Working Paper No. 358

Threshold Effects in the U.S. Budget Deficit

by
Philip Arestis
Levy Economics Institute of Bard College, New York
p.arestis@levy.org
Andrea Cipollini
Queen Mary and Westfield College, University of London
a.cipollini@qmul.ac.uk
Bassam Fattouh
CeFiMS, SOAS, University of London
bf11@soas.ac.uk

We are grateful to an anonymous referee for helpful comments. We are also grateful to Bruce Hansen for helpful suggestions, and for making available his Gauss Programme to estimate and make inferences on the TAR model. The usual disclaimer applies.

Introduction

There has been an ongoing debate on whether the U.S. fiscal policy is sustainable in the long run. In addressing this issue, a number of studies have examined whether the U.S. public finances are compatible with the government's intertemporal solvency constraint (Hamilton and Flavin, 1986; Trehan and Walsh, 1988; Hakkio and Rush, 1991, amongst others). The requirement of budget processes to be sustainable implies effectively that Ponzi games are ruled out as a viable option of government finance. In other words, further new borrowing cannot be used indefinitely as a method of financing interest payments on existing debt. Therefore, the solvency constraint requires that any changes in taxes and government spending are followed by adjustment in future taxation, and or spending, which equals to the original change in present value. This solvency constraint imposes testable restrictions on the time series properties of key fiscal aggregates.

In this paper we present new evidence on this ongoing solvency debate utilising an empirical framework which allows us to test whether there have been threshold effects in the U.S. deficit. Unlike existing empirical studies that focus only on the identification of regime shifts in U.S. fiscal policy, we also offer an explanation and provide evidence as to why these regime shifts might occur. Specifically, we argue that for fiscal authorities to be able to meet the solvency constraint, they would intervene through deficit cuts only when the government budget deficit becomes very large. Therefore, we expect a mean reverting dynamic behaviour for deficits only when they are above some threshold value. We test for this hypothesis using the threshold unit root empirical methodology recently developed by Caner and Hansen (2001).

The paper is organized as follows. Section 2 reviews the empirical evidence. Section 3 describes the empirical methodology, while Section 4 presents the empirical results which are analysed and interpreted. Section 5 summarises and concludes.

2. Intertemporal Solvency Condition: Empirical Underpinnings

One strand of the empirical literature focuses on the stationarity of the stock of public debt. Uctum and Wickens (1997) deal with stochastic interest rates and primary surpluses that are allowed to be either exogenous or endogenous to the stock of public debt. They show that a necessary and sufficient condition for the intertemporal solvency condition to hold is a stationary discounted stock of public debt. Another strand of the empirical literature has concentrated on the dynamics of the undiscounted inclusive-of-interest deficit, or alternatively on the long-run relationship between government spending and tax revenues. Trehan and Walsh (1988) show that government spending, inclusive of interest payments, and government revenues should be cointegrated with a cointegrating vector equal to [-1, 1]'. They present evidence that supports this restriction. Hakkio and Rush (1991) point out that a necessary condition for intertemporal solvency is the existence of cointegration between government expenditure, inclusive of interest payments, and government revenues, with a cointegrating vector equal to [1, -β]', and 0<β< 1.

Quintos (1995) expands on Hakkio and Rush (1991) and shows that cointegration is not necessary for the intertemporal balance condition to hold. Specifically, she distinguishes between a weak and a strong sustainability condition. The former implies that the government solvency holds, but the undiscounted debt process is exploding at a rate that is less than the growth rate of the economy. Although this case is consistent with deficit sustainability, it is inconsistent with the ability of the government to market its debt in the long run, especially if the focus is on the debt to GNP, or on the debt per capita (see, also, Hakkio and Rush, 1991). In contrast, strong sustainability implies that the undiscounted public debt is finite in the long run. As Quintos (1995) shows, strong sustainability occurs when government spending, inclusive of interest payments, and government revenues are cointegrated, with a cointegrating vector that is equal to [-1, β]', with β= 1. Alternatively, strong sustainability occurs if the total government deficit series (inclusive of interest payments) is stationary. Weak solvency occurs when 0 < β < 1, regardless of the order of integration of the estimated cointegrating regression residuals.

Empirical research on U.S. public deficit sustainability has emphasised the importance of taking into account the possibility of regime shifts in U.S. fiscal policy. Most of the studies that tested for the presence of these shifts, consider models with structural breaks, where the breakpoints are either chosen arbitrarily (e.g. they are exogenous) or are endogenously determined. Specifically, in the empirical works of Hakkio and Rush (1991) and of Tanner and Liu (1994), the breakpoint is exogenously determined. Hakkio and Rush (1991) use a sample that spans from the first quarter of 1952 to the fourth quarter of 1988 and find that the intertemporal budget constraint is satisfied in sub-samples up until 1976, but that for the subsample 1976:1 to 1988:4 the constraint is violated. Tanner and Liu (1994) find evidence of strong solvency, by imposing a dummy variable for a level shift in the cointegration vector at pre-defined points, i.e. 1982(1) and 1981(4). With the level dummy incorporated, they find evidence of cointegration and acceptance of the null that $\beta$is equal to unity. Haug (1995) and Quintos (1995) allow for the breakpoint to be endogenously determined. Using a recursive test for the stability of the cointegrating vector [1, -β]', Haug (1995) does not detect any structural change in the conduct of U.S. fiscal policy over the last four decades. Quintos (1995) uses a statistical procedure which is very similar to that used by Haug (1995) with only the null hypothesis being different.

Moreover, we attempt to take the literature on U.S. deficit with structural breaks one step further by offering an explanation, and providing evidence for the occurrence of regime shifts in the budget deficit. For this purpose, we follow Bertola and Drazen (1993) who show that the government solvency requirement implies that any increase in current government spending (which has the effect of increasing the current fiscal deficit) has a non-linear effect on the expected present discounted value of the future government spending. Under these circumstances linear techniques may not be accurate in characterising the deficit process. Specifically, the authors develop a framework in which there exist trigger points in the process of budget cuts, such that significant reduction in budget deficits may take place only when the ratio of deficit to output reaches a certain threshold. This may reflect the existence of political constraints blocking deficit cuts which are relaxed only when the budget deficit reaches a sufficiently high level, deemed to be unsustainable (Bertola and Drazen, 1993; Alesina and Drazen, 1991). Consequently, in our empirical model we expect the fiscal process to follow different dynamics depending on whether the change in the deficit is below or above that threshold. Mean reverting dynamic behaviour for deficits is present only when the change in the deficit is above a certain threshold. For this reason, we use a Threshold Autoregressive model (TAR), where the deviations of a threshold variable (observable) from a trigger point (to be estimated) are used to explain the occurrence of (possible) regime shifts in the U.S. deficit. It is customary to test the null of linearity against the alternative of a specific non-linear process (a TAR model, in this paper) only after finding evidence of a stationary time series, on the basis of standard linear unit root tests. This sequential testing procedure may affect the inference results.[1] In this paper we follow Caner and Hansen (2001), the first to allow for a joint test for linearity and stationarity in the time series under investigation.

Our empirical results reveal the following important findings. The U.S. budget deficit is sustainable in the long run but it has undergone regime shifts. These two results are consistent with those of many existing empirical studies. We also find evidence that these regime shifts can be explained by the extent of the change in the deficit. More specifically, only when the increase in deficit per capita reaches a certain threshold, will fiscal authorities intervene to reduce the deficit. Hence, we are able to identify two regimes, with the budget deficit following different dynamics in each one of them. These results provide support for the existence of trigger points in the U.S. fiscal policy adjustment.

3.Empirical Methodology

Standard unit root and cointegration estimation procedures assume that there is a tendency for a variable to move towards a long-run equilibrium in every time period. However, as Balke and Fomby (1997) observe, movements towards the long-run equilibrium need not occur in every period. The presence of fixed costs of adjustment, for instance, may imply that only when the deviation from the equilibrium exceeds a critical threshold, do the benefits exceed the costs of adjustment, and, therefore, economic agents act to move back to equilibrium. This type of discrete adjustment process has been used to describe many economic phenomena, such as the behaviour of inventories or investment. In this paper we attempt to characterise the discrete adjustment in the policymakers attitude towards public finance solvency, in terms of a threshold stationary process for the government deficit.

The threshold autoregressive model (TAR) was introduced by Tong (1978) who considered the possibility of a mean reverting time series only after hitting a certain threshold. Chan (1993) demonstrated that the least squares estimates of the threshold is super-consistent, and found its asymptotic distribution, while Hansen (1997) developed an alternative approximation to the asymptotic distribution of the threshold. Chan (1991) and Hansen (1996) studied the asymptotic distribution of the likelihood ratio test for a threshold. Balke and Fomby (1997) extended the idea of threshold effects to a long-run equilibrium (cointegrating) relationship among different series. All these studies rely on the assumption of stationary data, and, therefore, do not discriminate between non-stationary and/or non-linear time series. The work of Caner and Hansen (2001) is the first to provide statistical tests for the null of a stationary threshold autoregressive process, which simultaneously allow for both effects.

The Econometric framework

The TAR model can be described by:

$$\Delta y_t = \theta_1' x_{t-1} I(Z_{t-1} < \lambda) + \theta_2' x_{t-1} I(Z_{t-1} \geq \lambda) + u_t$$

where $t = 1,....,T$; $x_{t-1} = (int, y_{t-1}, \Delta y_{t-1} \ldots \Delta_{t-k})'$; $I(.)$ is the indicator function, and u_t is an i.i.d error. The threshold variable $Z_{t-1} = y_{t-1} - y_{t-m-1}$, is predetermined and strictly stationary, and m is the delay order[2]. The deterministic component *int* stands for an intercept; $\Delta y_{t-j} = y_{t-j} - y_{t-j-1}$ is the first order difference at lag j . The threshold λ is unknown, and it takes values in the interval $\lambda \epsilon \Lambda = [\lambda_1, \lambda_2]$, where λ_1, λ_2 are chosen so that $Prob(Z_t \leq \lambda_1) = \pi_1 > 0$ and $Prob(Z_t \leq \lambda_2) = \pi_2 < 1$.[3] In so far as the description of the empirical methodology is concerned, it is convenient to consider the vectors $\theta_1 = (\rho_1\ \beta_1\ \alpha_1)'$ and $\theta_2 = (\rho_2$

β_2 α_2)', where ρ_1 and ρ_2 are the slopes on the lagged levels, β_1 and β_2 are the slopes on the deterministic component, and α_1 and α_2 are the slopes on the lagged differences in the two regimes.

When estimating the TAR model, for each $\lambda \epsilon \Lambda$, we estimate θ_1 and θ_2 by Least Squares (LS). The LS point estimate of the threshold λ and of the corresponding vectors θ_1 and θ_2, are those which minimise the residual sum of squares. To test the null hypothesis of linearity, that is $H_0: \theta_1 = \theta_2$, against the alternative of threshold effects, we use a Wald test statistic, W_T. The latter is not identified under the null, and its asymptotic distribution, under the assumption of stationary data, was investigated by Davies (1987), Chan (1991), Andrews and Ploberger (1994) and Hansen (1996). Caner and Hansen (2001) find that under the restriction of a unit root process, the asymptotic distribution of W_T depends on the data structure which implies that critical values cannot be tabulated, and, therefore, the authors suggest two bootstrap methods to approximate the asymptotic distribution of W_T. One method is appropriate for the stationary case, and the other is appropriate for the unit root case. If the true order of integration is unknown, then Caner and Hansen (2001) suggest to calculate the bootstrap critical values and p-values both ways, and base inference on the more conservative (larger) p-value. Furthermore, the authors suggest that, setting the bounds of the trimming region to $\pi_1 = 0.15$ and $\pi_2 = 0.85$, provides a reasonable trade-off between the power and size properties of the test for threshold effects (see Caner and Hansen 2001, for details).[4]

Unit Root tests

As for the unit root tests, the following statistics are employed:

1. two-sided Wald test statistic, R_{2T}, for the null of unit root, $H_0: \rho_1 = \rho_2 = 0$, against the alternative $H_1: \rho_1 \neq 0$ or $\rho_2 \neq 0$;
2. a one-sided Wald test statistic, R_{1T}, for the null of unit root, $H_0: \rho_1 = \rho_2 = 0$, against the alternative $H_1: \rho_1 < 0$ *or* $\rho_2 < 0$;
3. a one-sided Wald test statistic, t_1, for the null of unit root, $H_0: \rho_1 = \rho_2 = 0$, against the alternative of stationarity only in regime 1, that is, $H_1: \rho_1 < 0$ and $\rho_2 = 0$;
4. a one-sided Wald test statistic, t_2, for the null of unit root, $H_0: \rho_1 = \rho_2 = 0$, against the alternative of stationarity only in regime 2 that is, $H_1: \rho_1 = 0$ and $\rho_2 < 0$.

The first two test statistics allow us to test whether the series under study is stationary or not. In the context of this study, these test statistics can be used to examine whether the U.S. budget deficit is sustainable in the long run. The third and fourth test statistics allow us to identify which of the regimes is stationary. This identification is important since it allows us to examine whether budget deficits follow different dynamics after they reach a certain threshold. As mentioned in the introduction, it could be the case that fiscal authorities will intervene to cut the budget deficit only when the increase in the per capita deficit has reached a certain threshold.

If there are no threshold effects (unidentified case), the asymptotic distribution of each of the four statistics is found to be dependent on the data structure. However, asymptotic bounds, free of nuisance parameters other than the trimming range, are found. Consequently, critical values can be tabulated and p-values can be computed (see Table 3 in Caner and Hansen, 2001). As for the asymptotic distribution of each of the four statistics in the presence of threshold effects (identified case), the authors find that the Dickey-Fuller tabulated critical values provide a conservative bound for the t_1 and t_2 tests. Furthermore, they find that the expression derived for the asymptotic distributions of the *one-sided* R_1 test and of the *two-sided* R_2 test, are only useful for computing the critical values for the latter.[5] Finally, improved finite sample inference may be conducted using a bootstrap distribution for both the unidentified and identified threshold cases.

4. Empirical Results

The data set used in this study comprises quarterly observations over the period 1947:2 to 2002:1. Our focus is on the dynamics of the total (inclusive of interest payments) real per capita U.S. government surplus (see Figure 1).[6] Hence, we concentrate only on the strong sustainability condition (see Quintos, 1995, for details). Following Franses and Van Dijk (2000), we use the minimisation of the Akaike (1973) (AIC) and Hannan-Quinn (1979) (HQ) statistics[7] to detect the appropriate lag order k and delay parameter m for the threshold variable. It is clear from Tables 1 and 2 that both criteria suggest a lag order of 4 and a delay parameter of order 2 for the TAR process in (3). Hence, the semi-annual change in the surplus per capita $(y_{t-1} - y_{t-3})$ is the chosen threshold variable.

In Table 3, we report the threshold test and the unit root tests for the TAR model with a delay parameter of order two. The bootstrap p-values (both unrestricted and under the unit root assumption) corresponding to the Wald tests W_T indicate that we can reject the null hypothesis of linearity in favor of the alternative that there is a threshold effect in the U.S. budget surplus per capita. This indicates that the U.S. fiscal policy has undergone structural shifts, a conclusion reached by most empirical studies (Martin, 2000; Quintos, 1995, amongst others).

The unit root tests show some interesting results regarding the time-series properties of the U.S. real per capita budget surplus. Both the *one-sided* R_1 and the *two-sided* R_2 Wald tests give evidence of stationary process. However, the individual *t-ratios* tests, t_1 and t_2, identify the presence of a unit root only in the second regime[8]. In fact, these results are confirmed by the Least Squares estimates of the TAR model reported in Table 4 where there is evidence of a statistically significant mean reverting dynamics only in the lower regime, that is, when the (semi-annual) change in the surplus per capita reaches the estimated threshold point estimate of -0.313.[9] These results suggest that the dynamics of the budget surplus per capita are different depending on whether the (semi-annual) change in surplus per capita is above or below the estimated threshold value.

In Figure 2 we construct a confidence interval for the threshold parameter following the method of Hansen (2000). The bounds of this confidence interval are given by the intersection of the likelihood ratio sequence plot and a flat line at the appropriate asymptotic critical value.[10] For the 90% confidence interval, the critical value is 5.94 (see Hansen, 2000, Table 1), and the lower and upper bounds are -1.78 and -0.01 respectively. This confidence region would suggest that politicians would become sensitive to budget deficits only when the semi-annual decrease in the real surplus per capita (or increase in the real deficit per capita) is either 'bad enough' or shows up in the books (given that the upper bound is close to zero). A 95% confidence interval would confirm these conclusions.[11]

There is also evidence of delays in the fiscal stabilization plans, since the lag order of the TAR process is equal to 4 (quarters) and policymakers react, only after half a year, to an undesirable decrease in the surplus per capita. This delay can be explained by taking into account political and institutional frictions. More concretely, in the U.S. case, all rules and regulations according to which budgets are drafted, approved and implemented, can imply long delays in modifying budget deficits. Alesina and Perotti (1996a, 1996b) review the effect of voting procedures in the preparation and implementation of the budget on delays in fiscal stabilization plans and conclude in a similar vein.

Further Discussion

In Figure 3, we plot the deviations of the semi-annual change in the surplus per capita (threshold variable) from the estimated threshold point estimate over the sample period. Note that in this figure positive values identify the upper regime while negative values identify the lower regime (which includes 40% of the total sample observations). We impose horizontal lines in the figure to identify the dates of the most significant switches during the period under study.[12] During the 1950s and early 1960s, regimes shifts were mainly related to recessions and/or wartime. Figure 3 shows a major shift in U.S. budgetary policy during the

Korean War where there is an abrupt shift in 1951:2 from a large real surplus per capita to a large real deficit per capita, reaching a peak in 1951:3. During Eisenhower's presidency, there were two minor shifts in 1957:3 and 1960:4 caused mainly by two short-lived recessions (Evans, 1997). During the 1960s, the figure shows three main switches from surplus to deficit (1964:2, 1967:1 and 1969:3), occurring during Johnson's Presidency. These switches correspond in part to the tax cut proposals initiated by Kennedy and promoted and signed as law during Johnson's presidency. They also correspond to the budgetary requirements of the Vietnam War, which increased significantly especially in 1967, and the sharp increases in domestic spending associated with the Great Society Programs (Ippolito, 1990).[13] A major shift from per capita surplus to per capita deficit occurs in 1974:3, peaks in 1975:2, and corresponds to the 1973 oil crisis, which plunged the U.S. economy into a deep recession.[14] This switch into per capita deficit is bigger than the one associated with the second oil crisis period of 1979:4, and peaking in 1980:2.

Figure 3 clearly shows another major shift in 1981:3. This date is quite significant because it is the largest switch experienced in period of peace, tranquility and without external shocks. This shift, which occurs at the beginning of the Reagan presidency, corresponds to the legislation passed by the Congress aimed at cutting personal income taxes over the next three years (the 1981 Economic Recovery Tax Act). Since the tax cuts were not met by equal cuts in government spending, the federal budget went into deficit and remained so for a considerable period of time. It is interesting to note that many regime shifts (1985:3, 1986:3, and 1989:3) occurred during Reagan's presidency, mainly in his second term. These regime switches reflect in part the intensive political and economic debate in the Congress and in the media and the efforts to reduce the large and growing budget deficit. These efforts were manifested in the Tax Reform Act of 1986 and the Balanced Budget and Emergency Deficit Control Act, which called for progressive reduction in the deficit and the achievement of a balanced budget by the early 1990s (Ippolito, 1990). Despite efforts to balance the budget, other major shifts from per capita surplus to per capita deficit occurs in 1989:3 and in 1991:3. These switches occur during the Senior Bush presidency and correspond closely to a recession that plunged the economy at the beginning of Bush's term and later to the budgetary requirements of the Gulf War. These deficits did not last long and in 1993:1 the budget went into surplus and remained so for the rest of the 1990s. This coincided with President Clinton's move to the White House and the importance he attached to balancing the budget in his economic policy. Finally, the graph shows another switch from a per capita surplus to a per capita deficit in 2001:1. This switch can be explained by resorting to the slowing of the economy after a period of exceptional growth. The peak of the deficit in 2001:3 can be justified by referring, in addition to the continuing slowdown, to the preparations for the war in Afghanistan. It is of note that out of the fourteen switches identified in the graph, five of them took place during Democrat Presidency and nine during Republican Presidency.

5. Summary and Conclusions

In this paper, we have been concerned with investigating empirically the long-run sustainability of fiscal policy in the U.S. during the period spanning from 1947 to 2002. The theoretical framework has been provided by the government intertemporal budget constraint. The latter states that the discounted value of the public debt stock tends to zero over time and it has testable implications on the process driving the total government deficit. By contrast to previous studies which focused on the possibility of structural breaks in the deficit dynamics, we provide an explanation and gauge evidence for the occurrence of regime shifts in U.S. fiscal policy. Using a threshold unit root estimation procedure recently proposed by Caner and Hansen (2001), we provide evidence that the U.S. fiscal policy has undergone structural shifts in the last four decades. Our findings suggest that these switches are driven by asymmetries in the adjustment process, an issue not addressed previously in the literature. More specifically, government authorities would intervene by cutting deficits only when they have reached a certain threshold.

Data Appendix

The data for nominal government spending (net of interest payments) and government revenues are the National Income and Product Accounts (NIPA) figures, in billions of U.S. dollars. Real government spending and real government revenues are obtained by dividing the nominal values of government spending and government revenues by the Gross National Product (GNP) price deflator (1996 is the base year). Both nominal GNP and real GNP (in 1996 dollars) were obtained from NIPA. To calculate the interest payments, we first deflated the previous period's market value of the privately held gross federal debt by the GNP deflator. The market value of the debt was obtained from the Federal Reserve Bank of Dallas. We then multiplied this by the appropriately deflated nominal interest rates, following Hakkio and Rush (1991). Since we focus on real surplus per capita, we use the ex-post real interest rate minus the rate of population growth. For the nominal interest rate, we use the treasury long-term bond yields obtained from the Federal Reserve Statistical Release. Population figures are the mid-period estimation obtained from NIPA.

REFERENCES

Akaike, A. 1973. "Information Theory And An Extension of The Maximum Likelihood Principle." In Petrov, B. N., and Saki, F. L.eds. Second International Symposium of Information Theory. Budapest.

Alesina, A. and Drazen, R. 1991. "Why are Stabilisation Delayed?" *American Economic Review* 82: 1170-1188.

Alesina, A. and R. Perotti. 1996a. "Budget Deficits and Budget Institutions." *NBER* , working paper no. 5556.

-----. 1996b "Fiscal Discipline and the Budget Process." *American Economic Review* 86: 401-407.

Andrews, D.W.K., and W. Ploberger. 1994. "Optimal Tests When a Nuisance Parameter is Present Only under the Alternative." *Econometrica* 62: 1383-1414.

Balke, N.S. and T.B. Fomby. 1997. "Threshold Cointegration." *International Economic Review* 38: 627-645.

Bertola, G. and A. Drazen. 1993. "Trigger Points and Budget Cuts: Explaining the Effects of Fiscal Austerity." *American Economic Review* 83(1): 11-26.

Caner, M., and B.E. Hansen. 2001. "Threshold Autoregression with A Unit Root." *Econometrica* 69: 1555-1596.

Chan, K.S. 1991. "Percentage Points of Likelihood Ratio Tests for Threshold Autoregression." *Journal of the Royal Statistical Society* Series B 53: 691-696.

-----. 1993. "Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model." *The Annals of Statistics* 21: 520-533.

Chan, K.S., and R.S. Tsay. 1998. "Limiting Properties of the Least Squares Estimator of a Continuous Threshold Autoregressive Model." *Biometrika* 45: 413-426.

Cipollini, A. 2001. "Testing for Government Intertemporal Solvency: A Smooth Transition Error Correction Model Approach." *The Manchester School* 69 (6): 643-655.

Davies, R.B. 1987. "Hypothesis Testing When A Nuisance Parameter Is Present Only Under The Alternative." *Biometrika* 74: 33-43.

Evans, G.R. 1997. "Red Ink: The Budget, Deficit, And Debt of the U.S. Government." San Diego: Academic Press.

Franses, P.H. and Van Dijk, D. 2000: *Non-Linear Time Series Models In Empirical Finance.* Cambridge University Press.

Hakkio, C.S. and M. Rush. 1991. "Is The Budget Deficit 'Too Large'?" *Economic Inquiry* 29: 429-445.

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

Hannan, E. J., and Quinn, B. G. 1979. "The Determination of the Order of an Autoregression" *Journal of the Royal Statistical Society* B (41): 190-195.

Hansen, B.E. 1992. "Tests for Parameter Instability in Regressions with I(1) Processes." *Journal of Business and Economic Statistics* 10: 321-335.

-----. 1996. "Inference When A Nuisance Parameter Is Not Identified Under The Null Hypothesis." *Econometrica* 64: 413-430.

-----. 1997. "Inference in TAR Models." *Studies in Nonlinear Dynamics and Econometrics* 1: 119-131.

-----. 2000. "Sample Splitting and Threshold Estimation." *Econometrica* 68: 575-603.

Haug, A. 1995. "Has Federal Budget Deficit Policy Changed in Recent Years?" *Economic Inquiry* 33: 104-118.

Ippolito, D.S. 1990. *Uncertain Legacies: Federal Budget Policy from Roosevelt Through Reagan*. Charlottesville: University Press of Virginia.

Martin, G.M. 2000. "U.S. Deficit Sustainability: A New Approach Based on Multiple Endogenous Breaks." *Journal of Applied Econometrics* 15: 83-105.

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

Sarno, L. 2001. "The Behavior of U.S. Public Debt: A Non-Linear Perspective." *Economics Letters* 74: 119-125.

Tanner, E. and P. Liu. 1994. "Is The Budget Deficit 'Too Large': Some Further Evidence." *Economic Inquiry* 32: 511-518.

Tong, H. 1978. "On a Threshold Model." In C.H. Chen, ed., *Pattern Recognition and Signal Processing.* Amsterdam: Sijoff and Noordhoff.

Trehan, B., and C.E. Walsh. 1988. "Common Trends, The Government Budget Constraint And Revenue Smoothing." *Journal of Economic Dynamics and Control.* 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* 21: 291-306.

Uctum, M. and Wickens, M. 1997. "Debt And Deficit Ceilings, And Sustainability of Fiscal Policies: An Intertemporal Analysis." *CEPR Discussion Paper* no. 1612.

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.

Table 1: AIC Lag and Delay Orders Choice								
m \ k	1	2	3	4	5	6	7	8
1	0.892	0.951	0.917	0.852	0.853	0.942	0.980	0.992
2	-	0.933	0.902	0.835	0.854	0.938	0.949	0.989
3	-	-	0.904	0.870	0.879	0.963	0.974	0.999
4	-	-	-	0.906	0.887	0.955	0.980	1.01
5	-	-	-	-	0.914	0.947	0.970	0.989
6	-	-	-	-	-	0.988	0.986	1.01
7	-	-	-	-	-	-	0.982	1.02
8	-	-	-	-	-	-	-	1.10
Note: m is the delay order in the long difference: y_{t-1}-y_{t-m-1}, and k denotes the lag length.								

Table 2: HQ Lag and Delay Orders Choice								
m \ k	1	2	3	4	5	6	7	8
1	1.02	1.08	1.05	0.980	0.982	1.07	1.11	1.12
2	-	1.05	1.02	0.951	0.970	1.05	1.06	1.10
3	-	-	1.01	0.972	0.981	1.06	1.08	1.10
4	-	-	-	0.995	0.976	1.04	1.07	1.10
5	-	-	-	-	0.990	1.02	1.05	1.06
6	-	-	-	-	-	1.05	1.05	1.08
7	-	-	-	-	-	-	1.03	1.07
8	-	-	-	-	-	-	-	1.14
Note: m is the delay order in the long difference: y_{t-1}-y_{t-m-1}, and k denotes the lag length.								

Table 3: Threshold and Unit Root Tests for the TAR Model

Bootstrap Threshold Test [a]		
Wald Statistics = 25.8	Boot. P-value= 0.015	Asym p-value = 0.015
Two-Sided Wald Test for Unit Root [b]		
Wald Statistics = 16.6	Boot. P-value= 0.019	Asym p-value = 0.016
One-Sided Wald Test for Unit Root [b]		
Wald Stat = 16.2	Boot. P-value= 0.019	Asym p-value = 0.014
t_1 Test for Stationarity [c]		
t- Stat = 4.03	Boot. P-value= 0.004	Asym p-value = 0.006
t_2 Test for Stationarity [d]		
t- Stat = -0.587	Boot. P-value= 0.899	Asym p-value = 0.861

Notes: The p-values for the threshold and unit root tests were obtained by 5000 bootstrap replications. The results reported in the Table correspond to a trimming region (which defines the bounds π_1 and π_2 within which the threshold fall) given by the values [0.15, 0.85]. The results (available upon request from the authors) are robust to different trimming regions. We have used Gauss for all estimations reported in Table 3. (a) The Wald statistic tests the null of linearity against the alternative of threshold effect (the bootstrap p-values reported are for the unrestricted and unit root restriction cases). (b)he two-sided R_2 and one-sided R_1 Wald statistics test the null of unit root against the alternative that either the first or the second regime is threshold stationary (both the asymptotic and the bootstrap p-values are reported). (c)The t_1 statistic tests the null of unit root against the alternative that the first regime is stationary (both the asymptotic and the bootstrap p-values are reported).
(d) The t_2 statistic to test the null of unit root against the alternative that the second regime is stationary (both the asymptotic and the bootstrap p-values are reported).

Table 4: Estimation of TAR		
Variable	Regime 1	Regime 2
Intercept	-0.831 (0.262)	0.382 (0.220)
Y(t-1)	-0.137 (0.034)	0.018 (0.032)
DY(t-1)	-0.082 (0.107)	-0.238 (0.123)
DY(t-2)	0.010 (0.106)	0.221 (0.120)
DY(t-3)	0.155 (0.094)	-0.109 (0.101)
DY(t-4)	-0.237 (0.097)	-0.172 (0.095)
Notes: Standard errors in parenthesis.		

Figure 1: The Real Surplus Per Capita, 1947:2-2002:1

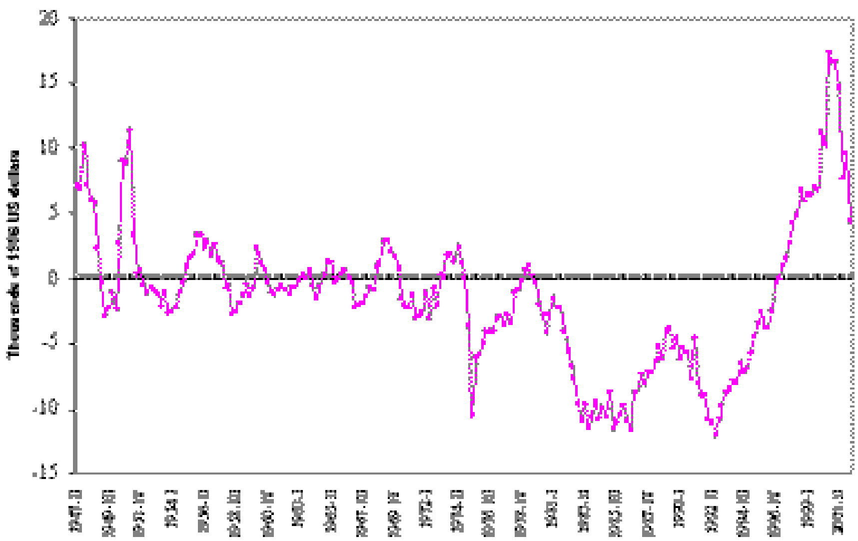


Figure 2: Confidence Interval Construction for Threshold

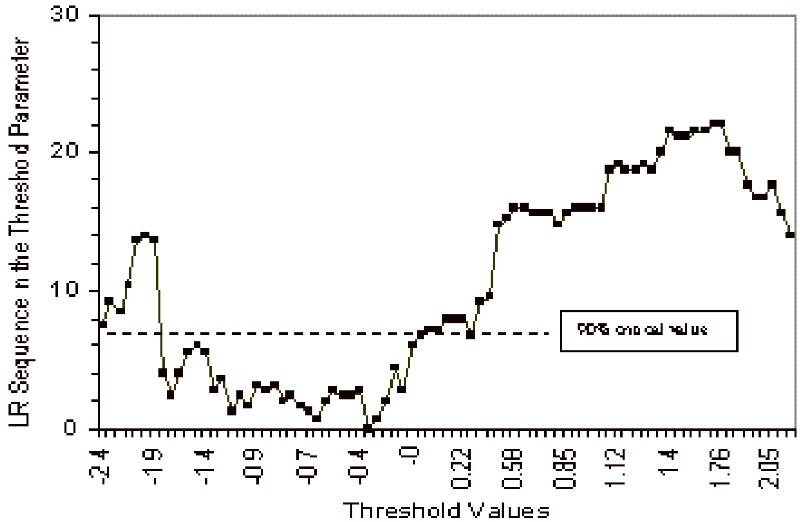
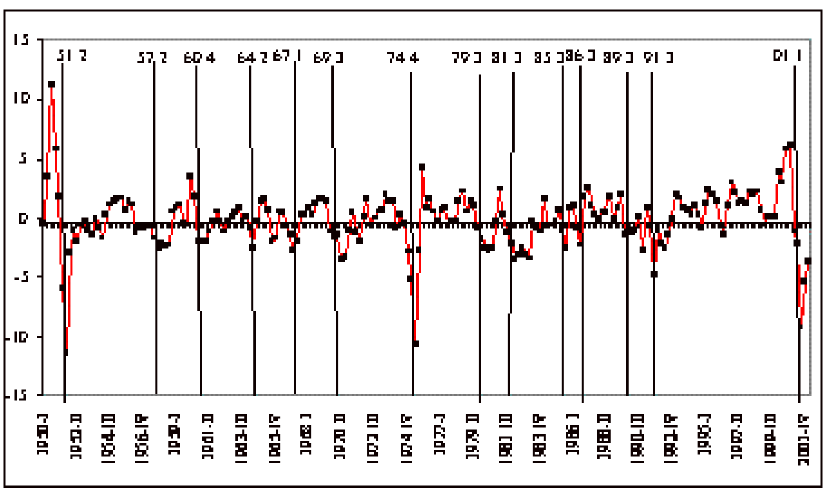


Figure 3: Identification of the Threshold Regimes



Note: This Figure plots the semi-annual change in the real surplus per capita, classified by the threshold regime. In other words, this Figure plots the deviation of threshold variables from the threshold point estimate (-0.313). Positive values identify the upper regime while negative values identify the lower regime. The lines superimposed on the graph represent the major switches from the upper to the lower regime. The figures next to the lines correspond to the dates of these major switches.