# BUDGET BALANCE THROUGH REVENUE OR SPENDING ADJUSTMENTS? SOME HISTORICAL EVIDENCE FOR THE UNITED STATES

by

**Henning Bohn**

**3-91**
(revision of 28-89)

RODNEY L. WHITE CENTER FOR FINANCIAL RESEARCH
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

# Budget balance through revenue or spending adjustments? Some historical evidence for the United States

by Henning Bohn*

Department of Finance, The Wharton School, University of Pennsylvania, Philadelphia, PA 19104-6367, USA

First Draft: July 1989. Current Version: January 1991

## ABSTRACT

The paper provides a historical perspective on the issue of whether budget deficits are typically eliminated by increased taxes or by reduced spending. By examining U.S. budget data from 1792-1988, I conclude that about 50-65% of all deficits due to tax cuts and about 65-70% of all deficits due to higher government spending have been eliminated by subsequent spending cuts, while the remainder was eliminated by subsequent tax increases. In contrast to previous studies, the empirical analysis uses error-correction models in a way that the intertemporal budget constraint is imposed in the estimation stage.

## 1. Introduction

Provided government policy is subject to an intertemporal budget constraint, high budget deficits must eventually be followed by higher taxes or by lower spending. The question which of these two methods the government should choose to eliminate high budget deficits has recently received considerable academic as well as political attention.[1] For forward-looking individuals, the question which method will be chosen is important for computing optimal consumption plans and for many other private sector decisions.

This paper examines how the U.S. government has adjusted taxes and spending in reaction to budget imbalances in the past, using data from 1792 to 1988. A long-term, historical perspective seems to be appropriate for two reasons. First, the intertemporal budget constraint does not prevent the government from running high deficits for substantial periods of time, provided the policy is changed eventually. The longer the sample period, the better are the chances of finding evidence that the government is starting to adjust to this constraint. Second, government debt is a variable that often moves very slowly but changes rapidly in rare instances such as wars. A long sample period permits averaging over episodes that might appear special in shorter samples.[2]

In estimating the government's response to deficits, econometric issues relating to the non-stationarity of time series are important. The data do not reject the non-stationarity of government spending, taxes, and government debt, even when these variables are measured as shares of GNP.[3] Since the deficit inclusive of interest payments turns out to be stationary—as required by the intertemporal budget constraint—taxes, spending, and debt are cointegrated. Error-correction models, or certain vector autoregressions which recognize this cointegration constraint, must be used.[4] These methods of estimation automatically impose the intertemporal budget constraint.

The question of how the budget is balanced is closely related to questions about the statistical causality between taxes and spending that have been examined previously.[5] The issues are whether or not tax changes are followed by spending changes, i.e., if the "tax-and-spend" hypothesis applies, and/or whether the converse, the "spend-and-tax" hypothesis, holds. I examine these hypotheses by separately estimating the adjustment processes following unexpected movements in taxes and in spending. The methodology differs from previous work in that the intertemporal budget constraint is imposed in the estimation stage.[6] The intertemporal budget constraint requires that any change in taxes and/or spending is followed by adjustments in future taxes and future spending that are equal to the original change in present value. The issue is therefore what fraction of the adjustment falls on the spending side and on the tax side of the budget. The estimated percentages of tax and spending adjustments can be taken as compact measures of the economic significance of tax-and-spend and spend-and-tax interactions.

The main results are that 50-65% of all deficits "caused" (in the statistical sense) by unexpected tax cuts and about 65-70% of all deficits caused by higher government spending are eliminated by subsequent spending cuts. The remainder is eliminated by tax changes. Thus, the data provide significant evidence in favor of both the tax-and-spend and the spend-and-tax hypotheses. Regardless of what was the original shock, a high deficit has typically been corrected by a combination of spending cuts and tax increases.

For interpreting these results, it is important that they are about average government behavior. In specific instances, e.g., in case of explicitly temporary tax rebates, all parties may know with certainty how the deficit will be eliminated. Then historical averages provide little guidance for decisions.[7] On the other hand, if the government has not announced specific plans to deal with a deficit (or surplus), or if its plans are not believed, it may be appropriate for individuals to use the average historical behavior of the government as a benchmark for predicting future government behavior.[8] Given no other information, individuals should interpret a $1 tax cut as signaling an increase in permanent after-tax income by 50-65 cents. Each $1 spending increase may be interpreted as signalling a tax increase with an annuity value of about 30-35 cents.

In terms of the Ricardian approach to debt policy (see Barro (1974), (1989)), these results reinforce Feldstein's (1982) warning that, even if consumers are "Ricardian," it is important to hold current and future government spending constant in assessing the effects of fiscal policy. Forward looking individuals will cut consumption in response to an unexpected tax increase, unless they can be convinced that the tax increase does not signal a rise in government spending. If expectations about future government behavior are based on the historical record of U.S. government behavior, Ricardian consumers should consider about 50-65% of any tax cut as an increase in net wealth. Thus, if policy-makers are simply interested in the question of whether consumption will rise when they cut taxes (without making any other announcements), the answer should be yes. This is the same answer that would be obtained, for other reasons, in non-Ricardian models. On the other hand, such consumption behavior cannot be taken as evidence against the Ricardian equivalence theorem.

The estimates have several other significant implications for public finance. The stationarity of the deficit (as share of GNP) suggests that government behavior has been consistent with the intertemporal budget constraint over the sample period of 1800-1988. This is interesting, since mixed results were obtained for shorter sample periods.[9] The result that deficits signal future tax changes may be taken as evidence against the random walk hypothesis for tax rates implied by tax-smoothing theories of government behavior.[10] The fact that tax increases are typically followed by spending increases has several possible interpretations (since it is reduced form evidence). It is consistent with a structural link between the government's financial position and government spending,[11] but also with exogenous spending and a signalling effect, as suggested by the tax smoothing model.

The paper is organized as follows. Section 2 derives the restrictions on fiscal policy imposed by the intertemporal budget constraint and reviews the econometric issues. Section 3 briefly summarizes the data. Section 4 presents the empirical results and Section 5 contains concluding remarks.

## 2. The Intertemporal Budget Constraint

This section will review the intertemporal budget constraint and its implications for the stochastic processes of taxes, spending, and debt.

### 2.1. Notation and Time Series Implications

The basic variables in the government budget are tax revenues, $T_t$, spending, $G_t$, and the beginning-of-period government debt, $B_t$. Government spending excludes interest payments and tax revenues include seignorage.[12] These variables are linked by the budget equation

$$B_{t+1} = G_t - T_t + (1+r) \cdot B_t + \varepsilon_{t+1}. \tag{1}$$

where r is the interest rate. The error term $\varepsilon_{t+1}$ allows for the fact that the observed time series of $T_t$, $G_t$, and $B_t$ do not satisfy an exact linear constraint. If the expected real return on government debt is constant, as assumed in much of the literature (see Barro (1979), Hamilton and Flavin (1986), Trehan and Walsh (1988)), r is the expected real return and $\varepsilon_{t+1}$ is uncorrelated. If not, $\varepsilon_{t+1}$ is the error made in approximating the return on debt by a fixed number r (perhaps best interpreted as notional interest charge). Then $\varepsilon_{t+1}$ may be autocorrelated. For the statistical analysis, it is important that $\varepsilon_{t+1}$ is stationary and that $r>0$.[13]

In principle, the government budget constraint and all the analysis below can be stated in terms of real variables, nominal variables, or relative to a scale variable like GNP, provided r is interpreted appropriately. In this theoretical section, it is convenient simply to refer to taxes, spending, and debt without being more specific. The empirical analysis will use GNP-shares.

The objective of the paper is to explore how future fiscal policy reacts to current economic and fiscal changes. In statistical terms, the issue is how the expected values of future taxes, spending, and debt are affected by innovations in current variables. To be precise about expectations, let $X_t$ be the vector of variables in the information set at time t and let $E_t[\cdot]$ be the expectation conditional on $X_t$ and its lags. The vector $X_t$ is assumed to include $T_t$, $G_t$, and $B_t$, but not $B_{t+1}$ and $\varepsilon_{t+1}$,[14] and it may include additional variables like GNP or inflation. The number of elements in $X_t$ will be denoted by n.

In projecting future policy variables on the period-t information set, it is important to recognize that the government budget constraint restricts the joint movements of fiscal variables. Provided the standard transversality constraint holds, the evolution of the vector $(T_t, G_t, B_t, \varepsilon_{t+1})$ is subject to the intertemporal budget constraint

$$(1+r) \cdot B_t = \sum_{j\geq 0} \rho^j \cdot E_t[T_{t+j}-G_{t+j}-\varepsilon_{t+j+1}], \tag{3}$$

where $\rho=1/(1+r)$ is the discount factor. Trehan and Walsh (1988) have shown that this intertemporal budget constraint requires that government debt is stationary in first differences. The stationarity of $\Delta B_t$ implies a cointegration restriction on the vector of fiscal variables $X_t^f=(T_t, G_t, B_t)$: From equation (1), the co-integrating vector is $(1,-1,-r)$ and the stationary linear combination is the familiar budget deficit (inclusive of interest),

$$DEF_t = G_t - T_t + r \cdot B_t \tag{4}$$

If the vector $X_t$ has additional elements that are all difference stationary and if one assumes that $X_t$ contains no other stationary linear combination, $X_t$ will also be co-integrated with co-integrating vector $\beta = (1,-1,-r,0,..,0)$.

Engle and Granger (1987) have shown that such a cointegrated process X has an error-correction representation

$$A(L)\ \Delta X_t = - \alpha \cdot \beta' X_{t-1} + u_t, \tag{5}$$

or

$$A(L)\ \Delta X_t = \alpha \cdot DEF_{t-1} + u_t, \tag{6}$$

where $A(0)=I$, $A(1)$ is finite, and $u_t$ is a stationary disturbance vector.[15] I will also assume that $u_t$ has zero autocorrelation and that $A(L)$ is of finite order k.

Projections of future policy on the information set can be estimated in several ways. First, if one rearranges system (6) as a first order system augmented by the identity

$$DEF_t = - \beta' \Delta X_t + DEF_{t-1}, \tag{7}$$

one obtains the first order stochastic difference equation

$$X^*_t = A^* \cdot X^*_{t-1} + u^*_t,$$

where $X^*_t=(\Delta X_t',\ldots,\Delta X_{t-k-1}',DEF_{t-k})'$ and $u^*_t=(u_t',0,\ldots,0)'$ are $(n \cdot k+1)$-vectors. The matrix $A^*$ is a $(n \cdot k+1)x(n \cdot k+1)$ matrix of coefficients of the form

$$A^* = \begin{pmatrix} A_1^* & A_2^* & \ldots & A_k^* & \alpha \\ I & 0 & \ldots & 0 & 0 \\ \vdots & & & \vdots & \vdots \\ 0 & \ldots & I & 0 & 0 \\ 0 & \ldots & 0 & -\beta' & 1 \end{pmatrix},$$

where the elements of the n x n sub-matrices $A_i^*$ are functions of the coefficients in A(L), $\alpha$, and $\beta$.[16] The j-period ahead response of a variable in X to a period-t innovation can be found in the appropriate row of the matrix $(A^*)^j$.[17]

A useful alternative has been derived by Campbell (1987). By replacing one element of $\Delta X^f$ by DEF, using (7) repeatedly, one obtains a vector autoregression (VAR) representation

$$C(L)\, Y_t = v_t, \tag{8}$$

where $Y_t$ contains DEF and the n-1 remaining elements of $\Delta X_t$. C(L) is a $k^{th}$ order polynomial and $v_t$ is a linear transformation of $u_t$. Campbell shows that unconstrained estimates of this VAR automatically satisfy the cointegration restrictions. The VAR yields impulse-response functions in the usual way.

Finally, in some cases I will use a simplified estimation strategy motivated by the fact that $B_t$ enters into the cointegrating vector only with weight r, which may be small. If the impact of $B_t$ is indeed small (in the limit $r \rightarrow 0$), the dimensionality of the estimated system of equations can be reduced by approximating DEF by the primary deficit (G-T) and by excluding B from X.

In all econometric specifications, we are interested in the response of future taxes and future spending to current innovations in taxes and in spending, respectively. Support for the tax-and-spend (or spend-and-tax) hypothesis would be found if one were able to reject the hypothesis that a tax change has no effect on spending (or a spending change has no effect on taxes) in any future period. For each specification one would have to examine several impulse response functions. The next subsection will show, however, that there is a more compact and perhaps equally informative way of assessing the intertemporal connections between taxes and spending.

### 2.2. Implications for Present Values

The intertemporal budget constraint requires that an increase in government spending (not accompanied by a change in current taxes) must either be reversed or eventually be followed by a tax increase. The future tax increases and the future spending reductions together must be equal to the initial shock in present value. This suggests that the fraction of the future changes accounted for by taxes may be taken as measure of the spend-and-tax aspect of government behavior. Similarly, the present value of the spending response as fraction of an impulse to taxes may be taken as measure of the tax-and-spend effect.

To start, it is useful to make explicit how the intertemporal budget constraint (3) restricts the responses of future spending and taxes. For any variable z, denote the discounted sum of future realizations by $PV(z)_t=\sum_{j\geq 1}\rho^j\cdot z_{t+j}$ and denote the innovations in a variable and its present value by $\hat{z}_t = z_t - E_{t-1}z_t$ and $\widehat{PV}(z)_t = E_tPV(z)_t - E_{t-1}PV(z)_t$, respectively. The budget constraint (3) can then be rewritten as

$$(1+r)\cdot B_t = T_t + E_tPV(T)_t - [G_t + E_tPV(G)_t + E_t(PV(\varepsilon)_t)/\rho], \tag{9}$$

In terms of innovations, this implies

$$\hat{T}_t + \widehat{PV}(T)_t = \hat{G}_t + \widehat{PV}(G)_t + (1+r)\cdot\hat{B}_t + \Omega_t/\rho \tag{10}$$

for all possible realizations of $\hat{T}_t$, $\hat{G}_t$, $\hat{B}_t$, where $\Omega_t = E_t[PV(\varepsilon)_t] - E_{t-1}[PV(\varepsilon)_t]$ is an error term. Since most variables in (10) have unit roots, it is useful to restate the equation in terms of stationary first differences. Using the the identity

$$(1-\rho)\cdot[z_t + PV(z)_t] = z_t + PV(\Delta z)_t \tag{11}$$

and the fact that $\Delta\hat{X}_t=\hat{X}_t$,[18] equation (10) is equivalent to

$$\Delta\hat{T}_t + \hat{PV}(\Delta T)_t = \Delta\hat{G}_t + \hat{PV}(\Delta G)_t + r \cdot \Delta\hat{B}_t + r \cdot \Omega_t. \qquad (12)$$

Over time, fiscal policy evolves as a series of innovations in $\Delta\hat{T}_t$, $\Delta\hat{G}_t$, and $\Delta\hat{B}_t$. Whenever a set of policy innovations is realized, plans about future policy must change in a way consistent with (10) and (12). Note that the expressions $\hat{PV}(\Delta T)_t$ and $\hat{PV}(\Delta G)_t$ are discounted sums of the values taken by the impulse response functions for different horizons and that the error term $\Omega_t$ is simply a linear combination of elements in $\Delta\hat{X}_t$.

The expressions in (12) have a straightforward interpretation in terms of permanent and temporary policy changes. In general, the permanent component of a variable z is defined as $z_t^p = (1-\rho)\cdot\sum_{j\geq 0}\rho^j\cdot E_t z_{t+j}$. Using the identity (11), its innovation is $\hat{z}_t^p = \Delta\hat{z}_t + \hat{PV}(\Delta z)_t$. The expressions $\Delta\hat{T}_t + \hat{PV}(\Delta T)_t$ and $\Delta\hat{G}_t + \hat{PV}(\Delta G)_t$ in (12) are therefore the innovations in permanent taxes and permanent spending. The equation shows that changes in permanent taxes and permanent spending must be equal, unless initial debt or the error term change. The important implication for fiscal policy is that any innovation in current taxes or spending must lead to changes in the other variable (spending or taxes) to the extent that the change is permanent. Thus, the tax-and-spend hypothesis is closely linked to the permanence of tax changes, and the spend-and-tax hypothesis is similarly linked to the permanence of spending changes.

To be more explicit about these links, consider an innovations in current spending, $\Delta\hat{G}_t$, not accompanied by a change in taxes and debt and, for now, ignore the error term. (The analysis of an innovation in taxes is analogous. The error term is discussed below.) Equation (12) shows that the sum of later downward revisions in spending, $-\hat{PV}(\Delta G)_t$, and later tax increases, $\hat{PV}(\Delta T)_t$, must add up to the known amount $\Delta\hat{G}_t$. If spending had no impact on taxes ($\Delta\hat{T}_{t+i}=0$ for all i), one would find $\hat{PV}(\Delta T)_t=0$ so that $\hat{PV}(\Delta G)_t=-\Delta\hat{G}_t$ would have to hold. If all spending changes were permanent (i.e., $\hat{PV}(\Delta G)_t = 0$), the impact of spending on taxes would have to be large, $\hat{PV}(\Delta T)_t=\Delta\hat{G}_t$. In intermediate cases, the present value of the tax response measures the strength of the spend-and-tax link in fiscal policy.

For the estimation, note that the present values of future changes in taxes and spending are linear combinations of the current innovations. For any variable z defined on the process X, let f(z) be the vector of projection coefficients for the present value of its future changes,

$$\hat{PV}(\Delta z)_t = f(z) \cdot \Delta\hat{X}_t = f_{X_1}(z) \cdot \Delta\hat{X}_{t1} + \ldots + f_{X_n}(z) \cdot \Delta\hat{X}_{tn} \tag{13}$$

where $f_{X_i}(z)$ indicates the marginal effect of an innovation in the $i^{th}$ element of $X_t$, $\Delta\hat{X}_{ti}$, on $\hat{PV}(\Delta z)_t$. Thus, the spend-and-tax measure can be computed as the coefficient estimate $f_G(T)$, which indicates the marginal impact of a unit movement in $\Delta\hat{G}_t$ on $\hat{PV}(\Delta T)_t$.[19]

In general, the interpretation is slightly complicated by the presence of the error term $\Omega_t$ in equation (12). If one writes $\Omega_t$ as linear combination of the innovations in $\Delta X_t$, $\Omega_t = \omega \cdot \Delta\hat{X}_t$ with elements $\omega_{X_i}$, the present value effects $\hat{PV}(\Delta T)_t$ and $-\hat{PV}(\Delta G)_t$ will not exactly add up to $\Delta G_t$ whenever $\omega_G \neq 0$. In terms of projections, the spend-and-tax measure $f_G(T)$ and the permanent effect of an innovation in spending on future spending, $1+f_G(G)$, will not be equal. Still, if the error is small, $f_G(T)$ should be close to $1+f_G(G)$; the discrepancy may be taken as an indicator of how well the approximation of bond returns by the constant r in equation (1) fits the data.[20]

Similar arguments apply for the analysis of innovations in taxes and in other variables. Estimates of $f_T(G)$ can be taken as measuring the effect of taxes on spending. The value of $f_T(G)$ will be close to $1+f_T(T)$, if the error $\omega_T$ is small.[21] Non-zero values of $f_T(G)$ would support the tax-and-spend hypothesis.

In the context of tax forecasts, it should be noted that the tax-smoothing theory of debt policy (Barro (1979)) implies that tax rates are a random walk. Provided marginal and average tax rates are monotonically and approximately linearly related, one should be unable to predict changes in future tax revenues.[22] In this sense, non-zero estimates of $f_G(T)$ or $f_T(T)$ can be interpreted as evidence against tax-smoothing.

The present value projections are particularly interesting as measures of fiscal policy in an economy with "Ricardian" consumers. A rational, forward-looking individual (or dynasty) that satisfies the Ricardian equivalence theorem will set consumption proportional to permanent disposable income (see, e.g., Campbell (1987)). Permanent income depends on fiscal policy only through the permanent component of taxes.[23] Since the innovation in permanent taxes is $\hat{\Delta T}_t + \hat{PV}(\Delta T)_t$, the impact of an unexpected current change in taxes on consumption is proportional to $1+f_T(T) \approx f_T(G)$ and the impact of an unexpected current change in spending is proportional to $f_G(T)$. The present value projections are "sufficient statistics" in the sense that they characterize the aspects of fiscal policy relevant to Ricardian consumers. This provides an additional motivation for using them as concise measures of the interaction between current and future policy variables. In the empirical analysis, much space can be saved by displaying and discussing such summary statistics instead of the complete impulse-response functions.

To compute the present value projections in the context of the error-correction model (5), note that each projection is the discounted sum of impulse-response terms. If the vector $h_s$ selects $\Delta X_{ts}$ from $X^*_t$ (where $X_s$=T and $X_s$=G will be needed), one has

$$\hat{PV}(\Delta X_s)_t = h_s \cdot \sum_{j\geq 1}(\rho A^*)^j \cdot u^*_t = h_s \rho A^* \cdot [I - \rho A^*]^{-1} \cdot u^*_t.$$

Since the period-t innovation in variable $X_i$ (again $X_i$=T and $X_i$=G will be needed) is the i-th element in $u^*_t$, the marginal effect of an innovation in $X_i$ on the present value of $\Delta X_s$ is the element (s,i) in this matrix expression, i.e.,

$$f_{X_i}(X_s) = \{\rho A^* \cdot [I - \rho A^*]^{-1}\}_{si} \tag{14}$$

where $\{M\}_{si}$ denotes the element in row s and column i of a matrix M. An analogous formula holds for the VAR-system (8).

## 3. Data

The intertemporal budget constraint does not prevent the government from running high budget deficits or surpluses for long periods of time. The transversality constraint restricts only the asymptotic behavior of government debt. Given the considerable inertia in the level of debt (with the notable exception of war years), one should suspect that a very long-term analysis is needed to obtain significant insights about government debt policy. Therefore, this study uses the longest data set available for the United States.

Early budget data, from 1792-1970, are available in the Historical Statistics of the United States. The basic budget series are federal government receipts, federal government outlays, outlays for interest payments, and debt. The series were updated and extended to 1988, using the Historical Tables, Budget of the United States and other supplementary sources. (Details are in the appendix.) The use of receipt and outlay series means that government spending and taxation are defined in a wide sense; for example, government outlays for transfers are included on the spending side of the budget and not netted-out on the revenue side.[24]

The budget data are annual and collected on a fiscal year basis. From 1789-1842, fiscal years are calendar years. Fiscal year (FY) 1843 covers January to June 1843. From FY1844 to FY1976, the fiscal year dated t covers the period from July of calendar year t-1, to June of year t. After the transition quarter of July-September 1976, fiscal years from FY1977 onward begin in October of year t-1 and end in September of year t. The 1976 transition quarter was included in FY1976, as many statistical tables do. Flow variables for FY1843 and FY1976 were annualized to obtain a consistent fiscal-year series from FY1792 through FY1988. A transformation to calendar years (or to a common starting month) was not attempted, because the fiscal year is the basic time interval for policy decisions and because of time aggregation problems implied by any transformation. For government debt, only par values are available. But since the focus is on the long-run, temporary differences between market and par values should not be critical. (The algebra of budget constraints is valid for par as well as for market values, provided r and $\varepsilon_{t+1}$ are interpreted appropriately. In any case, an error term is needed in equation (1) and its impact is an empirical question.)

As noted above, the theoretical analysis does not distinguish between real, nominal, or otherwise rescaled variables. Since the nation's productive activity provides the basis for all taxation and since the government sector is presumably bounded by the size of the whole economy, it seems appropriate to use ratios of fiscal variables to GNP. The use of GNP-shares also mitigates the heteroskedasticity problems that would be severe in long-run series if unscaled variables were used. The parameter r must then be interpreted as the difference of the real discount rate and the real rate of economic growth.

Unfortunately, a timing problem is inevitable in defining fiscal variables as GNP-ratios, because GNP is only available on a calendar year basis.[25] For most of the period, fiscal year (FY) t starts before calendar year t. I take the start of a fiscal year as the relevant time for budgetary decisions and deflate revenue and spending of FY t and debt at the beginning of FY t by the GNP of calendar year t-1. Fortunately, the timing issue does not seem to be important for the results (see estimate No.12 in Table 6 below). But because of the mismatch in the timing of GNP and revenue series, negative results on tax-smoothing must be interpreted very cautiously.

I did not attempt any adjustments to the budget data, in spite of the widespread dissatisfaction with federal accounting practices (see Eisner (1989)). The reasons is that the analysis of official budget data should be most informative about government behavior if policy-makers are primarily influenced by the official data. This is likely, since the official numbers are the most easily and most widely available data set. Thus, it seems appropriate to use the official budget numbers as they were originally defined and collected (even if their use may be more questionable for other purposes; see Eisner).

To sum up, the empirical analysis will use the following three series: (1) Federal outlays net of interest payments deflated by GNP, which is the empirical representation of the spending variable $G_t$; (2), federal receipts deflated by GNP, which is the empirical representation of the tax variable $T_t$; and (3), debt at the beginning of a fiscal year, deflated by GNP, which is the empirical counterpart to the variable $B_t$. Complete data are available for 1792-1988. To have enough lags in the estimation, the sample period 1800-1988 is used throughout the empirical analysis. The series are graphed in Figures 1 and 2. Summary statistics are in Table 1.[26]

In Figures 1 and 2, one can see the impact of the three major wars (Civil War, World War I, and World War II), which led to sudden increases in outlays and debt. The share of government activity in GNP changed significantly over the sample period, even excluding wartime peaks, from close to zero in the late 18th century to about 20% currently. For the empirical analysis, the graphs suggest that all three variables may be non-stationary and heteroskedastic.

## 4. Empirical Results

In this section, the time series properties of the data are analyzed and the reactions of future taxes and future spending to changes in current fiscal policy are estimated.

### 4.1. Time Series Properties

Preliminary, but important issues are the questions of stationarity and cointegration. Even though the tax-GNP and spending-GNP ratios are bounded by [0,1], they do not have to be stationary (see Ahmed and Yoo (1989)).

Several tests for stationarity are displayed in Table 2. For each series, two augmented Dickey-Fuller tests (ADF) with different lag lengths were computed (see Fuller (1976), Dickey and Fuller (1981)). The first lag length was based on the Akaike information criterion computed in AR(k)-models of the differenced series (i.e. under $H_0$ that the series is non-stationary). Second, the "long lag" version suggested by Said and

Dickey (1985) and Schwert (1987) was used (with k=13). A problem with this type of test is that the significance levels are not valid for heteroskedastic data. Therefore an adjusted version due to Phillips and Perron (1986) and Phillips (1987), which is valid under heteroskedasticity, was also computed.[27] All these tests can be coefficient-based or based on a t-statistic. For the Phillips-Perron test, both statistics are displayed and denoted by $Z\rho$ and $Z\tau$, respectively. For the ADF tests, the two t-type statistics are displayed under ADF(·); the coefficient-based statistics are omitted to save space. (They produced the same inferences.) Table 2 shows that one cannot reject non-stationarity of taxes, spending, and debt at the 5% level in any test. In contrast, non-stationarity of the first differences can be rejected with high significance in all tests.

The cointegration of taxes, spending, and debt can be tested by analyzing the stationarity of the budget deficit, since the deficit series DEF defined in (4) should be the stationary linear combination of $(T_t,G_t,B_t)$. Table 2 displays the test results for different values of r.[28] All of them indicate stationarity. Alternative tests for cointegration are obtained by examining more generally whether $X^f_t$ has stationary linear combinations. Maximum likelihood estimates of the cointegrating vectors and tests for the number of cointegrating vectors based on Johansen (1988) are computed in Table 3, Panel A. One cannot reject the hypothesis that the number of cointegrating vectors is 1, while one can reject the hypothesis that there are zero cointegration restrictions. After normalization, the estimated vector $\beta=(1,-0.877,-0.060)$ is close to the theoretically expected vector of the form (1,-1,-r).[29] Specifically, one cannot reject the hypothesis that $\beta=(1,-1,-r)$ for a wide range of values for r, between -1% and 14%.[30] The point estimate of r subject to the restriction that $\beta$ has the form (1,-1,-r) is r=2.90%.

Since r=0 was not rejected, the properties of the bivariate system $(T_t,G_t)$ were also explored. As Panel B of Table 3 shows, one cannot reject the hypothesis that the system has one cointegrating vector. The point estimate of $\beta=(1,-0.985)$ is not significantly different from (1,-1).

As another alternative, Engle and Yoo (1987) suggest a Dickey-Fuller test on the residual of a regression of one of the potentially cointegrated variable on the others. Here, an augmented Dickey-Fuller test (with one lag) of the residual from regressing $T_t$ on $G_t$ and $B_t$ yields a t-statistic of 3.783, which is barely significant.[31] Since the cointegrating vector was not measured very precisely, this approach was not pursued further.

Overall, one may conclude that $X^f_t$ is non-stationary and cointegrated with a cointegrating vector of the form (1,-1,-r) for a range of discount rates r. As Engle and Granger (1987) show, the statistically valid estimation and testing of such a vector process should use error-correction models.

### 4.2. Error-Correction Estimates

Having verified that the fiscal data are co-integrated, error-correction models (ECMs) of the form (6) and VARs of the type (8) are estimated to compute impulse response functions. Estimates were computed for different interest rates, for several different information sets, and for the limiting case $r \to 0$. Since the results were remarkably similar across all specifications, I will only describe one set of estimates in detail and present summary statistics of the others.

Table 4 displays the results of estimating (6) with the minimal information set $X=X^f$ that includes only the three fiscal variables, using r=3%.[32] Coefficient estimates with ordinary and heteroskedasticity-robust t-statistics (see White (1980)) are in Panel A. One can see that the deficit, the co-integrating variable, has a highly significant positive effect on taxes and a negative (though weaker) effect on spending. This is very much in line with the intuition that deficits generate stabilizing long-term tax and spending adjustments. The F-tests of the null hypothesis that all lags of a variable are zero indicate that spending and taxes appear to influence each other and debt, while debt has little direct effect on taxes and spending (though it enters indirectly through DEF).

To visualize the combined effect of the lags and the co-integrating vector, impulse response functions are graphed in Figures 3 and 4. Figure 3 displays the i-period ahead effects of a unit innovation in $\Delta T_t$ on $\Delta T_{t+i}$ and on $\Delta G_{t+i}$ in Panel A and the corresponding responses to a unit innovation in $\Delta G_t$ in Panel B. Figure 4 shows the implied effects on the levels of taxes and spending, which may be more interesting. One can see in Panel A that an innovation in taxes is partially reversed in subsequent periods, though about one half of the impulse persists. The tax innovation is followed by a strong spending reaction one period later, which is also partially reversed. Panel B shows that a spending innovation is typically followed by further spending increases (yielding a response above one) and that most of the increase is reversed. Only about 25% of the increase seems persistent. The lagged response of taxes to higher spending is positive and highly persistent with the level of taxes staying about 0.30 above the initial level.[33]

For a more compact assessment of the tax and spending interactions, Table 4B shows estimated present value projections $f_{X_i}(z)$ with their asymptotic standard errors and t-statistics. One can see that an unit innovation in taxes will, on average, be followed by a decrease in taxes of $-f_T(T)=0.49$ and an increase in spending of $f_T(G)=0.50$. Both values are significantly different from both zero and one. It is reassuring that the sum $-f_T(T) + f_T(G) = 0.99$ is close to 1.0, leaving only -0.01 to the error term $\Omega_t$ (defined in section 2.2). Also, the present value projections match up well with the permanent effects of a tax impulse suggested by the impulse-response graph.

The result that about half of all tax increases are permanent and signal spending increases provides strong support for the tax-and-spend hypothesis. The permanence of tax cuts also implies that Ricardian consumers should react rather strongly to current tax changes (unless, of course, they have case-specific information that a change is unusually temporary), namely with a marginal propensity to consume of about 0.50 in a traditional consumption function framework (see Bernheim (1987)). The 50% temporary component in taxes is somewhat troubling for tax smoothing as a positive theory of policy. Though a large fraction of all tax changes are explained as anticipation of spending changes, which gives the tax-smoothing model some explanatory power, some tax changes are apparently reversed.[34]

Turning to government spending, Table 4B shows that a unit innovation in government spending increases future taxes on average by only 0.30 and that it is followed by spending cuts of 0.64.[35] Thus, most changes in spending seem to be temporary. Once heteroskedasticity is taken into account, one cannot even reject the hypothesis that spending does not cause tax changes. (But significant effects were found in other specifications; see below.)

To examine the robustness of the results, several other specifications were estimated. To save space, only the key projection coefficients and their (robust) t-statistics are displayed. Table 5 shows the results for other specifications based on the information set ($T_t$, $G_t$, $B_t$). For comparison with Table 4, regression No.1 replicates the results of the initial specification. Regressions No.2-5 show that the results are not sensitive to the interest rate r or to the lag length. Regression No.6 uses an alternative definition of seignorage.[36] Regression No.7 uses the assumption that end of period debt, $B_{t+1}$, is in the period-t information set. Regressions 8-10 use the VAR specification (8), as suggested by Campbell (1987). None of the modifications appear to change the qualitative results from Table 4.[37]

Projections for the limiting case of $r \to 0$ are displayed in Table 6. The approximation $r \approx 0$ seems justified by the fact that $\Delta B$ appears to be rather unimportant in the three-variable system. The limiting case has several nice features. The three variable system reduces to two variables, which makes the estimates even more precise. The error terms in (12) vanishes so that the adjustments of spending and taxes add up to exactly 1.0. A small system is also a convenient starting point to explore the addition of other variables to the information set without using up too many degrees of freedom.

Table 6 displays the results for the limiting case. Regressions No.1-3 estimate the basic bivariate ECM and the two related VAR specifications. The results are roughly similar to those based on the full three-equation models in Tables 4 and 5. The response of future taxes to an initial tax change is even smaller than in the full model, only about 0.35, but still significant. The spending side accounts for 65% of the adjustment. The response of future taxes to an impulse to spending is very similar to the previous estimates, namely between 26% and 32%, depending on the specification. Here the response is clearly significant. Thus, there is significant evidence for the spend-and-tax direction of (statistical) causality.

For all these projections, it should be acknowledged that individuals and the government likely observe a much larger set of variables than any econometrician could ever put into a VAR. Still, one can explore whether the results are sensitive to the inclusion of plausible additional variables into the information set X. Previous studies of government behavior have focussed on output and inflation as likely determinants (e.g., Barro (1979) and (1986), Kremers (1989), Horrigan (1986)). Taking the basic bivariate ECM No.1 as starting point, regressions No.4-7 in Table 6 show the effects of including real GNP growth and inflation, each with or without the current value.[38] In regressions 4-7, the estimated effects of growth and inflation on the present value of taxes (which, because of the budget constraint, are equal to the effects on the present value of spending) are 0.04 (t=0.5), 0.05 (t=0.7), 0.065 (t=0.2), and -0.05 (t=-0.2), respectively. None of these values is significant, though inflation and growth cannot be excluded from the $\Delta T$ and $\Delta G$ equations. (In each specification, the exclusion of all their lags is rejected by the F-test for at least one of the two equations.) The projections for tax and spending impulses remain virtually unchanged.

The result that no variable other than G and T seems to matter is likely due to the fact that the present value projections largely reflect long-run adjustments and not short-term or cyclical dynamics, for which output and inflation may well be important. For the long-term analysis, it is important that proportional movements of fiscal variables with growth and inflation are already included through the definition of fiscal variables as GNP-shares (which differs from the studies referenced above). Economically, a significant value for any variable other than T and G in the projections would mean that that variable permanently changed the share of the government sector in the economy. Growth could possibly cause such shifts, if the demand for government spending had an income elasticity different from one. But the effect turns out to be small. For inflation, a large effect would have suggested significant real effects of inflation.

Time-variation in government behavior could be a problem for the interpretation of all the long-run estimates. Regressions 8-11 in Table 6 therefore re-run the basic bivariate ECM No.1 for four sub-periods. Regressions 8 and 9 split the sample in 1917, which marks a shift in taxation from a tariff-based to an income-based system (see Gardner and Kimbrough (1989)). The period 1918-1988 is also close to the period considered by previous fiscal policy studies (e.g., Barro (1979) and (1986)). Regression 10 excludes the Civil war period and prior years, for which data may not be as reliable. Regression 11 explores whether war periods are somehow special by excluding the three major wars (One might have suspected that the "big movements" during the wars drive the previous results.)

As Table 6 shows, estimates No.8-11 do not differ significantly from those for the overall sample. The exclusion of war periods seems to reduce the estimated effect of government spending on taxes slightly (to 0.20), but it reduces the standard error even more so that the effect ends up being more significant than in model No.1.

For completeness, I also looked at the post-war period, using a completely separate data set of quarterly National Income and Product Account data for 1954:1 to 1988:3. The estimates based on the basic bivariate ECM were $f_T(T)=-0.43$ and $f_G(T)=0.37$. Thus, more recent data seem to be consistent with the long-run estimates. Since the long-run estimates should be preferable for the reasons outlined before, a more detailed analysis of postwar data was not undertaken.

To explore whether the misalignment between fiscal and calendar years is important, I also examined a set of fiscal data that were rescaled in different way. Fiscal variables were deflated by a fiscal year GNP series obtained by interpolating annual values, instead of using the timing conventions of Section 3. The estimates based on the alternative data set are shown as regression No.12. The method of scaling seems to make little difference. Overall, the basic results seem to hold up in all specifications.

To summarize, if one asks how federal deficits per se have typically been eliminated, the quantitative answer depends somewhat on the source of the deficit, whether it was due to tax or spending innovations, and on the statistical model. In all cases, the point estimates suggest a "political compromise" solution of both tax increases and spending cuts. Depending on the source of the shock and on the model, the adjustments have been between 50% and 70% on the spending side (near 70% for spending innovations, 50-65% for tax innovations) and between 30% and 50% on the revenue side.

## 5. Conclusions

The paper has analyzed long-term U.S.-budget data to explore the responses of taxes and spending to changes in fiscal policy. Tax changes are found to signal substantial spending changes, implying that on average 50%-65% (depending on the model specification) of a deficit triggered by a negative shock to taxes has been eliminated by reduced spending and 35%-50% by higher future taxes. The hypothesis that the ratio of taxes to GNP is a random walk is rejected. About 30-35% of all changes in government spending are found to be permanent and accommodated by subsequent changes in taxes.

## Footnotes

[*] An earlier version of this paper was presented at the 1990 National Bureau of Economic Research Summer Institute and at the 1990 Western Economics Association Conference. I wish to thank Ben Bernanke, Andy Abel, Joe Haubrich, Robert Inman, Jim Poterba, Jeremy Siegel, and an anonymous referee for many valuable comments and suggestions. I retain, however, full responsibility for all remaining errors.

[1] See, e.g., the Spring 1989 issue of the Journal of Economic Perspectives. Money creation is sometimes mentioned as third alternative (see, e.g., Sargent and Wallace (1982)). Here, inflation taxes are included on the tax side.

[2] For example, studies using post-1929 (start of National Income and Product Accounts (NIPA)) or post-1916 data (used by Barro (1979) and others), can easily be dominated by World-War II. On the other hand, if World-War II is excluded, one is left with a peace-time sample that may not be representative of U.S. history, which does include wars.

[3] This is consistent with Ahmed and Yoo (1989). Many earlier studies of fiscal policy used detrended data, e.g. von Furstenberg, Green, and Jeong (1986), Horrigan (1986), and Barro (1986).

[4] See Trehan and Walsh (1988), Engle and Granger (1987), Campbell (1987).

[5] See von Furstenberg et.al. (1986), Miller and Russek (1990), and the references therein.

[6] Von Furstenberg et.al. (1986) use VAR-techniques. Miller and Russek (1990) estimate error-correction models, but they do not impose the intertemporal budget constraint.

[7] Recall, however, that the 1975 tax changes were originally introduced as tax rebate and then extended indefinitely (see Blinder (1981)). Thus, one may doubt whether individuals should take announcements about future tax policy at face value.

[8] The implicit stationary assumption may be unobjectionable, if government behavior reflects stable preferences, say, of a representative individual. But it may be worth exploring (currently being studied by the author) whether government behavior towards

deficits is changing over time, e.g., due to institutional factors. It is reassuring, though, that results do not differ much across sub-periods (see Section 4).

[9] See Hamilton and Flavin (1986), Hakkio and Rush (1986), Trehan and Walsh (1988), Wilcox (1989), and Kremers (1989).

[10] See Barro (1979, 1986), Sahasakul (1987), and Bohn (1990a).

[11] See Bohn (1988) for a structural model. Modern social choice theory (e.g., Tabellini and Alesina (1990)) also suggests such a link.

[12] Interest payments are taken out of spending, because the primary deficit is relevant for the intertemporal budget constraint (see McCallum (1984)). On the tax side, it would only be distracting to treat the inflation tax as a special case.

[13] Following the literature (Trehan and Walsh (1988) in particular), I treat the validity of these assumptions as an empirical issue. See Bohn (1990b) for a theoretical analysis.

[14] The alternative of known $B_{t+1}$ (the debt at the end of period t) will also be considered in the empirical part; the formal analysis is similar and therefore omitted.

[15] All vectors should be interpreted as column vectors. A prime (') indicates a transpose. The first element of $\beta$ is normalized to 1.0 and the first three elements in X are T, G, and B (in this sequence), which implies $DEF_t = - \beta'X_t$.

[16] Elements of A* are obtained from (5) by using (7) repeatedly until $DEF_{t-1}$ is replaced by $DEF_{t-k-1}$. The n x 1 column vectors $\alpha$ and $\beta$ are the same as in (5).

[17] That is, let the variable z be the $s^{th}$ element in X and let $h_s$ be the 0-1 vector that selects element $\Delta z$ from X*, $\Delta z_{t+j}=h_s \cdot X^*_{t+j}$. Then $E_t\Delta z_{t+j} =h_s \cdot (A^*)^j \cdot X^*_t$ and $E_t\Delta z_{t+j}-E_{t-1}\Delta z_{t+j} = h_s \cdot (A^*)^j \cdot u^*_t$. Thus, the elements of the vector $h_s \cdot (A^*)^j$ indicate the marginal impact of period-t innovations on $\Delta z_{t+j}$. Responses of variable z in levels can be obtained by summing up the responses of $\Delta z$.

[18] The identity is obtained by re-arranging the infinite sum (see Campbell (1987)). Equation (12) is obtained from (10) by multiplying by $(1-\rho)$ and then using the identity, the

fact that $(1-\rho)\cdot(1+r)=r$, and the fact that $\Delta\hat{X}_t=\hat{X}_t$ because $X_{t-1}$ is in the period t-1 information set.

[19] Note that neither an orthogonalization of innovations nor an assumption about the ordering of VAR-variables is needed to compute these coefficients.

[20] To be precise, the relation $f_G(T) - f_G(G) - r\cdot\omega_G = 1$ follows from inserting (13) into (12) for the particular innovation $\Delta\hat{X}_t=(0,1,\ldots,0)$.

[21] To be precise, (12) and (13) with $\Delta\hat{X}_t=(1,0,\ldots,0)$ imply $f_T(T) - f_T(G) - r\cdot\omega_T = -1$.

[22] In the empirical section, taxes will be scaled by GNP so that $T_t$ can be interpreted as average tax rate. In principle, marginal and average tax rates may move in different directions, if arbitrary government policies (e.g., including tax-reforms) are studied. But Barro's tax-smoothing model postulates an optimizing government, which will presumably rank different sources of revenue by their distortionary effects and turn to increasingly distortionary taxes as revenue needs rise. Then a monotonic relation between the average and marginal tax rate (the latter taken as index of marginal excess burden) should be expected.

[23] However, one has to assume either that government spending does not affect utility or that private and government spending enter separably into the utility function. Otherwise, changes in government spending may have direct effects on private consumption.

[24] For an analysis of government behavior, it seems best to use tax and spending concepts that capture as much of government activity as possible without netting out anything. For example, if taxation is distortionary, all revenue raising government activity imposes a burden on taxpayers, even if some revenues are returned to individuals—typically others—as transfers. The public policy issue is on which side of the budget adjustments are made in response to a deficit. On the other hand, a focus on spending on goods and services, i.e., on outlays and receipts net of transfers, may be preferable for studying the macroeconomic effects of government activity. Here I follow von Furstenberg et.al. (1986) in using

receipts and outlays. An analysis with the net concepts may be an interesting topic for future research.

[25] Annual GNP data for 1789-1973 have been compiled by Berry (1988). Since Berry uses NIPA data from 1929 on, I use his series for 1789-1928 (with upward adjustment of 0.81% to match revised NIPA data in 1929) and NIPA data thereafter.

[26] The series and a more detailed description of the data are in the appendix. Concerning the positive average of the primary deficit series, it should be noted that neither positive primary deficits nor the fact that economic growth has exceeded the interest rate on government bonds for long periods provide evidence against the inequality $r>0$ (see Bohn (1990b)). The reason is that risky tax revenues (which depend on the state of the economy) should likely be discounted at a higher than the safe rate. See Abel et.al. (1989) on the dynamic efficiency issues arising in this context and Bohn (1990b) for a more general discussion of discounting; Kremers (1989) argues similarly.

[27] An issue in this test is the size of the lag window. Since there is autocorrelation in the differenced data, it should not be chosen too small. I picked $k=10$ arbitrarily.

[28] As noted in Section 3, the parameter r should be interpreted as difference of real discount rate and growth rate. Since the choice of the discount rates is a non-trivial issue (see fn. 26 above), DEF was computed for several values of r. Fortunately, the results are almost identical for a wide range of values so that a precise determination was not needed.

[29] The scale can be normalized arbitrarily; here the first element was set to 1.0. Johansen's normalization (see Table 3A) is $\beta=(0.797,-0.691,-0.048)$.

[30] That is, tests of $H_0$: $r=r_0$ were executed as specified in Johansen (1988) for a grid (integer percentages) of $r_0$-values. Rejections were obtained for -2% and below and for 15% and higher.

[31] Engle and Yoo's simulations (of a somewhat different model) yield critical 1% (5%) values of 4.35 (3.78).

[32] Throughout, lag lengths were chosen by the Akaike information criterion; r=3% is the integer percentage closest to the estimated cointegrating vector.

[33] Since the spending response far exceeds the tax response in the first four periods, the higher permanent tax increase (0.30 as compared to 0.25 for spending) is consistent with the intertemporal budget constraint.

[34] Before rejecting tax-smoothing, one should recognize, however, that the tax series is some steps removed from the theoretical ideal of a marginal tax rate (as discussed above).

[35] Again, the numbers add up to about one and they match up with the impulse response graphs in Figure 4. The smaller persistent effects of spending as compared to taxes are consistent with the smaller long-term responses seen in Panel B of Figure 4 as compared to Panel A.

[36] After 1916, Federal reserve transfers to the Treasury were taken as measure of seignorage, since the Federal reserve system is off the federal budget. Before 1916, seignorage was estimated in two ways, as money stock times nominal interest rate and as change in the money stock. The first definition is appropriate if money is part of the debt, the second, if money is not considered a liability. The difference is non-trivial for 1860-1879. I generally chose the first version, but here I explore the alternative.

[37] The VAR-models seem to suffer from multicollinearity problems. If anything, taxes are less permanent and the effect of taxes on spending is larger than the corresponding ECM specification (No.1) suggested.

[38] The models are ECMs with 4 lags of each variable and cointegrating vector $\beta=(1,-1,0)$. Because fiscal and calendar years are not perfectly aligned, growth and inflation in the current fiscal year may or may not be considered known. Hence, two sets of regressions were run with each variable. Nominal GNP and the GNP-deflator (as measure of inflation) are from Berry (1988) and NIPA data (see Section 3). The deflator was adjusted by Barro's (1986) correction for World War II price controls.

# References

Abel, A., G. Mankiw, L. Summers and R. Zeckhauser, 1989, Assessing Dynamic Efficiency: Theory and Evidence, Review of Economic Studies 56, 1-20.

Ahmed, S. and B.S. Yoo, 1989, Fiscal Trends and Real Business Cycles, Working Paper, Pennsylvania State University.

Barro, R. J., 1974, Are Government Bonds Net Wealth? Journal of Political Economy 82, 1095-1117.

———, 1979, On the Determination of Public Debt, Journal of Political Economy 87, 940-971.

———, 1986, U.S. Deficits Since World War I, Scandinavian Journal of Economics 88, 195-222.

———, 1989, The Ricardian Approach to Budget Deficits, Journal of Economic Perspectives 3, 37-54.

Bernheim, D., 1987, Ricardian Equivalence: An Evaluation of Theory and Evidence, in: S. Fischer, ed., NBER Macroeconomics Annual 1987 (MIT Press, Cambridge) 263-304.

———, 1989, A Neoclassical Perspective on Budget Deficits, Journal of Economic Perspectives 3, 55-72.

Berry, T., 1988, Production and Population since 1789: Revised GNP Series in Constant Dollars, Bostwick Papers No.6 (The Bostwick Press, Richmond).

Blinder, A., 1981, Temporary Income Taxes and Consumer Spending, Journal of Political Economy 89, 26-53.

Board of Governors of the Federal Reserve System, 1943, Banking and Monetary Statistics (Federal Reserve, Washington).

Bohn, H., 1988, Endogenous Government Spending and Ricardian Equivalence, Working Paper, Rodney White Center for Financial Research (13-88), The Wharton School of the University of Pennsylvania.

———, 1990a, Tax-Smoothing with Financial Instruments, American Economic Review 80, 1217-1230.

———, 1990b, The Sustainability of Budget Deficits in a Stochastic Economy, Working Paper, Rodney White Center for Financial Research (6-90), The Wharton School of the University of Pennsylvania.

Campbell, J., 1987, Does Savings Anticipate Declining Labor Income? An Alternative Test of the Permanent Income Hypothesis, Econometrica 55, 1249-1273.

Dickey, D. and W. Fuller, 1981, Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root, Econometrica 49, 1057-1072.

Eisner, R., 1989, Budget Deficits: Rhetoric and Reality, Journal of Economic Perspectives 3, 73-93.

Engle, R. and C. Granger, 1987, Cointegration and Error-Correction: Representation, Estimation, and Testing, Econometrica 55, 251-276.

——— and B.S. Yoo, 1987, Forecasting and Testing in Cointegrated Systems, Journal of Econometrics 35, 143-159.

Feldstein, M., 1982, Government Deficits and Aggregate Demand, Journal of Monetary Economics 9, 1-20.

Firestone, J., 1960, Federal Receipts and Expenditures During Business Cycles, 1879-1958 (Princeton University Press, Princeton).

Fuller, W., 1976, Introduction to Statistical Time Series (Wiley, New York).

von Furstenberg, G., J. Green and J.-H. Jeong, 1986, Tax and Spend, or Spend and Tax? Review of Economics and Statistics 68, 179-188.

Gardner, G. and K. Kimbrough, 1989, The Behavior of U.S. Tariff Rates, American Economic Review 79, 211-218.

Gramlich, E., 1989, Budget Deficits and National Savings: Are Politicians Exogenous? Journal of Economic Perspectives 3, 37-54.

Hakkio, C. and M. Rush, 1986, Co-Integration and the Government's Budget Deficit, Working Paper, Federal Reserve Bank of Kansas City.

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

Hepburn, B., 1915, A History of Currency in the United States (MacMillan, New York).

Horrigan, B., 1986, The Long-Run Behavior of the Public Debt in the United States, Federal Reserve Bank of Philadelphia Working Paper 86-15.

Johansen, S., 1988, Statistical Analysis of Cointegration Vectors, Journal of Economic Dynamics and Control 12, 231-254.

Kremers, J., 1989, U.S. Federal Indebtedness and the Conduct of Fiscal Policy, Journal of Monetary Economics 23, 219-238.

McCallum, B., 1984, Are Bond-financed Deficits Inflationary? A Ricardian Analysis, Journal of Political Economy 92, 123-135.

Miller, S. and F. Russek, 1990, Co-Integration and Error-Correction Models: The Temporal Causality between Government Taxes and Spending, Southern Economic Journal 57,221-229.

Phillips, P., 1987, Time Series Regressions with Unit Root, Working Paper, Yale University.

_________ and P. Perron, 1986, Testing for a Unit Root in Time Series Regression, Working Paper, Yale University.

Sahasakul, C., 1987, The U.S.-Evidence of Optimal Taxation Over Time, Journal of Monetary Economics 18, 251-275.

Said, S. and D. Dickey, 1985, Hypothesis Testing in ARIMA(p,1,q) Models, Journal of the American Statistical Association 80, 369-374.

Sargent, T. and N. Wallace, 1982, Some Unpleasant Monetarist Arithmetic, Federal Reserve Bank of Minneapolis Quarterly Review 5, 1-17.

Schwert, W., 1987, Effects of Model Specification on Tests for Unit Roots in Macroeconomic Data, Journal of Monetary Economics 20, 73-103.

Tabellini, G. and A. Alesina, 1990, Voting on the Budget Deficit, American Economic Review 80, 37-49.

Trehan, B. and C. Walsh, 1988, Common Trends, The Government Budget Constraint, and Revenue Smoothing, Journal of Economic Dynamics and Control 12, 425-444.

United States Bureau of the Budget, Historical Tables, Budget of the United States Government, Fiscal Year 1990 (U.S. Government Printing Office, Washington).

United States Bureau of the Census, 1976, Historical Statistics of the United States, Colonial Times to 1970 (U.S. Department of Commerce, Bureau of the Census, Washington).

White, H., 1980, A Heteroskedasticity-Consistent Covariance Matrix Estimator and Direct Test for Heteroskedasticity, Econometrica 48, 817-838.

Wilcox, D., 1989, The Sustainability of Government Deficits: Implications of the Present-Value Borrowing Constraint, Journal of Money, Credit and Banking 21, 291-306.

# Table 1: Summary Statistics

| Series | Mean | Standard Deviation | Autocorrelations at lags: | | | | |
|---|---|---|---|---|---|---|---|
| | | | 1 | 2 | 3 | 4 | 5 |
| B | 21.446 | 21.19 | 0.977 | 0.930 | 0.877 | 0.826 | 0.783 |
| T | 7.083 | 7.16 | 0.977 | 0.949 | 0.926 | 0.901 | 0.870 |
| G | 7.490 | 9.12 | 0.927 | 0.806 | 0.695 | 0.618 | 0.580 |
| $\Delta B$ | 0.126 | 4.31 | 0.560 | 0.164 | -0.037 | -0.147 | -0.138 |
| $\Delta T$ | 0.096 | 1.18 | 0.154 | -0.133 | 0.030 | -0.093 | -0.177 |
| $\Delta G$ | 0.098 | 3.34 | 0.351 | -0.078 | -0.259 | -0.294 | -0.211 |
| $DEF^0$ | 0.407 | 4.37 | 0.738 | 0.414 | 0.155 | -0.033 | -0.107 |

Legend: The variables are B = federal debt, T = federal receipts, G = federal non-interest outlays, all measured as fractions of GNP. The $\Delta$'s refer to first differences; $DEF^0 = G - T$ is the primary deficit. The sample period is 1800-1988, the number of observation is 189.

## Table 2: Stationarity Tests

| Variable | k | ADF(k) | ADF(13) | $Z\tau$ | $Z\rho$ |
|---|---|---|---|---|---|
| T | 5 | -0.17 | -0.16 | -0.44 | -0.96 |
| G | 6 | -1.49 | -0.80 | -2.11 | -9.39 |
| B | 2 | -1.95 | -2.12 | -1.74 | -6.34 |
| $\Delta T$ | 0 | -11.72** | -3.89** | -11.57** | -129.79** |
| $\Delta G$ | 0 | -9.48** | -5.06** | -9.01** | -66.35** |
| $\Delta B$ | 0 | -7.27** | -3.27* | -6.45** | -57.32** |
| $DEF^0$ | 5 | -4.69** | -3.50* | -4.78** | -38.57** |
| $DEF^2$ | 5 | -4.41** | -3.21* | -4.71** | -37.99** |
| $DEF^4$ | 5 | -4.11** | -2.94* | -4.59** | -36.79** |

Legend: The variables are B = federal debt, T = federal receipts, G = federal non-interest outlays, all measured as fractions of GNP. The $\Delta$'s refer to first differences. The variables $DEF^r = G - T + r \cdot B$ refer to federal deficits computed for different interest rates r. The sample period is 1800-1988, the number of observation is 189.

The test statistics are ADF(k) = the augmented Dickey-Fuller t-statistic computed with lag length k (where k was either selected by the Akaike information criterion or fixed at k=13), $Z\tau$ = Phillips' adjustment of the Dickey-Fuller t-statistic, and $Z\rho$ = Phillips' adjustment of the coefficient estimate in the Dickey-Fuller regression multiplied by the sample size.

The symbols * and ** indicate significant rejections of non-stationarity at the 5% and 1% levels, respectively. Critical values for ADF(·) and $Z\tau$ at the 5% level are 2.88 and 2.89 for N=100 and 250 observations, respectively, or 3.46 and 3.51 at the 1% level, respectively. Critical values for $Z\rho$ at the 5% level are 13.7 and 14.0 for N=100 and 250 observations, respectively, or 19.8 and 20.3 at the 1% level, respectively. Here N=189, so that the critical values are in between those for 100 and 250.

## Table 3: Co-Integration Tests

**Panel A:** Test on vector $X_t = (T_t, G_t, B_t)$, Sample 1800-1988, Regressions with 6 lags

Squared canonical correlations between X and $\Delta X$:

| | | |
|---|---|---|
| 0.1266 | 0.0492 | 0.0029 |

Canonical variates (in columns):

| | | |
|---|---|---|
| 0.787 | -0.710 | 0.490 |
| -0.691 | 0.781 | -0.268 |
| -0.048 | -0.078 | -0.026 |

Maximum number of co-integrating vectors, p:

| p | Test value | 95% interval |
|---|---|---|
| 2 | 0.542 | [ 0.0, 5.3] |
| 1 | 10.081 | [ 1.0,13.9] |
| 0 | 35.666* | [ 7.0,26.1] |

| | | |
|---|---|---|
| Restriction to (1.0,-1.0,-r): | $\chi^2(1)=3.67$ | Significance level = 5.55% |
| Restriction to (1.0,-1.0,0): | $\chi^2(2)=4.77$ | Significance level = 9.21% |

## Table 3 continued:

**Panel B**: Test on vector $X_t = (T_t,G_t)$, Sample 1800-1988, Regressions with 4 lags

Squared canonical correlations between X and $\Delta$X:

| | |
|---|---|
| 0.2201 | 0.0001 |

Canonical variates (in columns):

| | |
|---|---|
| 0.520 | 0.233 |
| -0.514 | -0.082 |

Maximum number of co-integrating vectors, p:

| p | Test value | 95% interval |
|---|---|---|
| 1 | 0.023 | [ 0.0, 5.3] |
| 0 | 47.016* | [ 1.0,13.9] |

Restriction to (1.0,-1.0): $\chi^2(1)=0.07$ Significance level = 79%

Legend: The variables are B = federal debt, T = federal receipts, G = federal non-interest outlays, all measured as fractions of GNP. Maximum likelihood estimates of canonical correlations and canonical variates as well as the test statistics are computed as in Johansen (1988). A * indicates a rejection of the hypothesis that the number of cointegrating vectors is less or equal p at the 95% confidence level. The $\chi^2$-statistic under "restriction to" tests the null hypothesis that the cointegrating vector can be written as specified by the restriction.

## Table 4: An Error Correction Model

### Panel A: Estimates

| Regressor Variable and lag | Equation 1 | | | Equation 2 | | | Equation 3 | | |
|---|---|---|---|---|---|---|---|---|---|
| | Dependent Variable: $\Delta T$ | | | Dependent Variable: $\Delta G$ | | | Dependent Variable: $\Delta B$ | | |
| | Coef. | t-value | Robust t-value | Coef. | t-value | Robust t-value | Coef. | t-value | Robust t-value |
| $\Delta T$ 1 | -0.369 | -4.702 | -3.634 | 1.200 | 4.592 | 3.385 | -1.178 | -6.087 | -4.124 |
| $\Delta T$ 2 | -0.790 | -8.229 | -6.631 | -0.488 | -1.527 | -1.535 | 0.034 | 0.143 | 0.126 |
| $\Delta T$ 3 | -0.228 | -2.270 | -2.026 | 0.431 | 1.292 | 1.545 | -0.770 | -3.111 | -3.277 |
| $\Delta T$ 4 | -0.284 | -3.089 | -2.814 | -0.136 | -0.444 | -0.562 | -0.546 | -2.408 | -1.892 |
| $\Delta G$ 1 | 0.302 | 12.575 | 5.573 | 0.215 | 2.692 | 1.361 | 0.743 | 12.535 | 8.828 |
| $\Delta G$ 2 | 0.206 | 5.101 | 3.424 | -0.335 | -2.488 | -1.649 | 0.604 | 6.069 | 5.597 |
| $\Delta G$ 3 | 0.187 | 4.000 | 3.421 | -0.056 | -0.362 | -0.322 | 0.518 | 4.500 | 3.398 |
| $\Delta G$ 4 | 0.187 | 3.827 | 3.211 | -0.203 | -1.251 | -0.859 | 0.626 | 5.209 | 3.761 |
| $\Delta B$ 1 | -0.028 | -0.968 | -0.749 | -0.213 | -2.176 | -1.645 | 0.378 | 5.227 | 4.466 |
| $\Delta B$ 2 | 0.039 | 1.244 | 1.437 | 0.012 | 0.110 | 0.144 | -0.107 | -1.372 | -1.040 |
| $\Delta B$ 3 | -0.004 | -0.117 | -0.102 | 0.053 | 0.533 | 0.462 | 0.052 | 0.693 | 0.549 |
| $\Delta B$ 4 | -0.024 | -0.864 | -0.736 | -0.004 | -0.049 | -0.036 | 0.224 | 3.297 | 2.294 |
| $DEF^3$ | 0.170 | 4.825 | 3.301 | -0.192 | -1.663 | -1.789 | 0.224 | 2.581 | 2.017 |
| Constant | -0.010 | -0.152 | -0.185 | 0.247 | 1.146 | 1.507 | -0.181 | -1.133 | -1.289 |
| Std. Error | | 0.807 | | | 2.682 | | | 1.986 | |
| $R^2$ | | 0.565 | | | 0.398 | | | 0.803 | |
| Adjusted $R^2$ | | 0.533 | | | 0.353 | | | 0.788 | |
| F-test on $\Delta T$ | | 0.000 | | | 0.000 | | | 0.000 | |
| F-test on $\Delta G$ | | 0.000 | | | 0.009 | | | 0.000 | |
| F-test on $\Delta B$ | | 0.562 | | | 0.237 | | | 0.000 | |

## Table 4 continued:

**Panel B**: Projections

| Innovation in:<br>Effect on: | T<br>PV(ΔT) | T<br>PV(ΔG) | G<br>PV(ΔT) | G<br>PV(ΔG) |
|---|---|---|---|---|
| Estimate | -0.4889 | 0.4997 | 0.3036 | -0.6444 |
| Standard error | 0.1230 | 0.1421 | 0.1679 | 0.1940 |
| Robust std. error | 0.1176 | 0.1419 | 0.2669 | 0.3125 |
| T-value | -3.974 | 3.515 | 1.808 | -3.322 |
| Robust t-value | -4.157 | 3.522 | 1.138 | -2.062 |

Legend: Variables are as described in Table 2 (with $DEF^3$ denoting the with-interest deficit based on an interest rate of r=3%). The error correction model is estimated by Ordinary Least Squares for the sample 1800-1988 (189 observations). Robust t-statistics and standard errors indicate heteroskedasticity-consistent estimates (see White (1980)). The F-tests indicates at what significance level the variable (all four lags) can be excluded from the regression. The standard errors of the projections are only valid asymptotically.

## Table 5: Summary of Projections based on other Regressions

| Model | Projections of: T on PV(ΔT) | T on PV(ΔG) | G on PV(ΔT) | G on PV(ΔG) |
|---|---|---|---|---|
| 1 | -0.49 (-4.157) | 0.50 ( 3.522) | 0.30 ( 1.138) | -0.64 (-2.062) |
| 2 | -0.39 (-5.063) | 0.58 ( 6.803) | 0.30 ( 1.470) | -0.66 (-2.903) |
| 3 | -0.45 (-1.955) | 0.56 ( 2.023) | 0.33 ( 1.314) | -0.61 (-2.026) |
| 4 | -0.49 (-3.539) | 0.51 ( 3.647) | 0.29 ( 0.955) | -0.71 (-2.281) |
| 5 | -0.48 (-4.872) | 0.47 ( 3.250) | 0.31 ( 1.352) | -0.56 (-1.771) |
| 6 | -0.50 (-3.699) | 0.49 ( 3.001) | 0.29 ( 1.242) | -0.66 (-2.367) |
| 7 | -0.50 (-4.070) | 0.48 ( 3.292) | 0.25 ( 0.660) | -0.72 (-1.594) |
| 8 | -0.39 (-0.165) | 0.58 ( 0.224) | 0.30 ( 0.063) | -0.66 (-0.126) |
| 9 | -0.51 (-6.965) | NA | 0.29 ( 0.756) | NA |
| 10 | NA | 0.56 ( 1.005) | NA | -0.22 (-0.095) |

Legend: The variables are as described in Table 2. Model numbers refer to different sets of regressions. Robust t-statistics are in paretheses. Model No.1 is the error correction model (ECM) described in Table 4 (estimated with 4 lags, $DEF^r$ computed with r=0.03). The other models differ from No.1 as follows: No.2 and No.3 are estimated with 2 lags and 6 lags, respectively (and r=0.03). No.4 and No.5 use $DEF^r$ with r = 0.005 and r=0.06, respectively (and 4 lags). No.6 uses a different definition of seignorage revenue in computing T, as described in the text. No.7 assumes that end-of-period debt $B_{t+1}$ is in the information set, i.e. replaces $\Delta B_t$ by $\Delta B_{t+1}$ and computes $DEF^r$ using $B_{t+1}$. Models No.8-10 are Vector Autoregressions (4 lags, using r=0.03) of the type described in equation (8). No.8 uses the variable vector ($\Delta T$, $\Delta G$, $DEF^r$), No.9 uses ($\Delta T$, $\Delta B$, $DEF^r$), and No.10 uses ($\Delta G$, $\Delta B$, $DEF^r$). For models No.9 and 10, projections of the excluded variables ($\Delta G$ and $\Delta T$, respectively) cannot be computed (NA).

## Table 6: Projections based on the Limiting Case r→0

| Model | Projections of: | | | |
|---|---|---|---|---|
| | T on PV(ΔT) | T on PV(ΔG) | G on PV(ΔT) | G on PV(ΔG) |
| 1 | -0.35 (-4.113) | 0.65 ( 7.680) | 0.28 ( 2.328) | -0.72 (-6.000) |
| 2 | -0.35 (-4.249) | 0.65 ( 7.928) | 0.26 ( 3.408) | -0.74 (-9.549) |
| 3 | -0.37 (-1.521) | 0.63 ( 2.630) | 0.32 ( 0.390) | -0.68 (-0.835) |
| 4 | -0.40 (-4.748) | 0.60 ( 7.269) | 0.28 ( 1.973) | -0.72 (-4.991) |
| 5 | -0.34 (-3.545) | 0.66 ( 7.004) | 0.28 ( 1.918) | -0.72 (-4.957) |
| 6 | -0.38 (-3.026) | 0.62 ( 4.920) | 0.26 ( 2.083) | -0.74 (-5.893) |
| 7 | -0.43 (-5.590) | 0.57 ( 7.389) | 0.32 ( 1.870) | -0.68 (-3.995) |
| 8 | -0.44 (-3.788) | 0.56 ( 4.865) | 0.35 ( 2.239) | -0.65 (-4.249) |
| 9 | -0.33 (-2.946) | 0.67 ( 6.103) | 0.24 ( 2.104) | -0.76 (-6.584) |
| 10 | -0.33 (-2.901) | 0.67 ( 5.989) | 0.26 ( 2.059) | -0.74 (-5.811) |
| 11 | -0.29 (-3.353) | 0.71 ( 8.284) | 0.20 ( 3.174) | -0.80 (-12.523) |
| 12 | -0.37 (-4.976) | 0.63 ( 8.310) | 0.27 ( 2.138) | -0.73 (-5.730) |

Legend: The variables are as described in Table 2. Model numbers refer to different sets of regressions. Unless specifically noted below, all models are estimated by Ordinary Least Squares with 2 lags over the sample period 1800-1988 (189 observations).

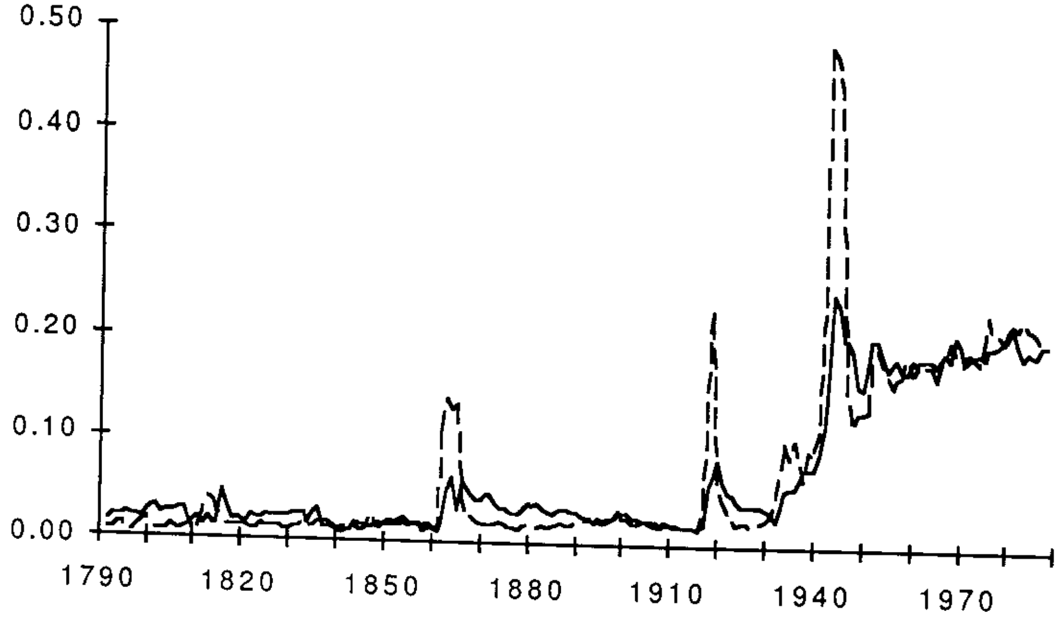
Model No.1 is an Error Correction Model (ECM) with variable vector ΔX = (ΔT,ΔG) and cointegrating variable $DEF^0$=G-T. Models No.2 and No.3 are Vector Autoregressions with variable vectors ΔX = (ΔT,$DEF^0$) and ΔX = (ΔT,$DEF^0$), respectively, where No.3 is estimated with 6 lags.

Models No. 4-7 are ECMs identical to model No.1, except that an additional variable is included into the vector X: No.4 adds current GNP-growth, No.5 adds lagged GNP-growth, No.6 adds current inflation, and No.7 adds lagged inflation as regressor to both ECM equations, each with 2 of its lagged values.

Models No.8-11 are ECMs identical to model No.1, except for different sample periods: No.8 is estimated for 1800-1917 (118 observations), No.9 for 1918-1988 (71 observations), No.10 for 1879-1988 (110 observations) and No.11 for the full 1800-1988 sample period, but excluding years of major wars (1861-1866, 1917-1919, 1941-1947), i.e. using 173 observations.

Model No.12 is an ECM identical to model No.1, except that the variables T and G are defined as federal revenues and federal non-interest outlays deflated by a value for fiscal year GNP that is obtained by interpolating calendar year GNP.

**Figure 1:** Federal Receipts and Non-interest Outlays as Fractions of GNP



Legend:

---- Outlays

—— Receipts

**Figure 2:** Public Debt as Fraction of GNP

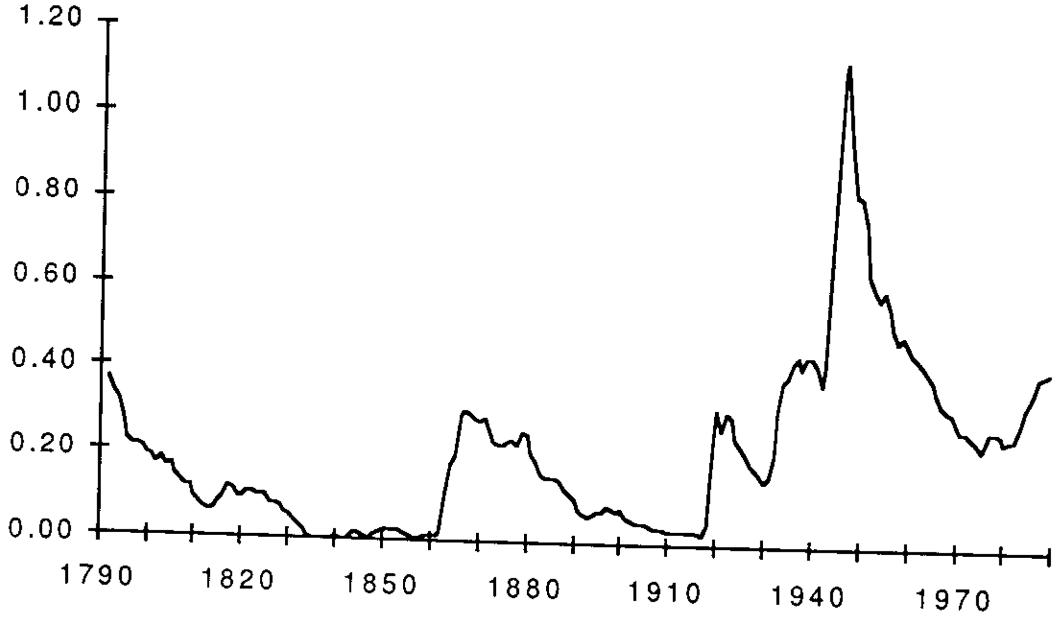
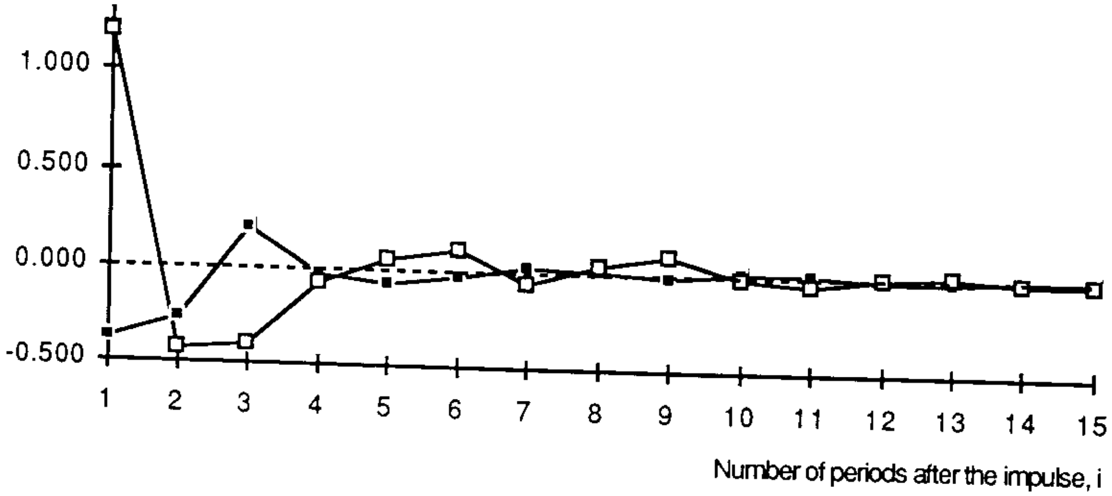
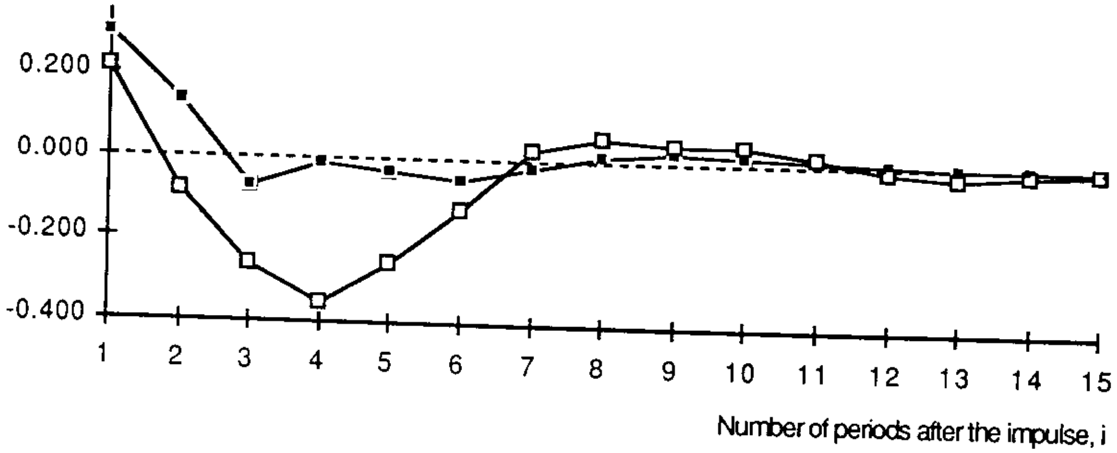


**Figure 3:** Impulse Response Functions - Differences

**Panel A:** Unit shock to receipts, responses of ΔT and ΔG after i periods:



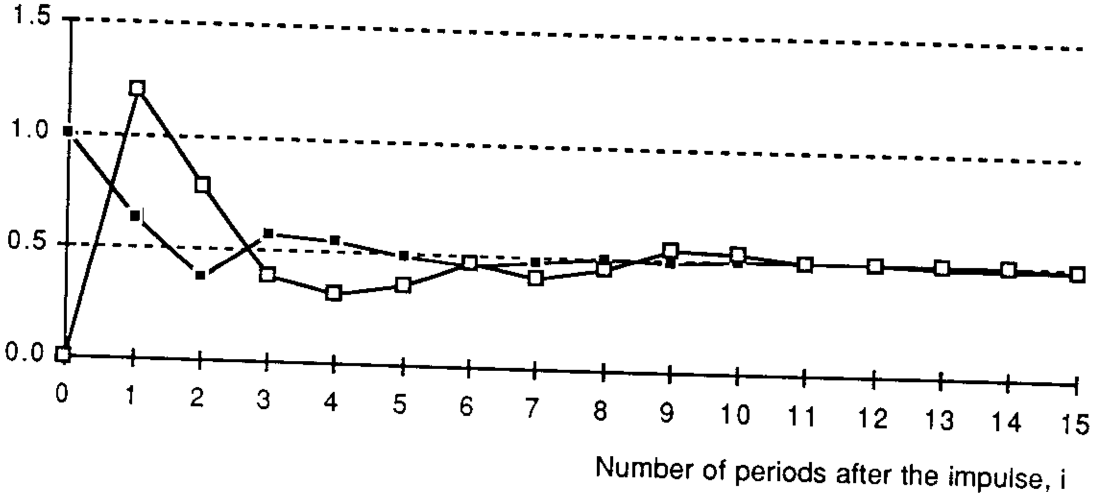
**Panel B:** Unit shock to non-interest outlays, responses of ΔT and ΔG after i periods:



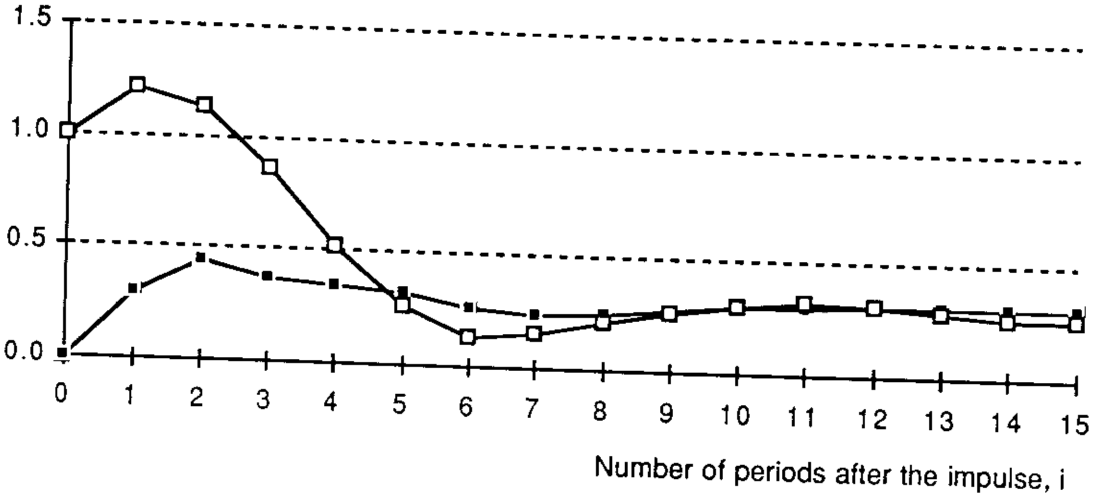
Legend: The graphs shows the responses in series ΔT, the first difference of federal receipts as a fraction of GNP, and series ΔG, the first difference of federal non-interest outlays as a fraction of GNP, after an impulse to receipts and outlays, respectively. Responses of ΔT are marked by -■-, responses of ΔG are marked by -□-. All impulse-response functions are based on estimates from the error correction model of Table 4.

**Figure 4**: Impulse Response Functions - Levels

**Panel A:** Unit shock to taxes, responses of T and G after i periods:



**Panel B:** Unit shock to spending, responses of T and G after i periods:



Legend: The graphs shows the responses in series T, the ratio of federal receipts to GNP, and series G, the ratio of federal non-interest outlays to GNP, after an impulse to receipts and outlays, respectively. Responses of T are marked by -■-, responses of G are marked by -□-. All impulse-response functions are based on estimates from the error correction model of Table 4.

# U.S. Budget Data from 1792-1988

## Appendix to:

## Budget Balance through Revenue or Spending Adjustments? Some Historical Evidence for the United States

Henning Bohn

Department of Finance
The Wharton School
University of Pennsylvania
Philadelphia, PA 19104-6367

The basic sources for U.S. budget data are the Historical Statistics of the United States, Colonial Times to 1970 and the Historical Tables, Budget of the United States Government, Fiscal Year 1990. The Historical Statistics have data up to 1970. The Historical Tables have data for 1940-1988. The series used in the empirical analysis were constructed as follows.

### A.1. The series for revenues, $T_t$:

Components are ordinary receipts and seignorage.

**Series (1), total government receipts**

1792-1963: Historical Statistics Series Y352, Y339, Y343.

1963-1988: U.S. Government Budget.

For the late 1960s, where the two sources show slight differences, the more recent source was used.

**Series (2), Seignorage**

While it is clear that seignorage must be considered somehow, there are several ways to proceed. (Impatient readers may skip the section on seignorage. The definition of seignorage is conceptually tricky, but the issue is quantitatively absolutely insignificant.)

The first choice involves the Federal Reserve system. I follow the unified budget definitions in leaving the Federal Reserve off the balance sheet. Thus, since Federal reserve payments to the Treasury are included in total receipts (series (1)), the bulk of seignorage since 1916 is already accounted for.[1] But seignorage was significant in the pre-Federal reserve period, in particular during and after the Civil War.

---

[1] Seignorage on Treasury money since 1917 is so small--probably less than the rounding errors--that it was omitted.

The second issue concerns the appropriate definition of government liabilities. To be specific, consider the budget equation with money, comparable to equation (1) in the paper:

$$\bar{B}^n_{t+1} + M^n_{t+1} = G^n_t - \bar{T}^n_t + (1 + r^n) \cdot \bar{B}^n_t + M^n_t + \varepsilon^n_{t+1} \, , \qquad (1')$$

where M = money, $\bar{B}$ = bonds, $\bar{T}$ = receipts excluding seignorage, and n indicates nominal variables. It leads to an intertemporal budget constraint (in GNP-shares) either of the form

$$(1 + r) \cdot \bar{B}_t = \sum_{j\geq 0} \rho^j \cdot E_t\left[\bar{T}_{t+j} - G_{t+j} - \varepsilon_{t+j+1} - (\Delta M^n_t)/Y_t\right] \qquad (2')$$

(with $Y_t$ = GNP) or of the form

$$(1 + r) \cdot (\bar{B}_t + M_{t+1}) = \sum_{j\geq 0} \rho^j \cdot E_t\left[\bar{T}_{t+j} - G_{t+j} - \varepsilon_{t+j+1} - r^n \cdot M_t\right] \, . \qquad (2'')$$

In (2'), money creation is immediately counted as net gain without obligation. Total receipts must be measured as $T_t = \bar{T}_t + (\Delta M^n_t)/Y_t$. The corresponding definition of debt (see below) includes only interest-bearing debt, $B_t = \bar{B}_t$. In (2"), money is a liability, leading to a definition of debt as bonds plus money, $B_t = \bar{B}_t + M_t$. Only the non-payment of nominal interest that would have been paid on bonds is considered a gain, leading to a measure of total receipts $T_t = \bar{T}_t + r^n \cdot M_t$. The choice between the two sets of definitions is a matter of interpretation. Over a long sample, the two measures of seignorage yield about the same receipts in present value terms, but the timing of receipts is radically different.

Here the period at issue is 1862-1917, since the Federal Reserve system has been excluded. Large amounts of paper money (U.S. notes, greenbacks) were printed during the Civil War and remained outstanding in roughly constant nominal quantity since then ($350-400 million). The question is whether to account for the seignorage in the 1860's or to recognize the flow of "interest savings" over the entire period 1862-1917.

Before 1933, government money was either convertible to gold or an eventual resumption of convertibility was expected. Money was clearly a liability. Definitions motivated by (2") will therefore be the preferred concepts for this study. Still, since there are different opinions on this issue,[2] I also derive series based on the alternative method. In detail, the series are as follows:

(2a) <u>Seignorage according to equation (2")</u>: Seignorage is the product of the following two series minus an adjustment:

(2a-i) The nominal interest rate, defined as yield on government debt, is computed as the ratio of interest charge (Historical Statistics Series Y498) and interest-bearing government debt (Series Y497).

(2a-ii) The value of non-interest bearing government debt, Historical Statistics Series Y496, with 2 corrections:

1862-1878: For these years, greenbacks were not convertible into gold and their value was far below par. The par value was therefore multiplied by the average value of greenbacks in gold, from Hepburn (1915).

1880-1899: The Gold Standard Act of 1900 specified that a gold reserve had to be held against U.S. notes. Since 1900, the series shows only the net amount. No reserves were specified for the period before 1900, leading to a break in the published series in

---

[2] For a sample period starting in 1890, Trehan and Walsh (1988) measure seignorage as change in high powered money, which is closer to (2'). While this may be appropriate for the period of fiat money (they have the Federal Reserve on-budget), it probably overestimates for the early part of the sample, since high powered money included gold and currency backed by gold.

1900. However, maintaining some non-interest bearing reserves seems to be a necessary cost of earning seignorage in a period with convertibility. Thus, I assumed that the reserves (a total of $150 million in 1900) were built up continuously since resumption in 1879 (over 20 year) in equal increments of $7.5 million. Deduction of these increasing reserves leads to an uninterrupted adjusted series.

(2a-iii) If gold reserves were viewed as a necessary condition of earning seignorage, the cost of its buildup should be charged against seignorage. Thus, the product of series (2a-i) and (2a-ii) was reduced by $7.5 million annually for 1880-1899 in computing the final series for seignorage.

(2b) Seignorage according to equation (2'): The change in non-interest bearing debt (Series Y496) was taken as measure of seignorage. The change in U.S. notes was negligible outside the period 1862-78 so that the series was only computed for this period and set to zero otherwise.

The preferred final series for $T_t$ is then the sum of series (1) and (2a), deflated by GNP as described in the text. It is used unless specifically noted. The alternative is the sum of series (1) and (2b), deflated by GNP, which is used for model No. 6 in Table 5.

## A.2. The series for government spending, $G_t$:

Components are total outlays and interest outlays, which must be deducted from total outlays, before deflating by GNP.

**Series (3), total government outlays**

1792-1963: Historical Statistics Series Y457. The values for FY 1888-1891 are higher than those in Hepburn (1915) and Firestone (1960). Therefore these 4 observations were replaced by Hepburn's data.

1963-1988: U.S. Government Budget.

For the late 1960s, where the two sources show slight differences, the more recent source was used.

**Series (4), interest outlays**

1789-1915: Historical Statistics Series Y461 (interest payments).

1916-1940: Historical Statistics Series Y461, adjusted by the ratio of total interest-bearing debt to publicly held interest-bearing debt, from Federal Reserve Banking and Monetary Statistics, Table No. 149.

1940-1988: U.S. Government Budget, net interest payments. In principle, the two observations for 1940 could differ, if the yield on total interest-bearing debt, which includes trust fund holdings, differed from the yield on publicly held interest-bearing debt. In fact, the values coincide with 3 digit accuracy.

## A.3. The series for government debt, $B_t$:

Consistent with the definition of seignorage (series (2a) above), gross public debt, which includes private and Federal reserve holdings, but not government debt in government accounts, was used in general. For regressions using the alternative definition of seignorage (series (2b) above), non-interest-bearing debt (series (2a-ii) above) was subtracted.

The definition of the debt variable is problematic. Even excluding issues related to government real investment and asset accumulation (see Eisner (1989)), one should deduct cash balances and monetary gold reserves held by the Treasury from debt. For the period from 1916 on, the exclusion of

Federal Reserve accounts takes care of some gold holdings; and Treasury cash is generally a small fraction of debt. Before 1916, the issue is much more complex and unsettled, involving not only gold but also silver. Since consistent data are not available, changes in gold stock and Treasury cash were generally excluded (with one exception: the gold reserve against U.S. notes, see series (2a-ii) above). These missing variables are implicitly reflected in the error $\varepsilon_t$. Fortunately, if the statistical relevance of $\varepsilon_t$ provides a measure of how serious the problem is, it does not seem to be serious.

As noted in the text, interest-bearing debt is measured at par, due to the lack of consistent data on market values and because of the higher visibility of par values in government statistics. The value of "greenbacks" has been adjusted to prevent gross errors in measurement. The timing convention in published data is to use end of period value. The variable $B_t$ uses the lagged values to obtain beginning of period values. The series was constructed as follows.

**Series (5), publicly held gross debt at the end of a fiscal year**

1792-1915: Historical Statistics, Series Y493 (total debt) with three adjustments:

a. Data for 1879-1890 are much higher than those in Hepburn (1915) with an obvious outlier in 1879. Hepburn's data were used for these 12 years.

b. For 1892-1878, the value of the non-interest bearing component, Series Y496 (largely "greenbacks") is overstated. Total debt is adjusted downwards by the difference of Series Y496 and my series (2a-ii). That is, total debt is the par value of interest bearing plus matured debt and series (2a-ii).

c. For 1880-1899, the component Series Y496 was similarly replaced by series (2a-ii), which reflects some buildup in gold reserves.

1916-1940: Federal Reserve Banking and Monetary Statistics, sum of interest bearing debt held privately, interest bearing debt held by the Federal Reserve (both Table 149), and non-interest bearing debt (Table 146), with an adjustment described below.

1940-1988: U.S. Government Budget, gross debt publicly held.

While Historical Statistics and Federal Reserve data match in 1916, the debt series has a break in 1940, where the new series (privately held gross debt) exceeds the historical series by 1.452 billion dollar or about 3.5%. Looking at the components, I suspect that the new series includes additional publicly held agency and federally guaranteed securities. Federal Reserve data on such securities are available since 1933 (where they start at zero), but they add up to a smaller amount in 1940. For 1933-1939, I assume that the new definition would have lead to an upward revision by the same percentage as in 1940 (14.8%), and revised the series accordingly. Again, this approximation involves only a small fraction of gross debt.

**Other remarks**

The GNP-data are from Berry (1988) and the Commerce department, as described in the text.

For fiscal years 1843 and 1976, which do not have 12 months, revenues and spending have been annualized, while debt is the value at the end (or with lag, start) of these fiscal years. To obtain an annualized series for the change in debt $\Delta B$, the values for 1843 and 1976 must therefore be multiplied by 2.0 and 0.8, respectively. The final series for revenues, spending, debt, and the annualized change in debt are printed in table A1 below.

**Table A1: Data**

| | $T_t$ | $G_t$ | $B_{t+1}$ | $\Delta B_{t+1}$ |
|---|---|---|---|---|
| 1792 | 1.767 | 0.905 | 34.187 | -3.132 |
| 1793 | 2.059 | 0.744 | 31.862 | -2.325 |
| 1794 | 2.155 | 1.385 | 26.910 | -4.952 |
| 1795 | 2.097 | 1.490 | 22.872 | -4.038 |
| 1796 | 2.375 | 0.717 | 21.260 | -1.612 |
| 1797 | 2.205 | 0.718 | 21.659 | 0.400 |
| 1798 | 2.084 | 1.235 | 20.566 | -1.093 |
| 1799 | 1.961 | 1.680 | 18.876 | -1.690 |
| 1800 | 2.612 | 1.782 | 19.147 | 0.271 |
| 1801 | 2.983 | 1.151 | 17.153 | -1.994 |
| 1802 | 3.100 | 0.777 | 18.152 | 1.000 |
| 1803 | 2.488 | 0.911 | 16.677 | -1.475 |
| 1804 | 2.562 | 0.957 | 17.039 | 0.362 |
| 1805 | 2.674 | 1.262 | 14.449 | -2.590 |
| 1806 | 2.732 | 1.071 | 13.361 | -1.088 |
| 1807 | 2.895 | 0.874 | 12.236 | -1.126 |
| 1808 | 3.017 | 1.155 | 12.203 | -0.033 |
| 1809 | 1.454 | 1.385 | 9.439 | -2.764 |
| 1810 | 1.553 | 0.888 | 8.119 | -1.321 |
| 1811 | 2.201 | 0.849 | 7.203 | -0.915 |
| 1812 | 1.471 | 2.686 | 6.526 | -0.677 |
| 1813 | 2.071 | 4.057 | 6.774 | 0.248 |
| 1814 | 1.352 | 3.641 | 8.395 | 1.621 |
| 1815 | 1.620 | 2.781 | 9.782 | 1.387 |
| 1816 | 4.674 | 2.294 | 12.224 | 2.442 |
| 1817 | 3.178 | 1.479 | 11.612 | -0.612 |
| 1818 | 2.029 | 1.298 | 9.568 | -2.044 |
| 1819 | 2.274 | 1.507 | 9.888 | 0.320 |
| 1820 | 1.851 | 1.367 | 10.733 | 0.845 |
| 1821 | 1.719 | 1.262 | 11.076 | 0.343 |
| 1822 | 2.490 | 1.206 | 10.433 | -0.644 |
| 1823 | 2.292 | 1.093 | 10.329 | -0.104 |
| 1824 | 2.202 | 1.738 | 10.008 | -0.320 |
| 1825 | 2.423 | 1.265 | 8.198 | -1.811 |
| 1826 | 2.471 | 1.272 | 8.309 | 0.111 |
| 1827 | 2.356 | 1.293 | 7.490 | -0.819 |
| 1828 | 2.506 | 1.346 | 6.590 | -0.900 |
| 1829 | 2.424 | 1.240 | 5.538 | -1.052 |
| 1830 | 2.356 | 1.252 | 4.578 | -0.960 |
| 1831 | 2.688 | 1.309 | 3.445 | -1.134 |
| 1832 | 2.808 | 1.456 | 1.987 | -1.457 |
| 1833 | 2.776 | 1.856 | 0.529 | -1.458 |
| 1834 | 1.647 | 1.391 | 0.357 | -0.172 |
| 1835 | 2.637 | 1.305 | 0.002 | -0.355 |
| 1836 | 3.203 | 1.935 | 0.002 | 0.000 |
| 1837 | 1.362 | 2.031 | 0.019 | 0.017 |
| 1838 | 1.465 | 1.887 | 0.182 | 0.163 |
| 1839 | 1.735 | 1.460 | 0.531 | 0.349 |

Legend: All values are percentages of GNP.

## Table A1: Data *Continued*

| | $T_t$ | $G_t$ | $B_{t+1}$ | $\Delta B_{t+1}$ |
|---|---|---|---|---|
| 1840 | 0.994 | 1.231 | 0.207 | -0.324 |
| 1841 | 0.970 | 1.513 | 0.293 | 0.086 |
| 1842 | 1.127 | 1.377 | 0.825 | 0.531 |
| 1843 | 1.007 | 1.343 | 2.011 | 2.373 |
| 1844 | 1.803 | 1.261 | 1.342 | -0.669 |
| 1845 | 1.712 | 1.251 | 0.832 | -0.510 |
| 1846 | 1.554 | 1.410 | 0.761 | -0.071 |
| 1847 | 1.301 | 2.760 | 1.659 | 0.898 |
| 1848 | 1.528 | 1.839 | 2.179 | 0.520 |
| 1849 | 1.447 | 1.919 | 2.765 | 0.586 |
| 1850 | 1.910 | 1.564 | 2.460 | -0.304 |
| 1851 | 2.040 | 1.708 | 2.616 | 0.155 |
| 1852 | 1.909 | 1.540 | 2.424 | -0.192 |
| 1853 | 2.255 | 1.629 | 1.944 | -0.480 |
| 1854 | 2.399 | 1.784 | 1.251 | -0.693 |
| 1855 | 1.937 | 1.701 | 0.967 | -0.284 |
| 1856 | 2.011 | 1.838 | 0.835 | -0.132 |
| 1857 | 1.799 | 1.725 | 0.693 | -0.142 |
| 1858 | 1.126 | 1.753 | 1.206 | 0.514 |
| 1859 | 1.437 | 1.787 | 1.469 | 0.263 |
| 1860 | 1.408 | 1.505 | 1.617 | 0.148 |
| 1861 | 1.036 | 1.560 | 2.167 | 0.550 |
| 1862 | 1.916 | 11.041 | 10.766 | 8.599 |
| 1863 | 5.066 | 14.171 | 18.000 | 7.234 |
| 1864 | 6.340 | 13.046 | 20.213 | 2.213 |
| 1865 | 3.314 | 13.582 | 29.594 | 9.381 |
| 1866 | 6.001 | 4.285 | 30.680 | 1.086 |
| 1867 | 5.619 | 2.580 | 30.106 | -0.575 |
| 1868 | 4.845 | 2.691 | 28.784 | -1.322 |
| 1869 | 4.226 | 2.141 | 28.357 | -0.427 |
| 1870 | 4.344 | 2.010 | 28.799 | 0.442 |
| 1871 | 4.713 | 1.969 | 27.356 | -1.444 |
| 1872 | 4.712 | 1.887 | 23.486 | -3.870 |
| 1873 | 3.875 | 1.972 | 22.612 | -0.874 |
| 1874 | 3.414 | 2.055 | 22.967 | 0.355 |
| 1875 | 3.472 | 1.824 | 23.201 | 0.234 |
| 1876 | 3.300 | 1.774 | 23.702 | 0.500 |
| 1877 | 3.079 | 1.604 | 22.959 | -0.743 |
| 1878 | 2.892 | 1.465 | 25.665 | 2.707 |
| 1879 | 3.422 | 1.921 | 25.345 | -0.320 |
| 1880 | 3.959 | 1.981 | 20.984 | -4.361 |
| 1881 | 3.721 | 1.788 | 19.039 | -1.946 |
| 1882 | 3.876 | 1.762 | 16.077 | -2.961 |
| 1883 | 3.506 | 1.785 | 15.426 | -0.651 |
| 1884 | 3.189 | 1.698 | 15.039 | -0.386 |
| 1885 | 3.056 | 1.932 | 15.078 | 0.039 |
| 1886 | 3.285 | 1.830 | 14.620 | -0.457 |
| 1887 | 3.575 | 2.069 | 12.667 | -1.953 |
| 1888 | 3.352 | 1.862 | 11.304 | -1.364 |
| 1889 | 3.220 | 1.967 | 9.924 | -1.380 |

Legend: All values are percentages of GNP.

**Table A1: Data *Continued***

| | $T_t$ | $G_t$ | $B_{t+1}$ | $\Delta B_{t+1}$ |
|---|---|---|---|---|
| 1890 | 3.249 | 2.081 | 8.477 | -1.446 |
| 1891 | 3.000 | 2.406 | 7.376 | -1.102 |
| 1892 | 2.630 | 2.358 | 6.727 | -0.648 |
| 1893 | 2.703 | 2.476 | 6.883 | 0.156 |
| 1894 | 2.217 | 2.433 | 7.994 | 1.111 |
| 1895 | 2.577 | 2.556 | 7.813 | -0.182 |
| 1896 | 2.426 | 2.256 | 9.125 | 1.312 |
| 1897 | 2.612 | 2.447 | 8.327 | -0.798 |
| 1898 | 2.766 | 2.754 | 7.948 | -0.379 |
| 1899 | 3.339 | 3.643 | 8.213 | 0.265 |
| 1900 | 3.287 | 2.747 | 6.705 | -1.508 |
| 1901 | 3.158 | 2.614 | 5.865 | -0.840 |
| 1902 | 2.735 | 2.189 | 5.421 | -0.444 |
| 1903 | 2.618 | 2.248 | 5.028 | -0.393 |
| 1904 | 2.376 | 2.426 | 4.932 | -0.097 |
| 1905 | 2.390 | 2.356 | 4.471 | -0.461 |
| 1906 | 3.375 | 2.156 | 3.948 | -0.523 |
| 1907 | 2.321 | 1.916 | 3.742 | -0.206 |
| 1908 | 1.985 | 2.081 | 4.219 | 0.476 |
| 1909 | 2.184 | 2.406 | 3.540 | -0.678 |
| 1910 | 2.100 | 2.073 | 3.411 | -0.130 |
| 1911 | 2.103 | 1.992 | 3.340 | -0.070 |
| 1912 | 2.021 | 1.932 | 3.174 | -0.166 |
| 1913 | 1.914 | 1.839 | 3.029 | -0.145 |
| 1914 | 1.854 | 1.784 | 3.235 | 0.206 |
| 1915 | 1.875 | 1.970 | 3.050 | -0.186 |
| 1916 | 1.965 | 1.767 | 2.438 | -0.612 |
| 1917 | 2.210 | 3.845 | 4.921 | 2.484 |
| 1918 | 6.032 | 20.665 | 15.904 | 10.982 |
| 1919 | 6.680 | 23.279 | 31.859 | 15.956 |
| 1920 | 8.358 | 6.722 | 26.881 | -4.978 |
| 1921 | 6.219 | 4.553 | 31.660 | 4.780 |
| 1922 | 5.401 | 3.110 | 30.174 | -1.486 |
| 1923 | 5.165 | 2.821 | 25.277 | -4.897 |
| 1924 | 4.459 | 2.284 | 23.677 | -1.600 |
| 1925 | 4.125 | 2.341 | 21.682 | -1.995 |
| 1926 | 4.123 | 2.312 | 19.235 | -2.447 |
| 1927 | 4.074 | 2.136 | 18.268 | -0.967 |
| 1928 | 4.018 | 2.335 | 16.897 | -1.371 |
| 1929 | 3.902 | 2.513 | 15.407 | -1.490 |
| 1930 | 3.850 | 2.602 | 16.634 | 1.227 |
| 1931 | 3.513 | 3.274 | 21.399 | 4.765 |
| 1932 | 2.618 | 5.338 | 32.325 | 10.926 |
| 1933 | 3.590 | 6.717 | 39.070 | 6.746 |
| 1934 | 5.536 | 10.585 | 40.252 | 1.182 |
| 1935 | 5.793 | 8.733 | 43.131 | 2.879 |
| 1936 | 5.769 | 10.601 | 44.427 | 1.296 |
| 1937 | 6.739 | 8.354 | 41.886 | -2.541 |
| 1938 | 7.667 | 6.511 | 44.464 | 2.578 |
| 1939 | 7.728 | 9.394 | 44.701 | 0.238 |

Legend: All values are percentages of GNP.

### Table A1: Data *Continued*

| | $T_t$ | $G_t$ | $B_{t+1}$ | $\Delta B_{t+1}$ |
|---|---|---|---|---|
| 1940 | 7.535 | 8.934 | 42.602 | -2.100 |
| 1941 | 9.165 | 12.263 | 38.425 | -4.177 |
| 1942 | 12.032 | 26.283 | 42.612 | 4.187 |
| 1943 | 15.784 | 48.955 | 66.303 | 23.691 |
| 1944 | 24.815 | 48.141 | 87.415 | 21.112 |
| 1945 | 23.728 | 45.029 | 110.207 | 22.792 |
| 1946 | 20.402 | 26.343 | 113.871 | 3.663 |
| 1947 | 20.495 | 16.346 | 95.382 | -18.488 |
| 1948 | 19.284 | 12.166 | 82.672 | -12.710 |
| 1949 | 15.893 | 13.361 | 82.305 | -0.367 |
| 1950 | 15.722 | 13.338 | 75.971 | -6.334 |
| 1951 | 18.519 | 13.633 | 64.285 | -11.686 |
| 1952 | 20.399 | 18.177 | 61.080 | -3.205 |
| 1953 | 20.334 | 19.614 | 58.768 | -2.312 |
| 1954 | 18.762 | 17.782 | 60.268 | 1.500 |
| 1955 | 17.576 | 17.090 | 55.831 | -4.438 |
| 1956 | 18.366 | 16.108 | 51.881 | -3.949 |
| 1957 | 18.681 | 16.671 | 48.630 | -3.252 |
| 1958 | 17.658 | 17.067 | 49.548 | 0.918 |
| 1959 | 17.349 | 18.902 | 47.338 | -2.210 |
| 1960 | 18.655 | 17.200 | 45.962 | -1.376 |
| 1961 | 18.317 | 17.675 | 44.653 | -1.309 |
| 1962 | 18.673 | 18.719 | 43.162 | -1.491 |
| 1963 | 18.552 | 18.075 | 41.848 | -1.314 |
| 1964 | 18.553 | 18.174 | 39.527 | -2.321 |
| 1965 | 17.975 | 16.868 | 36.985 | -2.543 |
| 1966 | 18.551 | 17.744 | 34.160 | -2.825 |
| 1967 | 19.275 | 19.072 | 32.659 | -1.501 |
| 1968 | 18.741 | 20.457 | 32.435 | -0.224 |
| 1969 | 20.937 | 19.144 | 28.852 | -3.582 |
| 1970 | 20.002 | 18.801 | 27.888 | -0.965 |
| 1971 | 18.424 | 19.238 | 27.481 | -0.406 |
| 1972 | 18.799 | 19.518 | 26.581 | -0.900 |
| 1973 | 19.030 | 18.828 | 25.080 | -1.501 |
| 1974 | 19.363 | 18.241 | 23.336 | -1.743 |
| 1975 | 18.950 | 20.984 | 24.693 | 1.357 |
| 1976 | 20.000 | 22.944 | 27.794 | 2.480 |
| 1977 | 19.946 | 21.277 | 27.586 | -0.208 |
| 1978 | 20.075 | 21.264 | 26.987 | -0.599 |
| 1979 | 20.594 | 20.487 | 25.507 | -1.480 |
| 1980 | 20.616 | 21.465 | 25.962 | 0.456 |
| 1981 | 21.936 | 22.308 | 25.709 | -0.253 |
| 1982 | 20.238 | 21.644 | 29.035 | 3.326 |
| 1983 | 18.961 | 22.695 | 33.210 | 4.176 |
| 1984 | 19.570 | 21.750 | 34.461 | 1.251 |
| 1985 | 19.461 | 21.655 | 37.345 | 2.884 |
| 1986 | 19.156 | 21.279 | 40.944 | 3.599 |
| 1987 | 20.142 | 20.424 | 41.711 | 0.767 |
| 1988 | 20.081 | 20.153 | 42.148 | 0.437 |

Legend: All values are percentages of GNP.