Journal of Business & Economic Statistics, October 1995, Vol. 13, No. 4

Sustainability of the Deficit Process With Structural Shifts

Carmela E. Quintos
John M. Olin School of Business, Washington University, St. Louis, MO 63130-4899

This article discusses the condition for deficit sustainability and searches for shifts in the structure of U.S. deficit policy. I show that, contrary to the current literature, cointegration between revenues and expenditures inclusive of debt payment is not a necessary but a sufficient condition for a strict interpretation of deficit sustainability. The necessary condition requires that debt grow slower than the borrowing rate. For tests on structural shifts, I use a test for structural change that searches for shifts in the rank of the cointegrating matrix. I find that there is a shift in deficit policy in the early 80s so that cointegration between revenue and expenditure inclusive of interest payments holds only up to 1980. I show, however, that the deficit process is still sustainable despite the failure of cointegration in the 80s. The finding of shifts in the deficit process is more uncertain when I normalize by gross national product or population, but the deficit is shown to be sustainable regardless of whether or not structural breaks exist in the data.

KEY WORDS: Cointegration; Government debt; Structural change; Unit roots.

Recent concern over the government deficit has raised the issue of sustainability—that is, whether the federal government can continue to operate under its current fiscal policy indefinitely. Fiscal policy is constrained by the need to finance the deficit, and virtually any pattern of deficits would be sustainable if it were possible to borrow without restraint and to finance the interest on debt by additional borrowing. But governments that operate in dynamically efficient economies have limits to borrowing and face a present-value borrowing constraint. In other words, a government that operates in a dynamically efficient economy balances its budget intertemporally by setting the current market value of debt equal to the discounted sum of expected future surpluses. A violation of intertemporal budget balance would indicate that the fiscal policy cannot be sustained forever because the value of debt would explode over time at a rate faster than the growth rate of the economy. Thus a sustainable fiscal policy is one that would cause the discounted value of debt to go to 0 at the limit so that the present-value borrowing constraint would hold.

Hamilton and Flavin (1986) investigated the testable implications of the present-value budget constraint. Their framework for testing suggested that the stationarity of the discounted debt would indicate a sustainable deficit policy. Using U.S. data, they found stationarity of the debt processes and concluded that there is no violation of the government's borrowing constraint. Later work in the area found that results on deficit sustainability depend on the existence of structural breaks in the deficit process. For example, Wilcox (1989) extended Hamilton and Flavin's analysis over the period 1960–1984 and found strong evidence of a shift in the structure of fiscal policy. For the period prior to 1974, he found no evidence of a violation of the borrowing constraint, but the stationarity of the discounted deficit failed to hold for the post-1974 period.

Hakkio and Rush (1991), Haug (1991), Smith and Zin (1991), and Trehan and Walsh (1988, 1991) developed an alternative framework to testing the borrowing constraint. Their condition for deficit sustainability was that revenues and expenditures inclusive of interest payments be cointegrated if each are I(1) processes. Trehan and Walsh and Haug supported Hamilton and Flavin's results for the full sample period. Imposing breaks, Hakkio and Rush obtained results similar to those of Wilcox. Using the sample period 1950–1988, they found cointegration between revenues and expenditures in the earlier years but no cointegration in the years starting from the mid-70s. They interpreted this result to mean that the deficit has become a problem only in recent years and is not sustainable.

The articles by Hakkio and Rush and Wilcox discussed previously chose the structural breaks in the model by exogenously imposing the break date on the data. An alternative study by Haug (1992) endogenously picked the break date using parameter stability tests developed by Hansen (1992) and found that there are no breaks in the data. In this article, I endogenously pick the break, as did Haug (1992), using a sequential Chow test with I(1) processes, and I find break dates that are surprisingly close to those used by Hakkio and Rush and Wilcox. My results support Hakkio and Rush and Wilcox's conclusion that cointegration holds for only the pre-break period. I show, however, that this does not mean that the deficit is not sustainable in the post-break period if one considers a "weaker" condition for deficit sustainability.

Thus this article extends the empirical literature on deficit sustainability in two ways. First, I introduce "strong" and "weak" conditions for deficit sustainability to hold. The "strong" requirement corresponds to Hamilton and Flavin's necessary and sufficient condition that the debt process be stationary for the bubble term to go to 0. Alternatively, this "strong" requirement also pertains to Trehan and Walsh's necessary and sufficient condition that revenue and expenditures be cointegrated. The weaker condition that I introduce in this article allows the bubble term to go to 0 at a rate slower than the stronger version. With this weaker condition, cointegration is only a *sufficient* condition with the necessary and sufficient condition being that the debt process grow slower than the growth rate of mean interest rates. Indeed, the deficit process can be integrated or even mildly explosive and the deficit will still be sustainable as long as the growth rate of debt does not exceed the growth rate of the economy. Although a debt process that satisfies the weaker condition is said to be sustainable, I show that it has serious policy implications because the government will have difficulty in marketing its debt in the long run. Contrary to Hakkio and Rush's results, I show that the undiscounted debt in recent years is mildly explosive but is still sustainable (though in the weaker sense).

Second, I search for structural shifts in the *rank* of the cointegrating matrix as opposed to the parameters of the cointegrating vector (as considered, for example, by Haug 1992). The difference is that the former searches for shifts in the long-run relationship itself as opposed to the shifts in the parameters that characterize the long-run relationship. I test the null hypothesis of cointegration for the full period against the alternative that cointegration holds for only part of the period using tests developed by Quintos (1994a). Such tests are called *rank constancy tests* because the rank of the cointegrating matrix (or the number of cointegrating relationships) changes over time.

Section 1 outlines the model of the government's intertemporal budget constraint under consideration and states the condition for deficit sustainability. Section 2 contains the formulation of the rank constancy test and its limiting distribution, and Section 3 contains the empirical application of the test to U.S. data. Section 4 contains the conclusions.

Some words on notation are required. The rate at which a stochastic sequence converges to a nonstochastic sequence in probability is denoted in the usual notation as $Op(.)$. The symbol $\xrightarrow{d}$ signifies convergence in distribution, the notation $[m]$ denotes the largest integer that does not exceed m, and the symbol $\rho(\alpha)$ denotes the rank of some matrix α.

1. THE MODEL

I construct the dynamic budget constraint following Hakkio and Rush (1991). The government's one-period budget constraint is given by

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

where B_t is the market value of federal debt, $G_t^r = G_t + r_t B_{t-1}$ is government expenditure inclusive of interest payment, and R_t denotes tax revenues. The interest rate r_t is assumed to be stationary around the mean r and (1) can be written as

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

where $E_t = G_t + (r_t - r)B_{t-1}$ is G_t^r with interest rates taken around a zero mean. Because (2) holds for every period, forward substitution yields

$$B_t = \sum_{j=0}^{\infty} \gamma^{j+1}(R_{t+j} - E_{t+j}) + \lim_{j\to\infty} \gamma^{j+1} B_{t+j}, \tag{3}$$

where I have set $\gamma = (1+r)^{-1}$. The representation of (3) in terms of the difference ΔB_t will be used later in empirical testing and is written as

$$G_t^r - R_t = \sum_{j=0}^{\infty} \gamma^{j-1}(\Delta R_{t+j} - \Delta E_{t+j}) + \lim_{j\to\infty} \gamma^{j+1} \Delta B_{t+j}, \tag{4}$$

where (4) is derived by applying the difference operator Δ in (3) and using (1).

For (3) or (4) to impose a constraint analogous to the intertemporal budget constraint faced by an individual, it must hold that

$$E_t \lim_{j\to\infty} \gamma^{j+1} \Delta B_{t+j} = 0 \tag{5}$$

in (4) [or $E_t \lim_{j\to\infty} \gamma^{j+1} B_{t+j} = 0$ in (3)]. If (5) is satisfied, then intertemporal budget balance or deficit sustainability holds because this would require that the government run future surplus equal, in expected present-value terms, to its current market value of debt.

To test Condition (5), the procedure in the literature is to test for the stationarity of ΔB_t, or alternatively to test for the stationarity of $G_t^r - R_t$ [if they are each I(1)] with cointegrating vector $(1, -1)$ imposed. An equivalent procedure is to test for cointegration in the regression equation

$$R_t = \mu + b G_t^r + \epsilon_t \tag{6}$$

and to test that $b = 1$. Note that Hakkio and Rush (1991) relaxed this condition by showing that cointegration and $0 < b \leq 1$ are necessary conditions for a strict interpretation of deficit sustainability [i.e., that (5) holds]. I argue that the condition $0 < b \leq 1$ is a necessary and sufficient condition for deficit sustainability and that cointegration is only a sufficient condition. Note, however, that, although our results show that $0 < b < 1$ is sufficient for the deficit to be sustainable, it is inconsistent with the government's ability to market its debt in the long run. In other words, the condition $0 < b < 1$ has serious policy implications because a government that continues to spend more than it earns has a high risk of default and would have to offer higher interest rates to service its debt.

Theorem 1.1. Assume that interest rates are constant at r. If ΔB_t is stationary, then the limit term in (5) behaves like

$$E_t \lim_{T\to\infty} \exp(-Tk) = 0 \tag{7}$$

for some constant k. If ΔB_t is nonstationary, then (5) behaves like

$$E_t \lim_{T\to\infty} \exp(-Tk) T^{1/2} = 0. \tag{8}$$

Proof. Write the limit term in (5) as

$$\begin{aligned} \lim_{T\to\infty} \gamma^T \Delta B_T &= \lim_{T\to\infty} \exp(T \ln(\gamma)) \Delta B_T \\ &= \lim_{T\to\infty} \exp(-Tk) \Delta B_T, \end{aligned} \tag{9}$$

where $k = -\ln(\gamma)$. Then the result follows by noting that $\Delta B_t = Op(1)$ if ΔB_t is stationary and $\Delta B_t = Op(T^{1/2})$ if it is I(1) (assuming that an invariance principle holds for ΔB_t).

Theorem 1.1 shows that stationarity of ΔB_t is a sufficient condition for the bubble term to go to 0. Note, however, that the bubble term in (7) goes to 0 faster than that in (8) when cointegration does not hold. I call (7) and (8) the "strong" and "weak" requirement for deficit sustainability, respectively. Condition (7) corresponds to Hamilton and Flavin's (1986) requirement for deficit sustainability. I shall show that the slower rate of convergence to 0 (8) translates to the interpretation that, although intertemporal budget balance is satisfied in the strict sense (because the bubble term goes to 0), it can have serious policy implications because the condition $0 < b < 1$ is sufficient for (8) to hold. I start by considering the testable implications of (7) and (8) in light of the regression equation (6).

Note that the error process $\{\epsilon_t\}$ is I(0) if R_t and G_t^r are cointegrated and I(1) if cointegration is not satisfied. Writing (6) into (1), the undiscounted debt is written as

$$B_t = (1 + (1-b)r_t)B_{t-1} + (1-b)G_t - \mu - \epsilon_t, \qquad (10)$$

or in terms of G_t^r,

$$\Delta B_t = (1-b)G_t^r - \mu - \epsilon_t. \qquad (11)$$

Note that if $0 < b < 1$ in (11) then ΔB_t is I(1) [because G_t^r is I(1)] regardless of whether ϵ_t is I(1) or I(0) (i.e., regardless of whether R_t or G_t^r are cointegrated). Therefore ΔB_t is stationary iff (a) $b = 1$ and (b) ϵ_t is I(0). Thus (a) and (b) are necessary and sufficient conditions for (7) to hold.

I now look at the necessary condition for the weaker form (8) to hold. Solving (10) forward and setting $r_t = r$, Condition (5) translates to the condition that

$$E_t \lim_{j\to\infty} \gamma^{j+1}\Delta B_{t+j} = E_t \lim_{j\to\infty}\left\{\sum_{k=0}^{j}[1+(1-b)r]^{j-k}\gamma^{j+1}\Delta S_{t+k} + [1+(1-b)r]^j\gamma^{j+1}\Delta B_{t-1}\right\} = 0, \qquad (12)$$

where spending is given by $S_t = (1-b)G_t - \mu - \epsilon_t$. For (12) to behave like (8), the necessary and sufficient conditions are that (a) $0 < b \le 1$ and that (b) $\Delta B_T = Op(T^{1/2})$ [note that $\Delta S_t = Op(1)$]. For (b) to hold, it is sufficient that ΔB_T be I(1) or even mildly explosive as long as the explosive roots behave as $\exp(C/T)$ with $C > 0$ (see Phillips 1988). [Note that the behavior of ΔB_T is not limited to that of an integrated or mildly explosive process. The process ΔB_T can be dominated by a trend (e.g., a unit root with drift) and the bubble term will still go to 0, although at a rate slower than (8).]

In the empirical framework, in which I assume that R_t and G_t^r are I(1), the necessary and sufficient condition for a *strict* interpretation of budget balance is the condition $0 < b \le 1$ with cointegration being only a *sufficient* condition. If $0 < b < 1$, then (8) is automatically satisfied whether or not G_t^r and R_t are cointegrated because, ΔB_t is I(1) from (11). If $b = 1$, then (7) is satisfied if G_t^r and R_t are cointegrated, whereas (8) is satisfied if cointegration fails to hold.

Testing therefore proceeds as follows. If R_t and G_t^r are I(1), I run Regression (6) and test the null that $b = 0$ against the one-sided alternative that $b > 0$. If the null is accepted, then the deficit is not sustainable, but if it is rejected, I test the hypothesis that $b = 1$ against the alternative that $b < 1$. If the null is rejected, then $0 < b < 1$ and (8) is satisfied. If the null is accepted so that $b = 1$, the condition of cointegration between R_t and G_t^r is important. If ϵ_t is I(0) in (11), then (7) is satisfied; otherwise the weaker form (8) holds.

Using U.S. data, Hakkio and Rush (1991) found cointegration between R_t and G_t^r for the period 1950(2)–1988(4), but b was shown to be significantly less than 1. This would imply that the weaker form (8) is satisfied. When the sample is split at 1964, however, the first sample is cointegrated but with b still less than 1 [and thus (8) is still satisfied], and the second sample shows a lack of cointegration. They interpreted this result to mean that the deficit has become a problem in recent years and is not sustainable. The finding of no cointegration, however, does not "strictly" imply that the deficit is not sustainable as discussed previously. The deficit could still be sustainable as long as $0 < b \le 1$. Note that Haug (1991, 1992), Smith and Zin (1991), and Trehan and Walsh (1988) used the stronger condition of having a cointegrating vector of $(1, -b) = (1, -1)$ as a test for deficit sustainability as well.

I extend the work in the literature by using the weaker condition (8) as a test for deficit sustainability when the data fail to satisfy the condition of having a cointegrating vector of $(1, -1)$. I also show how to apply tests for structural breaks in long-run relationships. A brief discussion of these tests is given in Section 2.

2. RANK CONSTANCY TESTS FOR A CHANGE IN COINTEGRATION

This section presents the test for structural change in the rank of the cointegrating matrix. A full exposition on the subject is contained in Quintos (1994a). Here, I simplify the case to one in which trends do not enter the vector autoregression (VAR), although the marginal processes are allowed to contain a linear trend.

Consider a test for a change in cointegration between an n-dimensional I(1) process $\{X_t\}$ that satisfies the following error-correction formulation:

$$\Delta X_t = \mu_t + \Pi_j^*(L)\Delta X_{t-1} + \Pi X_{t-1} + \epsilon_t, \qquad (13)$$

where $\mu' = (\mu_a, \mu_b)$, $\Pi_j^*(L) = \sum_{i=1}^{j-1}\Pi_i^* L^{i-1}$, and $\Pi_i^* = -\sum_{1=i+1}^{j}\Pi_1$ for $i = 1, \ldots, j-1$. The hypothesis of interest is that Π and the rank of Π or the number of cointegrating relationships remains stable over time:

$$H_o^q: \rho(\Pi_t)_t = q, \qquad \Pi_t = \Pi \qquad (14)$$

for all t, where $0 \le q < n$ and q is the number of cointegrating relations under the null model. If the breakpoints are known (or chosen exogenously), the hypothesis (14) can be written as follows. Suppose that there are J known structural shifts assumed to occur at times $t = m_j + 1$ for $j = 1, \ldots, J$. I denote the subsample periods as $T_k = \{m_{k-1}+1, \ldots, m_k\}$, $k = 1, \ldots, J+1$, and I initialize the time at $m_0 = 0$ and end it at $m_{J+1} = T$. For instance, if $J = 1$, the single breakpoint occurs at time $t = m_1 + 1$, and the pre- and post-break periods are $T_1 = \{1, \ldots, m_1\}$ and $T_2 = \{m_1+1, \ldots, T\}$, respectively. The null hypothesis with a single break is written as

$$H_o^q: \rho(\Pi_1)_1 = q = \rho(\Pi_2)_2, \quad \Pi_1 = \Pi_2 = \Pi, \tag{15}$$

where subscripts denote the periods before and after the breakpoint. More generally, the null hypothesis of parameter and rank constancy is written as

$$\begin{aligned} H_o^q: \rho(\Pi_1)_1 &= \cdots = \rho(\Pi_{J+1})_{J+1} = q \\ \Pi_1 &= \cdots = \Pi_{J+1} = \Pi. \end{aligned} \tag{16}$$

The alternative hypothesis to (16) is that both the parameters and the number of cointegrating relationships are unstable,

$$\begin{aligned} H_a^{q_1,\ldots,q_{J+1}}: \rho(\Pi_1)_1 = q_1, \ldots, \rho(\Pi_{J+1})_{J+1} = q_{J+1}, \\ \Pi_1 \neq \cdots \neq \Pi_{J+1} \neq \Pi. \end{aligned} \tag{17}$$

I summarize the results on testing the null H_o^q but omit their proofs because they follow closely the proofs of Quintos (1994a).

From (13), define $\Delta X_{kt} = \mathrm{vec}(\Delta X_{t-1}, \ldots, \Delta X_{t-k})$ and let $\Delta X_{kt}^* = (\Delta X_{kt}', 1)'$. Furthermore, I denote by R_{ot}, R_{nt}, and ν_t the residuals of the regression of ΔX_t, X_{t-1}, and ϵ_t on ΔX_{kt}^*, respectively. Write (13) in terms of these residual vectors:

$$R_{ot} = \Pi R_{nt-1} + \nu_t, \tag{18}$$

from which the rank of Π and the reduced rank estimator $\widehat{\Pi}$ of Π are found. Suppose that an invariance principle holds for $\{\epsilon_t\}$ (and hence $\{\nu_t\}$) so that

$$T^{-1/2} \sum_{t=1}^{[Ts]} \epsilon_t \xrightarrow{d} \Omega_a^{1/2} W_a(s) = B_a(s), \tag{19}$$

where $W_a(s)$ is a Wiener process and $B_a(s)$ is an n-dimensional Brownian motion with covariance matrix Ω_a. Then the likelihood ratio (LR) test for testing H_o^q against $H_a^{q_1,\ldots,q_J}$ is a version of Johansen's (1991) rank test with structural shifts.

For a single shift ($J = 1$), Quintos (1994a) showed that the LR statistic for a single shift is given by

$$\begin{aligned} \mathrm{LR}_1^+ = T \sum_{i=1}^{q} \ln(1 - \widehat{\lambda}_i) \\ - \left(p_1 \sum_{i=1}^{q_1} \ln(1 - \widehat{\lambda}_{1i}) + p_2 \sum_{i=1}^{q_2} \ln(1 - \widehat{\lambda}_{2i}) \right), \end{aligned} \tag{20}$$

where $\widehat{\lambda}_1$ and $\widehat{\lambda}_2$ are the squared canonical correlations between R_{nt} and R_{ot} found from the pre- and post-break periods. The limiting distribution of the test (20) is given in the following theorem.

Theorem 2.1. Suppose that $q \geq q_1, q_2$; then the distribution of the LR test statistic is given by

$$\mathrm{LR}_1^+ \xrightarrow{d} \chi_{qn}^2. \tag{21}$$

Note that the distributional result (21) requires one to set $q_1 = q_2 = q$. If one uses the test (20) without imposing this equality, then LR_1^+ is a conservative test. If I consider the case $q < q_1, q_2$, I have

Theorem 2.2. If $q = \min(q_1, q_2)$, then

$$\begin{aligned} \mathrm{LR}^+ \xrightarrow{d} \chi_{qn}^2 + \mathrm{tr}\left\{ \int_0^{\tau_1} dP_1 G_1' \left(\int_0^{\tau_1} G_1 G_1' \right)^{-1} \int_0^{\tau_1} G_1 dP_1' \right\} \\ + \mathrm{tr}\left\{ \int_{\tau_1}^{1} dP_2 G_2' \left(\int_{\tau_1}^{1} G_2 G_2' \right)^{-1} \int_{\tau_1}^{1} G_2 dP_2' \right\}, \end{aligned} \tag{22}$$

where G_1 and G_2 are each of dimension $q_1 - q$ and $q_2 - q$, $P_j = W_a(s)$, and

$$G_j = \begin{pmatrix} W_c(s) \\ s - 1/2 \end{pmatrix}$$

for $j = 1, 2$ with $W_c(s) = W_a(s) - \int_0^1 W_a(r)$.

Here LR^+ has a mixture distribution this is not invariant to the breakpoint, and the nonstationary parts G_1 and G_2 are dependent components because G_2 is the tail-end distribution of G_1.

3. EMPIRICAL RESULTS

In this section, I apply the tests presented previously to U.S. data over the period 1947–1992. I address the question of whether there has been a shift in the government's fiscal policy and, if there is evidence of a shift in deficit policy, its implications for recent government laws on revenue generation and spending. I work with Equation (6) in real variables

$$\mathrm{rr}_t = \mu h_t + b g_t^r + \epsilon_t, \tag{23}$$

where rr and g^r denote real revenues and real government expenditures inclusive of interest paid on debt, respectively. The function h_t is a polynomial of order p in time and characterizes the order of the deterministic trend in the regression.

I use quarterly data over the period 1947(2)–1992(3). Data for federal government revenues and expenditure net of interest payments are taken from the National Income Product Accounts. Nominal interest rates are approximated by the 10-year U.S. government bond yield at constant maturity. Real variables are constructed by deflating nominal variables using the gross national product (GNP) price deflator. When variables are normalized by population, real-interest rates are constructed following Hakkio and Rush (1991).

I begin with a full-sample analysis of Equation (23) in Section 3.1 and then consider the analysis with structural breaks later in the article.

3.1 Model Without Breaks

3.1.1 Tests for Data Nonstationarity. Figure 1 plots the real revenues (rr) and expenditures (g^r) series. The increase in the gap between the series is notable from the mid-70s

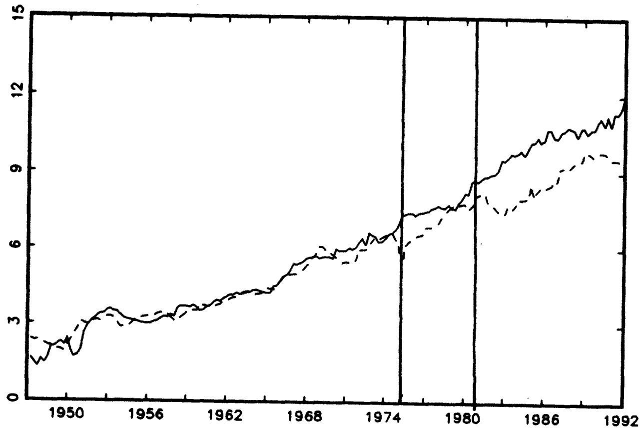

Figure 1. Real Government Revenue and Expenditure: – – –, Revenue; ———, Expenditure.

onward. As a test for stationarity of the data, I use the Phillips and Ploberger (1994) posterior information criterion (PIC) and the augmented Dickey–Fuller (ADF) test. The PIC finds the lag order (l) and trend degree (p) in an autoregressive model of the form

$$\Delta y_t = \alpha y_{t-1} + \phi_1 \Delta y_{t-1} + \cdots + \phi_l \Delta y_{t-l} + \gamma_0 + \gamma_1 t + \cdots + \gamma_p t^p + \epsilon_t \quad (24)$$

and gives the odds in favor of a unit root (i.e., choose the unit-root model with $\alpha = 0$ if PIC value >1). The major difference between the PIC criterion and the ADF test on (24) is that the PIC procedure selects the lag order and trend degree and therefore avoids overspecification or underspecification of the model structure under the null hypothesis.

For the real tax revenue series, the results favor a unit root with PIC = 19.25, $l = 1$, $p = -1$ (i.e., no intercept) and the coefficient on the first lag being 1.005. I also conduct standard classical tests for a unit root using the ADF test with the lag length and trend degree set to the values chosen by the PIC criterion. The test takes on the value of ADF = −.11, which falls below the 5% critical value and favors the unit-root hypothesis.

A similar set of tests is conducted on the real expenditure series with the results given as follows: PIC = 3.069e−05, $l = 2$, and $p = -1$, with the coefficient on the first lag in levels being 1.0097. The PIC value clearly does not favor a unit-root model, and a coefficient of 1.0097 suggests an explosive root. An ADF test on the data gives a value of ADF = .40, which indicates that a test for a unit root should be done against an alternative of an explosive root. I tabulate critical values for the ADF test with an explosive root alternative because tables in the literature are given with the stationary alternative. The 5% and 10% critical values are given by .68 and .28, respectively. With an ADF value of .40, the unit-root hypothesis is accepted at the 5% level.

The results are favorable to the unit-root hypothesis as well when I use real per capita data. The results for per capita real revenues are PIC = 115.63, $l = 1$, and $p = -1$, which clearly favors a unit-root model. An ADF value of ADF = −.84 also favors the unit-root hypothesis at the 5% level (critical value = −2.87). The results on per capita real expenditure are just as agreeable to the unit-root model with the test values of PIC = 10.28, $l = 1$, $p = -1$, and ADF = .79.

3.1.2 Test for Deficit Sustainability. To test for deficit sustainability, the current practice would be to first test for cointegration between revenues and expenditures. I have shown, however, that the necessary condition to be tested is whether $0 < b \leq 1$. The test for cointegration is important only if $b = 1$. If $b = 1$ and revenues and expenditure are cointegrated, then the strong form of deficit sustainabilty is satisfied. Otherwise, the lack of cointegration with $b = 1$ indicates that the weaker condition holds.

I perform the fully modified regression of Phillips and Hansen (1990) on (23). I use a modified t test on b, which I denote by t^+, to test the null that $b = 0$ against the alternative that $b > 0$. If the null is rejected, I then mount a t^+ test for the null that $b = 1$ against the alternative that $b < 1$. If the null is rejected, then $0 < b < 1$, which implies that the deficit could be sustainable.

The modified t test that I use is the square root of the modified Wald-type test of Phillips and Hansen (1990) for a model with a single regressor. I have the regression equation (23) with $\Delta g_t^r = \epsilon_{2t}$. Let $w_t = (\epsilon_t, \epsilon_{2t}')'$ and set

$$\lim_{T\to\infty} T^{-1} \sum_{1}^{T} E(w_t w_t') = \Omega = \begin{pmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{21} & \sigma_{22} \end{pmatrix}. \quad (25)$$

To test the linear parameter restriction $\widehat{b}^+ = c$, where $\widehat{b}^+$ denotes the fully modified estimator of b, the t^+ test is given by

$$t^+ = (\widehat{b}^+ - c)/\left(\widehat{\sigma}_{1.2}^2\left(\overline{g}^{r\prime}\overline{g}^r\right)^{-1}\right)^{-1/2},$$

where $\widehat{\sigma}_{1.2}^2 = \widehat{\sigma}_{11} - \widehat{\sigma}_{12}\widehat{\sigma}_{22}^{-1}\widehat{\sigma}_{21}$ is estimated using the quadratic spectral kernel with an automatic bandwidth parameter and is a consistent estimator of the conditional variance $\sigma_{1.2}^2$ (see Andrews 1991). A vector autoregressive model with one lag was used to pre-whiten the residuals. The matrix $\overline{g}^r$ is g_t^r regressed on the columns of h_t.

Table 1 contains the modified t tests on the parameter b. The symbols t_0^+ and t_1^+ denote tests of the null $b = 0$ and $b = 1$, respectively. The t^+ tests of $t_0^+ = 31.39$ and $t_1^+ = -3.45$

Table 1. Modified t Tests

Regression	*Parameter estimate*	t_0^+	t_1^+
Real rev on exp			
47(2)–92(3)	.90	31.39	−3.45
Break = 75(2)			
47(2)–75(2)	.97	39.45	−1.04
75(3)–92(3)	.86	26.16	−4.08
Break = 80(4)			
47(2)–80(4)	.96	48.64	−2.09
81(1)–92(3)	.84	47.95	−8.84
Per cap real rev on exp			
47(2)–92(3)	.91	30.47	−2.95
Break = 80(4)			
47(2)– 80(4)	.97	46.69	−1.39
81(3)– 92(3)	.85	50.11	−9.08

NOTE: 5% critical value for t_0^+ = 1.65 and t_1^+ = −1.65.

Table 2. Explosive Roots of VAR(4)

Model	Variables	Eigenvector	Roots > 1
	(a) Real variables		
Full sample			
47(2)–92(3)	Tax	−1.39	1.007
	Expend*	−1.39	
	Debt	−2.66	
	Rates	−0.11	
2nd sample			
75(2)–92(3)	Tax	−.72	1.002
	Expend*	−.61	
	Debt	−3.06	
	Rates	.004	
81(1)–92(3)	Tax	−1.85	1.008
	Expend*	−1.75	
	Debt	−5.84	
	Rates	−3.25e-05	
	(b) Real per capita variables		
Full sample			
47(2)–92(3)	Tax	−2.18	1.005
	Expend*	−2.38	
	Debt	−3.75	
	Rates	−.004	
2nd sample			
81(1)–92(3)	Tax	2.25	1.001
	Expend*	2.35	
	Debt	8.99	
	Rates	−.002	

*Expenditure exclusive of interest payments.

clearly reject the nulls of $b = 0$ and $b = 1$ so that $0 < b < 1$ at the 5% level (critical values are 1.65 and −1.65). This implies that the deficit is sustainable in the full sample although not in the "strong" sense.

Note that the result that $0 < b < 1$ implies that the undiscounted debt process is explosive from (10). Therefore one way to verify the result that $0 < b < 1$ would be to test whether the debt process has an explosive root. An analysis of the roots of the VAR on the system of four variables—revenues, government expenditure exclusive of interest payments, debt, and interest rates—shows that there is one explosive root in the system (see Table 2). The eigenvector corresponding to this root shows that the explosiveness comes mostly from movements in the debt process as expected. A univariate analysis of the debt process alone favors a unit root with PIC = 59.96, $l = 5$, and $p = -1$; however, the coefficient on the first lag in levels is $\widehat{\alpha} = 1.004$ and is indicative of an explosive root. An ADF value of ADF = 2.1 clearly favors the alternative of an explosive root against the null of a unit root (5% critical value is .68). Note that from Table 3 the deficit is sustainable because the rate that the debt explodes (2%) is still less than the mean interest rate of 5%.

I also test whether the deficit is a problem in a growing economy by measuring real revenue and spending relative to population. The t^+ tests on b are given in Table 1, and once again I have that $0 < b < 1$ at the 5% level (viz., $t_0^+ = 30.47$ and $t_1^+ = -2.95$). This implies an explosive behavior in the undiscounted debt process as affirmed by the VAR analysis on Table 2 and by an ADF value of ADF = .67

Table 3. Growth Rate of Debt and Mean Rates (for full and post-break samples)

	$\hat{b}$	$(1 - \hat{b})r$	r
Rev. on exp			
47(2)–92(3)	.90	.0205	.051
75(2)–92(3)	.865	.0431	.079
80(4)–92(3)	.844	.0551	.087
Per cap rev. on exp			
47(2)–92(3)	.91	.0169	.047
80(4)–92(3)	.846	.0476	.076

(PIC = 301.54, $l = 5$, $p = -1$, with $\widehat{\alpha} = 1.002$), but the deficit is still sustainable because debt explodes at a rate of 1.6% as compared to a 4.7% growth is the economy.

3.2 Model With Breaks

I found that there is evidence of an explosive root in the undiscounted debt process but that the deficit is still sustainable in the weak sense (i.e., it goes to 0 at a slower rate). In this section, I take issue with the possible misspecification of Equation (23). The possibility of a structural break is particularly clear in Figure 1, in which revenues and expenditures continue to diverge in the mid-70s. Hakkio and Rush (1991) showed that the presence of structural breaks in (23) altered their results to one in which deficit is *not sustainable* (in the "strong" sense) in the years after the break occurred.

3.2.1 Choice of the Break Date. The issue of whether structural breaks exist at all in the deficit process is still an unsettled issue in the literature. Haug (1992) found no breaks in the deficit process, whereas Hakkio and Rush (1991), Hamilton and Flavin (1986), and Wilcox (1989) found evidence of structural breaks. In all of these articles that found breaks, the dates were exogenously chosen as occurring in the mid-70s or early 80s (viz., Hakkio and Rush, 1964(1) and 1976(3); Hamilton and Flavin, 1981; and Wilcox, 1974). Here, I search for breaks endogenously following Haug (1992).

To test for the significance of the breakpoints, I do a modified Wald test [also called a G test by Phillips and Hansen (1990)] on the parameter δ in the regression equation

$$\mathrm{rr}_t = \mu + b g_t^r + \delta\left(D_t g_t^r\right) + \epsilon_{1t}, \qquad (26)$$

where

$$\begin{aligned} D_t &= 1 \quad \text{if} \quad t \in T_1 = \{1 \cdots m\} \\ &= 0 \quad \text{if} \quad t \in T_2 = \{m+1 \cdots T\} \end{aligned}$$

and m denotes the time of the breakpoint. The null hypothesis is that $\delta = 0$. This is in contrast to Haug (1992), who applied tests directly to the parameter b. From Hansen (1990), a test on δ is χ^2 distributed at the limit if rr and g^r are I(1), and I denote the test statistic by G_m, where the subscript m denotes the year of the break date. If rr and g^r are near integrated [i.e., with roots equal to $\exp(C/T)$, $C > 0$], then the statistic can be shown to be χ^2 distributed also because the rates of convergence are the same when $C = 0$ and $C > 0$ (see Phillips 1988).

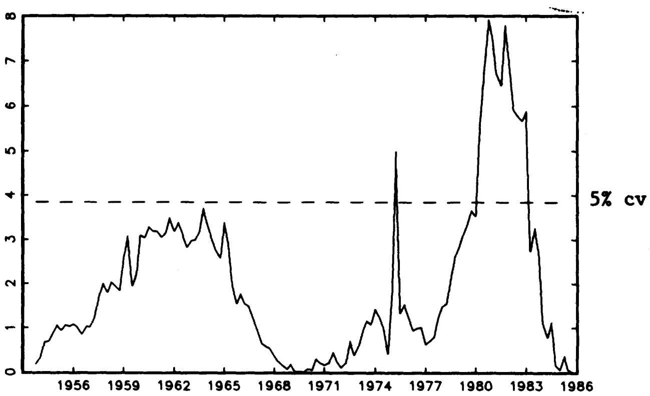

Figure 2. Sequential Chi-squared Test for Real Variables.

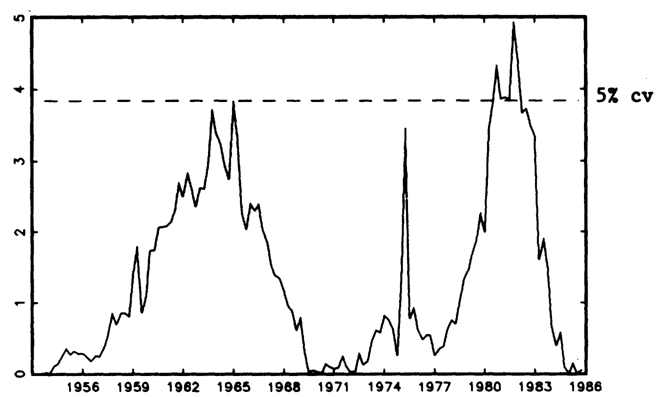

Figure 3. Sequential Chi-squared Test for Real per Capita Variables.

Figure 2 graphs the pointwise chi-squared test through time when I have trimmed the initial and final parts of the sample by 15% following Andrews (1993). Observe from Figure 2 that the modified Wald test rejects the null of stability (i.e., that $H_o: \delta = 0$) at the point 1975(2) (with a value of $G = 4.99$) and at several points in the early 80s. Because breakpoints cannot fall too close together the break date I take in the 80s is 1980(4) (with a value of $G_{80} = 7.93$), which is also the sup of the Wald test over time. With a critical value of 3.84, both points 75(2) and 80(4) are significant at the 5% level. Note that it was in the early 80s that legislative tax changes in the U.S. were introduced (e.g., the Kemp–Roth tax cut). Note also that the breakpoints chosen coincide roughly with those in the literature. Vertical lines in Figure 1 indicate the positions of the break dates in the data.

I also calculate the Schwartz and Akaike information criteria (SIC and AIC, respectively) and give the results in Table 4. Both the AIC and SIC indicate that the appropriate model has a break date at 1975(2).

Although the model-selection criteria (AIC and SIC) pick 1975(2) as the break date, I perform the analysis treating 1980(4) as a break date as well because I found it to be significant in the previous analysis. I work with a single break model in the article [either 1975(2) or 1980(4)] because I have too few data points to accommodate both break dates.

I also search for breakpoints in per capita variables, and a sequential plot of the G tests is given in Figure 3. Here, significant breakpoints occur at 1980(4), 1981(4), and 1982(1), with test values of $G_{80} = 4.33$, $G_{81} = 4.93$, and $G_{82} = 4.41$, respectively. Although these breaks are significantly different from 0 at the 5% level, results of the model-selection criteria in Table 4 indicate that the appropriate model is the full-sample model without breaks, and the model with a break at 80(4) is a close second.

Table 4. Model Selection

Model	AIC	SIC
Real rev on exp		
No break	3.63	3.66
Break = 75(2)	4.97	5.02
Break = 80(4)	3.76	3.82
Break = 75(2) and 80(4)	4.52	4.59
Per cap real rev on exp		
No break	6.91	6.94
Break = 80(4)	6.67	6.73
Break = 81(4)	6.51	6.56
Break = 82(1)	6.55	6.61

3.2.2 Tests for Deficit Sustainability. Tests for data nonstationarity on subperiod data for breakpoints 75(2) and 80(4) (for real and real per capita variables) were conducted as in Section 3.1.1. All subperiod data supported the unit-root model at the 5% level.

I perform subperiod fully modified regressions on Equation (23) and use the modified t^+ tests, t_0^+ and t_1^+, as before. The results are contained in Table 1. The results are striking: For the first subsample period, the model $b = 1$ is appropriate, whereas for the second subsample I have $0 < b < 1$ at the 5% level for both breakpoints at 75(2) and 80(4). Furthermore, tests for cointegration in Table 5 show that cointegration holds for the first subsample periods (although note that the post-break samples have a short span and cointegration tests may have low power for this period). The test statistic LR denotes Johansen's LR test for the rank of the

Table 5. Tests for Cointegration

	ADF $k = 1$	Z_t	Z_α	LR Trace $q = 0$	LR Trace $q \le 1$	LR Max $q = 0$	LR Max $q \le 1$
(a) Between real revenues and expenditure							
Full sample							
47(2)–92(3)	−3.65*	−3.73*	−27.53*	13.6	.7	12.9	.7
Break = 75(2)							
47(2)–75(2)	−3.57	−4.55*	−49.05*	16.3	1.0	15.3	1.0
73(3)–92(3)	−1.92	−2.04	−7.19	4.6	.3	4.6	.3
Break = 80(4)							
47(2)–80(4)	−4.54*	−4.41*	−35.33*	21.3	.9	20.4	.9
81(1)–92(3)	−1.61	−1.87	−5.47	5.8	.1	5.6	.1
(b) Between per capita real revenues and expenditure							
Full sample							
47(2)–92(3)	−4.33*	−4.29*	−34.84*	19.6	.4	19.2	.4
Break = 80(4)							
47(2)–80(4)	−4.76*	−4.49*	−36.56*	24.5	.1	24.4	.1
81(1)–92(3)	−1.56	−1.81	−5.05	14.9	.7	14.2	.7

NOTE: 5% and 10% critical values for ADF and Z_t are −3.40 and −3.01, for Z_α = −19.19 and −16.19. Critical values for the LR test from table A2, Johansen and Juselius (1990). *— reject null of no cointegration at the 5% level.

cointegrating matrix with the lag length ($l = 1$) and order of deterministic trend ($p = 0$) in the error-correction regression chosen using the AIC and SIC criteria. The Z_α, Z_t, and ADF tests are the unit-root tests used in Section 3.1.1. The results mean that the deficit is sustainable in the "strong" sense in the first period (because $b = 1$ and there is cointegration), but it is sustainable only in the "weak" sense in the post-break period (because $0 < b < 1$). For the post-break period, this implies an explosive root in the undiscounted debt process, and this is substantiated by ADF tests on the debt process [ADF = 1.45 for a break in 75(2) with PIC = 3.46e–06, $l = 1$, $p = -1$, and $\widehat{\alpha} = 1.02$; ADF = .34 (significant at 10% level) for a break in 80(4) with PIC = .24, $l = 1$, $p = 0$, and $\widehat{\alpha} = 1.002$] and in Table 2 and Table 3. These results hold as well for the model with real per capita variables.

3.3 Model With or Without Breaks?

From Table 5 and Table 1, I have cointegration and $0 < b < 1$ for the full-sample, cointegration and $b = 1$ for the first period [regardless of whether the break date is at 75(2) or 80(4)], and no cointegration and $0 < b < 1$ in the post-break period. These results translate into testing whether the full period with cointegration—that is, H_0^1—fits the data better than the split model with cointegration only in the first period—that is, $H_a^{1,0}$. In this section, I apply the rank constancy tests of Section 2 to determine which model is appropriate.

I first note the power of the rank constancy test with the breakpoint specified. The data-generating mechanism follows Quintos (1994b) and is given by

$$\begin{aligned} y_t &= bx_t + q_t \\ q_t &= \rho_q q_{t-1} + e_{qt} \quad \text{with} \quad |\rho_q| < 1, \\ \Delta x_t &= e_{xt} \end{aligned}$$

where

$$\begin{bmatrix} e_{qt} \\ e_{xt} \end{bmatrix} \sim \text{iid } N\left[0, \begin{bmatrix} 1 & \nu\sigma \\ \nu\sigma & \sigma^2 \end{bmatrix}\right].$$

I use for b the parameters found in Table 1 for a breakpoint at 75(2); to wit, $(b, \rho_q, \sigma, \nu) = (.97, .5, .5, -.5)$ for the first period and $(b, \rho_q, \sigma, \nu) = (.86, .5, .5, -.5)$ for the post-break period. The sample size is the length of our data set, $T = 182$. With a break located at the same proportion as the point 75(2) is relative to sample size (i.e., I set the break at 62% of the sample), the test rejects the null of stability correctly 94% of the time. As the break gets closer to the end of the sample [e.g., a break at 80(4) with $(b, \rho_q, \sigma, \nu) = (.96, .5, .5, -5)$ for the first period and $(b, \rho_q, \sigma, \nu) = (.84, .5, .5, -5)$ for the post-break period] the power falls to 79%. I have also done a Monte Carlo study on the behavior of the test as sample size increases, while fixing the breakpoint at the middle of the sample. The LR_1^+ test has good power properties at $T = 60$ with power equal to 72%, $T = 100$ with power equal to 92%, and $T = 120$ with power at 94%.

I apply the test to data, and my null hypothesis of interest is H_o^1, which I test against the alternative $H_a^{1,0}$. The form of the test I use is given in (20) and in Theorem 2.1,

$$\mathrm{LR}_1^+ = T \ln(1 - \widehat{\lambda}_1) - p_1 \ln(1 - \widehat{\lambda}_{11}). \tag{27}$$

For a break at 75(2), I get $\mathrm{LR}_1^+ = 2.91$, and for a break at 80(4) I have $\mathrm{LR}_1^+ = 8.8$. My critical value is $\chi_2^2 = 5.99$. Thus the preferred model indicates that there is a shift in deficit policy in 80(4) and that cointegration between revenues and expenditures fails to hold in the early 80s. Note that if I treat the breakpoint as unknown and look at the sup of the LR^+ test over time, I get that the break occurs at 81(1) with a value of $\mathrm{LR}^+ = 9.09$. This coincides roughly with the break chosen using the parameter instability test. (The 5% critical value for the sup test is 15.98 for a sample size of 182 and 500 iterations. Using this critical value, the sup value of 9.09 would fail to reject the null of rank stability. The purpose of the sup test, however, is to look for one-time shifts in the time series, and it has low power in detecting gradual shifts over time. Figure 1 shows that the spread in revenues and expenditures is more a gradual than a one-time occurrence.) A similar test on per capita variables gives the same conclusion, with $\mathrm{LR}_1^+ = 5.6$ at the 10% level (critical value = 4.6). When the breaks are treated as unknown, the break occurs at 81(4) with the value of 6.5. Thus, the tests generally show that the instability in the deficit process occurs in the early 80s.

Note that the evidence for a shift in deficit policy seems weaker when I use per capita data, both for the model-selection results in Section 3.2.1 and for the preceding LR^+ tests (i.e., I find structural shifts only at the 10% level). An alternative formulation is to account for a growing economy by using GNP rather than population as recommended by Kremers (1989). My results accord with Hakkio and Rush (1991): Real revenue and expenditure inclusive of interest payments normalized by GNP are stationary, in which case (7) is satisfied. Furthermore, a sequential Chow test for parameter stability shows that there are no breaks in the deficit process, as shown by Haug (1992). This result leaves an open question as to whether there are shifts in deficit policy when variables are normalized to account for a growing tax base. In both cases in which GNP or population is used, however, the deficit is still sustainable with the strong form holding for variables normalized by GNP and the weaker form with population.

A final note regarding the choice of my start date. It is well known that unit-root and cointegration tests are not robust to the length of the data set. It turns out that I get similar results when I use 1950 as the start date (as did Hakkio and Rush) or 1960 (as did Hamilton and Flavin). Using the modified Wald test to search for breaks, I have a break at 1980(4) for a start date at 1950 and 1981(1) for a start date at 1960. These coincide roughly with my break when I use an initial point at 1947. Furthermore, the results agree with those of a 1947 start date; that is, I have cointegration and $0 < b < 1$ for the full-sample, cointegration and $b = 1$ for the pre-break sample, and no cointegration and $0 < b < 1$ for the post-break sample. For the rank constancy test, I get similar results for the 1950 start date with $\mathrm{LR}^+ = 7.56$ (reject H_o^1) but not for the 1960 start date with $\mathrm{LR}^+ = 4.22$ (do not reject H_o^1, although one is close to rejecting the null at the 10% level). Of course, the results may differ drastically if I include the interwar years as done by Kremers (1989), but the results seem robust for the postwar period.

4. CONCLUSION

In this article, I tested whether U.S. fiscal policy is consistent with intertemporal budget balance and whether there has been a structural change in deficit policy. I introduced a weaker condition for deficit sustainability that allows the bubble term in the debt process to go to 0 at a slower rate. Under this weaker condition, cointegration between revenues and expenditures is a sufficient condition for deficit sustainability. Although intertemporal budget balance is satisfied in the strict sense, a government that satisfies its intertemporal budget constraint under this weaker condition may have difficulty marketing its debt because spending continually exceeds revenues earned.

Using rank tests for a change in the long-run relationships, I found that there is a shift in deficit policy in the early 80s. In other words, I found that revenues and expenditures inclusive of interest paid have cointegrating vector $(1, -1)$ for the pre-break period, but they are not cointegrated in the post-break period with $0 < b < 1$. The finding of a break in the deficit process is more uncertain when variables are normalized by GNP and population. Regardless of whether or not there is a shift in policy I have shown that the deficit is sustainable.

ACKNOWLEDGMENTS

This article is a chapter of my dissertation written at Yale University. My thanks go to John Chao, Chris Sims, and Peter Phillips for helpful discussions and comments. My thanks also go to two anonymous referees and to Michael Cox of the Federal Reserve Bank of Dallas for providing the data on the market value of federal debt. Fellowship support was provided by the Cowles Foundation at Yale University.

[Received January 1994. Revised February 1995.]

REFERENCES

Andrews, D. W. K. (1991), "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, 59, 817–858.

——— (1993), "Tests for Parameter Instability and Structural Change With Unknown Change Point," *Econometrica*, 61, 821–856.

Hakkio, C. S., and Rush, M. (1991), "Is the Budget Deficit 'too Large'?" *Economic Inquiry*, 24, 429–445.

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

Hansen, B. E. (1990), "Testing for Structural Change of Unknown Form in Models With Non-stationary Regressors," unpublished manuscript.

——— (1992), "Tests for Parameter Instability in Regression With I(1) Processes," *Journal of Business & Economic Statistics*, 10, 321–335.

Haug, A. (1991), "Cointegration and Government Borrowing Constraints: Evidence for the United States," *Journal of Business & Economic Statistics*, 9, 97–101.

——— (1992), "Has Federal Budget Deficit Policy Changed in Recent Years?" unpublished manuscript, University of Saskatchewan, Dept. of Economics.

Johansen, S. (1991), "Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models," *Econometrica*, 59, 1551–1580.

Johansen, S., and Juselius, K. (1990), "Maximum Likelihood Estimation and Inference on Cointegration—With Application to the Demand for Money," *Oxford Bulletin of Economics and Statistics*, 52, 169–210.

Kremers, J. J. M. (1989), "U.S. Federal Indebtedness and the Conduct of Fiscal Policy," *Journal of Monetary Economics*, 23, 219–238.

Phillips, P. C. B. (1988), "Regression Theory for Near Integrated Time Series," *Econometrica*, 56, 1021–1044.

Phillips, P. C. B., and Hansen, B. E. (1990), "Statistical Inference in Instrumental Variable Regression With I(1) Processes," *Review of Economic Studies*, 57, 99–125.

Phillips, P. C. B., and Ploberger, W. (1994), "Posterior Odds Testing for a Unit Root With Data Based Model Selection," *Econometric Theory*, 10, 774–808.

Quintos, C. E. (1994a), "Rank Constancy Tests in Cointegrating Regressions," unpublished manuscript, Yale University, Dept. of Economics.

——— (1994b), "Analysis of Cointegration Vectors Using the GMM Approach," unpublished manuscript, Washington University, John M. Olin School of Business.

Smith, G. W., and Zin, S. E. (1991), "Persistent Deficits and the Market Value of Government Debt," *Journal of Applied Econometrics*, 6, 31–44.

Trehan, B., and Walsh, C. E. (1988), "Common Trends, the Government's Budget Constraint, and Revenue Smoothing," *Journal of Economic Dynamics and Control*, 12, 425–444.

——— (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.

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