statistic is not significant at the conventional levels. The critical values are taken from Shin (1994), Table 1, for m=1.
(iii)   The number of leads and lags selected was q=5≈$INT(T^{1/3})$, as proposed in Stock and Watson (1993). The long-run variance of the cointegrating regression residuals has been estimated using the Bartlett window with l=12≈$INT(T^{1/2})$, as proposed in Newey and West (1987).

Table 3
Kejriwal-Perron tests for structural change

sup $F_T(1)$	sup $F_T(2)$	sup $F_T(3)$	UD max	Number of breaks selected
9.18	5.80	4.69	9.18	0

Note: No test statistic is significant at the conventional levels. The critical values are taken from Kejriwal and Perron (2010), Table 1.10, trending case.

Table 4
Hansen-Seo tests for threshold cointegration

	sup LM^0	sup LM
Test statistic value	28.69	29.45
Calculated p-values	0.004	0.001
Threshold parameter	4.38	4.52
Estimate of the cointegrating vector	1.00	1.01