Debt and Deficits: Fiscal Analysis with Stationary Ratios

John Y. Campbell Can Gao Ian W. R. Martin*

October 2023

Abstract

We introduce a new measure of a government's fiscal position that exploits cointegrating relationships among fiscal variables and output. The measure is a loglinear combination of tax revenue, government spending and the market value of government debt that—unlike the debt-GDP ratio—is stationary in the US and the UK since World War II. Fiscal deterioration forecasts a long-run decline in spending rather than increased tax revenue or low returns for bondholders. Fiscal adjustment to tax and spending shocks occurs through mean-reversion in tax and spending growth, with a negligible contribution from debt returns.

*John Campbell: Harvard University and NBER, https://scholar.harvard.edu/campbell/home. Can Gao: University of St. Gallen and Swiss Finance Institute, https://sites.google.com/view/can-gao/home. Ian Martin: London School of Economics, https://personal.lse.ac.uk/martiniw.We acknowledge helpful comments from Henning Bohn, Martin Ellison, Hanno Lustig, Greg Mankiw, Casey Mulligan, Emi Nakamura, Jon Steinsson, Andrew Scott, Tuomo Vuolteenaho, and seminar participants at Copenhagen Business School, Harvard, London School of Economics, NBER Summer Institute 2023, University of British Columbia, and Washington University.

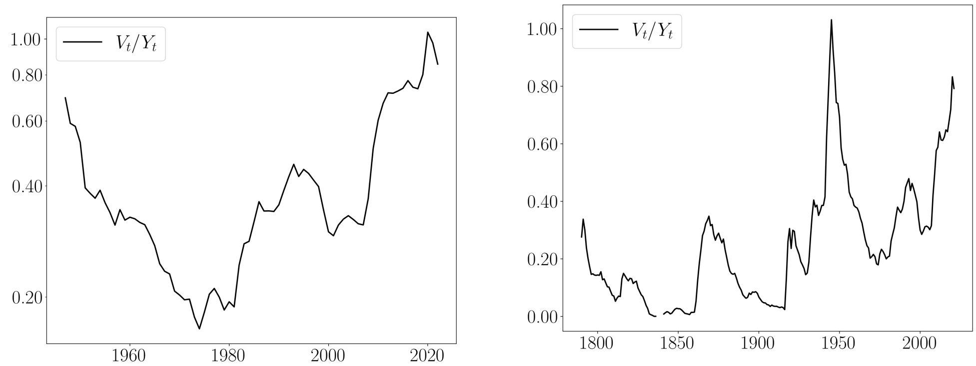
When a government is in a weak fiscal position, then over the long run holders of government debt must earn low returns, or taxes must rise, or spending must fall; or some combination of all three possibilities must occur. As we will show, this follows essentially as a matter of accounting. But which of the three channels is most important empirically?

Any answer to this question requires a suitable definition of the "fiscal position." We will argue that some seemingly natural definitions are problematic. Certainly the primary surplus of a government is an essential ingredient. The primary surplus—the excess of tax revenue over government expenditure—is the flow of resources that the government devotes to servicing its debt. When it is positive, the growth rate of the value of the debt is less than the return on the debt. When it is negative—that is, when the government runs a primary deficit—the debt grows at a faster rate than the return on debt. Under the standard assumption that the expected return on the debt exceeds its growth rate, the value of the debt is the expected discounted value of the primary surpluses that will service it in the future.

To be useful in fiscal analysis, the primary surplus must be scaled in some way so that the resulting ratio is stationary. A common approach is to divide both the primary surplus and the value of debt by GDP to create the surplus-GDP and debt-GDP ratios. If either of these two ratios is stationary, the other should also be because of the present value relation that links surpluses and the value of debt. Many papers treat both ratios as stationary and ask what forces return the debt-GDP ratio to its unconditional mean (see, for example, Henning Bohn (1998, 1991, 2008), John H. Cochrane (2001, 2022, 2023), Olivier Blanchard (2019), and Zhengyang Jiang, Hanno Lustig, Stijn Van Nieuwerburgh and Mindy Z. Xiaolan (2021*b*)).

Contrary to this approach, we find that the debt-GDP ratio does not behave like a stationary time series in US data since World War II. As Figure 1, Panel a, shows, it has drifted persistently up and down for long periods of time. As one would expect, it shows no upward or downward trend; but it also shows no strong tendency to return to a constant mean. A unit root test fails to reject the null hypothesis that the debt-GDP ratio

Figure 1: The debt-GDP ratio is nonstationary in US data.



(a) NIPA data, 1947–2022. Log scale.

(b) Long sample, 1790–2022. Linear scale.

has a unit root, and cointegration tests fail to find statistically significant evidence that government debt is cointegrated with GDP. This nonstationarity helps to explain the (at first sight puzzling) finding in this literature that the debt-GDP ratio is not a successful predictor of fiscal outcomes.

From a theoretical perspective, the nonstationarity of debt-GDP is not particularly surprising: for example, Robert J. Barro (1979) writes, "There is no force that causes the ratio of debt to income to approach some target value".[1] Even if one believes that economic forces act to make the primary surplus-GDP ratio and the debt-GDP ratio truly stationary in the very long run—and the longer series shown in Figure 1, Panel b does not support this view—the persistence of these time series implies that it is inadvisable to model them using the standard techniques of stationary time-series analysis (John Y. Campbell and Pierre Perron, 1991).[2]

An alternative approach is to scale the primary surplus by the value of debt, and to work with the primary surplus-debt ratio. In an economy in which the return on the debt and the growth rate of the debt are stationary, the primary surplus-debt ratio should also be stationary.[3]

[1]By contrast, a trend in debt-GDP would be surprising since it would imply arbitrarily large or small values for this ratio in the distant future.

[2]Appendix A.1.1 describes our data sources.

[3]Indeed, standard unit root tests reject the null hypothesis that the primary surplus-debt ratio has a unit root in favor of the alternative that it is stationary. However, this is also true of the primary surplus-GDP ratio as we show in Appendix Table A.2.

The primary surplus-debt ratio is analogous in the fiscal context to the dividend-price ratio on a stock. Just as a corporation pays dividends to the owners of its stock, so the government pays primary surpluses to the owners of its debt. This suggests the possibility of analyzing the primary surplus-debt ratio using a John Y. Campbell and Robert J. Shiller (1988) loglinearization to relate it to future log returns on debt and log growth rates of primary surpluses.

Two problems arise in doing so, and both result from the fact that the primary surplus can be negative. First, the log growth rate of the primary surplus is ill-defined when the surplus is negative. Second, an exogenous increase in the debt, which worsens the fiscal position of the government, can either raise or lower the primary surplus-debt ratio depending on whether the primary surplus is positive or negative. Thus, the effect of a given shock to the primary surplus-debt ratio depends on the sign of the ratio. Both these problems also afflict the standard analysis of the primary surplus-GDP ratio.

In this paper we develop an alternative loglinear analysis, related to the work of Chryssi Giannitsarou, Andrew Scott and Eric M. Leeper (2006) and Antje Berndt, Hanno Lustig and Şevin Yeltekin (2012), that solves these problems. Our approach is to approximate the primary surplus-debt ratio in a way that can be loglinearly related to the growth rates of tax revenue and of government expenditure. Both revenue and expenditure are always positive, so their log growth rates are well defined; and our loglinear approximation to the primary surplus-debt ratio has the appealing property that an increase in debt always reduces it, whether the primary surplus is currently positive or negative.

The approximations developed by Giannitsarou, Scott and Leeper (2006) and Berndt, Lustig and Yeltekin (2012) are similar in spirit but rely on the assumption that the tax revenue-debt and government expenditure-debt ratios are stationary, so that one can approximate around their means. In the US data we find to the contrary that neither of these ratios are stationary. Instead, their logs are cointegrated with a cointegrating vector that

Primary surpluses are noisy enough that nonstationary dynamics in scaled surplus may be hard for standard tests to detect. For this reason we do not emphasize unit root test results for ratios with the primary surplus in the numerator.

is close to but not equal to a unit vector. We use this finding of cointegration to develop an approximation, related to the work of Can Gao and Ian W. R. Martin (2021), that does not rely on inappropriate stationarity assumptions.

As the resulting measure of the fiscal position is stationary, it is a useful predictor variable for fiscal analysis. We use it to explore the dynamics of debt, tax revenue, and government expenditure in US data since World War II. By contrast with the nonstationarity of the debt-GDP ratio, we find that the tax revenue-GDP ratio appears to be stationary. The government expenditure-GDP ratio, on the other hand, appears nonstationary, which implies nonstationary primary surplus-GDP and debt-GDP ratios. Given these results, we estimate a vector autoregression (VAR) model including two growth rates—the return on debt and the growth rate of tax revenue—and two stationary ratios: the primary surplus-debt ratio (or rather our loglinearized approximation to it) and the tax revenue-GDP ratio.

Our main empirical findings are as follows. First, expected returns on government debt, while time-varying, are not variable or persistent enough to contribute importantly to the dynamics of the primary surplus-debt ratio. Second, the mean-reversion of the primary surplus-debt ratio occurs in the short run through changes in tax revenue, but in the longer run more than all the adjustment occurs through changes in government expenditure. This finding relies critically on the inclusion of the tax revenue-GDP ratio in the VAR model, and it reflects the fact that faster growth of tax revenue raises the tax revenue-GDP ratio, predicting slower growth of GDP and eventually slower growth in tax revenue. Third, fiscal adjustment to shocks in tax revenue and government expenditures occurs almost entirely through mean-reversion in the growth rates of taxes and expenditures. Expected returns on government debt again have little importance, and the same is true for unexpected returns on debt contemporaneous with tax and expenditure shocks.

We repeat the analysis for the UK and find similar results. While the evidence is not completely decisive, the surplus-debt ratio appears stationary and the debt-GDP ratio nonstationary, as in the US. We also find that the tax-GDP ratio is stationary whereas the spending-GDP ratio is nonstationary; again, these mirror our findings for the US, though the tax-GDP ratio is somewhat more persistent in the UK. We therefore estimate the same VAR system as we do for the US, and find the same sign pattern on statistically significant coefficients; as in the US, the variance decomposition reveals that shocks to the fiscal position are resolved, in the long run, by movements in spending rather than in taxes or returns.

Two caveats should be kept in mind when interpreting our results. First, because we conduct a reduced-form time-series analysis, we cannot make causal statements about fiscal dynamics. For example, our finding that an increase in the tax revenue-GDP ratio predicts slower GDP growth does not prove that high taxes cause lower growth as argued by Carmen M. Reinhart and Kenneth S. Rogoff (2010) and Alberto Alesina, Carlo Favero and Francesco Giavazzi (2020).

For the same reason we cannot resolve the debate about the fiscal theory of the price level (Thomas Sargent and Neil Wallace, 1981; Eric M. Leeper, 1991; Christopher A. Sims, 1994; Michael Woodford, 1995; Cochrane, 2001, 2023). According to traditional analysis, the ability of the primary surplus-debt ratio to predict future fiscal adjustment is causal, in that a given value of the debt forces the government to run future primary surpluses that will pay it off. According to the fiscal theory of the price level, the predictive relationship reflects reverse causality: the debt has the value that is consistent with an exogenous path of future surpluses, as in a forward-looking asset pricing model of the sort analyzed by John Y. Campbell and Robert J. Shiller (1987); Campbell and Shiller (1988). If the debt promises to make fixed nominal payments, the required adjustment in value can occur largely through changes in the price level, although also in part through changes in long-term nominal interest rates (Cochrane, 2001).

Second, we take the returns on government debt as given, measuring them in the data without requiring them to satisfy the restrictions of any asset pricing model other than the weak restriction that they are high enough to rule out the existence of a bubble in government debt. We do not address the question, studied by Zhengyang Jiang, Hanno Lustig, Stijn Van Nieuwerburgh and Mindy Z. Xiaolan (2021*a*), of whether the measured return is too low to be consistent with the risk of the government debt, or the related question, discussed by Robin Greenwood, Samuel G. Hanson and Jeremy C. Stein (2015), Arvind Krishnamurthy and Annette Vissing-Jorgensen (2012), Ricardo Reis (2022), and Atif R. Mian, Ludwig Straub and Amir Sufi (2022), of whether government debt offers a convenience yield that investors value separately from its return.

The organization of the paper is as follows. In Section 1 we present a simple steady-state analysis of the primary surplus-debt ratio. This motivates the dynamic framework for fiscal analysis introduced in Section 2. We apply the framework empirically to US data in Section 3, and to UK data in Section 4. Section 5 concludes. An online appendix (John Y. Campbell, Can Gao and Ian W. R. Martin, 2023) presents supplementary details.

1 The primary surplus-debt ratio in steady state

By definition, the gross return on government debt is

$$R_{t+1} = \frac{V_{t+1} + T_{t+1} - X_{t+1}}{V_t}\,. \tag{1}$$

Here R_{t+1} is the return on debt (including money) from time t to $t+1$, V_t is the total value of debt (including money) in period t, T_{t+1} is tax income and X_{t+1} is expenditure. Everything is in real terms. We define the surplus as $S_t = T_t - X_t$ and assume throughout that the gross return R_{t+1} is strictly positive. Note that the debt return R_{t+1} should only be interpreted as a riskless interest rate in the special case in which all government debt is short-term real debt. We allow debt to be risky: the realized return on debt is low if, for example, real yields rise, or if there is a sudden unexpected inflation or explicit default.

As a first step toward a simple benchmark, let us imagine that conditional expectations of growth in tax, spending, and the debt are all equal to some constant, G.[4] Similarly, let us suppose that the conditionally ex-

[4] This assumption is not unreasonable for unconditional expectations. Table A.1 shows that the sample averages of log tax growth, log spending growth, and log debt growth are all approximately equal in our sample period. They are also all approximately equal to log GDP growth, consistent with the absence of a trend in the log debt-GDP

pected return on debt equals R. Equation (1) then implies

$$R = \mathbb{E}_t \frac{V_{t+1}}{V_t} + \mathbb{E}_t \frac{T_{t+1}}{T_t}\frac{T_t}{V_t} - \mathbb{E}_t \frac{X_{t+1}}{X_t}\frac{X_t}{V_t} = G\left(1 + \frac{S_t}{V_t}\right). \tag{2}$$

It follows that the primary surplus-debt ratio is a constant:

$$\log\left(1 + \frac{S_t}{V_t}\right) = \log R - \log G. \tag{3}$$

We write the ratio in this form for comparability with the more general analysis below. When $R > G$, the government must run primary surpluses to pay off its debt. By contrast, if $R \le G$ the government need not run surpluses: even an unexpected increase in debt—for example, to fight a war—never needs to be paid off. In this case, the value of the debt reflects the presence of a rational bubble. In our more general analysis of Section 2, we will rule out this possibility a priori.

Equation (3) exhibits the primary surplus-debt ratio as a natural quantity of interest, analogous to the dividend-price ratio in the Gordon growth model. Figure 2 shows the evolution of the surplus-debt ratio, S_t/V_t, in the US from 1947 to 2022. As the surplus can take negative values, we plot the series on a linear scale. (We provide a detailed description of our data sources in Appendix A.1.1, and summary statistics are provided in Table A.1.) Although surplus-debt is not constant as it would be in a Gordon-growth-type model, it does appear to be stationary.[5]

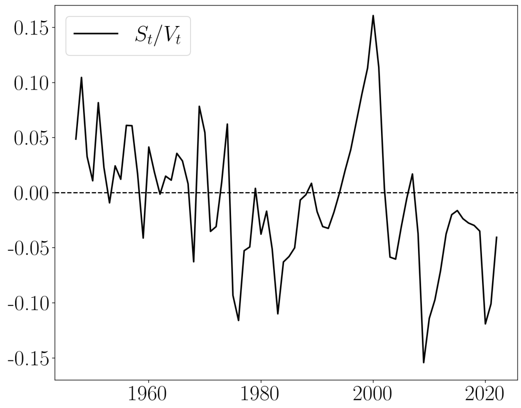
2 A framework for fiscal analysis

The simple benchmark (3) is unrealistic in various important ways: for one thing, it implies that surplus cannot switch sign. To set up an empirically useful framework, we will have to account for the fact that conditionally

ratio. Of course conditional expectations vary in the data, as we discuss later.

[5]This impression is supported by an augmented Dickey–Fuller (ADF) test, reported in Table A.2, which rejects the presence of a unit root at the 99% confidence level. Although unit root tests can have poor finite-sample properties for ratios with noisy numerators such as the primary surplus, this finding, together with the theoretical presumption that the surplus-debt ratio should be stationary, gives us confidence to base our analysis on a stationarity assumption.

Figure 2: The surplus-debt ratio is stationary in postwar data. Linear scale. NIPA data, 1947–2022.



expected tax growth, spending growth, debt growth, and so on, vary over time. We now present a general approach to doing so. In our framework all these quantities are allowed to be time-varying, though their *unconditional* means are all equal to each other (and to unconditionally expected GDP growth) so that tax-debt, spending-debt, and debt-GDP ratios do not trend upwards or downwards over time.

To make a start, rewrite equation (1) as

$$R_{t+1} = \frac{V_{t+1}}{V_t}\left(1 + \frac{S_{t+1}}{V_{t+1}}\right). \tag{4}$$

Taking logs of (4), we have

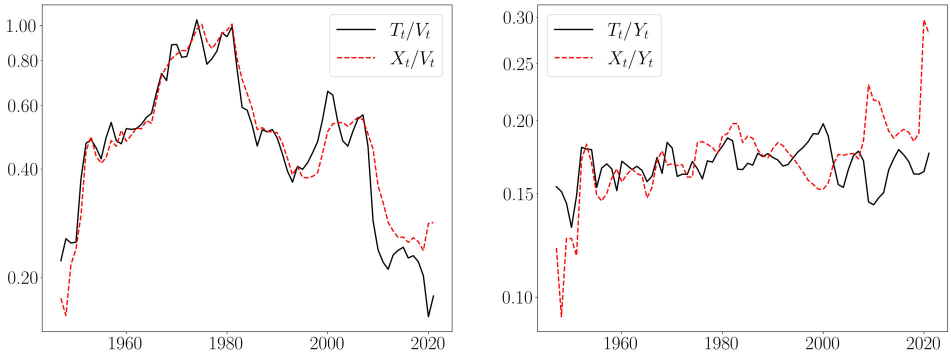
$$r_{t+1} = \Delta v_{t+1} + \log\left(1 + \frac{S_{t+1}}{V_{t+1}}\right). \tag{5}$$

An uncomfortable feature of the post-war data is that the time-series average of the surplus-debt ratio is negative over the sample period, as illustrated in Figure 2. If we believe that this sample average is an accurate measure of the true population average, then it follows from identity (5) that

$$\underbrace{\mathbb{E}\, r_{t+1}}_{\text{“}R\text{”}} - \underbrace{\mathbb{E}\,\Delta v_{t+1}}_{\text{“}G\text{”}} = \mathbb{E}\log\left(1 + \frac{S_{t+1}}{V_{t+1}}\right) < 0\,. \tag{6}$$

This is an “$R < G$” condition. But, as we will show, if the expected

Figure 3: The spending-debt and tax-debt ratios are nonstationary. The spending-GDP ratio is also nonstationary, but the tax-GDP ratio is stationary.



log return on the debt is less than its expected log growth rate, then we are forced to conclude that the value of the debt reflects the presence of a rational bubble. We rule out this possibility by imposing a positive population mean $\mathbb{E} \log(1 + S_t/V_t) > 0$ in our empirical work.

The left panel of Figure 3 breaks the primary surplus $S_t = T_t - X_t$ into its constituent parts, plotting the tax-debt and spending-debt ratios separately. Again, the impression which emerges from these figures is confirmed by ADF tests: neither $\tau v_t = \log T_t/V_t$ nor $xv_t = \log X_t/V_t$ is stationary, despite the fact that the surplus-debt ratio is stationary. These facts place important constraints on how we set up our analysis.

The right panel of Figure 3 plots tax-GDP and spending-GDP over time. By now it may come as no surprise that spending-GDP is not stationary. But tax-GDP *is* stationary. (We report ADF tests in Table A.2.) This important empirical fact supplies us with another stationary variable to take into account when we analyze fiscal dynamics.

2.1 A loglinear measure of the fiscal position

The measure of the surplus-debt ratio that appears on the right-hand side of (5) echoes the dividend-price ratio measure, $\log(1 + D_{t+1}/P_{t+1})$, used by Gao and Martin (2021). It allows surplus to go negative; moreover, the measure is in natural units, in the sense that $\log(1 + S_{t+1}/V_{t+1})$ is approximately equal to S_{t+1}/V_{t+1} if surplus-debt is small. It can be written in terms of the log tax-debt ratio, $\tau v_t = \log(T_t/V_t)$, and the log spending-debt ratio, $xv_t = \log(X_t/V_t)$, as

$$\log\left(1 + \frac{S_{t+1}}{V_{t+1}}\right) = \log\left(1 + e^{\tau v_{t+1}} - e^{xv_{t+1}}\right) . \tag{7}$$

To construct a tractable measure of the fiscal position, we linearize equation (7) in τv_t and xv_t. In doing so, we exploit the fact that while neither tax-debt, τv_t, nor spending-debt, xv_t, is stationary over the postwar sample (as discussed in the previous section and shown in Figure 3) they do appear to be cointegrated. Table A.3 reports results of Johansen tests that indicate a cointegrating relationship: that is, $\tau v_t - \beta\, xv_t$ is stationary for some constant β. The estimates of β are close to but slightly less than one. Likewise, $\log(1 + S_t/V_t)$ is stationary, as discussed in the previous section. We use these facts to guide our linearization.

Specifically, linearizing $\log\left(1 + e^{\tau v_{t+1}} - e^{xv_{t+1}}\right)$ around $(\tau v_{t+1}, xv_{t+1}) = (\log a, \log b)$, where a and b are both positive, we have

$$\log\left(1 + e^{\tau v_{t+1}} - e^{xv_{t+1}}\right) = k + \frac{1}{1+a-b}\left(a\, \tau v_{t+1} - b\, xv_{t+1}\right) \tag{8}$$

up to higher order terms in τv_{t+1} and xv_{t+1}, where

$$k = \log\left(1 + a - b\right) + \frac{b \log b - a \log a}{1 + a - b} . \tag{9}$$

We choose a and b to satisfy two conditions. First, we want to linearize around the unconditional mean of $\log(1 + S_{t+1}/V_{t+1})$: that is, we require

$$\log(1 + a - b) = \mathbb{E} \log\left(1 + \frac{S_t}{V_t}\right) . \tag{10}$$

We write $\mathbb{E}\log(1+S_t/V_t) = -\log\rho$, where $\rho > 0$. As noted in the discussion following equation (6), we assume that $\mathbb{E}\log(1+S_t/V_t) > 0$, or equivalently that $\rho < 1$. This is equivalent to imposing an a priori constraint that the government must ultimately pay off its debt. Thus, to summarize, ρ must lie between zero and one. In this notation, equation (10) becomes

$$1 + a - b = \frac{1}{\rho} . \tag{11}$$

Second, we want the right-hand side of (8) to be stationary, as the left-hand side is. Given the cointegrating relationship between τv_t and xv_t, this requires that

$$\frac{b}{a} = \beta\,. \tag{12}$$

Equations (11) and (12) jointly determine a and b in terms of β and ρ. We have

$$a = \frac{1}{1-\beta}\frac{1-\rho}{\rho} \quad \text{and} \quad b = \frac{\beta}{1-\beta}\frac{1-\rho}{\rho}\,. \tag{13}$$

As a and b are positive, and $0 < \rho < 1$ by assumption, we must have $0 < \beta < 1$. Plugging these choices of a and b back into (8), we have our linearization

$$\log\left(1 + \frac{S_{t+1}}{V_{t+1}}\right) = \log\left(1 + e^{\tau v_{t+1}} - e^{xv_{t+1}}\right) = \underbrace{k + \frac{1-\rho}{1-\beta}\left(\tau v_{t+1} - \beta\, xv_{t+1}\right)}_{sv_{t+1}}, \tag{14}$$

where the first equality follows from the definition of surplus. Here k is as in equation (9) with a and b given by (13).

We will refer to the quantity on the far right-hand side of equation (14) as sv_{t+1} and will use it as our measure of the government's fiscal position. That is, we define

$$sv_t = k + \frac{1-\rho}{1-\beta}\left(\tau v_t - \beta\, xv_t\right) \tag{15}$$

where

$$k = \rho \log \rho + (1-\rho)\log\frac{1-\rho}{1-\beta} - \frac{1-\rho}{1-\beta}\beta\log\beta\,, \tag{16}$$

so that sv_t is a linearization of $\log(1 + S_t/V_t)$ that, like $\log(1 + S_t/V_t)$, is stationary.

The two quantities differ in one important way, however. As the level of debt rises with surplus held fixed, sv_t declines whether the surplus is positive or negative. This follows from the definition (15), given that ρ and β lie between zero and one. Similarly, sv_t declines when tax falls or when spending rises with other quantities held fixed. Thus we can think of sv_t as a measure of the fiscal position: it is high when the government is in a strong fiscal position, and low when the government is in a weak fiscal position. By contrast, the more conventional measures S_t/V_t and

$\log(1 + S_t/V_t)$ are harder to interpret: as the debt grows, they go *down* if surplus is positive, but *up* if the surplus is negative.

2.2 A present value model for the fiscal position

The linearity of sv_t allows us to relate it to fundamentals in a linear present value framework. Inserting the linearization (14) into the exact identity (5), we have

$$r_{t+1} = \Delta v_{t+1} + sv_{t+1}\,. \tag{17}$$

Taking differences of (15) and rearranging, we have

$$(1-\rho)\Delta v_{t+1} = \frac{1-\rho}{1-\beta}\Delta\tau_{t+1} - \beta\frac{1-\rho}{1-\beta}\Delta x_{t+1} - \Delta sv_{t+1}\,. \tag{18}$$

We use (18) to eliminate Δv_{t+1} from (17), giving, after some rearrangement,

$$sv_t = (1-\rho)\left[r_{t+1} - \frac{1}{1-\beta}\Delta\tau_{t+1} + \frac{\beta}{1-\beta}\Delta x_{t+1}\right] + \rho\, sv_{t+1}\,. \tag{19}$$

We now solve forward in the usual way, to find that

$$sv_t = (1-\rho)\sum_{j=0}^{T-1}\rho^j\left[r_{t+1+j} - \frac{1}{1-\beta}\Delta\tau_{t+1+j} + \frac{\beta}{1-\beta}\Delta x_{t+1+j}\right] + \rho^T sv_{t+T}\,. \tag{20}$$

Stationarity implies that sv_t is not explosive, so that $\lim_{T\to\infty}\rho^T sv_{t+T} = 0$. In the limit as $T \to \infty$, we therefore have the dynamic generalization of the static present value formula (3) that we were seeking:

$$sv_t = (1-\rho)\sum_{j=0}^{\infty}\rho^j\left[r_{t+1+j} - \frac{1}{1-\beta}\Delta\tau_{t+1+j} + \frac{\beta}{1-\beta}\Delta x_{t+1+j}\right]\,. \tag{21}$$

In other words, if the government is in a strong fiscal position (sv_t is high), then either the holders of government debt will earn high log returns, or taxes will grow slowly, or government expenditure will grow rapidly, or some combination of the above will occur, at some point in the future. This relationship is a loglinear approximation to an accounting identity, so it holds ex post. It also holds ex ante for rational expectations, and indeed for any subjective expectations that respect identities.

Four further points about equation (21) are worth noting. First, as $(1-\rho)\sum_{j=0}^{\infty}\rho^j = 1$, the right-hand side of (21) can be interpreted as a weighted average. This means that we have the unconditional relationship

$$\mathbb{E}\, sv_t = \mathbb{E}\, r_t - \frac{1}{1-\beta}\,\mathbb{E}\,\Delta\tau_t + \frac{\beta}{1-\beta}\,\mathbb{E}\,\Delta x_t\,. \tag{22}$$

As noted at the beginning of Section 2, we must have equal unconditional growth rates of tax, spending, and debt so that fiscal ratios do not trend upwards or downwards over time. (This is borne out in postwar US data: log tax growth, log spending growth, and log debt growth have means of, respectively, 0.031, 0.030, and 0.030.) Writing $\mathbb{E}\,\Delta\tau_t = \mathbb{E}\,\Delta x_t = \mathbb{E}\,\Delta v_t = g$, equations (17) and (22) each imply the relationship

$$\mathbb{E}\, sv_t = \mathbb{E}\, r_t - g\,, \tag{23}$$

analogous to an unconditional Gordon growth model.

Second, the discounting with discount factor $\rho < 1$ implies that the longer the various sources of fiscal adjustment are delayed, the larger they must ultimately be. This effect is stronger when ρ is low, as will be the case when returns on government debt are high relative to growth.

Third, the multiplication of tax growth by $1/(1-\beta)$ and of spending growth by $\beta/(1-\beta)$—which are large numbers given that β is close to one—reflects the fact that when the average primary surplus is small relative to the average levels of tax revenue and government expenditure, small percentage changes in either taxes or spending have large proportional effects on the primary surplus and hence on our measure of the fiscal position.

Fourth, when we use sv_t as a forecasting variable with the property that

$$sv_t = (1-\rho)\sum_{j=0}^{\infty}\rho^j\,\mathbb{E}_t\left[r_{t+1+j} - \frac{1}{1-\beta}\Delta\tau_{t+1+j} + \frac{\beta}{1-\beta}\Delta x_{t+1+j}\right], \tag{24}$$

as follows on taking conditional expectations of (21), we should bear in mind that it is expected *log* returns that matter.[6] As Gao and Martin

[6]Related, Narayana R. Kocherlakota (2023) shows, in models driven by a discrete-time time-homogeneous Markov process, that the limiting yield on an infinitely long-term zero-coupon bond dictates whether infinite debt rollover can in principle be sus-

(2021) note, we can write

$$\mathbb{E}_t\, r_{t+1+j} = \log \mathbb{E}_t\, R_{t+1+j} - \frac{1}{2}\,\mathrm{var}_t\, r_{t+1+j} - \sum_{n=3}^{\infty} \frac{\kappa_t^{(n)}(r_{t+1+j})}{n!}\,,$$

where $\kappa_t^{(n)}(r_{t+1+j})$ is the nth conditional cumulant of the log return. If debt returns are conditionally lognormal, then the higher cumulants $\kappa_t^{(n)}(r_{t+1+j})$ are zero for $n \geq 3$, but even in this case, low expected log returns—a potential resolution of a scenario in which fiscal health is poor, i.e. sv_t is low—may be consistent with *high* expected simple returns if returns are volatile; and the gap between the two may be wider still if log returns are right-skewed or fat-tailed.

3 Empirical results in US data

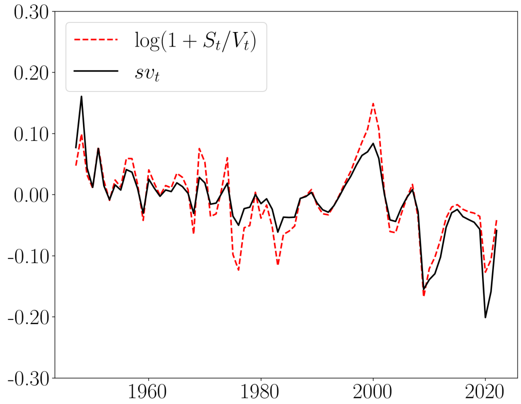
3.1 Parameter calibration

As $\mathbb{E}\log(1+S_t/V_t) = -\log\rho$, we could in principle use the sample mean of $\log(1+S_t/V_t)$ to pin down ρ, given a sufficiently long sample. In postwar data, however, the average surplus-debt ratio is negative, so this procedure would set ρ greater than one, and would bake in an "$R < G$" assumption. In order to impose a restriction that the government must pay off its debt, we therefore set ρ less than one as an a priori choice.

In our baseline analysis, we set $\rho = 0.999$ so that the implied unconditional expectation of $\log(1+S_t/V_t)$ is not too far from its sample mean in postwar data. For consistency with equation (23) and the surrounding discussion, we de-mean returns, tax growth and the fiscal position in our VAR estimation using "theory means" $\mathbb{E}\, r_t = 0.031$, $\mathbb{E}\,\Delta\tau_t = 0.030$, and $\mathbb{E}\, sv_t = 0.001$, and estimate with zero intercepts. (That is, we set $\mathbb{E}\, sv_t$ equal to $-\log\rho = 0.001$ and $\mathbb{E}\, r_t$ equal to $\mathbb{E}\, sv_t + g$, and we impose

tained. The sustainability of infinite debt rollover is therefore intimately linked to the properties of log returns, given that Ian W. R. Martin and Stephen A. Ross (2019) show, in the finite-state Markov chain setting, that the infinitely long yield is constant and equal to the unconditional expected log return on the long bond. For our purposes, the relevant quantity is the log return on the debt considered as a whole, rather than the log return on a hypothetical infinitely long zero-coupon bond, as the government does not in practice finance itself through long-horizon zero-coupon borrowing.

Figure 4: sv_t and $\log(1 + S_t/V_t)$.



$\mathbb{E}\,\Delta\tau_t = \mathbb{E}\,\Delta x_t = \mathbb{E}\,\Delta v_t = g$ throughout. As the sample means of $\Delta\tau_t$, Δx_t, and Δv_t are 0.031, 0.030, and 0.030, respectively, we set g equal to 0.030.) In the case of the log tax-GDP ratio, we de-mean using the sample mean over the period 1947–2022, $\mathbb{E}\,\tau y_{t+1} = -1.784$.

We then choose β so that sv_t optimally approximates $\log(1 + S_t/V_t)$ in a least-squares sense. That is, β is chosen to solve the problem

$$\min_{\beta} \sum_t \left(\log\left(1 + S_t/V_t\right) - \underbrace{\left[k + \frac{1-\rho}{1-\beta} \left(\tau v_t - \beta\, x v_t\right) \right]}_{sv_t} \right)^2 , \tag{25}$$

where k is given in equation (16). With $\rho = 0.999$, this procedure sets $\beta = 0.997$. The resulting time series of sv_t is shown in Figure 4, together with $\log(1 + S_t/V_t)$ which it approximates.

In Appendix A.2.2, we conduct a sensitivity analysis by setting $\rho = 0.99$, and show that this choice has little effect on our conclusions. Finally, let us emphasize that our approach allows for the possibility that there are extended periods in which the *conditional* expectation of $\log(1 + S_t/V_t)$ is negative; what we want to rule out is the possibility that the mean is negative unconditionally, for all time.

3.2 A VAR system

The approximate identity (21) relates our measure of the fiscal position, sv_t, to future debt returns, tax growth, and spending growth. It formalizes the fact that when the government is in a weak fiscal position (i.e., sv_t is low) we must subsequently have some combination of low debt returns, high tax growth, and low spending growth.

To determine which of these channels is most important empirically, we estimate a VAR for the variables r_t, $\Delta\tau_t$, sv_t, and the log tax-GDP ratio, $\tau y_t = \log T_t/Y_t$. A Johansen test, reported in Table A.4, confirms that the VAR is well specified.

By including τy_t, we ensure that the VAR takes into account the stationary relationship between tax and output. We do not include Δx_t as it is mechanically related to the first three included variables via the approximate identity (19). Indeed, we treat the identity as holding exactly, so that we can infer Δx_{t+j} from (19),

$$\frac{\beta}{1-\beta}\Delta x_{t+j} = \frac{sv_{t+j-1} - \rho sv_{t+j}}{1-\rho} - r_{t+j} + \frac{1}{1-\beta}\Delta\tau_{t+j}\,. \tag{26}$$

Similarly, we do not include GDP growth Δy_t because it is mechanically related to included variables by the identity

$$\Delta y_{t+j} = \Delta\tau_{t+j} - \Delta\tau y_{t+j}\,. \tag{27}$$

The estimated VAR is shown in the first four columns of Table 1. The more persistent variables—tax growth, the fiscal position, and tax-GDP ratio—are relatively predictable, with R^2 between 40% and 65%, and are each strongly predicted by their lags. Restricting to coefficients with t-statistics above three for the purposes of discussion, we see that a high tax-GDP ratio τy_t predicts high returns on debt and low tax growth; high tax growth predicts a high future tax-GDP ratio; and high debt returns predict a lower future tax-GDP ratio.

The last two columns of Table 1 show imputed coefficients for forecasts of GDP growth and spending growth. The explanatory power of the model for GDP growth is modest, but the tax-output ratio does enter significantly and predicts slow GDP growth.

Table 1: VAR coefficient estimates for $(r_t, \Delta\tau_t, sv_t, \tau y_t)$, US data 1947–2022.

OLS standard errors are reported in square brackets. The last columns show the imputed coefficients of GDP growth and spending growth based on the multi-variable OLS.

	r_{t+1}	$\Delta\tau_{t+1}$	sv_{t+1}	τy_{t+1}	Δy_{t+1}	Δx_{t+1}
r_t	0.188	−0.249	−0.242	−0.310	0.061	0.477
	[0.113]	[0.107]	[0.073]	[0.094]	[0.049]	[0.173]
$\Delta\tau_t$	−0.097	0.356	−0.031	0.364	−0.009	0.452
	[0.101]	[0.096]	[0.065]	[0.085]	[0.044]	[0.155]
sv_t	0.226	−0.217	0.725	−0.257	0.040	0.609
	[0.115]	[0.109]	[0.074]	[0.096]	[0.049]	[0.176]
τy_t	0.197	−0.408	0.020	0.680	−0.088	−0.469
	[0.097]	[0.093]	[0.063]	[0.082]	[0.042]	[0.150]
R^2	19.0%	41.59%	65.32%	64.67%	7.80%	25.20%

3.3 Decomposing the variance of the fiscal position

We can use the VAR to understand what fluctuations in the fiscal position, sv_t, imply about the subsequent evolution of debt returns, tax growth, and spending growth. Stacking the variables into a vector $\boldsymbol{z}_{t+1} = (r_{t+1}, \Delta\tau_{t+1}, sv_{t+1}, \tau y_{t+1})'$ and arranging the entries of Table 1 into a coefficient matrix $\boldsymbol{A}$, we have $\mathbb{E}_t\, \boldsymbol{z}_{t+j} = \boldsymbol{A}^j \boldsymbol{z}_t$. If we write $\boldsymbol{e}_n$ for a vector with one in the nth entry and zeroes elsewhere, we therefore have $\mathbb{E}_t\, r_{t+j} = \boldsymbol{e}_1' \boldsymbol{A}^j \boldsymbol{z}_t$, $\mathbb{E}_t\, \Delta\tau_{t+j} = \boldsymbol{e}_2' \boldsymbol{A}^j \boldsymbol{z}_t$, and so on.

We can use identity (20) to derive finite-horizon variance decompositions in the form

$$1 = \frac{\operatorname{cov}(sv_t, (1-\rho)\sum_{j=0}^{T-1} \rho^j\, \mathbb{E}_t\, r_{t+1+j})}{\operatorname{var}\, sv_t} + \frac{\operatorname{cov}(sv_t, -(1-\rho)\sum_{j=0}^{T-1} \rho^j\, \mathbb{E}_t\, \frac{1}{1-\beta}\Delta\tau_{t+1+j})}{\operatorname{var}\, sv_t} + \\ + \frac{\operatorname{cov}(sv_t, (1-\rho)\sum_{j=0}^{T-1} \rho^j\, \mathbb{E}_t\, \frac{\beta}{1-\beta}\Delta x_{t+1+j})}{\operatorname{var}\, sv_t} + \frac{\operatorname{cov}(sv_t, \rho^T\, \mathbb{E}_t\, sv_{t+T})}{\operatorname{var}\, sv_t}. \tag{28}$$

This decomposition can be derived by taking the time-t conditional expec-

Table 2: A variance decomposition for the fiscal position, sv_t.

Panel A: Variance decomposition for sv_t				
Horizon	return	tax	spending	future sv
1	0.0%	8.4%	16.2%	76.8%
3	0.1%	26.5%	44.0%	30.8%
10	0.0%	2.0%	95.9%	3.4%
30	0.0%	2.1%	99.2%	0.0%
∞	0.0%	2.1%	99.2%	0.0%
Panel B: Bootstrap intervals				
Horizon	return	tax	spending	future sv
1	[0.0%, 0.0%]	[1.6%, 15.1%]	[5.6%, 26.9%]	[62.7%, 90.1%]
3	[0.0%, 0.1%]	[11.0%, 41.8%]	[19.1%, 64.2%]	[4.5%, 60.8%]
10	[-0.1%, 0.1%]	[-14.9%, 13.1%]	[76.1%, 108.3%]	[-4.8%, 26.5%]
30	[-0.1%, 0.2%]	[-24.1%, 13.2%]	[88.1%, 123.0%]	[-0.2%, 2.7%]
∞	[-0.1%, 0.2%]	[-26.2%, 13.1%]	[88.2%, 127.6%]	[-0.0%, 0.0%]

tations of both sides of (20), computing covariances with sv_t and, finally, scaling by var sv_t so that the four terms on the right-hand side of (28) add up to 100%. The decomposition tells us the relative contribution of future debt returns, tax growth, spending growth, and persistent variation in the fiscal position to explaining the variability of the fiscal position at any given horizon. As we let the horizon increase, the contribution of the long-horizon future fiscal position declines to zero and we are left with a three-variable infinite-horizon variance decomposition for the fiscal position.

Table 2, Panel A reports the results of this exercise over various different horizons T. In each row of the table, the four entries correspond to the four terms on the right-hand side of (28); notice that we include a minus sign inside the second covariance term on the right-hand side of (28), so that positive entries in the column labelled "tax" indicate a negative covariance between the fiscal position and subsequent tax growth.

Panel B reports bootstrapped 95% confidence intervals for these estimates. Each bootstrap sample is computed by first drawing a new VAR coefficient matrix using the point estimates and the covariance matrix of the estimated coefficients. Using this VAR coefficient matrix, we generate the news series and do the variance decomposition. We repeat this procedure 10,000 times and report the 2.5% and 97.5% quantiles.

At short horizons, variation in sv_t is largely reflected in short-run future sv_t: if the fiscal position is weak this year, it probably will be next year too. At all horizons, there is essentially no relationship between the fiscal position and expected real returns; this contrasts with the evidence that dividend yields do forecast returns on the stock market. As a result, the fiscal position sv_t must in the long run forecast tax growth or spending growth, or both. We find that a poor fiscal position (low sv_t) is associated with high expected tax growth and low expected spending growth over the medium run. Over the long run, though, essentially all of the burden of adjustment falls on spending: a weak fiscal position forecasts spending cuts, not tax growth.

3.4 Decomposing the fiscal adjustment to tax and expenditure shocks

As our framework allows us to analyze the behavior of tax and spending separately, we can also ask whether deficits driven by shocks to taxes look different from deficits driven by shocks to spending.

We address this question by using the identity (20) to explore the implications of unexpected shocks to taxes or spending. Applying the "news operator", $\Delta\,\mathbb{E}_{t+1} = \mathbb{E}_{t+1} - \mathbb{E}_t$, to both sides of (20) and rearranging, we have

$$\underbrace{\Delta\,\mathbb{E}_{t+1}\,\tau_{t+1}}_{\text{short-run tax news}} = (1-\beta)\underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j r_{t+1+j}}_{\text{return news}} - \underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=1}^{T-1}\rho^j \Delta\tau_{t+1+j}}_{\text{long-run tax news}} + \\ + \beta\underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j \Delta x_{t+1+j}}_{\text{spending news}} + \frac{1-\beta}{1-\rho}\underbrace{\Delta\,\mathbb{E}_{t+1}\,\rho^T sv_{t+T}}_{\text{future fiscal position news}}. \tag{29}$$

This identity allows us to trace out the consequences of an unexpected shock to taxes. We refer to such a shock as short-run tax news, $N_{\text{SR tax},t+1} = \Delta\,\mathbb{E}_{t+1}\,\tau_{t+1}$. A positive short-run tax shock must be reflected in some combination of (i) news about returns, $N_{\text{return},t+1} = \Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j r_{t+1+j}$; (ii) news about declines in long-run tax growth, $N_{\text{LR tax},t+1} = \Delta\,\mathbb{E}_{t+1}\sum_{j=1}^{T-1}\rho^j \Delta\tau_{t+1+j}$; (iii) news about spending growth, $N_{\text{spending},t+1} = \Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j \Delta x_{t+1+j}$; and/or (iv) news about the future fiscal position, $N_{\text{future sv},t+1} = \Delta\,\mathbb{E}_{t+1}\,\rho^T sv_{t+T}$. This last term becomes negligible once the horizon, T, is sufficiently long.

Taking covariances of both sides of (29) with short-run tax news, $N_{\text{SR tax},t+1} = \Delta\,\mathbb{E}_{t+1}\,\tau_{t+1}$, and rearranging, we have

$$1 = \frac{\operatorname{cov}\left((1-\beta)N_{\text{return},t+1}, N_{\text{SR tax},t+1}\right)}{\operatorname{var} N_{\text{SR tax},t+1}} + \frac{\operatorname{cov}\left(-N_{\text{LR tax},t+1}, N_{\text{SR tax},t+1}\right)}{\operatorname{var} N_{\text{SR tax},t+1}}$$
$$+ \frac{\operatorname{cov}\left(\beta N_{\text{spending},t+1}, N_{\text{SR tax},t+1}\right)}{\operatorname{var} N_{\text{SR tax},t+1}} + \frac{\operatorname{cov}\left(\frac{1-\beta}{1-\rho}N_{\text{future sv},t+1}, N_{\text{SR tax},t+1}\right)}{\operatorname{var} N_{\text{SR tax},t+1}}. \tag{30}$$

Panel A of Table 3 reports the four terms on the right-hand side of the identity (30) for a range of horizons, T; the four terms in each row would add up to precisely 100% if our loglinear approximation were exact. Panel B shows bootstrapped 95% confidence intervals calculated as in the variance decomposition reported in Table 2.

In the very short run, at horizon $T = 1$, unexpected declines in tax are associated with unexpected contemporaneous *increases* in spending. This movement is in the "wrong" direction, which exacerbates the shock to the fiscal position (hence the entry greater than 100% in the rightmost column of Panel A). At longer horizons, an unexpected short-run tax cut forecasts rises in long-run tax growth, but as these do not fully offset the effect of the original tax cut it also forecasts declines in long-run spending. That said, we should note that as the confidence intervals are wide, our results are not decisive about the relative importance of tax and spending adjustment.

We can carry out a similar exercise for spending rather than taxes,

Table 3: A variance decomposition for short-run tax news.

Panel A: Variance decomposition for short-run tax news				
T	return	LR tax	spending	future sv
1	-0.1%	—	-33.9%	135.5%
3	0.1%	47.6%	-21.1%	74.9%
10	0.1%	76.8%	18.7%	5.9%
30	0.2%	83.0%	18.3%	0.0%
∞	0.1%	83.0%	18.3%	—
Panel B: Bootstrap intervals				
T	return	LR tax	spending	future sv
1	[-0.1%, -0.0%]	—	[-42.2%, -25.7%]	[127.3%, 143.8%]
3	[-0.1%, 0.2%]	[13.7%, 73.3%]	[-66.5%, 18.3%]	[26.5%, 136.5%]
10	[0.0%, 0.3%]	[48.7%, 98.5%]	[-4.9%, 48.2%]	[-20.4%, 37.3%]
30	[0.0%, 0.3%]	[55.4%, 98.2%]	[3.1%, 44.9%]	[-0.9%, 2.9%]
∞	[0.0%, 0.3%]	[54.5%, 98.2%]	[3.1%, 46.9%]	—

rewriting the identity (29) as

$$\underbrace{\Delta\,\mathbb{E}_{t+1}\,x_{t+1}}_{\text{short-run spending news}} = -\frac{1-\beta}{\beta}\underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j r_{t+1+j}}_{\text{return news}} + \frac{1}{\beta}\underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=0}^{T-1}\rho^j \Delta\tau_{t+1+j}}_{\text{tax news}} + \\ -\underbrace{\Delta\,\mathbb{E}_{t+1}\sum_{j=1}^{T-1}\rho^j \Delta x_{t+1+j}}_{\text{long-run spending news}} - \frac{1-\beta}{\beta(1-\rho)}\underbrace{\Delta\,\mathbb{E}_{t+1}\,\rho^T sv_{t+T}}_{\text{future fiscal position news}}. \tag{31}$$

We write $N_{\text{tax},t+1}$ for the tax news term that appears on the right-hand side of identity (31). This is the sum of short-run tax news and long-run tax news, as defined in (29): $N_{\text{tax},t+1} = N_{\text{SR tax},t+1} + N_{\text{LR tax},t+1}$. Similarly, we write $N_{\text{SR spending},t+1}$ for short-run spending news and $N_{\text{LR spending},t+1}$ for long-run spending news, so that $N_{\text{spending},t+1}$ as defined after identity (29) is equal to the sum $N_{\text{SR spending},t+1} + N_{\text{LR spending},t+1}$.

We can now decompose the variance of short-run spending news as the

Table 4: A variance decomposition for short-run spending news.

Panel A: Variance decomposition for short-run spending news				
T	return	tax	LR spending	future sv
1	-0.1%	-14.7%	—	116.1%
3	-0.0%	-4.1%	39.5%	65.9%
10	-0.0%	-21.8%	117.4%	5.8%
30	-0.0%	-21.3%	122.6%	0.0%
∞	-0.01%	-21.3%	122.6%	—

Panel B: Bootstrap intervals				
T	return	tax	LR spending	future sv
1	[-0.1%, -0.0%]	[-18.1%, -11.2%]	—	[112.6%, 119.5%]
3	[-0.1%, 0.1%]	[-20.2%, 10.1%]	[17.7%, 58.2%]	[41.7%, 95.9%]
10	[-0.2%, 0.2%]	[-41.0%, -8.3%]	[91.0%, 134.8%]	[-8.0%, 37.2%]
30	[-0.2%, 0.2%]	[-56.1%, -7.8%]	[109.0%, 154.4%]	[-0.2%, 3.6%]
∞	[-0.2%, 0.2%]	[-57.9%, -7.8%]	[109.1%, 159.3%]	—

sum of its covariances with news about returns, about tax growth, about long-run spending growth, and about the long-run fiscal position:

$$1 = \frac{\text{cov}\left(-\frac{1-\beta}{\beta} N_{\text{return},t+1}, N_{\text{SR spending},t+1}\right)}{\text{var } N_{\text{SR spending},t+1}} + \frac{\text{cov}\left(\frac{1}{\beta} N_{\text{tax},t+1}, N_{\text{SR spending},t+1}\right)}{\text{var } N_{\text{SR spending},t+1}} + \frac{\text{cov}\left(-N_{\text{LR spending},t+1}, N_{\text{SR spending},t+1}\right)}{\text{var } N_{\text{SR spending},t+1}} + \frac{\text{cov}\left(-\frac{1-\beta}{\beta(1-\rho)} N_{\text{future sv},t+1}, N_{\text{SR spending},t+1}\right)}{\text{var } N_{\text{SR spending},t+1}}. \tag{32}$$

Panel A of Table 4 reports the four terms on the right-hand side of the identity (32) for a range of horizons, T. Panel B reports the corresponding bootstrapped 95% confidence intervals.

In the very short run, at horizon $T = 1$, unexpected increases in spending are associated with unexpected contemporaneous *decreases* in tax. Again, this movement is in the "wrong" direction, which exacerbates the shock to the fiscal position.

At longer horizons, a deterioration in the fiscal position due to an unexpected rise in short-run spending does not forecast an increase in tax over the long run (as was the case for an unexpected decline in short-run tax) but a *decline*. As a result, a positive spending news shock forecasts a large decline in long-run spending growth that more than offsets the original increase, as indicated by the entry greater than 100% in the column labelled "LR spending."

We note, finally, that whether the fiscal position worsens due to unexpected declines in tax or unexpected increases in spending, Tables 3 and 4 show that there is almost no association with news about returns, either contemporaneously or in the long run.

3.5 The importance of the tax-GDP ratio

These results depend critically on our inclusion of the stationary variable τy_t in our VAR model. As we will now see, a three-variable VAR in $(r_t, \Delta\tau_t, sv_t)$—which does not "know" that the log tax-output ratio is stationary—would suggest that variations in sv_t are largely resolved by future tax growth.

Table A.6 reports the results of a VAR that includes returns, r_{t+1}, tax growth, $\Delta\tau_{t+1}$, and the fiscal position, sv_{t+1}. (See Table A.5 for the associated Johansen test.) The coefficient estimates are consistent with those reported in Table 1, but returns and tax growth are substantially less predictable in an R^2 sense when the tax-GDP ratio is not included.

Table A.7 reports the result of a variance decomposition of the fiscal position that applies the identity (28) using this new VAR system. At the shortest horizon, $T = 1$, the results echo our baseline findings shown in Table 2: variation in the fiscal position is largely reflected in the short-run future fiscal position. But at longer horizons the picture is very different: the VAR suggests that variation in the fiscal position is resolved more through tax than through spending adjustment. This conclusion, which differs sharply from our earlier finding that the fiscal position is entirely resolved by adjustments in spending, reflects the fact that the VAR does not take into account the stationarity of the tax-GDP ratio.

The variance decompositions of tax and spending news also look quite different if we neglect the importance of the tax-GDP ratio. Using the identities (30) and (32) to decompose the variance of unexpected shocks to taxes or to spending, as before, we find results shown in Tables A.8 and A.9. These should be compared to Tables 3 and 4.

When the tax-GDP ratio is not included in the VAR, roughly 30% of variation in short-run tax news is accounted for by adjustments in long-run tax news, and roughly 70% by long-run spending adjustments. This contrasts with our earlier findings, in Table 3, which attributed a larger fraction of the variation to adjustments in long-run taxes. The reason is that following positive tax news (an unexpected rise in taxes), the full system understands that this drives up the tax-GDP ratio; as this ratio is stationary, the full system predicts a greater role for the offsetting decrease in taxes in the long run.

The reduced system attributes the variance of short-run spending news shocks roughly equally to adjustments in tax and spending, though with wide confidence intervals that include zero both for tax and spending. Here there is a sharp contrast with our baseline results, which attribute the variance of spending news (more than) entirely to long-run spending—to the extent that the 95% confidence intervals in Tables 4 and A.9 for long-run spending at horizon $T = \infty$ do not overlap.

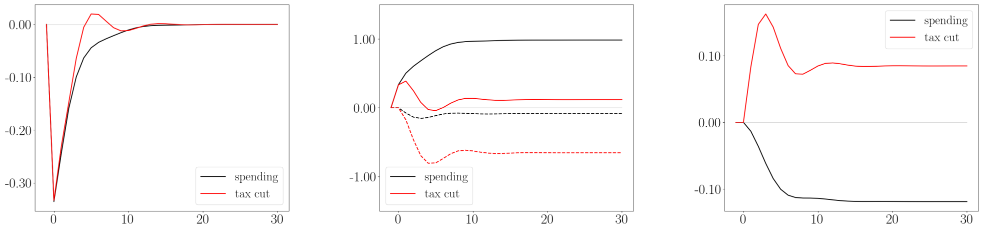
3.6 Impulse response functions

We now use the estimated VAR coefficient matrix shown in Table 1 (i.e., the full system including the tax-GDP ratio τy_t) to plot impulse response functions.

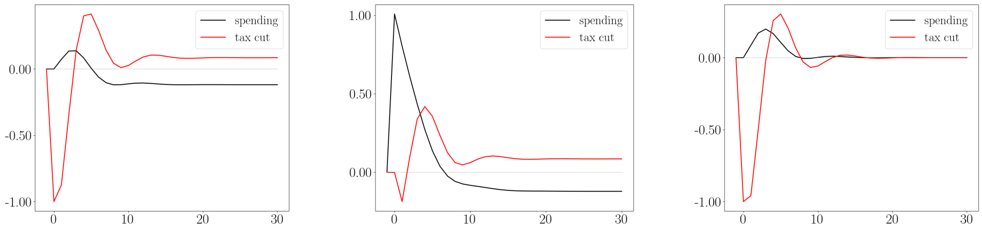
Figure 5 shows how the system evolves following a debt-financed increase in spending (black lines), or a debt-financed decline in taxes (red lines). These shocks have identical effects on surplus, so analyses that focus directly on surplus without separating into its constituent parts, tax and spending, would impose identical dynamic responses to the two shocks by construction. By contrast, our framework generates very different responses to the two shocks.

The black lines in the panels of Figure 5 indicate the response to a sudden increase in spending (x_t) at time zero that is financed by an increase

Figure 5: Debt-financed spending or tax cut, 4D system, US data.

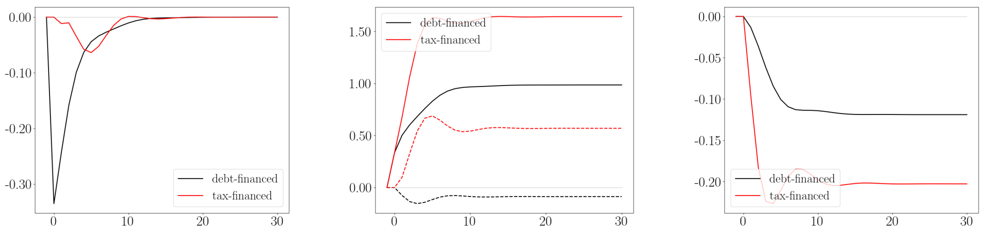


(a) sv_t (b) v_t (c) y_t

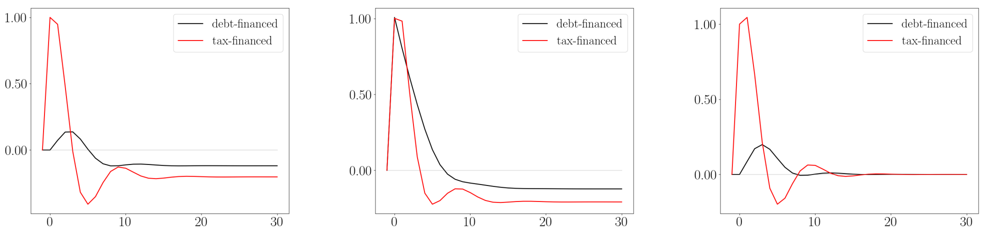


(d) τ_t (e) x_t (f) τy_t

Figure 6: Debt-financed or tax-financed spending, 4D system, US data.



(a) sv_t (b) v_t (c) y_t



(d) τ_t (e) x_t (f) τy_t

in debt (v_t) in such a way that the debt return (r_t) is held constant at $t = 0$.

At impact, the increase in debt represents a sharp deterioriation in the fiscal position (panel a). The effect is reasonably persistent, with a half-life on the order of three years. In part this is because subsequent bond returns are expected to be positive. The dotted black line in panel b shows the effect of cumulated bond returns on the value of the debt, and the solid black line shows the path of the debt value itself. Both returns and, more importantly, cumulated primary deficits swell the debt over the next 10–15 years. There is a small short-time rise in tax revenue (panel d) which reverses in the long-run; the reversion of tax-GDP, τy_t, to its mean (panel f) is achieved by a decline in output (panel c).

The red lines in Figure 5 trace out the corresponding responses to a decline in taxes (τ_t) that is financed by an increase in debt in such a way that the debt return is held constant, as before.

At impact, there is again a sharp deterioration in the fiscal position (panel a), but the recovery is faster for a decline in taxes than it was for an increase in spending. There are two contributing factors. First, the dotted red line in panel b shows that bond returns are expected to be negative following the shock, so the value of outstanding debt does not increase as much. Second, a positive shock to GDP, shown in panel c, rapidly restores the level of tax revenue—and, indeed, boosts it above its level prior to the shock—so that the primary surplus recovers more rapidly.

To emphasize the differential impact of the way in which a spending shock is financed, Figure 6 shows the response to a spending shock that is either financed by debt or by taxes, in such a way that the primary surplus is unaffected. The former case is shown as a black line and is identical to the black line in Figure 5. The latter case is shown as a red line. In frameworks that analyze surplus directly, a tax-financed spending shock must have zero impact by construction. In our framework, the red lines in Figure 6 show that the fiscal position is unchanged at impact of a tax-financed spending shock, but the fiscal position deteriorates over time (panel a). There are two reasons for this deterioration. First, debt returns turn positive (panel b). Second, output declines (panel c), and as the tax-GDP ratio is stationary, taxes and spending are lower in the long run (panels d and e).

3.7 The impact of the average surplus-debt ratio

Our analysis started from an assumption that the government debt does not have a bubble component. This implies that the unconditional average surplus-debt ratio $\mathbb{E}\log(1 + S_t/V_t) = -\log\rho$ must be positive. In our baseline analysis, we set $\rho = 0.999$ so that the implied unconditional expectation of $\log(1 + S_t/V_t)$ is not too far from its sample mean in postwar US data, but it is reasonable to ask how sensitive our results are to this assumption.

Our major conclusions are not sensitive to the choice of ρ. To demonstrate this fact, we reproduce the variance decompositions of Sections 3.3 and 3.4 for a range of values between $\rho = 0.75$ and $\rho = 0.999$. These different values of ρ represent different assumptions about the true unconditional population expectation, $\mathbb{E}\log(1+S_t/V_t)$, ranging from 28.8% when $\rho = 0.75$ to 0.1% when $\rho = 0.999$. Lower values of ρ are associated with higher values of $\mathbb{E}\log(1+S_t/V_t)$; loosely speaking, lower ρ represents higher "$R - G$", so that issuing debt is more burdensome. We emphasize that we do not consider values of ρ below about 0.96 to be reasonable: we include them merely to show how our results would change in a world in which R is much higher than G.

As ρ influences the choice of β in problem (25) and the linearized variable sv_t in our VAR, we recalculate β and reestimate the VAR for each value of ρ. As in our baseline VAR, we impose consistency on our model by de-meaning with theoretical means, as discussed in Section 3.1.

The variance decomposition for sv_t is shown in Table A.10. The first four columns report the various values of ρ together with the associated implied value of β and the approximation error in (25), and the maximum eigenvalue of the coefficient matrix. This quantity must be smaller than one in magnitude in order that the estimated system does not have explosive dynamics.

The rightmost three columns report the resulting variance decomposition (at the infinite horizon, $T = \infty$). They show the share of variation in fiscal health attributable to movements in returns, tax growth, and spending growth. As ρ declines, both returns and taxes have a somewhat greater role to play. The increasing importance of returns with lower ρ is consistent with the fact that our VAR model predicts time-varying near-term returns but predicts almost constant returns in the distant future. The total weight on forecasts of all future returns is invariant to ρ in the identity (28), but as ρ declines the identity places relatively more weight on near-term forecasts which are those that vary over time. Although the variance share of returns increases from 0.0% at $\rho = 0.999$ to 16.4% at $\rho = 0.96$ and 25.7% at $\rho = 0.75$, our main conclusion—that variability in fiscal health, sv_t, is predominantly resolved by changes in spending—survives.

Tables A.11 and A.12 repeat this sensitivity analysis for the variance decompositions for short-run tax and spending news. Again returns become somewhat more important as ρ declines, but they never play more than a minor supporting role in the fiscal adjustment to tax and spending shocks.

3.8 Local projections

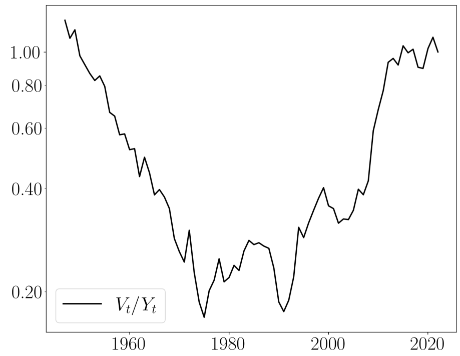
Our approximate identities (21) and (24) make no assumptions about the data-generating process. When we implement the variance decomposition (28), however, we are assuming that the VAR(1) system estimated in Table 1 accurately summarizes the data. Òscar Jordà (2005), Mikkel Plagborg-Møller and Christian K. Wolf (2021), and Dake Li, Mikkel Plagborg-Møller and Christian K. Wolf (2022) have argued for a local projection approach that imposes less structure on the underlying multivariate dynamic system. We therefore report results for an approach based on local projections at horizons of 1, 3, and 10 years in Section A.4 of the Appendix. Again, the central conclusion is unchanged.

4 Empirical results in UK data

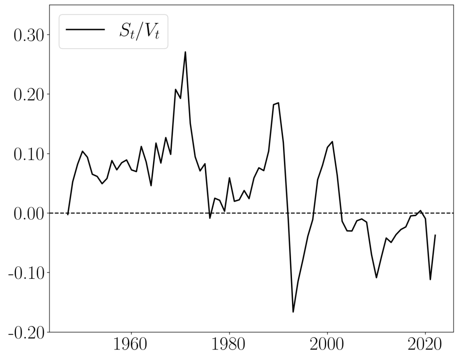
We now repeat the analysis for the UK.

To set the scene, Figure 7, Panel a shows that the debt-GDP ratio also appears nonstationary in postwar UK data. This visual impression is confirmed by an ADF test which fails to reject the presence of a unit root in the log debt-GDP ratio (with a p-value of 0.666: see Table A.14 of the Appendix). Panel b of the same figure shows the surplus-debt ratio, for which the ADF test does reject the presence of a unit root (with a p-value

Figure 7: The debt-GDP and surplus-debt ratios in postwar UK data, 1947–2022.



(a) The debt-GDP ratio. Log scale.



(b) The surplus-debt ratio. Linear scale.

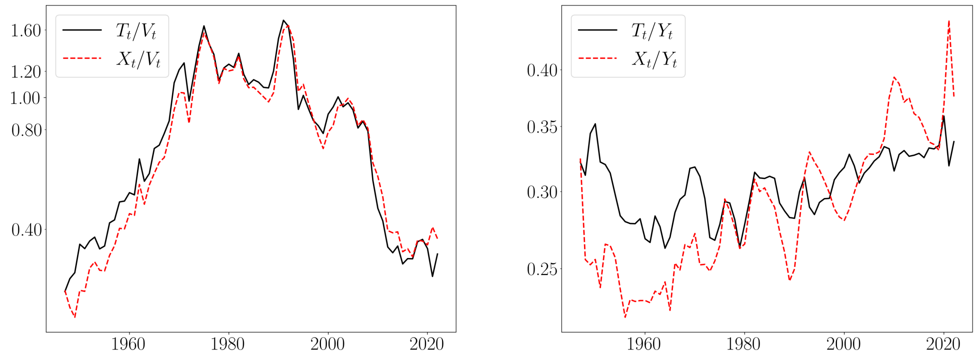
of 0.012, also reported in Table A.14).

Figure 8, Panel a plots the tax-debt ratio and spending-debt ratio for the UK on a log scale. Both ratios appear nonstationary, as confirmed by ADF tests reported in Table A.14 with p-values of 0.559 and 0.398, respectively. Figure 8, Panel b plots the tax-GDP ratio and spending-GDP ratio, also on a log scale. The spending-GDP ratio appears nonstationary, as in the US, and the ADF test fails to reject the presence of a unit root (p-value 0.747). The evidence for the tax-GDP ratio is more mixed: the ADF test gives a p-value of 0.175. Given our earlier results for the US, we proceed under the assumption that the tax-GDP ratio is stationary.

As before, we approximate the surplus-debt ratio by sv_t, as in equation (15). The sample mean of the surplus-debt ratio is positive over our sample period, so we set $-\log\rho$ equal to the sample mean of $\log(1+S_t/V_t)$, which is reported (together with other summary statistics for the UK data) in Table A.13. Having done so, we pick β to minimize (25), as we did for the US. This procedure sets $\rho = 0.967$ and $\beta = 0.952$. Figure 9 shows the resulting series sv_t, together with the surplus-debt ratio, $\log(1 + S_t/V_t)$, which it approximates.

Table 5 shows the estimated coefficients in VAR featuring r_t, $\Delta\tau_t$, sv_t, and τy_t; as we did in Section 3, we de-mean all variables by their "theory means", in this case $\mathbb{E}\, r_t = 0.084$, $\mathbb{E}\,\Delta\tau_t = 0.050$, $\mathbb{E}\,\tau y_t = -1.192$, and

Figure 8: The spending-debt and tax-debt ratios are nonstationary in the UK. The spending-GDP ratio is also nonstationary; the evidence for the tax-GDP ratio is mixed.



(a) Tax- and spending-debt ratios. Log scale.

(b) Tax- and spending-GDP ratios. Log scale.

Figure 9: sv_t and $\log(1 + S_t/V_t)$ in the UK.

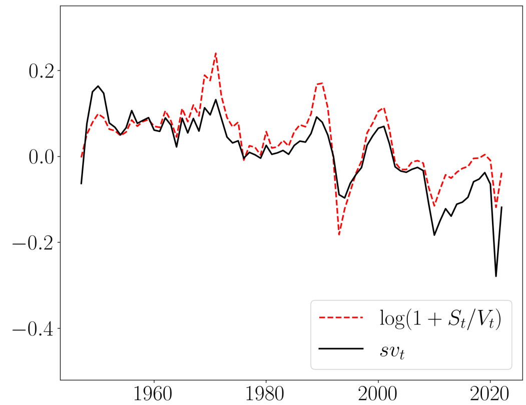


Table 5: VAR coefficient estimates. UK data, 1947–2022.

OLS standard errors are reported in square brackets. The last columns show the imputed coefficients of GDP growth and spending growth based on the multi-variable OLS.

	r_{t+1}	$\Delta\tau_{t+1}$	sv_{t+1}	τy_{t+1}	Δy_{t+1}	Δx_{t+1}
r_t	−0.172	−0.073	−0.049	−0.137	0.064	0.005
	[0.110]	[0.038]	[0.052]	[0.041]	[0.033]	[0.067]
$\Delta\tau_t$	0.907	0.449	0.000	0.347	0.102	0.432
	[0.272]	[0.093]	[0.129]	[0.101]	[0.081]	[0.167]
sv_t	0.079	−0.131	0.821	−0.163	0.032	0.159
	[0.142]	[0.048]	[0.067]	[0.052]	[0.042]	[0.087]
τy_t	0.009	−0.314	−0.164	0.835	−0.149	−0.087
	[0.152]	[0.052]	[0.072]	[0.056]	[0.045]	[0.093]
R^2	17.17%	47.11%	73.97%	79.13%	26.10%	22.50%

$\mathbb{E}\, sv_t = 0.034$. Johansen test results are reported in Tables A.15–A.17.

In the US VAR, we singled out certain coefficients whose t-statistics were above three: a high tax-GDP ratio τy_t predicted high returns on debt and low tax growth; high tax growth predicted a high future tax-GDP ratio; and high debt returns predicted a lower future tax-GDP ratio. With just one exception, the corresponding coefficients for the UK are also significant, with t-statistics above three and with the same sign. The exception is that high τy_t forecasts high returns on debt, but the estimate is not significantly different from zero. Similarly—though less surprisingly—tax growth, the fiscal position and the tax-GDP ratio were all strongly predicted by their lags in the US data, and we find that the same is true for the UK.

4.1 Variance decompositions and impulse response functions

As before, we use the estimated VAR, together with the identity (20), to derive variance decompositions for the UK surplus-debt ratio, sv_t.

The results, which are shown in Table 6, are strikingly similar to the corresponding results reported for the US in Table 2. Very little of the

Table 6: Variance decomposition for sv_t. UK data, 1947–2022.

Panel A: Variance decomposition for sv_t				
Horizon	return	tax	spending	future sv
1	0.8%	-2.7%	18.9%	84.3%
3	1.1%	-5.6%	42.3%	63.5%
10	3.5%	-54.6%	116.7%	35.8%
30	6.7%	-119.0%	206.1%	7.5%
∞	7.6%	-136.0%	229.8%	0.0%
Panel B: Bootstrap intervals				
Horizon	return	tax	spending	future sv
1	[-0.0%, 1.6%]	[-8.6%, 3.2%]	[12.2%, 29.0%]	[72.3%, 93.5%]
3	[-0.8%, 3.1%]	[-21.3%, 10.0%]	[25.9%, 65.9%]	[41.2%, 81.6%]
10	[-1.6%, 8.8%]	[-92.8%, -15.8%]	[73.9%, 161.2%]	[9.5%, 58.3%]
30	[-2.2%, 17.9%]	[-226.7%, -35.4%]	[130.9%, 298.0%]	[-0.0%, 25.7%]
∞	[-2.5%, 23.9%]	[-319.2%, -36.6%]	[136.4%, 398.9%]	[0.0%, 0.0%]

variation in surplus-debt ratio is explained by variation in returns on the debt at any horizon. At long horizons, variation in the surplus-debt ratio is resolved by adjustments in spending. Indeed, the evidence suggests that movements in tax go in the "wrong" direction: a weak fiscal position is associated, in the long run, with *reduced* tax growth. As a result, spending growth must contract even more to resolve the weak fiscal position.

Table A.18 uses the identity (30) to understand the correlates of tax news, as Table 3 did for the US, and with similar results. A positive tax news shock forecasts a decline in future tax growth, an increase in spending, and an increased return on the debt: quantitatively, our point estimates imply that about 89 per cent of the variance of tax news shocks is explained by adjustments in long-run tax growth, about 10 per cent by adjustments in long-run spending growth, and about 1 per cent by adjustments in returns.

Table A.19 uses the identity (32) to do the analogous exercise for shocks to spending, as Table 4 did for the US. At short horizons, we find that there is some compensating adjustment in both tax and spending in response to an initial spending shock; at long horizons, we find, as in the US, that the burden of adjustment falls more than entirely on long-run spending, so that, for example, a rise in short-run spending is offset by a larger decline in long-run spending.

As we did for the US, we look at impulse responses in the appendix. Figure A.1 shows the impulse responses following a debt-financed increase in spending (black lines), or a debt-financed decline in taxes (red lines). We find, as in the corresponding plots for the US (Figure 5), quite different responses to these two different ways in which the fiscal position can deteriorate. In some respects the results are similar to those we found for the US. Spending increases have a more persistent negative effect on the fiscal position, and a larger positive effect on the size of the debt, than do tax declines; and spending increases forecast long-run declines in tax, spending, and GDP. In the case of the UK, however, the point estimates suggest that debt-financed reductions in tax also forecast long-run declines in tax, spending, and GDP: we found that the opposite was true in the US.

Figure A.2 shows the responses to a given spending shock that is financed either with debt (as in Figure A.1) or with taxes. A tax-financed spending shock forecasts a deteriorating fiscal position (panel a) for the same two reasons as in the US: debt returns turn positive (panel b), and GDP declines over the medium to long run (panel c), so that both taxes and spending decline in the long run due to the stationarity of the tax-GDP ratio.

4.2 Robustness

As we did in the US, we estimate a three-variable VAR in $(r_t, \Delta\tau_t, sv_t)$ and use it to assess the importance for our findings of the stationarity of τy_t. Table A.20 reports the results of a VAR that includes returns, r_{t+1}, tax growth, $\Delta\tau_{t+1}$, and the fiscal position, sv_{t+1}. The coefficient estimates are consistent with those reported in Table 5, but tax growth is substantially less predictable when the tax-GDP ratio is not included.

Table A.21 reports a variance decomposition of the fiscal position that applies the identity (28) using this VAR. At longer horizons, the VAR attributes more of the variation in the fiscal position to tax adjustment than was the case in Table 6. This conclusion, which contrasts with our earlier finding that variation in the fiscal position is entirely resolved by adjustments in spending, reflects the fact that the VAR does not take into account the stationarity of the tax-GDP ratio. These results are consistent with our findings for the US.

Finally, we confirm that (as in the US) variance decompositions of tax and spending news also look quite different if we neglect the importance of the tax-GDP ratio. Using the identities (30) and (32) to decompose the variance of unexpected shocks to taxes or to spending, as before, we find results shown in Tables A.22 and A.23. These should be compared to the results for the full VAR system that are reported in Tables A.18 and A.19.

Once again, the results are consistent with the corresponding results for the US. The full system suggests that shocks to short-run tax news are resolved, to a considerable degree, by offsetting adjustments in long-run taxes. The reduced system, which does not appreciate the stationarity of the tax-GDP ratio, attributes more of the adjustment to spending.

Similarly, the reduced system attributes the variance of short-run spending news shocks roughly equally to adjustments in tax and spending. Again there is a sharp contrast with our baseline results, which attribute the variance of spending news (more than) entirely to long-run spending.

We explore the sensitivity of our UK conclusions to changes in ρ in Tables A.24, A.25, and A.26. As in the US, returns on government debt become somewhat more important as ρ declines, but they never play a dominant role in any variance decomposition, and our conclusions about the relative importance of tax and spending adjustment are robust to plausible variation in the level of ρ.

Finally, we report local projection estimates for the UK in section A.4 of the Appendix. The UK variance decomposition for sv_t, like the US variance decomposition, is robust to this variation in our methodology.

5 Conclusion

Conventional tests do not reject the presence of a unit root in the debt-GDP ratio in US postwar data. We have presented a framework for fiscal analysis that takes this uncomfortable fact into account by making the surplus-debt ratio—which does appear to be stationary in postwar data—the central object of interest.

Our framework allows us to analyze the contributions of taxes and spending to surplus separately, and so to draw a distinction between, say, declines in tax revenue and increases in government expenditure. There are good economic reasons to think that these two variables might not have symmetrical properties: spending might exhibit occasional spikes at times of war, for example, whereas we might expect tax revenue to evolve relatively smoothly over time. Concretely, we find that despite the non-stationarity of the surplus-GDP ratio and the expenditure-GDP ratio, the tax-GDP ratio does appear to be stationary, a fact that has important implications for our empirical findings.

We organize our empirical work by deriving a loglinear approximation to the surplus-debt ratio that summarizes the fiscal position of the government. Our key identity relates the fiscal position to future returns on government debt and to future tax and spending growth rates, just as the identities derived by Campbell and Shiller (1988) relate the dividend yield on a security to that security's future returns and dividend growth rates. A weak fiscal position must be followed by some combination of low long-run returns on government debt, high long-run tax growth, and low long-run spending growth.

We use this identity to interpret variation in the fiscal position over time in postwar data from the US and the UK. The fiscal position has almost no forecasting power for future returns; instead, it forecasts adjustment in the primary surplus. More specifically we find that in the long run the burden of adjustment falls essentially entirely on *spending*, with a weak fiscal position predicting long-run declines in spending.

These findings contrast with the results of papers that study the ratio of debt to GDP, a nonstationary ratio that has little ability to predict fiscal adjustment and mostly predicts its own future value (Jiang et al. (2021*b*)). These findings also differ sharply from those reported in the literature that carries out variance decompositions for stock market returns, following John Y. Campbell (1991), and where it is generally argued that valuation ratios have more forecasting power for returns than for cashflow growth.

The comparative importance of cashflow growth in our context reflects, in part, the simple fact that while surplus—tax minus expenditure—is a relatively small number, tax and expenditure are *large* numbers. Thus, say, a 1% change in the level of spending can have a very large proportional impact on the surplus. Meanwhile, the limited role for tax by comparison with spending reflects the fact that taxes are linked to GDP via stationarity of the tax-GDP ratio; and fiscal variables do not strongly predict GDP growth.

We also use our identity to analyze the fiscal adjustment to tax and spending shocks. Again we find that debt returns, both unexpected returns at the time the shocks occur and subsequent predictable returns, play almost no role in fiscal adjustment. Instead, mean-reverting tax and spending growth satisfy the government's intertemporal budget constraint allowing debt value to remain stable. While our framework does not allow us to say which variables are exogenous and which are endogenous, this pattern does tell us that if, as the fiscal theory of the price level asserts, debt value is endogenous, postwar governments in the US and the UK have chosen fiscal policies that avoid large predictable or unpredictable returns to debtholders, perhaps because large swings in the value of the debt are politically risky for incumbent policymakers.[7]

It is possible, perhaps even probable, that our framework would attribute a more significant role to debt returns in countries that have experienced turbulent macroeconomic crises. A priority for future research should be to apply our analysis to other countries where data are available on the market value (as opposed to the face value) of the public debt.

[7] As an example, the market reaction to unexpectedly large tax cuts in the September 2022 "mini-budget" in the United Kingdom led to the departure of the Chancellor of the Exchequer and of the Prime Minister.

References

Alesina, Alberto, Carlo Favero, and Francesco Giavazzi. 2020. *Austerity: When It Works and When It Doesn't.* Princeton University Press.

Barro, Robert J. 1979. "On the Determination of the Public Debt." *Journal of Political Economy*, 87(5): 940–971.

Berndt, Antje, Hanno Lustig, and Şevin Yeltekin. 2012. "How Does the US Government Finance Fiscal Shocks?" *American Economic Journal: Macroeconomics*, 4(1): 69–104.

Blanchard, Olivier. 2019. "Public Debt and Low Interest Rates." *American Economic Review*, 109(4): 1197–1229.

Bohn, Henning. 1991. "Budget Balance Through Revenue or Spending Adjustments?: Some Historical Evidence for the United States." *Journal of Monetary Economics*, 27(3): 333–359.

Bohn, Henning. 1998. "The Behavior of US Public Debt and Deficits." *Quarterly Journal of Economics*, 113(3): 949–963.

Bohn, Henning. 2008. "The Sustainability of Fiscal Policy in the United States." *Sustainability of Public Debt*, 15–49.

Campbell, John Y. 1991. "A Variance Decomposition for Stock Returns." *Economic Journal*, 101(405): 157–179.

Campbell, John Y., and Pierre Perron. 1991. "Pitfalls and Opportunities: What Macroeconomists Should Know about Unit Roots." *NBER Macroeconomics Annual*, 6: 141–201.

Campbell, John Y., and Robert J. Shiller. 1987. "Cointegration and Tests of Present Value Models." *Journal of Political Economy*, 95(5): 1062–1088.

Campbell, John Y., and Robert J. Shiller. 1988. "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors." *Review of Financial Studies*, 1(3): 195–228.

Campbell, John Y., Can Gao, and Ian W. R. Martin. 2023. "Online Appendix to Debt and Deficits: Fiscal Analysis with Stationary Ratios." *Working paper.*

Cochrane, John H. 2001. "Long-term Debt and Optimal Policy in the Fiscal Theory of the Price Level." *Econometrica*, 69(1): 69–116.

Cochrane, John H. 2022. "The Fiscal Roots of Inflation." *Review of Economic Dynamics*, 45: 22–40.

Cochrane, John H. 2023. *The Fiscal Theory of the Price Level.* Princeton University Press.

Ellison, Martin, and Andrew Scott. 2020. "Managing the UK National Debt 1694–2018." *American Economic Journal: Macroeconomics*, 12(3): 227–57.

Gao, Can, and Ian W. R. Martin. 2021. "Volatility, Valuation Ratios, and Bubbles: An Empirical Measure of Market Sentiment." *Journal of Finance*, 76(6): 3211–3254.

Giannitsarou, Chryssi, Andrew Scott, and Eric M. Leeper. 2006. "Inflation Implications of Rising Government Debt." In *NBER International Seminar on Macroeconomics.* 393–439.

Greenwood, Robin, Samuel G. Hanson, and Jeremy C. Stein. 2015. "A Comparative-Advantage Approach to Government Debt Maturity." *Journal of Finance*, 70(4): 1683–1722.

Hall, George J., and Thomas J. Sargent. 2021. "Debt and Taxes in Eight US Wars and Two Insurrections." *The Handbook of Historical Economics*, 825–880.

Jiang, Zhengyang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan. 2021*a*. "The US Public Debt Valuation Puzzle." *NBER Working Paper No. 26583.*

Jiang, Zhengyang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan. 2021*b*. "What Drives Variation in the US

Debt/Output Ratio? The Dogs that Didn't Bark." *NBER Working Paper No. 29351.*

Johnston, Louis, and Samuel H. Williamson. 2023. "What Was the US GDP Then?"

Jordà, Òscar. 2005. "Estimation and Inference of Impulse Responses by Local Projections." *American Economic Review*, 95(1): 161–182.

Kocherlakota, Narayana R. 2023. "Infinite Debt Rollover in Stochastic Economies." *Econometrica*, 91(5): 1629–1658.

Krishnamurthy, Arvind, and Annette Vissing-Jorgensen. 2012. "The Aggregate Demand for Treasury Debt." *Journal of Political Economy*, 120(2): 233–267.

Leeper, Eric M. 1991. "Equilibria under Active and Passive Monetary and Fiscal Policies." *Journal of Monetary Economics*, 27(1): 129–147.

Li, Dake, Mikkel Plagborg-Møller, and Christian K. Wolf. 2022. "Local Projections vs. VARs: Lessons from Thousands of DGPs." National Bureau of Economic Research.

Martin, Ian W. R., and Stephen A. Ross. 2019. "Notes on the Yield Curve." *Journal of Financial Economics*, 134: 689–702.

Mian, Atif R., Ludwig Straub, and Amir Sufi. 2022. "A Goldilocks Theory of Fiscal Deficits." *National Bureau of Economic Research.*

Plagborg-Møller, Mikkel, and Christian K. Wolf. 2021. "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89(2): 955–980.

Reinhart, Carmen M., and Kenneth S. Rogoff. 2010. "Growth in a Time of Debt." *American Economic Review*, 100(2): 573–78.

Reis, Ricardo. 2022. "Debt Revenue and the Sustainability of Public Debt." *Journal of Economic Perspectives*, 36(4): 103–24.

Sargent, Thomas, and Neil Wallace. 1981. "Some Unpleasant Monetarist Arithmetic." *Quarterly Review*, 5(3): 1–17.

Sims, Christopher A. 1994. "A Simple Model for Study of the Determination of the Price Level and the Interaction of Monetary and Fiscal Policy." *Economic Theory*, 4(3): 381–399.

Woodford, Michael. 1995. "Price-level Determinacy without Control of a Monetary Aggregate." Vol. 43, 1–46, Elsevier.

For Online Publication

Appendix of "Debt and Deficits: Fiscal Analysis with Stationary Ratios"

John Y. Campbell   Can Gao   Ian W. R. Martin[1]

Abstract

This appendix presents supplementary material and results not included in the main body of the paper.

[1]John Campbell: Harvard University and NBER, https://scholar.harvard.edu/campbell/home. Can Gao: University of St. Gallen and Swiss Finance Institute, https://sites.google.com/view/can-gao/home. Ian Martin: London School of Economics, https://personal.lse.ac.uk/martiniw.

A.1 Data Sources

A.1.1 US

In Figure 1, Panel b, the debt value is from George J. Hall and Thomas J. Sargent (2021). GDP data before 1930 is from Louis Johnston and Samuel H. Williamson (2023); after 1930, GDP data is from the FRED series FYGDP.

For tax and spending, NIPA/OMB provides annual data of total receipts, outlays and interest payments from 1947 on the FRED website. We use total receipts as T_t, and the difference between total outlays and interest payments as X_t.

According to the OMB description, the governmental receipts are taxes and other collections from the public. For example, social security taxes are counted as taxes, and therefore social security benefit payments must be treated as outlays.[8] Outlays are the measure of Government spending. They are payments that liquidate obligations.[9] The OMB budget data records outlays when obligations are paid, in the amount that is paid. The Federal Government also collects income from the public through market-oriented activities. Collections from these activities are subtracted from gross outlays, rather than added to taxes and other governmental receipts.[10] For example, premiums for healthcare benefits is counted as off-settings in outlays rather than components of the receipts. The difference between governmental receipts and outlays plus the interest payment, which is provided by OMB (we use FRED website's data), is the primary surplus or deficit.

For the market value of debt, the Dallas Fed provides the market value of marketable debt, V_t, from the 1930s.

For GDP and inflation, we use NIPA data from the FRED website.

[8]See table 17.1 in https://www.whitehouse.gov/wp-content/uploads/2023/03/ap_17_receipts_fy2024.pdf for list of the source for receipts account.

[9]See chapter *Outlays* in https://www.whitehouse.gov/wp-content/uploads/2023/03/ap_15_concepts_fy2024.pdf

[10]See table 18.1 in https://www.whitehouse.gov/wp-content/uploads/2023/03/ap_18_offsetting_fy2024.pdf for details.

A.1.2 UK

For tax and spending, we use the dataset from website of office for budget responsibility https://obr.uk/data/.

For GDP, we take the nominal GDP data from Johnston and Williamson (2023). For inflation, we take CPI inflation from BOE (pre 2017) and world bank (after 2017) from FRED website.

For the market value of debt, we use the results in Martin Ellison and Andrew Scott (2020) for 1947-2017. And we update the last 5 data points till 2022 using the same source files the authors used from David Wilkie and Andrew Cairns's webpage https://www.macs.hw.ac.uk/~andrewc/gilts/. We thank the authors of Ellison and Scott (2020) for sharing with us the public source of UK gilt price and quantity data.

A.2 Tables and figures: US

Table A.1: Summary statistics of US (NIPA) data, 1947–2022

sv_t is computed with parameters $\rho = 0.999$, $\beta = 0.997$.

Variable	mean	std	skew	kurt	median	max	min	auto-corr
r_t	0.018	0.062	-0.536	1.663	0.018	0.188	-0.180	0.223
Δx_t	0.030	0.119	-1.429	14.62	0.028	0.416	-0.628	0.191
$\Delta\tau_t$	0.031	0.068	-0.119	1.553	0.038	0.231	-0.188	0.226
Δy_t	0.030	0.024	-0.254	0.040	0.030	0.084	-0.028	0.079
Δv_t	0.030	0.091	0.186	1.166	0.015	0.288	-0.226	0.493
τv_t	-0.773	0.475	-0.393	-0.640	-0.717	0.038	-1.860	0.960
xv_t	-0.744	0.443	-0.244	-0.408	-0.689	0.010	-1.853	0.972
sv_t	-0.012	0.056	-0.653	2.772	-0.006	0.161	-0.201	0.747
S_t/V_t	-0.010	0.060	-0.029	0.288	-0.008	0.149	-0.167	0.664
$\log(1+S_t/V_t)$	-0.011	0.061	-0.232	0.359	-0.008	0.139	-0.183	0.660
T_t/Y_t	0.168	0.012	-0.28	0.385	0.169	0.198	0.132	0.664
X_t/Y_t	0.175	0.029	1.103	5.184	0.174	0.295	0.093	0.801
S_t/Y_t	-0.007	0.03	-1.549	4.654	-0.003	0.059	-0.132	0.744
V_t/Y_t	0.405	0.202	1.263	0.996	0.342	1.044	0.164	0.969
τy_t	-1.784	0.075	-0.539	0.745	-1.778	-1.622	-2.028	0.662
xy_t	-1.755	0.165	-0.272	4.002	-1.746	-1.222	-2.379	0.814
vy_t	-1.011	0.454	0.400	-0.478	-1.074	0.043	-1.808	0.976
T_t/C_t	0.307	0.029	-0.551	-0.141	0.312	0.364	0.235	0.800
X_t/C_t	0.317	0.046	0.508	4.726	0.317	0.498	0.169	0.754
S_t/C_t	-0.011	0.052	-1.383	4.268	-0.005	0.107	-0.224	0.734
V_t/C_t	0.721	0.320	1.223	1.089	0.625	1.763	0.317	0.965
τc_t	-1.187	0.100	-0.781	0.206	-1.164	-1.01	-1.449	0.809
xc_t	-1.158	0.149	-0.814	4.924	-1.149	-0.697	-1.778	0.781
vc_t	-0.414	0.409	0.384	-0.448	-0.470	0.567	-1.149	0.973

Table A.2: ADF tests (lag = AIC) for US data, 1947–2022

All tests include a free constant term. Number of lags are chosen to minimize the corresponding AIC information criterion. sv_t is computed with parameters $\rho = 0.999$, $\beta = 0.997$. The last column ("p-value*") reports the p-value of a constrained ADF test in which the time series is demeaned by the theoretical average and no constant term is included in the test.

Variable	test-stat	90%	95%	99%	p-value	p-value*
r_t	−6.98	−2.59	−2.90	−3.52	0.000	0.000
Δx_t	−10.95	−2.59	−2.90	−3.52	0.000	0.000
$\Delta \tau_t$	−5.87	−2.59	−2.90	−3.53	0.000	0.000
τv_t	−1.47	−2.59	−2.90	−3.52	0.549	—
xv_t	−2.57	−2.59	−2.90	−3.52	0.100	—
sv_t	−4.27	−2.59	−2.90	−3.52	0.001	0.000
S_t/V_t	−3.89	−2.59	−2.90	−3.52	0.002	—
$\log(1 + S_t/V_t)$	−3.92	−2.59	−2.90	−3.52	0.002	0.000
T_t/Y_t	−4.59	−2.59	−2.90	−3.52	0.000	—
X_t/Y_t	−1.71	−2.59	−2.90	−3.52	0.426	—
S_t/Y_t	−4.02	−2.59	−2.90	−3.52	0.001	—
V_t/Y_t	−0.62	−2.59	−2.90	−3.52	0.866	—
τy_t	−4.64	−2.59	−2.90	−3.52	0.000	0.000
xy_t	−2.15	−2.59	−2.90	−3.52	0.224	—
vy_t	−1.09	−2.59	−2.90	−3.52	0.720	—
T_t/C_t	−3.66	−2.59	−2.90	−3.52	0.004	—
X_t/C_t	−2.77	−2.59	−2.90	−3.52	0.062	—
S_t/C_t	−4.04	−2.59	−2.90	−3.52	0.001	—
V_t/C_t	−0.71	−2.59	−2.90	−3.52	0.843	—
τc_t	−1.69	−2.59	−2.90	−3.53	0.436	—
xc_t	−3.12	−2.59	−2.90	−3.52	0.024	—
vc_t	−1.36	−2.59	−2.90	−3.52	0.602	—

Table A.3: Johansen test for $(\tau v_t, xv_t)$, US (NIPA) data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test. 'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true.

Null	alternative	test-stat	90%	95%	99%
$r = 0$	$r \geq 1$	36.91	13.43	15.49	19.93
$r = 1$	$r \geq 2$	3.97	2.71	3.84	6.63
$r = 0$	$r \geq 1$	32.94	12.3	14.26	18.52
$r = 1$	$r \geq 2$	3.97	2.71	3.84	6.63

Table A.4: Johansen test for $(r_t, \Delta\tau_t, sv_t, \tau y_t)$, US (NIPA) data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test. 'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true. All the time series are demeaned by the theoretical average, and no constant term is included in the test.

Null	alternative	test-stat	90%	95%	99%
$r = 0$	$r \geq 1$	108.84	37.03	40.17	46.57
$r = 1$	$r \geq 2$	45.42	21.78	24.28	29.51
$r = 2$	$r \geq 3$	20.01	10.47	12.32	16.36
$r = 3$	$r \geq 4$	1.58	2.98	4.13	6.94
$r = 0$	$r \geq 1$	63.42	21.84	24.16	29.06
$r = 1$	$r \geq 2$	25.41	15.72	17.80	22.25
$r = 2$	$r \geq 3$	18.43	9.47	11.22	15.09
$r = 3$	$r \geq 4$	1.58	2.98	4.13	6.94

Table A.5: Johansen test for $(r_t, \Delta\tau_t, sv_t)$, US (NIPA) data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test.'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true. All the time series are demeaned by the theoretical average, and no constant term is included in the test.

Null	alternative	test-stat	90%	95%	99%
$r=0$	$r \geq 1$	79.27	21.78	24.28	29.51
$r=1$	$r \geq 2$	25.54	10.47	12.32	16.36
$r=2$	$r \geq 3$	4.37	2.98	4.13	6.94
$r=0$	$r \geq 1$	53.73	15.72	17.8	22.25
$r=1$	$r \geq 2$	21.17	9.47	11.22	15.09
$r=2$	$r \geq 3$	4.37	2.98	4.13	6.94

A.2.1 Results based on the 3D system

Table A.6: VAR coefficient estimates. US (NIPA) data, 1947–2022

OLS standard errors are reported in square brackets. The second last column shows the imputed coefficients spending growth based on the multi-variable OLS.

	r_{t+1}	$\Delta\tau_{t+1}$	sv_{t+1}	Δx_{t+1}
r_t	0.270	−0.419	−0.234	0.282
	[0.108]	[0.112]	[0.068]	[0.171]
$\Delta\tau_t$	−0.024	0.204	−0.024	0.277
	[0.097]	[0.101]	[0.061]	[0.154]
sv_t	0.293	−0.356	0.732	0.450
	[0.113]	[0.117]	[0.071]	[0.178]
R^2	14.25%	26.89%	65.22%	14.90%

Table A.7: Variance decomposition for sv_t based on the system $(r_t, \Delta\tau_t, sv_t)$.

Panel A: Variance decomposition				
Horizon	return	tax	spending	future sv
1	0.0%	8.6%	15.9%	76.8%
3	0.1%	36.1%	33.9%	31.2%
10	0.1%	60.9%	40.0%	0.3%
30	0.1%	61.3%	39.9%	0.0%
∞	0.0%	62.7%	38.7%	0.0%
Panel B: Bootstrap intervals				
Horizon	return	tax	spending	future sv
1	[0.0%, 0.0%]	[0.9%, 16.3%]	[5.1%, 27.5%]	[63.1%, 90.3%]
3	[0.0%, 0.1%]	[13.9%, 57.7%]	[6.7%, 56.6%]	[7.0%, 63.6%]
10	[0.0%, 0.2%]	[29.2%, 102.2%]	[-5.4%, 67.6%]	[-3.5%, 18.4%]
30	[0.0%, 0.2%]	[31.1%, 108.6%]	[-7.8%, 70.1%]	[-0.0%, 0.6%]
∞	[0.0%, 0.2%]	[31.1%, 109.1%]	[-8.0%, 70.1%]	[-0.0%, 0.0%]

Table A.8: Variance decomposition for short-run tax news based on the system $(r_t, \Delta\tau_t, sv_t)$.

Panel A: Variance decomposition for short-run tax news				
T	return	LR tax	spending	future sv
1	-0.1%	—	-1.9%	103.4%
3	-0.1%	-13.1%	52.5%	62.0%
10	0.02%	28.5%	71.8%	1.04%
30	0.02%	29.6%	71.8%	-0.0%
∞	0.02%	29.6%	71.8%	—

Panel B: Bootstrap intervals				
T	return	LR tax	spending	future sv
1	[-0.1%, -0.0%]	—	[-9.0%, 4.4%]	[97.1%, 110.5%]
3	[-0.2%, 0.1%]	[-44.1%, 12.2%]	[17.4%, 87.7%]	[27.0%, 107.5%]
10	[-0.1%, 0.2%]	[-9.9%, 61.5%]	[34.5%, 108.5%]	[-8.0%, 20.7%]
30	[-0.8%, 0.2%]	[-8.0%, 67.6%]	[33.2%, 109.3%]	[-0.0%, 062%]
∞	[-0.1%, 0.2%]	[-7.9%, 68.2%]	[33.2%, 109.3%]	—

Table A.9: Variance decomposition for short-run spending news based on the system $(r_t, \Delta\tau_t, sv_t)$.

Panel A: Variance decomposition for short-run spending news				
T	return	tax	LR spending	future sv
1	-0.0%	-0.9%	—	102.3%
3	0.0%	21.7%	25.1%	54.5%
10	0.1%	61.6%	38.9%	0.7%
30	0.1%	62.4%	38.8%	-0.0%
∞	0.1%	62.4%	38.8%	—

Panel B: Bootstrap intervals				
T	return	tax	LR spending	future sv
1	[-0.1%, -0.0%]	[-0.0%, 0.0%]	—	[99.2%, 105.7%]
3	[-0.1%, 0.1%]	[6.0%, 38.3%]	[3.3%, 43.8%]	[32.4%, 82.0%]
10	[-0.1%, 0.3%]	[29.5%, 104.6%]	[-12.2%, 67.3%]	[-5.2%, 25.3%]
30	[-0.1%, 0.3%]	[31.7%, 116.4%]	[-15.2%, 69.5%]	[-0.0%, 1.1%]
∞	[-0.1%, 0.3%]	[31.7%, 116.8%]	[-16.2%, 69.5%]	—

A.2.2 Alternative values of ρ

Table A.10: Sensitivity Analysis of ρ. US (NIPA) data, 1947–2022

Variance decomposition of sv_t.

ρ	β	approx. error	λ_{max}	return	tax	spending
0.999	0.997	0.053	0.700	0.0%	2.1%	99.2%
0.995	0.984	0.059	0.762	0.4%	2.2%	98.7%
0.990	0.970	0.079	0.835	1.5%	2.5%	97.4%
0.980	0.945	0.160	0.916	5.9%	4.0%	91.4%
0.970	0.927	0.284	0.945	11.5%	6.4%	83.5%
0.960	0.914	0.439	0.958	16.4%	8.6%	76.3%
0.950	0.904	0.614	0.966	20.6%	10.4%	70.3%
0.900	0.870	1.630	0.983	30.7%	14.3%	56.3%
0.850	0.847	2.674	0.990	31.8%	14.8%	54.8%
0.800	0.826	3.650	0.995	29.3%	14.4%	57.6%
0.750	0.805	4.541	0.998	25.7%	13.9%	61.8%
0.700	0.784	5.356	1.001	22.1%	13.3%	65.9%

Table A.11: Sensitivity Analysis of ρ. US (NIPA) data, 1947–2022

Variance decomposition of short run tax news.

ρ	β	approx. error	λ_{max}	return	LR tax	spending
0.999	0.997	0.053	0.700	0.1%	83.0%	18.3%
0.995	0.984	0.059	0.762	0.8%	82.1%	19.3%
0.990	0.970	0.079	0.835	2.1%	81.2%	19.5%
0.980	0.945	0.160	0.916	7.0%	80.6%	16.5%
0.970	0.927	0.284	0.945	12.7%	81.0%	11.0%
0.960	0.914	0.439	0.958	17.5%	81.3%	6.0%
0.950	0.904	0.614	0.966	21.4%	81.2%	2.1%
0.900	0.870	1.630	0.983	29.3%	75.9%	-3.4%
0.850	0.847	2.674	0.990	28.5%	67.6%	1.0%
0.800	0.826	3.650	0.995	24.8%	59.1%	7.8%
0.750	0.805	4.541	0.998	20.6%	51.0%	14.3%

Table A.12: Sensitivity Analysis of ρ. US (NIPA) data, 1947–2022

Variance decompositon of short run spending news.

ρ	β	approx. error	λ_{max}	return	tax	LR spending
0.999	0.997	0.053	0.700	-0.0%	-21.3%	122.6%
0.995	0.984	0.059	0.762	0.2%	-21.4%	122.6%
0.990	0.970	0.079	0.835	1.2%	-21.0%	121.2%
0.980	0.945	0.160	0.916	6.3%	-19.2%	114.2%
0.970	0.927	0.284	0.945	12.9%	-16.6%	104.9%
0.960	0.914	0.439	0.958	18.8%	-14.3%	96.3%
0.950	0.904	0.614	0.966	23.8%	-12.5%	89.0%
0.900	0.870	1.630	0.983	35.1%	-9.4%	69.8%
0.850	0.847	2.674	0.990	34.5%	-10.1%	62.8%
0.800	0.826	3.650	0.995	29.1%	-11.2%	58.7%
0.750	0.805	4.541	0.998	22.4%	-11.8%	53.9%

A.3 Tables and figures: UK

Table A.13: Summary statistics of UK data, 1947–2022

sv_t is computed with parameters $\rho = 0.967$, $\beta = 0.952$.

Variable	mean	std	skew	kurt	median	max	min	auto-corr
r_t	0.083	0.102	0.513	0.632	0.081	0.430	-0.116	-0.065
Δx_t	0.048	0.085	-2.535	15.078	0.046	0.263	-0.442	0.348
$\Delta \tau_t$	0.052	0.045	0.275	0.750	0.053	0.165	-0.055	0.410
Δy_t	0.053	0.031	-0.313	2.613	0.050	0.128	-0.064	0.591
Δv_t	0.050	0.117	0.341	0.091	0.042	0.376	-0.176	0.135
τv_t	-0.357	0.551	-0.223	-1.373	-0.203	0.537	-1.349	0.972
xv_t	-0.409	0.563	-0.252	-1.234	-0.252	0.501	-1.527	0.976
sv_t	0.010	0.083	-0.808	0.867	0.026	0.164	-0.279	0.834
S_t/V_t	0.037	0.077	-0.193	0.448	0.053	0.240	-0.182	0.814
$\log(1 + S_t/V_t)$	0.034	0.075	-0.458	0.679	0.052	0.215	-0.201	0.808
T_t/Y_t	0.304	0.023	0.021	-0.784	0.310	0.359	0.262	0.844
X_t/Y_t	0.292	0.048	0.792	0.256	0.285	0.450	0.223	0.911
S_t/Y_t	0.013	0.038	-0.752	1.884	0.015	0.097	-0.131	0.805
V_t/Y_t	0.515	0.309	0.788	-0.820	0.381	1.239	0.169	0.978
τy_t	-1.192	0.075	-0.110	-0.845	-1.173	-1.025	-1.338	0.848
xy_t	-1.245	0.160	0.463	-0.505	-1.257	-0.799	-1.501	0.921
vy_t	-0.835	0.586	0.275	-1.268	-0.965	0.214	-1.780	0.977

Table A.14: ADF tests (lag = AIC) for UK data, 1947–2022

All tests include a free constant term. Number of lags are chosen to minimize the corresponding AIC information criterion. sv_t is computed with parameters $\rho = 0.967$, $\beta = 0.952$. The last column ("p-value*") reports the p-value of a constrained ADF test in which the time series is demeaned by the theoretical average and no constant term is included in the ADF test.

Variable	test-stat	90%	95%	99%	p-value	p-value*
r_t	−9.00	−2.59	−2.90	−3.52	0.000	0.000
Δx_t	−7.59	−2.59	−2.90	−3.52	0.000	0.000
$\Delta \tau_t$	−5.69	−2.59	−2.90	−3.52	0.000	0.000
τv_t	−1.45	−2.59	−2.90	−3.52	0.559	—
xv_t	−1.76	−2.59	−2.90	−3.52	0.398	—
sv_t	−2.40	−2.59	−2.90	−3.52	0.142	0.021
S_t/V_t	−3.37	−2.59	−2.90	−3.52	0.012	—
$\log(1 + S_t/V_t)$	−3.10	−2.59	−2.90	−3.52	0.027	0.002
T_t/Y_t	−2.31	−2.59	−2.90	−3.52	0.167	—
X_t/Y_t	−1.48	−2.59	−2.90	−3.52	0.542	—
S_t/Y_t	−2.65	−2.59	−2.90	−3.52	0.082	—
V_t/Y_t	−1.79	−2.59	−2.90	−3.52	0.385	—
τy_t	−2.29	−2.59	−2.90	−3.52	0.175	0.020
xy_t	−1.02	−2.59	−2.90	−3.52	0.747	—
vy_t	−1.22	−2.59	−2.90	−3.52	0.666	—

Table A.15: Johansen test for $(\tau v_t, xv_t)$, UK data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test. 'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true.

Null	alternative	test-stat	90%	95%	99%
$r = 0$	$r \geq 1$	25.13	13.43	15.49	19.93
$r = 1$	$r \geq 2$	2.59	2.71	3.84	6.63
$r = 0$	$r \geq 1$	22.54	12.30	14.26	18.52
$r = 1$	$r \geq 2$	2.59	2.71	3.84	6.63

Table A.16: Johansen test for $(r_t, \Delta\tau_t, sv_t, \tau y_t)$, UK data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test. 'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true. All the time series are demeaned by the theoretical average, and no constant term is included in the test.

Null	alternative	test-stat	90%	95%	99%
$r = 0$	$r \geq 1$	80.71	37.03	40.17	46.57
$r = 1$	$r \geq 2$	31.20	21.78	24.28	29.51
$r = 2$	$r \geq 3$	11.00	10.47	12.32	16.36
$r = 3$	$r \geq 4$	0.60	2.98	4.13	6.94
$r = 0$	$r \geq 1$	49.51	21.84	24.16	29.06
$r = 1$	$r \geq 2$	20.20	15.72	17.8	22.25
$r = 2$	$r \geq 3$	10.40	9.47	11.22	15.09
$r = 3$	$r \geq 4$	0.60	2.98	4.13	6.94

Table A.17: Johansen test for $(r_t, \Delta\tau_t, sv_t)$, UK data 1947–2022

Top panel is the trace test, bottom panel is the eigenvalue test. 'r' is short for 'rank'. When the test statistic is higher than the x% confidence criteria, there is x% confidence that the 'alternative' is true. All the time series are demeaned by the theoretical average, and no constant term is included in the test.

Null	alternative	test-stat	90%	95%	99%
$r = 0$	$r \geq 1$	70.55	21.78	24.28	29.51
$r = 1$	$r \geq 2$	26.66	10.47	12.32	16.36
$r = 2$	$r \geq 3$	2.84	2.98	4.13	6.94
$r = 0$	$r \geq 1$	43.89	15.72	17.80	22.25
$r = 1$	$r \geq 2$	23.82	9.47	11.22	15.09
$r = 2$	$r \geq 3$	2.84	2.98	4.13	6.94

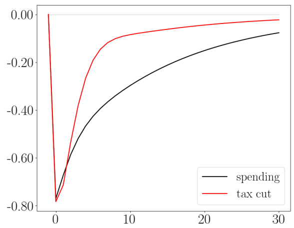
Table A.18: Variance decompositions for short-run tax news. UK data, 1947–2022.

Panel A: Short-run tax news				
T	return	LR tax	spending	future sv
1	0.4%	—	-5.1%	106.4%
3	4.9%	20.6%	46.6%	29.7%
10	1.3%	93.8%	4.7%	1.9%
30	1.4%	90.8%	9.1%	0.4%
∞	1.5%	89.8%	10.4%	—
Panel B: Bootstrap intervals				
T	return	LR tax	spending	future sv
1	[-0.3%, 1.2%]	—	[-12.9%, 2.7%]	[98.2%, 114.8%]
3	[1.5%, 8.6%]	[-13.4%, 46.2%]	[2.5%, 89.1%]	[-8.8%, 82.7%]
10	[-4.4%, 5.8%]	[55.8%, 141.4%]	[-46.6%, 39.7%]	[-19.4%, 31.7%]
30	[-5.8%, 7.4%]	[23.2%, 169.7%]	[-60.8%, 67.7%]	[-6.5%, 8.2%]
∞	[-6.9%, 8.2%]	[6.2%, 185.1%]	[-78.5%, 88.4%]	—

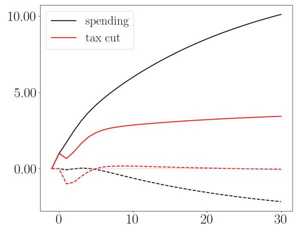
Table A.19: Variance decompositions for short-run spending news. UK data, 1947–2022.

Panel A: Short-run spending news				
T	return	tax	LR spending	future sv
1	-0.2%	-1.9%	—	102.6%
3	-0.1%	17.9%	17.7%	64.8%
10	1.0%	-6.7%	72.7%	33.4%
30	4.0%	-66.7%	156.1%	6.9%
∞	4.8%	-82.5%	178.1%	—
Panel B: Bootstrap intervals				
T	return	tax	LR spending	future sv
1	[-0.6%, 0.3%]	[-5.0%, 1.1%]	—	[99.5%, 105.6%]
3	[-1.8%, 1.6%]	[4.1%, 32.2%]	[0.4%, 37.4%]	[44.1%, 83.5%]
10	[-3.7%, 5.6%]	[-39.7%, 31.1%]	[29.8%, 108.8%]	[11.2%, 58.1%]
30	[-4.1%, 13.1%]	[-156.2%, 10.5%]	[84.1%, 231.3%]	[0.0%, 24.2%]
∞	[-4.3%, 18.7%]	[-237.7%, 7.9%]	[92.7%, 323.8%]	—

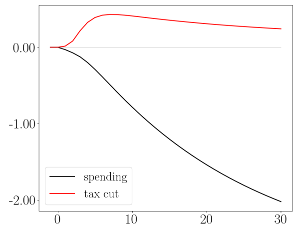
Figure A.1: Debt-financed spending or tax cut, 4D system, UK data.



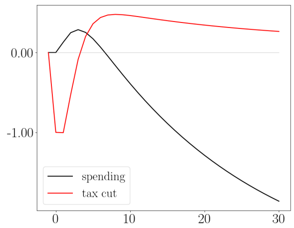
(a) sv_t



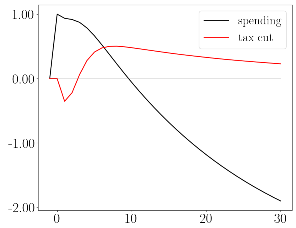
(b) v_t



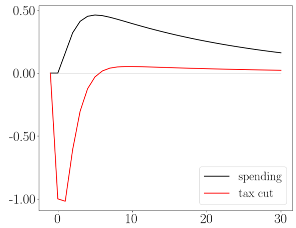
(c) y_t



(d) τ_t

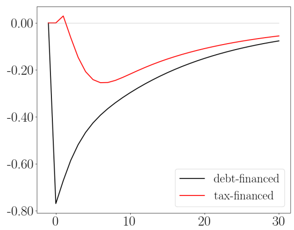


(e) x_t

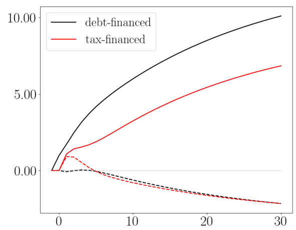


(f) τy_t

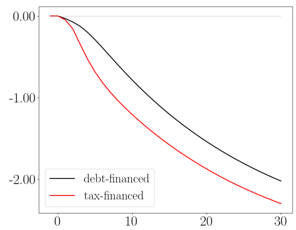
Figure A.2: Debt-financed or tax-financed spending, 4D system, UK data.



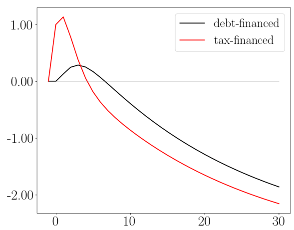
(a) sv_t



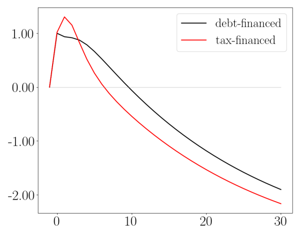
(b) v_t



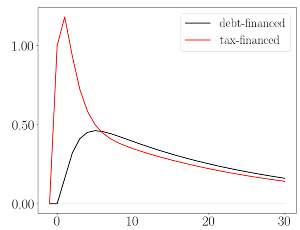
(c) y_t



(d) τ_t



(e) x_t



(f) τy_t

A.3.1 Results based on the 3D system

Table A.20: VAR coefficient estimate of the system $(r_t, \Delta\tau_t, sv_t)$. UK data, 1947–2022.

OLS standard errors are reported in square brackets. The second last column shows the imputed coefficients spending growth based on the multi-variable OLS.

	r_{t+1}	$\Delta\tau_{t+1}$	sv_{t+1}	Δx_{t+1}
r_t	−0.172	−0.079	−0.051	0.004
	[0.11]	[0.046]	[0.054]	[0.067]
$\Delta\tau_t$	0.906	0.495	0.025	0.445
	[0.271]	[0.113]	[0.133]	[0.166]
sv_t	0.077	−0.060	0.858	0.178
	[0.137]	[0.057]	[0.068]	[0.084]
R^2	17.16%	21.39%	72.04%	21.60%

Table A.21: Variance decomposition for sv_t based on the system $(r_t, \Delta\tau_t, sv_t)$, UK data 1947–2022.

Panel A: Variance decomposition for sv_t				
Horizon	return	tax	spending	future sv
1	0.8%	-1.8%	18.7%	83.6%
3	0.9%	7.0%	37.0%	56.4%
10	0.6%	28.1%	57.5%	15.2%
30	0.4%	35.9%	64.7%	0.4%
∞	0.4%	36.1%	64.9%	0.0%
Panel B: Bootstrap intervals				
Horizon	return	tax	spending	future sv
1	[-0.1%, 1.6%]	[-8.8%, 5.9%]	[9.3%, 28.7%]	[71.1%, 94.6%]
3	[-1.1%, 2.9%]	[-18.5%, 32.8%]	[11.5%, 59.7%]	[32.6%, 82.0%]
10	[-3.8%, 5.1%]	[-37.9%, 88.1%]	[-7.2%, 106.3%]	[0.4%, 55.1%]
30	[-5.5%, 7.1%]	[-57.6%, 120.7%]	[-20.8%, 144.4%]	[-0.0%, 17.9%]
∞	[-5.8%, 7.3%]	[-64.4%, 129.0%]	[-24.6%, 160.8%]	[-0.0%, 0.0%]

Table A.22: Variance decompositions for short-run tax news based on the system $(r_t, \Delta\tau_t, sv_t)$. UK data, 1947–2022.

Panel A: Short-run tax news				
T	return	LR tax	spending	future sv
1	0.1%	—	-34.8%	135.6%
3	2.9%	-7.9%	17.3%	88.6%
10	3.6%	57.7%	42.0%	-2.5%
30	3.6%	55.8%	41.5%	0.0%
∞	3.6%	55.8%	41.5%	—
Panel B: Bootstrap intervals				
T	return	LR tax	spending	future sv
1	[-0.3%, 0.5%]	—	[-41.1%, -28.1%]	[128.9%, 142.0%]
3	[0.2%, 5.5%]	[-40.3%, 18.0%]	[-20.5%, 51.7%]	[53.0%, 135.9%]
10	[-0.1%, 8.6%]	[13.3%, 103.8%]	[-7.0%, 86.7%]	[-11.6%, 9.4%]
30	[0.0%, 8.7%]	[13.2%, 102.1%]	[-8.1%, 86.0%]	[-0.1%, 0.1%]
∞	[0.0%, 8.7%]	[13.2%, 102.1%]	[-8.1%, 86.0%]	—

Table A.23: Variance decompositions for short-run spending news based on the system $(r_t, \Delta\tau_t, sv_t)$. UK data, 1947–2022.

Panel A: Short-run spending news				
T	return	tax	LR spending	future sv
1	2.1%	-15.3%	—	114.8%
3	4.2%	27.7%	24.4%	45.3%
10	4.1%	65.3%	33.4%	-1.1%
30	4.1%	64.3%	33.2%	0.0%
∞	4.1%	64.3%	33.2%	—

Panel B: Bootstrap intervals				
T	return	tax	LR spending	future sv
1	[1.9%, 2.4%]	[-18.3%, -12.1%]	—	[111.6%, 117.8%]
3	[2.9%, 5.9%]	[9.9%, 44.5%]	[0.9%, 44.6%]	[21.8%, 74.1%]
10	[1.6%, 8.5%]	[36.6%, 108.6%]	[-14.4%, 61.5%]	[-4.6%, 6.7%]
30	[1.7%, 8.6%]	[37.2%, 109.5%]	[-15.2%, 62.0%]	[-0.1%, 0.1%]
∞	[1.7%, 8.6%]	[37.2%, 109.5%]	[-15.2%, 62.0%]	—

A.3.2 Alternative values of ρ

Table A.24: Sensitivity Analysis of ρ. UK data, 1947–2022

Variance decomposition of sv_t.

ρ	β	approx. error	λ_{max}	return	tax	spending
0.999	0.998	0.131	0.94	0.6%	-176.9%	277.6%
0.995	0.991	0.122	0.941	2.2%	-173.3%	272.4%
0.990	0.982	0.125	0.943	3.6%	-168.9%	266.6%
0.980	0.968	0.137	0.948	5.2%	-156.1%	252.2%
0.970	0.956	0.158	0.954	6.8%	-140.9%	235.5%
0.960	0.945	0.184	0.96	9.1%	-124.6%	216.8%
0.950	0.935	0.215	0.965	12.0%	-108.8%	198.2%
0.900	0.894	0.414	0.981	30.2%	-57.5%	128.6%
0.850	0.858	0.684	0.99	44.5%	-37.6%	94.4%
0.800	0.824	1.031	0.995	51.7%	-29.2%	78.8%
0.750	0.790	1.468	0.997	53.7%	-24.4%	72.1%
0.700	0.756	2.002	0.998	52.4%	-20.9%	69.8%
0.600	0.684	3.390	0.999	45.3%	-15.6%	71.6%

Table A.25: Sensitivity Analysis of ρ. UK data, 1947–2022

Variance decomposition of short run tax news.

ρ	β	approx. error	λ_{max}	return	LR tax	spending
0.999	0.998	0.131	0.940	0.1%	106.8%	-5.5%
0.995	0.991	0.122	0.941	0.3%	104.0%	-2.7%
0.990	0.982	0.125	0.943	0.4%	100.5%	0.9%
0.980	0.968	0.137	0.948	0.5%	94.8%	6.5%
0.970	0.956	0.158	0.954	1.1%	90.7%	10.0%
0.960	0.945	0.184	0.960	2.3%	87.9%	11.4%
0.950	0.935	0.215	0.965	4.1%	85.9%	11.4%
0.900	0.894	0.414	0.981	19.6%	78.0%	2.8%
0.850	0.858	0.684	0.990	38.3%	66.8%	-5.7%
0.800	0.824	1.031	0.995	53.5%	53.8%	-9.5%
0.750	0.790	1.468	0.997	63.6%	41.7%	-9.3%
0.700	0.756	2.002	0.998	68.9%	31.5%	-6.9%
0.600	0.684	3.390	0.999	70.5%	16.9%	1.0%

Table A.26: Sensitivity Analysis of ρ. UK data, 1947–2022

Variance decomposition of short run spending news.

ρ	β	approx. error	λ_{max}	return	tax	LR spending
0.999	0.998	0.131	0.940	0.5%	-125.0%	225.8%
0.995	0.991	0.122	0.941	1.6%	-120.6%	220.2%
0.990	0.982	0.125	0.943	2.5%	-115.3%	213.9%
0.980	0.968	0.137	0.948	3.3%	-102.1%	199.7%
0.970	0.956	0.158	0.954	4.3%	-87.4%	183.6%
0.960	0.945	0.184	0.960	6.0%	-71.9%	166.0%
0.950	0.935	0.215	0.965	8.6%	-57.2%	148.4%
0.900	0.894	0.414	0.981	26.6%	-11.7%	82.3%
0.850	0.858	0.684	0.990	42.4%	2.2%	48.0%
0.800	0.824	1.031	0.995	49.8%	4.3%	30.0%
0.750	0.790	1.468	0.997	48.1%	2.9%	19.8%
0.700	0.756	2.002	0.998	39.4%	1.1%	13.4%
0.600	0.684	3.390	0.999	18.0%	-0.7%	6.0%

A.4 Local projections

Consider the following regressions for a given value of T:

$$(1-\rho)\sum_{j=0}^{T-1}\rho^j r_{t+1+j} = \alpha_{r,T} + \beta_{r,T} sv_t + \varepsilon_{r,t+T}, \tag{A.1}$$

$$-(1-\rho)\sum_{j=0}^{T-1}\rho^j \frac{1}{1-\beta}\Delta\tau_{t+1+j} = \alpha_{\tau,T} + \beta_{\tau,T} sv_t + \varepsilon_{\tau,t+T}, \tag{A.2}$$

$$(1-\rho)\sum_{j=0}^{T-1}\rho^j \frac{\beta}{1-\beta}\Delta x_{t+1+j} = \alpha_{x,T} + \beta_{x,T} sv_t + \varepsilon_{x,t+T} sv_{t+T}, \tag{A.3}$$

$$\rho^T sv_{t+T} = \alpha_{sv,T} + \beta_{sv,T} sv_t + \varepsilon_{sv,t+T}. \tag{A.4}$$

For any T, equation (28) implies that the coefficients satisfy the restriction

$$\beta_{r,T} + \beta_{\tau,T} + \beta_{x,T} + \beta_{sv,T} = 1\,. \tag{A.5}$$

Table A.27 reports estimates of the betas from the above regressions together with bootstrapped standard errors.

Table A.27: Local projections, US data 1947-2022.

Standard errors, in square brackets, are computed as follows: 1) We simulate the time series of $(r_t, \Delta\tau_t, sv_t, \tau y_t)$ using the VAR; 2) we impute the time series of Δx_t that satisfies our one-period equation (26); 3) we compute local projections at horizons of 1, 3, and 10 years; 4) we generate a bootstrapped sample by repeating this procedure 10,000 times and report its standard deviation.

Horizon	return	tax	spending	future sv
1	0.0%	8.5%	17.8%	73.7%
	[0.0%]	[5.1%]	[6.3%]	[7.4%]
3	0.0%	16.5%	39.7%	43.8%
	[0.0%]	[14.5%]	[11.0%]	[16.1%]
10	-0.0%	15.9%	61.2%	22.9%
	[0.1%]	[16.9%]	[20.5%]	[17.7%]

Table A.28: Local projections, UK data 1947-2022.

Standard errors, in square brackets, are computed as follows: 1) We simulate the time series of $(r_t, \Delta\tau_t, sv_t, \tau y_t)$ using the VAR; 2) we impute the time series of Δx_t that satisfies our one-period equation (26); 3) we compute local projections at horizons of 1, 3, and 10 years; 4) we generate a bootstrapped sample by repeating this procedure 10,000 times and report its standard deviation.

Horizon	return	tax	spending	future sv
1	0.9%	-3.1%	20.0%	81.6%
	[0.5%]	[6.9%]	[7.3%]	[9.0%]
3	1.4%	-11.2%	48.1%	59.8%
	[1.3%]	[20.3%]	[15.2%]	[17.5%]
10	4.6%	-62.6%	110.2%	43.0%
	[3.4%]	[44.4%]	[38.0%]	[18.7%]