US deficit sustainability revisited: A multiple structural change approach*

Oscar Bajo-Rubio

Universidad de Castilla-La Mancha
and Instituto de Estudios Fiscales

Carmen Díaz-Roldán

Universidad de Castilla-La Mancha

Vicente Esteve

Universidad de Valencia

September 2005

Abstract

In this paper we re-examine the long-run sustainability of US budget deficits, using Bai and Perron's multiple structural change approach. While the deficit would have been weakly sustainable over the full sample (1947:1-2004:3), strong sustainability would appear only between 1982:1 and 1995:4.

Keywords: Fiscal policy; Sustainability; US budget deficit.

JEL classification: E62, H62.

*The authors acknowledge 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).

1 Introduction

The US government finances have experienced a remarkable turnaround in recent years, with large budget deficits in the 1980s and early 1990s. This has led to a substantial amount of empirical work aimed to examining their long-run sustainability. Later on, the record surpluses in the late 1990s and early 2000s become record deficits after 2002, with budget projections showing large federal deficits over the next decade. As a result, the US general government deficit is now among the highest in the OECD, and its sustainability has become again a highly relevant issue.

When analyzing the sustainability of budget deficits, the traditional approach has consisted of testing whether the government's intertemporal budget constraint (IBC) holds, that is, whether the current market value of debt equals the discounted sum of expected future surpluses. However, empirical tests on sustainability are largely 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, focusing on either the univariate properties of the government's deficit and debt (Hamilton and Flavin, 1986; Wilcox, 1989), or the presence of a long-run cointegration relationship between government revenues and expenditures (Trehan and Walsh, 1988, 1991; Haug, 1991; Smith and Zin, 1991). Further, the eventual occurrence of a structural break in the cointegrating relationship has been examined in Hakkio and Rush (1991), who assumed that the break point was exogenously given; and in Haug (1995), Quintos (1995), and Martin (2000), where the break point was endogenously derived. Overall, the results of these studies suggest that the US deficit would have undergone a shift in recent times, with the deficit being either unsustainable or only weakly sustainable in the post-break period.

In this paper we re-examine the sustainability of US budget deficits, using a new approach developed by Bai and Perron (1998, 2003a). This procedure allows to test endogenously for the presence of multiple structural changes in an estimated relationship, and has a number of advantages over previous approaches. In particular, the underlying assumptions are less restrictive, confidence intervals for the break dates can be calculated, the data and errors are allowed to follow different distributions across segments, and the sequential method used in the application can allow for the presence of serial correlation in the errors and heterogeneous variances across segments; see Bai and Perron (2005) for details. As a further contribution, as compared with previous studies on the subject [e.g., Quintos (1995) or Martin (2000), where the sample ends at 1992.3], our period of analysis includes the most recent developments in the evolution of the US budget deficit, extending from 1947:1 to 2004:3.

The rest of the paper is organized as follows. A brief description of the underlying theoretical framework is provided in section 2, the methodology and empirical results are presented in section 3, and the main conclusions are summarized in section 4.

2 Theoretical framework

Assuming that budget deficits are financed using bonds of one-period maturity, in any single period a government faces the following budget constraint:

$$\Delta B_t = G_t - R_t \tag{1}$$

where B_t, G_t, and R_t denote, respectively, the real market value of government debt, real government expenditure inclusive of interest payments, and real tax revenues. Hence, $G_t - R_t$ defines the total (i.e., real interest inclusive) budget deficit, with $G_t = GE_t + r_t B_{t-1}$, being GE_t the real government expenditure exclusive of interest payments, and r_t the one-period real interest rate.

The interest rate r_t is assumed to be stationary around a mean r so that, defining $EXP_t = GE_t + (r_t - r)B_{t-1}$, the constraint (1) can be written as:

$$B_t = EXP_t - R_t + (1+r)B_{t-1} \tag{2}$$

And, since (2) holds every period, solving for B_t and iterating forward over an infinite horizon yields the IBC:

$$B_t = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^{j+1} (R_{t+j+1} - EXP_{t+j+1}) + \lim_{j\to\infty} \left(\frac{1}{1+r}\right)^{j+1} B_{t+j+1} \tag{3}$$

Then, defining E_t as the expectations operator, conditional on information at time t, fiscal sustainability involves:

$$\lim_{j\to\infty} \left(\frac{1}{1+r}\right)^{j+1} E_t B_{t+j+1} = 0 \tag{4}$$

which implies that the current value of outstanding government debt is equal to the present value of future budget surpluses. In other words, the budget deficit would be sustainable if and only if the stock of debt is expected to grow no faster on average than the mean real rate of interest, the latter taken as a proxy of the growth rate of the economy.

The cointegration framework for testing the IBC would appear once first differences are taken in (3):

$$\begin{aligned} \Delta B_t &= G_t - R_t = \sum_{j=0}^{\infty} \left(\frac{1}{1+r}\right)^{j+1} (\Delta R_{t+j+1} - \Delta EXP_{t+j+1}) \\ &\quad + \lim_{j\to\infty} \left(\frac{1}{1+r}\right)^{j+1} \Delta B_{t+j+1} \end{aligned} \tag{5}$$

so that sustainability would require:

$$\lim_{j\to\infty} E_t \left(\frac{1}{1+r}\right)^{j+1} \Delta B_{t+j+1} = 0 \tag{6}$$

Under a no-Ponzi scheme rule, the IBC imposes some restrictions on the time series properties of government revenues and expenditures. These restrictions follow from the specification of the right-hand side of equation (5), which will be stationary as long as government revenues and expenditures, and the stock of public debt, are all stationary in first differences. In order to test for condition (6), the usual procedure consists of testing for the stationarity of $G_t - R_t$, provided that both of them are I(1), with a cointegration relationship $(1, -1)$, in a regression model of the form:

$$R_t = \alpha + \beta G_t + \varepsilon_t \tag{7}$$

and then testing the linear restriction $\beta = 1$. From here, Quintos (1995) shows that:

(i) The fiscal deficit is strongly sustainable if and only if R_t and G_t are cointegrated and $\beta = 1$.

(ii) The fiscal deficit is only weakly sustainable if R_t and G_t are cointegrated and $0 < \beta < 1$.

(iii) The fiscal deficit is unsustainable if $\beta \leq 0$.

3 Methodology and empirical results

In this section we provide a test of the sustainability of the US budget deficit, over the period 1947:1 to 2004:3. The data on federal government revenues and expenditures, inclusive of interest paid on debt, are taken from the National Income Product Accounts (NIPA, Table 3.1), and real values are calculated using the GDP deflator (NIPA, Table 1.1.4).

As a first step of the analysis, we test for the order of integration of the series. To that end, we use 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 more conventional tests for unit roots. Table 1 show the results of the three tests, $\bar{M}Z_{\alpha}^{GLS}$, $\bar{M}Z_{t}^{GLS}$, and ADF^{GLS}. As shown in the table, the null hypothesis of non stationarity cannot be rejected, independently of the test, for the two series in levels; and the presence of two unit roots is clearly rejected at the 1% significance level. Accordingly, both series would be concluded to be I(1).

Next, we perform a cointegration analysis of equation (7) over the whole sample, with no breaks included. The estimation is made using the method of Dynamic Ordinary Least Squares (DOLS), which provides 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. This method, proposed by Stock and Watson (1993) and extended by Shin (1994), is implemented in two stages. The first step involves the estimation of a long-run dynamic equation including leads and lags of the explanatory variables in equation (7), i.e., the so-called DOLS regression:

$$R_t = \alpha + \beta G_t + \sum_{j=-q}^{q} \gamma_j \Delta G_{t-j} + \upsilon_t \qquad (8)$$

where υ_t is an error term. Then, in a second step, Shin's test is performed 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 from the DOLS estimation and the Shin test are reported in the first column of Table 2. The null of deterministic cointegration between R_t and G_t is not rejected at the 1% level of significance, and the estimated value for β is 0.93, significantly different from zero at the 1% level. However, this estimate would be 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 by W_{DOLS} in Table 2. Accordingly, since R_t and G_t would be cointegrated and $0 < \hat{\beta} < 1$, the US fiscal deficit would be only weakly sustainable over the full sample, which would confirm, over a more extended sample, previous results by, e.g., Quintos (1995) and Martin (2000).

The main objective of this section, though, is estimating equation (7) through a multiple endogenous break model, making use of the approach of Bai and Perron (1998, 2003a). As a key feature, Bai and Perron's procedure allows testing for multiple breaks at unknown dates, so that each break point is successively estimated by using a specific-to-general strategy in order to determine consistently the number of breaks. More specifically, Bai and Perron (1998, 2003a) propose three methods to determine the number of breaks: a sequential procedure, SP (Bai and Perron, 1998); the Schwarz modified criterion, LWZ (Liu, Wu and Zidek, 1997); and the Bayesian information criterion, BIC (Yao, 1988). Also, the authors suggest several statistics in order to identify the break points:

- The $\sup F_T(k)$ test, i.e., a $\sup F$-type test of the null hypothesis of no structural break *versus* the alternative of a fixed (arbitrary) number of breaks k.

- Two maximum tests of the null hypothesis of no structural break *versus* the alternative of an unknown number of breaks given some upper bound, i.e., $UD\max$ test, an equal weighted version, and $WD\max$ test, with weights that depend on the number of regressors and the significance level of the test.

- The $\sup F_T(l+1|l)$ test, i.e., a sequential test of the null hypothesis of l breaks *versus* the alternative of $l+1$ breaks.

The results of applying the Bai-Perron tests to the relationship between R_t and G_t, allowing up to 5 breaks, are shown in Table 3. Both the $UD\max$ and $WD\max$ tests are highly significant, which implies that at least one break is present. Next, all the $\sup F_T(k)$ tests are significant, with k running between 1 and 5, so that at least one break would be present in this relationship. In turn, the $\sup F_T(l+1|l)$ test is not significant for any $l \geq 3$, so the sequential procedure selects three breaks. Hence, the results of the Bai-Perron tests would suggest a model of four regimes, with the dates of the breaks estimated at 1955:2, 1982:1, and 1996:1; their confidence intervals are shown in Table 3.

Finally, once the three break dates have been identified by means of the Bai and Perron procedure, we proceed to estimate the cointegration equation (8) for the four sub-samples, and the results are shown in the last four columns of Table 2. As can be seen, in the first and second regimes (1947:1-1955:1 and 1955:2-1981:4) the null of deterministic cointegration is not rejected at the 1% level, and the restriction on the estimate of β being equal to one is clearly rejected, which implies that the US budget deficit would have been only weakly sustainable as in the whole sample. In turn, in the third regime (1982:1-1995:4) the null of deterministic cointegration is again not rejected at the 1% level, but now the estimate of β would not be significantly different from one according to the Wald test, so that the US budget deficit would have been strongly sustainable during this period. Finally, in the fourth regime (1996:1-2004:3) no long-run relationship between public revenues and expenditures would appear, since the null of deterministic cointegration is now rejected at the 10% level, and with an estimate of β well above one, reflecting the fact that the US budget deficit would have registered a large surplus during an important part of this period. The above results are summarized in Table 4.

4 Conclusions

In this paper we have re-examined the long-run sustainability of US budget deficits, using the multiple structural change approach of Bai and Perron (1998, 2003a). We found evidence of weak sustainability of the deficit over the full sample (1947:1-2004:3), a result in line with previous findings by Quintos (1995) and Martin (2000) for the period ending at 1992.3. Further, we have detected up to three breaks (estimated at 1955:2, 1982:1, and 1996:1) along the whole sample period, so that the US budget deficit would have been strongly sustainable only in the third regime (1982:1-1995:4), weakly sustainable in the first and second regimes (1947:1-1955-1 and 1955:2-1981:4, respectively), and a surplus would have prevailed over the final regime (1996:1-2004:3).

References

[1] Bai, J. and Perron, P. (1998): "Estimating and testing linear models with multiple structural changes", *Econometrica,* 66, 47-78.

[2] Bai, J. and Perron, P. (2003a): "Computation and analysis of multiple structural change models", *Journal of Applied Econometrics*, 18, 1-22.

[3] Bai, J. and Perron, P. (2003b): "Critical values for multiple structural change tests", *Econometrics Journal*, 6, 72-78.

[4] Bai, J. and Perron, P. (2005): "Multiple structural change models: A simulation analysis", in P.D. Corbae, S.N. Durlauf and B.E. Hansen (eds.): *Econometric Theory and Practice. Frontiers of Analysis and Applied Research*, Cambridge University Press, forthcoming.

[5] Hakkio, C.S. and Rush, M. (1991): "Is the budget deficit 'too large'?", *Economic Inquiry*, 29, 429-445.

[6] Hamilton, J.D. and Flavin, M.A. (1986): "On the limitations of government borrowing: A framework for empirical testing", *American Economic Review*, 76, 808-819.

[7] Haug, A.A. (1991): "Cointegration and government borrowing constraints: Evidence for the U.S.", *Journal of Business and Economic Statistics*, 9, 97-101.

[8] Haug, A.A. (1995): "Has the federal deficit policy changed in recent years?", *Economic Inquiry*, 33, 104-118.

[9] Liu, J., Wu, S. and Zidek, J.V. (1997): "On segmented multivariate regressions", *Statistica Sinica*, 7, 497-525.

[10] Martin, G. (2000): "US deficit sustainability: A new approach based on multiple endogenous breaks", *Journal of Applied Econometrics*, 15, 83-105.

[11] Newey, W.K. and West, K.D. (1987): "A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix", *Econometrica*, 55, 703-708.

[12] Ng, S. and Perron, P. (2001): "Lag length selection and the construction of unit root tests with good size and power", *Econometrica,* 69, 1529-1554.

[13] Perron, P. and Ng, S. (1996): "Useful modifications to some unit root test with dependent errors and their local asymptotic properties", *Review of Economic Studies,* 63, 435-465.

[14] Quintos, C.E. (1995): "Sustainability of the deficit process with structural shifts", *Journal of Business and Economic Statistics*, 13, 409-417.

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

[16] Smith, G.W. and Zin, S.E. (1991): "Persistent deficits and the market value of government debt", *Journal of Applied Econometrics*, 6, 31-44.

[17] Stock, J.H. and Watson, M.W. (1993): "A simple estimator of cointegrating vectors in higher order integrated systems", *Econometrica,* 61, 783-820.

[18] Trehan, B. and Walsh, C.E. (1988): "Common trends, the government's budget balance, and revenue smoothing", *Journal of Economic Dynamics and Control*, 12, 425-444.

[19] Trehan, B. and Walsh, C.E. (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.

[20] 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.

[21] Yao, Y.-C. (1988): "Estimating the number of change-points via Schwarz's criterion", *Statistics and Probability Letters*, 6, 181-189.

Table 1
Ng-Perron tests of unit roots

I(2) *vs.* I(1)	Case: $p = 0, \bar{c} = -7.0$		
Variable	MZ_{α}^{GLS}	MZ_{t}^{GLS}	ADF^{GLS}
ΔR_t	-41.76^{*}	-4.56^{*}	-6.00^{*}
ΔG_t	-63.15^{*}	-5.61^{*}	-7.37^{*}

I(1) *vs.* I(0)	Case: $p = 1, \bar{c} = -13.5$		
Variable	MZ_{α}^{GLS}	MZ_{t}^{GLS}	ADF^{GLS}
R_t	-1.46	-0.74	-0.75
G_t	0.75	0.52	0.50

Notes:

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

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

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

Parameter estimates	Full sample 1947:1-2004:3	First regime 1947:1-1955:1	Second regime 1955:2-1981:4	Third regime 1982:1-1995:4	Fourth regime 1996:1-2004:3
α	0.36	1.02	0.30	-0.54	-19.1
	(2.42)	(2.53)	(2.15)	(-0.61)	(-16.8)
β	0.93	0.82	0.95	1.05	3.41
	(47.4)	(12.59)	(48.00)	(9.40)	(23.8)
$\bar{R}^2$	0.99	0.91	0.99	0.97	0.98
$\hat{\sigma}^2$	0.045	0.056	0.035	0.031	0.010
C_μ	0.087	0.088	0.062	0.072	0.265^{***}
W_{DOLS}	10.15^{*}	6.98^{*}	6.29^{**}	0.23	283.31^{*}

Notes:

[a] *, **, and *** denote significance at the 1%, 5%, and 10% levels, respectively. The critical values for the Shin test are taken from Shin (1994), Table 1, for $m = 1$.

[b] t-statistics in parentheses.

[c] The number of leads and lags selected was $q = 3 \simeq INT\left(T^{1/3}\right)$, as proposed in Stock and Watson (1993). The long-run variance of the cointegrating regression residuals was estimated using the Bartlett window with $l = 5 \simeq INT\left(T^{1/2}\right)$, as proposed in Newey and West (1987).

Table 3
Bai-Perron tests of multiple structural changes in the long-run relationship

Tests statistics:

$UD\max$ 117.82*	$WD\max$ 123.76*			
$\sup F_T(1)$ 117.82*	$\sup F_T(2)$ 83.04*	$\sup F_T(3)$ 82.69*	$\sup F_T(4)$ 62.82*	$\sup F_T(5)$ 50.27*
$\sup F_T(2\|1)$ 38.47*	$\sup F_T(3\|2)$ 58.69*	$\sup F_T(4\|3)$ 5.01	$\sup F_T(5\|4)$ 0.0	

Break dates estimates:

T_1	1955:2
	[1954:4-1957:2]
T_2	1982:1
	[1981:1-1982:2]
T_3	1996:1
	[1995:4-1996:2]

Notes:

[a] * denotes significance at the 1% level. The critical values are taken from Bai and Perron (1998), tables I and II; and from Bai and Perron (2003b), tables 1 and 2.

[b] The number of breaks (in our case, three) has been determined according to the sequential procedure of Bai and Perron (1998), at the 5% size for the sequential test $\sup F_T(l+1|l)$.

[c] 95% confidence intervals in brackets.

Table 4
Sustainability of the US public deficit: Summary results

	Full sample 1947:1-2004:3	First regime 1947:1-1955:1	Second regime 1955:2-1981:4	Third regime 1982:1-1995:4	Fourth regime 1996:1-2004:3
Cointegration	Yes	Yes	Yes	Yes	No
Estimate of β	0.93	0.82	0.95	1.05	3.41
Null $\hat{\beta} = 1$	No	No	No	Yes	No
Sustainability	Yes (weak)	Yes (weak)	Yes (weak)	Yes (strong)	–