ON TESTING SUSTAINABILITY OF GOVERNMENT DEFICITS IN A STOCHASTIC ENVIRONMENT

by

HENNING BOHN

19-91

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

On Testing the Sustainability of Government Deficits in a Stochastic Environment

by Henning Bohn*

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

First Draft: May 1991

Current Version: August 1991

* The paper was written while I was visiting the Center for Economic Studies at the University of Munich (Germany). Financial support from the Geewax-Terker Program in Financial Instruments is gratefully acknowledged.

Abstract

In recent years, a number of empirical studies have examined the long-run sustainability of U.S. debt policy. Some studies conclude that U.S. fiscal policy has been sustainable, others disagree. This paper argues that the issue should be reexamined, because the traditional sustainability tests explicitly or implicitly assume that the rate of return on government debt is "on average" above the rate of economic growth, a condition that does not hold for historical U.S. data. The paper derives and implements a new test for sustainability that does not rely on a particular relation between interest rates and growth rates. I conclude that U.S. fiscal policy has historically satisfied a sufficient condition for sustainability.

1. Introduction

In recent years, a number of empirical studies have examined the question whether U.S. fiscal policy has historically been on a sustainable path, i.e., whether it is consistent with the government's intertemporal budget constraint. Some conclude that U.S. fiscal policy has been sustainable, others disagree. (See Hamilton and Flavin 1986, Hakkio and Rush 1986, Kremers 1989, Trehan and Walsh 1988 and 1991, and Wilcox 1989.)[1]

This paper reexamines the sustainability issue and implements a new test for the sustainability of U.S. fiscal policy. A reexamination is needed because the existing tests are based on theoretical models that explicitly or implicitly assume an interest rate on government bonds above the average growth rate of the economy. In fact, interest rates on U.S. government bonds have been below the average growth rate for long periods, including the periods that were studied in the literature. The sustainability condition tested in this paper does not require a specific inequality between average growth rates and interest rates.

The empirical analysis is based on the theoretical work of Bohn (1991c), which derives constraints on sustainable policies for a class of stochastic, dynamically efficient economies with complete markets.[2] A sufficient condition for the sustainability of government policy in such economies is that the level of the primary deficit responds positively to marginal changes in the debt-GNP ratio. The empirical analysis shows that such a positive response exists in U.S. fiscal policy, which suggests that U.S. government policy has historically been sustainable. The empirical part of the paper also comments the high deficits during war periods and suggests that the resulting jumps in debt may have been a source of inference problems in time-series analyses of fiscal data.

---

[1] Corsetti and Roubini (1991) examine the same question for a cross-section of countries.

[2] The complete markets assumption precludes problems of the type examined by Blanchard and Weil (1990). The emphasis in Bohn (1991c) is on examining what form transversality conditions take in a stochastic as opposed to a deterministic setting, given that one has an economy in which Ponzi-schemes are impossible.

The paper is organized as follows. Section 2 reviews the theoretical issues and derives the sufficient condition for sustainability. Section 3 examines the empirical evidence. Section 4 concludes.

2. Conditions for Sustainability

This section reviews the government's intertemporal budget constraint and the transversality condition and then derives a sufficient condition for sustainability.

2.1. What is the appropriate Sustainability Condition?

The first issue is the appropriate form of the intertemporal constraint on fiscal policy. Traditionally, empirical studies have simply asserted that the path of government debt has to satisfy a constraint of the form

$$\lim_{N\to\infty} \frac{1}{(1+r)^N} \cdot E_t[D_{t+N}] = 0 \tag{1}$$

where r is the safe interest rate, D_{t+N} government debt at the start of period t+N, and E_t the conditional expectation at time t. As empirical proxies for the safe interest rate, some studies use the average return on government debt (e.g., Hamilton and Flavin, 1986), others use the N-period compound actual return on government debt (e.g., Wilcox 1989).[3]

However, such an approach is questionable for two reasons. From a theoretical perspective, the question whether or not a transversality constraint has to hold is a general equilibrium issue. In asserting a constraint without providing a general equilibrium setting, one has to rely implicitly on some other body of theory that may or may not be appropriate for the empirical analysis. From an empirical perspective, we know that real returns on U.S. government debt have been rather low, at least prior to

[3] Similar constraints are invoked by Corsetti and Roubini (1991), Hakkio and Rush (1986), and Trehan and Walsh (1988, 1991). The exact definition does not matter much for the argument below, as long as the discounting is done at a rate somewhere near the average rate of return on government bonds. I will use the label "safe interest rate" as shorthand for the precise definitions.

1980. The average real return on U.S. Treasury bills, for example, has been 0.23% for 1929-88 and 1.11% for 1954-1988 (Bohn 1991c). Both values are well below the average U.S. growth rate.

The relation between growth rate and interest rate is important, because the limit in equation (1) will be infinity whenever government debt grows at a rate above the safe interest rate. But in an economy in which the safe interest rate is below the growth rate of aggregate income (GNP) on average, the limit in (1) will be infinity even if the debt-GNP ratio is constant or slightly falling. That is, sustainability tests based on equation (1) would find a policy with constant debt-GNP ratio non-sustainable. Such a conclusion is clearly unwarranted, suggesting that a zero limit in (1) is not a necessary condition for sustainability. Unfortunately, equation (1) is the condition tested in the literature (Hamilton and Flavin 1986, Hakkio and Rush 1986, Kremers 1989, Trehan and Walsh 1988, 1991, Wilcox 1989, Corsetti and Roubini 1991).[4]

One might add that problems with equation (1) should not really come as a surprise, because the deterministic and certainty-equivalence models in which a constraint like (1) can be derived just do not allow steady states with dynamic efficiency combined with a real interest rate below the rate of economic growth. However, dynamic efficiency and a low real interest rate are both key characteristics of historical U.S. data.[5] Thus, theoretical models that can justify equation (1) do not provide an appropriate framework for empirical analysis.

To derive constraints on government policy that apply even if the safe interest rate is relatively low, Bohn (1991c) has examined a simple stochastic setting with risk-averse individuals. The model assumes dynamic efficiency, complete markets, and

---

[4] Kremers and Corsetti-Roubini also examine time series of debt-income (GNP or GDP) ratios, supposedly to test a stronger constraint. However, if the safe interest rate is below the average growth rate, a stationary debt-income ratio is in fact a weaker constraint than (1).

[5] See Abel et al. (1989) for empirical evidence on dynamic efficiency.

infinitely-lived agents (as in Lucas 1978). Because of risk aversion, the safe interest rate can be below the average rate of economic growth.[6] The model implies a transversality constraint on government debt that is similar to the deterministic constraint (1), except that risk-averse individuals discount future debt at a rate that depends on the intertemporal marginal rate of substitution. The relevant constraint on the path of government debt turns out to be

$$\lim_{N\to\infty} E_t[u_{t,N} \cdot D_{t+N}] = 0, \tag{2}$$

where $u_{t,N}$ is the marginal rate of substitution between periods t and t+N. The expectation in (2) can also be interpreted as product of debt times state-contingent claims prices summed over all states of nature. Equation (2) reduces to (1) if individuals are risk-neutral.

The key difference between constraints (1) and (2) is that, according to (2), the rate at which debt D_{t+N} is discounted depends on the probability distribution of debt across states of nature and on the correlation of debt with the marginal rate of substitution. The discount rate is generally not related to the interest rate on government debt, not even if government debt is perfectly safe. For example, if government policy is such that the amount of outstanding (safe) debt is asymptotically proportional to aggregate income, then the correct discount rate in the transversality condition is the rate at which claims on future income would be discounted. Since the discount rate of future income will be above the rate of income growth (as an implication of dynamic efficiency), a debt policy with stable debt-GNP ratio will be found sustainable in the sense of equation (2).

Overall, the point is that equation (2) and not equation (1) is the appropriate sustainability condition in a world with uncertainty. This is the condition that should be tested empirically.

---

[6] Dynamic efficiency is still satisfied because the relevant risky rate of return is above the growth rate on average; see Abel et al. (1989), Zilcha (1991).

2.2. A Strategy for Empirical Testing

At first sight, one might think that tests of condition (2) might proceed analogously to the familiar tests of condition (1), using estimates of the marginal rate of substitution instead of the safe interest rate.[7] Unfortunately, there is a more fundamental inference problem that suggests a different empirical strategy.

The main problem is that when real interest rates are below the growth rate, the debt-GNP ratio will have a downward drift in expectation even if the government is simply rolling-over all debt with interest. A policy of always rolling over debt with interest clearly violates equation (2), which one can see most easily by examining the intertemporal budget constraint associated with equation (2), which is

$$D_t = \sum_{n\geq 0} E_t[u_{t,n} \cdot (T_{t+n} - G_{t+n})]. \tag{3}$$

This constraint shows that initial debt must be balanced by later primary surpluses in at least some periods and some states of nature. The rolling-over strategy violates the transversality constraint (2) because there are some sample paths along which economic growth is (unexpectedly) below the interest rate. However, these sample paths may have very low probability. Most of the time, one will observe a downward drift in the debt-GNP series. The time series provides no indication of a sustainability problem in the sense that many sustainable policies could generate similar sample paths.

To provide a specific example, suppose government policy is to set taxes equal to non-interest outlays for low values of the debt-GNP ratio and to run a primary surplus only when the debt-GNP hits an upper bound. Debt will grow at the rate of interest until

---

[7] In the literature, three types of sustainability tests are common. A first set of papers estimates empirical proxies for the path of discounted government debt (e.g., Hamilton-Flavin 1986, Wilcox 1989). A second set of papers exploits cointegration properties of fiscal data that are implied by the intertemporal budget constraint (e.g., Hakkio and Rush 1986, Trehan and Walsh 1988, 1991). A third set of papers examines the path of debt-GNP ratios (e.g., Kremers 1989). All three approaches test equation (1). The comments of this section apply most directly to the first and third group of tests. Cointegration tests are discussed in Section 3.2.

the debt-GNP hits the upper bound, at which point taxes will be increased to prevent a further increase. Because of the upper bound on the debt-GNP ratio, this policy is sustainable in the sense of (2). Assuming the interest rate is below the average growth rate, debt will grow more slowly than aggregate income in expectation. Most of the time, debt will be rolled-over with interest. A tax increase will be needed only very rarely. But unless one of these rare states of nature occurs along the sample path that is realized, an outside observer cannot distinguish this policy from the non-sustainable policy of always rolling over debt.

More generally, an inference problem arise because the single sample path of the economy that we observe does not always provide enough information about the distribution of fiscal variables across different possible sample paths.[8] Since average returns on U.S. government debt have historically been below the U.S. growth growth rate, the possibility of such inference problems cannot be dismissed easily. To obtain empirical results about the sustainability of government policy, one has to hope that actual government policy belongs to a class of policies for which inferences can be made. Such inferences will have to be conditional on the assumption that actual government policy falls into the relevant class of policies.

This paper will derive such conditional results. The empirical strategy is to exploit the fact that policy parameters can be estimated from a finite sample, if the policy variables are <u>linear</u> functions of a small number of state variables and if there is sufficient sample variation in these state variables. A single sample path will then provide enough information to assess sustainability. The analysis will therefore focus on a class of linear policy rules. The next section will derive sufficient conditions for the

---

[8] The general inference problem and the above example are discussed in more detail in Bohn (1991b). Similar inference problems have been noted in the context of the "Peso problem" in the foreign exchange literature and in the context of the "excess volatility question" in the asset pricing literature.

sustainability of such policies. If an estimated policy satisfies these conditions, one may conclude that government policy is sustainable.[9] If not, one might have to examine whether the policy satisfies a less stringent sufficient condition, or consider the possibility that government policy might be non-linear.

2.3. *Sufficient Conditions for Sustainability*

The set of sufficient conditions for sustainability stated below is motivated by the fact that the government cannot run permanent primary budget deficits (see (3)). If the government runs primary deficits and if debt grows relative to GNP, the government must eventually respond by reducing the deficit. If this reaction happens only at high levels of debt, we may not see it along the observed sample path. (That is the inference problem noted above.) But if taxes and/or spending respond linearly to higher initial debt, one should see a positive reaction of the primary budget surplus to changes in the debt-GNP ratio at all levels of the debt-GNP ratio. One should be able to estimate it. The proposition below will show that such a positive reaction is sufficient for sustainability, provided the other determinants of the ratio of primary surpluses to GNP are bounded.

To state the result formally, some notation is useful. Denote aggregate income (GNP) by Y_t, the "tax rate" by $\tau_t = T_t/Y_t$, the ratio of government spending to income by $g_t=G_t/Y_t$, the debt-income ratio (at the beginning of a period) by $d_t = D_t/Y_t$, and the primary surplus relative to income by $s_t=\tau_t-g_t$. Let R_t be the return on government debt and y_t be the growth rate of income. Then the usual budget equation in levels

$$D_{t+1} = (1+R_{t+1}) \cdot [G_t - T_t + D_t] \tag{4}$$

implies a budget equation in ratio form

$$d_{t+1} = x_{t+1} \cdot [g_t - \tau_t + d_t] = x_{t+1} \cdot [d_t - s_t], \tag{5}$$

---

9 That is, the policy is sustainable if it is indeed a member of the class of policies for which the condition applies. This qualification has to be added because even if the policy appears linear in the sample, one cannot exclude the possibility that it has a non-linearity in some "rare" states of nature that remain unobserved and that may make the policy non-sustainable.

where $x_{t+1} = (1+R_{t+1})/(1+y_{t+1})$. Note that the return on debt has to satisfy the Euler condition $E_t[u_{t,1}\cdot(1+R_{t+1})]=1$. With this notation, the sufficient condition is as follows.

PROPOSITION 1:

If the stream of aggregate income Y_t has a finite present value and if the primary surplus can be written as

$$s_t = \mu_t + \rho \cdot d_t, \tag{6}$$

where μ_t is a bounded stochastic process and $\rho>0$, then government policy satisfies constraints (2) and (3), i.e., policy is sustainable in this sense. ♦

PROOF: See the appendix. ♦

The key element in this statement is the requirement that the government responds to increased initial debt by increasing the primary surplus, i.e., that s_t is linear in d_t with a strictly positive coefficient. Intuitively, the crucial question for testing sustainability is what the time series reveals about government behavior at "high" levels of the debt-income ratio. Given the bounds on μ_t, a positive coefficient implies that the primary surplus will turn positive if d_t is sufficiently high. The main point of the proposition is that even if μ_t is a large negative number so that the primary surplus is negative for a wide range of d_t-values, the condition $\rho>0$ is sufficient to satisfy the transversality constraint (2). The assumption of linearity allows one to draw inferences about the value of ρ from a sample that does not necessarily contain a large number of "high" values of the debt-income ratio, i.e., it eliminates the inference problem noted above.[10]

---

[10] To see the significance of this assumption, consider the policy $s_t = \mu_t + \rho\cdot\max(0,d_t-d^*)$, where $\rho>0$ and where $d^*>0$ is a "large" positive constant. This policy is sustainable. However, if the available sample has d_t-values below d^*, one would not be able to identify ρ. If one (erroneously) estimated (6), one would find $\rho=0$. On the other hand, if one estimates (6) even though the true policy has the non-linear form $s_t = \mu_t + \rho\cdot\max(d_t, d^*)$, one may infer $\rho>0$. But this non-linear policy is not necessarily sustainable.

The other assumptions are largely technical. The assumption of a finite present value of income seems reasonable, because otherwise the economy would have infinite wealth. The boundedness assumption could be replaced by a stationarity assumption combined with a restriction on the correlation between μ_t and the income process (see the appendix). But since the boundedness assumption does not appear to be very stringent—the bound may be as wide as one wishes—more complicated conditions would only be distracting. Thus, Proposition 1 is far from being necessary for sustainability, but it should suffice for empirical analysis. Note that, in contrast to the conditions derived in the literature, this condition does not require a positive rate of return on government debt or a positive safe interest rate. Also, it does not require government bond returns or safe interest rates above the rate of economic growth.

The objective of not imposing restrictions on interest rates also motivates why the proposition has been written in terms of the primary surplus as function of debt rather than as condition on the path of debt (cf. Kremers 1989). If the primary deficit satisfies (6), the debt-income ratio will have the law of motion

$$d_{t+1} = x_{t+1} \cdot (1-\rho) \cdot d_t + x_{t+1} \cdot \mu_t. \tag{7}$$

If the interest rate exceeds the growth rate on average in the sense that $E[x_{t+1}]=x>1$, it may be instructive to examine the path of debt and its unit root properties. One might be able to infer that $\rho>0$, if one finds a coefficient on lagged debt (i.e., an estimate of $x\cdot(1-\rho)$) less or equal to one. However, real interest rates on U.S. government bonds have historically been below the growth rate (at least up to 1980). In addition, the shift to higher real interest rates in the 1980s suggests that the process x_{t+1} may not even have an unconditional expectation and that it may have a conditional expectation that sometimes switches signs. If $x<1$, a precise estimate of how far the regression coefficient is below one and a precise estimate of the value of x would be required to infer a positive value of ρ from (7). If the expectation does not exist, inferences are even more difficult. Because of these complications, the empirical analysis here will focus on equation (6).

If the stochastic process μ_t is stationary, the policy rule (6) links the time series properties of the primary surplus and of debt. If debt has a unit root, the primary surplus should also have one, and (s_t, d_t) should be co-integrated with vector $(1,-\rho)$. If debt is stationary, the primary surplus should be stationary, too. Since the cointegration literature has emphasized such links (e.g., Trehan and Walsh 1991), unit root issues will be reviewed in the empirical section.

Finally, a note on potential additional constraints on the debt-income ratio may be appropriate. As in deterministic models (McCallum 1984), the transversality condition (2) does not require a bounded or even stationary debt-income ratio. However, there may be economies in which the government faces an upper bound on the debt-income ratio or on tax revenues (Blanchard 1984, Kremers 1989, Bohn 1991d). If there are such additional restrictions on government policy, constraint (2) will be necessary but not sufficient for the feasibility of government policy. Fortunately, in the case of $x<1$, a policy with $\rho>0$ should have a debt-income ratio that is usually declining rather than growing (see (7)), which suggests that constraints on the debt-income ratio may well be satisfied. In any case, the empirical analysis will focus on testing the necessary condition (2). The question whether U.S. government policy satisfies more stringent additional conditions would be beyond the scope of this paper and is left for future research.[11]

3. Empirical Evidence

This section will examine two types of empirical evidence. First, a structural model of government behavior based on Barro's tax-smoothing approach will be used to estimate a process for the primary surplus. Second, I will examine the time series properties of the fiscal data.

---

[11] The main problem would be that if one takes the notion of an upper bound on tax revenues serious, it implies an upper bound on the debt-income ratio for all states of nature, not just in expectation (Bohn 1991d; cf. Corsetti and Roubini 1991).

3.1. Estimating the Determinants of the Primary Surplus

The basic framework for the structural analysis is Barro's (1979, 1986a, 1986b) tax-smoothing model.[12] The model considers an optimizing government that minimizes the the cost of tax collection by smoothing marginal tax rates over time. Key features of the optimal policy are that tax rates should only depend on permanent government spending and on initial debt, i.e., not vary over the business cycle or with temporary fluctuations in spending. If one subtracts current non-interest outlays from taxes to obtain the primary surplus (all relative to GNP), the model implies that the level of temporary government spending GVAR and a business cycle variable YVAR are the determinants of the non-debt components of primary surplus, μ_t. When an approximation error ε_t is included, one obtains

$$\mu_t = \alpha_0 + \alpha_G \cdot GVAR_t + \alpha_Y \cdot YVAR_t + \varepsilon_t$$

and

$$s_t = \alpha_0 + \alpha_G \cdot GVAR_t + \alpha_Y \cdot YVAR_t + \rho \cdot d_t + \varepsilon_t \tag{8}$$

where the α's are coefficients. The variables GVAR and YVAR are taken from Barro (1986a) for 1916-83 and updated for 1984-89. The empirical analysis will focus on the question whether the estimate for ρ is positive.[13]

Estimates of equation (8) are in Table 1. All regressions use Ordinary Least Squares (OLS) estimation. White's (1980) robust standard errors are provided to address potential heteroskedasticity problems. Line 1 shows the results for the full sample period 1916-

[12] This model was also used in Kremer's (1989) sustainability paper. It was originally derived as a partial equilibrium model (Barro 1979), but it can easily be embedded in a general equilibrium setting (see the appendix). Regarding alternative approaches, note that a model of government behavior would also be needed for estimating discounted debt, because the discount rate in (2) depends on the stochastic process of debt. Traditional discounted-debt tests (Hamilton-Flavin 1986, Wilcox 1989) appear to be model-independent only because they do not address the discount rate issues discussed in Section 2.

[13] I will not examine the boundedness of μ_t in any formal way. The variable GVAR, the ratio of temporary spending to GNP, is by definition bounded in the interval [-1,1]. Boundedness of YVAR is plausible, because it is just the ratio of actual to potential GNP.

1989. The regression in line 2 excludes the two world wars. Lines 3 and 4 show results for the sample period 1916-83, for which Barro's original regressors are available. Line 3 has results for the entire period, while line 4 focuses on the sample period 1920-82 excluding 1941-47, which is the period Barro (1986a) used. Line 5 shows results for the post-war period 1948-89 and line 6 shows results for the period 1960-84, which is the sample period used by Hamilton and Flavin (1986) and by Wilcox (1989).

The coefficients on d_t are significantly positive in all regressions and they are quantitatively reasonable. For example, the ρ-value of 0.056 in line 1 means that a marginal increase in initial government debt increases the primary surplus by 5.6% on an annual basis (or $5.60 per $100 initial debt). All estimates for ρ are between +3.0% and +5.6%. Thus, there is strong evidence that the U.S. government has satisfied the sustainability condition of Proposition 1.

The variables GVAR and YVAR enter negatively in all regressions, as the Barro model predicts. As in Barro (1986a), the YVAR coefficients are above one in absolute value, which suggests that tax revenues fall more than GNP in a recession instead of falling at the same rate. This is not quite consistent with the basic tax smoothing model, but it might be consistent with optimal policy if distortions (e.g. labor supply elasticities) vary over the business cycle.

Table 2 displays sample averages of interest rates, growth rates, and of the primary surplus. The average primary surplus is negative for the full sample and only slightly positive in those subperiods that do not include major wars. As explained in Bohn (1991c), average primary deficits would be difficult to reconcile with sustainability under risk-neutrality, but they are consistent with sustainability in a model with risk-aversion. The return on government bonds $r=E[R_t]$ is below the average growth rate $y=E[y_t]$ for all sample periods.[14] The fact that $r<y$ holds in the data is noteworthy for two reasons. First,

[14] The return estimates in the table are annual net federal interest payments divided by debt at the start of the fiscal year, minus inflation measured by the GNP-deflator. I also computed a series for

it underscores the doubts about the applicability of models in which dynamic efficiency is linked to the inequality r>y. Tests based on such models might find U.S. government policy non-sustainable, even if it had a constant or declining debt-GNP ratio.[15] Second, the fact that r-y<0 and also ρ>0 suggests that the debt-income ratio has mean reversion (see (7)). This is confirmed in Table 3. Table 3 reports regressions similar to those in table 1, but with s_t replaced by Δd_{t+1}. The coefficient on d_t is significantly negative in all cases. The YVAR and the GVAR variables enter positively in all regressions, as predicted by the model, though not always significantly.

The mean reversion result for debt is interesting in light of Barro's (1979) prediction that the debt-income ratio should be non-stationary, i.e. that $\rho \cong r\text{-}y$.[16] Here, one would have to be seriously concerned about sustainability if the equality $\rho \cong r\text{-}y$ were indeed true, because that would imply $\rho<0$. Still, one may ask whether $\rho \neq r\text{-}y$ implies a rejection of the tax-smoothing model. Fortunately (for the model), the prediction $\rho \cong r\text{-}y$ does not extend to a stochastic setting with risk aversion. Though a general analysis of optimal taxation in stochastic models is clearly beyond the scope of this paper, the appendix shows for a special case that the optimal policy will satisfy $\rho>0$ irrespective of the sign of r-y.

---

x_{t+1} by dividing d_{t+1} by $(d_t\text{-}s_t)$ and obtained a mean of 0.987, which implies $r\text{-}y \cong \text{-}1.2\%$. More data on U.S. interest rates and growth rates are in Bohn (1991c).

[15] In the U.S., the debt-GNP ratio for 1989 is above the pre-war ratios but below the ratios of 1948 and 1960. The ratio in 1983 is above the ratio in 1916, but below the ratios of 1920, 1948, and 1960. The frequently used Hamilton-Flavin debt series (covering 1960-84) also implies a declining debt-income ratio. Both Hamilton-Flavin (1986) and Wilcox (1989) work with interest rates below the U.S. growth rate.

[16] In contrast to this paper, Barro finds no evidence for mean reversion. He uses (scaled) changes in *nominal* government debt as left hand side variable and adds a proxy for expected inflation on the right. One may wonder whether his inability to find mean reversion is due to problems associated with estimating inflation. (See also Kremers 1989.) Note that all data used in lines 3 and 4 are taken directly from Barro (1986a) so that data problems cannot explain the differences.

The results in Tables 1 and 3 are remarkably uniform across different sample periods. In contrast, the policy discussion about the recent "deficit-problems" suggests that U.S. fiscal policy of the 1980s might differ from the previous pattern. To explore this issue, Figure 1 compares actual and predicted primary surpluses. The predictions are based on the regression in Table 1, line 1, which uses the full sample period.[17] The graph confirms that the 1980s are somewhat unusual: Actual primary surpluses have been below the predicted values since 1984. The primary surplus for 1983 is almost exactly as predicted. But from 1983 to 1986, the primary surplus failed to increase as much as the model predicts; it even declined in 1985 and 1986. Since 1986, however, the actual surplus increased has started to catch up, so that the prediction error declined from 1.5% of GNP in 1986 to 0.7% in 1989. Given that the underlying regression has a standard error of 1.4%, these prediction errors are not particularly large. It is therefore not surprising that tests for structural stability do not reveal significant differences between pre- and post-1983 fiscal policy.[18]

Another unusual feature of the 1980s is that real interest rates were much higher than in previous decades. This affects the illustrative sample averages in Table 2 (where the 1980s are dominated by other periods in every sample), but not the main results in Table 1. In contrast to the standard sustainability condition (1), the sustainability condition tested here does not rely on assumptions about interest rates. In addition, Proposition 1 has been formulated in terms of the primary surplus, which excludes interest payments.[19]

---

[17] Predictions based on the other regressions yield similar results.

[18] I tested for changes in the intercept, in the intercept and the ρ-coefficient, and in all four coefficients. Note that the increase in the predicted surplus from 1983-89 is mainly due to cyclical factors; only 0.6% of the predicted 4.2% increase is due to higher debt.

[19] High real interest rate have an indirect influence on the regression of Tables 1 and 3 in that d_t increases faster than it would have otherwise.

Overall, the estimates suggest that the sufficient condition for sustainability has historically been satisfied for the United States. The primary deficit has responded positively to increased debt. Moreover, there is evidence for mean reversion in the debt-income ratio.

3.2. Time Series Properties of Fiscal Data

This section will discuss the time series properties of government debt and of the primary surplus. I will also examine how the above results relate to the literature that uses unit root and cointegration results to draw inferences about sustainability. The main purpose of the section is to defend the above conclusions against potential objections based on unit root tests.

The non-structural analysis of fiscal variables has been deferred so far, because the raw data are difficult to interpret without some reasonable null hypotheses derived from a structural model. A theoretical framework is particularly important here, because statistical tests cannot easily distinguish a process with unit root from a stationary series with high autocorrelation. Unfortunately, these are the relevant alternatives for the series of debt-income ratios. An additional problem for data analysis is created by the war periods, which introduce significant heteroskedasticity into the government spending, deficit, and debt series. Major movements in these series occur over short periods that are special in many ways.

In the literature, inferences about unit roots are usually based on augmented Dickey-Fuller tests that test the null hypothesis that the series are integrated of order one (Dickey and Fuller 1981, Fuller 1976). More recently, Phillips (1987) and Phillips and Perron (1988) have developed versions that are robust against heteroskedasticity. Table 4 shows such unit root tests on d_t and s_t for the sample period 1916-1989, for a longer sample period 1800-1988, and for subperiods that exclude the major wars. The 1800-1988 sample is provided because the unit roots issue concerns the long-run properties of the data, about which a long sample should be most informative.[20] To save space, only the augmented Dickey-Fuller "t-statistics" and the Phillips-Perron "Zτ-statistics" are displayed.[21] Using the full 1916-89 and full 1800-1988 samples, one cannot reject the null hypotheses that government debt is I(1) and that the deficit is I(0), where I(·) denotes the order of integration. Similar results have been obtained in the literature (e.g., Trehan and Walsh 1988, 1991).[22]

The result that debt and the primary deficit have different orders of integration is somewhat disturbing, because it is difficult to reconcile with the sustainability condition (6), unless one assumes that the μ_t process is non-stationary but still bounded. However, I will argue that the unit root result for d_t should not be accepted at face value. The unit root result for d_t is rather odd in light of equation (5) and in light of the previous finding that the interest rate on government debt is below the growth rate, i.e., that $E[x_{t+1}] = x < 1$ (see Table 2). If s_t is stationary and $x<1$, then equation (5) implies that d_t should <u>not</u> have a unit root. However, since x is close to 1, the debt-income ratio should be very highly autocorrelated. Thus, a highly autocorrelated stationary process is the appropriate null hypothesis for d_t. But this is not the hypothesis tested in the unit root tests, which have in fact notoriously low power against highly autocorrelated alternatives. Given the overwhelming evidence against a unit root in s_t, the failure to reject the null hypothesis of the unit root tests should therefore not be seen as evidence that d_t is I(1) but rather as

---

20 The long run data are from Bohn (1991a). The sample period for the structural analysis was constrained by the availability of Barro's regressors.

21 Similar results were obtained with non-augmented Dickey-Fuller tests and with the Phillips-Perron "Zρ-statistics." The Phillips-Perron values are based on a lag window of 3, but the choice did not matter much.

22 A difference to the cointegration literature is that the variables are GNP-shares, not levels. This is done partly for econometric reasons—to eliminate growth trends and potential heteroskedasticity—and partly for plausibility. If fiscal variables were defined as stationary stochastic processes in levels while GNP has a trend or unit root component, the ratio might become implausibly small or large.

evidence that the unit root tests have insufficient power against the relevant highly autocorrelated alternative.

In the fiscal policy context, one should be particularly cautious with statistical inferences because of the special importance of war-periods. Table 4 shows that a unit root in the debt-GNP ratio can be rejected strongly, when the two world wars are excluded from the 1916-89 sample and when the three major wars (Civil War, World War I and World War II) are excluded from the 1800-1988 sample. The unit root rejections for s_t become even stronger.

The impression that the war periods deserve special scrutiny is reinforced when one examines Figure 2. The debt-GNP ratio tends to move towards zero during peacetime periods (with some exceptions), but these movements are interrupted by upward jumps during wars. The wartime jumps are in line with the tax-smoothing model, which predicts that temporary military spending should be largely debt-financed. But the war periods create problems in unit root tests, because high deficits often occur in the later years of a war, i.e., right after debt has already been driven up by the deficits in the early years of the war. This produces a positive link between the change in debt and the level of debt during war years, which obscures the longer run mean reversion—or at least raises the standard errors so much that a unit root cannot be rejected.

To illustrate this effect, Figure 3 plots the change in the debt-income ratio, Δd_{t+1}, against initial debt, d_t. This is the relation fitted by the unit root tests. If one abstracts from the obvious "outliers" (some labeled with years), there appears to be a negative relation. However, the position of the outliers—the 1942-45 values in particular—is such that no significant relation is found when they are included. A common characteristic of all the outliers is that they refer to war-periods. As the non-war results in Table 4 show, one obtains a significantly negative relation between level and subsequent change in debt when the war-years are excluded.

A nice feature of the structural model of Section 3.1 is that the GVAR variable absorbs most of the short-term disturbances created by wars. To illustrate the contrast to the non-structural approaches, Figures 4 and 5 show two different scatterplots of the primary surplus against initial debt. Figure 4 shows the raw data. The war-time outliers are as visible as in Figure 3, perhaps even more clearly. In Figure 5, the primary surplus is adjusted for the estimated "other" systematic determinants of the primary deficit, the variable μ_t. The graph plots $s_t-\mu_t$, the non-cyclical, non-war related component of the primary surplus, against d_t. In contrast to the other graphs, the World War II observations do not show up as outliers (except in that debt was high), indicating that the outliers in the other graphs reflect an omitted variables problem. After adjusting for war-related and cyclical factors, there is a clear positive relation between debt and primary surplus. Separately, it may be worth noting that the 1980s do not show up as outliers in any of these plots.

To conclude, there is no convincing evidence for a unit root in the debt-GNP ratio. A unit root is strongly rejected for non-war periods. Failures to reject a unit root for periods that include wars are likely due to the inability to distinguish between high autocorrelation and a unit root.[23]

This conclusion about the debt-GNP ratio is significant for several reasons. First, the finding that both s_t and d_t are stationary assures that the results of Section 3.1 do not suffer from a spurious regression problem. It also implies that these regressions cannot be interpreted as cointegrating. The difficulty of drawing inferences in the context of war-

---

[23] A caveat may that one should hesitate to make definitive statements about a variable that is strongly affected by war-related spending, given that there are relatively few observations on major wars. Even if the debt-GNP ratio appears stationary over non-war periods, it may be non-stationary as a whole, if war-periods dominate the long-run path of the series. Depending on ones prior beliefs about the frequency and cost of future wars, one may arrive at different judgements about the importance of war-time spending. To argue against stationarity of the debt-GNP series one might, e.g., highlight the fact that the military spending-GNP ratios and the peak debt-GNP ratios increased from Civil War to WWI to WWII (see Figure 2).

related noise explains why a structural model was used for estimating the effect of d_t on s_t.

Second, the unit root results provide a connection to Kremer's (1989) study of the sustainability of government debt.[24] He finds a unit root in the logarithm of the debt-income ratio. The contrast suggests that the choice of levels versus logarithms matters for small sample inferences. Also, note that if one takes logarithms in equation (7) and if μ_t is relatively small, one might find a unit root in $\log(d_{t+1})$ because the multiplicative term involving d_t will (approximately) be transformed into a sum with a unit coefficient on $\log(d_t)$. This will be true even if ρ is far above zero.

Third, the result that there is neither a unit root in the primary deficit nor in government debt raises some questions about the interpretation of sustainability tests based on co-integration. Most of these tests—e.g., Trehan and Walsh (1988,1991), Hakkio and Rush (1986)—use real levels of fiscal variables and find unit roots in government spending, debt, and taxes. From the results here, it appears that the unit root in real debt is either not really there or due to a unit root in GNP. The unit roots in spending and taxes seem to exist in both definitions.

For long sample periods (say, 1800-1988), an analysis in GNP-shares is clearly preferable to an analysis in levels, because there would be extreme heteroskedasticity if levels of fiscal variables were used. For shorter sample periods, this heteroskedasticity problem may not be as obvious, but it remains troubling. If the analysis is done terms of GNP-shares, the stationarity of s_t and d_t implies that regressions involving debt and deficits cannot be interpreted as co-integrating regressions. The non-stationarity of τ_t and g_t together with the stationarity of s_t implies cointegration between τ_t and g_t. This can be

24 In the structural analysis, Kremers also uses Barro's model as theoretical framework, and—like Barro—focuses on the growth of nominal debt, not on the primary surplus or the debt-income ratio as this paper. The above comments on Barro's approach apply analogously.

exploited to obtain insights about the behavior of governments in response to high deficits (Bohn 1991a), but it does not provide information about sustainability.[25]

4. Conclusions

The paper has derived a sufficient condition for the sustainability of government policy that is applicable even if the interest rate on government bonds is below the average rate of economic growth. Based on this condition, the paper concludes that U.S. fiscal policy has historically been sustainable. The paper also suggests that some previous tests of the sustainability of government debt may have to be interpreted cautiously, because they rely on models that implicitly assume a safe interest rate above the average rate of economic growth, an inequality which is not satisfied in the data.

---

[25] Trehan and Walsh (1991) also question the relevance of cointegration restrictions, but because of the sensitivity of these restrictions to assumptions about interest rates. More generally, based on the examples in Section 2, I am not convinced that the stationarity of debt and/or deficits by itself can provide decisive evidence on sustainability (cf. Trehan and Walsh 1991). Regarding Bohn (1991a), the results here provide a new justification for estimating error-corrections models with (τ_t, g_t) without including d_t.

References

Abel, Andrew, Gregory Mankiw, Larry Summers, and Richard Zeckhauser, 1989, Assessing Dynamic Efficiency: Theory and Evidence, Review of Economic Studies 56, 1-20.

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

_______, 1981, Output Effects of Government Purchases, Journal of Political Economy 89, 1086-1121.

_______, 1986a, U.S. Deficits since World War I, Scandinavian Journal of Economics 88, 195-222.

_______, 1986b, The Behavior of United States Deficits, in R. Gordon, ed., The American Business Cycle: Continuity and Change, University of Chicago Press, Chicago, Ill.

Blanchard, Olivier, 1984, Current and Anticipated Deficits, Interest Rates and Economic Activity, European Economic Review 25, 7-27.

_______, and Philippe Weil, 1990, Dynamic Efficiency and Debt Ponzi-Games under Uncertainty, Working Paper, Massachusetts Institute of Technology.

Bohn, Henning, 1991a, Budget Balance through Revenue or Spending Adjustments? Some Historical Evidence for the United States, forthcoming Journal of Monetary Economics.

_______, 1991b, Inference Problems in Assessing the Sustainability of Government Policy, Working Paper, University of Pennsylvania.

_______, 1991c, The Sustainability of Budget Deficits in a Stochastic Economy, Working Paper, University of Pennsylvania.

_______, 1991d, The Sustainability of Budget Deficits with Lump-Sum and with Income-Based Taxation, forthcoming Journal of Money, Credit, and Banking.

Corsetti, Giancarlo, and Nouriel Roubini, 1991, Fiscal Deficits, Public Debt, and Government Solvency: Evidence from OECD Countries, NBER Working Paper No. 3658.

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

Fuller, Wayne, 1976, Introduction to Statistical Time Series, Wiley, New York.

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

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

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

Lucas, Robert, 1978, Asset Prices in an Exchange Economy, Econometrica 46, 1429-1445.

McCallum, Bennett, 1984, Are Bond-financed Deficits Inflationary? A Ricardian Analysis, Journal of Political Economy 92, 125-135

Phillips, P., 1987, Time Series Regressions with Unit Root, Econometrica 55, 277-301.

_______, and P. Perron, 1988, Testing for a Unit Root in Time Series Regression, Biometrica 75, 335-46.

Sahasakul, Chaipat, 1986, The U.S.-Evidence of Optimal Taxation Over Time, Journal of Monetary Economics 18, 251-275.

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

Trehan, Bharat, and Carl Walsh, 1991, Testing Intertemporal Budget Constraints: Theory and Applications to U.S. Federal Budget and Current Account Deficits, Journal of Money, Credit and Banking 23, 210-223.

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

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

Zilcha, Itzhak, 1991, Characterizing Efficiency in Stochastic Overlapping Generations Models, Working Paper, Tel-Aviv University.

Appendix

A1. Proof of Proposition 1

We have to show that $z_n = E_t[u_{t,n} \cdot D_{t+n}]$ converges to zero as $n \to \infty$. That is, given any $\varepsilon > 0$, there must be a value N^* such that $|z_n| < \varepsilon$ for all $n \geq N^*$. By iterating on (5) and (6), one obtains

$$d_{t+n} = (1-\rho)^n \left(\prod_{j=1}^{n} x_{t+j}\right) \cdot d_t + \sum_{i=1}^{n} (1-\rho)^{n-i} \cdot \left(\prod_{j=i}^{n} x_{t+j}\right) \cdot \mu_{t+i-1}$$

Using the relations $z_n = Y_t \cdot E_t u_{t,n} \cdot \prod_{j=1}^{n} (1+y_{t+j}) \cdot d_{t+n}$ and $E_t[u_{t+i,1} \cdot (1+R_{t+i+1})] = 1$, this implies

$$z_n / Y_t = (1-\rho)^n \cdot d_t + \sum_{i=1}^{n} (1-\rho)^{n-i} \cdot E_t[a_{i-1}], \text{ where } a_k = u_{t,k} \cdot \prod_{j=1}^{k} (1+y_{t+j}) \cdot \mu_{t+k}.$$

By assumption, the present value of future income, $V_t = Y_t \cdot \sum_{k \geq 1} E_t u_{t,k} \cdot \prod_{j=1}^{k} (1+y_{t+j})$ is finite. Finiteness of this sum implies that the elements in the sum must converge to zero, i.e., $E_t u_{t,k} \cdot \prod_{j=1}^{k} (1+y_{t+j}) \to 0$ as $k \to \infty$. Combined with a bound on μ_t, this implies $E_t a_k \to 0$ as $k \to \infty$. That is, for any $\delta > 0$ there is an N such that $|E_t a_k| \leq \delta$ for all $k \geq N$. Let $\sum_{i=1}^{N} (1-\rho)^{n-i} \cdot E_t[a_{i-1}] = \Omega$, then for $n \geq N$ we have

$$\left| z_n / Y_t - [(1-\rho)^n \cdot d_t + (1-\rho)^{n-N} \cdot \Omega] \right| = \sum_{i=N+1}^{n} (1-\rho)^{n-i} \cdot |E_t[a_{i-1}]| \leq \delta / \rho$$

Since $(1-\rho)^n \to 0$ as $n \to \infty$, the absolute value of z_n will be less than ε for high enough n, provided one picks $\delta < \varepsilon \cdot \rho / Y_t$. Q.E.D.

To generalize the proposition, suppose μ_t follows the stationary stochastic process $\mu_t = B^* + B(L) e_t$, where e_t is white noise. No boundedness assumption is imposed. To show that $E_t a_k \to 0$ still holds, one needs that the sum

$$E_t \sum_{n=0}^{k-1} u_{t,k} \cdot \prod_{j=1}^{k} (1+y_{t+j}) \cdot e_{t+k-n} \cdot B_n$$

converges to zero. (The B^*-part is unproblematic by the same argument used for μ^* above.) Since B_n and $E_t u_{t,k} \cdot \prod_{j=1}^{k} (1+y_{t+j})$ by themselves converge to zero, sufficiently low where r_t=r is determined by $E_t[(1+r)\cdot\beta\cdot(1+y_{t+1})^{-\alpha}] = 1$. To examine what this condition implies for debt service, consider the class of linear polices $\tau_t = g + \rho\cdot d_t$. Since this class of policies implies $d_{t+1} = (1+r)/(1+y_{t+1})\cdot(1-\rho)\cdot d_t$, substitution into (A2) yields

$$(1-\rho)\cdot E_t\left[(1+r)^2\cdot\beta\cdot(1+y_{t+1})^{-\alpha-1}\right] = 1, \tag{A3}$$

For comparison, note that the optimal policy in a hypothetical economy with complete markets will also fall into this linear class (suggesting that the linearity restriction is not unreasonable) with a parameter $\rho_v = v/(1+v)>0$, where v is defined by $E_t[(1+v)\cdot\beta\cdot(1+y_{t+1})^{1-\alpha}] = 1$. (The value v can be interpreted as ratio of income to the present value of income, which is positive because of dynamic efficiency.) Using the lognormality assumption, one has

$$\log(1+r) = -\log(\beta) + \alpha x - \alpha^2/2\cdot\sigma^2$$

$$\log(1+v) = -\log(\beta) + (\alpha+1)x - (\alpha+1)^2/2\cdot\sigma^2$$

$$\log(1-\rho)+2\cdot\log(1+r) = -\log(\beta) + (\alpha-1)x - (\alpha-1)^2/2\cdot\sigma^2$$

where (x,σ^2) are the mean and variance of $\log(1+y_t)$. After some algebra, this implies

$$\log(1-\rho) + \log(1+v) = -\sigma^2 > 0,$$

hence $\rho>\rho_v>0$. That is, regardless of the interest rate r, the optimizing government will run a primary budget surplus that is at least as large as the surplus it would have run under the optimal policy with complete markets. In particular, the government will not try to exploit an interest rate on safe debt below the average growth rate to run primary deficits.

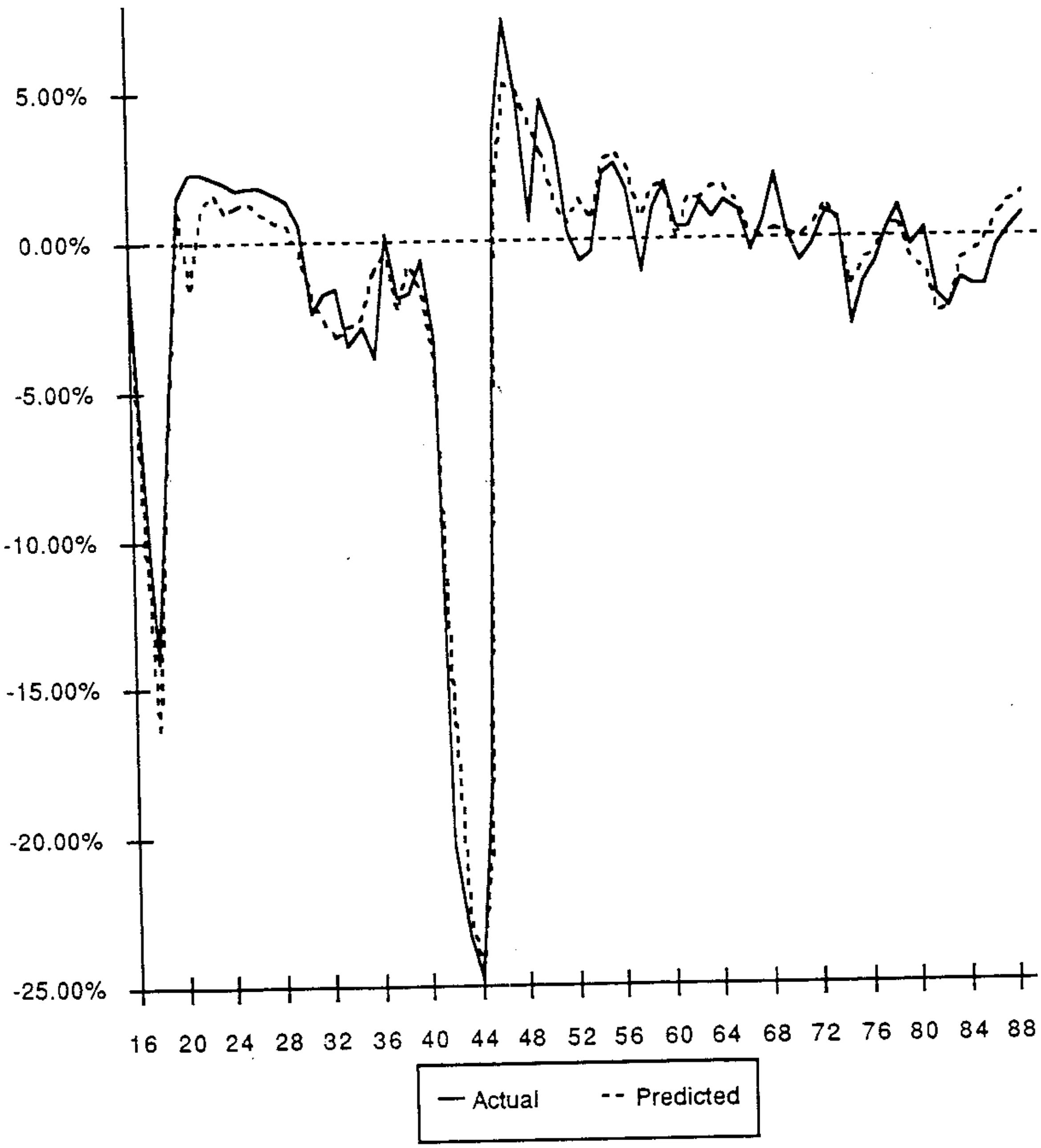
A3. Description of the Data

Except for the budget surplus s_t, the data used in Tables 1 and 3 are based on Barro (1986a). For 1916-83, the series YVAR and GVAR were taken directly from Barro (1986a), except that an adjustment was made for YVAR in 1925 and 1930 where the values did not match those in Barro (1986b). The variables were updated for 1984-89 using the methods explained in Barro (1981, equation 14; 1986a, p. 204) and Sahasakul (1986; equations 20, 21). The resulting values for 1984-89 are 0.0042, 0.0037, 0.0033, 0.0015, 0.0, and -0.0004 for YVAR, and -0.0028, -0.0032, -0.0043, -0.0087, -0.0137, and -0.0159 for GVAR, respectively. The debt series d_t is the ratio of privately held public debt (from the WEFA database and Federal Reserve Banking and Monetary Statistics) to GNP (from the WEFA database and pre-1929 Commerce Department data). It differs lightly from Barro's series on the private debt-GNP ratio because of data revisions (mainly in GNP) and because of the inclusion of minor amounts of non-interest bearing debt. In lines 3 and 4 of Tables 1 and 3, Barro's original debt-GNP series was used for comparison.

The primary budget surplus s_t for calendar years 1929-89 was constructed by dividing the difference of federal receipts and non-interest outlays by nominal GNP, where all series were taken from the National Income and Product Accounts (WEFA database). The surplus for 1916-28 was obtained by interpolating a fiscal year surplus series (with sources described below).

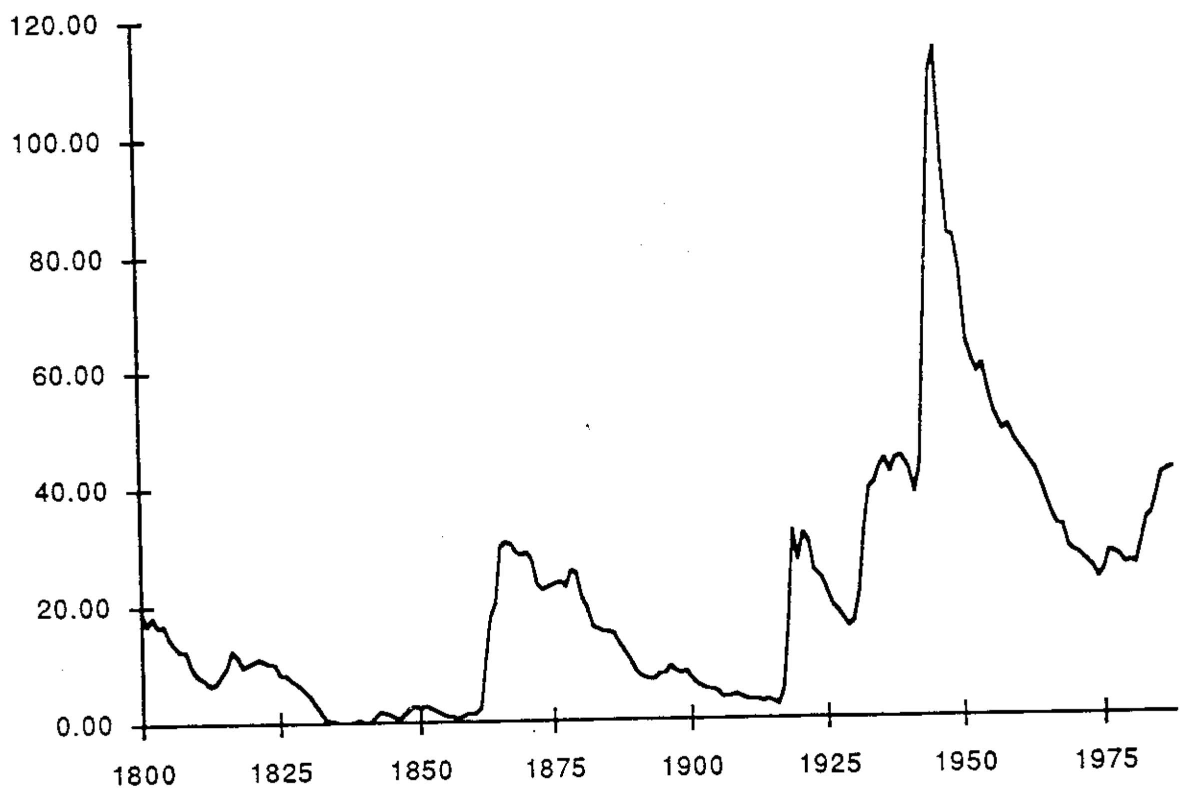
In Table 4 and in Figure 2, the fiscal-year debt series for 1800-1988 refers to public debt. All fiscal year series were taken from Bohn (1991a). They are largely based on the Historical Statistics of the United States and The Budget of the United States. A description of this data set is available from the author. In Table 2, y is average real GNP growth and r is the average ratio of federal interest outlays divided by initial debt on a fiscal year basis, minus the growth rate of the GNP deflator. The GNP deflator was taken as inflation measure to maintain comparability to y.

Figure 1: Actual and Predicted Primary Surpluses



Legend:
The figure shows the primary surplus of the U.S government as ratio to GNP and and it shows the primary surpluses that are predicted by the statistical model of Table 1, line 1.

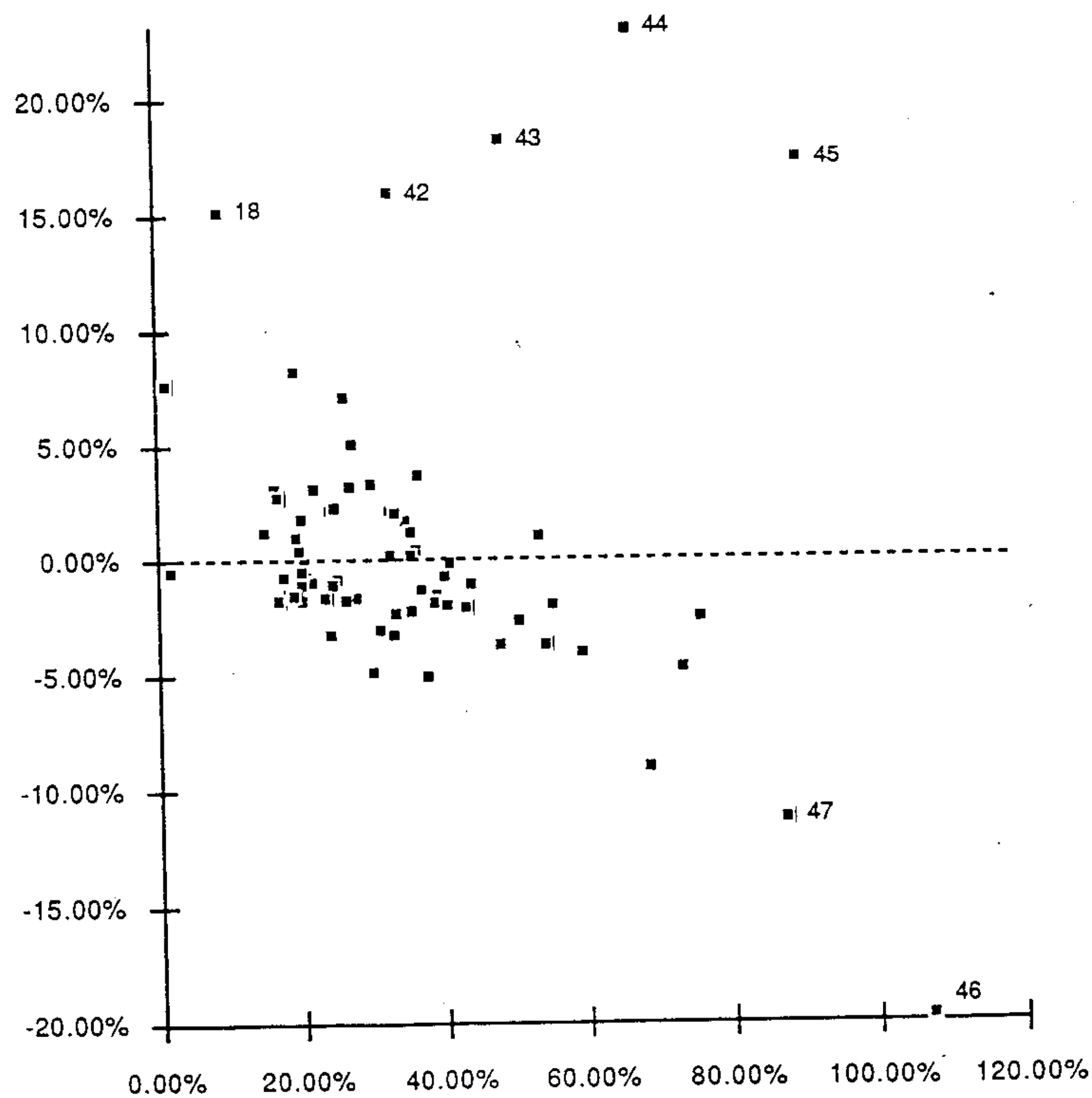
Figure 2: The Ratio of U.S. Government Debt to GNP



Legend:

The figure shows the publicly-held debt of the federal government divided by GNP at the start of each fiscal year (in percent). Data sources are described in the appendix.

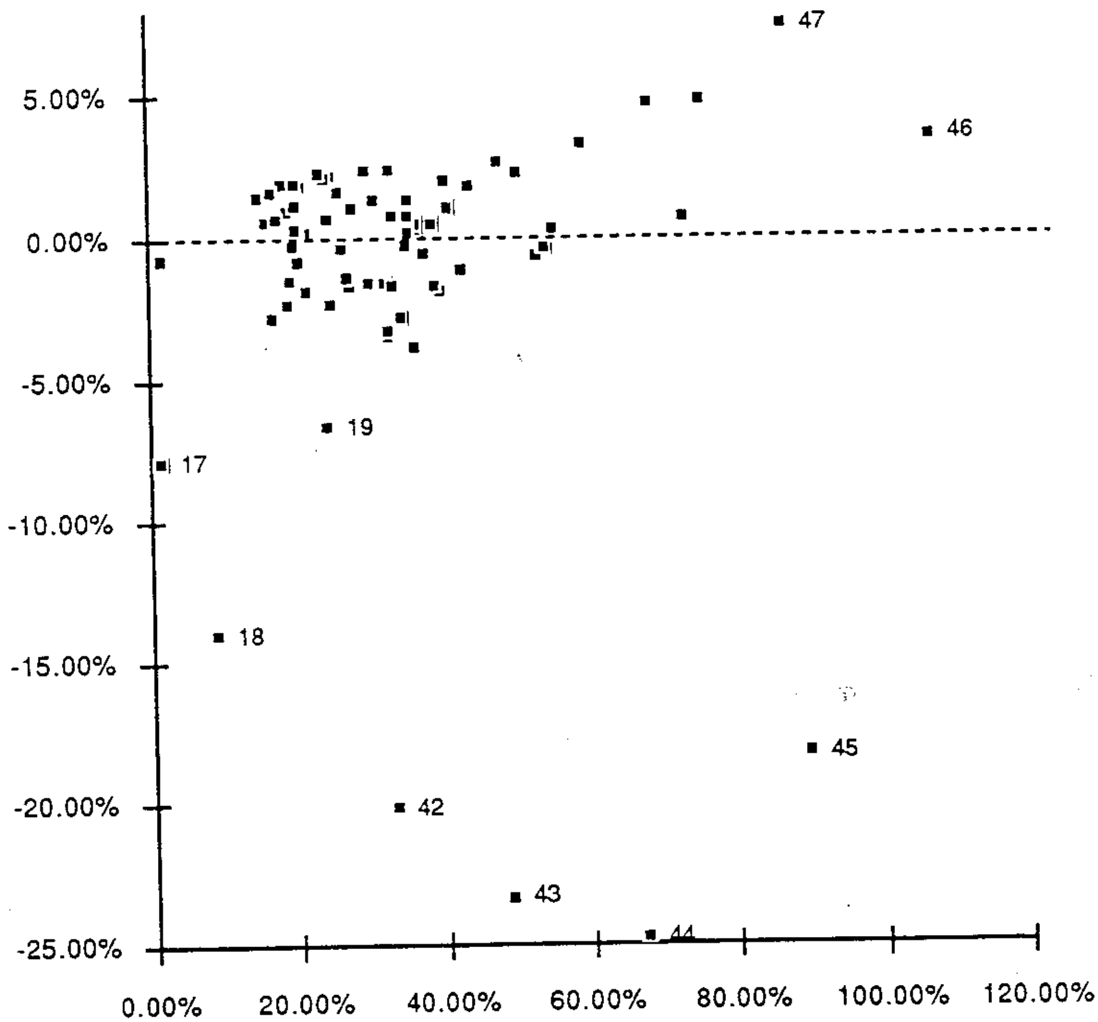
Figure 3: Change in Debt versus Initial Debt



Legend:

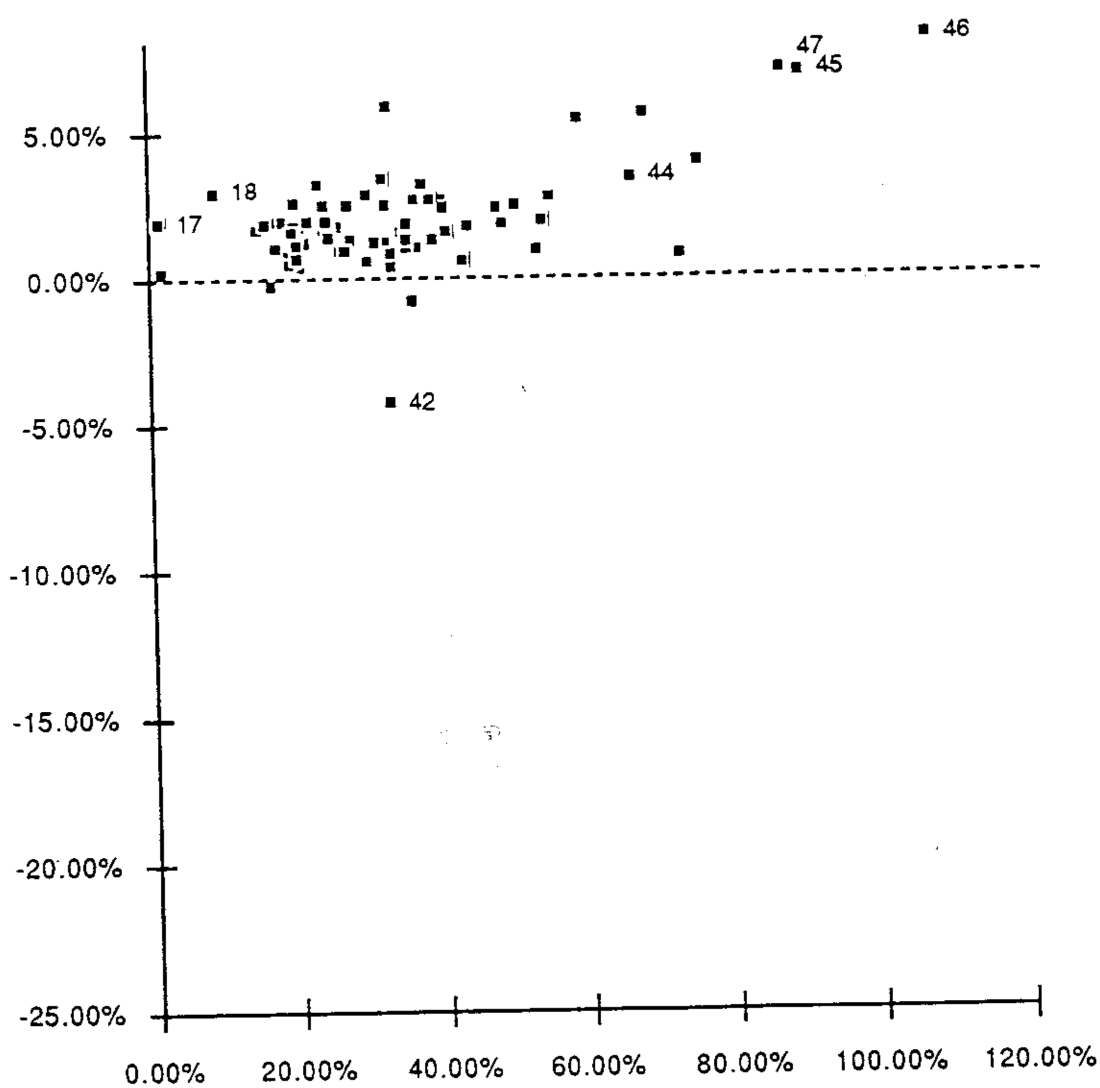
The figure plots privately-held government debt/GNP at the start of a period on the horizontal axis against the change in this variable over the subsequent year on the vertical axis. The plot is for 1916-89, with the labels referring to specific years. Data sources are described in the appendix.

Figure 4: Primary Surplus versus Initial Debt



Legend:
The figure plots privately-held government debt/GNP at the start of a period on the horizontal axis against the primary budget surplus/GNP on the vertical axis. The plot is for 1916-89, with the labels referring to specific years. Data sources are described in the appendix.

Figure 5: Adjusted Primary Surplus versus Initial Debt



Legend:

The figure plots privately-held government debt/GNP at the start of a period on the horizontal axis against the difference of primary surplus/GNP and its "other" determinants μ_t, on the vertical axis. The other determinants, μ_t, are estimated from the regression in Table 1, line 1. The plot is for 1916-89, with the labels referring to specific years. Data sources are described in the appendix.

Table 1: Determinants of the Budget Surplus

Dependent Variable: Budget Surplus s_t

Sample	Constant	GVAR	YVAR	d_t	R^2	σ	DW
1916-89	-0.019	-0.798	-1.503	0.056	0.939	0.014	1.47
	(-5.255)	(-32.348)	(-3.600)	(6.090)			
	[-4.739]	[-22.573]	[-4.340]	[4.639]			
1920-89	-0.009	-0.532	-2.034	0.030	0.635	0.011	1.45
excl. 40-47	(-1.882)	(-3.705)	(-4.758)	(2.806)			
	[-1.548]	[-2.886]	[-4.081]	[2.151]			
1916-83	-0.018	-0.800	-1.446	0.055	0.942	0.015	1.54
	(-4.906)	(-31.672)	(-3.360)	(5.998)			
	[-4.611]	[-22.557]	[-4.154]	[4.693]			
1920-82	-0.008	-0.532	-1.954	0.031	0.630	0.012	1.56
excl. 40-47	(-1.713)	(-3.613)	(-4.439)	(2.817)			
	[-1.442]	[-2.878]	[-3.857]	[2.203]			
1948-89	-0.015	-0.570	-2.135	0.040	0.663	0.010	1.63
	(-3.294)	(-3.741)	(-4.087)	(3.599)			
	[-2.520]	[-3.138]	[-3.573]	[2.425]			
1960-84	-0.012	-0.393	-2.158	0.044	0.730	0.007	1.44
	(-2.057)	(-2.017)	(-4.381)	(2.036)			
	[-2.139]	[-2.097]	[-3.702]	[2.623]			

Legend: () = t-statistics, [] = White's (1980) robust t-statistics, σ = standard error, DW = Durbin-Watson statistic. The variable s_t is the primary U.S. budget surplus divided by GNP, d_t is privately-held debt/GNP at the start of the year, GVAR and YBAR are measures of temporary government spending and of cyclical variations in output, respectively, which are based on Barro (1986a). All estimates are OLS with annual data. Details and data sources are described in the appendix.

Table 2: Sample Averages

Sample Periods	1916-89	1920-89 excl. 40-47	1916-83	1920-82 excl 40-47	1948-89	1960-84
Surplus s_t	-1.25%	0.27%	-1.30%	0.42%	0.40%	0.02%
GNP-growth y	3.29%	2.84%	3.23%	2.71%	3.30%	3.14%
Real Returns r	-0.17%	1.39%	-0.57%	1.02%	0.26%	-0.14%
Difference r-y	-3.46%	-1.44%	-3.81%	-1.69%	-3.05%	-3.28%
t-value on r-y	-3.22	-1.45	-3.29	-1.51	-5.59	-6.26

Legend:
Averages are taken over annual observations. The variable s_t is the primary U.S. budget surplus divided by GNP, GNP-growth is the growth rate of real GNP. Real returns are net interest payments on the public debt divided by initial debt, minus inflation. Data sources are described in the appendix.

Table 3: Determinants of Changes in the Debt-GNP Ratio

Dependent Variable: Change in the Debt-GNP ratio Δd_{t+1}

Sample	Constant	GVAR	YVAR	d_t	R^2	σ	DW
1916-89	0.038	0.753	1.443	-0.131	0.779	0.029	1.90
	(5.160)	(15.060)	(1.705)	(-7.060)			
	[3.305]	[15.668]	[2.189]	[-3.507]			
1920-89	0.019	0.071	2.506	-0.083	0.385	0.023	1.54
excl. 40-47	(2.016)	(0.245)	(2.889)	(-3.755)			
	[2.256]	[0.218]	[2.951]	[-3.755]			
1916-82	0.037	0.779	1.230	-0.133	0.794	0.031	2.12
	(4.767)	(14.863)	(1.352)	(-6.963)			
	[3.278]	[16.018]	[1.743]	[-3.621]			
1920-82	0.017	0.085	2.349	-0.085	0.371	0.025	1.90
excl. 40-47	(1.759)	(0.272)	(2.514)	(-3.650)			
	[1.840]	[0.247]	[2.596]	[-3.318]			
1948-89	0.009	0.161	3.786	-0.050	0.662	0.012	1.27
	(1.203)	(0.584)	(4.967)	(-2.462)			
	[2.030]	[0.656]	[8.165]	[-2.634]			
1960-84	0.016	0.518	3.201	-0.076	0.861	0.007	2.19
	(2.766)	(2.715)	(6.641)	(-3.568)			
	[3.649]	[4.143]	[9.877]	[-3.442]			

Legend: () = t-statistics, [] = White's (1980) robust t-statistics, σ = standard error, DW = Durbin-Watson statistic. The variable d_t is publicly-held debt/GNP at the start of the year, GVAR and YBAR are measures of temporary government spending and of cyclical variations in output, respectively. All estimates are OLS with annual data. Details and data sources are described in the appendix.

Table 4: Unit Root Tests

Panel A: Augmented Dickey-Fuller Tests

Samples:	Calendar Years 1916-89		Fiscal Years 1800-1988	
	All Years	Excl. Wars†	All Years	Excl. Wars††
Debt d_t	2.20	3.07*	2.08	3.68**
Surplus s_t	4.34**	5.12**	6.21**	13.51**

Panel B: Phillips-Perron Tests

Samples:	Calendar Years 1916-89		Fiscal Years 1800-1988	
	All Years	Excl. Wars†	All Years	Excl. Wars††
Debt d_t	2.24	4.58**	1.86	4.16**
Surplus s_t	3.56**	6.24**	5.69**	14.38**

Legend: The test statistics in Panel A are absolute t-values from augmented Dickey-Fuller regressions estimated with two lags. Phillips-Perron "$Z\tau$-statistics" are in Panel B. Both test the null hypothesis of a unit root. The critical values for both statistics are 2.89 at the 5% level and 3.45 at the 1% level (based on 100 observations). The variables s_t and d_t are the primary U.S. budget surplus and U.S. government debt (private debt for 1916-89, public debt for 1800-1988), respectively, each divided by GNP. The 1916-89 sample has calendar year variables (74 observations), the 1800-1988 sample has fiscal year values (189 observations). Details and data sources are described in the appendix.

† = sample period 1920-40 and 1948-89; 63 observations.

†† = sample period 1800-1988, excluding 1861-65, 1917-19, and 1941-47; 174 observations.