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, Max Croce, Francois Gourio, James Hamilton, Lars Hansen, 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.

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 forecast 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 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 (lower nominal returns, growth, or inflation) 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) 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.

Our approach delivers a model-free test of fiscal sustainability defined as mean reversion in the debt/output ratio driven by adjustments in future returns or surpluses. At horizons of up to 10 years, our variance decomposition attributes only 8% of the variation in the current debt/output ratio in the post-war U.S. sample to variation in either future surpluses or future returns. Rather, the future debt/output ratio 10 years from now accounts for 92% of the variation in today's debt/output ratio.

First, higher debt/output does not predict lower real growth-adjusted returns on U.S. government debt. Discount rates do not explain variation in the U.S. debt/output ratio. Second, a higher debt/output ratio does not forecast higher 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. The confidence intervals for the cash flow component are wider at longer horizons, implying that we can neither prove nor disprove some role for cash flows.

We implement a forward-looking decomposition of the government debt/output ratio which splits the market value of debt into components due to future returns and future surpluses.[2] In earlier work, Hall and Sargent (2011); Berndt and Yeltekin (2015); Hall and Sargent (2021) implement a backward-looking decomposition of the variation in the U.S. debt/GDP ratio, imputing

[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]See also Berndt, Lustig, and Yeltekin (2012); Cochrane (2021a). In earlier work, Gourinchas and Rey (2007) decompose the variation in the U.S. net foreign asset position using the same approach.

changes in the ratio to contemporaneously realized inflation, growth and returns. Our ex-ante approach provides a different and complementary perspective.

Our findings are also different from those in the literature.[3] Studying a shorter sample that ends in the mid-1990s, Bohn (1998) finds evidence that the primary surplus increases when the debt/output ratio is high. Cochrane (2021a,b) finds evidence that the debt/output ratio predicts nominal returns on the government debt portfolio. We find no evidence that the debt/output ratio predicts surpluses nor real growth-adjusted returns over the full post-WW-II sample. Results from a longer sample that starts in 1842 confirm the post-WW-II results. Bond prices today are insufficiently responsive to news about future macro fundamentals, i.e., government surpluses and real growth-adjusted returns.

We reach a different conclusion because of a statistical inference challenge that plagues the fiscal sustainability literature. This issue is well known in the asset pricing literature (Stambaugh, 1999), and at least as relevant in our context. Specifically, there is a small-sample bias in the slope coefficients of the return and surplus predictability regressions, leading one to overstate the degree of predictability. The issue arises because of the combination of (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 (surpluses and returns). An increase in bond risk premia will tend to lower realized returns, the dependent variable, and simultaneously lower the market value of debt to 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, violating the classical OLS assumptions. 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. The unbiased coefficients display substantially weaker surplus and return predictability than what is indicated by the OLS estimates.

This evidence indicates that the debt/output ratio is subject to permanent innovations, inconsistent with fiscal sustainability. As corroborating evidence, we simulate data from a model under the null hypothesis that the debt/output ratio has a unit root, i.e., a model where there is no role for fundamentals. This model exactly replicates our empirical findings. We find spurious evidence

[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). Sargent (2012) and D'Erasmo, Mendoza, and Zhang (2016) provide comprehensive reviews of the literature. This literature largely sidesteps the issue of discount rate variation by assuming constant discount rates.

of surplus and return predictability using unadjusted OLS slope estimates, and bias-corrected estimates that recover the true values indicating no predictability.[4]

However, we cannot definitively conclude from this evidence that U.S. is on a fiscally unsustainable path. There is a long literature on the low power of unit root tests in short samples against close-to-unit-root alternatives (Schwert, 1987; Lo and MacKinlay, 1989, in finance). Even in the sample that starts in 1842, we lack the power to rule out that the debt/output ratio is (slowly) mean-reverting. Given the higher short-run volatility of surpluses than of Treasury returns, what we can say is that, if there is mean reversion in the debt/output ratio, then debt adjustment is more likely to occur via adjustment of surpluses than via adjustment of real growth-adjusted returns.[5] Our findings are consistent with this view. While the statistical evidence against the discount rate channel is strong at all horizons, the confidence intervals for the surplus predictability coefficients are much wider. Hence, 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. The longer sample results in somewhat tighter confidence intervals on the long-horizon surplus predictability coefficient.

One type of permanent shocks to the debt/GDP ratio are structural breaks. There may have been a structural break in the relation between the debt/output ratio, bond returns, and surpluses. In stock markets, there also is evidence of structural breaks in the relation between the dividend yield, stock returns, and dividend growth (see Lettau and Van Nieuwerburgh, 2008; Smith and Timmermann, 2020).[6] We estimate a structural break in the debt/output ratio in 2007. Bond investors may have revised their estimate of the long-run debt/output ratio upwards around the start of the Great Financial Crisis (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 means before and after 2007, we find stronger evidence for surplus predictability, 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. Removing a structural break removes a low-frequency component in the debt/output ratio; the resulting transitory component in the debt/GDP ratio is less persistent. The predictive coefficients have much smaller small-sample biases. Of course, this analysis leaves the large, permanent increase in the debt/output ratio (as well as its timing) unexplained.

One candidate explanation for the low-frequency change in the relationship between debt and

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

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

[6]Such structural breaks can lead to instability in the predictive coefficients (Lettau and Van Nieuwerburgh, 2008) and may help to account for the poor out-of-sample performance of these predictors (Goyal and Welch, 2008).

fundamentals is the rise in foreign and Central Bank purchases os U.S. Treasurys. If we use the domestic private sector holdings of debt scaled by GDP as the predictor, we also find larger surplus predictability, similar to the evidence that uses the transitory component of the debt/output ratio as predictor. This suggests that inelastic demand by the Fed and the rest of the world may be a contributing factor to the insensitivity of government bond prices to macro fundamentals.

Mispricing is another candidate explanation for the structural break. 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 under her subjective beliefs. If bond 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 generating a stationary debt/output ratio under the subjective measure. Our paper connects to the growing literature on the role of subjective beliefs in asset pricing (Barberis, 2018; Bordalo, Gennaioli, LaPorta, and Shleifer, 2019; Bordalo, Gennaioli, Ma, and Shleifer, 2020). For the stock market, De La O and Myers (2021); Wang (2020) impute a much larger share of variation in the price/dividend ratio to the cash flow component under subjective than under objective beliefs. Mistakes in interest rate and growth forecasts also contribute to our findings.[7] We show evidence from Congressional Budget Office (CBO) projections that the structural break in the debt/output ratio at the start of the GFC has been partly fueled by persistently biased forecasts of surpluses. These forecasts impute more mean-reversion to the predicted debt/output ratio under the subjective than under the actual measure.[8] After 2007, the CBO consistently overestimated future surpluses and underestimated the government's effective cost of funding. For example, in 2010, which is after the GFC, the CBO projections of 10-year cumulative surpluses were 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). This evidence that

[7] See Piazzesi, Salomao, and Schneider (2015); Cieslak (2018) for evidence on the gap between statistical risk premia measured by econometricians and subjective risk premia expected by investors in the bond market.

[8] The CBO projections do not take into account any future legislative action that could implement a fiscal correction. The source of bond market investor optimism about future surpluses may be the expectation of future fiscal rectitude through such legislation. A comparison of the average private-sector forecast of one-year ahead surpluses from Consensus Economics to the CBO forecast shows that the two forecasts are close in both levels and dynamics.

high valuations revert back to the mean through low subsequent returns is common across most asset classes. Not so for the entire U.S. government bond portfolio. 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 lower returns on the entire bond portfolio push the debt/GDP ratio back to its long-run mean following a high value. The fluctuations in the market value of the entire U.S. government bond portfolio is too small compared to the fluctuations in its fundamentals; there is excess smoothness as opposed to excess volatility in the stock market. The cash flows and the discount rates are the dogs that did not bark when there is 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 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 government bond portfolio is overpriced. 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.

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).[9] We do not rule out non-stationary behavior of the debt/output ratio, because we are interested in testing fiscal sustainability.We implement the Campbell and Yogo (2006) testing procedure, valid under general assumptions, which confirms that there is neither evidence of surplus nor return predictability.

We start in section 1 by deriving and implementing the decomposition of the debt/output ratio. Section 2 does some robustness checks. 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.

[9] 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.

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

We use the log-linear Campbell and Shiller (1988) 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 (cash flows), expected future real growth-adjusted bond returns (discount rates), and the expected future value of the debt/output ratio. After we correct for the Stambaugh (1999) bias, we find that most of the variation in the debt/output ratio cannot be attributed to cash flows or discount rates over the next 10 years.

1.1 Realized Treasury Returns and Surpluses

We compute the market value of all marketable U.S. Treasurys using data from CRSP Treasurys. Following Hall and Sargent (2011), 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.

Let r_{t+j} denote the nominal return on the government debt portfolio (in logs), x_{t+j} real GDP growth (in logs), and π_{t+j} log inflation. Let $\widetilde{r}_{t+j} = r_{t+j} - x_{t+j} - \pi_{t+j}$ denote the adjusted log bond return, which is the nominal return on the government bond portfolio minus output growth and inflation. We refer to $\widetilde{r}$ as the "real growth-adjusted return" or simply the "return" when there is no scope for confusion. Let $s_{t+j} = sy_{t+j}/exp(v)$ 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:

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

where ρ is a constant of linearization. We set $\rho = \exp(-(r - x - \pi)) = 1$. This decomposition expands the debt/output ratio around the unconditional average $r = x + \pi$.[10] 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 evidence for $r < x + \pi$ is concentrated in the first half of the sample. The real growth-adjusted returns on the entire

[10] The details of the derivation are in the appendix of Cochrane (2021a) on pages 3 and 4. We do robustness to different values of ρ in section 2.2.

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

	$\tilde{r}$	r	x	π	$x+\pi$	sy
1950-1959	-3.8%	2.7%	4.1%	2.4%	6.5%	1.6%
1960-1969	-2.8%	3.9%	4.4%	2.3%	6.7%	2.0%
1970-1979	-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%
1980-1989	4.1%	11.8%	3.0%	4.6%	7.6%	-2.2%
1990-1999	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%

bond portfolio, reported in the first column, are strongly negative in the 1950s (-3.8%), due to low real returns and strong economic growth. This pattern continues in the 1960s, with real growth-adjusted returns of -2.6%. In the 1970s, bondholders were surprised by persistently high inflation, delivering growth-adjusted returns of -2.5%. Overall, in the first three decades after WW-II, the average gap between returns and growth $\tilde{r}$ is -3.5%; r was much lower than $x+\pi$. In the next four decades, the average real growth adjusted return $\tilde{r}$ is positive (1.5%); r was higher than $x+\pi$. The 1980s represents a radical departure from what came before. Bondholders earned high real returns despite high inflation (4.1%). Even in the 1990s and 2000s, bondholders earned high real growth-adjusted returns of 1.6% and 0.8%, respectively.

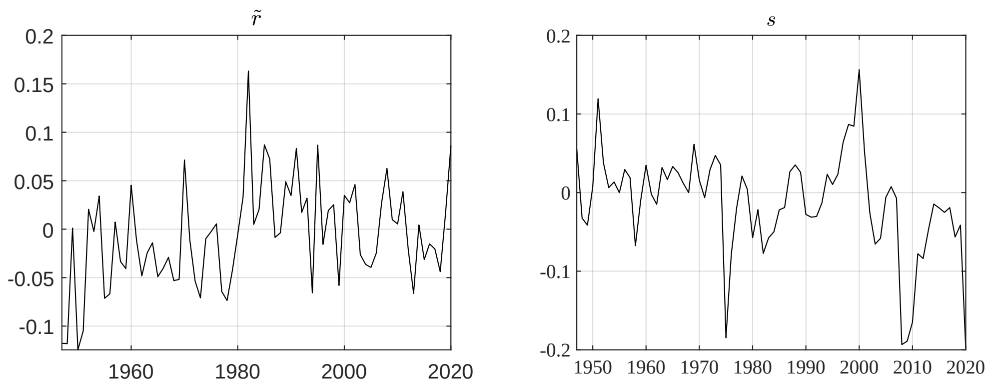
Surpluses, reported in the last column, are modestly positive in the 1950s, 1960s, and 1990s, and negative in the five other decades. Surpluses are most negative over the past twenty years.

Figure 1 plots the variables we want to forecast: the real growth-adjusted returns $\tilde{r}_t$ in the left panel and the surplus s_t in the right panel. As discussed above, returns were low from the start of the sample until the 1980s. After 2000, real growth-adjusted returns declined again. The surpluses s are high in the 1950s and 1990s and decline precipitously in the last two decades. We note that the surpluses s_t, which is backed out from equation (1), has a high correlation of 83% with the raw surplus/output ratio. The implied surplus s_t is more volatile (standard deviation of 6.47%) than the actual surplus/output ratios (standard deviation of 2.6%).

The solid line in Figure 2 is the predictor variable, the demeaned log debt/output ratio. 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 GFC.

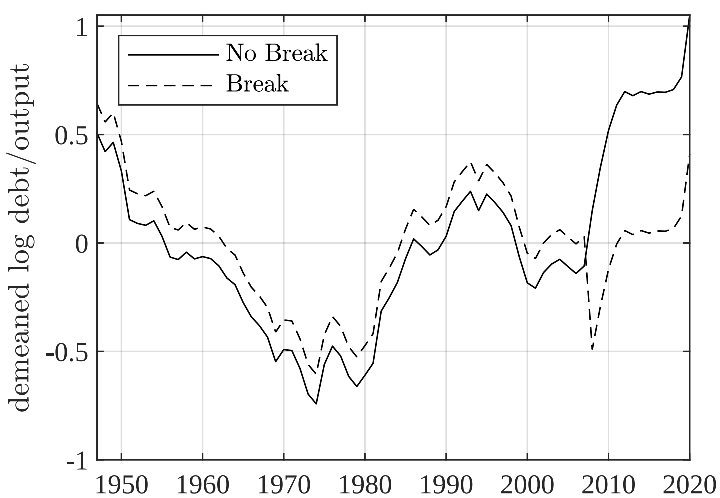
Hall and Sargent (2011) provide a detailed analysis of the ex-post realized returns on govern-

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.

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.

ment debt and how these impact the U.S. debt dynamics. We study how debt dynamics affect the variation in long-run future surpluses and returns, such as the decadal averages shown in Table 1.

1.2 Campbell-Shiller Decomposition of the Debt/Output Ratio

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

$$v_t = \mathbb{E}_t \sum_{j=1}^{T} \left(s_{t+j} - \widetilde{r}_{t+j} \right) + \mathbb{E}_t v_{t+T}. \tag{2}$$

The debt/output ratio reflects either expected future surpluses, expected future real growth-adjusted returns, or the expected future debt/output ratio. Since we only iterate forward on the budget constraint T times, we do not impose the transversality condition (TVC).[11]

The decomposition in equation (2) allows for the presence of convenience yields on government debt.[12] Convenience yields would increase bond prices and lower $\widetilde{r}$.

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:

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

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 to T periods from now. If v_t is non-stationary, then we replace the unconditional population moments with the finite-sample moments in the variance decomposition:[13]

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

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

[11]Imposing a TVC requires taking a stand on 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). If the TVC is satisfied and if v_t is stationary, then $\mathbb{E}_t[v_{t+T}]$ converges to its unconditional mean.

[12]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 contributes a component to the valuation of public debt.

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

tor:

$$
\begin{aligned}
\sum_{j=1}^{T} s_{t+j} &= a_s + b_T^s v_t + \varepsilon_{t+T}^s, \\
\sum_{j=1}^{T} \widetilde{r}_{t+j} &= a_r + b_T^r v_t + \varepsilon_{t+T}^r, \\
v_{t+T} &= \phi_0 + \phi_T v_t + \varepsilon_{t+T}^v. \qquad (5)
\end{aligned}
$$

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 the covariance between the debt/output ratio and future returns and surpluses. We estimate separate univariate regressions over longer horizons to infer the long-run dynamics rather than constraining the predictability across horizons through a VAR.[14]

Just as their counterparts from predictability regressions using stock returns, these regression coefficients measure the fraction of the variance in the debt/output ratio at each horizon T that is attributable to each component: future surpluses, future growth-adjusted returns, and future debt/output ratios:

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

Note that the three predictability 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 case of an AR(1) process for the debt/output ratio. The variance decomposition is then 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. If the debt/output approaches a unit-root process ($\phi_1 \to 1$), v_{t+T} accounts for all of the variation in v_t at horizon T.

We define fiscal sustainability over a finite horizon T as $\phi_T < 1$. We define fiscal sustainability as $\lim_{T\to\infty} \phi_T = 0$ over long horizons. If this condition is not satisfied, then the surplus/output ratio will inherit a unit root, which means that surpluses can drift arbitrarily far apart from output, which is not fiscally sustainable.

[14] Long-run VAR forecasts are more susceptible to misspecification of the dynamics (Jordà, 2005).

What do we expect to find in this variance decomposition? To anchor ideas, consider a world where the debt/output ratio is stationary ($\phi_T < 1$). In this world, the entire variance of ex-post realized returns is attributable to cash flows in the long run (see Appendix B):

$$\lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}] = \lim_{k\to\infty} \frac{1}{k} var[s_{t\to t+k}]. \tag{6}$$

A similar equation for stocks implies that, if the dividend-price ratio is stationary, the variance of stock returns is driven only by dividend growth in the long run. Consider two polar cases in this world with a stationary debt/gdp ratio.

In the first case, growth-adjusted returns are unpredictable. If growth-adjusted returns are i.i.d., with a variance ratio of one at all horizons (including $\lim_{k\to\infty} \frac{1}{k} var[\tilde{r}_{t\to t+k}] = var[\tilde{r}_t]$), then equation (6) 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 ($var[\tilde{r}_t] < var[s_t]$). The only channel pushing the debt/output ratio back to the mean following an increase in the debt/output ratio is through larger surpluses: $b^s_T = 1 - \phi_T > 0$. In the 1947-2020 sample, the standard deviation of s_t is 6.47% which exceeds the standard deviation of $\tilde{r}_t$ of 5.30%. The small difference in short-run variances of surpluses and returns suggests only weak mean reversion.

In the second case, surpluses are unpredictable. If the surpluses are i.i.d., then returns must do the adjustment to bring the debt/output ratio back down after an increase: $-b^r_T = 1 - \phi_T > 0$. However, given the empirical constellation of short-run variances of surpluses and returns, the stationarity of the debt/output ratio would imply long-run mean aversion in returns, which is at odds with the data.

This back-of-the-envelope evidence on the relative volatility of surpluses versus returns seems to rule out a world with a stationary debt/output ratio and no surplus predictability, but potentially allows for a world with a stationary debt/output ratio and no return predictability.[15]

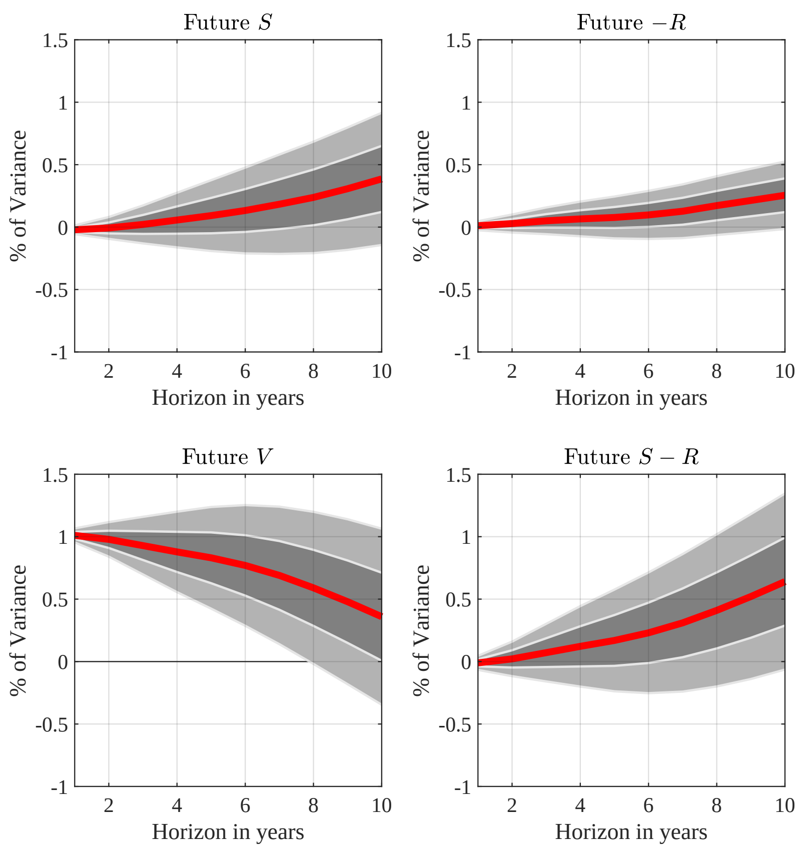
If there is no predictability of surpluses nor of real growth-adjusted returns, then the debt/output ratio inherits a unit root, which violates fiscal sustainability.

[15] Treasurys are different from stocks. In the case of stocks, the short run variance of returns exceeds that of dividend growth. 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.

1.3 Variance Decomposition of the U.S. Debt/Output Ratio

Figure 3 plots the baseline OLS regression results. Each panel plots the OLS slope coefficients in (5), where the dependent variable is future surpluses (top left), future real growth-adjusted returns (top right), and future debt/output ratios (bottom left) for horizons from $T = 1$ to $T = 10$ years. The bottom right panel has surpluses minus growth-adjusted returns as the dependent variable. We will refer to this dependent variable as "fundamentals." Table 2 reports these same predictability results in table format.

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 real growth-adjusted returns (with a minus sign) $\sum_{j=1}^{T} -\tilde{r}_{t+k}$, future log debt/output ratio v_{t+T}, and the difference of the future surpluses and returns. Sample is annual, 1947—2020. We plot 1 s.e. (dark) and 2 s.e. (light) confidence intervals. 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.

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.[16] 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. At the 10-year horizon, the other two channels start to matter: 39% of the variation is attributed to fluctuations in expected future surpluses and 25% of the variation is attributed to fluctuations in expected future returns. The terminal debt/output ratio in the 10-year horizon accounts for the remaining 36% of the variation.

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

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: $p_t^N = -\sum_{i=0}^{N-1} hpr_{t+i+1}^{N-i}$. But the predictability effect on nominal returns is almost completely offset by a similar-sized predictability effect on inflation, as shown in the third panel of Table 3. The net effect of the debt/output ratio on future real growth-adjusted returns, shown in the bottom panel, is small and statistically insignificant.

To generate confidence intervals for these predictability coefficients, we construct 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 such that the budget constraint identity is enforced. Here is the full system of equations for the bootstrap:

$$v_{t+1} = \psi_0 + \psi_1 v_t + \psi_2 v_{t-1} + u_{t+1}^v,$$

[16] 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 such unit root tests.

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

This table further decomposes the adjusted bond return $\widetilde{r}_{t+j}$ into its three components: nominal bond return r_{t+j}, GDP growth x_{t+j}, and inflation π_{t+j}. The adjusted bond return $\widetilde{r}_{t+j}$ 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 $\widetilde{v}_t$. 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 $\sum_{j=1}^{T} \widetilde{r}_{t+j}$					
$b_T^{\tilde{r}}$	-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

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

The bootstrapped standard errors are under the null that the debt/output ratio is mean-reverting and there is no role for cash flow or discount rates to affect the debt dynamics. For bootstrapping purposes, we use two lags because the debt/output ratio in the data fits an AR(2) structure well, and delivers white noise estimated residuals to bootstrap from. The confidence intervals (CIs) imply considerable uncertainty about the point estimate. Even at the 10-year horizon, the two-standard-deviation CIs for the b_j^s and b_j^r coefficients contain zero, so that the explanatory power of future surpluses and growth-adjusted returns are indistinguishable from zero.

1.4 Small-Sample Bias Correction

The high persistence in the explanatory variable v_t raises concern of small-sample bias. The OLS slope coefficients in the return predictability relationship, $|b_j^r|$, 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). Intuitively, an increase in bond risk premia will tend to induce lower realized returns and lower the ratio of the market value of debt to output. That gives rise to positive correlation between the regressor innovations and the return innovations. The positive biases tend to increase for long-horizon predictability regressions (Boudoukh, Israel, and Richardson, 2020).

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^s 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 in small samples. One can construct unbiased coefficients 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:

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

where ϕ denotes the first-order autocorrelation of v_t , T denotes the predictability 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 returns are both positive. OLS will tend to overstate the importance of both the cash flow and the discount rate component in accounting for the variation in the debt/output ratio. Holding fixed 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, 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 :

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

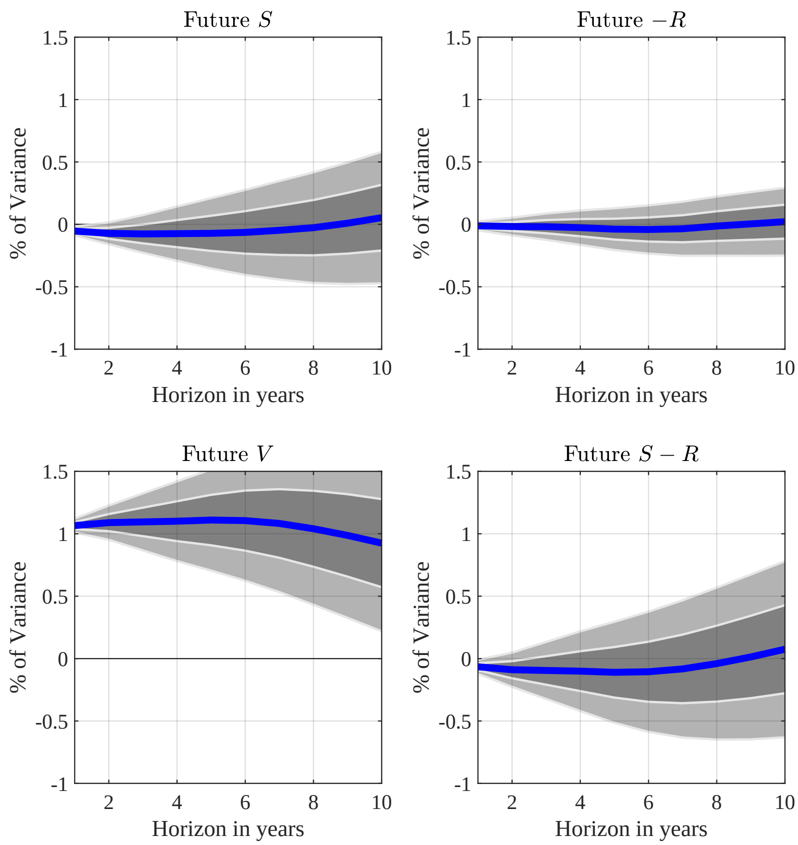
where $\mathbb{E}[\widehat{\phi} - \phi]$ is roughly $-(1+3\phi)/N$. This expectation is taken over all possible values of true autocorrelation ϕ. If instead we are willing to restrict $\phi \leq 1$, then the bias attains an upper bound at $\phi = 1$ (Lewellen, 2004). In the case of stock return predictability, Lewellen (2004) shows that the null of no predictability can be rejected more often if this stationarity restriction is imposed. Imposing that $\phi \leq 1$ is equivalent to imposing $b_1^r + b_1^s \geq 0$, i.e., assuming that there is return or cash flow predictability. We do not impose any ex-ante restrictions on the true autocorrelation of 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. Table 2 reports the bias-corrected coefficient estimates in the rows labeled *unbiased*. The three predictability coefficients 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 five-year horizon. As a result, 111% is accounted for by the future debt/output ratio. At the ten-year horizon, we still attribute 92% of the variance to the future debt/output ratio, after correcting for the small-sample bias. Variation in future surpluses over the next ten years accounts for 5% and future returns for 2% of the variation in the debt/output ratio. The latter two numbers are far below the 39% and 25% OLS point estimates, showing a very large small-sample bias. The bias-corrected point estimates combined with the standard errors imply that there is no 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 fundamentals, the difference between surpluses and growth-adjusted returns. This is the main result in the paper.

The confidence intervals for the discount rate contribution in the top right panel are quite 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 role at all. 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 50% at the ten-year horizon. Similarly, we cannot definitively rule out that there is significant mean-reversion at the ten-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, a well-documented feature of unit root tests.

At this stage, we consider two important robustness checks. First, our surplus variable s_t, as defined in equation (1), is defined based on the log-linearization that enforces the linearized budget constraint exactly. Under this definition, the cross-equation restriction on the three predictability coefficients always holds. As noted earlier, s_t has a correlation with the actual surplus/GDP ratio s_t^{Actual} of 0.83, but it is more volatile. We check whether the log debt/GDP ratio predicts the

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 $\sum_{j=1}^{T} \tilde{r}_{t+k}$, 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.

sum of actual government surplus/GDP ratios:

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

We report the regression coefficients in Appendix Figure A1. Consistent with our main result, the log debt/GDP ratio does not predict the actual government surplus/GDP ratios either.

Second, we extend the sample in an effort to improve the power of our tests. Using the Hall, Payne, and Sargent (2018) historical data, we construct a longer U.S. sample that starts in 1842.

The variance decomposition with and without bias adjustment is plotted in Figure A2. The bias-corrected point estimates are similar to those obtained for the post-WW-II sample. However, the longer sample shrinks the CIs around the point estimates. At the 10-year horizon, we can now definitively rule out that future V accounts for less than 50% of the variation. Conversely, we can rule out that fundamentals account for more than 50% of the variation in the debt/output ratio. 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 10 years. After the bias correction, future v ten years out explains about 85% of the variation in today's v.

1.5 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 general assumptions on 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 three additional robustness checks by revisiting the shorter Bohn (1998) sample, considering different steady-state values for the adjusted returns on the debt portfolio, and considering a different definition for the predictor variable.

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. Our paper re-examines this evidence taking into account the small-sample bias issue, which gives rise to spuriously large point estimates for surplus predictability at longer horizons. We repeated our exercise on Bohn (1998)'s 1948—1995 sample. Before correcting for the bias, there is indeed a larger role for fundamentals in this subsample, as shown in Figure A3, consistent with Bohn's findings. However, this evidence mostly disappears after the bias correction.

2.2 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 (2) 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 A4 reports the estimates with and without bias adjustment for $\rho = \exp(1\%)$. Figure A5 reports the estimates for $\rho = \exp(-1\%)$. The results after the bias correction for both cases are similar to the ones obtained for the benchmark $\rho = 0$ case.

2.3 Domestic Holdings of Treasury Debt

Foreign investors and the Federal Reserve Bank are considered to be relatively price-inelastic investors (Jiang, Krishnamurthy, and Lustig, 2022). Foreign holdings rise starting in the mid 1980s while the Fed holdings increase dramatically following the GFC and Covid crises. To examine the extent to which our results are driven by these holdings, we construct a measure of the debt/output ratio that excludes the holdings of the Fed and the rest of the word. We denote the ratio of the debt held by the domestic private sector and output by v^d_t. Figure A6 shows that there has indeed been an increasing gap between domestic private sector holdings of Treasury debt and the total amount of Treasury debt, especially since 2007.

Do domestic private holdings of Treasury debt have stronger predictive power for future surpluses or discount rates? We regress future surpluses and future adjusted debt returns on v^d_t:

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

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^d_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 A7 as functions of the forecast horizon T. The predictor v^d_t explains around 40% of the variation in the future surpluses $\sum_{j=1}^{T} \rho^{j-1} s_{t+j}$ at the 10-year horizon. This suggests that future surpluses may respond more to the debt that has to be absorbed by the more price-elastic domestic investors. However, after the bias correction, we still cannot reject the null that surpluses are not predictable.

3 Permanent Shocks to the Debt/Output Ratio

3.1 Simulation Evidence

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. Under this null, there is no contribution from return or 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. The system is given by the following equations:

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

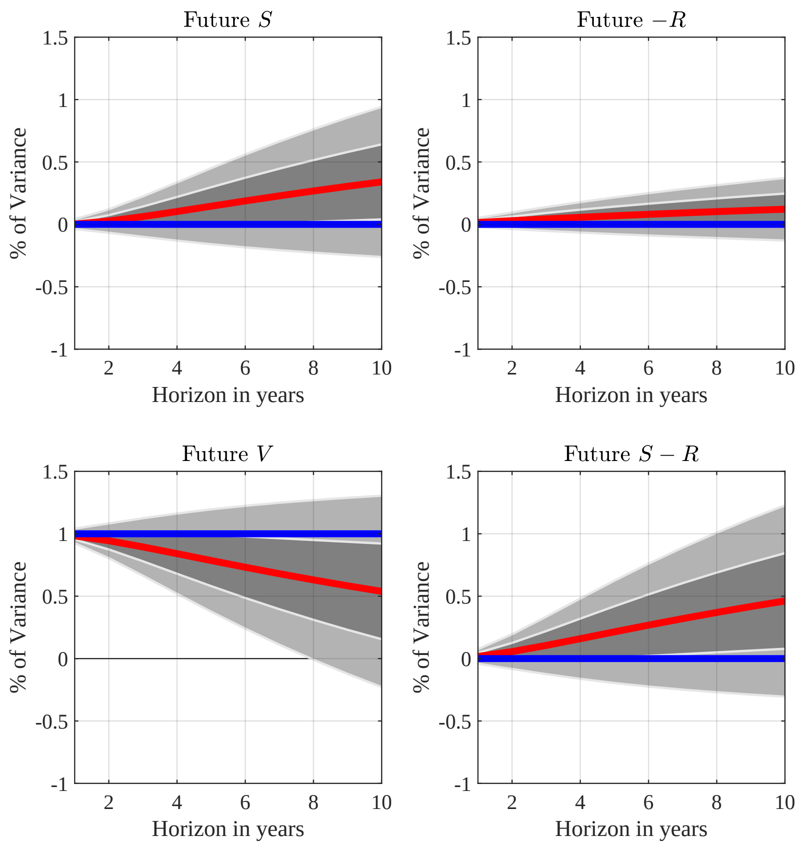
As always, we infer the surplus from the budget constraint: $s_{t+1} = \widetilde{r}_{t+1} - \Delta v_{t+1}$. In each round of simulation, we draw with replacement from the joint distribution of residuals $(\varepsilon_{t+1}^v, \varepsilon_{t+1}^r)$. Then, we estimate the forecasting regressions (5) by OLS in each simulated sample. We plot the mean slope coefficients 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 they do not by assumption. Figure 5 reports the variance decomposition implied by the simulated samples.[17] The average OLS slope coefficients generated by the unit-root model imply variance decompositions that are very close to our point estimates from the historical sample in Figure 3. The 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 ϕ_T are one at all horizons T.

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 log debt/output ratio is about $36\% + 50\% = 86\%$ in the 10-year horizon. This result is quite close to the 92% bias-corrected number we obtained in Figure 4 and Table 2. The simulation exercise

[17] 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.

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 surpluses $\sum_{j=1}^{T} s_{t+j}$, future real growth-adjusted returns (with a minus sign) $\sum_{j=1}^{T} -\tilde{r}_{t+k}$, the future log debt/output ratio v_{t+T}, and the difference between the future surpluses and returns. 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 OLS slope coefficients in red. We plot 1 s.e. (dark) and 2 s.e. (light) CIs. We also plot the long-sample slope coefficients in blue from a single simulation of 100,000 periods.

provides corroborating evidence for the presence of a unit root in the log debt/output ratio.

3.2 Structural Breaks

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 GFC 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. A Chow test for structural breakpoints rejects the null hypothesis of no structural break at the 1% level in 2007 and at no other date.

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 as the dashed line. This approach removes a low-frequency component from the debt/output ratio. 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. We recompute the surpluses by feeding $\widetilde{v}_t$ into equation (1) so that the cross-equation restriction still holds. The slope coefficients now provide a variance decomposition of the transitory variation in debt/output. Taking covariances with $\widetilde{v}_t$ on both sides of the previous equation, we obtain the following expression for the variance of the transitory component of debt/output ratio $\widetilde{v}_t$:

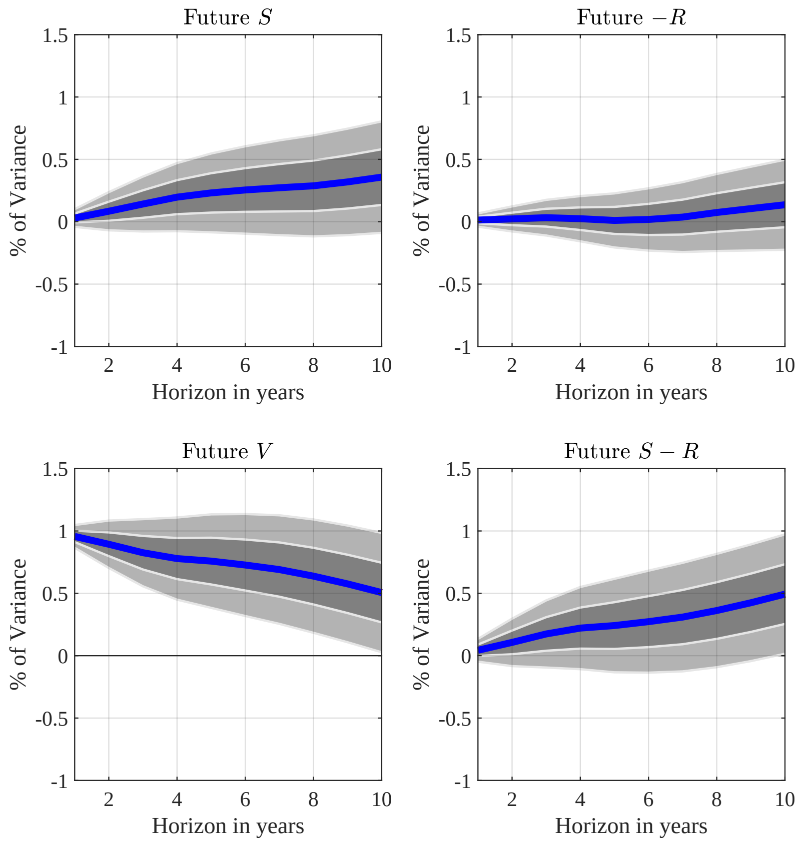
$$var(\widetilde{v}_t) = cov\left(\sum_{j=1}^{T} s_{t+j}, \widetilde{v}_t\right) - cov\left(\sum_{j=1}^{T} \widetilde{r}_{t+j}, \widetilde{v}_t\right) + cov(\widetilde{v}_t, \widetilde{v}_{t+T}). \quad (10)$$

The resulting break-adjusted series is much less persistent than the original series (lower ϕ_T). Because of the cross-equation restrictions $b_T^s - b_T^r + \phi_T = 1$, this imputes a larger role to fundamentals in explaining the variation in the transitory component of debt/output.

Figure 6 decomposes the variance of the transitory component of the debt/output ratio. The transitory component of the debt/output ratio explains $R^2 = 36\%$ of the variation in five-year cumulative surpluses. When the transitory component of the debt/output ratio is high, surpluses tend to increase to push the debt/output ratio back down. Table 4 shows that future surpluses and returns combine to explain 24% (50%) of the variation in transitory component of debt/output at the five-year (ten-year) horizon after bias correction. We can reject the null that fundamentals do not play a role at the ten-year horizon. 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 explanation for the large and permanent increase in the debt/output ratio in the post-2007 sample. One candidate explanation for the structural break in the relationship between debt/GDP and fundamentals is the rise in foreign and Fed holdings of Treasurys discussed above. The large rise in these holdings comes after 2007, so the timing fits with the estimated time of the structural break. The next section discusses another candidate explanation for the structural break.

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.

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. One possible explanation for these results is that the U.S. government debt portfolio is mispriced, perhaps because investors have been overly optimistic about the U.S. fiscal

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
					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	0.12	0.14	0.15	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.	0.04	0.08	0.11	0.14	0.16	0.18	0.19	0.2	0.21	0.22
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	0.1	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

situation.

Suppose investors evaluate equation (2) using their own subjective expectation, denoted by the operator $\mathbb{F}$, which differs from the true 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 the econometrician parses out the variations in the debt/GDP ratio under the objective measure, she obtains:

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

Comparing the two equations, we obtain:

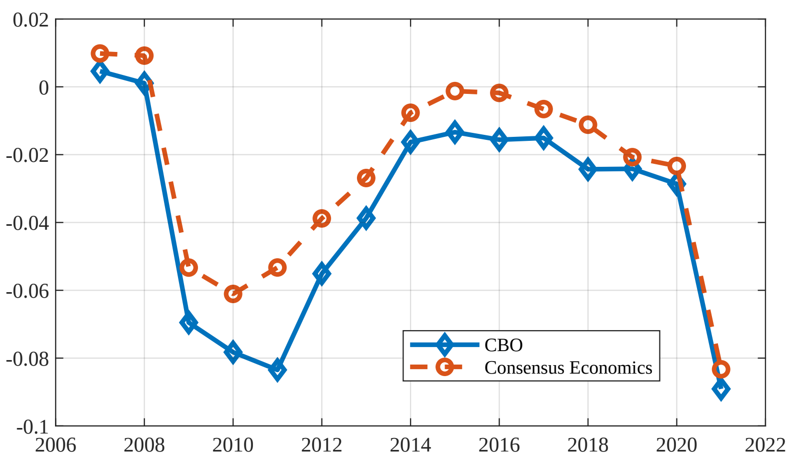
$$\mathbb{E}_t v_T = \mathbb{F}_t v_{t+T} + (\mathbb{F}_t - \mathbb{E}_t) \sum_{j=1}^{T} (s_{t+j} - \widetilde{r}_{t+j}),$$

where the last term denotes the investors' forecast errors about fundamentals. This setup allows for the possibility that the debt/GDP ratio is explained fully by fundamentals under the subjective measure, i.e., $cov(v_t, \mathbb{F}_t v_{t+T}) = 0$, while the econometricians finds the opposite, i.e., $cov(v_t, \mathbb{E}_t v_{t+T}) \approx var[v_t]$. In this setting, what the econometrician interprets as a unit root in the debt/GDP process reflects investors' forecast errors that covary with the debt/output ratio.

To assess this possibility empirically, we conjecture that government bond market forecasters turn to the CBO's detailed and sophisticated projections to make their own subjective forecasts about future debt/GDP and surplus/GDP ratios. Private sector survey evidence from Consensus Economics supports this conjecture. Specifically, for the 2007–2021 period we have annual one-year ahead forecasts for the surplus/GDP ratio made at the start of the year from both the CBO and Consensus Economics. These forecasts have a 98% correlation and are very close in both levels and timing, as shown in Figure 7. If anything, the private sector forecasters are slightly more optimistic about future surpluses than the CBO. The main advantage of working with the CBO projections is that they go out ten years, so that we can compute the expected cumulative surpluses over the next ten years. Consensus Economics forecasts are only available for the next year.

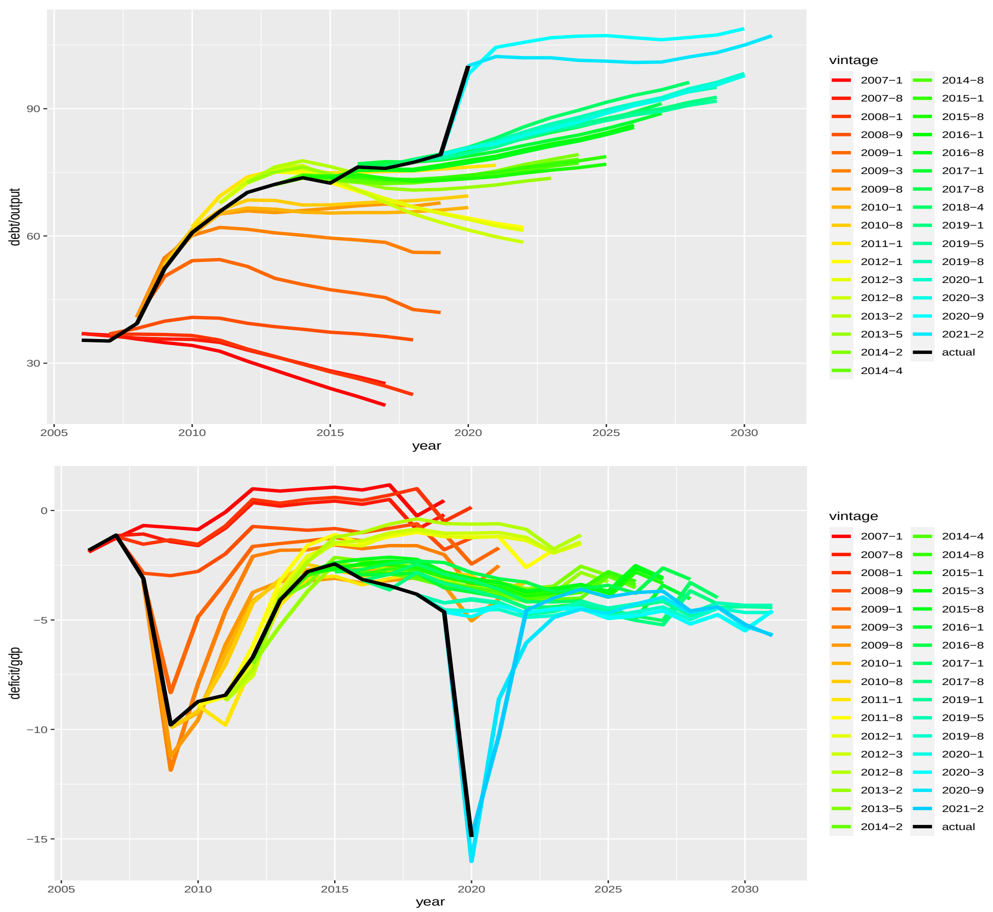
While the CBO forecasts GDP, inflation, and interest rates, it makes projections of future surpluses based on current law. There is a large and consistently negative (positive) forecast error in its debt/GDP (surplus/GDP) projections. The top panel of Figure 8 plots the CBO ten-year projections $\mathbb{F}_t v_{t+T}$ for the debt/GDP ratio against the actual time series. The CBO has been consistently under-predicting the debt/GDP ratio. This bias arises mainly because the CBO has been over-predicting future surpluses as a fraction of GDP, as shown in the bottom panel. If the CBO

Figure 7: Comparing CBO and Private-Sector Surplus Forecasts



This figure plots the forecast, made at the start of the year (in January), for the current calendar year surplus/GDP ratio. The solid blue line with diamonds is for the Congressional Budget Office forecast while the dashed red line with circles line is the mean forecast from Consensus Economics, a dataset of private sector forecasts. Each line combines forecasts for the government surplus, debt service, real GDP growth, and inflation from the respective data sets. Since the CBO data starts in 2007, we use the longest overlapping sample of 2007—2021 for this graph.

Figure 8: CBO Projections for Debt/GDP and Surplus/GDP



This figure reports the ten-year U.S. federal government debt/GDP (top panel) and total federal government surplus/GDP (bottom panel) projections made by the CBO at the start of each year from 2007–2020. The solid black line is the ex-post realized debt/GDP (top panel) and surplus GDP/ratio (bottom panel).

had assumed future fiscal corrections (beyond those in current law) when making its debt/output projections, its projections would have been even farther from the realized debt/output ratios. In sum, 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 the projections remain biased well after the GFC. Given the size and persistence of the CBO projection errors, and based on the evidence from one-year surplus forecast in Figure 7, 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 ten-year surplus/output ratio using the CBO projections, $\sum_{j=1}^{T} s_{t+j}$, where $s_{t+j} = sy_{t+j}/e^{v}$ and v is the 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.

We consider the case where there is no return predictability under either the objective or the subjective measure.[18] To explain our findings that the future debt/output ratio v_T accounts for most of the variation without any surplus predictability under the objective measure, we need a large positive covariance between surplus forecast errors and the debt/output ratio under the subjective measure:

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

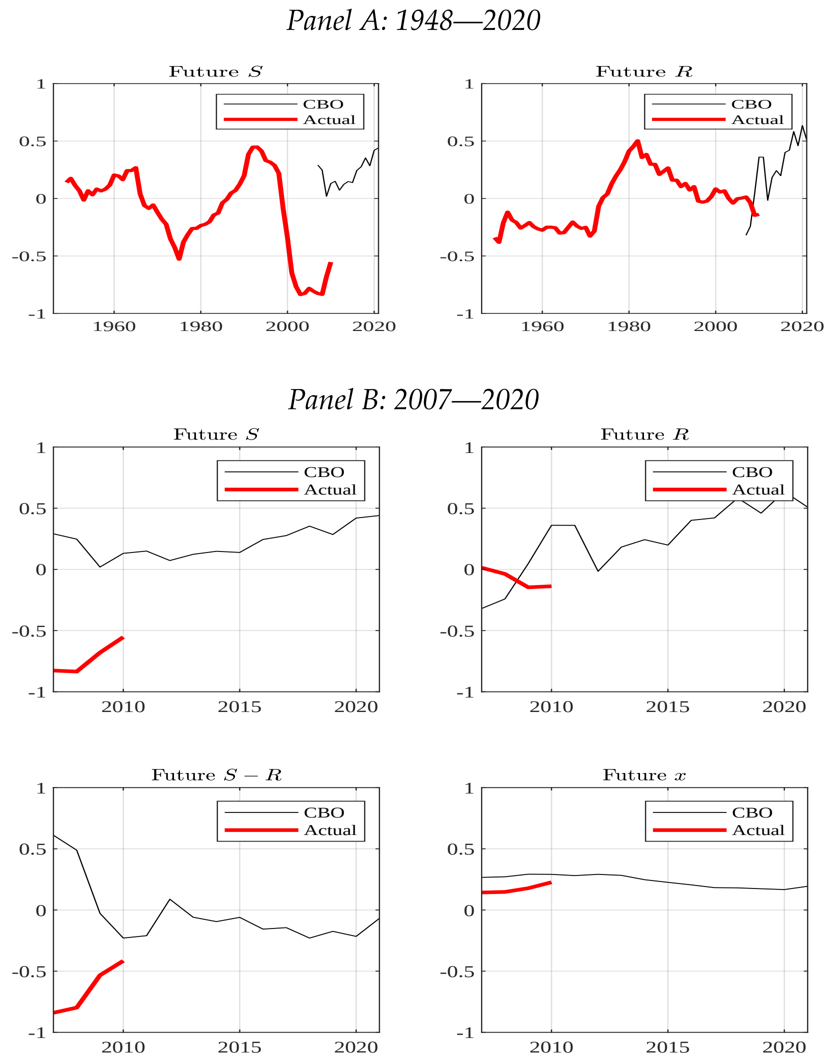
That is, if the first covariance term is not as large under the subjective than under the objective measure, then the second covariance term must be positive and make up the difference. The evidence in Figure 8 indeed shows that, as the fiscal situation worsened, the CBO projections persistently over-projected surpluses. Likewise, bond market investors may have been pricing in larger cumulative surpluses than materialized after 2007.

Figure 9 analyzes the difference between subjective expectations of ten-year future surpluses and actual realized surpluses. Panel A is for the full sample 1948–2020, while Panel B focuses on the post-2007 sample. The left plot in Panel A shows that the realized surplus series declines steeply in 1998 as the GFC enters the forward-looking sum. In 2007, just before the GFC, the CBO projections were far too optimistic: $(\mathbb{F}_t - \mathbb{E}_t)\sum_{j=1}^{T} s_{t+j} \gg 0$. The CBO projected a positive ten-year surplus of 29 log points (roughly 2.9% per annum). 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 (-8.2% per annum). There is a 111 log point forecast error (11.1% per annum). The sign of the gap may not be surprising, but the size is. The top left graph of Panel B shows that the surplus projections remain too optimistic during and after the GFC. 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.

Similarly, the return projections were too optimistic. The CBO-projected real growth-adjusted returns over the next 10 years were -32 log points or -3.2% per annum in 2007, shown in the top right graph in Panel B of Figure 9. The realized returns were 1.3 log points. As shown in the bottom right panel, overly rosy GDP growth predictions contributed to the under-prediction of growth-adjusted returns. Once the inflation expectations are included, the CBO projected $r - x - \pi$ to be 2.2 % lower per annum than what actually realized. Hence, there is an additional 22 log point contribution to the forecast error of fundamentals coming from the discount rate component. This component shrinks and eventually reverses by 2010 since the CBO raised its projections for the effective long-run cost of government funding.

[18]This can be justified by the tighter confidence intervals around a zero point estimate we obtained for the return predictability coefficient.

Figure 9: Decomposition for CBO Projections



The plot shows 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 $\sum_{j=1}^{T} \widetilde{r}_{t+k}$, for $T = 10$. We also report future real growth $\sum_{j=1}^{T} \widetilde{x}_{t+k}$.

The bottom left graph in panel B plots the projected fundamentals $S - R$. There is a total 135 log point gap in 2007 between the projected and realized fundamentals. The projected fundamental over the next ten years equals the projected decrease in the log debt/output ratio over the next ten years. As a result, the realized debt/output ratio in 2017 was 135 log points higher than it was projected to be in 2007. In 2007, the CBO projected a more than 50 log points decline in the debt/output ratio over the next ten years which failed to materialize. Even in 2009, it still projected a constant debt/output ratio over the next decade.

The empirical evidence from the CBO and private forecasters is consistent with mispricing.

The CBO seems to overestimates long-run 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/output ratio. If investors' subjective expectations are aligned with CBO projections, an assumption supperted in the data, they think they live in a world without permanent shocks to debt/output. But their systematic forecast errors induce permanent shocks in the debt/output ratio from the perspective of an econometrician given the historical sample.

5 Conclusion

The bond market's valuation of a claim to surpluses is surprisingly insensitive to news about future surpluses or returns from the perspective of an econometrician looking at the U.S. historical sample. This is a direct result of the debt/output ratio's persistence. This persistence plagues small-sample predictability regressions that aim to uncover the extent to which high debt/output episodes are followed by higher government surpluses, higher growth, higher inflation, or lower bond returns. After correcting the small-sample bias there is no evidence of such adjustments. To be clear, we do not claim to have proven that the U.S. is on an unsustainable fiscal path. What we have shown is that there is no statistical evidence to conclude that the U.S. is on a fiscally sustainable path. Our 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 but the bond market investor does. We show evidence from fiscal projections that suggests that investors may have been over-predicting future surpluses after 2007. As a result, bond investors anticipated mean reversion in the debt/output ratio that failed to materialize. Systematic forecast errors could be a source of the persistence in the debt/output ratio. Price-inelastic purchases of Treasurys by foreign investors and the Central Bank are another potential source of persistence in the debt/output ratio.

Even though the pricing of individual bonds does not allow for arbitrage opportunities, except in times of market disruptions, the entire portfolio of U.S. federal government debt may be persistently mispriced from the perspective of rational investors. Limits to arbitrage (Shleifer and Vishny, 1997) would plausibly prevent rational arbitrageurs from taking advantage of potential overpricing in the last two 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.

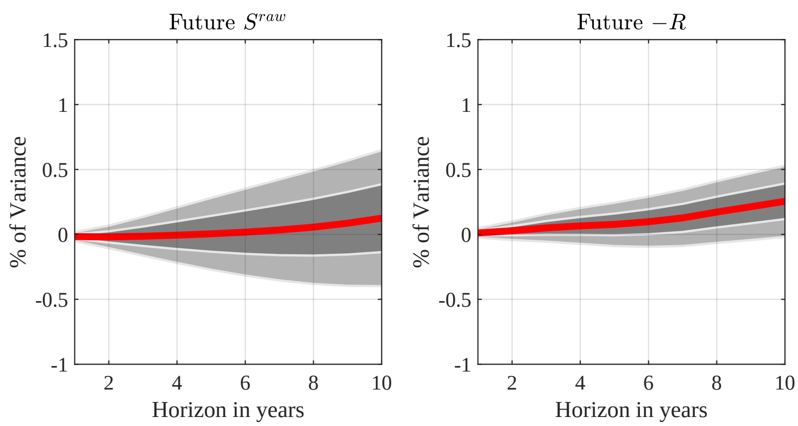
Finally, we cannot rule out the possibility of a peso-event, a large fiscal correction not observed in our sample, but anticipated and priced in by investors. Hitherto, even the most sophisticated budget projections which abstract from future fiscal corrections have proven too optimistic.

Appendix

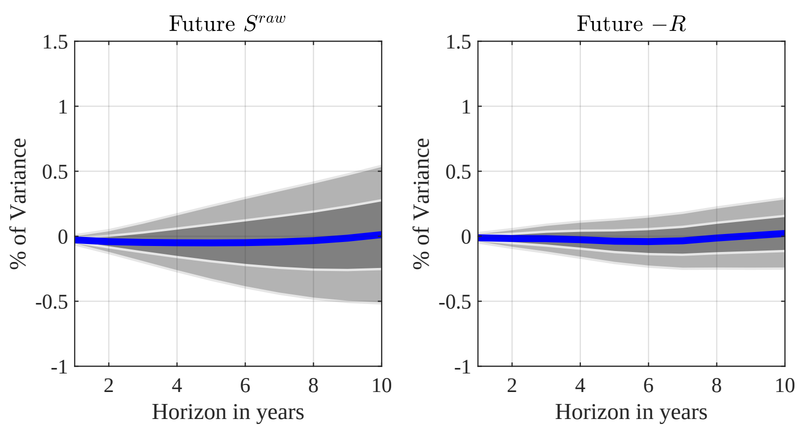
A Robustness Tables

Figure A1: Forecasting Power of Log Debt/Output Ratio for Actual Government Surplus

Panel A: Before Bias Correction



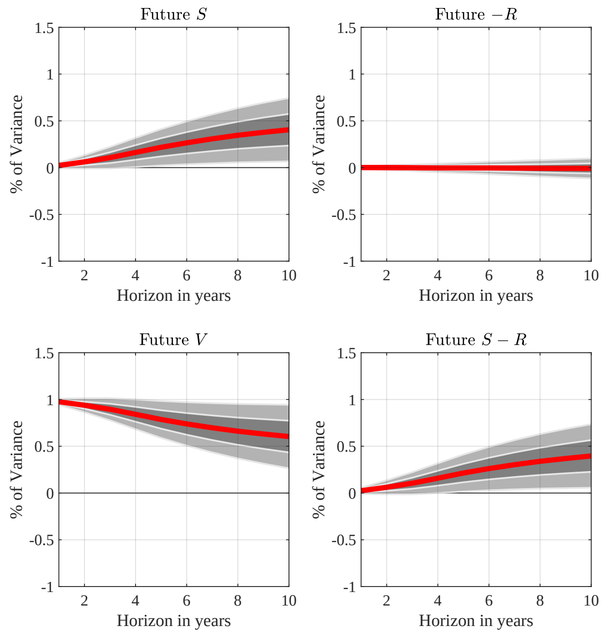
Panel B: After Bias Correction



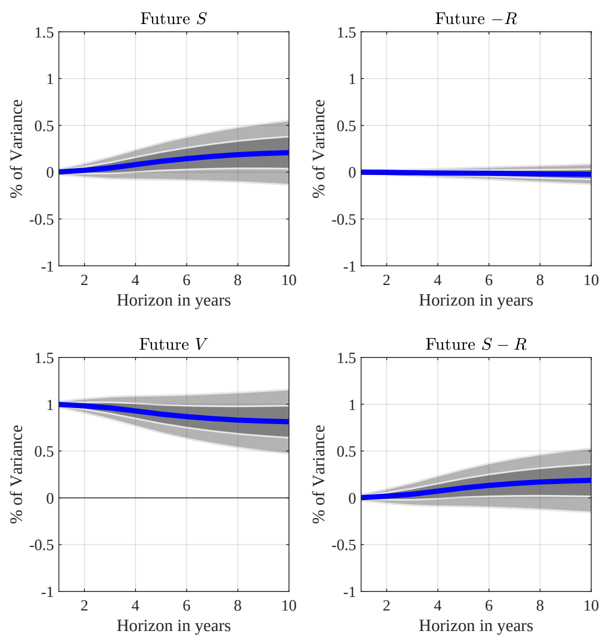
This figure reports regression coefficients associated with the log debt/output ratio. S^{raw} is based on the actual government surplus/GDP ratio. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, and Richardson (2020) small-sample bias correction for the OLS 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.

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

Panel A: Before Bias Correction



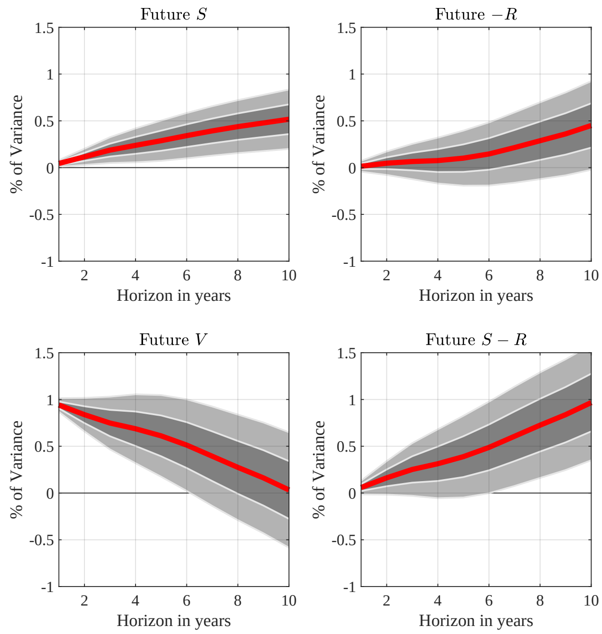
Panel B: After Bias Correction



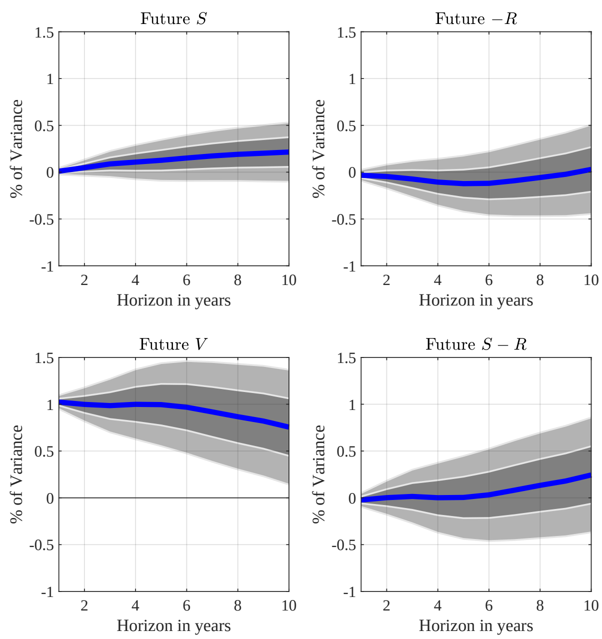
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, 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: Shorter Bohn Sample 1948—1995

Panel A: Before Bias Correction

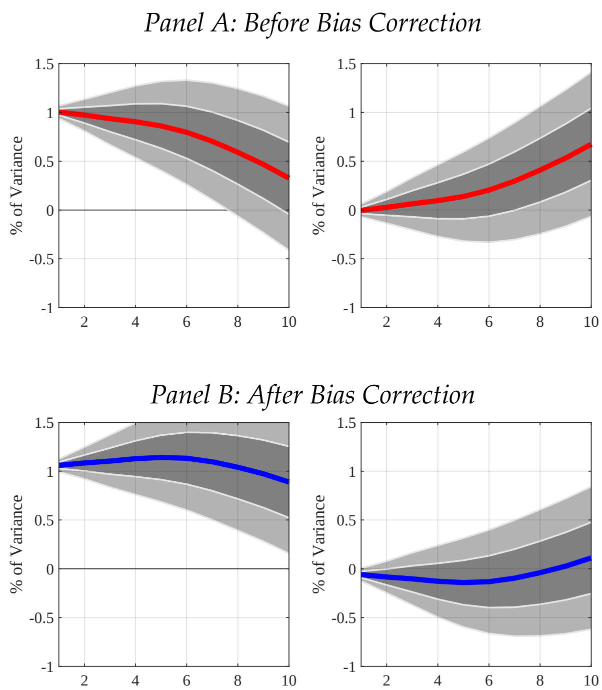


Panel B: After Bias Correction



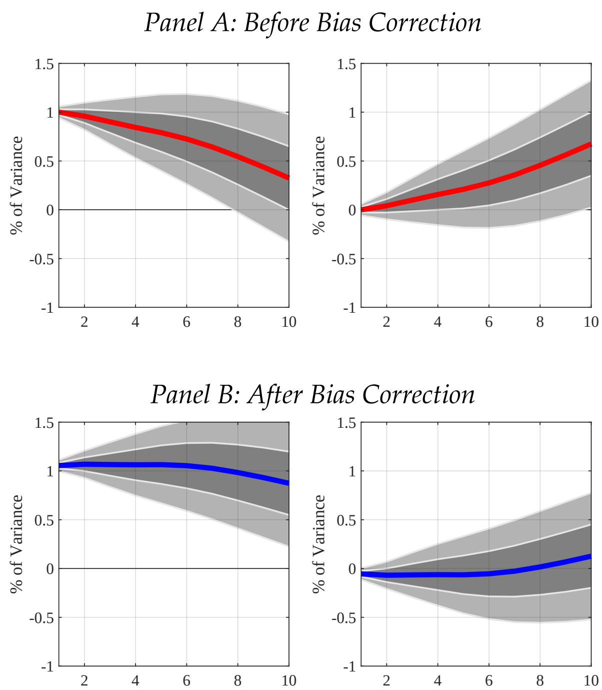
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio v_{t+T} and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, Israel, 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 A4: Variance Decomposition of Log Debt/Output Ratio: with $\rho = \exp(1\%)$



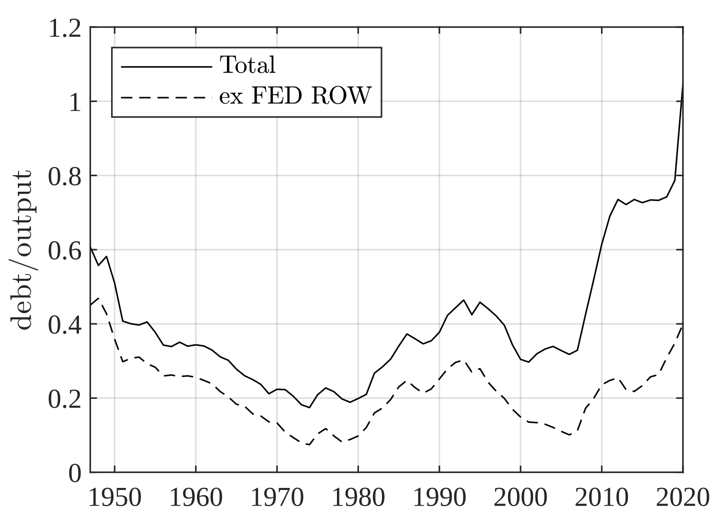
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, 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: Variance Decomposition of Log Debt/Output Ratio: with $\rho = \exp(-1\%)$



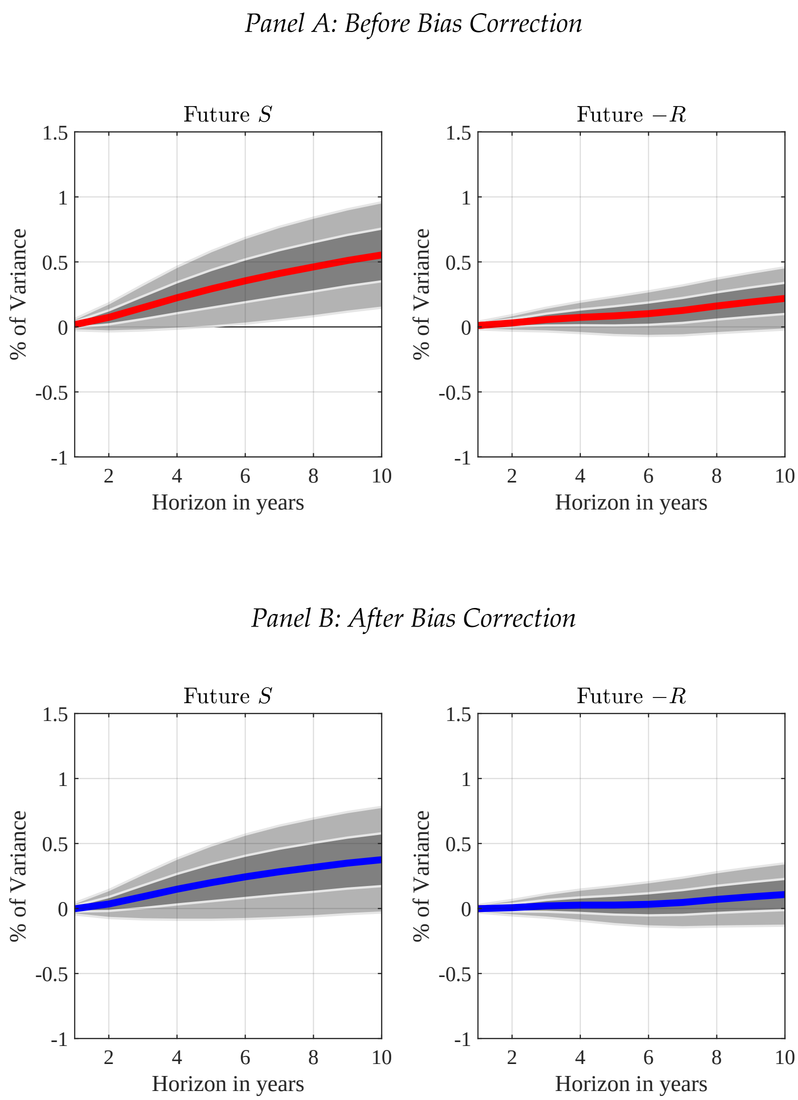
This figure reports the variance decomposition of the log debt/output ratio v_t into components due to future log debt/output ratio $\rho^T v_{t+T}$, and the combination of the future government surpluses and discount rates. We generate bias-corrected slope coefficients using the Stambaugh (1999); Boudoukh, 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 A6: Debt/Output Ratio



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

Figure A7: Forecasting Power of Log Debt/Output Ratio Held by the Domestic Sector



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.

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 choose $\rho = \exp(-(r - x - \pi)) = 1$.

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 $\lim_{k\to\infty} \frac{1}{k} cov(\Delta v_{t\to t+k}, s_{t\to t+k}) = 0$.

Consider a first case in which the returns are i.i.d.. Then the long-horizon variance ratio of returns is one:

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

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

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

If $var[\tilde{r}_t] < var[s_t]$ then it follows from the fact that the variance ratio of long-horizon returns is one (see equation A2) and from the fact that the variance of returns and surpluses converge over long horizons (see equation 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.

Now consider a second case in which the surpluses are i.i.d.. Then the long-horizon variance ratio of surpluses is one:

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

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 surpluses converges to one (see equation A3) and from the fact that the variance of returns and surpluses converge over long horizons (see equation A1) that the variance ratio of surpluses is smaller than one:

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

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

References

BARBERIS, N. (2018): *Handbook of Behavioral Economics: Foundations and Applications*chap. Psychology-Based Models of AssetPrices and Trading Volume, pp. 80–175. Elsevier.

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.

BORDALO, P., N. GENNAIOLI, R. LAPORTA, AND A. SHLEIFER (2019): "Diagnostic Expectations and Stock Returns," *Journal of Finance*, 74(6), 2839–2874.

BORDALO, P., N. GENNAIOLI, Y. MA, AND A. SHLEIFER (2020): "Overreaction in Macroeconomic Expectations," *American Economic Review*, 110(9), 2748–82.

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).

D'ERASMO, P., E. G. MENDOZA, AND J. ZHANG (2016): *Handbook of Macroeconomics*chap. What is a Sustainable Public Debt?, pp. 2493–2597. Elsevier.

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.

GOYAL, A., AND I. WELCH (2008): "A Comprehensive Look at The Empirical Performance of Equity Premium Prediction," *Review of Financial Studies*, 21(4), 1455–1508.

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*vol. Rational Expectations Econometrics, chap. 5. Westview Press.

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.

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

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.

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.

Wang, R. (2020): "Subjective Return Expectations," *Available at SSRN 3740792*.