NBER WORKING PAPER SERIES

WHAT DRIVES VARIATION IN THE U.S. DEBT/OUTPUT RATIO? THE DOGS THAT DIDN'T BARK

Zhengyang Jiang
Hanno Lustig
Stijn Van Nieuwerburgh
Mindy Z. Xiaolan

Working Paper 29351
http://www.nber.org/papers/w29351

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2021, Revised May 2022

We thank Bernard Dumas, Mac Croce, Francois Gourio, James Hamilton, Zhiguo He, Bryan Kelly, Ralph Koijen, Howard Kung, Stavros Panageas, Wenyi Shen, Alan Timmerman, and participants at the ASU Sonoran Winter Finance Conference and the Adam Smith Conference for detailed comments, as well as seminar participants at the NY Fed, the Fed Board of Governors, LBS, UCSD, UT Austin, Adam Smith, 2022 GSU CEAR-Finance Conference. We are grateful to Thomas Sargent and George Hall for making their historical data available. We gratefully acknowledge financial support from NSF award 2049260. This paper was first circulated as part of another paper 'Quantifying U.S. Treasury Investor Optimism'. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2021 by Zhengyang Jiang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

What Drives Variation in the U.S. Debt/Output Ratio? The Dogs that Didn't Bark
Zhengyang Jiang, Hanno Lustig, Stijn Van Nieuwerburgh, and Mindy Z. Xiaolan
NBER Working Paper No. 29351
October 2021, Revised May 2022
JEL No. E62,G12

ABSTRACT

Higher U.S. government debt/output ratios do not reliably predict higher future surpluses or lower real returns on Treasurys. Neither future cash flows nor discount rates account for the variation in the current debt/output ratio. The market valuation of Treasurys is surprisingly insensitive to the macro fundamentals. Instead, the future debt/output ratio accounts for most of the variation. Systematic surplus forecast errors may help to account for part of these findings. Since the start of the GFC, surplus projections have anticipated a large fiscal correction that failed to materialize.

Zhengyang Jiang
Kellogg School of Management
Northwestern University
2211 Campus Drive
Evanston, IL 60208
and NBER
zhengyang.jiang@kellogg.northwestern.edu

Hanno Lustig
Stanford Graduate School of Business
655 Knight Way
Stanford, CA 94305
and NBER
hlustig@stanford.edu

Stijn Van Nieuwerburgh
Columbia University
Graduate school of Business
Uris Hall, office 809
3022 Broadway
New York, NY 10027
and NBER
svnieuwe@gsb.columbia.edu

Mindy Z. Xiaolan
University of Texas at Austin
2110 Speedway B6600
Austin, TX 78703
mindy.xiaolan@mccombs.utexas.edu

A data appendix is available at http://www.nber.org/data-appendix/w29351

There is an ongoing debate in the U.S. and other advanced economies about fiscal sustainability (see, e.g., Croce, Nguyen, and Schmid, 2012; Chernov, Schmid, and Schneider, 2020; Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2019; Brunnermeier, Merkel, and Sannikov, 2020; Reis, 2021; Mian, Straub, and Sufi, 2021a). Some economists have pointed out that lower inflation- and growth-adjusted returns on government debt can rationalize the recent increase in the U.S. debt/output ratio (Blanchard, 2019; Furman and Summers, 2020; Cochrane, 2021a), while others have argued that the run-up in debt could reflect higher future surpluses (Bohn, 1998; Cochrane, 2020).

To analyze the empirical validity of these claims, we approach the valuation of the U.S. federal government debt using standard tools from asset pricing. We apply a Campbell and Shiller (1988)-style decomposition to the market value of the U.S. federal government's debt divided by the U.S. gross domestic product. A higher debt/output ratio has to be followed by lower real growth-adjusted returns, higher surplus/output ratios, or, absent adjustment in fundamentals, higher future debt/output ratios.

At horizons of up to 10 years, in the post-war U.S. sample, our variance decomposition attributes none of the variation in the current debt/output ratio to variation in future surpluses or future returns. Instead, the future debt/output ratio 10 years from now accounts for 92% of the variation in today's debt/output ratio. Importantly, we cannot definitively rule out that the market prices in a large, fiscal corrections that occurs outside of our sample with a low probability. To guard against this, we extend our sample to the Civil War, and we extend the horizon from 10 years to 25 years in a longer sample. The results are unchanged.

First, higher debt/output does not predict lower real growth-adjusted returns on U.S. government debt in sample. Discount rates do not explain variation in the U.S. debt/output ratio. Second, higher debt/output ratio does not forecast larger surpluses. Cash flows do not explain variation in the debt/output ratio either. Instead, we cannot rule out that the U.S. debt/output ratio follows a unit-root process. Fundamentals fail to push the debt/output ratio back to the mean, which violates fiscal sustainability.[1] The confidence intervals for the discount rate component are narrow around zero, even at longer horizons, allowing us to rule out a significant role for discount rates at any horizon, but the confidence intervals for the cash flow component are wider at longer horizons, implying that we cannot definitively rule out some role for cash flows.

We implement a forward-looking decomposition of the government debt/output ratio which decomposes the valuation into components due to future returns and future surpluses (see Berndt, Lustig, and Yeltekin, 2012; Cochrane, 2021a).[2] In earlier work, Hall and Sargent (2011); Berndt and

[1]The small role for fundamentals (cash flows and discount rates) is implied by the persistence of the debt/output ratio. The first-order auto-correlation of the debt/output ratio is 0.99 at annual frequencies. As we get closer to a unit root, the fraction of the variance accounted for by fundamentals decreases to zero % at all horizons.

[2]In earlier work, Gourinchas and Rey (2007) decompose the variation in the U.S. net foreign asset position using the same approach.

Yeltekin (2015); Hall and Sargent (2021) implement an ex post, backward-looking decomposition of the variation in the U.S. federal debt/GDP ratio, imputing changes in the ratio to contemporaneously realized inflation, growth and returns. Our ex ante approach provides a different, complementary perspective.

Our findings are also different from those in the literature.[3] For example, Bohn (1998), studying a shorter sample that ends in the mid-90s, finds evidence that the primary surplus increases when the debt/output ratio is high, while Cochrane (2021a,b) finds evidence that the debt/output ratio predicts nominal returns on the government debt portfolio. We replicate these findings, but we find no evidence that the debt/output ratio predicts real growth-adjusted returns or surpluses. Bond prices today are not sufficiently responsive to news about future macro fundamentals, returns and surpluses.

We reach a different conclusion because of a statistical inference challenge, well known in asset pricing, that is pervasive in the fiscal sustainability literature. There is a small-sample bias in the slope coefficients of the return and surplus predictability regressions due to (i) the high persistence of the debt/output ratio, the predictor, and (ii), the high correlation between the innovations to the predictor and the predicted variables (Stambaugh, 1999). For example, an increase in bond risk premia will tend to lower realized returns, the dependent variable, and lower the debt/output ratio, the predictor. As a result, the expectation of the residual in the return predictability regression conditional on the debt/output ratio this year and last year cannot be zero. Thus, the classical OLS assumptions are violated. The OLS estimator of the slope coefficient will tend to be too high as a result. Because of this mechanical link between realized returns and the predictor, OLS will find evidence of return predictability, even where there is none. The same small-sample bias also affects the surplus predictability regressions. Governments must issue more debt when they run large deficits. The negative correlation between the innovation in the debt/output ratio and the surplus similarly produces spurious evidence of surplus predictability.

We cannot rule out that the debt/output ratio is subject to permanent innovations, which is not consistent with fiscal sustainability. As a check, we simulate data from a model in which the debt/output ratio has a unit root and there is no role for fundamentals. This model exactly replicates our empirical findings. We find spurious evidence of surplus and return predictability using unadjusted OLS slope estimates, but bias-corrected estimates that recover the true values indicating no predictability.[4]

However, we cannot definitively conclude from this evidence that U.S. is not on a fiscally sus-

[3] There is a large literature in macro-economics that addresses the question of government debt sustainability (Hamilton and Flavin, 1986; Trehan and Walsh, 1988, 1991; Bohn, 1998, 2007; Croce, Nguyen, and Schmid, 2012; Croce, Nguyen, and Raymond, 2021; Mian, Straub, and Sufi, 2021b), starting with the seminal work by Hansen and Sargent (1980); Hansen, Roberds, and Sargent (1991) (see Sargent, 2012, for a comprehensive review). This literature largely sidesteps the issue of discount rate variation. Instead, most of the models assume constant discount factors.

[4] Croce, Nguyen, and Schmid (2012); Croce, Nguyen, and Raymond (2021) warned about the high persistence of the U.S. debt/output ratio. They explore the effects on fiscal uncertainty in an equilibrium model.

tainable path. There is a long literature on the low power of unit root tests in short historical samples against close-to-unit-root alternatives (Schwert, 1987; Lo and MacKinlay, 1989). Even in the longer sample, we lack the power to completely rule out that the debt/output ratio is mean-reverting, albeit slowly. If it is, given the higher short-run volatility of Treasury cash flows and the lower short-run volatility of Treasury returns, mean reversion in returns on the Treasury portfolio without predictability of surpluses, is not a likely scenario, but mean reversion in surpluses is.[5] Our findings are indeed consistent with this view. While the statistical evidence against the discount rate channel is strong at all horizons, we cannot rule out some role for the cash flow channel at long horizons, even though the bias-adjusted point estimates are close to zero, simply because the confidence intervals are wider.[6]

Structural breaks are one source of permanent shocks. There may have been a structural break in the relation between bond returns, surpluses and the debt/output ratio.[7] We find evidence of a structural break in the debt/output ratio around 2007. Bond investors may have revised their estimate of the long-run debt/output ratio upwards around the start of the GFC. If we allow for a permanent 0.78 increase (in log scale) in the debt/output ratio in 2007, and use the debt/ouput ratio adjusted for a different subsample mean before and after 2007, we find stronger evidence for surplus, but not return predictability. Fundamentals now account for about 50% of the variation in the transitory component of the debt/output ratio at the 10-year horizon. This structural break reduces the debt/output ratio's persistence and creates more room for fundamentals in explaining the–now more transitory nature of the—variation in the debt/output ratio. Of course, this analysis leaves the large, permanent increase in the debt/output ratio (as well as its timing) unexplained. Similar evidence obtains when we use the domestically and privately held debt/output ratio as the predictor, instead of the transitory component of the debt/output ratio, suggesting that inelastic demand by the Fed and the Treasury may be a contributing factor.

Mis-pricing is a candidate explanation for these structural breaks. An econometrician with access to the U.S. sample does not predict higher surpluses or lower returns when the debt/output ratio rises, but the bond investor may. If investors systematically over-predict surpluses and under-predict returns as the debt/output ratio increases, this could impute a unit root the debt/output ratio under the actual measure, while under the subjective measure, the debt/output ratio is stationary. Our paper connects to the growing literature on the role of subjective beliefs in asset pricing. de La O and Myers (2021) implement a Campbell-Shiller decomposition under the subjective measure of U.S.stock valuations. Under the subjective measure, they impute a large share

[5]In this sense, Treasurys as an asset class are quite different from stocks. For stocks, the cash flow volatility is much smaller than discount rate volatility.

[6]In fact, when use the transitory component of the debt/output ratio as the predictor, we find a statistically significant role for surplus predictability.

[7]In stock markets, there also is evidence of structural breaks in the relation between returns and the dividend yield (see Lettau and Van Nieuwerburgh, 2008; Smith and Timmermann, 2020). This may help to account for the poor out-of-sample performance of these predictors (Welch and Goyal, 2008).

of variation in the price/dividend ratio and price/earnings ratio on U.S./ stocks to the cash flow component. Renxuan (2020) finds that investors in the stock market all overstate the importance of expected cash flow growth fluctuations. Mistakes in interest rate and growth forecasts also contribute to our findings (see Piazzesi, Salomao, and Schneider, 2015; Cieslak, 2018, for evidence on the gap between statistical bond risk premia measured by econometricians and subjective bond risk premia expected by investors).

We show evidence from CBO (Congressional Budget Office) projections that the permanent break in the debt/output ratio at the start of the GFC (Great Financial Crisis) has been partly fueled by persistently biased forecasts of surpluses, imputing more mean-reversion to the predicted debt/output ratio under the subjective than the actual measure.[8] Throughout and after the GFC, the CBO consistently overestimated future surpluses and underestimated the government's effective cost of funding. Even by 2010, the CBO projections of 10-year surpluses were still overshooting realized surpluses by roughly 6.8% of GDP per annum. Under the investors' subjective measure, more of the variation in the debt/output ratio can be attributed to perceived future surpluses.

Our paper builds on the growing literature that examines the predictability of returns in asset markets. There is substantial evidence that stock returns are predictable (see Koijen and Nieuwerburgh, 2011, for a survey of the literature on stock return predictability). When the valuations are high in stock markets, they revert back to the mean mostly through lower subsequent returns (Campbell and Thompson, 2007; Cochrane, 2008), partly through higher dividend growth (Binsbergen and Koijen, 2010; Golez and Koudijs, 2018). Overall, the discount rates on stocks are remarkably volatile (Hansen and Jagannathan, 1991), and the valuation of stocks seems excessively volatile compared to its fundamentals (LeRoy and Porter, 1981; Shiller, 1981).

Typically, in asset markets, high valuations revert back to the mean through low subsequent returns. Not so for the whole U.S. government bond market. Even though there is evidence of return predictability for individual bonds (see Fama and Bliss, 1987; Campbell and Shiller, 1991; Cochrane and Piazzesi, 2005; Ludvigson and Ng, 2009; Cochrane, 2011), we find no evidence that return predictability for the entire bond portfolio pushes the debt/GDP ratio back to its long-run mean. The valuation of the entire U.S. government bond portfolio is too smooth compared to its fundamentals. The cash flows and the discount rates are the dogs that did not bark here. The valuation of all Treasurys is not sensitive enough to news about future surpluses or returns. Instead, we only find a statistically significant role for the future debt in accounting for the entire value of debt today.

Our paper does not impose no-arbitrage restrictions, but only uses accounting identities to

[8] These projections do not take into account any future legislative action to implement a fiscal correction, which may have been anticipated by investors. The source of bond market investors optimism may have been the expectation of future fiscal rectitude through legislation.

understand variation in the debt/output ratio, rather than the level. In other work, Jiang, Lustig, Van Nieuwerburgh, and Xiaolan (2019) price a claim to future surpluses in a no-arbitrage model, and they conclude that the debt is overpriced.[9]

There is a large literature on the pitfalls of predicting returns with persistent predictors, going back to Nelson and Kim (1993); Stambaugh (1999); Lewellen (2004); Torous, Valkanov, and Yan (2004); Campbell and Yogo (2006); Boudoukh, Israel, and Richardson (2020). The Stambaugh correction can be refined if the econometrician rules out ex ante that the predictor variable has a true autocorrelation larger than one (Lewellen, 2004).[10] We do not rule out non-stationary behavior of the debt/output ratio, because we are interested in testing fiscal sustainability.[11] We also implement the Campbell and Yogo (2006) testing procedure, valid under general assumptions, and we conclude there is neither evidence of surplus or return predictability.

We start in section 1 by deriving and implementing the decomposition of the debt/output ratio. Section 2 looks at a longer sample and longer horizons. The results are robust. Section 3 allows for permanent shocks to the debt/output ratio. Finally, section 4 argues that systematic forecast errors may play a role in accounting for debt/output dynamics.

1 What Drives Variation in the U.S. Debt/Output Ratio?

To illustrate the Stambaugh bias, we use the log-linear Campbell-Shiller decomposition of the government's debt/output ratio in Cochrane (2021a,b). This approach decomposes the variation in the debt/output ratio into three components: expected future government surpluses, expected future discount rates, and expected terminal value of the debt/output ratio in the future. After we correct for the Stambaugh bias, we find that most of the variation in the debt/output ratio cannot be attributed to subsequent government surpluses or discount rates even 10 years out.

1.1 Campbell-Shiller Decomposition of the Debt/Output Ratio

Let r_{t+j} denote the return on the government debt portfolio, x_{t+j} real GDP growth (in logs), and π_{t+j} log inflation. We use annual data. The government debt data is from CRSP Treasurys. We compute the market value of all marketable U.S. Treasurys. Following Hall and Sargent (2011),

[9] In no-arbitrage models, all of the information for forecasting bond returns is embedded in the yield curve, except when the underlying factors cannot be inverted from the yields (see Duffee, 2011). Some authors report evidence that macro factors have incremental forecasting ability for bond returns (Cooper and Priestley, 2008; Ludvigson and Ng, 2009; Joslin, Priebsch, and Singleton, 2014). Bauer and Hamilton (2017) conclude that the evidence for macro variables predicting bond returns, after controlling for bond yields, against the spanning hypothesis, is weaker than previously thought, citing similar small-sample biases in the presence of persistent predictors.

[10] If the econometrician imposes that the true autocorrelation of the debt/output ratio is smaller than one, that effectively implies that either surpluses or returns are predictable.

[11] In a Monte-Carlo experiment, we implement the Stambaugh-correction on simulated data generated from a unit-root data generating process for the debt/ouput ratio. We find spurious evidence of surplus and return predictability. However, the bias-corrected estimated slope coefficients recover the true values.

we compute the return on government debt as the sum of the principal and coupon payments less new issuance, plus the market value at the end of the period, divided by the market value at the end of the previous period. We exclude non-marketable debt which is mostly held in intra-governmental accounts. The inflation rate is the annual log change in the CPI, taken from the BLS. Nominal GDP is from the Bureau of Economic Analysis. Our sample of annual data comprises the period from 1947 to 2020.

We implement Cochrane (2021a)'s version of a Campbell-Shiller decomposition for the log of the debt/GDP ratio v_t. Let $\widetilde{r}_{t+j} = r_{t+j} - x_{t+j} - \pi_{t+j}$ denote the adjusted log bond return, which is the nominal log return on the government bond portfolio minus output growth and inflation (in logs). Let $s_{t+j} = sy_{t+j}/e^v$ denote the adjusted government surplus/output ratio, where sy_t denotes the actual surplus/output ratio and v denotes the average log debt/output ratio.

We start from the log-linearized return equation implied by the government budget constraint:

$$r_{t+1} - \pi_{t+1} - x_{t+1} = \rho v_{t+1} - v_t + s_{t+1},$$

where ρ is a constant of linearization. We will choose $\rho = \exp(-(r - x - \pi)) = 1$. This decomposition expands the debt/output ratio around the unconditional average $r = x + \pi$. [12] We back out the surplus variable s_t from this equation:

$$s_{t+1} \equiv \widetilde{r}_{t+1} - \Delta v_{t+1}. \tag{1}$$

Table 1 reports the decade-by-decade averages for these variables. The $r < g$ evidence is concentrated at the start of the sample, when the Fed and the Treasury were engaged in yield curve control. The inflation-and-growth adjusted returns on the entire bond portfolio, reported in the first column, are strongly negative in the decade after the war (-3.8%), partly due to low nominal returns combined with high inflation, partly because of strong economic growth. This pattern continues in the 60s, with growth-and-inflation-adjusted returns of -2.6%. In the 70s, bondholders were surprised by persistently high inflation, delivering adjusted returns of -2.5%. Overall, in the first three decades after WW-II, the average gap between returns and growth $\tilde{r}$ is -3%. r was much lower than g.

In the next four decades, the average inflation-and-growth adjusted returns $\tilde{r}$ are positive (1.5%). r was much higher than g. The 80s represents a radical departure from what came before. Bondholders earned high real returns in spite of high inflation. Growth-and-inflation-adjusted returns surge to 4.1% in the 80s. Even in the 90s, bondholders earned high real returns. Growth-and-inflation-adjusted returns are still positive in the 90s (1.6%) and the 00s (0.8%).

Hall and Sargent (2011) provide a detailed analysis of the ex post realized returns on govern-

[12] We relax this in section 2.3. The details of the derivation are in the appendix of Cochrane (2021a) on page 3 and 4.

ment debt and how these impact the U.S. debt dynamics. As we show in section B of the Appendix, when the debt/output ratio is stationary, the variance of ex post realized adjusted returns is completely attributable to cash flows in the long run:

$$\lim_{k\to\infty}\frac{1}{k}var[\tilde{r}_{t\to t+k}] = \lim_{k\to\infty}\frac{1}{k}var[s_{t\to t+k}]. \tag{2}$$

A similar equation for stocks implies that the variance of stock returns is driven only by dividend growth in the long run, if the dividend yield is stationary. But the case of Treasurys is quite different. If adjusted Treasury returns are i.i.d., with a variance ratio of one at all horizons ($\lim_{k\to\infty}\frac{1}{k}var[\tilde{r}_{t\to t+k}] = var[\tilde{r}_t]$), this equality dictates that there is mean reversion in the surpluses, i.e., the variance ratio of surpluses is smaller than one at long horizons,

$$\lim_{k\to\infty}\frac{1}{k}var[s_{t\to t+k}] < var[s_t]$$

provided that the returns are less volatile than the surpluses in the short run ($var[\tilde{r}_t] < var[s_t]$). In the 1947-2020 sample, the standard deviation of s_t is 6.47%, while the standard deviation of $\tilde{r}_t$ is 5.30%. However, the mean reversion will be small because the difference in short-run variances of cash flows and returns is small. On the other hand, if the surpluses are i.i.d., then the stationarity of the debt/output ratio would imply mean aversion in returns, which is implausible. This back-of-the-envelope evidence seems to rule out a world without surplus predictability, but potentially allows for a world without return predictability, because the short run volatility of cash flows is high relative to returns.

Treasurys are different from the case of stocks. In the case of stocks, the short run variance of stock returns exceeds that of cash flows, and i.i.d. dividend growth leads to mean reversion in returns, while i.i.d. stock returns would lead to mean aversion in dividend growth. In the case of stocks, the prima facie evidence rules out a world without return predictability, because the short run volatility of returns is so high.

By iterating eqn. (1) forward and taking expectations, we obtain the following expression for the debt/output ratio:

$$v_t = \mathbb{E}_t\sum_{j=1}^{T}\left(s_{t+j} - \widetilde{r}_{t+j}\right) + \mathbb{E}_t v_{t+T}. \tag{3}$$

The debt/output ratio reflects either expected future surpluses or expected future returns after adjusting for inflation and growth. Our decomposition allows for convenience yields on government debt.[13] These convenience yields would lower $\widetilde{r}$ in our decomposition. This accounting

[13]In models developed by Bassetto and Cui (2018); Brunnermeier, Merkel, and Sannikov (2020); Reis (2021); Chien and Wen (2019, 2020); Kocherlakota (2021); Mian, Straub, and Sufi (2021a), government debt allows agents to self-insure against idiosyncratic risk and provides liquidity services. The resulting convenience yield contribute a separate

identity leaves a future debt/output ratio term in the Campbell-Shiller decomposition.[14] If the TVC is satisfied and if v_t is stationary, then $\mathbb{E}_t v_{t+T}$ converges to its unconditional mean.

Taking covariances with v_t on both sides of the previous equation, we obtain the following expression for the variance of the debt/output ratio :

$$var(v_t) = cov\left(\sum_{j=1}^{T} s_{t+j}, v_t\right) - cov\left(\sum_{j=1}^{T} \widetilde{r}_{t+j}, v_t\right) + cov(v_t, v_{t+T}). \tag{4}$$

The log debt/output ratio varies because it either predicts future surpluses, future returns, or the future debt/output ratios. In other words, the adjustment to an increase in the debt/output ratio either happens through an increase in future surpluses or a decrease in real growth-adjusted returns, or the debt is simply rolled over T periods from now. If v_t is non-stationary, then we replace the unconditional population moments with the finite sample moments in this variance decomposition.[15]

$$\hat{var}_N(v_t) = \hat{cov}_N\left(\sum_{j=1}^{T} s_{t+j}, v_t\right) - \hat{cov}_N\left(\sum_{j=1}^{T} \widetilde{r}_{t+j}, v_t\right) + \hat{cov}_N(v_t, v_{t+T}). \tag{5}$$

To compute the variance decomposition, we directly estimate a system of univariate forecasting regressions for $\sum_{j=1}^{T} s_{t+j}, \sum_{j=1}^{T} \widetilde{r}_{t+j}, v_{t+j}$ using the lagged debt/output ratio as a predictor:

$$\begin{aligned}
\sum_{j=1}^{T} s_{t+j} &= a_s + b_T^s v_t + \varepsilon_{t+T}^s, \\
\sum_{j=1}^{T} \widetilde{r}_{t+j} &= a_r + b_T^r v_t + \varepsilon_{t+T}^r, \\
v_{t+T} &= \phi_0 + \phi_T v_t + \varepsilon_{t+T}^v.
\end{aligned} \tag{6}$$

Cochrane (2008); Lettau and Van Nieuwerburgh (2008) adopt the same approach to implementing a Campbell-Shiller decomposition of the price/dividend ratio for stocks. Our claim is not that the debt/output ratio is the only predictor of bond returns or surpluses. We simply want to compute covariance between the debt/output ratio and future returns/surpluses.

Just as their counterparts from predictability regressions using stock returns, these regression

component to the valuation of public debt.

[14]We do not impose the transversality condition (TVC) . That requires taking a stand on what the right discount rate is to eliminate the terminal value v_{t+T} term (see Giglio, Maggiori, and Stroebel, 2016; Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2019). There are equilibrium models that generate violations of the TVC including Samuelson (1958); Diamond (1965); Blanchard and Watson (1982); Hellwig and Lorenzoni (2009). Most of these models abstract from aggregate risk premia which would be priced into the terminal value and are likely to enforce the TVC (Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2020; Barro, 2020).

[15]If v_t is non-stationary, the unconditional variance is infinite. However, in any finite sample, we can still compute the sample moments, and compute this in-sample variance decomposition using the sample moments.

coefficients can be interpreted as the variance of the explanatory variable explained by each component for a certain horizon j:

$$\begin{aligned}
\frac{cov(\sum_{j=1}^{T} s_{t+j}, v_t)}{var(v_t)} &= b_T^s, \\
\frac{cov(-\sum_{j=1}^{T} \widetilde{r}_{t+j}, v_t)}{var(v_t)} &= -b_T^r, \\
\frac{cov(v_{t+T}, v_t)}{var(v_t)} &= \phi_T.
\end{aligned}$$

As discussed, the slope coefficients measure the fraction of the variance in the debt/output ratio at each horizon that is attributable to each component, the future returns, the future surpluses and future debt/output ratios respectively. Note that the coefficients sum to one at each horizon. The cross-equation restriction $b_T^s - b_T^r + \phi_T = 1$ is automatically satisfied, so that these three components jointly explain 100% of the variation in the debt/output ratio v_t.

To develop intuition for how the persistence in the debt/output ratio impacts the variance decomposition, consider the simpler case in which the debt/output ratio follows an AR(1): The variance decomposition is given by $b_1^s(1-\phi_1)^T/(1-\phi_1)$, $-b_1^r(1-\phi_1)^T/(1-\phi_1)$ and ϕ_1^T. In this case, if the debt/output approaches a unit-root process (namely, $\phi_1 \to 1$), v_{t+T} accounts for all of the variation in v_t at horizon T.

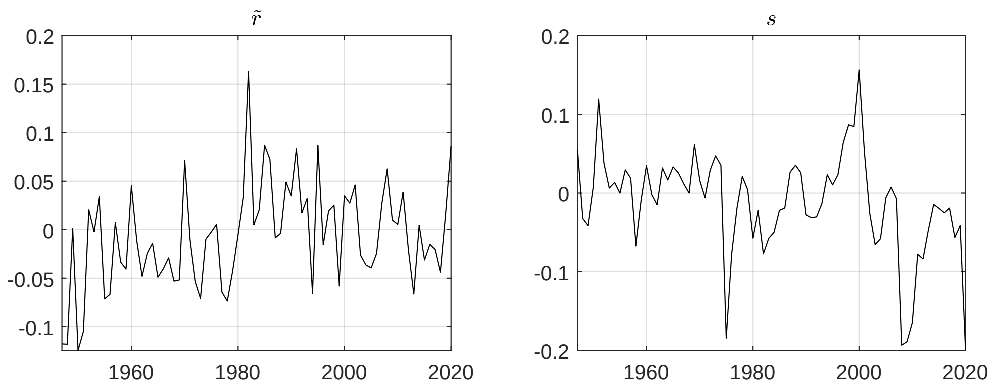
We define fiscal sustainability (over horizon T) as $\phi_T < 1$. At long horizons, we define fiscal sustainability as $\lim_{T\to\infty} \phi_T = 0$. If this condition is not satisfied, then the surplus/output ratio will inherit a unit root, which is not sustainable.

What do we expect to find in this variance decomposition? We consider a few different cases. First, we consider the case of risk-free or zero beta debt. If the debt is fiscally sustainable at horizon T, and the debt portfolio has zero beta, the real risk-free rate is i.i.d. and the growth rate of the economy is i.i.d, a reasonable benchmark, then the cash flows have to adjust: $b_T^s = 1 - \phi_T > 0$. The only channel pushing the debt/output ratio back to the mean is through larger surpluses, following an increase in the debt/output ratio. Over long horizons, we expect surpluses to account for all of the variation: $\lim_{T\to\infty} b_T^s = 1$.[16]

Second, we consider the case of risky debt. If the debt is fiscally sustainable at horizon T, and the surpluses are not predictable, then the growth-and-inflation adjusted returns have to adjust: $-b_T^r = 1 - \phi_T > 0$. The only channel pushing the debt/output ratio back to the mean is through lower growth-and-inflation adjusted returns following an increase in the debt/output ratio. Over

[16] What about the option of accumulating debt until it hits an upper bound? This policy is not consistent with zero-beta debt. When the economy hits the bound, the debt cannot be risk-free unless the surpluses live on a bounded support. In general, when the debt is zero-beta and the debt/output ratio follows an autoregressive process, the persistence of the surpluses is highly limited, unless the debt/output process has higher-order autoregressive dynamics (Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2020). In any case, we expect the market value of debt to respond before the economy hits the bound.

Figure 1: Debt Returns and Government Surpluses



This figure plots the inflation-and-growth-adjusted log debt returns $\tilde{r}_t$ and the surplus/output ratio s_t.

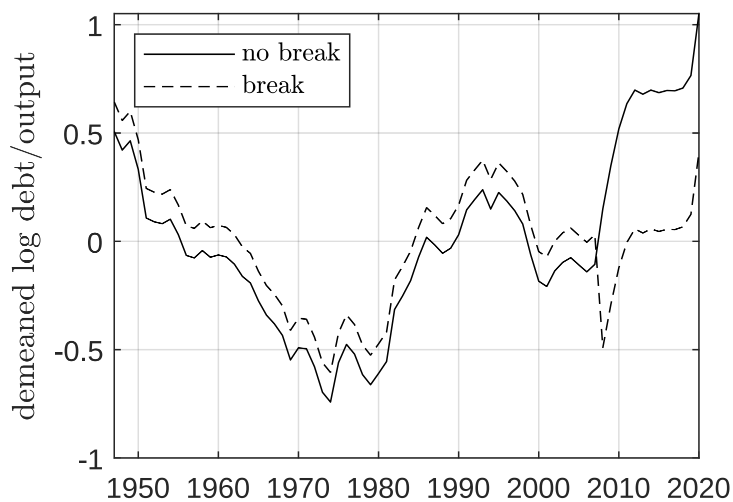
long horizons, we expect returns to account for all of the variation: $\lim_{T\to\infty} -b^r_T = 1$. If there is no predictability of the surpluses, then the deflated returns have to predictable. If there is no predictability of the adjusted returns or surpluses, then the debt/output ratio inherits a unit root, which violates fiscal sustainability.

Table 1: Summary Stats: Decade-by-Decade Averages

	$\tilde{r}$	r	x	π	$x+\pi$	s/y
			Forecasting $\sum_{j=1}^{T} r_{t+j}$			
1950-59	-3.8%	2.7%	4.1%	2.4%	6.5%	1.6%
1960-69	-2.8%	3.9%	4.4%	2.3%	6.7%	2.0%
1970-79	-2.5%	7.0%	3.2%	6.3%	9.5%	-1.4%
1947-1979	-3.5%	3.9%	3.6%	3.8%	7.4%	0.3%
1980-89	4.1%	11.8%	3.0%	4.6%	7.6%	-2.2%
1990-99	1.6%	6.9%	3.2%	2.2%	5.3%	1.9%
2000-2009	0.8%	4.9%	1.9%	2.2%	4.1%	-3.3%
2010-2019	-1.3%	2.6%	2.2%	1.7%	3.9%	-5.5%
2010-2020	-0.4%	2.9%	1.7%	1.6%	3.3%	-5.5%
1980-2019	1.3%	6.5%	2.6%	2.7%	5.2%	-2.3%
1980-2020	1.5%	6.5%	2.4%	2.6%	5.1%	-2.7%
1947-2019	-0.9%	5.4%	3.0%	3.2%	6.2%	-1.1%
1947-2020	-0.7%	5.4%	3.0%	3.2%	6.1%	-1.3%

Figure 1 plots the variables we want to forecast: the adjusted log returns $\tilde{r}_t$ in the left panel and the surplus/output ratio in the right panel. Returns were low from the start of the sample until the 1980s. After 2000, real growth-adjusted returns declined again. The surpluses backed out from eqn. (1) plotted on the right are more volatile, with a standard deviation of 6.4% than the actual surplus output ratios, with a standard deviation of 2.6%, but the two series are highly correlated

Figure 2: Debt/Output Ratio



The full line is the demeaned log debt/output ratio. The dashed line is the demeaned log debt/output ratio, demeaned by two different sub-sample means before and after 2007.

(0.85). [17]

We are interested in whether the debt/output ratio can predict the decade-by-decade variation in growth-and-inflation-adjusted returns or surpluses. The full line in Figure 2 is the predictor variable, the debt/output ratio, demeaned and logged. There was a marked decline in the debt/output ratio between 1948 and 1974, followed by a gradual increase in the eighties and nineties, and a surge in the wake of the Great Financial Crisis.

1.2 Results without Bias Correction

Figure 3 plots the baseline regression results. These figures plot the slope coefficients in regressions of future returns, surpluses and debt/output ratios on the current debt/output ratio (see regressions in eqn. (6)). At the one-year horizon, 101% of the variance is attributed to the next year's debt/output ratio. That is to say, the log debt/output ratio is highly persistent. The first-order autocorrelation is 1.01.[18] Even at the 5-year horizon, 83% of the debt/output ratio fluctuations can be attributed to the future debt/output ratio. The R^2 in this debt/output predictability regression exceeds 50% at the 5-year horizon—see Table 2, which reports these predictability results for horizons from 1 to 10 years.

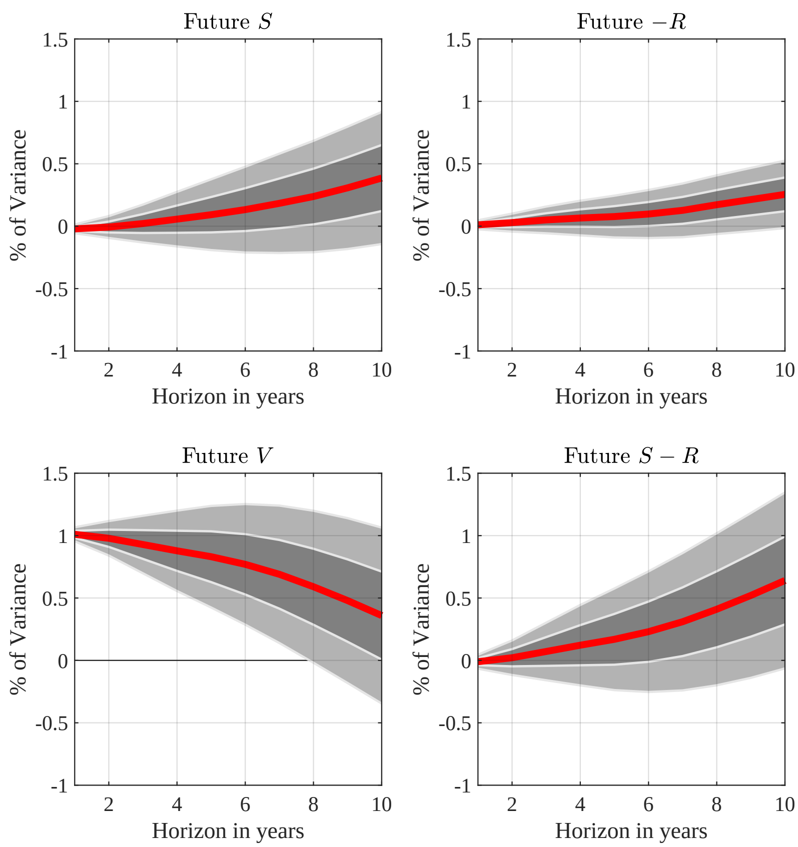
[17] We check below that our results extend to the actual surplus/output series, but we use s_t as the main variable to enforce the cross-equation restriction.

[18] In fact, using the Augmented Dickey-Fuller test, we cannot reject the null hypothesis of the presence of the unit root in the log debt/output ratio in our sample period. However, we need to acknowledge the low power of these unit-root tests.

At the 10-year horizon, the other two channels start to matter: 39% of the variation is attributed to variations in expected government surpluses and 25% of the variation is attributed to variations in discount rates. However, we cannot reject the null that the fraction is zero. The terminal debt/output ratio in the 10-year horizon only accounts for 36% of the variation.

Why is the discount rate channel so weak? The debt/output ratio does predict nominal bond returns with the right sign, as emphasized by Cochrane (2021a): higher debt/output predicts lower nominal returns. This effect is mechanical because bond prices are determined by the nominal log returns over the maturity of the bond: $p_t^N = -\sum_{i=0}^{N-1} hpr_{t+i+1}^{N-i}$. This effect is almost completely offset by a similar effect on inflation, as shown in Table 3. The top panel looks at forecasts

Figure 3: Variance Decomposition of Log Debt/Output Ratio



This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future government surpluses $\sum_{j=1}^{T} s_{t+j}$, future discount rates $\sum_{j=1}^{T} \tilde{r}_{t+k}$, future log debt/output ratio v_{t+T}, and the combination of the future government surpluses and discount rates. Sample is annual, 1947—2020. We plot 1 s.e. (dark) and 2 s.e. (light) CIs. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t.

of nominal bond returns, the second panel looks at forecasts of real GDP growth, the third panel considers inflation, and the fourth panel considers real-growth adjusted returns. The net effect of the debt/output ratio on future real returns is small and statistically insignificant as a result.

Table 2: Forecasting Returns and Surpluses with log Debt/Output ratio

OLS Regression of $\sum_{j=1}^{T} s_{t+j}, \sum_{j=1}^{T} \tilde{r}_{t+j}, v_{t+T}$ on v_t. Annual data. Sample: 1947—2020. Standard errors generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t, and we set $s_{t+1} = \tilde{r}_{t+1} - \Delta v_{t+1}$.

Horizon	1	2	3	4	5	6	7	8	9	10
					Forecasting $\sum_{j=1}^{T} -\tilde{r}_{t+j}$					
$-b_T^r$	0.01	0.03	0.05	0.07	0.08	0.1	0.13	0.17	0.21	0.25
s.e.	0.02	0.04	0.05	0.07	0.08	0.09	0.11	0.12	0.13	0.13
R^2	0.01	0.02	0.03	0.04	0.04	0.05	0.06	0.08	0.1	0.12
unbiased	-0.01	-0.02	-0.02	-0.03	-0.04	-0.04	-0.04	-0.01	0	0.02
					Forecasting $\sum_{j=1}^{T} s_{t+j}$					
b_T^s	-0.02	-0.01	0.02	0.06	0.09	0.13	0.18	0.24	0.31	0.39
s.e.	0.02	0.04	0.08	0.11	0.14	0.17	0.2	0.22	0.24	0.26
R^2	0.02	0	0	0.01	0.02	0.03	0.05	0.06	0.09	0.11
unbiased	-0.05	-0.07	-0.08	-0.07	-0.07	-0.06	-0.05	-0.03	0.01	0.05
					Forecasting v_{t+T}					
ϕ	1.01	0.98	0.93	0.88	0.83	0.77	0.69	0.59	0.48	0.36
s.e.	0.03	0.07	0.11	0.16	0.2	0.24	0.27	0.3	0.33	0.35
R^2	0.95	0.85	0.74	0.64	0.54	0.43	0.32	0.22	0.13	0.07
unbiased	1.07	1.09	1.1	1.1	1.11	1.11	1.08	1.04	0.99	0.92

Table 3: Forecasting Nominal Returns and Inflation with Log Debt/Output Ratio

This table further decomposes the adjusted bond return $\tilde{r}_{t+j}$ into its three components: nominal bond return r_{t+j}, GDP growth x_{t+j}, and inflation π_{t+j}. The adjusted bond return $\tilde{r}_{t+j}$ is equal to $r_{t+j} - x_{t+j} - \pi_{t+j}$. We run OLS Regression of $\sum_{j=1}^{T} r_{t+j}, \sum_{j=1}^{T} x_{t+j}, \sum_{j=1}^{T} \pi_{t+j}$ on $\tilde{v}_t$. Annual data. Sample: 1947—2020. Standard errors generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, x_t, π_t, v_t) on two lags of v_t.

Horizon	1	2	3	4	5	6	7	8	9	10
					Forecasting $\sum_{j=1}^{T} r_{t+j}$					
b_T^r	-0.05	-0.1	-0.16	-0.22	-0.28	-0.34	-0.42	-0.5	-0.59	-0.67
s.e.	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.08	0.09	0.11
R^2	0.19	0.28	0.34	0.38	0.41	0.44	0.47	0.5	0.53	0.55
unbiased	-0.04	-0.08	-0.12	-0.17	-0.21	-0.26	-0.32	-0.4	-0.47	-0.54
					Forecasting $\sum_{j=1}^{T} x_{t+j}$					
b_T^x	0	0	0	0	0	0	0	0.01	0.01	0.02
s.e.	0.01	0.02	0.02	0.03	0.04	0.05	0.05	0.06	0.07	0.07
R^2	0	0	0	0	0	0	0	0	0	0.01
unbiased	-0.01	-0.02	-0.03	-0.04	-0.05	-0.06	-0.07	-0.07	-0.07	-0.08
					Forecasting $\sum_{j=1}^{T} \pi_{t+j}$					
b_T^π	-0.04	-0.07	-0.11	-0.15	-0.2	-0.24	-0.29	-0.34	-0.39	-0.44
s.e.	0.01	0.01	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
R^2	0.44	0.51	0.54	0.55	0.59	0.62	0.63	0.63	0.63	0.63
unbiased	-0.04	-0.07	-0.11	-0.15	-0.2	-0.24	-0.29	-0.34	-0.39	-0.44
					Forecasting $\sum_{j=1}^{T} \tilde{r}_{t+j}$					
$b_T^{\tilde{r}}$	-0.01	-0.03	-0.05	-0.07	-0.08	-0.1	-0.13	-0.17	-0.21	-0.25
s.e.	0.02	0.04	0.06	0.07	0.09	0.1	0.11	0.12	0.13	0.14
R^2	0.01	0.02	0.03	0.04	0.04	0.05	0.06	0.08	0.1	0.12
unbiased	0.01	0.02	0.02	0.03	0.04	0.04	0.04	0.01	0	-0.02

To generate CIs for these coefficients, we construct the standard errors by bootstrapping 10,000 samples from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We draw the surplus innovation such that the identity is enforced. Here is the full system of equations for the bootstrap:

$$
\begin{aligned}
s_{t+1} &= \widetilde{r}_{t+1} - \Delta v_{t+1}, \\
\widetilde{r}_{t+1} &= a_r + b_1^r v_t + b_2^r v_{t-1} + u_{t+1}^r, \\
v_{t+1} &= \psi_0 + \psi_1 v_t + \psi_2 v_{t-1} + u_{t+1}^v.
\end{aligned} \tag{7}
$$

For bootstrapping purposes, we use two lags because the debt/output ratio fits an AR(2) structure well, which delivers white noise estimated residuals to bootstrap from. The surplus residual is backed out from the cross-equation restriction. The CIs imply considerable uncertainty in the point estimate. Even at the 10-year horizon, the two-standard-deviation CIs for the b_j^s and b_j^r contain zero, so that the explanatory power of future government surpluses and discount rates are indistinguishable from zero.

1.3 Bias Correction

The high persistence in the explanatory variable v_t raises concern of small-sample bias. In particular, the OLS slope coefficients, $|b_j^r|$, are biased upwards in absolute value, because the innovations to the returns ε^r are positively correlated with the regressor innovations ε^v, and the regressor is highly persistent (Stambaugh, 1999). For example, an increase in bond risk premia will tend to induce lower realized returns and lower debt/output ratio. That gives rise to positive correlation between the regressor innovations and the return innovations. These positive biases tend to increase for long-horizon predictability regressions (Boudoukh, Israel, and Richardson, 2020).[19]

We find a similar positive bias for the surplus predictability regression. An increase (decrease) in debt issuance tends to coincide with the government running large deficits (surpluses). As a result, there is a strong, negative correlation between ε^r and the regressor innovations ε^v. This, combined with the persistence of the debt/output ratio, induces a large upward bias in b_T^s as well. In both case, the bias leads us to find too much predictability.

The unbiased coefficients were obtained by applying the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in the predictability regression with horizon j:

$$
\begin{aligned}
bias_T^r &= \mathbb{E}\left(\widehat{b}_T^r - b_T^r\right) = \frac{1}{N}\left[T(1+\phi) + 2\phi\frac{1-\phi^T}{1-\phi}\right] \times -\frac{cov(\varepsilon^v, \varepsilon^r)}{var(\varepsilon^v)}, \\
bias_T^s &= \mathbb{E}\left(\widehat{b}_T^s - b_T^s\right) = \frac{1}{N}\left[T(1+\phi) + 2\phi\frac{1-\phi^T}{1-\phi}\right] \times -\frac{cov(\varepsilon^v, \varepsilon^s)}{var(\varepsilon^v)},
\end{aligned}
$$

[19]Note that we plot minus the slope coefficient $-b_T^r$ in the figures.

where ϕ denotes the first-order autocorrelation of v_t , T denotes the horizon, and N denotes the size of the sample. We note that $corr(\varepsilon^v, -\varepsilon^r) = -0.75$ and $corr(\varepsilon^v, \varepsilon^s) = -0.85$, so the implied biases for the coefficient b_T^s associated with surplus and for the coefficient $-b_T^r$ associated negative debt are both positive. Fixing the volatilities of these innovations, the small-sample bias grows as the true autocorrelation $\phi \to 1$, and as $corr(\varepsilon^v, -\varepsilon^r) \to 1$ and $corr(\varepsilon^v, \varepsilon^s) \to -1$. Given the persistence of the debt/output ratio, and the size of the residual correlations, it is fair to say that the bias is close to its upper bound.

To better understand the bias, we can restate the bias of the coefficients at horizon $T = 1$ as follows :

$$
\begin{aligned}
bias_1^r &= \mathbb{E}[\widehat{\phi} - \phi] \times -\frac{cov(\varepsilon^v, \varepsilon^r)}{var(\varepsilon^v)}, \\
bias_1^s &= \mathbb{E}[\widehat{\phi} - \phi] \times -\frac{cov(\varepsilon^v, \varepsilon^s)}{var(\varepsilon^v)},
\end{aligned}
$$

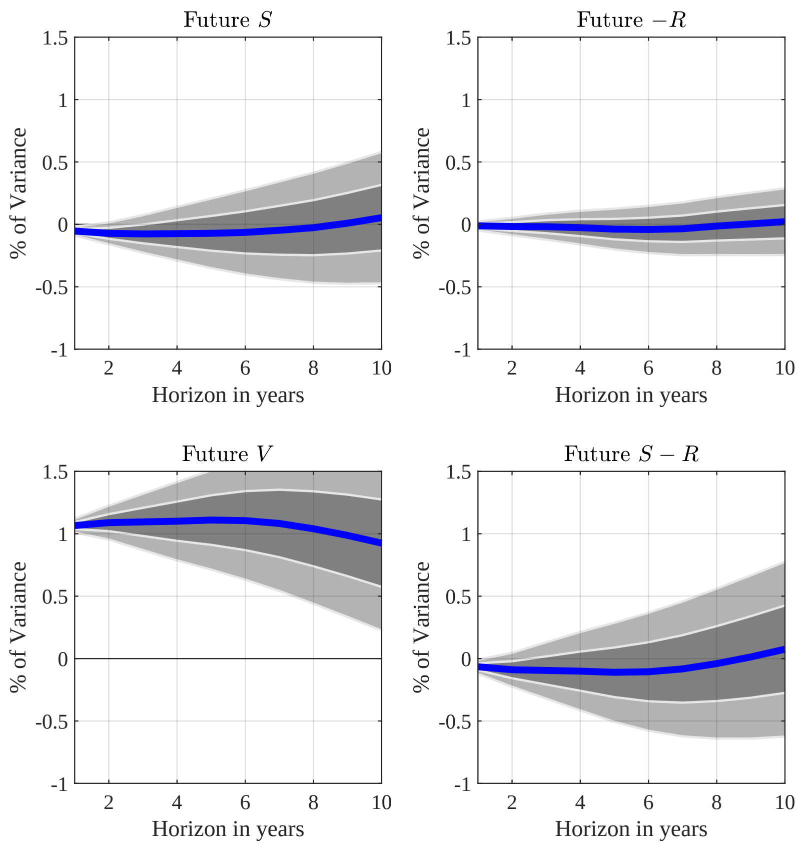
where $\mathbb{E}[\widehat{\phi} - \phi]$ is roughly $-(1 + 3\phi)/N$. This expectation is taken over all possible values of true autocorrelation ϕ. If instead we are willing to restrict $\phi \leq 1$, then the bias attains an upper bound at $\phi = 1$ (Lewellen, 2004). In the case of stock return predictability, Lewellen (2004) shows that the null of no predictability can be rejected more often if this stationarity restriction is imposed. This is not surprising: imposing that $\phi \leq 1$ is equivalent to imposing $b_1^r + b_1^s \geq 0$, assuming ex ante that there is return or cash flow predictability. We do not impose any ex ante restrictions on the true autocorrelation of the the debt/output ratio. We do not want to rule out that the U.S. is on a fiscally unsustainable path.

Figure 4 reports the bias-adjusted regression coefficients, which still sum to one. The bias-corrected variance decomposition attributes -4% and -7% of the debt/output ratio variance to the discount rate and cash flow channel respectively at the 5-year horizon. As a result, 111% is accounted for by the future debt/output ratio. At the 10-year horizon, we still attribute 92% of the variance to the future debt/output ratio, after correcting for the small-sample bias.

These point estimates imply that there is no statistical evidence that the debt/output ratio predicts either future government surpluses or future bond returns. The null hypothesis that there is no predictability in future government surpluses or future bond returns cannot be rejected at any horizon. At the 10-year horizon, we cannot even reject the joint null that the debt/output ratio does not predict the sum of surpluses and returns.

The confidence intervals for the discount rate contribution in the top right panel are narrow. We can rule out that discount rates play a quantitatively significant role in imputing mean reversion to the debt/output ratio, if there is any. However, the confidence intervals for the cash flow channel in top left panel are wider at longer horizons. Even though the point estimates are close to zero, the 95% confidence interval includes values of more than 50% at the 10-year horizon.

Figure 4: Variance Decomposition of Log Debt/Output Ratio after Bias Correction



This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future government surpluses $\sum_{j=1}^{T} s_{t+j}$, future discount rates $\sum_{j=1}^{T} \widetilde{r}_{t+k}$, future log debt/output ratio v_{t+T}, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1947—2020. We plot 1 s.e. (dark) and 2 s.e. (light) CIs. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t.

Similarly, we cannot definitively rule out that there is significant mean-reversion at the 10-year horizon, even though the point estimate for the future V contribution, shown in the bottom left panel, is close 100 %. We have low power at longer horizons, also a well-documented feature of unit root tests.

1.4 Tests of Predictability Under Local-to-Unity Asymptotics

Standard return predictability tests, such as the t-test of the slope of the OLS predictability coefficient may be inappropriate when the predictor is persistent and its innovations are highly correlated with returns. In such cases, large-sample theory provides a poor approximation to the finite-sample distribution of tests statistics. Campbell and Yogo (2006) offer an alternative test that is valid under general assumptions about the predictor dynamics, even when the largest root is larger than one, and the distribution of innovations. Campbell and Yogo (2006) develop a pre-test to diagnose whether the conventional t-test leads to valid inference. Their Dickey-Fueller Generalized Least Squares test is based on the CI for the largest autoregressive root of the predictor variable. Applying their method to our context, we obtain a 95% CI of $[0.958, 1.061]$ for the persistence of the debt/output variable v_t. Since this CI contains the unit root, this indicates that standard t-tests are not valid.

Campbell and Yogo go on to develop an asymptotically valid and efficient Q-test, which results in a Bonferroni CI for the predictive coefficient of interest. When we apply their procedure to the surplus predictability regression, we find a 90% CI of $[-0.051, 0.019]$. This CI includes zero, so that we fail to reject the null hypothesis that the lagged debt/gdp ratio v_t does not forecast the surplus/gdp ratio s_{t+1} ($H_0 : b_1^s = 0$). We repeat the analysis for the adjusted return on the debt portfolio, $\widetilde{r}_{t+1}$ and find a 90% Bonferroni CI of $[-0.013, 0.044]$. We fail to reject the null hypothesis that the lagged debt/gdp ratio does not forecast the government debt return ($H_0 : b_1^r = 0$). This analysis confirms our findings: once the high degree of persistence of the debt/gdp ratio is taken into account, the evidence for predictability of future surpluses or future returns by the lagged debt/gdp ratio is very weak.

2 Robustness

We conduct robustness by looking at the shorter Bohn (1998) sample and at a longer sample from 1842. We also consider different steady-state values for the adjusted returns on the debt portfolio, and consider different variable definitions for the predictor (debt/GDP ratio) and the predicted variable (government surpluses).

2.1 Shorter Bohn Sample

In a classic paper on this topic, Bohn (1998) interprets his finding that the debt/output ratio predicts surpluses as evidence against the unit root hypothesis, but our results show that small-sample bias will give rise to large point estimates at longer horizons. We repeated our exercise on Bohn (1998)'s 1948—1995 sample. Before correcting for the bias, there is a much larger role for fundamentals in this subsample, as shown in Figure A1, consistent with Bohn's findings. However, this evidence mostly disappears after the bias correction.

2.2 Longer Sample

Using the Hall, Payne, and Sargent (2018) historical data, we construct a longer U.S. sample that starts in 1842. With the longer sample, we also report results with forecast horizon up to 25 years instead of 10 years. The variance decomposition with and without bias adjustment is plotted in Figure A2. We focus on the results after the bias correction. The point estimates are similar to those obtained in the post-war sample. However, the longer sample shrinks the CIs around the point estimates. At the 10-year horizon, we can now rule out that future V accounts for less than 50% of the variation, or, conversely, we can rule out that fundamentals account for more than 50% of the variation.

There is no statistical evidence of either surplus or return predictability. Furthermore, we cannot reject the null that the fundamentals do not explain any variation, even after 25 years. After the bias correction, we find that future v still explains about 70% of the variation in today's v even 25 years out.

2.3 Approximation around Different Steady-State Values

Recall that we back out the linearized government surplus s_{t+1} from the following relation,

$$\tilde{r}_{t+1} = \rho v_{t+1} - v_t + s_{t+1},$$

where $\rho = \exp(-(r - x - \pi))$ is a constant of linearization. In the benchmark case, we chose $\rho = 1$ to linearize the equation system at $r = x + \pi$. Here, we re-derive our equations for general $\rho = \exp(-(r - x - \pi))$, and report a robustness result using $r - (x + \pi) = -1\%$ so that the risk-free rate is below the output growth rate by 1% per annum, in the Blanchard (2019) region of the paremeter space.

For the case of an arbitrary ρ, equation (3) becomes

$$v_t = \mathbb{E}_t \sum_{j=1}^{T} \rho^{j-1} \left(s_{t+j} - \widetilde{r}_{t+j} \right) + \mathbb{E}_t \rho^T v_{t+T}. \tag{8}$$

The corresponding regression equations become

$$\begin{aligned} \sum_{j=1}^{T} \rho^{j-1} s_{t+j} &= a_s + b_T^s v_t + \varepsilon_{t+T}^s, \\ \sum_{j=1}^{T} \rho^{j-1} \widetilde{r}_{t+j} &= a_r + b_T^r v_t + \varepsilon_{t+T}^r, \\ \rho^T v_{t+T} &= \phi_0 + \phi_T v_t + \varepsilon_{t+T}^v. \end{aligned} \tag{9}$$

so that the cross-equation restriction $b^s_T - b^r_T + \phi_T = 1$ is still satisfied.

Since we use the same persistent predictor v_t on the right-hand side, we apply the same Stambaugh-bias adjustment formula as in the main text. Figure A3 reports the estimates with and without bias adjustment for $\rho = \exp(1\%)$. Figure A4 reports the estimates for $\rho = \exp(-1\%)$. The results after the bias correction are quite similar to the ones obtained in the benchmark $\rho = 0$ case.

2.4 Domestic Holdings of Treasury Debt

Foreign investors and the Fed are considered to be relatively price inelastic investors (Jiang, Krishnamurthy, and Lustig (2022)), and their holdings increase dramatically in the past 30 years. To examine whether our result is solely driven by their holdings, we exclude their holdings and construct the ratio between the debt held by the domestic sector and the output, which we denote by $v^{domestic}_t$. Figure A5 plots this ratio (in level as opposed to in log), and shows that there has indeed been an increase in the gap between domestic holdings of Treasury debt and the total amount of Treasury debt.

Do the domestic holdings of Treasury debt have a stronger predictive power for future surpluses or discount rates? We regress future government surpluses and future adjusted debt returns on this predictor:

$$\begin{aligned}\sum_{j=1}^{T} \rho^{j-1} s_{t+j} &= a_s + \hat{b}^s_T v^{domestic}_t + \varepsilon^s_{t+T},\\ \sum_{j=1}^{T} \rho^{j-1} \widetilde{r}_{t+j} &= a_r + \hat{b}^r_T v^{domestic}_t + \varepsilon^r_{t+T},.\end{aligned}$$

Since the domestically held debt/output ratio is not equal to future government surpluses minus adjusted debt returns, the adding-up constraint between $\hat{b}^s_T$, $\hat{b}^r_T$ and the autocorrelation of $v^{domestic}_t$ does not hold. Nevertheless, we can still study these regression coefficients and apply small-sample bias adjustments. We report the regression coefficients in Figure A6 as functions of the forecast horizon T. $v^{domestic}_t$ explains around 40% of the variation in the future government surpluses $\sum_{j=1}^{T} \rho^{j-1} s_{t+j}$ at the 10-year horizon, suggesting that future surpluses may respond more to the debt that has to be absorbed by the more elastic domestic investors. However, after the bias correction, we still cannot reject the null that surpluses are not predictable.

2.5 Actual Government Surpluses

Finally, we note that our surplus variable s_t, as defined in equation (1), is defined based on the log-linearization that allows the add-up constriant to hold. It is highly correlated with the actual surplus/GDP ratio s_t^{Actual} with a correlation of 0.83. To confirm the robustness of this approximation, we examine whether the sum of actual government surplus/GDP ratios can be predicted by the log debt/GDP ratio:

$$\sum_{j=1}^{T} \rho^{j-1} s_{t+j}^{Actual} = a_s + \hat{b}_T^s v_t + \varepsilon_{t+T}^s.$$

We report the regression coefficients in Figure A7. Consistent with our main result, the log debt/GDP ratio does not predict the actual government surplus/GDP ratios, either.

3 Permanent Shocks to the Debt/Output Ratio

Our evidence is consistent with a unit root in the debt/output ratio. Next, we evaluate the accuracy of the small-sample bias correction term by simulating 10,000 samples with the same length as the actual sample. We simulate under the null that there is no mean reversion in the debt/output ratio. As a result, there is no contribution from return/surplus predictability (the fundamentals) either ($b_T^s - b_T^r = 0 = 1 - \phi_T$) at all horizons T.

We assume that returns are i.i.d. and that the debt/output ratio follows a unit root process. We estimate an ARIMA(1,1,0) process for the debt/output dynamics. Hence, the system is given by the following equations:

$$\begin{aligned} \Delta v_{t+1} &= \psi_0 + \psi_1 \Delta v_t + \varepsilon_{t+1}^v, \\ \tilde{r}_{t+1} &= r_0 + \varepsilon_{t+1}^r, \\ v_{t+1} &= v_t + \Delta v_{t+1}. \end{aligned}$$

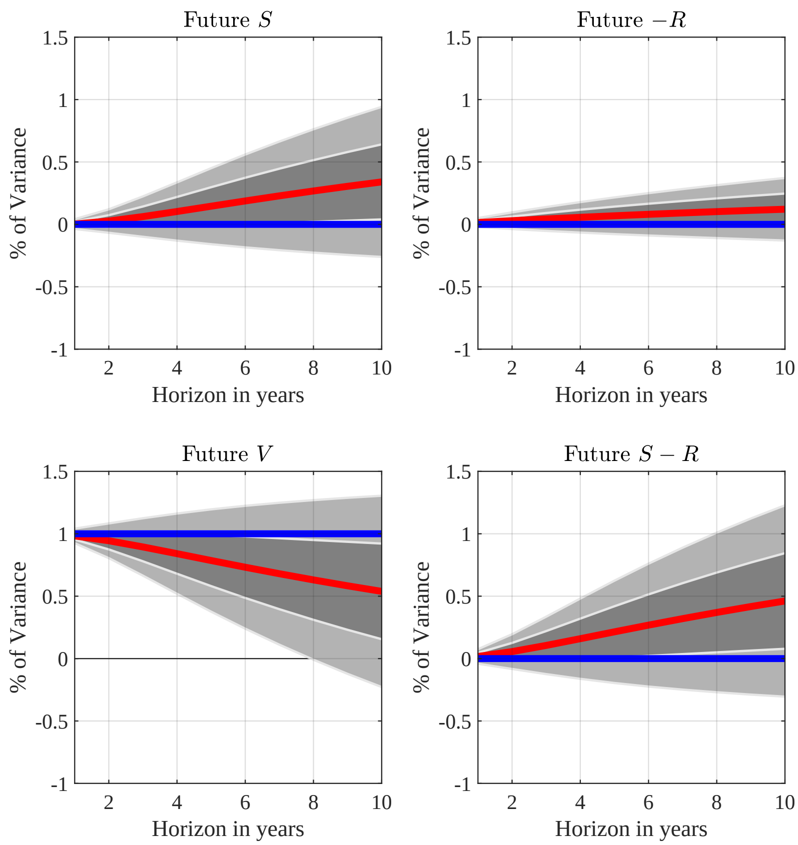
We infer the government surplus from the accounting identity:

$$s_{t+1} \equiv \widetilde{r}_{t+1} - \Delta v_{t+1}.$$

In each round of simulation, we draw with replacement from the joint distribution of the debt/output residual and the bond return $(\varepsilon_{t+1}^v, \varepsilon_{t+1}^r)$. Then, we run the forecasting regressions in (6) in each simulated sample. We plot the mean slope coefficients obtained running the forecasting regressions and the two-standard-deviation CIs around these mean slope coefficient estimates.

In the simulated data, we find that future government surpluses and discount rates appear to explain variations in the debt/output ratio, even though we know that they do not. Figure 5 reports the variance decomposition implied by the simulated samples. The average slope coefficients obtained from the unit root model imply variance decompositions between the fundamentals and the future debt/output ratios that are very close to our point estimates in Figure 3. The

Figure 5: Variance Decomposition of Log Debt/Output Ratio under Unit Root



This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future government surpluses $\sum_{j=1}^{T} s_{t+j}$, future discount rates $\sum_{j=1}^{T} \tilde{r}_{t+k}$, future log debt/output ratio v_{t+T}, and the combination of the future government surpluses and discount rates. The samples are generated by simulation under the null that the debt/output ratio has a unit root. We plot the mean of the small-sample slope coefficients in red. We also plot the long-sample slope coefficients in blue from a single simulation of 100,000 periods. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

OLS estimates for the autoregressive coefficient are severely biased downwards in small samples when the true model has a unit root (Hamilton, 1994, p. 217). As a result, we find spurious evidence of mean reversion that creates a large role for fundamentals over longer horizons, in cases where there is no mean-reversion. The true slope coefficients are one at all horizons.[20]

Quantitatively, the simulation result suggests that the downward bias in the variance explained by the future log debt/output ratio v_{t+T} is about 50% at the 10-year horizon. If we adjust the variance decomposition result in Figure 3, the bias-adjusted variance explained by the future

[20]The variance decomposition itself is not well defined for a unit root process, because the unconditional variance is not well defined. However, we can still run the estimation in the simulated small samples.

log debt/output ratio is about 36% + 50% = 86% in the 10-year horizon. This result is quite close to Figure 4 that adjusts the bias using the Stambaugh (1999) formula, and it fails to reject the null that the future log debt/output ratio explains 100% of the variance in today's log debt/output ratio.

3.1 Structural Breaks and Variance Decomposition of the Transitory Component of Debt/Output

One way of allowing for permanent shocks to the debt/output ratio, consistent with the unit-root evidence, is to allow for a structural break. There may have been structural shifts in the relation between the valuation of debt and the fundamentals. A major contributor to the small role of fundamentals is the large run-up in government debt during the Great Financial Crisis which was not followed by commensurate increases in surpluses or decreases in returns. Consequently, we consider a structural break in the log debt/output ratio in 2007.

Following Lettau and Van Nieuwerburgh (2008)'s work on stock return predictability, we allow for a structural break in the log debt/output ratio by demeaning the log debt/output ratio $\widetilde{v}_t = v_t - \bar{v}_t$ with a lower pre-2007 sample mean ($\bar{v}_t, t < 2007$) and a higher post-2007 sample mean ($\bar{v}_t, t \geq 2007$). The structural break introduces a 0.78 (in log scale) permanent increase in the debt/output ratio. Figure 2 plots the resulting series. This approach removes a low-frequency component from the debt/output ratio.[21] Obviously, when we allow for this break, we introduce permanent innovations in the debt/output ratio.

We re-estimate the forecasting regressions using the new predictor. The slope coefficients now provide a variance decomposition of the transitory variation in debt/output. Taking covariances with v_t on both sides of the previous equation, we obtain the following expression for the variance of the transitory component of debt/output ratio :

$$var(\widetilde{v}_t) = cov\left(\sum_{j=1}^{T} s_{t+j}, \widetilde{v}_t\right) - cov\left(\sum_{j=1}^{T} \widetilde{r}_{t+j}, \widetilde{v}_t\right) + cov(\widetilde{v}_t, \widetilde{v}_{t+T}). \tag{10}$$

The resulting break-adjusted series is much less persistent than the original series, and hence, because of the cross-equation restrictions $b^s_T - b^r_T + \phi_T = 1$, will mechanically impute a larger role to fundamentals in explaining the variation in the transitory component, by lowering ϕ_T.

The transitory component of the debt/output ratio now explains 23% of the variation in surpluses at the 5-year horizon (R^2). Figure 6 decomposes the variance of the transitory component into its fundamental components. When the transitory component of the debt/output ratio is high, surpluses tend to increase to push the debt/output ratio back down. As reported in Table 4, the fundamentals now explain 24% (49%) of the variation in transitory component of debt/output

[21] We recompute the surpluses by feeding $\widetilde{v}_t$ into eqn. (1) so the cross-equation restriction still holds.

at the 5-year (10-year) horizon. We can reject the null that fundamentals do not play a role at 10 years. This approach restores a role for fundamentals after accounting for the small-sample bias, but only in explaining the transitory variation in the debt/output ratio. This variance decomposition punts on the large and permanent increase in the debt/output ratio in the post-2007 sample.

Table 4: Forecasting Returns and Surpluses with Break-Adjusted log Debt/Output ratio

OLS Regression of $\sum_{j=1}^{T} s_{t+j}, \sum_{j=1}^{T} \widetilde{r}_{t+j}, v_{t+T}$ on v_t. $\widetilde{v}_t$ with structural break in 2007. Annual data. Sample: 1947—2020. Standard errors generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t, and we set $s_{t+1} = \widetilde{r}_{t+1} - \Delta v_{t+1}$.

Horizon	1	2	3	4	5	6	7	8	9	10
	Structural Break									
	Forecasting $\sum_{j=1}^{T} -\widetilde{r}_{t+j}$									
$-b_T^r$	0.03	0.05	0.07	0.07	0.07	0.08	0.11	0.16	0.2	0.24
s.e.	0.03	0.05	0.07	0.09	0.11	0.13	0.14	0.16	0.17	0.18
R^2	0.02	0.03	0.04	0.03	0.02	0.02	0.04	0.06	0.08	0.1
unbiased	0.01	0.02	0.03	0.02	0.01	0.02	0.04	0.07	0.11	0.14
	Forecasting $\sum_{j=1}^{T} s_{t+j}$									
b_T^s	0.07	0.16	0.25	0.34	0.41	0.46	0.51	0.56	0.62	0.68
s.e.	0.03	0.07	0.11	0.13	0.16	0.17	0.19	0.2	0.21	0.23
R^2	0.04	0.12	0.2	0.29	0.36	0.42	0.47	0.5	0.53	0.57
unbiased	0.03	0.08	0.14	0.2	0.23	0.25	0.27	0.29	0.32	0.36
	Forecasting v_{t+T}									
ϕ	0.91	0.79	0.68	0.59	0.53	0.45	0.38	0.29	0.19	0.08
s.e.	0.05	0.09	0.13	0.16	0.19	0.2	0.22	0.23	0.23	0.24
R^2	0.86	0.7	0.55	0.44	0.35	0.27	0.19	0.11	0.05	0.01
unbiased	0.96	0.89	0.83	0.78	0.76	0.73	0.69	0.64	0.58	0.51

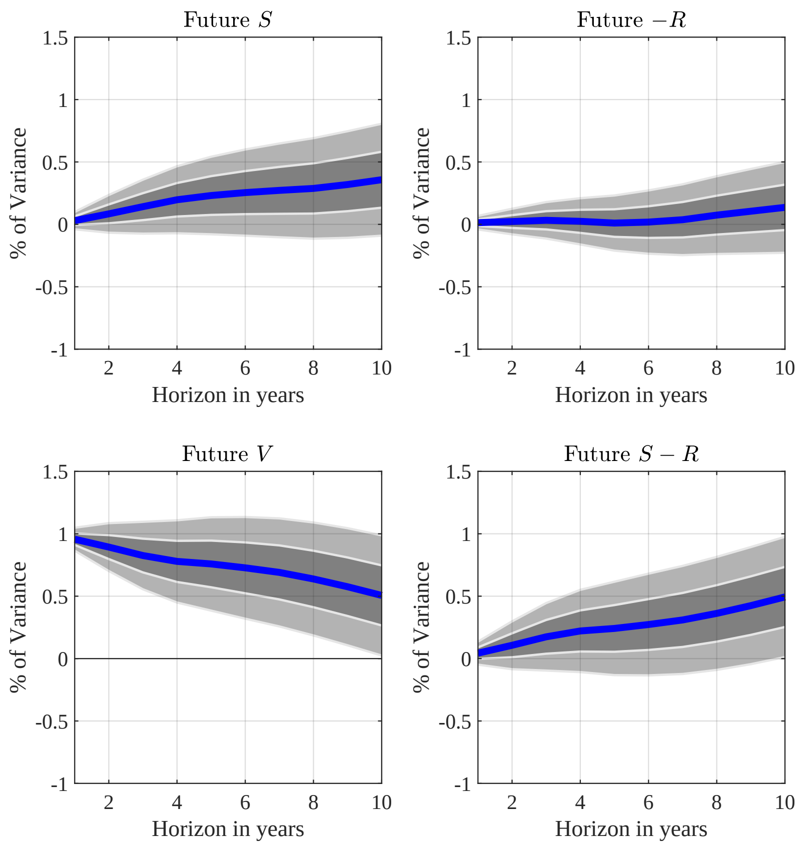
4 Systematic Forecast Errors as a Source of Permanent Shocks to the Debt/Output Ratio

We find a large Stambaugh bias in the fiscal predictability regression. After we correct this bias, future government surpluses and discount rates do not explain variation in the level of outstanding debt/output ratio. Our results suggest that the debt is mis-priced, perhaps because investors have been overly optimistic about the U.S. fiscal situation. Suppose investors evaluate Eq. (3) using their own subjective expectation, denoted by the operator $\mathbb{F}$, which may differ from the conditional expectation operator $\mathbb{E}$:

$$v_t = \mathbb{F}_t \sum_{j=1}^{T} \left(s_{t+j} - \widetilde{r}_{t+j}\right) + \mathbb{F}_t v_{t+T}.$$

When we, as econometricians, parse out the variations in the debt/GDP ratio under the objec-

Figure 6: Variance Decomposition of Log Debt/Output Ratio with Break



This figure reports the variance decomposition of the transitory component of the log debt/output ratio $v_t - \bar{v}_t$ with a lower pre-2007 sample mean ($\bar{v}_t, t < 2007$) and a higher post-2007 sample mean ($\bar{v}_t, t \geq 2007$) into components due to future government surpluses $\sum_{j=1}^{T} s_{t+j}$, future discount rates $\sum_{j=1}^{T} \widetilde{r}_{t+k}$, future log debt/output ratio $v_{t+T} - \bar{v}_{t+T}$, and the combination of the future government surpluses and discount rates. Sample is annual, 1947—2020. We impose a structural break in the log debt/output ratio by shifting its average level before and after 2007. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of $(r_t, \widetilde{v}_t)$ on two lags of $\widetilde{v}_t$. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

tive measure, we have

$$v_t = \mathbb{E}_t \sum_{j=1}^{T} \left(s_{t+j} - \widetilde{r}_{t+j}\right) + \mathbb{E}_t v_T.$$

Comparing the two equations, we obtain

$$\mathbb{E}_t v_T = \mathbb{F}_t v_{t+T} + (\mathbb{F}_t - \mathbb{E}_t) \sum_{j=1}^{T} (s_{t+j} - \widetilde{r}_{t+j}),$$

where $(\mathbb{F}_t - \mathbb{E}_t) \sum_{j=1}^{T} (s_{t+j} - \widetilde{r}_{t+j})$ denotes the investors' forecast errors about fundamentals (including government surpluses and debt returns). As a result, it is possbile that the debt/GDP ratio is explained only by the fundamentals under the subjective measure, i.e., $\lim_{T\to\infty} cov(v_t, \mathbb{F}_t v_{t+T}) = 0$, while the econometricians find the opposite, i.e., $\lim_{T\to\infty} cov(v_t, \mathbb{E}_t v_{t+T}) > 0$. In this setting, *what the econometricians find as a unit root in the debt/GDP process can be interpreted as the investors' forecast errors*.

We conjecture that government bond market forecasters turn to the CBO's detailed and sophisticated projections to make their forecasts about future debt/GDP. While the CBO forecasts GDP, inflation, and interest rates, it makes projections of future surpluses based on current law. There is a remarkably large and consistently negative bias in the debt/GDP projections, even well after the GFC. Note that if the CBO had assumed future fiscal corrections when making its debt/output projections in the past, its projections would have been even farther from the realized debt/output ratios.

Panel A of Figure 7 plots the CBO 10-year projections $\mathbb{F}_t v_{t+T}$ for the debt/GDP ratio against the actual time series. This bias arises mainly because the CBO has been over-predicting future surpluses as a fraction of GDP, as shown in panel B. As a result, the CBO projections consistently impute too much mean-reversion to the debt/output ratio. The GFC and the COVID-19 pandemic are large unanticipated shocks with a huge fiscal impact, but these projections remain biased throughout the GFC and even well after the GFC. Given the size and persistence of the CBO projection errors, it is conceivable that private investors have similarly biased forecasts of future surpluses.

We reconstruct the Campbell-Shiller decomposition of the debt/output ratio using the CBO projections. We compute the projected $s_{t+j} = sy_{t+j}/e^{v}$ adjusted government surplus/output ratio using the CBO projections. For v, we use average log debt/output ratio between 2007 and 2019. We then back out the implied future discount rates $\sum_{j=1}^{T} \widetilde{r}_{t+k}$, for $T = 10$ from the projected v_{t+T} and the projected sum of the surpluses.

Consider the benchmark of no predictability under the real measure: $cov(\sum_{j=1}^{T} \left(s_{t+j} - \widetilde{r}_{t+j}\right), v_t) = 0$. As a result, in this case, the debt/output ratio has a unit root. Below, we compute sample moments:

$$\hat{cov}_N(v_t, v_{t+T}) = \hat{cov}_N\left(v_t, \mathbb{F}_t v_{t+T}\right) + \hat{cov}_N\left(v_t, (\mathbb{F}_t - \mathbb{E}_t)\sum_{j=1}^{T} s_{t+j}\right) = \hat{var}_N(v_t),$$

where we assumed (i) $\mathbb{F}_t \sum_{j=1}^{T} \widetilde{r}_{t+j}$ is constant, and (ii) $cov(\sum_{j=1}^{T} \left(s_{t+j} - \widetilde{r}_{t+j}\right), v_t) = 0$, implying that in sample $\hat{cov}_N(v_t, v_{t+T}) = \hat{var}_N(v_t)$. Assumption (ii) implies a unit root under the real measure. Assumption (i) implies no predictability under the subjective measure. To explain our findings that the terminal value accounts for most of the variation, without any predictability under the real measure, we need a large positive covariance between surplus forecast errors on the part of bond market investors and the debt/output ratio. As the fiscal situation worsened, we find that the CBO projections persistently over-projected surpluses and under-projected deflated debt returns. Bond market investors were presumably pricing in a much larger surpluses than realized throughout the GFC.

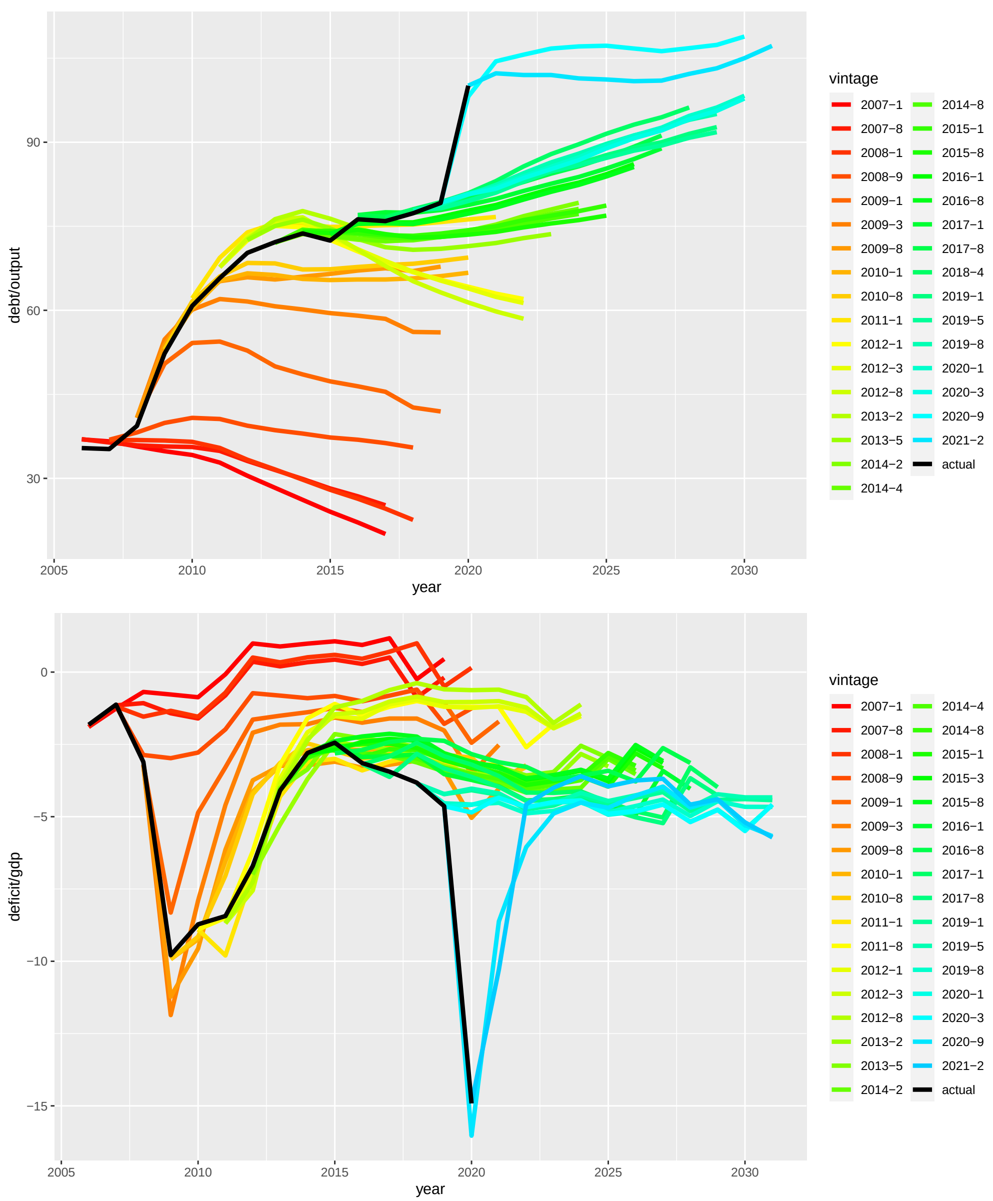
Panel A in Figure 8 plots the entire sample for the 10-year sum of realized surpluses s_t and deflated returns, $\tilde{r}_t$. Panel B focusses on the 2007—2020 sample. The realized surplus series declines steeply as the GFC enters the forward-looking sum in 1997—1998. Just before the start of the GFC, the CBO projections were consistently too optimistic: $(\mathbb{F}_t - \mathbb{E}_t) \sum_{j=1}^{T} s_{t+j} \gg 0$. In 2007, at the start of the GFC, the CBO projected a positive 10-year surplus of 29 log points or roughly 2.9% per annum, shown in the top left graph of Panel B in Figure 8. The units on the y-axis are log points of output. The 2007 projection was one standard deviation above the average of realized surpluses over the entire U.S. post-war history. The actual realized number was -82 log points, or -8.2% per annum. There is a 111 log points forecast error, or 11.1% per annum. The sign of the gap is not surprising, but the size is. Throughout the GFC, the projections were much too optimistic. In 2008, there is still a 108 log points gap between the projected 10 year surplus and the realized one. The gap shrinks only slowly. In 2010, well after the end of the GFC, there is still a 68 log points or 6.8% per annum gap between projected and realized surpluses.[22]

The bottom left graph in panel B plots the projected $S - R$, the projected decrease in the log debt/output ratio over the next 10 years. There is a total 135 log points gap in 2007 between the projected and realized fundamentals. As a result, the realized debt/output ratio in 2017 was 145 log points higher than projected in 2007. In 2007, the CBO projected a more than 50 log points decline in the debt/output ratio. By 2009, it still projected a constant debt/output ratio over the next 10 years.

Our empirical evidence is certainly consistent with mis-pricing. The CBO seems to overestimates future surpluses when the economy enters a recession, and the debt/output ratio increases. This evidence is consistent with a large positive covariance between forecast errors and the debt. If investors' expectations were aligned with CBO projections, they think they live in a world

[22]Similarly, the return projections were too optimistic. The CBO-projected deflated returns over the next 10 years were -32 log points or -3.2% per annum in 2007, shown in the top right figure of Panel B in Figure 8, but the realized returns were 1.3 log points, mainly because realized growth was much lower than projected. Hence, there is an additional 22 log points gap contributed by the discount rate components. The CBO projected $r - x - \pi$ to be 2.2 % lower per annum than realized. However, the CBO has significantly raised its projections for the effective cost of government funding since 2010.

Figure 7: CBO Projections for Debt/GDP and Surplus/GDP

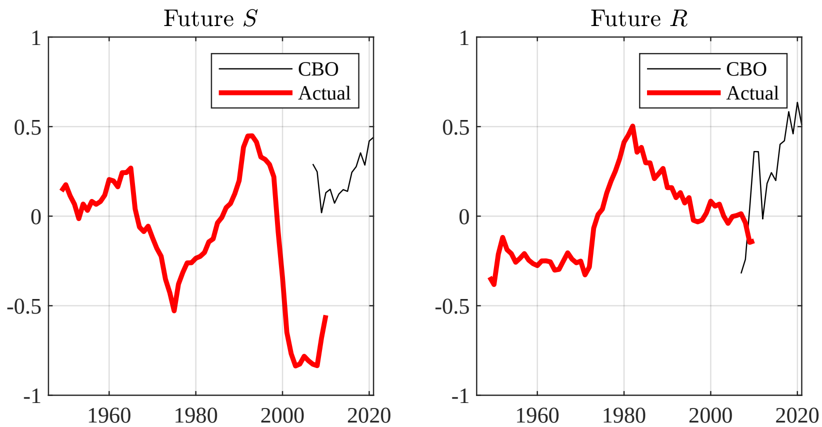


This figure reports the CBO 10-year U.S. federal government debt/GDP and total federal government surplus/GDP projections for each forecast date, plotted against the actual debt/GDP and surplus GDP/ratio.

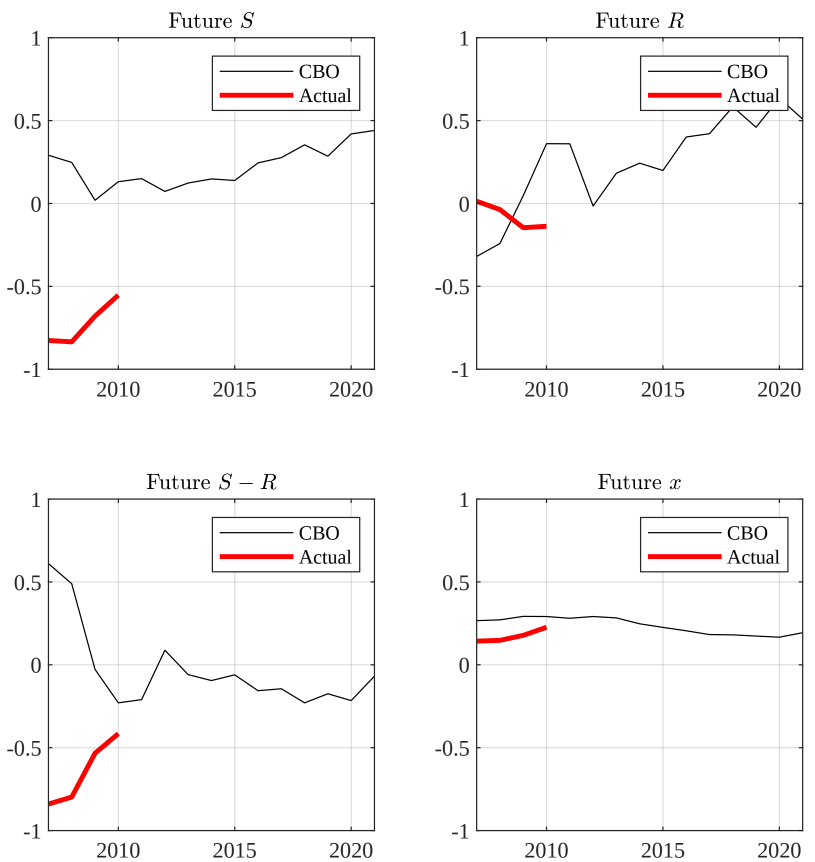
without permanent shocks to debt/output, but their forecast errors induce permanent shocks in debt/output, at least from an econometrician's perspective.

Figure 8: Decomposition for CBO Projections

Panel A: 1948—2020



Panel B: 2007—2020



Panel A and B show the decomposition of the log debt/output ratio v_t into components due to CBO-projected (and realized) future government surpluses $\sum_{j=1}^{T} s_{t+j}$, future discount rates $\sum_{j=1}^{T} \tilde{r}_{t+k}$, for $T = 10$. We also report future real growth $\sum_{j=1}^{T} \tilde{x}_{t+k}$.

5 Conclusion

The U.S. bond market's valuation of a claim to U.S. surpluses is surprisingly insensitive to news about future surpluses or returns, at least from the perspective of an econometrician. This is a direct result of the debt/output ratio's persistence. To be clear, we cannot claim to have shown that U.S. is on an unsustainable fiscal path, or that the debt/output ratio is a unit-root process. What we have shown is that there is no statistical evidence to conclude that the U.S. is on a fiscally sustainable path.

Our model-free exercise suggests that the bond market's assessment of future surpluses may diverge from the econometrician's. The econometrician does not forecast larger surpluses or lower discount rates when the debt/output ratio rises, not based on the entire U.S. sample, but, obviously, the bond market investor does, unless there is a violation of the no-bubble condition. We show evidence from CBO projections that investors may have been over-predicting surpluses and under-predicting returns on debt throughout and well after the GFC. As a result, bond investors anticipated mean reversion in the debt/output ratio that failed to materialize.

Even though the pricing of individual bonds does not allow for arbitrage opportunities, except in times of market disruptions, the U.S. federal government debt as a whole may be persistently mis-priced. It is plausible that limits to arbitrage (Shleifer and Vishny, 1997) would prevent rational arbitrageurs from taking advantage of potential over-pricing in the last decades. Treasurys benefit from safe asset demand. Safe assets tend to appreciate in high marginal utility states of the world, when arbitrageurs are more likely to be constrained.

Of course, we cannot definitely rule out the possibility of a peso-event, a large fiscal correction, not observed in our sample, but anticipated and priced in by investors. However, even the most sophisticated budget projections which abstract from future fiscal corrections have been far too optimistic over the past decades.

References

Barro, R. J., 2020, "r Minus g," NBER Working Paper 28002.

Bassetto, M., and W. Cui, 2018, "The Fiscal Theory of the Price Level in a World of Low Interest Rates," *Joural of Econonmic Dynamics and Control*, 89, 5–22.

Bauer, M. D., and J. D. Hamilton, 2017, "Robust Bond Risk Premia," *Review of Financial Studies*, 31(2), 399–448.

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

Berndt, A., and Ş. Yeltekin, 2015, "Monetary Policy, Bond Returns and Debt Dynamics," *Journal of Monetary Economics*, 73, 119–136.

Binsbergen, J. H. V., and R. S. J. Koijen, 2010, "Predictive Regressions: A Present-Value Approach," *Journal of Finance*, 65(4), 1439–1471.

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

Blanchard, O. J., and M. Watson, 1982, "Bubbles, Rational Expectations, and Financial Markets," in *Crises in the Economic and Financial Structure*, ed. by P. Wachtel. Lexington Books.

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

———, 2007, "Are Stationarity and Cointegration Restrictions Really Necessary for the Intertemporal Budget Constraint?," *Journal of Monetary Economics*, 54(7), 1837 – 1847.

Boudoukh, J., R. Israel, and M. P. Richardson, 2020, "Biases in Long-Horizon Predictive Regressions," NBER Working Paper 27410.

Brunnermeier, M., S. Merkel, and Y. Sannikov, 2020, "The Fiscal Theory of Price Level with a Bubble," NBER Working Paper 27116.

Campbell, J. Y., and R. J. Shiller, 1988, "The Dividend-Price Ratio and Expectations of Future Dividends and Discount Factors," *Review of Financial Studies*, 1, 195–227.

———, 1991, "Yield Spreads and Interest Rate Movements: A Bird's Eye View," *Review of Economic Studies*, 58, 495–514.

Campbell, J. Y., and S. B. Thompson, 2007, "Predicting Excess Stock Returns Out of Sample: Can Anything Beat the Historical Average?," *Review of Financial Studies*, 21(4), 1509–1531.

Campbell, J. Y., and M. Yogo, 2006, "Efficient Tests of Stock Return Predictability," *Journal of Financial Economics*, 81, 27–60.

Chernov, M., L. Schmid, and A. Schneider, 2020, "A Macrofinance View of U.S. Sovereign CDS Premiums," *Journal of Finance*, 75(5), 2809–2844.

Chien, Y., and Y. Wen, 2019, "The Determination of Public Debt under Both Aggregate and Idiosyncratic Uncertainty," Federal Reserve Bank of St. Louis Working Paper.

——— , 2020, "Time-Inconsistent Optimal Quantity of Debt," Federal Reserve Bank of St. Louis Working Paper.

Cieslak, A., 2018, "Short-Rate Expectations and Unexpected Returns in Treasury Bonds," *Review of Financial Studies*, 31(9), 3265–3306.

Cochrane, J. H., 2008, "The Dog That Did Not Bark: A Defense of Return Predictability," *Review of Financial Studies*, 21(4), 1533–1575.

——— , 2011, "Presidential Address: Discount Rates," *Journal of Finance*, 66(4), 1047–1108.

——— , 2020, "The Surplus Process," Accessed: 2020-10-24.

Cochrane, J. H., 2021a, "The Fiscal Roots of Inflation," *Review of Economic Dynamics*.

——— , 2021b, "A Fiscal Theory of Monetary Policy with Partially-Repaid Long-Term Debt," *Review of Economic Dynamics*.

Cochrane, J. H., and M. Piazzesi, 2005, "Bond Risk Premia," *American Economic Review*, 95(1), 138–160.

Cooper, I., and R. Priestley, 2008, "Time-Varying Risk Premiums and the Output Gap," *Review of Financial Studies*, 22(7), 2801–2833.

Croce, M., T. T. Nguyen, and S. Raymond, 2021, "Persistent government debt and aggregate risk distribution," *Journal of Financial Economics*, 140(2), 347–367.

Croce, M. M., T. T. Nguyen, and L. Schmid, 2012, "The market price of fiscal uncertainty," *Journal of Monetary Economics*, 59(5), 401–416, Carnegie-NYU-Rochester Conference Series on Public Policy - Robust Macroeconomic Policy at Carnegie Mellon University on November 11-12, 2011.

de La O, R., and S. Myers, 2021, "Subjective Cash Flow and Discount Rate Expectations," *Journal of Finance*, (jofi.13016).

Diamond, P. A., 1965, "National Debt in a Neoclassical Growth Model," *American Economic Review*, 55(5), 1126–1150.

Duffee, G. R., 2011, "Information in (and not in) the Term Structure," *Review of Financial Studies*, 24(9), 2895.

Fama, E. F., and R. H. Bliss, 1987, "The Information in Long-Maturity Forward Rates," *American Economic Review*, 77(4), 680–692.

Furman, J., and L. Summers, 2020, "A Reconsideration of Fiscal Policy in the Era of Low Interest Rates," https://www.brookings.edu/wp-content/uploads/2020/11/furman-summers-fiscal-reconsideration-discussion-draft.pdf, Accessed: 2020-12-27.

Giglio, S., M. Maggiori, and J. Stroebel, 2016, "No-Bubble Condition: Model-Free Tests in Housing Markets," *Econometrica*, 84(3), 1047–1091.

Golez, B., and P. Koudijs, 2018, "Four Centuries of Return Predictability," *Journal of Financial Economics*, 127(2), 248–263.

Gourinchas, and Rey, 2007, "International Financial Adjustment," *Journal of Political Economy*, 115(4), 665–703.

Hall, G., J. Payne, and T. J. Sargent, 2018, "US Federal Debt 1776-1960: Quantities and Prices," *Manuscript*.

Hall, G. J., and T. J. Sargent, 2011, "Interest Rate Risk and Other Determinants of Post-WWII US Government Debt/GDP Dynamics," *American Economic Journal: Macroeconomics*, 3(3), 192–214.

———, 2021, "Debt and Taxes in Eight U.S. Wars and Two Insurrections," *The Handbook of Historical Economics*.

Hamilton, J. D., 1994, *Time Series Analysis*. Princeton University Press, Princeton, NY.

Hamilton, J. D., and M. A. Flavin, 1986, "On the Limitations of Government Borrowing: A Framework for Empirical Testing," *American Economic Review*, 76(4), 808–819.

Hansen, L. P., and R. Jagannathan, 1991, "Restrictions on Intertemporal Marginal Rates of Substitution Implied by Asset Returns," *Journal of Political Economy*, 99, 225–262.

Hansen, L. P., W. Roberds, and T. J. Sargent, 1991, *Time Series Implications of Present Value Budget Balance*Westview Press, vol. Rational Expectations Econometrics, chap. 5.

Hansen, L. P., and T. J. Sargent, 1980, "Formulating and Estimating Dynamic Linear Rational Expectations Models," *Journal of Economic Dynamics and Control*, 2, 7–46.

Hellwig, C., and G. Lorenzoni, 2009, "Bubbles and Self-Enforcing Debt," *Econometrica*, 77(4), 1137–1164.

Jiang, Z., A. Krishnamurthy, and H. Lustig, 2022, "The Rest of the World's Dollar-Weighted Return on U.S. Treasurys," .

Jiang, Z., H. Lustig, S. Van Nieuwerburgh, and M. Z. Xiaolan, 2019, "The U.S. Public Debt Valuation Puzzle," NBER Working Paper 26583.

Jiang, Z., H. Lustig, S. Van Nieuwerburgh, and M. Z. Xiaolan, 2020, "Manufacturing Risk-free Government Debt," NBER Working Paper 27786.

Joslin, S., M. Priebsch, and K. J. Singleton, 2014, "Risk Premiums in Dynamic Term Structure Models with Unspanned Macro Risks," *Journal of Finance*, 69(3), 1197–1233.

Kocherlakota, N., 2021, "Public Debt Bubbles in Heterogeneous Agent Models with Tail Risk," *University of Rochester Manuscript*.

Koijen, R., and S. V. Nieuwerburgh, 2011, "Predictability of Returns and Cash Flows," *Annual Review of Financial Economics*, 3(1), 467–491.

LeRoy, S. F., and R. D. Porter, 1981, "The Present-Value Relation: Tests Based on Implied Variance Bounds," *Econometrica*, 49(3), 555–574.

Lettau, M., and S. Van Nieuwerburgh, 2008, "Reconciling the Return Predictability Evidence," *Review of Financial Studies*, 21(4), 1607–1652.

Lewellen, J., 2004, "Predicting Returns with Financial Ratios," *Journal of Financial Economics*, 74(2), 209–235.

Lo, A. W., and A. MacKinlay, 1989, "The size and power of the variance ratio test in finite samples: A Monte Carlo investigation," *Journal of Econometrics*, 40(2), 203–238.

Ludvigson, S. C., and S. Ng, 2009, "Macro Factors in Bond Risk Premia," *Review of Financial Studies*, 22(12), 5027–5067.

Mian, A., L. Straub, and A. Sufi, 2021a, "A Goldilocks Theory of Fiscal Policy," Working Paper.

——— , 2021b, "A Goldilocks Theory of Fiscal Policy," NBER Working Paper No. 29351.

Nelson, C. R., and M. J. Kim, 1993, "Predictable Stock Returns: The Role of Small Sample Bias," *Journal of Finance*, 48(2), 641–661.

Piazzesi, M., J. Salomao, and M. Schneider, 2015, "Trend and Cycle in Bond Premia," Working Paper.

Reis, R., 2021, “The Constraint on Public Debt when $r < g$ but $g < m$,” London School of Economics Working Paper.

Renxuan, W., 2020, “Subjective Return Expectations,” *Available at SSRN 3740792*.

Samuelson, P. A., 1958, “An Exact Consumption-Loan Model of Interest with or without the Social Contrivance of Money,” *Journal of Political Economy*, 66(6), 467–482.

Sargent, T. J., 2012, “Nobel Lecture: United States Then, Europe Now,” *Journal of Political Economy*, 120(1).

Schwert, G., 1987, “Effects of model specification on tests for unit roots in macroeconomic data,” *Journal of Monetary Economics*, 20(1), 73–103.

Shiller, R. J., 1981, “Do Stock Prices Move Too Much to be Justified by Subsequent Changes in Dividends?,” *American Economic Review*, 71, 421–436.

Shleifer, A., and R. W. Vishny, 1997, “The Limits of Arbitrage,” *Journal of Finance*, 52(1), 35–55.

Smith, S. C., and A. Timmermann, 2020, “Break Risk,” *Review of Financial Studies*, 34(4), 2045–2100.

Stambaugh, R. F., 1999, “Predictive Regressions,” *Journal of Financial Economics*, 54, 375–421.

Torous, W., R. Valkanov, and S. Yan, 2004, “On Predicting Stock Returns with Nearly Integrated Explanatory Variables,” *The Journal of Business*, 77(4), 937–966.

Trehan, B., and C. E. Walsh, 1988, “Common Trends, the Government’s Budget Constraint, and Revenue Smoothing,” *Journal of Economic Dynamics and Control*, 12(2), 425–444.

——— , 1991, “Testing Intertemporal Budget Constraints: Theory and Applications to U. S. Federal Budget and Current Account Deficits,” *Journal of Money, Credit and Banking*, 23(2), 206–223.

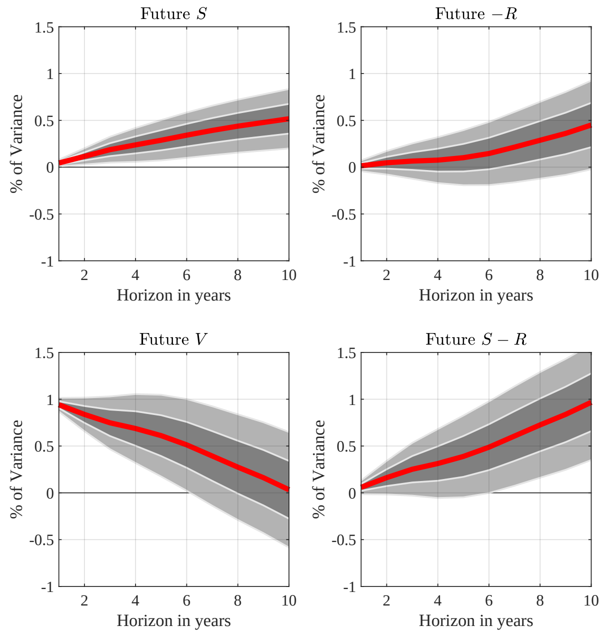
Welch, I., and A. Goyal, 2008, “A Comprehensive Look at The Empirical Performance of Equity Premium Prediction,” *Review of Financial Studies*, 21(4), 1455–1508.

Appendix

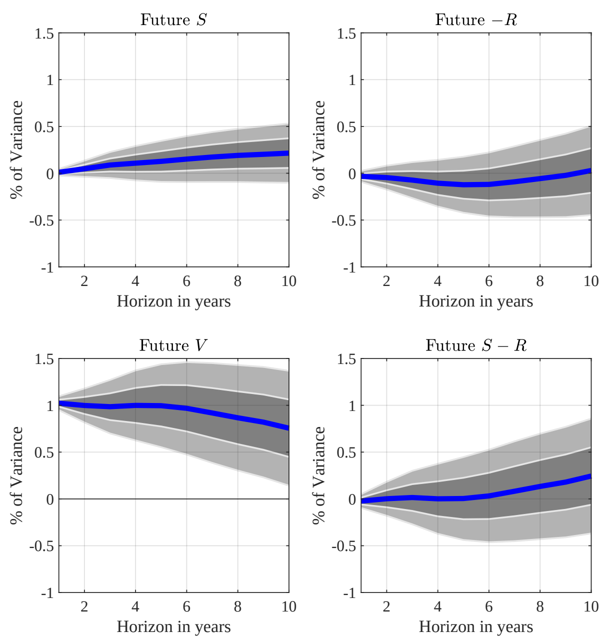
A Robustness Tables

Figure A1: Variance Decomposition of Log Debt/Output Ratio: Shorter Bohn Sample 1948—1995

Panel A: Before Bias Correction



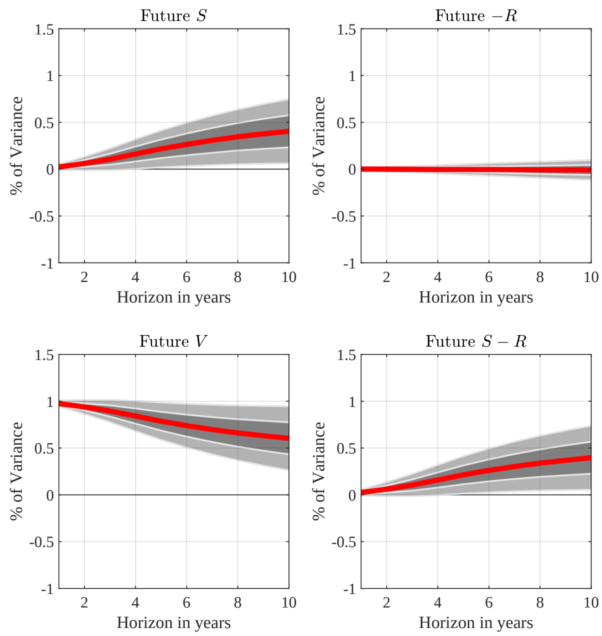
Panel B: After Bias Correction



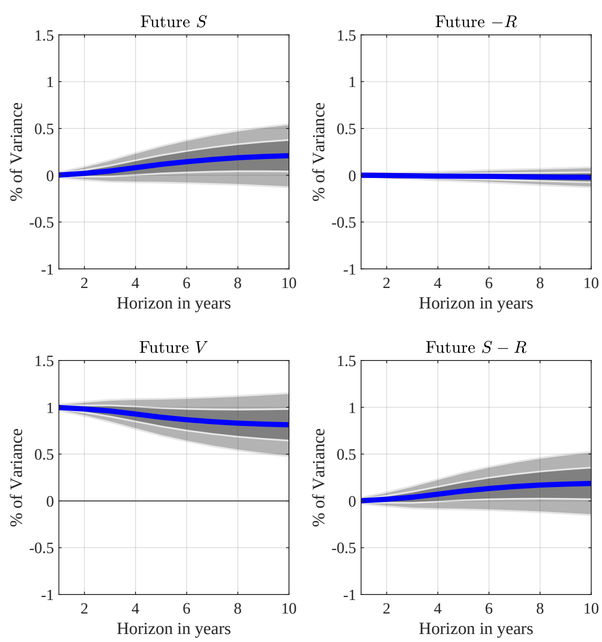
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio v_{t+T} and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1948—1995. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

Figure A2: Variance Decomposition of Log Debt/Output Ratio: Longer Sample 1842—2020

Panel A: Before Bias Correction



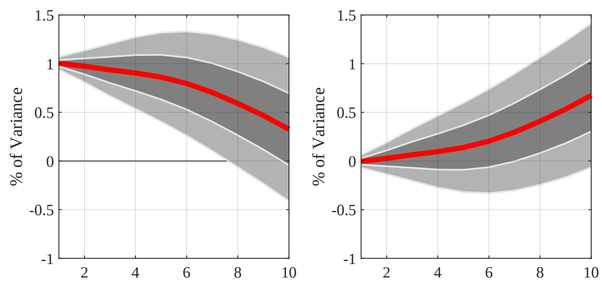
Panel B: After Bias Correction



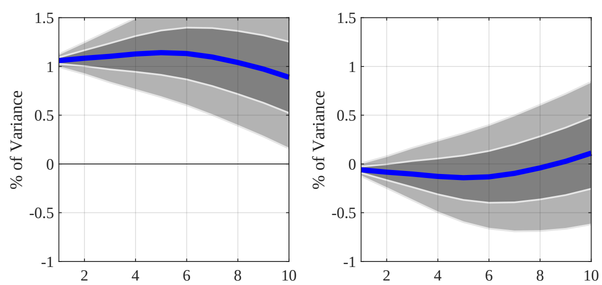
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1842—2020. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

Figure A3: Variance Decomposition of Log Debt/Output Ratio: with $\rho = \exp(1\%)$

Panel A: Before Bias Correction



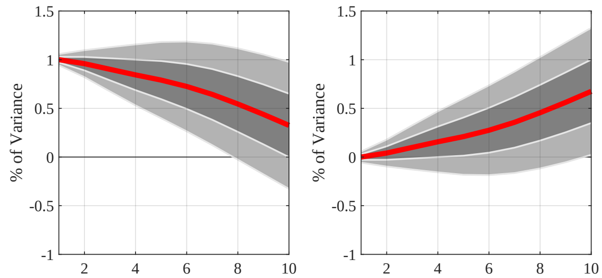
Panel B: After Bias Correction



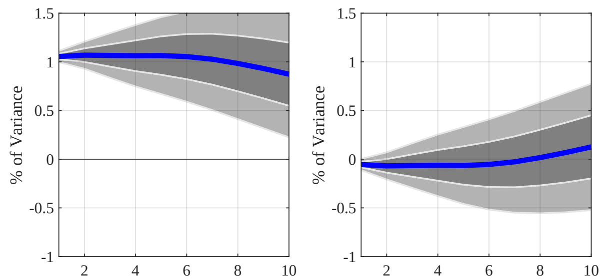
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1947—2020. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

Figure A4: Variance Decomposition of Log Debt/Output Ratio: with $\rho = \exp(-1\%)$

Panel A: Before Bias Correction

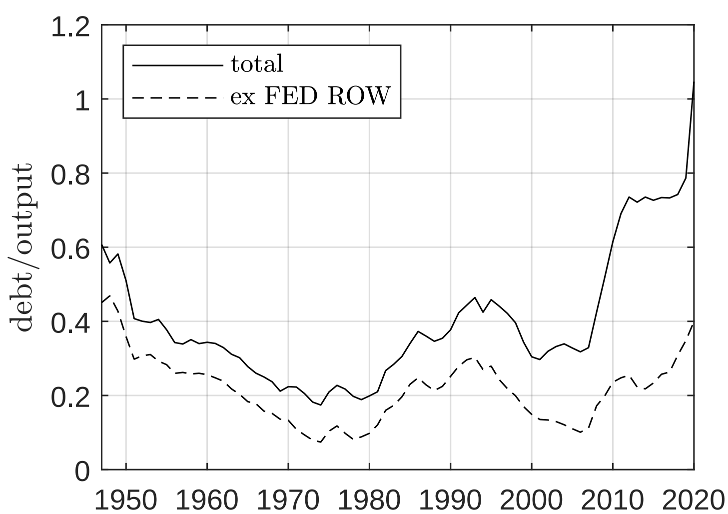


Panel B: After Bias Correction



This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1947—2020. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

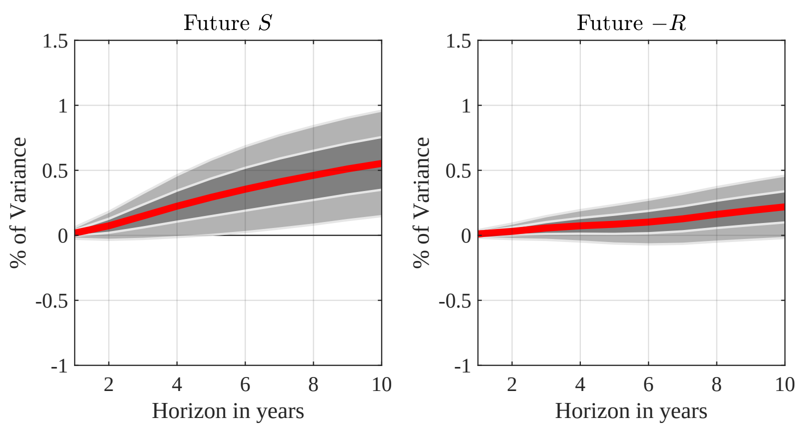
Figure A5: Debt/Output Ratio



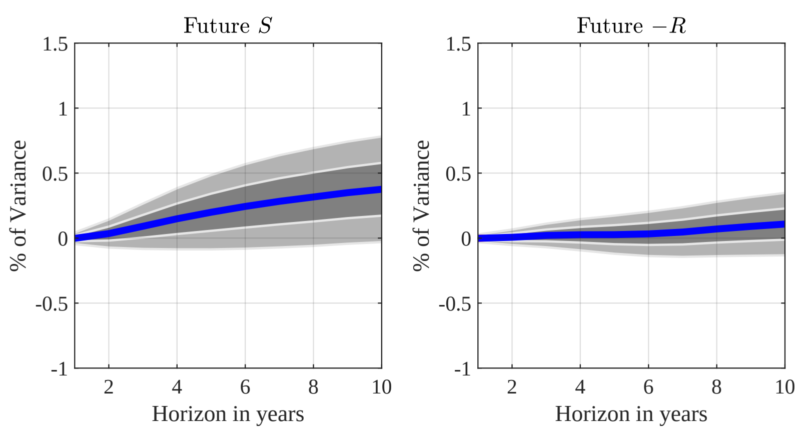
The full line is the debt/output ratio. The dashed line is the domestically and privately held debt/output ratio.

Figure A6: Forecasting Power of Log Debt/Output Ratio Held by the Domestic Sector

Panel A: Before Bias Correction



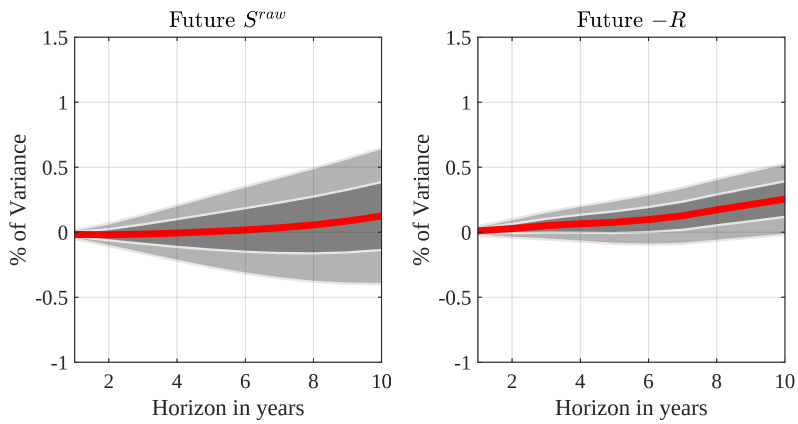
Panel B: After Bias Correction



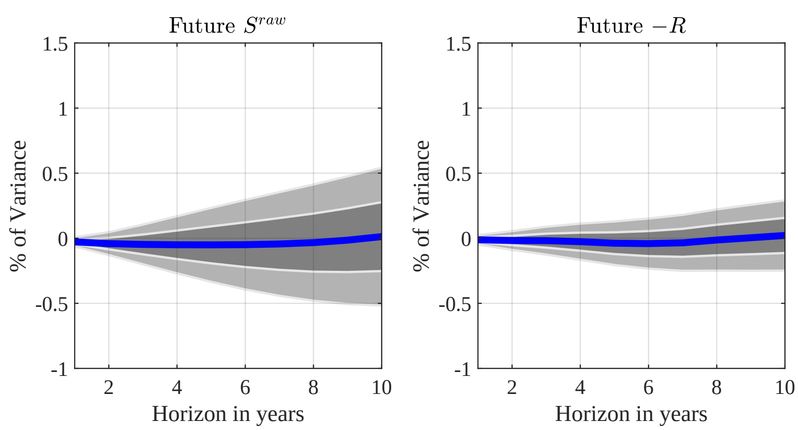
This figure reports regression coefficients associated with the log domesticly held debt/output ratio. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1947—2020. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

Figure A7: Forecasting Power of Log Debt/Output Ratio for Actual Government Surplus

Panel A: Before Bias Correction



Panel B: After Bias Correction



This figure reports regression coefficients associated with the log debt/output ratio. S^{raw} is based on the actual government surplus/GDP ratio. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS coefficients in long-horizon predictability regressions. Sample is annual, 1947—2020. Standard errors are generated by bootstrapping 10,000 draws from the joint residuals in a regression of (r_t, v_t) on two lags of v_t. We plot 1 s.e. (dark) and 2 s.e. (light) CIs.

B Mean Reversion of Returns and Surpluses

We start from the log-linearized return equation implied by the government budget constraint:

$$r_{t+1} - \pi_{t+1} - x_{t+1} = \rho v_{t+1} - v_t + s_{t+1},$$

where ρ is a constant of linearization. We will choose $\rho = \exp(-(r - x - \pi)) = 1$. We use $\rho = 1$. Over longer horizons k, the cumulative log return equals the change in the log of the debt/output ratio less the cumulative surplus over this horizon:

$$\tilde{r}_{t\to t+k} = \Delta v_{t\to t+k} - s_{t\to t+k}.$$

By taking variances on both sides, dividing by k and then taking the limit of the horizon to ∞, we obtain the following expression for the per period variance of the cumulative log returns:

$$\lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}] = \lim_{k\to\infty} \frac{1}{k} var[\Delta v_{t\to t+k}] + \lim_{k\to\infty} \frac{1}{k} var[s_{t\to t+k}] - 2 \lim_{k\to\infty} \frac{1}{k} cov(\Delta v_{t\to t+k}, s_{t\to t+k}).$$

If we impose that debt/output ratio s_t is stationary, then we end up with the implication that the variance of cumulative returns converges to the variance of cumulative surpluses:

$$\lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}] = \lim_{k\to\infty} \frac{1}{k} var[s_{t\to t+k}], \tag{A1}$$

where we have used that $\lim_{k\to\infty} \frac{1}{k} var[\Delta v_{t\to t+k}] = 0$ and $\lim_{k\to\infty} \frac{1}{k} cov(\Delta v_{t\to t+k}, s_{t\to t+k}) = 0$.
First, if the returns are i.i.d., then the variance ratio converges to one at long horizons:

$$\lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}] = var[\tilde{r}_t]. \tag{A2}$$

The variance of the cumulative surpluses per period converges to the one-period variance of returns:

$$var[\tilde{r}_t] = \lim_{k\to\infty} \frac{1}{k} var[s_{t\to t+k}].$$

Suppose $var[\tilde{r}_t] < var[s_t]$. Then it follows from the fact that the variance ratio of returns converges to one (see eqn. (A3)), and from the fact that the variance of returns and surpluses converge over long horizons (see eqn. (A1)), that the variance ratio of surpluses is smaller than one:

$$\lim_{k\to\infty} \frac{1}{k var[\tilde{r}_t]} var[\tilde{r}_{t\to t+k}] = 1 > \lim_{k\to\infty} \frac{1}{k var[s_t]} var[s_{t\to t+k}].$$

There is mean reversion in the surpluses when the debt/output ratio is stationary and the surpluses are more volatile than the returns. Surpluses are predictable.
Second, if the surpluses are i.i.d., then the variance ratio of surpluses converges to one at long horizons:

$$\lim_{k\to\infty} \frac{1}{k} var[s_{t\to t+k}] = var[s_t], \tag{A3}$$

which implies the following equality:

$$var[s_t] = \lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}].$$

If $var[\tilde{r}_t] < var[s_t]$, then we know from the fact that the variance ratio of returns converges to one (see eqn. (A3)), and from the fact that the variance of returns and surpluses converge over long horizons (see eqn. (A1)), that the variance ratio of surpluses is smaller than one:

$$\lim_{k\to\infty} \frac{1}{k var[\tilde{r}_t]} var[\tilde{r}_{t\to t+k}] > 1 = \lim_{k\to\infty} \frac{1}{k var[s_t]} var[s_{t\to t+k}].$$

This implies that there is mean aversion in the returns when the debt/output ratio is stationary, and the surpluses are i.i.d.