The U.S. Public Debt Valuation Puzzle[*]

Zhengyang Jiang
Northwestern Kellogg, NBER

Hanno Lustig
Stanford GSB, NBER, SIEPR

Stijn Van Nieuwerburgh
Columbia Business School, NBER, CEPR

Mindy Z. Xiaolan
UT Austin McCombs

April 23, 2024
First draft: March 2019

Abstract

The government budget constraint ties the market value of government debt to the expected present discounted value of fiscal surpluses. We find evidence that U.S. Treasury investors fail to impose this no-arbitrage restriction in the U.S. Both cyclical and long-run dynamics of tax revenues and government spending make the surplus claim risky. In a realistic asset pricing model, this risk in surpluses creates a large gap between the market value of debt and its fundamental value, the PDV of surpluses, suggesting that U.S. Treasurys may be overpriced.

Keywords: bond pricing, fiscal policy, term structure, convenience yield.

---

[*] Jiang: Finance Department, Kellogg School of Management, Northwestern University. 2211 Campus Drive, Evanston, IL 60208. `zhengyang.jiang@kellogg.northwestern.edu`. Lustig: Department of Finance, Stanford Graduate School of Business, Stanford CA 94305; `hlustig@stanford.edu`; `https://people.stanford.edu/hlustig/`. Van Nieuwerburgh: Department of Finance, Columbia Business School, Columbia University, 3022 Broadway, New York, NY 10027; `svnieuwe@gsb.columbia.edu`; Tel: (212) 854-1282. Xiaolan: McCombs School of Business, the University of Texas at Austin; `mindy.xiaolan@mccombs.utexas.edu`. The authors would like to thank Ravi Bansal, Marco Bassetto, Jules van Binsbergen (discussant), Philip Bond, Markus Brunnermeier (discussant), Max Croce (discussant), Tetiana Davydiuk (discussant), Peter DeMarzo, John Donaldson, Ben Hebert, Lars Hansen (discussant), Chris Hrdlicka, Narayana Kocherlakota, Nobu Kyotaki, Ralph Koijen, Yang Liu, Ian Martin, John Moore, Christian Moser, Jonathan Parker, Carolin Pflueger (discussant), Jean-Paul Renne, Tom Sargent, Lukas Schmid (discussant), Jesse Schreger, Ivan Werning, Pierre Yared, Steven Zeldes, and seminar and conference participants at the Joint Stanford-U.C. Berkeley finance seminar, Columbia University macro-economics, Kellogg finance, LSE, Chicago Booth finance, UT Austin finance, the Federal Reserve Board, the University of Washington, Stanford economics, Stanford finance, USC, UCLA Anderson, Shanghai Advanced Institute of Finance, the UCLA virtual finance workshop, Yale SOM, Banca d'Italia, NBIM, MIT Sloan-Harvard Business School joint seminar, the New York Federal Reserve Bank, the Society for Economic Dynamics meetings in St Louis, the Advances in Macro-Finance Tepper-LAEF Conference, NBER Summer Institute AP/MEFM, the Western Finance Association, the Midwest Finance Association, Minneapolis Federal Reserve Conference on Monetary/Fiscal Interactions, the American Finance Association, the Guanghua International Symposium on Finance, and the Vienna Symposium on Foreign Exchange Markets for insightful discussions. We gratefully acknowledge financial support from NSF award 2049260.

1 Introduction

In this paper, we price the entire portfolio of outstanding Treasury debt, rather than individual bond securities. In the absence of bubbles, the market value of outstanding debt should equal the present discounted value of current and future primary surpluses. By the same logic, the expected return on the debt portfolio has to reflect the risk profile of primary surpluses. What makes this a valuation equation, not an accounting identity, is that we insist on pricing the surplus claim in a manner that is consistent with the risk pricing in stock and bond markets. We find evidence of mispricing. The value of the bond portfolio exceeds the value of the surplus claim, a gap we label the government debt valuation puzzle. Conversely, yields on the Treasury bond portfolio are lower than the relevant "interest rate" bond investors ought to be earning given the risk in the surplus claim, a gap we label the government debt risk premium puzzle.

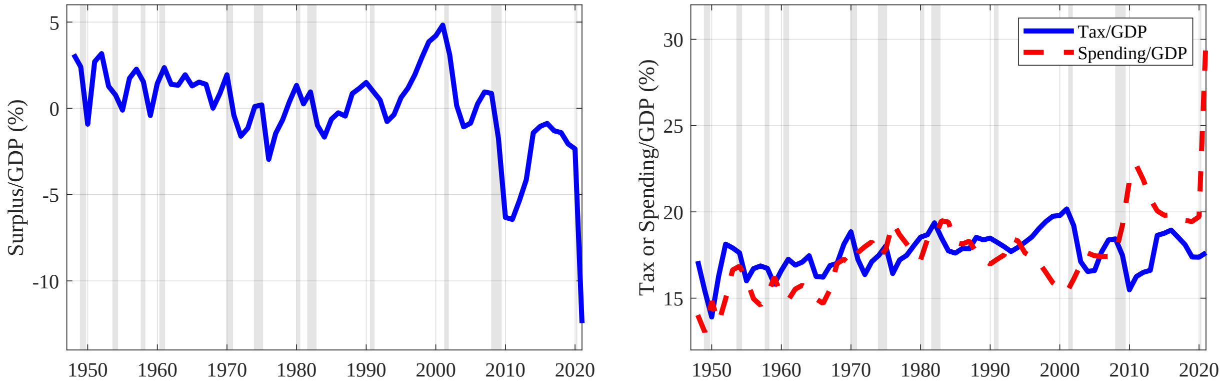
To understand why, consider a stock-pricing analogy. The price of a stock is the expected present discount value (PDV) of future dividends. Risk-free interest rates are below expected dividend growth rates, yet the price of the stock is finite. Since the stock's dividend growth is pro-cyclical, its cash flows are low when the investors' marginal utility is high. The relevant "interest rate" for the dividend claim contains a risk premium because of the risk exposure of its cash flows: the equity risk premium. Analogously, a portfolio strategy that buys all new Treasury issues and receives all Treasury coupon and principal payments has as its cash flow the primary surplus of the federal government. Primary surpluses are strongly pro-cyclical just like stock dividends, as shown in Figure 1. Spending by the federal government increases in recessions, while the tax system produces pro-cyclical revenue.[1] In recessions, when marginal utility is high, surpluses are negative and net government bond issuance is high. The Treasury portfolio cash flows have substantial business cycle risk. As explained below, and just like stocks, tax revenue and spending also have substantial long-run risk due to cointegration with GDP. Taken together, the relevant "interest rate" for surpluses contains a substantial risk premium reflecting compensation for both short- and long-run risk exposures.

The expected risk-adjusted PDV of surpluses governs a country's fiscal backing, i.e., how much debt it can issue. The expected risk-adjusted PDV of primary surpluses divided by output, is obtained as the difference between the expected risk-adjusted PDV of federal tax revenues, given by the tax/output ratio (τ_t) times the price/dividend ratio on a claim to future taxes, $\tau_t \exp(pd_t^T)$, and the expected risk-adjusted PDV of federal spending excluding debt service, given by the spend-

---

[1] Non-discretionary spending, including Social Security, Medicare and Medicaid, food stamps, and unemployment benefits, accounts for at least two-thirds of government spending. Many of these transfer payments rise automatically in recessions. In addition, the government often temporarily increases transfer spending in recessions, e.g., the extension of unemployment benefits in 2009 and 2020. On the revenue side, the progressive nature of the tax code generates strongly pro-cyclical variation in tax revenue as a fraction of GDP.

Figure 1: U.S. Government Surplus



(a) Surplus/GDP

(b) Tax and Spending/GDP

The figure plots the U.S. federal government primary surplus as a fraction of GDP. The primary surplus is defined as the current revenue minus current spending, excluding net interest spending. The data source is NIPA Table 3.2. The sample period is from 1947 to 2020.

ing/output ratio (g_t) times the price/dividend ratio on a claim on future spending, $g_t \exp(pd_t^G)$. In the long run, spending and taxes grow at the same rate as output because of co-integration. The pro-cyclicality of tax revenues makes the tax revenue claim risky and raises its discount rate; the average price/dividend ratio $\exp(pd_0^T)$ is low. The counter-cyclicality of government spending makes the spending claim safer; $\exp(pd_0^G)$ is high. As a result, when the country is running zero average primary surpluses ($\tau_0 = g_0$), as the U.S. has done over the past nine decades, the country's steady-state fiscal backing, $\tau_0 \exp(pd_0^T) - g_0 \exp(pd_0^G)$, cannot be positive, regardless of the difference between the risk-free rate and the growth rate of the economy. This result can only be overturned by rendering the tax claim safer than the spending claim, i.e. by shifting more aggregate risk onto taxpayers, who are short the tax claim, and onto transfer recipients, who are long the spending claim.

Our paper is the first to quantitatively evaluate the magnitude of the value of the surplus claim in a dynamic asset pricing model with priced aggregate risk, including the value of the seigniorage revenue derived from convenience yields on Treasurys. We develop a straightforward methodology to do so. We first do so by deriving a model-free upper bound on the PDV of surpluses. We then deploy a fully specified dynamic asset pricing model that matches a rich set of asset pricing moments for stocks and bonds. Using either method, we find a large wedge between the model estimate of fiscal backing and the actual debt/output ratio.

The surplus value can be decomposed as the present value of future surpluses, discounted using the risk-free term structure of interest rates, plus the covariance of future surpluses with the stochastic discount factor. Without aggregate risk, there is no covariance term. In this case, fiscal backing is unbounded when the average risk-free rate is lower than the average expected growth rate of the economy. Much of the literature, including recent work, has ignored the covariance term. However, in the presence of priced aggregate risk, the covariance term will typically lower the government's fiscal backing because surpluses move with the business cycle in the short run and are co-integrated with output in the long run. Our work is the first to estimate and quantify the covariance term in a realistic dynamic asset pricing model. When we insist that our model be consistent with key moments of asset prices, we find that fiscal backing is much lower than conventionally thought, and lower than the market value of outstanding debt.

The above argument relies on a realistic model of quantities and prices of risk. When modeling the quantity of risk in fiscal cash flows, adequately capturing the dynamics of government spending and tax revenue is crucial. We model the growth rates of tax revenues-to-GDP and government spending-to-GDP in a VAR alongside macro-economic and financial variables. This structure allows us to capture the cyclical properties of fiscal cash-flows. A second important feature of fiscal cash flows is that tax revenues and spending are co-integrated with GDP, so that revenues, spending, and GDP adjust when revenue-to-GDP or spending-to-GDP move away from their long-run relationships. This error correction imposes a form of long-run automatic stabilization.

The only unknowns in our valuation exercise are the risk premia on claims to taxes and spending. We deal with this in two ways. In our first approach, to guard against model misspecification, we derive a model-free upper bound on the PDV of surpluses. By setting both the risk premium on the tax and on the spending claim equal to the risk premium on a GDP claim, we obtain an upper bound on the former, a lower bound on the latter, and hence an upper bound on their difference. In the steady-state, the upper bound is given by the product of the surplus and the valuation ratio of a claim to output $(\tau_0 - g_0)\exp(pd_0^Y)$. If the output risk premium is 3%, an empirically plausible estimate of the unlevered equity risk premium, the valuation ratio $\exp(pd_0^Y)$ is 57. The government's fiscal backing, the upper bound on debt/GDP, increases by 57% per 1% of surplus/GDP. The steady-state upper bound is the average surplus-to-GDP, which is 0.05% in post-war U.S. data, times 57 or 3.15% of GDP. This is far below the average debt-to-GDP ratio in U.S. post-war data. The dynamic upper bound additionally reflects time variation in the cash-flow and the discount-rate components of the output valuation ratio. For most of the U.S. sample, the dynamic upper bound is lower than the actual debt-to-GDP ratio. At the end of 2020, the gap between the observed debt-to-GDP ratio and the upper bound exceeds 100% of GDP. This is a lower bound on the wedge with the observed debt-to-GDP ratio. The calculations incorporate additional conservatism by assuming that the spending-to-GDP ratio will converge back to its unconditional post-war mean, a strong assumption under current fiscal policies according to the Congressional Budget Office.

The U.S. Treasury earns a convenience yield on the debt it issues, making Treasury yields lower than the risk-free rate. Our paper is the first to quantitatively explore the implications of convenience yields for the budget constraint. Convenience yields λ_0 generate an additional source of revenue, $d_0\lambda_0$, which increase the surplus and steady-state fiscal backing to $(\tau_0 + d_0\lambda_0)\exp(pd_0^T) - g_0\exp(pd_0^G)$. Using the standard convenience yield estimates of Krishnamurthy and Vissing-Jorgensen (2012), the puzzle remains. We estimate seigniorage revenue of 0.11% of U.S. GDP, which raises the steady-state upper bound by only 6.26% (57 times 0.11%) of U.S. GDP. Even when we use higher convenience yield estimates, we cannot fully close the gap due to an offsetting discount rate effect. The higher the convenience yield, the higher the true risk-free rate for the same observed short-term Treasury yield. The valuation ratios of the spending and tax claims are correspondingly lower. Higher surpluses due to convenience are discounted at a higher rate.

In our second approach, we explicitly model and estimate risk premia. To do so, we posit a state-of-the-art stochastic discount factor (SDF) model. Rather than committing to a specific utility function, we use a flexible SDF that accurately prices the nominal and real term structure of Treasury bond yields, matches stock prices, and generates an equity risk premium. The SDF model's rich implications for the term structure of risk allow it to adequately price short- and long-run risk to spending and tax revenue.

Combining features from both quantities and prices of risk, the long-run discount rates on claims to tax revenues, spending, and GDP must all be equal. A claim to GDP is akin to an unlevered equity claim. In any reasonable asset pricing model with a large permanent component in the SDF, the unlevered equity risk premium exceeds the yield on a long-term government bond (Alvarez and Jermann, 2005; Hansen and Scheinkman, 2009; Borovička, Hansen, and Scheinkman, 2016; Backus, Boyarchenko, and Chernov, 2018). The discount rate for revenues and spending is high. Because of the dynamic government budget constraint, the relevant "interest rate" on the portfolio of government debt must also be high. In contrast, real-world Treasury investors seem willing to purchase government debt at low yields. The historical return on the U.S. government debt portfolio is only 1.57% in excess of the T-bill rate.

Quantitatively, we find an average surplus value equal to -320.04% of GDP, far below the average market value of outstanding government debt, 39.45% of GDP. The difference between these two, which is 359% on average over the last 75 years, quantifies the government debt valuation puzzle. The gap widens dramatically after the Great Financial Crisis (GFC) as the level of government debt/GDP rises while the valuation of the surplus claim falls. It peaks at 695% of GDP in 2020. The U.S. government has been issuing government debt while simultaneously decreasing the expected surpluses to back up the debt. The puzzle may deepen further as the government continues to incur large deficits in the wake of the COVID-19 pandemic and entitlement programs turn from surplus to deficit. The consideration of aggregate risk lowers fiscal backing substantially.

The last part of the paper studies several potential resolutions of the government bond valuation and risk premium puzzles. First, the valuation gap can be interpreted as a violation of the transversality condition (TVC) in the U.S. Treasury market, due to a rational bubble. Second, we explore the possibility of a future large fiscal correction that is absent from our sample, but present in the minds of investors who value the surplus claim. We back out from the market value of debt what annual probability investor need to assign to such an austerity event to justify the observed market value of debt. Specifically, the probability of austerity we back out implies that there is a 99.9% probability that the peso event would have realized over our sample period. This belies the nature of a peso event, and is not consistent with rational expectations. Our results can be interpreted as indirect evidence that U.S. Treasury investors seem to make systematic errors when they forecast future surpluses. We conclude that Treasurys may be mispriced relative to other asset classes. Third, the wedge can be interpreted as a measure of the value of unmeasured net government assets. Fourth, investors may be pricing in the possibility of future financial repression, i.e., an implicit tax imposed on bondholders through low real returns, analogous to what happened in the aftermath of WW-II.

Our puzzle is specific to the valuation of U.S. Treasurys in the post-war era. In follow-up work, Chen, Jiang, Lustig, Van Nieuwerburgh, and Xiaolan (2022) find no evidence of a similar bond valuation puzzle before WW-I in the U.S. or after WW-II in the U.K.

Related Literature A large literature seeks to relate the riskiness of bonds to macro-economic risks (see Lettau and Wachter, 2007; Baele, Bekaert, and Inghelbrecht, 2010; Gabaix, 2012; David and Veronesi, 2013; Campbell, Sunderam, and Viceira, 2017; Duffee, 2018; Campbell, Pflueger, and Viceira, 2020; Du, Pflueger, and Schreger, 2020). Our paper contributes to this literature by adding novel no-arbitrage restrictions on the aggregate Treasury portfolio, in addition to the no-arbitrage restrictions on individual bonds. Our contribution is to test a hitherto-untested implication of all no-arbitrage bond pricing models, namely that the value of the government bond portfolio equal the PDV of future surpluses.

Our paper contributes to the literature on the fiscal backing of the government (see D'Erasmo, Mendoza, and Zhang, 2016, for a recent review). One strand derives time-series restrictions on the government revenue and spending processes that enforce the government's inter-temporal budget constraint, starting with the seminal work of Hansen, Roberds, and Sargent (1991). Recently, the question of U.S. fiscal backing has received renewed attention (Bassetto and Cui, 2018; Blanchard, 2019; Furman and Summers, 2020; Mehrotra and Sergeyev, 2021; Mian, Straub, and Sufi, 2021;

Brunnermeier, Merkel, and Sannikov, 2024; Reis, 2021; Collin-Dufresne, Hugonnier, and Perazzi, 2023). We do not impose that debt is risk-free, as in Hansen et al. (1991). We infer large risk premia on government surpluses when no-arbitrage restrictions on bond and stock markets are imposed, resulting in lower estimates of U.S. fiscal backing.

Our paper is the first to study fiscal backing in a world with priced aggregate risk and to estimate the covariance term between the intertemporal marginal rate of substitution and the surplus, as suggested by Bohn (1995)'s seminal paper. Our main new qualitative insight is that the overall government bond portfolio is a risky asset since the government must issue debt in high marginal utility states of the world. In other words, the covariance term is negative, reducing fiscal backing. The main new quantitative result is that this covariance is large. Fiscal backing is much smaller due to this covariance term. The presence of a large amount of permanent risk in output, and by virtue of cointegration, in tax revenues, spending, and debt, is crucial for the quantitative result.

There is a parallel literature in asset pricing which tests the present value equation for stocks and other long-lived assets, starting with the seminal work by Shiller (1981); LeRoy and Porter (1981); Campbell and Shiller (1988). That work starts from the definition of a stock return to derive a testable relationship between stock prices and expected discounted dividend growth rates. Similarly, we start from the definition of the government budget constraint and derive a testable relationship between the market value of the government debt portfolio and expected discounted future surpluses. What makes this a testable restriction rather than an accounting identity is that we insist that the discount rates for surpluses be consistent with those for other securities, notably stocks. While the prices of stocks appear excessively volatile relative to their fundamentals, the value of the government debt portfolio seems excessively smooth relative to fiscal fundamentals.

Our work connects to the large literature on the convenience yield of U.S. government bonds (Longstaff, 2004; Krishnamurthy and Vissing-Jorgensen, 2012; Nagel, 2016; van Binsbergen, Diamond, and Grotteria, 2022). U.S. Treasurys are typically expensive relative to TIPS (Fleckenstein, Longstaff, and Lustig, 2014), corporate bonds (Bai and Collin-Dufresne, 2019), foreign sovereign bonds (Du, Im, and Schreger, 2018; Jiang, Krishnamurthy, and Lustig, 2021a; Koijen and Yogo, 2024), and duration-matched stocks (van Binsbergen, 2021). This paper finds that a portfolio of all U.S. Treasurys is expensive relative to the underlying collateral, a claim to surpluses, even after accounting for traditional convenience yields.

Bassetto and Cui (2018); Chien and Wen (2019); Angeletos, Collard, and Dellas (2023); Brunnermeier et al. (2024); Reis (2021) consider heterogeneous-agent models in which government bonds play a key role in allowing agents to smooth idiosyncratic risk, resulting in lower rates and larger convenience yields on Treasurys. The ability to retrade government debt in response to idiosyncratic shocks resolves the valuation puzzle in the model by Brunnermeier et al. (2024).

We contribute to a recent literature at the intersection of asset pricing and public finance. Chernov, Schmid, and Schneider (2020); Renne and Pallara (2024) argue that higher CDS premia for U.S. Treasuries since the financial crisis are related to the underlying fiscal fundamentals. Our puzzle holds in the presence of default: the value of defaultable sovereign debt is still backed by future surpluses. Liu, Schmid, and Yaron (2021) argue that increasing safe asset supply can be risky as more government debt increases corporate default risk premia despite providing more convenience. Croce, Nguyen, Raymond, and Schmid (2019) study cross-sectional differences in firms' exposure to government debt. Jiang (2021, 2022) relate fiscal cash flows to currency strength.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 presents theoretical results. Section 4 describes the VAR we use for forecasting spending and tax revenue. Section 5 derives a model-free upper bound on U.S. fiscal backing. Section 6 estimates the risk premium on taxes and spending in a dynamic asset pricing model and quantifies the puzzle. Section 7 discusses potential resolutions of the puzzle. Section 8 concludes and suggests avenues for further research. The appendix presents proofs of the propositions, and details of model derivation and estimation.

2 Data

We conduct our analysis at annual frequency, arguably a reasonable frequency to study the cash flow risk in fiscal revenues and outlays. We focus on the post-war period from 1947 until 2020. The primary surpluses are constructed using NIPA (2021) Table 3.2 from the Bureau of Economic Analysis (BEA). All variables are nominal and seasonally adjusted. Government spending is current expenditures (line 24) minus net interest payments (line 33 minus line 14). Tax revenue is the current government receipts (line 1). The primary surplus is defined as tax revenue minus government spending. Constant-maturity Treasury yields are from FRED (2021). Stock price and dividend data are from CRSP (2021); we use the CRSP value-weighted market portfolio. Dividends are summed across the months in the year. We obtain the time series of GDP from NIPA (2021) Table 1.1.5. Inflation is the change in the GDP price index from NIPA (2021) Table 1.1.4. Real GDP growth x_t is nominal GDP growth minus inflation.

We construct the market value and the total returns of the marketable government bond portfolio using CUSIP-level data from the CRSP (2021) Treasuries Monthly Series. At the end of each period, we multiply the nominal price of each CUSIP by its total amount outstanding (normalized by the face value), and sum across all issuances (CUSIPs). We exclude non-marketable debt which is mostly held in intra-governmental accounts.[2] Marketable debt includes the Treasury holdings

---

[2] The largest holders of non-marketable debt are the Social Security Administration (SSA) and the federal government's defined benefit pension plan (GDBPP). The revenues and spending from the SSA and GDBPP plan are included

of the Federal Reserve Bank. Hence, we choose not to consolidate the Fed and the Treasury, which would add reserves and subtract the Fed's Treasury holdings on the left-hand side of (1). Doing so would mainly change the duration of the bond portfolio. Following Hall and Sargent (2011) and extending their sample to 2020, we construct zero-coupon bond (strip) positions from all coupon-bearing Treasury bonds issued in the past and outstanding in the current period. This is done separately for nominal and real bonds. Since zero-coupon bond prices are also observable, we can construct the market value of outstanding marketable U.S. government debt, the left-hand side of eqn. (1) below. The average return in excess of the T-bill rate on the entire Treasury portfolio realized by an investor who buys all of the new issuances and collects all of the coupon and principal payments is 1.57% per year. The portfolio has an average duration of 3.61 years. Given the secular decline in interest rates over the forty years to 2020, the observed average realized return on the bond portfolio is, if anything, an over-estimate of investors' expected return.

3 Theoretical Results

We derive two theoretical results which are general in that they rely on the absence of arbitrage opportunities and two weak assumptions on government cash flows. The first assumption concerns the long run: tax revenues and government spending are cointegrated with GDP; they share a stochastic trend. The second assumption concerns the short-run: spending is counter-cyclical spending and tax revenues are pro-cyclical.

3.1 Valuation of Government Debt

Let G_t denote nominal government spending excluding interest expenses on the debt, T_t denote nominal government tax revenue, and $S_t = T_t - G_t$ denote the nominal primary surplus. Let $P_t^\$(h)$ denote the price at time t of a nominal zero-coupon government bond that pays \$1 at time $t+h$, where h is the maturity. There exists a multi-period stochastic discount factor (SDF) $M_{t,t+h}^\$ = \prod_{k=0}^{h} M_{t+k}^\$$, which is the product of the adjacent one-period SDFs $M_{t+k}^\$$. By no arbitrage, bond prices satisfy $P_t^\$(h) = \mathbb{E}_t\left[M_{t,t+h}^\$\right] = \mathbb{E}_t\left[M_{t+1}^\$ P_{t+1}^\$(h-1)\right]$. By convention, $P_t^\$(0) = M_{t,t}^\$ = M_t^\$ = 1$ and $M_{t,t+1}^\$ = M_{t+1}^\$$. The government bond portfolio is stripped into zero-coupon bond positions $Q_t^\$(h)$, where $Q_t^\$(h)$ denotes the outstanding face value at time t of the government bond payments due at time $(t+h)$. The outstanding debt reflects all past bond issuance decisions, i.e., all past primary deficits. Let D_t denote the nominal market value of the outstanding government debt portfolio.

---

in the federal government revenue and spending. As a result, we net out the SSA and GDBPP holdings of Treasuries, since they are an asset of one part of the consolidated government and a liability of the other part.

Proposition 1 (**Value Equivalence**)**.** In the absence of arbitrage opportunities and subject to a transversality condition, the market value of the outstanding government debt portfolio equals the expected risk-adjusted present discounted value of future primary surpluses:

$$D_t \overset{\text{def}}{=} \sum_{h=1}^{H} P_t^{\$}(h) Q_t^{\$}(h) = \mathbb{E}_t \left[\sum_{j=1}^{\infty} M_{t,t+j}^{\$} (T_{t+j} - G_{t+j}) \right] \overset{\text{def}}{=} P_t^T - P_t^G, \tag{1}$$

where the ex-dividend values of the tax claim and the spending claim are defined as:

$$P_t^T = \mathbb{E}_t \left[\sum_{j=1}^{\infty} M_{t,t+j}^{\$} T_{t+j} \right], \quad P_t^G = \mathbb{E}_t \left[\sum_{j=1}^{\infty} M_{t,t+j}^{\$} G_{t+j} \right].$$

The proof is given in Appendix A. The proof relies only on the existence of a SDF, i.e., the absence of arbitrage opportunities, not on the uniqueness of the SDF, i.e., complete markets. It imposes a transversality condition (TVC) that rules out a government debt bubble: $\mathbb{E}_t \left[M_{t,t+T}^{\$} D_{t+T} \right] \to 0$ as $T \to \infty$. Imposing the TVC rules out rational bubbles. We return to possible violations of the TVC in Section 7.2.

This valuation equation is not an accounting identity. The bond portfolio can be mispriced, just like a stock can be over- or under-valued. Eqn. (1) requires that the same SDF which prices individual government bonds and stocks also prices a claim to surpluses, i.e., the entire bond portfolio. Even when the SDF correctly prices individual bonds and stocks, this entire bond portfolio could be mispriced, for example, because agents have misspecified beliefs about future surpluses. Eqn. (1) is an accounting identity only when one does not impose any restrictions on discount rates.

When the government runs a deficit in a future date and state, it will need to issue new bonds to the investing public. If those dates and states are associated with a high value of the SDF for the representative bond investor, that debt issuance occurs at the "wrong" time. The representative investor who buys all debt issues and participates in all redemptions needs to be induced by low prices (high yields) to absorb that new debt. To see this, we can rewrite the intertemporal budget constraint, with finite horizon T, as:

$$D_t = \sum_{j=1}^{T} P_t^{\$}(j) \mathbb{E}_t \left[S_{t+j} \right] + \sum_{j=1}^{T} \text{Cov}_t \left(M_{t,t+j}^{\$}, T_{t+j} \right) - \sum_{j=1}^{T} \text{Cov}_t \left(M_{t,t+j}^{\$}, G_{t+j} \right) + \mathbb{E}_t \left[M_{t,t+T} D_{t+T} \right]. \tag{2}$$

The first term on the right-hand side is the present discounted value of all expected future surpluses, using the term structure of risk-free bond prices. It is the expected PDV for a risk-neutral investor. If the SDF is constant, this is the only term on the right-hand side. Fiscal backing is then constrained by the government's ability to generate current and future surpluses, and lower interest rates increase fiscal backing.

The second and third terms encode the riskiness of the government debt portfolio, and arise in the presence of risk-averse investors. If tax revenues tend to be high when times are good ($M^{\$}_{t,t+j}$ is low), then the second term is negative. If government spending tends to be high when times are bad ($M^{\$}_{t,t+j}$ is high), then the third term is positive. If both are true, then these covariance terms lower the government's fiscal backing. Put differently, the risk-neutral present value of future surpluses will need to be higher by an amount equal to the absolute value of the covariance terms to support a given, positive amount of government debt D_t. Our paper is the first to quantify these covariance terms, first derived by Bohn (1995) in a simple consumption-CAPM, in a realistic model of risk and return that is not subject to the equity risk premium puzzle. The covariance terms not only have the hypothesized sign, but they are also quantitatively important.

Discounting future surpluses using the term structure of risk-free interest rates, as typically done in the literature, is inappropriate. In fact, as $T \to \infty$, the first term will diverge if the average risk-free rate is lower than the average expected growth rate of the economy. Even when the debt D_{t+T} is risk-free, the last term will not converge to zero if we discount at the risk-free rate.

The valuation eqn. (1) holds ex-ante both in nominal and in real terms. Ex-post, the government can erode the real value of outstanding debt by creating surprise inflation. The same valuation equation holds allowing for sovereign default: the valuation of government debt is still backed by the value of future surpluses, but bond prices adjust to reflect the possibility of sovereign default. The proof is given in Appendix A.

3.2 Discount Rates

As tax revenue and government spending have different cyclical properties, their discount rates may differ. We define the holding period returns on the bond portfolio, the tax claim, and the spending claim as:

$$R^D_{t+1} = \frac{\sum_{h=1}^{H} P^{\$}_{t+1}(h-1)Q^{\$}_t(h)}{\sum_{h=1}^{H} P^{\$}_t(h)Q^{\$}_t(h)}, \quad R^T_{t+1} = \frac{P^T_{t+1} + T_{t+1}}{P^T_t}, \quad R^G_{t+1} = \frac{P^G_{t+1} + G_{t+1}}{P^G_t}.$$

The expected returns on these three assets are connected as follows:

Proposition 2 (**Risk Premium Equivalence**). Under the same assumptions of Proposition 1, we have:

$$\mathbb{E}_t\left[R^D_{t+1}\right] = \frac{P^T_t}{D_t}\mathbb{E}_t\left[R^T_{t+1}\right] - \frac{P^G_t}{D_t}\mathbb{E}_t\left[R^G_{t+1}\right], \tag{3}$$

where $D_t = P^T_t - P^G_t$ from Proposition 1.

The proof is given in Appendix A. The average discount rate on government debt is equal to the average discount rate on government assets, a claim to primary surpluses. Since the primary surpluses are tax revenues minus government spending, the discount rate on government debt equals the difference between the discount rates of tax revenues and government spending, appropriately weighted. By subtracting the risk-free rate on both sides, we can express the relationship in terms of expected excess returns, or risk premia. If the tax revenue claim is riskier than the spending claim and earns a higher risk premium, then the risk premium on government debt exceeds that on the revenue and the spending claims:

$$\mathbb{E}_t\left[R^D_{t+1} - R^f_t\right] > \mathbb{E}_t\left[R^T_{t+1} - R^f_t\right] > \mathbb{E}_t\left[R^G_{t+1} - R^f_t\right]. \tag{4}$$

We show below that the revenue claim is indeed riskier than the spending claim. The risk premium equivalence then implies that the portfolio of government debt ought to carry a positive risk premium. The right discount rate for government debt, given by (3), cannot be the risk-free rate.

To understand the riskiness of the debt claim, we study the short-run and long-run risk properties of the T- and G-claim. To do so, we study spending and revenue strips. A spending strip is a claim that pays off G_{t+j} at time $t+j$ and nothing at other times. A revenue strip similarly pays off T_{t+j}. Let $R^G_{t,t+j}(j)$ and $R^T_{t,t+j}(j)$ be the holding-period returns on these strips. At the short end of the maturity spectrum (business cycle frequencies j of 1–3 years), the risk premium on the revenue strip exceeds that on the corresponding-maturity spending strip:

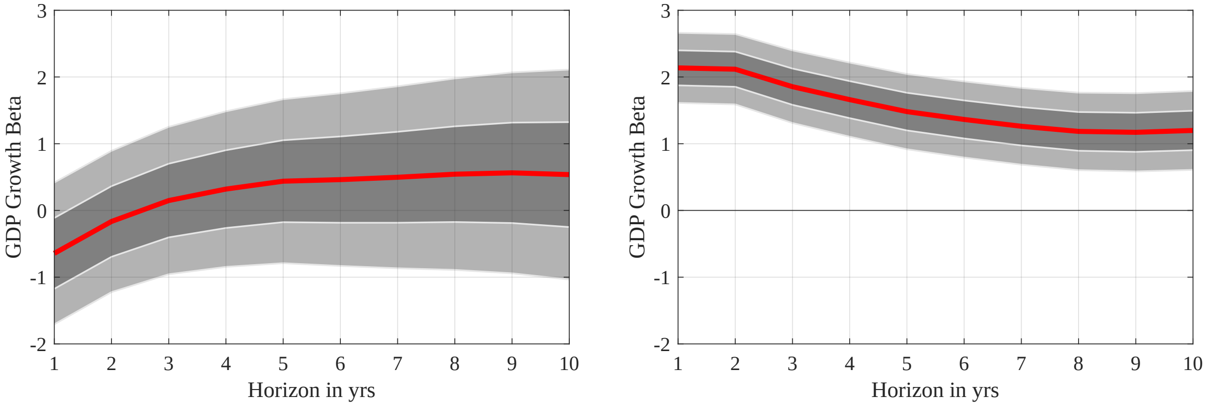
$$\mathbb{E}_t\left[R^T_{t,t+j}(j) - R^f_t\right] > \mathbb{E}_t\left[R^G_{t,t+j}(j) - R^f_t\right]. \tag{5}$$

The reason is that tax revenue is highly pro-cyclical while government spending is counter-cyclical. Figure 2 plots the cash-flow betas of U.S. government spending and taxes with respect to GDP growth over different horizons (in years). These betas are estimated from regressions of log cumulative spending or tax revenue growth on log cumulative GDP growth. At the 1-year horizon, the tax beta exceeds 2 while the spending beta is negative. At all horizons less than 10 years, the tax revenue beta exceeds the spending beta, i.e., eqn. (5) is satisfied. Since government debt investors have a long position in a riskier claim and a short position in a safer claim, the short end of the strip curve contributes to a positive risk premium on the government debt portfolio.

At the long end of the strip curve, we study the limit of the strip returns as $j \to \infty$. We denote log returns by lowercase letters.

Proposition 3. If the log of government spending/output ratio G/Y (revenue/output ratio T/Y) is stationary in levels, then the long-run expected log excess return on long-dated spending (rev-

Figure 2: Cash Flow Betas



(a) Spending Beta

(b) Tax Beta

The figure plots the GDP growth betas of log government spending and log tax revenue against the horizon (in years), computed by regressing cumulative log spending and tax revenue growth on cumulative log GDP growth. The sample period is from 1947 to 2020. Plotted with 1- and 2-standard error bands. Standard errors generated by bootstrapping 10,000 times from time-series model with cointegration for taxes (spending) and output, as well debt and output. The log of spending/output, the log of taxes/output, and the log of debt/output are AR-processes. Spending growth and tax revenue growth generated by bootstrapping with replacement from joint residuals.

enue) strips equals that on GDP strips:

$$\lim_{j\to\infty} \mathbb{E}_t \left[\log R^G_{t,t+j}(j)\right] = \lim_{j\to\infty} \mathbb{E}_t \left[\log R^T_{t,t+j}(j)\right] = \lim_{j\to\infty} \mathbb{E}_t \left[\log R^Y_{t,t+j}(j)\right] > y^\$_t(\infty). \tag{6}$$

The proof is given in Appendix A. Government spending and tax revenue have to be cointegrated with GDP; their ratio is stationary in levels. As a result, the cash-flow betas converge to one at long horizons, consistent with the empirical evidence in Figure 2. Under this realistic assumption on cash flows, expected returns on long-dated spending and tax revenue strips converge to the expected return on a long-dated GDP strip. In the presence of permanent shocks to marginal utility, the long-run discount rate on GDP is higher than the yield on long-term risk-free bonds (Alvarez and Jermann, 2005). This proposition implies that investors in the government bond portfolio have a net long position in a claim with the same long-run risk as the GDP claim. It follows immediately from this discount rate argument that the value of the long-run spending minus revenue strips will be smaller than what would be obtained when discounting with long-term bond yields.

To summarize, any model of government debt that is consistent with asset prices will have to confront two forces that push up the equilibrium return on government debt. First, there is short-run cash-flow risk that pushes the expected return on the revenue claim above the expected return on the spending claim. Second, the long-run discount rates are higher than the yield on a long-maturity bond, because of the long-run cash flow risk in the spending and revenue claims equals that of long-run GDP risk. Government debt investors have a net long position in a claim that is exposed to the same long-run cash flow risk as GDP. The low observed interest rate, or equivalently the high observed value of the government debt portfolio represents a challenge to standard dynamic asset pricing models in light of the fundamental risk of the cash flows backing that debt. Our paper is the first to highlight this tension.

In their seminal paper, Hansen et al. (1991) impose that debt has zero beta and is risk-free, i.e., the value of government debt cannot respond to any (e.g., fiscal) news. This restriction can only be satisfied if the tax revenue process is less risky than the spending process, i.e. the tax claim has a smaller beta than the spending claim (Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2020). But tax revenue is more pro-cyclical than spending as a fraction of GDP while both are cointegrated with GDP in the long run. We do not impose that government debt is risk-free. Instead, we derive the risk properties of the government bond portfolio, which follow from the measured riskiness of the actual surplus process.

3.3 Convenience Yields

Convenience yields may relieve this tension. U.S. government bonds occupy a privileged place in the world's financial system. They carry a convenience yield which makes Treasury yields lower than the safe rate of interest. The convenience yield produces an additional source of revenue, because the U.S. Treasury can sell its bonds for more than their fundamental value. The question is how far this explanation can go towards accounting for the bond valuation puzzle. We enrich the baseline model to account for the convenience of Treasury debt.

The convenience yield $\lambda_t(h)$ is the expected return on a government bond of maturity h that investors are willing to forgo under the risk-neutral measure:

$$e^{-\lambda_t(h)} = \mathbb{E}_t\left[M^{\$}_{t+1}\frac{P^{\$}_{t+1}(h)}{P^{\$}_t(h+1)}\right].$$

Proposition 4. If the TVC holds, the value of the government debt portfolio equals the value of future surpluses plus the value of future seigniorage revenue:

$$D_t = \mathbb{E}_t\left[\sum_{j=1}^{\infty} M^{\$}_{t,t+j}\left(T_{t+j}-G_{t+j}\right)\right] + \mathbb{E}_t\left[\sum_{j=0}^{\infty} M^{\$}_{t,t+j}\sum_{h=1}^{H} Q^{\$}_{t+j}(h)P^{\$}_{t+j}(h)(1-e^{-\lambda_{t+j}(h)})\right], \quad (7)$$

where D_t on the left-hand side denotes the value of the government debt portfolio at the end of period t, and $\sum_{h=1}^{H} Q^{\$}_{t+j}(h)P^{\$}_{t+j}(h)(1-e^{-\lambda_{t+j}(h)})$ on the right-hand side denotes the seigniorage revenue that the government earns from issuing the debt at a convenience yield.

The proof is given in Appendix A. When there is no convenience yield (i.e., $\lambda_t(h) = 0$), we end up back in the case of Proposition 1. If the quantity of current and future outstanding government debt is positive, then a positive convenience yield increases government revenue. This additional income is akin to seigniorage revenue and could potentially turn fiscal deficits into (broadly-defined) surpluses. At the same time, a higher convenience yield λ_t results in a lower discount factor $\mathbb{E}_t[M^{\$}_{t,t+j}]$ for a given observed bond price $P^{\$}_t(h+1) = e^{\lambda_t}\mathbb{E}_t[M^{\$}_{t,t+j}]$. This discount rate effect lowers the value of a given (broadly-defined) surplus stream. Hence, the introduction of convenience yields generates offsetting cash flow and discount rate effects on the valuation of the government bond portfolio. We investigate below which effect dominates.

4 Forecasting Tax Revenue and Government Spending

In order to quantify the value of the claims to tax revenue and government spending in eqn. (1), we need to take a stance on (i) the time-series properties of revenue and spending, and (ii) a stochastic discount factor $M_{t,t+j}$ to discount these cash flows. In this section, we describe how we model the tax and spending processes.

4.1 Cash Flow Dynamics

State Variables We assume that the $N \times 1$ vector of state variables $\boldsymbol{z}$ follows a Gaussian first-order VAR:

$$\boldsymbol{z}_t = \boldsymbol{\Psi}\boldsymbol{z}_{t-1} + \boldsymbol{u}_t = \boldsymbol{\Psi}\boldsymbol{z}_{t-1} + \boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\varepsilon}_t, \tag{8}$$

with $N \times N$ companion matrix $\boldsymbol{\Psi}$ and homoscedastic innovations $\boldsymbol{u}_t \sim i.i.d.\,\mathcal{N}(0, \boldsymbol{\Sigma})$. The Cholesky decomposition of the covariance matrix, $\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{\frac{1}{2}}\left(\boldsymbol{\Sigma}^{\frac{1}{2}}\right)'$, has non-zero elements on and below the diagonal. In this way, shocks to each state variable u_t are linear combinations of its own structural shock $\boldsymbol{\varepsilon}_t$, and the structural shocks to the state variables that precede it in the VAR, with $\boldsymbol{\varepsilon}_t \sim i.i.d.\,\mathcal{N}(0, I)$. These state variables are defined in Table 1, in order of appearance of the VAR. The vector $\boldsymbol{z}$ contains the state variables demeaned by their respective sample averages.

Our approach takes spending and tax policies as given. By including spending and taxes in the state vector, we assume that the government commits to a tax and spending policy that is affine in the state vector. Both policies are allowed to depend on a rich set of state variables with dependencies that are estimated from 74 years of data.

The VAR includes $\Delta \log \tau_t$ and $\Delta \log g_t$, the log change in tax revenue-to-GDP ratio and the log change in government spending-to-GDP ratio in its eight and tenth rows. It also includes the log level of revenue/GDP, τ_t, and spending/GDP, g_t, in its ninth and eleventh rows. This fiscal cash flow structure has three important features.

Table 1: State Variables

Position	Variable	Mean	Description	Sample Mean
1	π_t	π_0	Log Inflation	3.16%
2	$y_t^{\$}(1)$	$y_0^{\$}(1)$	Log 1-Year Nominal Yield	4.26%
3	$yspr_t^{\$}$	$yspr_0^{\$}$	Log 5-Year Minus Log 1-Year Nominal Yield Spread	0.58%
4	x_t	x_0	Log Real GDP Growth	2.95%
5	$\Delta \log d_t$	μ_d	Log Stock Dividend-to-GDP Growth	-0.18%
6	$\log d_t$	$\log d_0$	Log Stock Dividend-to-GDP Level	-3.65
7	pd_t^M	pd_0^M	Log Stock Price-to-Dividend Ratio	3.54
8	$\Delta \log \tau_t$	μ_τ	Log Tax Revenue-to-GDP Growth	0.02%
9	$\log \tau_t$	$\log \tau_0$	Log Tax Revenue-to-GDP Level	-1.74
10	$\Delta \log g_t$	μ_g	Log Spending-to-GDP Growth	0.65%
11	$\log g_t$	$\log g_0$	Log Spending-to-GDP Level	-1.75

First, our approach allows spending and revenue growth rates to depend not only on their own lags, but also on a rich set of macroeconomic and financial variables. Lagged inflation, GDP growth, interest rates, the slope of the term structure, the stock price-dividend ratio, and dividend growth all predict future revenue and spending growth. In addition, we allow innovations to the fiscal variables to be correlated with contemporaneous innovations in these macro-finance variables.

Second, we include the level variables τ_t and g_t. When there is a positive shock to spending, spending tends to revert to its long-run trend with GDP. Similarly, after a negative shock to tax revenue, future revenues tend to increase back to their long-run level relative to GDP. This mean reversion captures the presence of automatic stabilizers and of corrective fiscal action, as pointed out by Bohn (1998). By having spending/GDP growth $\Delta \log g_t$ (revenue/GDP growth $\Delta \log \tau_t$) depend on the lagged spending/GDP level $\log g_t$ (lagged revenue/GDP $\log \tau_t$) with a negative coefficient, the VAR captures this mean reversion. Mean reversion is amplified when spending-to-GDP growth $\Delta \log g_t$ ($\Delta \log \tau_t$) depends on lagged revenue-to-GDP $\log \tau_t$ ($\log g_t$) with a positive sign.

Formally, the inclusion of the levels of spending and tax revenue relative to GDP in the VAR is motivated by a cointegration analysis; the system becomes a vector error correction model. We take an a priori stance that the tax/GDP ratio $\log \tau$ and the spending/GDP ratio $\log g$ are stationary. That is, we assume a cointegration coefficient vector of $(1, -1)$ for both relationships.[3] In the absence of cointegration, all shocks to spending and tax revenues would be permanent rather than mean-reverting. Importantly, by imposing cointegration, we are being conservative about future fiscal rectitude. Large deficits relative-to-GDP, like the one that occurred in 2020, are assumed to auto-correct in the future. Thus, imposing cointegration raises the expected PDV of

---

[3] Appendix B.1 performs Johansen and Phillips-Ouliaris cointegration tests. The results support two cointegration relationships, one between log tax revenue and log GDP and one between log spending and log GDP.

future surpluses.[4]

Third, we include both the change and the level of the log dividend/GDP ratio $\log d_t$. The growth rate loads on the lagged level with a negative coefficient. This specification imposes cointegration of dividends and GDP. As a result, in the long run, the claims to taxes, spending, GDP, and aggregate dividends all earn the same risk premium because they are exposed to the same long-run risk.

In-Sample Trend in Spending/GDP One empirical issue requires further discussion. The spending-to-GDP ratio trends up in the U.S. data. As shown in Table 1, the sample average of $\Delta \log g_t$ is $\widehat{\mu}^G = 0.65\%$. Tax revenue-to-GDP is approximately stationary: the sample average of $\Delta \log \tau_t$ is $\widehat{\mu}^T = 0.02\%$. Because it is theoretically desirable to impose cointegration on the log tax/GDP and the log spending/GDP ratios, the true unconditional growth rates of the tax/GDP and the spending/GDP ratios have to be zero ($\mu_0^T = \mu_0^G = 0$).[5] To avoid biased estimates of the VAR coefficients, we cannot include trending variables in the VAR. Hence, when we estimate the dynamics of the state variables, and only then, we remove the sample averages from the growth rates. We reconstruct the log tax/GDP and log spending/GDP level variables that enter the VAR as follows:

$$\log \tau_t = \log \tau_1 + \sum_{k=1}^{t} (\Delta \log \tau_k - \widehat{\mu}^T), \quad \log g_t = \log g_1 + \sum_{k=1}^{t} (\Delta \log g_k - \widehat{\mu}^G).$$

The initial level $\log g_1$ ($\log \tau_1$) for the log spending/GDP (revenue/GDP) ratios do not affect the results. We set them to their 1947 values.

Importantly, when we price assets and value claims to spending and tax revenues, we always evaluate the state vector at the actual values of τ and g, not the detrended ones. This approach is conservative starting around 1980, because the actual spending/GDP ratio (tax/GDP) is well above (slightly below) the detrended one. Hence, the model's cash flow forecasts imply much larger future spending declines (slightly larger tax revenue increases) than we would obtain if we had used the detrended variables instead. The model predicts higher future surpluses in the last half of the sample, increasing the present-discounted value of future surpluses and hence the fiscal backing estimate.

Estimation We estimate the VAR system in eqn. (8) using OLS. The point estimates of $\mathbf{\Psi}$ are reported in Panel A of Table B.1. Lagged macro-finance variables affect fiscal variables, and vice

---

[4] The model implies that the large deficits observed at the end of our sample will mean revert, leading to a–potentially long–period where primary surpluses are high relative to GDP.

[5] Indeed, a trending spending/GDP ratio not offset by a trending tax revenue/GDP ratio would lead the surplus/GDP ratio to become arbitrarily negative given enough time. There are no instances of spending/GDP and tax revenue/GDP ratios in the U.S. or elsewhere that exceed a high level, e.g., 100%. That said, we acknowledge that a univariate statistical test for a unit root in the spending/DGP ratio cannot be rejected.

versa. Consistent with the error correction dynamics imposed by cointegration, the response of the dividend/GDP growth rate to the lagged dividend/GDP ratio ($\mathbf{\Psi}_{[5,6]}$), the response of the tax/GDP growth rate to the lagged tax/GDP level ($\mathbf{\Psi}_{[8,9]}$), and the response of the spending/GDP growth rate to the lagged spending/GDP level ($\mathbf{\Psi}_{[10,11]}$) are all negative and statistically significantly different from zero.

The dynamics of $\log d_t$, $\log \tau_t$, and $\log g_t$ in rows 6, 9, and 11 of the VAR are implied by the corresponding dynamics of their first differences $\Delta \log d_t$, $\Delta \log \tau_t$, and $\Delta \log g_t$ in rows 5, 8, and 10, respectively, with the exception of the autoregressive coefficient which is 1 minus the corresponding coefficient. There is no independent innovation to these level variables.

Panel B of Table B.1 reports the estimate of $\mathbf{\Sigma}^{\frac{1}{2}}$, the Cholesky decomposition of the innovation variance-covariance matrix. The innovation in tax revenue/GDP growth rate is positively correlated with the GDP growth rate innovation, while the spending/GDP growth shock is negatively correlated with the GDP growth shock. In other words, tax revenues are pro-cyclical and government spending is counter-cyclical, as anticipated by our earlier discussion.

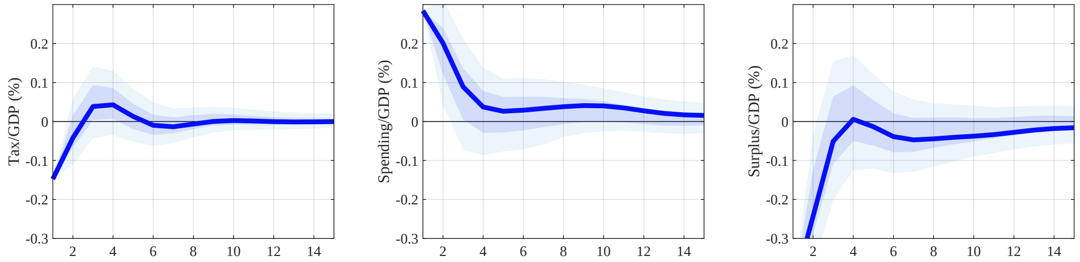
4.2 Implied Revenue and Spending Dynamics

Figure 3 plots the impulse-response functions (IRFs) of the tax revenue-to-GDP ratio (τ_t, left panels), government spending-to-GDP ratio (g_t, middle panels), and surplus-to-GDP ratio (s_t, right panels) to a GDP shock (top row), a revenue shock (middle row), and a spending shock (bottom row). The shocks are calibrated such that the log GDP growth decreases by 1%, the revenue-to-GDP ratio goes down by 1%, and the spending-to-GDP ratio goes up by 1%. The top row shows that the tax revenue-to-GDP ratio declines and the government spending-to-GDP ratio increases in response to a negative GDP shock. The surplus-to-GDP ratio is strongly pro-cyclical. The second and third rows show that mean reversion in spending and revenues brings the responses to their own shocks back to zero within about four years. The instantaneous response of the surplus to all three shocks is negative. There is some evidence of an S-shaped response of the surplus to a spending shock in the bottom right panel as the initial deficits turn into a small surplus after 3 years. However, these surpluses are short-lived, and the confidence intervals indicate that even the peak surplus response after 4-5 years is not significantly different from zero. All responses revert to zero in the long run because of cointegration between spending and GDP and between tax revenues and GDP.

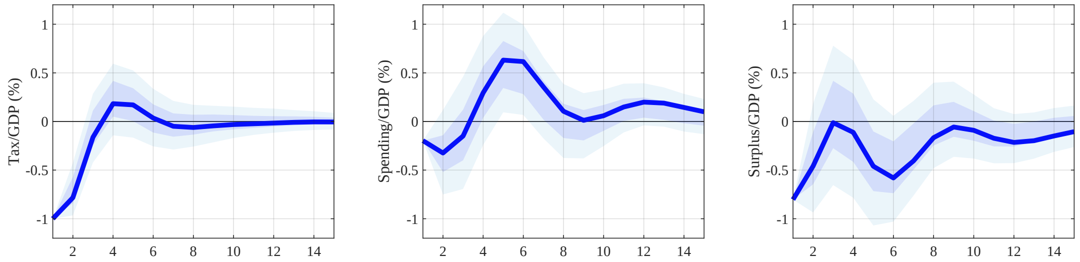
Figure 4 adds further credibility to the cash-flow projections by comparing expected cumulative spending and revenue growth over the next one, five, and ten years to realized future spending and revenue growth over those same horizons. To assess predictive accuracy, we compare the prediction of the VAR (evaluated at actual, un-detrended values of the state variables) to that of

Figure 3: Fiscal Impulse Responses

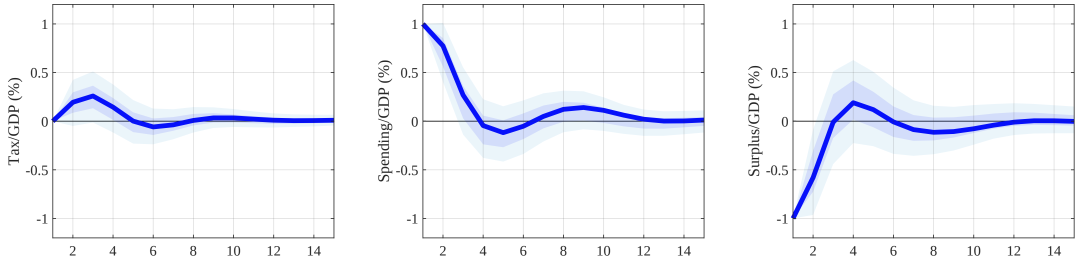
Panel A: -1% Shock to GDP Growth.



Panel B: -1% Shock to Tax-to-GDP.



Panel C: 1% Shock to Spending-to-GDP.



Solid blue line shows the impulse responses for the benchmark VAR. The impulse in the top row is a -1 percentage point shock to GDP growth x_t. The impulse in the middle row is a -1 percentage point shock to tax revenues. The impulse in the bottom row is a +1 percentage point shock to spending growth. We plot the one- and two-standard-deviation confidence intervals based on bootstrapping over 10,000 rounds.

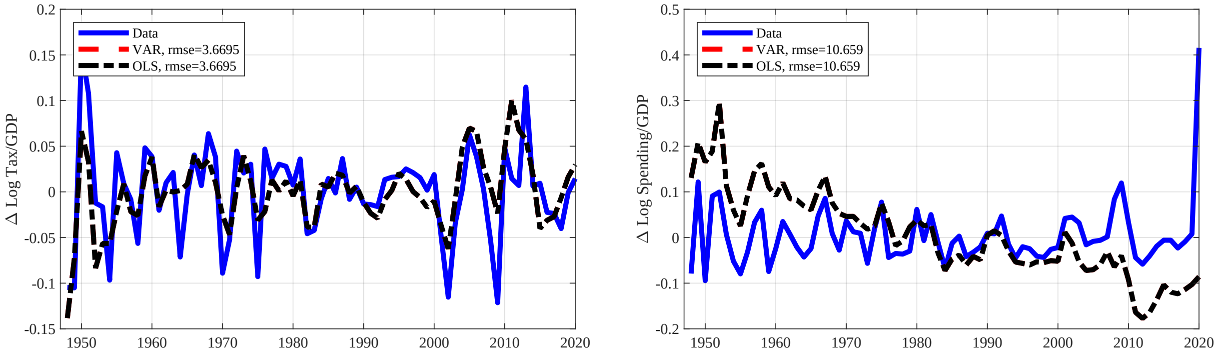
the best linear forecaster at that horizon. By design, the VAR prediction is the best linear forecaster at the one-year horizon, but not at the five- and ten-year horizons.[6] Predictive accuracy of the VAR for longer horizons is similar to that of the best linear forecast. The graph shows that the VAR implies reasonable behavior of long-run fiscal cash flows. Note how the long-run spending forecast from the VAR is low at the end of the sample. This implies that the VAR predicts too much

---

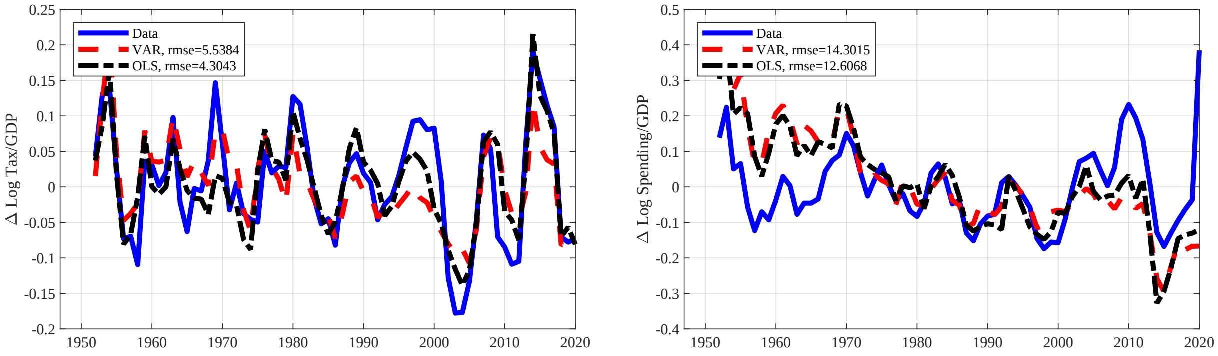
[6] Since we use the actual tax-to-GDP and spending-to-GDP series to compute the VAR predictions, but the companion matrix is estimated using the detrended series, the VAR series has a slightly higher RMSE than the OLS prediction at the one-year horizon.

Figure 4: Cash Flow Forecasts

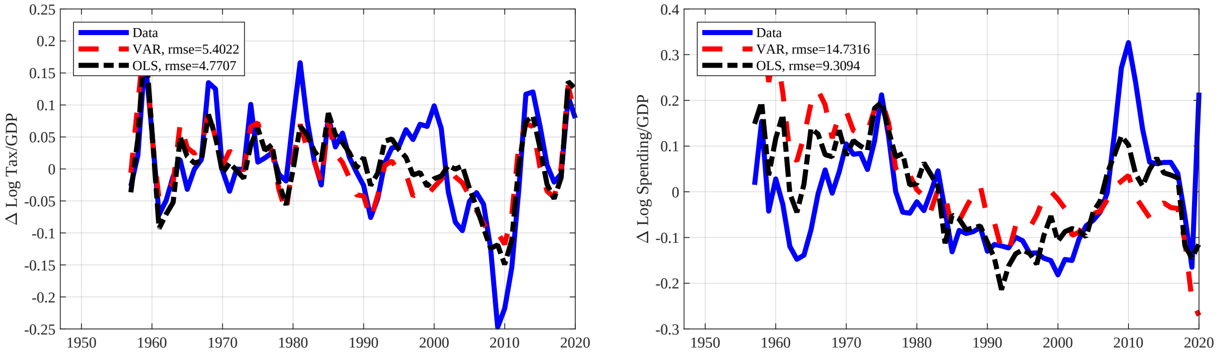
Panel A: Forecast of 1-Year Growth in Log Tax/GDP and Log Spending/GDP.



Panel B: Forecast of 5-Year Growth in Log Tax/GDP and Log Spending/GDP.



Panel C: Forecast of 10-Year Growth in Log Tax/GDP and Log Spending/GDP.



We plot the actual log tax and spending growth rates over 1-year, 5-year and 10-year rolling windows in solid blue lines. The value at each year represents the *k*-year growth rates that end at that year. We also plot these rates as forecasted by our VAR model in dashed red lines and these rates as forecasted by the OLS model in dash-dotted black lines. The value at each year represents the *k*-year growth rates condition on the (un-detrended) state variables *k* years ago. The RMSEs are in percentage points.

mean reversion in the surplus, especially compared to the high realized spending growth in the data. This understatement occurs because we evaluate the VAR forecasts at the actual value of the tax-to-GDP and spending-to-GDP ratios. The former is well below its long-run mean towards the end of the sample, while the latter is well above its mean. The error correction dynamics result in higher future tax revenue-to-GDP and lower future spending-to-GDP forecasts at the end of the sample. This forecast results in a higher present value of future surpluses and a smaller bond valuation puzzle, making it conservative.

5 Bounding Fiscal Backing

Our first approach is to develop an upper bound on fiscal backing without having to commit to a specific model of risk prices. This approach guards against misspecification of the asset pricing model and should be satisfied in any plausible asset pricing model.

By log-linearizing returns and iterating forward, we obtain the standard Campbell and Shiller (1988) decomposition of the log price/dividend ratios on the tax and spending claims:

$$\begin{aligned} pd_t^T &= \frac{\kappa_0^T}{1-\kappa_1^T} + \mathbb{E}_t\left[\sum_{j=1}^{\infty}(\kappa_1^T)^{j-1}\Delta\log T_{t+j}\right] - \mathbb{E}_t\left[\sum_{j=1}^{\infty}(\kappa_1^T)^{j-1}r_{t+j}^T\right], \\ pd_t^G &= \frac{\kappa_0^G}{1-\kappa_1^G} + \mathbb{E}_t\left[\sum_{j=1}^{\infty}(\kappa_1^G)^{j-1}\Delta\log G_{t+j}\right] - \mathbb{E}_t\left[\sum_{j=1}^{\infty}(\kappa_1^G)^{j-1}r_{t+j}^G\right], \end{aligned} \tag{9}$$

where κ_0^i and κ_1^i, for $i \in \{T, G\}$, are linearization coefficients which depend on the mean of the log price/dividend ratio pd_0^T. The derivations are detailed in Appendix C. We use rp_t^G and rp_t^T to denote the log risk premium on the spending and tax claim relative to the yield on a long-term bond. As a result, the expected log return on the tax and spending claims are given by the following expression: $\mathbb{E}_t[r_{t+1}^i] = y_t^\$(1) + yspr_t^\$ + rp_t^i$ for $i \in \{T, G\}$. Given that the short-term rate and the term spread are part of the VAR, the risk premia rp_t^T and rp_t^G are the only "free" parameters that remain to be pinned down.

Restating eqn. (1), the expected risk-adjusted PDV value of surpluses PV_t^S scaled by GDP Y_t is given by:

$$\frac{PV_t^S}{Y_t} = \frac{P_t^T}{Y_t} - \frac{P_t^G}{Y_t} = \tau_t \exp(pd_t^T) - g_t \exp(pd_t^G). \tag{10}$$

The upper bound on fiscal backing exploits two insights discussed in Section 3. First, we know from asset pricing theory that the lower bound on the expected return on the tax claim is given by the expected return on the GDP claim. The tax claim is exposed to the same long-run risk as the output claim, because of co-integration with output, but it is also more exposed to business cycle risk. Second, the upper bound on the expected return on the spending claim is given by the expected return on the GDP claim. The spending claim is exposed to the same long-run risk as the output claim, because of co-integration without output, but it is also much less exposed to business cycle risk. In summary:

$$rp_t^T \geq rp_t^Y \geq rp_t^G. \tag{11}$$

We assume that risk premia are constant at their unconditional levels rp_0^T and rp_0^G. Using the estimated VAR dynamics in eqn. (9), and denoting by $\boldsymbol{e}_z$ a selector vector that has a one in the position of some state variable z, we obtain the following expressions for the log price/dividend ratio on the tax claim:

$$pd_t^T = pd_0^T + [(\boldsymbol{e}_\pi + \boldsymbol{e}_x + \boldsymbol{e}_{\Delta\tau})'\boldsymbol{\Psi} - (\boldsymbol{e}_{y1} + \boldsymbol{e}_{yspr})'](I - \kappa_1^T \boldsymbol{\Psi})^{-1}\boldsymbol{z}_t,$$

where $(pd_0^T, \kappa_0^T, \kappa_1^T)$ solve the system of equations:

$$\begin{aligned} pd_0^T &= -\frac{(y_0^\$(1) + yspr_0^\$ + rp_0^T) - (x_0 + \pi_0)}{(1-\kappa_1^T)} + \frac{\kappa_0^T}{(1-\kappa_1^T)}, \\ \kappa_1^T &= \frac{e^{pd_0^T}}{e^{pd_0^T}+1}, \qquad \kappa_0^T = \log(1+\exp(pd_0^T)) - \kappa_1^T pd_0^T. \end{aligned} \tag{12}$$

Because of cointegration, nominal tax revenues grow at the same long-run growth rate as nominal GDP $(x_0 + \pi_0)$. The expressions for the price/dividend ratios on the spending claim and the GDP claim are analogous. We use the shorthand $\widetilde{CF}_t^T \overset{\text{def}}{=} (\boldsymbol{e}_\pi + \boldsymbol{e}_x + \boldsymbol{e}_{\Delta\tau})'\boldsymbol{\Psi}(I - \kappa_1^T\boldsymbol{\Psi})^{-1}\boldsymbol{z}_t$ and $\widetilde{DR}_t \overset{\text{def}}{=} (\boldsymbol{e}_{y1} + \boldsymbol{e}_{yspr})'(I - \kappa_1^T\boldsymbol{\Psi})^{-1}\boldsymbol{z}_t$ to denote the time-varying cash-flow and discount-rate components. They are mean-zero because $\boldsymbol{z}_t$ is mean-zero.

Using the valuation ratios of tax revenue and spending claims, we obtain the following expression for the surplus claim, i.e., the model-implied debt-to-GDP ratio:

$$\frac{PV_t^S}{Y_t} = \tau_t \exp(pd_0^T + \widetilde{CF}_t^T - \widetilde{DR}_t^T) - g_t \exp(pd_0^G + \widetilde{CF}_t^G - \widetilde{DR}_t^G). \tag{13}$$

To derive some intuition, we can evaluate eqn. (13) at $\boldsymbol{z} = 0$, i.e., when all state variables are at their unconditional mean. Then, the present value of government surplus is given by:

$$\frac{PV_t^S}{Y_t}(\boldsymbol{z} = 0) = \tau_0 \exp(pd_0^T) - g_0 \exp(pd_0^G).$$

This expression has important implications for fiscal backing. A country that runs steady-state deficits ($\tau_0 < g_0$) can only maintain positive debt capacity if the valuation of the tax exceeds that of the spending claim: $pd_0^T > pd_0^G$. But given that spending and taxes grow at the same rate as GDP in the long run, we can only satisfy this condition if $rp_0^T < rp_0^G$, i.e., if the tax revenue process is less risky than the spending process. However, that is inconsistent with eqn. (11), itself motivated by the properties of U.S. tax and spending data.

We derive an upper bound on the value of the surplus claim by maximizing the value of the tax claim and minimizing the value of the spending claim. This is accomplished by equating the expected returns on taxes and spending to the expected return on GDP:

$$rp_0^G = rp_0^Y = rp_0^T.$$

The upper bound on the present value of government surplus is given by:

$$\frac{\overline{PV}_t^S}{Y_t} = \tau_t \exp(pd_0^Y + \widetilde{CF}_t^T - \widetilde{DR}_t^T) - g_t \exp(pd_0^Y + \widetilde{CF}_t^G - \widetilde{DR}_t^G), \tag{14}$$

where the unconditional valuation ratios and the linearization constants $pd_0^Y = pd_0^T = pd_0^G$, $\kappa_0^Y = \kappa_0^T = \kappa_0^G$, and $\kappa_1^Y = \kappa_1^T = \kappa_1^G$ solve the system of equations in (12), and the mean-zero discount rate and expected cash-flow growth terms are given by:

$$\widetilde{DR}_t^T = \widetilde{DR}_t^G = \widetilde{DR}_t^Y = (\boldsymbol{e}_{y1} + \boldsymbol{e}_{yspr})'(I - \kappa_1^Y \boldsymbol{\Psi})^{-1} \boldsymbol{z}_t,$$
$$\widetilde{CF}_t^T = (\boldsymbol{e}_\pi + \boldsymbol{e}_x + \boldsymbol{e}_{\Delta\tau})' \boldsymbol{\Psi}(I - \kappa_1^Y \boldsymbol{\Psi})^{-1} \boldsymbol{z}_t, \qquad \widetilde{CF}_t^G = (\boldsymbol{e}_\pi + \boldsymbol{e}_x + \boldsymbol{e}_{\Delta g})' \boldsymbol{\Psi}(I - \kappa_1^Y \boldsymbol{\Psi})^{-1} \boldsymbol{z}_t.$$

When the state vector is evaluated at its unconditional mean ($\boldsymbol{z} = 0$), the upper bound on the debt/output ratio is given by:

$$\frac{\overline{PV}_t^S}{Y_t}(\boldsymbol{z} = 0) = \exp(pd_0^Y)\,(\tau_0 - g_0)\,. \tag{15}$$

First, this steady-state upper bound is positive if and only if steady-state surpluses are positive ($\tau_0 > g_0$). Second, the expression that matters for the (upper bound on) fiscal backing is not the risk-free rate minus the growth rate, $y_0^\$(1) - (x_0 + \pi_0)$, but the expected return on the GDP claim minus the growth rate, $(y_0^\$(1) + yspr_0^\$ + rp_0^Y) - (x_0 + \pi_0)$. When we assume a GDP risk premium of $rp_0^Y = 3\%$, $y_0^\$(1) - (x_0 + \pi_0)$ is -1.85% while $(y_0^\$(1) + yspr_0^\$ + rp_0^Y) - (x_0 + \pi_0)$ is 1.73%. A higher output risk premium lowers the valuation ratio pd_0^Y and lowers the upper bound on debt capacity. Third, an increase in pd_0^Y only increases the borrowing capacity of the government in the steady-state if the steady-state surplus is positive ($\tau_0 > g_0$). If the government runs a steady-state deficit, a higher pd_0^Y lowers fiscal backing.

In section 5.1, we evaluate the dynamic upper bound in the benchmark 1947–2020 sample. To check the robustness of the results, we explore other specifications in section 5.2, including one where we include debt in the VAR. In section 5.3 we re-estimate the VAR in a longer sample. We introduce convenience yields in section 5.4.

5.1 Benchmark Results

In our benchmark calibration, we set the GDP risk premium—defined relative to the average long-term bond yield—to 3% per year. A standard assumption in representative-agent asset pricing models is to proxy the return on the GDP claim, the total wealth return, by the unlevered return on the stock market. Unlevering a 4.5% equity risk premium on the aggregate stock market using a ratio of equity to assets of 2/3 results in a 3% unlevered risk premium.[7,8] With a GDP risk premium of $rp_0^Y = 3\%$, the expected real return on the output claim $y_0^{\$}(1) + yspr_0^{\$} + rp_0^Y - \pi_0 = 4.68\%$. In the 1947–2020 sample, the realized real growth rate of GDP x_0 is given by 2.95%. The linearization coefficient in (12) is $\kappa_1^Y = 0.98$ and the price/dividend ratio for the output claim $\exp(pd_0^Y) = 57.39$. Multiplying this steady-state price/dividend ratio with the steady-state primary surplus-to-GDP ratio of $\tau_0 - g_0 = 0.05\%$, we obtain the steady-state upper bound in eqn. (15) of 3.15% of GDP. The steady-state upper bound on the U.S. debt/output ratio is close to zero. The upper bound increases by 57.39% for each percentage point of surplus-to-GDP. For the steady-state upper bound to accommodate a debt-to-GDP ratio of 100%, the observed value in 2020, the steady-state primary surplus-to-GDP ratio would need to increase to 1.74%. This would be a dramatic reversal from the deficit in recent years, from the long-run average, and from the March 2021 Congressional Budget Office projections for the next twenty years which indicate an average annual primary deficit of 3.9% of GDP until 2050 under current law (CBO, 2021).

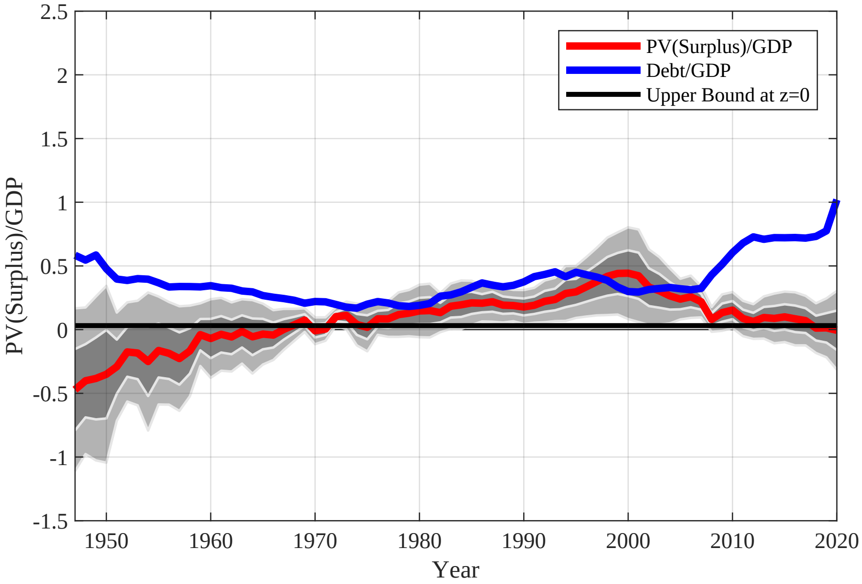
Figure 5 plots the dynamic upper bound in eqn. (14) on the U.S. debt/output ratio (in red) with two-standard error bands computed by bootstrap, as well as the steady-state upper bound of 3.15% discussed above (in black). The dynamic upper bound incorporates time-varying expected growth rates and discount rates. The latter are driven by long-term interest rates since the risk premium on the GDP claim is assumed to be constant.

---

[7] Given the 4.8% long-term bond yield, a 4.5% equity risk premium relative to the long-term bond implies a nominal expected stock return of 9.3%, close to the 9.2% historical average realized stock return in the 1947–2020 sample.

[8] More generally, aggregate dividends differ from the dividends on the current market portfolio since some of the future dividends in the economy will come from companies not yet in the current market portfolio (Panageas, 2020). If dividend growth of new entrants is as risky as the dividend growth of existing firms, then aggregate dividend growth is as risky as the dividend growth of existing firms, and the (unlevered) aggregate equity risk premium or GDP risk premium is equal to the (unlevered) risk premium on the current market portfolio. If new entrants are riskier (less risky) than existing listed firms, then the GDP risk premium is higher (lower) than the unlevered market risk premium. In Brunnermeier et al. (2024), the ratio of the equity risk premium and the GDP risk premium varies across the cycle and is counter-cyclical.

Figure 5: Upper Bound on the Value of Surpluses/GDP

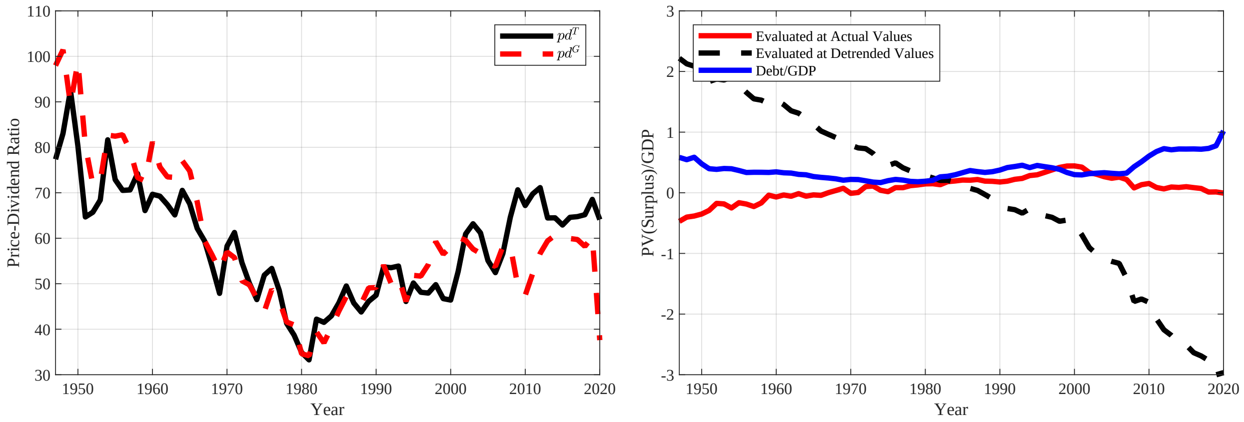


The red line plots the upper bound on the present value of government surpluses relative to GDP in eqn. (14), the solid black line the steady-state upper bound evaluated at $z = 0$ in eqn. (15), and the blue line the actual debt/output ratio. The benchmark model assumes a GDP risk premium $rp_0^Y = rp_0^G = rp_0^T$ of 3%. The sample period is from 1947 to 2020. One- and two-standard-error confidence intervals are shown in dark and light gray, respectively. To generate the standard errors for the upper bound we use a bootstrap with 10,000 samples. In each bootstrap iteration, we draw with replacement from the VAR residuals $\{\hat{u}_t\}_{t=1}^{T}$ and generate a new dataset using the VAR companion matrix. Then, we re-estimate the VAR, and re-compute the upper bound $\{\frac{\overline{PV}_t^S}{Y_t}\}_{t=1}^{T}$. We reject samples that imply negative equity risk premium.

The main result is that the dynamic upper bound on fiscal backing is below the market's valuation of U.S. Treasurys except for a brief period around the year 2000. Before the GFC, the observed market value of debt is often within the 95% confidence interval of the upper bound. However, this changes markedly after the GFC, when the debt/output ratio invariably exceeds the 95% confidence interval until the end of the sample. As of the end of our sample in 2020, the gap exceeds 100% of the GDP.

Importantly, we evaluate the upper bound $\frac{\overline{PV}_t^S}{Y_t} = \tau_t \exp(pd_t^T) - g_t \exp(pd_t^G)$ at the actual values for τ_t and g_t, including when we evaluate the valuation ratios (pd_t^T, pd_t^G). Because the VAR implies mean-reversion, this implicitly assumes that the spending-to-GDP ratio will revert to its unconditional mean of 17.40% from its 2020 value of 30.08%. As a result, the expected spending growth rate is extremely low, and therefore so is the 2020 valuation ratio of the spending claim. The left panel of Figure 6 shows the valuation ratio of the spending claim at 35 in 2020 (red dashed line). By the same logic, the valuation ratio of the tax claim at 62 is currently very high because it assumes higher future taxes (black line in left panel). If we instead had assumed that the tax-to-GDP and spending/ratios would converge to their in-sample trend, then the upper bound in 2020 would decrease from 0 to -3 times GDP, as shown by the dashed line in the right panel of Figure 6. This again illustrates the conservatism in our approach.

Figure 6: Valuation Ratios and Upper Bounds



The left panel plots the implied valuation ratios for the tax and spending claims in the benchmark case—evaluated at the actual values. We consider the upper bound case when the GDP risk premium $rp_0^Y = rp_0^G = rp_0^T$ is 3%. The sample period is from 1947 to 2020. The right panel plots the upper bound evaluated at the actual values (Benchmark) and the detrended values.

To understand this point better, note that when the average primary surplus is close to zero–as it is in the U.S. data–the cash-flow dynamics become primary determinants of fiscal backing. A first-order Taylor approximation of the upper bound evaluated at $\tau_t = g_t$ implies that:

$$\frac{\overline{PV}_t^S}{Y_t}(\tau_t = g_t) \approx \exp(pd_0^Y)\tau_t \left(\widetilde{CF}_t^T - \widetilde{CF}_t^G\right) = \exp(pd_0^Y)\tau_t \left(\boldsymbol{e}_{\Delta\tau} - \boldsymbol{e}_{\Delta g}\right)' \boldsymbol{\Psi}(I - \kappa_1^Y \boldsymbol{\Psi})^{-1} \boldsymbol{z}_t.$$

The dynamics of the short-term interest rate, the slope of the term structure, and the real GDP growth rate are irrelevant since they equally affect the expected cash-flow growth and discount rates of the T and G claims. The last term shows that what matters is the dynamics in tax-to-GDP and spending-to-GDP growth rates. The upper bound increases with expected future tax increases and spending cuts. The upper bound on fiscal backing is still around zero in 2020 despite the large primary deficit in that year because of large expected increases in future primary surpluses. As the right panel of Figure 6 illustrates, making a weaker–and possibly more realistic–assumption on the strength of that mean reversion leads to a much lower upper bound on fiscal backing.

Our method is not bound to deliver low values for the upper bound on fiscal backing. For comparison, the U.K. primary surplus over the 1947–2020 sample is 1.8%. Applying the U.S. output valuation ratio of 57 to the U.K. average surplus results in a steady-state upper bound of 102%. This bound comfortably accommodates the U.K.'s debt-to-GDP ratio in the last 70 years (see Chen et al., 2022, for detailed calculations for the U.K.).

We also study the wedge between the upper bound on fiscal backing in (14) and the market value of debt in first differences. Appendix Figure C.1 shows that, in recessions, fiscal fundamentals deteriorate and result in a downward revision to our measure of fiscal backing. At the same time, the market value of debt increases. Hence, the wedge between the two increases in recessions. The cyclical dynamics of the wedge provide a further rejection of (1).

As a robustness check, we consider different values for the GDP premium. A lower (higher) output risk premium increases (decreases) the upper bound on the PDV of surpluses. For example, a GDP risk premium of 2.5% increases the steady-state valuation ratio of the output claim $\exp(pd_0^Y)$ from 57 to 88. This change only has a small effect on the steady state upper bound, because the average deficits are close zero. Figure C.2 in the Appendix shows that the dynamic upper bound is similar as well, with wider confidence intervals. Nevertheless, the conclusion that the market value of debt substantially exceeds the upper bound after the GFC remains intact. Figure C.3 considers values as low as 1.5% for the GDP risk premium. The resulting fiscal backing bound accommodates the observed average debt/GDP ratio. However, the corresponding valuation ratios for the GDP and aggregate dividend claim are empirically implausible. In section 6, we estimate a fully-specified asset pricing model that implies an average GDP risk premium of 4.21%. Applying this higher discount rate lowers the upper bound on fiscal backing relative to the benchmark calibration.

As a further robustness check, we extend the analysis to time-varying GDP risk premia. Figure C.4 shows that the upper bound with constant and with time-varying GDP risk premia are similar, except in the late 1990s when the latter is higher.

To enforce the debt valuation in eqn. (13) as an equality, one could reverse-engineer rp_t^T:

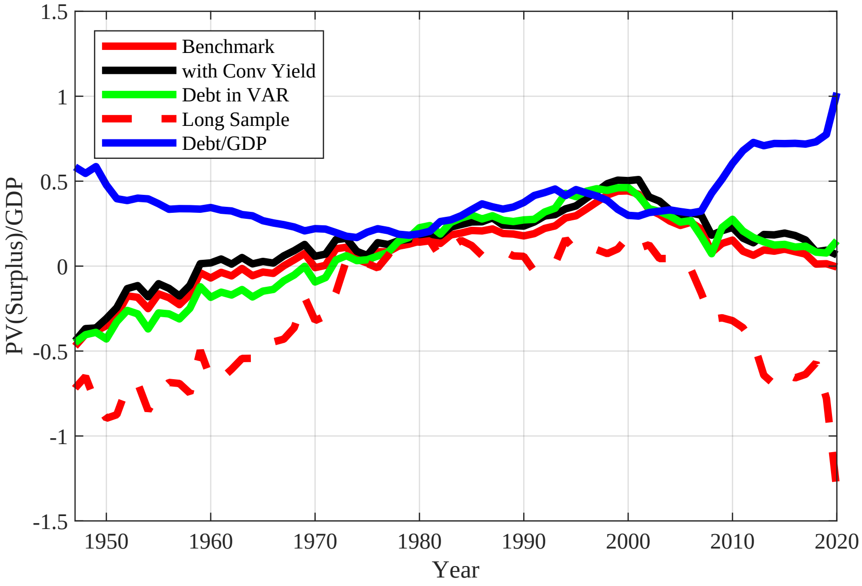
$$\frac{D_t}{Y_t} = \tau_t \exp\left(pd_0^T - \frac{rp_t^T - rp_0^Y}{1-\kappa_1^T} + \widetilde{CF}_t^T - \widetilde{DR}_t^T \right) - g_t \exp\left(pd_0^Y + \widetilde{CF}_t^G - \widetilde{DR}_t^G \right). \tag{16}$$

The right-hand side equals the upper bound in eqn. (14) for $rp_t^T = rp_0^Y$, $\forall t$. Whenever the debt-to-GDP ratio exceeds the upper bound in the benchmark model, this would require a risk premium on the tax claim that is lower than the risk premium on the GDP claim: $rp_t^T < rp_0^Y$. But this violates the risk premium condition in eqn. (11).

5.2 Debt in the VAR

Based on prior findings that highlight a fiscal response to the level of debt (Bohn, 1998; Cochrane, 2019, 2022), we also construct a version of the upper bound that includes the log debt-to-GDP ratio as a predictor variable in the state vector, allowing spending and revenue growth to depend on the lagged debt/output ratio. We include both the first-difference and the level of the log debt/GDP ratio in the VAR to impose the same error-correction dynamics as we did for spending-to-GDP and revenues-to-GDP. This way of incorporating debt in the VAR results not only in a better-behaved VAR system, but also in more realistic predictions for future debt and surplus dynamics. This approach is conservative in that it results in a stronger response of surpluses to an increase in the debt/GDP ratio than what we find in univariate tests (Jiang, Lustig, Van Nieuwerburgh, and Xiaolan, 2024). The state variables for this system are shown in Appendix Table C.1.

Figure 7: Upper Bound on the Value of Surpluses/GDP



The figure plots the upper bound on the present value of government surpluses in eqn. (14), the steady-state upper bound evaluated at $z = 0$ in eqn. (15), and the actual debt/output ratio. The GDP risk premium is 3%. We report the results from 4 different models: the benchmark (red line), the model with debt in the VAR (green line), the long sample (dashed red line), and the model with convenience yields (black line).

Figure 7 plots the actual debt/output ratio (blue line) against the dynamic upper bound implied by the model with debt in the VAR (green line). The upper bound is quite similar to that in the benchmark model (red line). It is slightly lower before 1975, slightly higher from 1975 to 2000, and slightly lower after the GFC. The upper bound on fiscal backing in the extended model with debt cannot accommodate the observed U.S. debt-to-GDP ratio between 1950 and 1970 or after the GFC in 2008, despite several sources of conservatism built into the construction of the upper bound. The similarity to the benchmark results arises from the fact that the VAR without debt already builds in a strong increase in expected future surpluses in response to a negative GDP shock (or a negative shock to the surplus). Adding debt in the VAR does not change the impulse-response function of primary surpluses much.

5.3 Longer Sample

We study robustness to a longer sample using U.S. data form 1929–2020. In the longer sample, the valuation ratio $\exp(pd^Y)$ is 53.90 and the steady state surplus $\tau_0 - g_0$ is given by -0.55%. The U.S. has been running a substantial average primary deficits over the past 92 years. As a result of the lower average surplus, we obtain a lower steady-state upper bound of -29.42%. For the steady-state upper bound to accommodate a debt/GDP ratio of 100%, the steady-state surplus-to-GDP ratio would need to increase by 2.40% to 1.86%.

Figure 7 plots the dynamic upper bound in the longer sample (dashed red line). The observed debt/output ratio now always exceeds the upper bound. Using these estimates, the gap between the upper bound and the current debt/output ratio increases to over 2 times GDP in 2020.

In the longer sample, the spending/output and the tax/output ratio and their growth rates have a different mean. This results in more muted mean-reversion dynamics and hence a larger gap between the surplus value and the observed debt-to-GDP ratio at the end of the sample. Appendix C.2 provides more details on the VAR estimation for the longer sample. Figure C.6 includes confidence intervals and explores different values of the output risk premium. The results are not very sensitive to the choice of output risk premium.

Chen et al. (2022) take an even longer-term perspective, using U.S. data going back to the 18th century and contrasting the U.S. and the U.K. experience. They find that there is no government bond valuation puzzle for the U.S. in the period before WW-II. In contrast, there is a substantial valuation puzzle for the U.K. in that period, when the U.K. was the global safe asset provider.

5.4 Convenience Yields

As discussed in Section 3.3, U.S. Treasurys earn a convenience yield, which produces additional revenue for the U.S. government. The question we address here is how far this explanation can go towards accounting for the bond valuation puzzle.

We use the variable k to represent the seigniorage revenue from convenience as a fraction of output. Building on Prop. 4, we can rewrite the intertemporal budget constraint as:

$$D_t = \mathbb{E}_t\left[\sum_{j=1}^{\infty} M^{\$}_{t,t+j}(T_{t+j} + K_{t+j})\right] + K_t - \mathbb{E}_t\left[\sum_{j=1}^{\infty} M^{\$}_{t,t+j} G_{t+j}\right].$$

The observed nominal Treasury yield $y^{\$}_t(1) = \rho^{\$}_t(1) - \lambda_t(1)$ equals the difference between the true one-period nominal risk-free rate $\rho^{\$}_t(1)$ and the convenience yield $\lambda_t(1)$. We include the one-period nominal risk-free rate as the second element of the VAR, replacing the nominal short rate: $z(2,t) = \rho^{\$}_t(1) = y^{\$}_t(1) + \lambda_t(1)$. We continue to assume that $rp^Y_0 = rp^G_0 = rp^T_0$ to derive the upper bound.

Evaluated at the steady-state, the upper bound on the debt/output ratio with convenience is:

$$\frac{\overline{PV}_t^S}{Y_t}(z=0) = \exp(pd_0^Y)\,(\tau_0 + k_0 - g_0) + k_0, \tag{17}$$

where pd_0^Y is the valuation ratio of the GDP claim computed above. A country may now be able to run steady-state deficits if the seigniorage revenue turns the primary deficit into a surplus.

To proxy for the convenience yield $\lambda_t(1)$, we first construct the spread cy_t between the 3-month Treasury yield and a risk-free benchmark, which is the 3-month CD rate from 1964 and the 3-month banker's acceptance rate before 1964. Panel A of Figure B.1 plots this spread. The average spread cy_0 is 0.56% per year over the period 1947–2020. Next, we assume that Treasury bonds with longer maturities earn lower convenience yields. Specifically, bills with maturities within 1 year earn 100% of cy_t, 1-year bonds earn 90% of cy_t, 2-year bonds earn 80% of cy_t, and so on. Treasury bonds with 10-year maturities no longer earn a convenience yield.[9] We construct the seigniorage revenue in each period based on the amount of bounds outstanding of each maturity. Panel B of Figure B.1 shows the resulting time series. We find an average seigniorage revenue k_0 of 0.11% of U.S. GDP. As a result, the steady-state surplus $(\tau_0 + k_0 - g_0)$ is 0.16%, compared to 0.05% without convenience.

A narrow measure of convenience yield only affects Treasuries, leaving expected returns on other assets such as stocks or a claim to GDP unchanged. A given expected stock (GDP claim) return–as measured in the data–would then imply a lower equity (GDP) risk premium in the presence of a higher true risk-free rate (observed short rate plus convenience yield). It follows that the measured average excess return on risky assets that do not earn a convenience yield puts an upper bound on the narrow convenience yield.

Hence, we decrease the output risk premium rp_0^Y by 0.28% per year, our estimate of the convenience yield at the 5-year horizon. The average valuation ratio $\exp(pd_0^Y)$ becomes 58.03, very close to its value in the benchmark model without convenience of 57.39.

When the state vector is evaluated at its long-run mean ($z = 0$), the upper bound on the debt/output ratio is given by:

$$\frac{\overline{PV}_t^S}{Y_t}(z=0) = \exp(pd_0^Y)\,(\tau_0 + k_0 - g_0) + k_0 = 58.03 \times 0.16\% + 0.11\% = 9.56\%.$$

This is a modest increase in fiscal backing compared to the benchmark estimate of 3.15%.

---

[9] In their shorter 2000–2016 sample, Du et al. (2018) estimate a convenience yield of 26 basis points at the 3-month maturity and 3.5 basis points at the 10-year maturity. The literature on swap spreads (Jermann, 2020; Du, Hébert, and Li, 2023) also points out that, if anything, long-term Treasuries are "inconvenient," as Treasury yields are actually higher than same-maturity swaps.

We compute the dynamic upper bound, by adding seigniorage revenue as part of tax revenues and including the risk-free rate instead of the nominal short rate in the VAR. The derivation is in Appendix C.2. The black solid line in Figure 7 plots the upper bound with convenience yields. It is only modestly above the benchmark upper bound. Since the GFC, the gap between the actual debt/output ratio and the upper bound with convenience yields has widened dramatically, just like in the benchmark model. In conclusion, standard measures of (narrow) convenience yields do not resolve the bond valuation puzzle.

5.4.1 A Broader Measure of Convenience

Recently, some have argued that convenience yields may be much larger than previously thought, as high as than 200 bps per annum (Jiang et al., 2021a; Koijen and Yogo, 2024). Large convenience yields are likely to be broad convenience yields, applying to assets beyond Treasuries, for example to all U.S. dollar-denominated assets of high credit quality.

In its starkest form, a broad convenience yield applies equally to government bonds and the GDP claim, raising the true risk-free rate and the true expected return on the GDP claim, purged of convenience yield effects, by the same amount equal to the convenience yield. The GDP risk premium is unaffected. The increase in the true expected return on the GDP claim lowers the valuation ratio pd_0^Y. As a result, the government generates less debt capacity per percentage point of (seigniorage revenue-augmented) surplus/output. Discounting both primary surpluses and seigniorage revenues from convenience at a higher discount rate lowers the expected PDV of surpluses inclusive of seigniorage revenues. This negative discount rate effect offsets the positive cash flow effect of higher seigniorage revenues from higher convenience yields.[10,11] This argument shows that higher convenience yields are not a panacea for resolving the valuation puzzle.

An intermediate case between the stark-form broad convenience yield and the narrow convenience yield may be more realistic. In such case, a broad convenience yield would have a larger effect on Treasuries than on other assets. A much higher convenience yield then results in a much higher risk-free rate and a modestly higher expected return on the GDP claim. The GDP risk premium falls less than one-for-one with the magnitude of the convenience yield.

As one increases the (broad) convenience yield λ_0 by enough, the discount rate effect comes to dominate the cash flow effect; the marginal effect of higher convenience yields on the upper

---

[10] In contrast, a narrow convenience yield measure only affects the expected return on Treasurys and not the expected return on other asset classes. An increase in the narrow convenience yield raises the true risk-free rate but not the expected return on other assets. The risk premium on other assets falls one-for-one with the size of the convenience yield. For the narrow measure of convenience, there is no offsetting discount rate effect. This was the case we considered in Figure 7.

[11] By the same token, a large broad convenience yield makes a TVC violation less likely.

bound turns negative. To see this, note that:

$$\frac{\partial}{\partial \lambda_0}\left(\frac{\overline{PV}_t^S}{Y_t}(z=0)\right) \leq 0 \iff \frac{\tau_0 + k_0 - g_0}{(1-\kappa_1^Y)}\left(1 + \frac{\partial}{\partial \lambda_0} rp_0^Y\right) \geq d_0\left(1 + \exp(-pd_0^Y)\right).$$

The proof is given in Appendix A. A marginally higher convenience yield lowers fiscal backing if the marginal discount rate effect, given by the term on the left of the second inequality, exceeds the marginal cash-flow effect, given by the debt/GDP ratio d_0. First consider the case where $\frac{\partial rp_0^Y}{\partial \lambda_0} = 0$, the stark form of the broad convenience yield. The term $(\tau_0 + k_0 - g_0)/(1-\kappa_1^Y)$ is approximately the PDV of the steady-state surplus including seigniorage. At low levels of the convenience yield λ_0, seigniorage revenue-to-GDP k_0 is low, making $(\tau_0 + k_0 - g_0)/(1-\kappa_1^Y)$ smaller and the derivative of the fiscal backing bound w.r.t. the convenience yield more likely to be positive. The cash-flow effect dominates. At high enough levels of the convenience yield, k_0 is high enough to produce $(\tau_0 + k_0 - g_0)/(1-\kappa_1^Y) > d_0\left(1 + \exp(-pd_0^Y)\right)$, and the derivative turns negative. The discount rate effect dominates. In sum, high broad convenience yields cannot generate high fiscal backing since fiscal backing shrinks in the convenience yield for high levels of the convenience yield. For the intermediate case of the broad convenience yield, $0 < 1 + \frac{\partial rp_0^Y}{\partial \lambda_0} < 1$, the argument continues to go through but the sign switch occurs at higher levels of the convenience yield.[12]

5.5 Global VAR

As a final robustness check, we add global ex-U.S. real GDP growth, inflation, and stock returns to the VAR. They are computed based on 17 major developed economies. Those foreign variables may be relevant for U.S. fiscal dynamics. Appendix E.3 shows that the upper bound on fiscal backing is very similar to that in the benchmark model. It also shows that the results for the asset pricing model, described in the next section, are similar in a model where these global state variables affect the pricing of U.S. stocks and bonds.

6 Measuring Fiscal Backing

The upper bound exercise in the previous section assumes equal and constant risk premia rp_0^T and rp_0^G on the tax and spending claims. In this section, we infer these risk premia from asset prices

---

[12] In the narrow convenience yield case, $1 + \frac{\partial rp_0^Y}{\partial \lambda_0} = 0$. There only is a cash-flow effect, so that a higher (narrow) convenience yield always increases the fiscal backing bound. A very large narrow convenience yield is implausible since it results in negative risk premia. If a 5% convenience yield only accrued to government bonds, then the implied risk premium on unlevered equity would be $3\% - 5\% = -2\%$.

by using a dynamic asset pricing model. This allows us to quantify the PDV of surpluses and gap with the value of debt.

6.1 Asset Pricing

We choose a flexible SDF model that only assumes no arbitrage, and use it to price the term structure of interest rates as well as stocks. This approach guarantees that our debt valuation is consistent with observed Treasury bond yields. It also results in an SDF that has enough permanent risk to account for the equity risk premium (Alvarez and Jermann, 2005).[13] The nominal SDF $M^{\$}_{t+1} = \exp(m^{\$}_{t+1})$ is conditionally log-normal:

$$m^{\$}_{t+1} = -y^{\$}_t(1) - \frac{1}{2}\boldsymbol{\Lambda}'_t\boldsymbol{\Lambda}_t - \boldsymbol{\Lambda}'_t\boldsymbol{\varepsilon}_{t+1}.$$

The real SDF is $M_{t+1} = \exp(m_{t+1}) = \exp(m^{\$}_{t+1} + \pi_{t+1})$, which is also conditionally Gaussian. The priced sources of risk are the structural innovations in the state vector $\boldsymbol{\varepsilon}_{t+1}$ from eqn. (8). We use the VAR with the state vector listed in Table 1.

These aggregate shocks are associated with a $N \times 1$ market price of risk vector $\boldsymbol{\Lambda}_t$ of the affine form:

$$\boldsymbol{\Lambda}_t = \boldsymbol{\Lambda}_0 + \boldsymbol{\Lambda}_1 \boldsymbol{z}_t,$$

The $N \times 1$ vector $\boldsymbol{\Lambda}_0$ collects the average prices of risk while the $N \times N$ matrix $\boldsymbol{\Lambda}_1$ governs the time variation in risk premia. This setting allows us to derive analytical solutions for bond yields and stock price-dividend ratios, as detailed in Appendix D. Asset pricing in this model amounts to estimating the market prices of risk in $\boldsymbol{\Lambda}_0$ and $\boldsymbol{\Lambda}_1$.

For asset pricing purposes, we cannot include the debt/output ratio, a near-unit-root variable (Jiang et al., 2024), in the state vector because it imputes too much persistence to long yields. Including it in the no-arbitrage model results in the debt/output ratio accounting for a counterfactually large share of the variation in long-term bond yields.[14]

---

[13] The asset pricing model combines a vector auto-regression model for the state variables as in Campbell (1996) with a no-arbitrage model for the (SDF) as in Duffie and Kan (1996); Dai and Singleton (2000); Ang and Piazzesi (2003). Using a similar approach, Lustig, Van Nieuwerburgh, and Verdelhan (2013) study the properties of a claim to aggregate consumption, the wealth-consumption ratio, and Gupta and Van Nieuwerburgh (2021) evaluate the performance of private equity funds. This empirical SDF is to be interpreted as the SDF that prices U.S. stocks and individual bonds. Since those assets are held globally, it is not the U.S. SDF. There is no presumption of rationality; the empirical SDF could encapsulate belief distortions or non-standard preferences.

[14] In a no-arbitrage model, including the market value of debt implies a process for the risk-neutral dynamics of the surpluses, which obviates the need for including the actual surpluses. A complementary exercise to the one in this paper is to back out these risk-neutral surplus dynamics in a model that does not include tax revenues and government spending in the VAR. Jiang, Lustig, Van Nieuwerburgh, and Xiaolan (2021b) shows that this exercise results in a similar conclusion.

Estimation We estimate the model's risk prices by minimizing the distance between several bond and stock price moments in model and data. Appendix E spells out the moments and reports the point estimates for the market price of risk parameters. It shows that the model provides a tight fit for the entire time series of nominal bond yields of the various maturities. The fit for real bond yields is reasonable. The model closely matches the dynamics of the nominal bond risk premium. Finally, the model produces a good fit for the equity risk premium levels and dynamics, for the time-series of the price-dividend ratio on the aggregate stock market, and for the dividend strip prices which discipline the pricing of short-term versus long-term risk.

6.2 Government Surplus Pricing

Recall that the model-implied present value of government surpluses is given by (10). When the government runs deficits, one needs a larger valuation ratio for the tax claim than the spending claim ($pd_t^T > pd_t^G$) to generate a positive value for the PDV of surpluses. However, our estimates will reveal that the tax claim is riskier than the spending claim. This discount-rate effect needs to be offset by generating higher expected growth of tax revenues in the short run when the government runs deficits. As we show, this cash-flow effect is not strong enough in the data.

We use $r_0^T = \mathbb{E}[r_t^T]$ to denote the unconditional expected log return on the tax claim. Using the log-linearized returns, the log of the valuation ratio on the tax claim is affine in the state vector:

$$pd_t^T = pd_0^T + (\bar{\boldsymbol{B}}^T)' \boldsymbol{z}_t, \tag{18}$$

where the constant valuation ratio is:

$$pd_0^T = -\frac{r_0^T - (x_0 + \pi_0)}{(1-\kappa_1^T)} + \frac{\kappa_0^T}{(1-\kappa_1^T)},$$

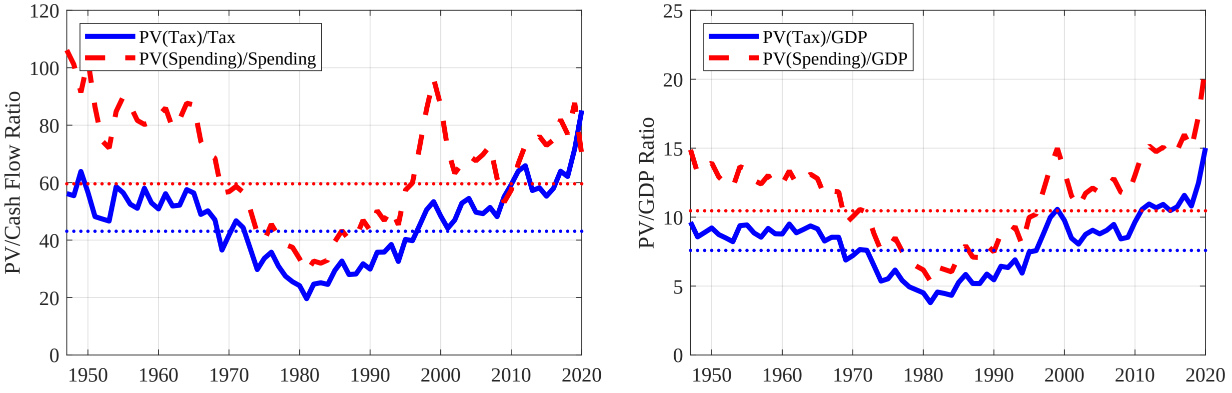
with κ_0^T and κ_1^T as defined above. A similar expression applies to the spending claim. The proof is in Appendix D.4.

Importantly, differences in the average valuation ratios are driven exclusively by discount rates effects, because the unconditional cash flow growth rates of tax revenue and spending claims are identical due to cointegration with GDP. If we equate the expected returns on the output, spending and tax revenue claims ($r_0^G = r_0^T = r_0^Y$), we recover the upper bound in eqn. (15).

Instead, we now use the estimated risk prices $\boldsymbol{\Lambda}_0$ in the expression for the unconditional risk premium on the tax claim:

$$r_0^T - y_0^{\$}(1) + Jensen = \left(\boldsymbol{e}_{\Delta\tau} + \boldsymbol{e}_x + \boldsymbol{e}_\pi + \kappa_1^T \bar{\boldsymbol{B}}^T\right)' \boldsymbol{\Sigma}^{\frac{1}{2}} \boldsymbol{\Lambda}_0, \tag{19}$$

Figure 8: DAPM Valuations of Taxes and Spending



(a) Normalized by Current Cash Flow

(b) Normalized by Current GDP

The figure plots the present values of tax revenues and of government spending. Both time series are scaled by their respective current cash flows in the left panel, and by the current GDP in the right panel. The sample is annual, 1947–2020. The dotted lines represent the values when the state variables are at $z_t = 0$.

The left-hand side of this equation is the unconditional expected excess log return with Jensen adjustment. The-right-hand side is the unconditional covariance of the log SDF with the log return. The Jensen's term is given by one half of the unconditional variance of the log returns: $\frac{1}{2}\left(\boldsymbol{e}_{\Delta\tau}+\boldsymbol{e}_x+\boldsymbol{e}_\pi+\kappa_1^T\bar{\boldsymbol{B}}^T\right)'\boldsymbol{\Sigma}\left(\boldsymbol{e}_{\Delta\tau}+\boldsymbol{e}_x+\boldsymbol{e}_\pi+\kappa_1^T\bar{\boldsymbol{B}}^T\right)$. Similarly, we can calculate the time-varying component of the risk premium on the tax claim:

$$\left(\boldsymbol{e}_{\Delta\tau}+\boldsymbol{e}_x+\boldsymbol{e}_\pi+\kappa_1^T\bar{\boldsymbol{B}}^T\right)'\boldsymbol{\Psi}-(\bar{\boldsymbol{B}}^T)'-(\boldsymbol{e}_{yn})'=\left(\boldsymbol{e}_{\Delta\tau}+\boldsymbol{e}_x+\boldsymbol{e}_\pi+\kappa_1^G\bar{\boldsymbol{B}}^T\right)'\boldsymbol{\Sigma}^{\frac{1}{2}}\boldsymbol{\Lambda}_1, \tag{20}$$

where $\boldsymbol{\Lambda}_1$ governs the time-varying component of the market prices of risk. The left-hand side is the time-varying component of the expected excess log return. The right-hand side is the time-varying component of the conditional covariance of the log SDF with the log return on the tax claim. Using this equation, we can recover the conditional expected return and price-dividend ratio for the tax claim. The same procedure applies to the spending claim.[15]

The left panel of Figure 8 plots the present value of tax revenue normalized by tax revenue, P_t^T/T_t. In steady-state (at $\boldsymbol{z} = 0$), investors are willing to pay 43.08 times annual tax revenue for the tax claim (horizontal dotted line). They are willing to pay larger multiple of 59.60 for the spending claim. The right panel of Figure 8 plots the present value of tax revenue normalized by GDP, P_t^T/GDP_t. Investors willing to pay 8.12 times the annual GDP on average for the right

---

[15] The implied value of surpluses over output, PV^S/Y, is not log-linear in the state vector. When the government adheres to an affine spending and tax rule, the surplus value is not affine. In the exponentially-affine framework, it is convenient to work with tax revenues/GDP and government spending/GDP, each of which is always strictly positive, so that logs can be taken. Since the surplus/GDP ratio is frequently negative in the data, its log is not well-defined.

to receive all current and future tax revenues. The value of the tax claim displays substantial time-variation. With the exception of the late 1990s, a V-shape arises, which is inherited from the inverse V-shape of long-term real interest rate. Real rates are high in the mid-1970s to mid-1980s and low at the beginning and end of the sample. Discounting future tax revenues by a low (high) long-term real rate results in a high (low) valuation ratio. The time-series average of the present value of government spending normalized by GDP, P_t^G/GDP_t, is 11.32. The higher valuation of the spending claim than the tax revenue claim reflects the higher risk in tax revenues than in spending, (in part) because of the pro–cyclicality of tax revenues and the counter-cyclicality of government spending in the short run.

Now we are in a position to value the claim to future government surpluses as the value of the tax claim minus the value of the spending claim, the right-hand side of eqn. (1). When evaluated at $z = 0$, the steady-state value of the debt/output ratio is:

$$\frac{PV_t^S}{Y_t}(z=0) = \tau_0 \exp(pd_0^T) - g_0 \exp(pd_0^G) = 17.60\% \times 43.08 - 17.55\% \times 59.60 = -287.64\%.$$

The gap between the steady-state value of -287.64% and the steady-state upper bound of 3.15% of GDP is entirely due to the higher risk premia for the tax claim than for the spending claim because there are only discount rate effects and no cash flow effects in the steady state.

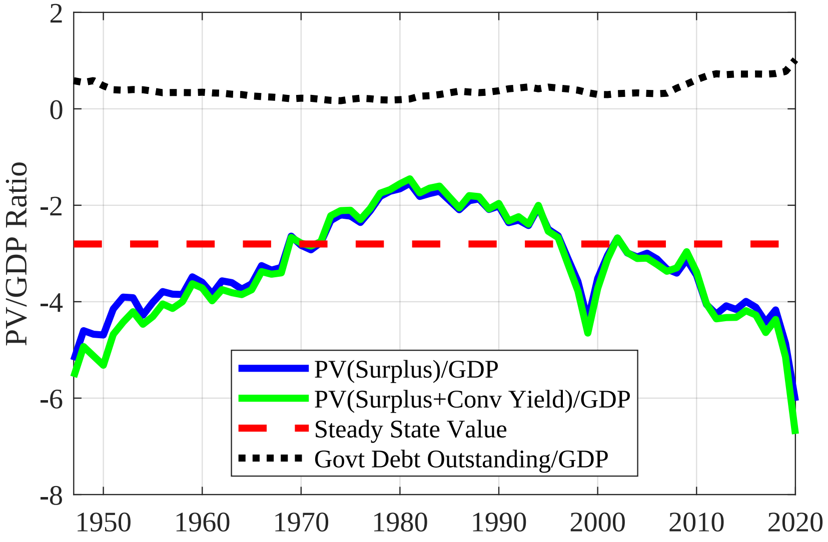
Figure 9 plots the present value of government surpluses as the solid blue line. Its average value of -320.04% is far below the average market value of outstanding government debt, 39.45% of GDP. We also include the dashed red line, which indicates the steady-state PDV of government surpluses. The valuation gap measures the difference between the market value government debt and the present value of surpluses. It quantifies the government debt valuation puzzle. The gap/GDP ratio is 359% on average. In the time series, the gap widens substantially in the last 20 years of the sample. The level of government debt rises to 55.21% of the GDP while the valuation of the surplus claim decreases to -374% of GDP. In other words, the U.S. government has been issuing government debt while simultaneously decreasing the expected surpluses to back up the debt. The effect is particularly stark in the last few years of the sample. The gap reaches 695% of the GDP in 2020.[16] The bond valuation puzzle may deepen further due to a large fiscal response to the covid pandemic in 2021 and 2022 and growing deficits from entitlement programs thereafter.

We reiterate that imposing cointegration between tax revenues and GDP and spending and GDP is not only imperative to accurately describe fiscal dynamics but also leads to conservative estimates for the gap/GDP ratio. Without the error correction dynamics present in our VAR sys-

---

[16] At the end of the sample, interest rates are low and expected returns on equity are low relative to expected dividend growth rates. This makes valuations of bonds and stocks, but also of the surplus claim, naturally sensitive to changes in the interest/discount rate. Expressing the present value of surpluses relative to the present value of GDP, shown in Figure E.11, displays substantially less sensitivity and shows less of a decline at the end of the sample.

Figure 9: Present Value of Government Surpluses and Market Value of Government Debt



The figure plots the present values of government surpluses and the market value of government debt. Both time series are scaled by the current U.S. GDP. We also estimate a model that incorporates bond convenience yields as in Section 5.4 and report the present value of government surpluses and seigniorage revenues.

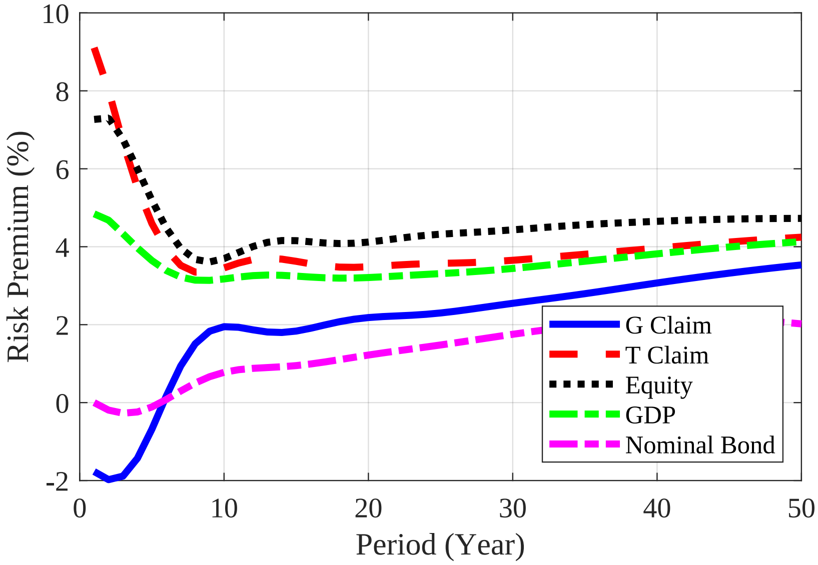
tem, an increase in government spending following a recession is not offset by future reductions in spending or future increases in tax revenues, but rather becomes permanent. In such a world, the spending claim would be much safer and the tax revenue claim much riskier, leading to a much more negative present value of government surpluses and a much larger valuation gap (recall the right panel of Figure 6).

Finally, we estimate our asset pricing model that incorporates bond convenience yields as in Section 5.4, and compute the model-implied present value of surpluses plus seigniorage revenues as the green line in Figure 9. We find a slightly larger gap between this measure of fundamental value and the market value of government debt, despite positive seigniorage revenues. The presence of bond convenience yields in Treasuries implies a higher ex-convenience risk-free yield. Given the same observed equity returns, the equity risk premium in excess of the true risk-free rate declines. The lower risk premia result in higher valuation ratios for tax and spending claims, as well as a larger difference between them. Therefore, consistent with our discussion in Section 5.4.1, the discount rate effect of the convenience yield dominates the cash flow effect, leading to a more negative present value of surpluses of -329.57% of GDP on average, compared to -320.04% of GDP in our baseline case.

6.3 Risk Premia on Tax and Spending Strips

To further understand the drivers of the negative value of the surplus claim, Figure 10 plots the risk premia on revenue and spending strips for maturities from 1 to 50 years predicted by the asset pricing model. It also plots the risk premia on GDP and aggregate dividend strips and on nominal bonds for comparison. At the short end of the maturity spectrum, risk premia on spending strips are very low (-1.77% at the one-year horizon). Because spending is counter-cyclical, these strips are a hedge. In contrast, short-maturity tax revenue strips have high risk premia (9.11% at the one-year horizon) because tax revenues are low in high marginal utility times, making the tax claim a risky asset. Hence, at the short end, the inequality hypothesized in eqn. (5) is satisfied.

Figure 10: Term Structure of Risk Premia on the T-Claim and the G-Claim



This figure plots the cumulative risk premia on the spending strips, the tax strips, equity strips, and the GDP strips in our benchmark model against the holding period. Each point is an annualized holding-period risk premium, as derived in eqn. (D.13) of the Appendix.

As we move to long maturities, risk premia on revenue and spending strips converge towards each other. As noted in eqn. (6), since tax and spending are cointegrated with the GDP, their risk premia also converge towards the risk premium on a GDP strip. By horizon of 50 years, most of this convergence in risk premia has taken place. In our sample, we estimate an output risk premium (relative to the long bond) rp_0^Y of 4.21%.

Claims to GDP are like unlevered equity claims. They have risk premia that are well in excess of real bond risk premia but below levered equity risk premia. At long horizons, dividend strip and GDP strip risk premia converge again by virtue of their cointegration.

We generate these discount rates while maintaining an excellent fit for the term structure of Treasury yields. The claim to surpluses reflects the risk of the government's future debt issuance strategy.

7 Potential Resolutions to the Valuation Puzzle

7.1 Austerity as a Peso Event

To interpret the magnitude of the valuation puzzle, we consider a simple variant of our model in which bond investors price in the possibility of a major, permanent government spending cut.[17] Such radical austerity never occurs in our 74-year sample but bond investors may think it could. We ask how large the spending cut probability needs to be in order to equate the market value of government debt to the present value of surpluses.

More precisely, we consider a permanent spending cut that lowers current and future government spending by 40%.[18] For a typical year with an average spending-to-GDP ratio of 17.55%, the cut lowers it to 10.53%. Given that the spending cut is permanent once it occurs, the long-run mean of spending-to-GDP, g_0, also falls by $\ell = 40\%$. We assume that the probability itself is i.i.d. over time; if the peso event did not happen this period, next period there is a new independent draw of the fiscal correction event. The dynamics of the demeaned state variables, including the demeaned log spending-to-GDP ratio, are still given by the benchmark VAR after the spending cut. As a result, the price of the spending claim is simply scaled by a factor of $1 - \ell = 60\%$ when the peso event happens. We assume that the peso event does not change the market prices of risk $\mathbf{\Lambda}_t$. We denote the probability of the spending cut that closes the valuation gap by ϕ_t, which satisfies:

$$\begin{aligned} D_t &= P_t^T - P_t^{G,peso} \\ P_t^{G,Peso} &= (1-\phi_t)\mathbb{E}_t[M_{t+1}^{\$}(G_{t+1} + P_{t+1}^{G,Peso})] + \phi_t(1-\ell)P_t^G \end{aligned}$$

where $P_t^{G,peso}$ is the value of the spending claim when the austerity event is possible. The blue line in Figure 11 reports the resulting time series of ϕ_t (Baseline Peso). To match the average 359% gap-to-GDP ratio, the annual probability of the spending cut has to be 8.49% on average. Using the time series of ϕ_t, we compute a cumulative probability of not seeing any peso event in our

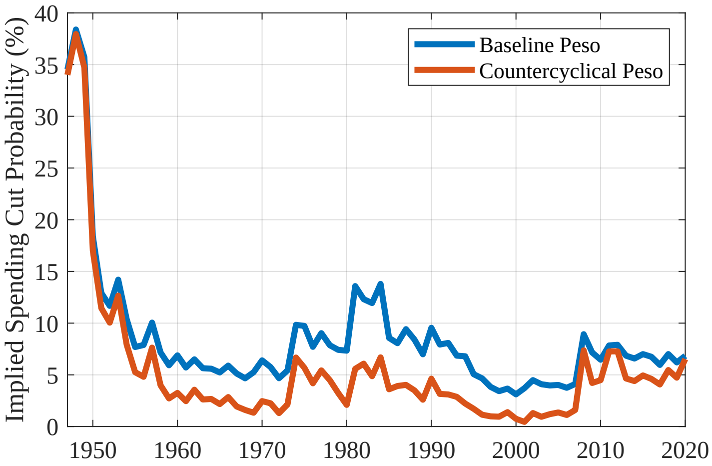
---

[17] The possibility of a large future increase in tax revenues is an alternative way to engineer a fiscal correction. We have confirmed that the results are similar.

[18] For context, U.S. defense spending accounted for 16% of the federal budget, Social Security for 23%, and Medicare for 15% in 2019.

74-year sample to be 0.11%. We interpret this result as a restatement rather than a resolution of the puzzle.

Figure 11: Probabilities of Spending Cut Implied by Debt-to-GDP Ratio



This figure reports the time series of probabilities of spending cuts implied by the debt-to-GDP ratio, ϕ_t. The baseline peso case is obtained from a constant probability of spending cut, whereas the countercyclical peso case is obtained from a time-varying probability of spending cut which increases when the GDP growth declines.

The above calculation assumes there is no correlation between SDF innovations and the occurrence of a fiscal correction. If instead the fiscal correction is more likely to take place in high marginal utility states, the implied probability of these fiscal corrections would likely be smaller.[19] To consider one such case, we assume alternatively that the spending cut is more likely to happen when GDP growth declines:

$$\mathbb{P}(Peso_{t+k}|\varepsilon_{t+k}) \overset{\text{def}}{=} \tilde{\phi}_{t+k} = 1 - (1 - \phi_t)\exp(-\frac{1}{2}\nu^2 + \nu\varepsilon^x_{t+k}), \qquad \forall k \geq 1,$$

where $\varepsilon^x_{t+1} = e'_x\varepsilon_{t+1}$ represents the orthogonalized GDP shock in our VAR system. We assume that $\nu = 2\%$, so that when the GDP growth is 2-standard-deviation lower (which is 3.67%), the probability of no spending cut decreases by 4%. In this way, while the ex-ante (physical) probability of the peso event is still $\mathbb{P}_t(peso_{t+1}) = \phi_t$, the response of the peso shock probability to the GDP shock affects how it is priced by investors. We report the derivations in Appendix D.5. The red line Figure 11 reports the resulting time series of ϕ_t (Countercyclical Peso) that closes the valuation gap in each period. The average probability of the spending cut is now 5.36%, naturally

---

[19] As an example, Elenev, Landvoigt, Shultz, and Van Nieuwerburgh (2021) allow for the possibility of a regime switch in fiscal policy once the debt-to-GDP crosses a boundary. Such a future regime change is shown to make the tax claim less risky at medium horizons.

lower than in the case where the peso event is uncorrelated with the state of the economy.[20] However, the cumulative probability of not seeing any peso event in our 74-year sample is still only 1.34%, leaving our conclusion on the implausibility of this resolution to the puzzle intact. Finally, Nakamura, Steinsson, Barro, and Ursúa (2013) find that real government bond returns suffer in disaster states. They are -3% per annum on average and, in 25% of the disaster cases, the realized real bond returns are as low as those on stocks. This suggests that disaster states are not typically accompanied by large fiscal adjustments (that represent good news for government bondholders).

7.2 Bubbles and Limits to Arbitrage

The valuation gap can result from a violation of the transversality condition (TVC), consistent with the presence of a rational bubble in government debt (Collin-Dufresne et al., 2023).[21] The TVC is violated if the value of debt in the far future does not converge to zero:

$$\lim_{T\to\infty} \mathbb{E}_t \left[M_{t,t+T} \frac{D_{t+T}}{Y_{t+T}} Y_{t+T} \right] \neq 0.$$

In standard infinite-horizon asset pricing models with realistic amounts of aggregate risk, i.e., a high enough GDP risk premium relative to the difference between the risk-free rate and the GDP growth rate, the TVC is satisfied (Jiang et al., 2020; Barro, 2020).

There is a large literature on rational bubbles in asset markets (Samuelson, 1958; Diamond, 1965; Blanchard and Watson, 1982; Woodford, 1990; Aiyagari and McGrattan, 1998, for seminal contributions). Recent work by Dumas, Ehling, and Yang (2021); Abel and Panageas (2022) creates a bubble component in government debt in overlapping-generations models. Most of this work ignores aggregate risk in output growth.

Bubbles on government debt can also arise in infinite-horizon models, where government debt plays a special role as a risk-sharing device (Bassetto and Cui, 2018; Chien and Wen, 2019; Angeletos et al., 2023; Brunnermeier et al., 2024; Reis, 2021). In response to our bond valuation puzzle, Brunnermeier et al. (2024) develop a model with long-lived investors where government debt allows households to partially insure counter-cyclical idiosyncratic production risk, subject to a risk retention requirement. Their model generates a negative PDV of surpluses, consistent with

---

[20] Intuitively, making the spending claim subject to a severe cut that occurs in bad times increases the riskiness of the spending claim. That lowers its value, compared to the case without spending cut (and to the case with an independent spending cut) and closes the valuation gap (for a smaller peso probability). Appendix Figure D.1 report the term structure of risk premium for the spending claim in the presence of the peso probability, confirming this intuition. The counter-cyclical peso event increases the annual risk premium on the spending claim by about 4% points, which is a very large effect relative to the risk premium on either the baseline spending claim or the tax claim. This shows that a value of $\nu = 0.02$ is substantial.

[21] This resolution is premised on the assumption that the asset pricing (SDF) model is the correct one. It may not be. Treasuries may be mispriced relative to other securities, something for which there is ample evidence as noted in the introduction. Mispricing is our preferred resolution of the valuation gap.

our analysis. However, debt also contains a bubble component that reflects the service flow from agents' ability to retrade the debt in response to idiosyncratic shocks. The value of this service flow is counter-cyclical and may be large enough on average to account for the valuation gap. This approach raises two new questions. First, why is the post-WW-II U.S. government able to mine the bubble while most other governments, including the U.S. pre-WW-I, are not? Second, how large is the implied narrow convenience yield and does it result in risk premia on a broad set of risky assets remaining positive?

7.3 Missing Government Assets

Another potential resolution of the government bond valuation puzzle is missing U.S. government assets. Bansal, Croce, Kiao, and Rosen (2019) study reallocation of resources away from private-sector R&D investment towards government investment in times of high uncertainty. In that interpretation, the wedge between government debt and the present value of surpluses (PVS) would be a measure of the market value of missing government assets. Several observations are in order. First, to the extent that the government assets in question are providing positive cash flows which contribute to the primary surplus, the PVS already accounts for these assets. This is the case for federal financial assets, such as student loans, which represent 7.4% of GDP in 2022 according to the Congressional Budget Office (CBO). Second, if they are not providing any positive cash flow, they are only valuable to the extent that they could provide a positive cash flow in the future. For example, a federal highway may currently be operating at a loss, but may be worth a positive amount if it were privatized. Third, we should only count federal government assets since we are studying federal government debt. Many assets like schools, utility services, public transportation assets are owned by local governments. Fourth, many federal government assets are not for sale for geopolitical reasons. This includes military assets but also critical infrastructure assets. Fifth, we need to balance the missing government assets against missing government liabilities. If these obligations increase the deficit in the future, then this acts like a negative government asset that is not yet considered in our analysis. Most notably, the cash flows from Social Security, Medicare, and Medicaid have been positive and declining over the course of our sample. They are about zero in the last year of our sample, 2020. The CBO calculates that in 2022, these programs added 2.1% of GDP to the federal budget deficit, and predicts that this deficit contribution will steadily grow over the next two decades. Given the size of these future obligations, our PVS is likely conservative.

7.4 Financial Repression

In the past, the U.S. Treasury and the Federal Reserve have engaged in coordinated efforts to lower the government's cost of funding, especially to help finance large wars. Examples include large-scale purchases of government bonds during and after WW-I and WW-II (Hall and Sargent, 2022; Acalin and Ball, 2022). Investors may be pricing in future episodes of financial repression. Recent rounds of large-scale asset purchases resemble these earlier episodes of financial repression.

8 Conclusion

Fiscally speaking, the U.S. is an outlier. The U.S. federal government has been running an average primary deficit for the past 90 years yet issues positive and growing amounts of debt. In a standard asset pricing framework, this constellation can only be sustained if the government provides insurance against aggregate risk to its bondholders, even when the risk-free rate is below the growth rate of the economy. In reality, the U.S. government tends to run larger deficits in recessions, times when bond investors face high marginal utility. The U.S. government must tap debt markets at inopportune times. Using a state-of-the-art dynamic asset pricing model, we quantify that the riskiness of surpluses dramatically lowers the U.S. government's fiscal backing. We conclude that the pricing of the portfolio of U.S. Treasury debt seems to violate the no-arbitrage restriction implied by the government budget constraint, a violation we call the U.S. government debt valuation puzzle.

This paper as well as follow-up work has begun to explore potential resolutions to this U.S.-specific puzzle, including convenience yields, the possibility of unprecedented future fiscal correction, and the presence of rational bubbles in government debt. Our current evidence points towards mispricing of the U.S. Treasury portfolio relative a large class of asset pricing models. It is well known in finance that assets can be subject to persistent mispricing when there are limits to arbitrage (Shleifer and Vishny, 1997). More work is needed to explore these and other potential explanations of our puzzle.

References

ABEL, A. B. AND S. PANAGEAS (2022): "Running Primary Deficits Forever in a Dynamically Efficient Economy: Feasibility and Optimality," Tech. rep., National Bureau of Economic Research.

ACALIN, J. AND L. BALL (2022): "Did the U.S. Really Grow out its WWII Debt?" .

AIYAGARI, S. R. AND E. R. MCGRATTAN (1998): "The optimum quantity of debt," *Journal of Monetary Economics*, 42, 447–469.

ALVAREZ, F. AND U. J. JERMANN (2005): "Using asset prices to measure the persistence of the marginal utility of wealth," *Econometrica*, 73, 1977–2016.

ANDREWS, S., R. COLACITO, M. M. M. CROCE, AND F. GAVAZZONI (2023): "Concealed carry," *Available at SSRN 3743205*.

ANG, A. AND M. PIAZZESI (2003): "A no-arbitrage vector autoregression of term structure dynamics with macroeconomic and latent variables," *Journal of Monetary Economics*, 50, 745–787.

ANGELETOS, G.-M., F. COLLARD, AND H. DELLAS (2023): "Public Debt as Private Liquidity: Optimal Policy," *Journal of Political Economy*, 131, 3233–3264.

BACKUS, D., N. BOYARCHENKO, AND M. CHERNOV (2018): "Term structures of asset prices and returns," *Journal of Financial Economics*, 129, 1–23.

BAELE, L., G. BEKAERT, AND K. INGHELBRECHT (2010): "The Determinants of Stock and Bond Return Comovements," *Review of Financial Studies*, 23, 2374–2428.

BAI, J. AND P. COLLIN-DUFRESNE (2019): "The CDS-bond basis," *Financial Management*, 48, 417–439.

BANSAL, R., M. M. CROCE, W. KIAO, AND S. ROSEN (2019): "Uncertainty-Induced Reallocations and Growth," CEPR Discussion Paper No. DP13964.

BARRO, R. J. (2020): "r Minus g," NBER Working Paper No. 28002.

BASSETTO, M. AND W. CUI (2018): "The fiscal theory of the price level in a world of low interest rates," *Journal of Economic Dynamics and Control*, 89, 5–22, fed St. Louis-JEDC-SCG-SNB-UniBern Conference, titled: "Fiscal and Monetary Policies".

BLANCHARD, O. (2019): "Public debt and low interest rates," *American Economic Review*, 109, 1197–1229.

BLANCHARD, O. J. AND M. W. WATSON (1982): "Bubbles, rational expectations and financial markets," *NBER Working Paper*.

BOHN, H. (1995): "The Sustainability of Budget Deficits in a Stochastic Economy," *J. Money Credit Bank.*, 27, 257–271.

——— (1998): "The behavior of US public debt and deficits," *Quarterly Journal of Economics*, 113, 949–963.

BOROVIČKA, J., L. P. HANSEN, AND J. A. SCHEINKMAN (2016): "Misspecified recovery," *Journal of Finance*, 71, 2493–2544.

BRUNNERMEIER, M., S. MERKEL, AND Y. SANNIKOV (2024): "Safe Assets," *Journal of Political Economy*, forthcoming.

CAMPBELL, J. Y. (1996): "Understanding risk and return," *Journal of Political Economy*, 104, 298–345.

CAMPBELL, J. Y., C. PFLUEGER, AND L. M. VICEIRA (2020): "Macroeconomic Drivers of Bond and Equity Risks," *Journal of Political Economy*, 128, 3148–3185.

CAMPBELL, J. Y. AND R. J. SHILLER (1988): "The dividend-price ratio and expectations of future dividends and discount factors," *The Review of Financial Studies*, 1, 195–228.

CAMPBELL, J. Y., A. SUNDERAM, AND L. M. VICEIRA (2017): "Inflation Bets or Deflation Hedges? The Changing Risks of Nominal Bonds," *Critical Finance Review*, 6, 263–301.

CBO (2021): "The 2021 Long-Term Budget Outlook," Accessed: 2023-5-22.

CHEN, Z., Z. JIANG, H. LUSTIG, S. VAN NIEUWERBURGH, AND M. Z. XIAOLAN (2022): "Exorbitant privilege gained and lost: Fiscal implications," Tech. rep., National Bureau of Economic Research.

CHERNOV, M., L. SCHMID, AND A. SCHNEIDER (2020): "A macrofinance view of US Sovereign CDS premiums," *Journal of Finance*, 75, 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.

COCHRANE, J. H. (2019): "The value of government debt," Tech. rep., National Bureau of Economic Research.

——— (2022): "The fiscal roots of inflation," *Review of Economic Dynamics*, 45, 22–40.

COCHRANE, J. H. AND J. SAA-REQUEJO (2000): "Beyond arbitrage: Good-deal asset price bounds in incomplete markets," *Journal of Political Economy*, 108, 79–119.

COLLIN-DUFRESNE, P., J. HUGONNIER, AND E. PERAZZI (2023): "Admissible Surplus Dynamics and the Government Debt Puzzle," *Swiss Finance Institute Research Paper*.

CROCE, M. M., T. T. NGUYEN, S. RAYMOND, AND L. SCHMID (2019): "Government debt and the returns to innovation," *Journal of Financial Economics*, 132, 205–225.

CRSP (2021): Center for Research in Security Prices, LLC. University of Chicago.

DAI, Q. AND K. J. SINGLETON (2000): "Specification analysis of affine term structure models," *Journal of Finance*, 55, 1943–1978.

DAVID, A. AND P. VERONESI (2013): "What Ties Return Volatilities to Price Valuations and Fundamentals?" *Journal of Political Economy*, 121, 682–746.

D'ERASMO, P., E. G. MENDOZA, AND J. ZHANG (2016): "What is a Sustainable Public Debt?" in *Handbook of Macroeconomics*, Elsevier, vol. 2, chap. 32, 2493–2597.

DIAMOND, P. A. (1965): "National debt in a neoclassical growth model," *American Economic Review*, 55, 1126–1150.

DU, W., B. M. HÉBERT, AND W. LI (2023): "Intermediary balance sheets and the treasury yield curve," *Journal of Financial Economics*, 150, 103722.

DU, W., J. IM, AND J. SCHREGER (2018): "The us treasury premium," *Journal of International Economics*, 112, 167–181.

DU, W., C. E. PFLUEGER, AND J. SCHREGER (2020): "Sovereign debt portfolios, bond risks, and the credibility of monetary policy," *Journal of Finance*, 75, 3097–3138.

DUFFEE, G. R. (2018): “Expected inflation and other determinants of treasury yields: Expected inflation and treasury yields,” *Journal of Finance*, 73, 2139–2180.

DUFFIE, D. AND R. KAN (1996): “A yield-factor model of interest rates,” *Mathematical Finance*, 6, 379–406.

DUMAS, B., P. EHLING, AND C. YANG (2021): “The debt capacity of a government,” Tech. rep., National Bureau of Economic Research.

ELENEV, V., T. LANDVOIGT, P. SHULTZ, AND S. VAN NIEUWERBURGH (2021): “Can Monetary Policy Creat Fiscal Capacity,” NBER Working Paper No. 29129.

FLECKENSTEIN, M., F. A. LONGSTAFF, AND H. LUSTIG (2014): “The TIPS-treasury bond puzzle,” *Journal of Finance*, 69, 2151–2197.

FRED (2021): Board of Governors of the Federal Reserve System. Data retrieved from FRED, Federal Reserve Bank of St. Louis.

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.

GABAIX, X. (2012): “Variable Rare Disasters: An Exactly Solved Framework for Ten Puzzles in Macro-Finance,” *The Quarterly Journal of Economics*, 127, 645–700.

GUPTA, A. AND S. VAN NIEUWERBURGH (2021): “Valuing private equity strip by strip,” *The Journal of Finance*, 76, 3255–3307.

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, 192–214.

——— (2022): “Three world wars: Fiscal–monetary consequences,” *Proceedings of the National Academy of Sciences*, 119, e2200349119.

HANSEN, L. P. (2012): “Dynamic valuation decomposition within stochastic economies,” *Econometrica*, 80, 911–967.

HANSEN, L. P., W. ROBERDS, AND T. J. SARGENT (1991): “Time series implications of present value budget balance and of martingale models of consumption and taxes,” *Rational Expectations Econometrics*, 121–61.

HANSEN, L. P. AND J. A. SCHEINKMAN (2009): “Long-term risk: An operator approach,” *Econometrica*, 77, 177–234.

JERMANN, U. J. (2020): “Negative swap spreads and limited arbitrage,” *The Review of Financial Studies*, 33, 212–238.

JIANG, Z. (2021): “US Fiscal cycle and the dollar,” *Journal of Monetary Economics*, 124, 91–106.

——— (2022): “Fiscal cyclicality and currency risk premia,” *The Review of Financial Studies*, 35, 1527–1552.

JIANG, Z., A. KRISHNAMURTHY, AND H. LUSTIG (2021a): “Foreign Safe Asset Demand and the Dollar Exchange Rate,” *Journal of Finance*, 76, 1049–1089.

JIANG, Z., H. LUSTIG, S. VAN NIEUWERBURGH, AND M. Z. XIAOLAN (2020): "Manufacturing Risk-free Government Debt," NBER Working Paper No. 27786.

——— (2021b): "Quantifying U.S. Treasury Investor Optimism," .

——— (2024): "What Drives Variation in the U.S. Debt/Output Ratio? The Dogs that Didn't Bark," *Journal of Finance*, forthcoming.

KOIJEN, R. S. J. AND M. YOGO (2024): "Exchange Rates and Asset Prices in a Global Demand System," SSRN Working Paper No. 3383677.

KRISHNAMURTHY, A. AND A. VISSING-JORGENSEN (2012): "The aggregate demand for treasury debt," *Journal of Political Economy*, 120, 233–267.

LEROY, S. F. AND R. D. PORTER (1981): "The present-value relation: Tests based on implied variance bounds," *Econometrica*, 555–574.

LETTAU, M. AND J. A. WACHTER (2007): "Why Is Long-Horizon Equity Less Risky? A Duration-Based Explanation of the Value Premium," *Journal of Finance*, 62, 55–92.

LIU, Y., L. SCHMID, AND A. YARON (2021): "The risks of safe assets," SSRN Working Paper No. 3699618.

LONGSTAFF, F. A. (2004): "The Flight-to-Liquidity premium in US treasury bond prices," *Journal of Business*, 77.

LUSTIG, H., S. VAN NIEUWERBURGH, AND A. VERDELHAN (2013): "The Wealth-Consumption Ratio," *Review of Asset Pricing Studies*, 3(1), 38–94., review of Asset Pricing Studies.

MEHROTRA, N. R. AND D. SERGEYEV (2021): "Debt sustainability in a low interest rate world," *Journal of Monetary Economics*, 124, S1–S18, the Real Interest Rate and the MarginalProduct of Capital in the XXIst CenturyOctober 15-16, 2020.

MIAN, A., L. STRAUB, AND A. SUFI (2021): "A Goldilocks Theory of Fiscal Policy," NBER Working Paper No. 29351.

NAGEL, S. (2016): "The liquidity premium of near-money assets," *Quarterly Journal of Economics*, 131, 1927–1971.

NAKAMURA, E., J. STEINSSON, R. BARRO, AND J. URSÚA (2013): "Crises and Recoveries in an Empirical Model of Consumption Disasters," *American Economic Journal: Macroeconomics*, 5, 35–74.

NIPA (2021): "National Income and Product Accounts," U.S. Bureau of Economic Analysis. Accessed: 2021.

PANAGEAS, S. (2020): "The implications of heterogeneity and inequality for asset pricing," *Foundations and Trends in Finance*, 12, 199–275.

REIS, R. (2021): "The Constraint on Public Debt when $r < g$ but $g < m$," Working Paper London School of Economics.

RENNE, J.-P. AND K. PALLARA (2024): "Fiscal fatigue and sovereign credit spreads," SSRN Working Paper No. 3475425.

SAMUELSON, P. A. (1958): "An exact consumption-loan model of interest with or without the social contrivance of money," *Journal of Political Economy*, 66, 467–482.

SHILLER, R. J. (1981): "Do stock prices move too much to be justified by subsequent changes in dividends?" *American Economic Review*.

SHLEIFER, A. AND R. W. VISHNY (1997): "The limits of arbitrage," *Journal of Finance*, 52, 35–55.

VAN BINSBERGEN, J. (2021): "Duration-Based Stock Valuation: Reassessing Stock Market Performance and Volatility," SSRN Working Paper No. 3611428.

VAN BINSBERGEN, J., M. BRANDT, AND R. KOIJEN (2012): "On the timing and pricing of dividends," *American Economic Review*, 102, 1596–1618.

VAN BINSBERGEN, J., W. DIAMOND, AND M. GROTTERIA (2022): "Risk-free interest rates," *Journal of Financial Economics*, 143, 1–29.

VAN BINSBERGEN, J. H. AND R. S. KOIJEN (2017): "The term structure of returns: Facts and theory," *Journal of Financial Economics*, 124, 1–21.

WOODFORD, M. (1990): "Public debt as private liquidity," *The American Economic Review*, 80, 382–388.