Monetary-Fiscal Interactions in the United States

Paul Bouscasse — Sciences Po
Seungki Hong* — Purdue University

First version: June 2023
This version: January 2026

Abstract

How does the fiscal side of the US government respond to monetary policy, and does it matter? We estimate the response of fiscal variables to monetary shocks and the counterfactual response of macroeconomic aggregates under different fiscal rules. Following an interest rate hike, the fiscal authority does not react: spending and transfers remain unchanged, tax receipts fall along with output, and interest payments and debt increase. Monetary policy would be more contractionary if fiscal policy were to stabilize debt through spending or taxes, but less contractionary if it used transfers. Indeed, transfer hikes reduce real debt by raising inflation.

Keywords: fiscal policy, monetary policy
JEL codes: E52, E63

*Paul Bouscasse: Department of Economics, Sciences Po (email: paul.bouscasse@sciencespo.fr) and CEPR. Seungki Hong: Department of Economics, The Mitch E. Daniels, Jr. School of Business, Purdue University (hong397@purdue.edu). We thank Adrien Auclert, Nicolas Caramp, Joshua Chan, Jeanne Commault, Miguel Faria-e-Castro, Axelle Ferriere, Nils Gornemann, Jennifer La'O, Gaston Navarro, Morten Ravn, Jón Steinsson, and Michael Woodford, as well as seminar and conference participants at various institutions for comments and suggestions. Some of the results of this paper were previously circulated under the title "Fiscal Adjustment to Monetary Shocks," by Paul Bouscasse (2021). Paul Bouscasse thanks the Janeway Institute at the University of Cambridge for financial support. Paul Bouscasse and Seungki Hong thank the Center for International Price Research, headed by Mario Crucini at Purdue University, for financial support.

1 Introduction

In many macroeconomic models, the effect of monetary policy sharply depends on how the fiscal side of the government reacts. For instance, in heterogeneous-agent new Keynesian (HANK) models, an interest rate hike increases payments on public debt, thus deteriorating the budget balance. Whether the fiscal authority clears its budget constraint by changing income taxes, transfers, spending, or issuing more debt shapes the response of output because it shifts the burden of adjustment to different households (Kaplan et al., 2018, Alves et al., 2020). However, there is little empirical evidence on how Congress responds to the decisions of the Federal Open Market Committee (FOMC) and whether that matters.

Our first contribution is to estimate the response of several fiscal variables to various monetary policy shock series. The first series is constructed in the spirit of Romer and Romer (2004). These shocks are interest rate changes purged from forecasts of output, inflation, and unemployment prepared by the staff of the Federal Reserve System. Since the FOMC might react to news about future fiscal policy, we also purge rate changes from forecasts of government receipts, expenditures, and surpluses. In addition, we rely on the series of Aruoba and Drechsel (2024) and Bauer and Swanson (2023b). Aruoba and Drechsel (2024) use some of the Federal Reserve's internal documents to construct sentiment indicators and purge interest rate changes from those indicators, along with a large number of forecasts of endogenous variables. Bauer and Swanson (2023b) use high-frequency changes in asset prices around monetary policy announcements to obtain monetary shocks. They then purge those shocks from recent news about the state of the economy to avoid contamination by the "response to news effect" (Bauer and Swanson, 2023a). Armed with these shocks, we estimate their effect on tax receipts, spending, transfers, interest payments, and real debt at the federal level. Our treatment of the data preserves the budget constraint of the government, so our results can be transposed into a theoretical model.

We find that, following an exogenous monetary policy tightening, receipts decrease, spending and transfers are constant, and interest payments and debt increase—all of these variables being expressed in real terms. For a 100 basis point increase in the nominal interest rate, tax receipts fall by 0.2–0.5% of trend GDP within one or two years and bounce back afterwards. Using a database on legislated tax changes (Romer and Romer, 2010), we show that this response is not driven by legislated changes in the tax schedule, but by the endogenous reaction of tax receipts to the fall in output. Perhaps surprisingly, government transfer payments do not exhibit a strong response. The explanation is simple: most transfers, such as Social Security and Medicare, are not automatic stabilizers. Unemployment insurance is, but it only accounts for a small share of transfers paid by the federal government. With receipts falling, roughly constant expenditures, and increased interest payments, the budget balance deteriorates and feeds an increase in federal debt.

Our second contribution is to estimate the response of the economy to a Romer–Romer-style monetary contraction under counterfactual rules for fiscal policy. We employ the semi-structural method of McKay and Wolf (2023) to construct these counterfactual scenarios: we make some structural assumptions, but we do not rely on a specific model. Their method requires the estimation of the response of the relevant variables to several shock paths for each fiscal instrument. We can then construct the counterfactual scenario by taking the weighted average of these impulse response functions that best enforces the counterfactual rule. This methodology is immune to the Lucas (1976) critique.

We find that, if fiscal policy stabilizes real debt by cutting spending or increasing taxes, the monetary contraction pushes the economy into a deeper recession. An increase in transfers, on the other hand, can make the recession milder, even though it stabilizes real debt. This surprising result stems from a simple fact: in the data, transfer hikes are expansionary and inflationary enough that they *reduce*—or at least do not increase—real debt: transfers seem to pay for themselves. Thus, the transfer instrument allows the fiscal authority to fight the monetary contraction while stabilizing real debt at the cost of substantial inflation.

Self-funded transfer shocks are reminiscent of the work of Bianchi and Ilut (2017) and Bianchi et al. (2023). They estimate dynamic stochastic general equilibrium models and find that some transfer shocks are unfunded and accommodated by the central bank. Bianchi et al. (2023), in particular, find that unfunded shocks are prevalent and explain most of inflation since WWII. Complementing these papers, we provide reduced-form evidence of this mechanism.

Our findings shed light on the HANK literature. Our spending and tax counterfactual scenarios are qualitatively consistent with prominent HANK models (Kaplan et al., 2018, Auclert et al., 2020): these models predict that, after a monetary contraction, stabilizing debt by cutting spending or raising taxes is more contractionary than letting debt adjust. Our transfer results are in contrast with these models since they tend to predict that using transfers is more contractionary than using taxes or debt.

Related literature: Kaplan et al. (2018) lament that “there is no empirical evidence that reveals what type of fiscal adjustment is the most likely to occur in practice, following a monetary shock.” Still, some papers have touched this question *en passant*. Using recursively-identified vector autoregression (VAR) shocks, Cochrane (1999) finds “not a shred of statistical evidence that federal-funds shocks forecast surpluses.” Using a VAR with high-frequency shocks, Sterk and Tenreyro (2018) estimate a response of real debt that is roughly consistent with ours. Using a VAR with recursive identification, Caramp and Silva (2018) find that fiscal revenues over GDP fall after a monetary shock, government purchases are constant and transfers slightly increase.

Some papers focus on how the effect of fiscal shocks depends on monetary policy. Ascari et al. (2023) argue that the effect of announcements of future government spending depends on whether the economy is in a fiscally- or monetary-led regime: they find that defense news shocks were expansionary during the Great Inflation (1960–79) but contractionary during the Great Moderation (1984–2007), suggesting that the economy switched from a fiscally-led to a monetary-led regime. Wolf (2023) applies the McKay-Wolf methodology to estimate the sensitivity of the fiscal multiplier to different monetary policy rules.

In contemporaneous work, Breitenlechner et al. (2025) and Kurcz (2025) also study the fiscal response to a monetary shock and ask different counterfactual questions, focusing on scenarios in which fiscal policy stabilizes tax collection or transfers. Closely related to McKay and Wolf (2023), Barnichon and Mesters (2023) leverage the similar theoretical insights to evaluate macroeconomic policy in a semi-structural way.

2 Reality: Methodology

2.1 Monetary Shocks

To identify monetary shocks, we use three different series.

The first is a variation of the measure developed by Romer and Romer (2004). They purge intended federal funds rate changes of forecasts of output, inflation, and unemployment to remove the component of monetary policy that is endogenous to economic conditions. The forecasts they use, known as the Greenbook forecasts, are prepared before each Federal Open Market Committee (FOMC) meeting by the staff of the Federal Reserve. It is plausible, however, that the monetary side of the US government should systematically react to the stance of its fiscal side, above and beyond the latter's effect on output, inflation and unemployment. For instance, the FOMC may monetize fiscal deficits or tighten in the face of those deficits as a show of independence. To mitigate this concern, we add Greenbook forecasts for receipts, expenditures and surplus of the federal government to the list of controls. Thus, we estimate

$$\begin{aligned}\Delta i_m =& \alpha + \beta i_{m-1} + \sum_{q=-1}^{2} \gamma^q \Delta\tilde{y}_m^q + \sum_{q=-1}^{2} \zeta^q \left(\Delta\tilde{y}_m^q - \Delta\tilde{y}_{m-1}^q\right) \\ &+ \sum_{q=-1}^{2} \eta^q \tilde{\pi}_m^q + \sum_{q=-1}^{2} \theta^q \left(\tilde{\pi}_m^q - \tilde{\pi}_{m-1}^q\right) + \iota \tilde{u}_m^0\end{aligned}$$

$$
\begin{aligned}
&+ \sum_{q=-1}^{2} \kappa^q \Delta \tilde{\text{rec}}_m^q + \sum_{q=-1}^{2} \lambda^q \left(\Delta \tilde{\text{rec}}_m^q - \Delta \tilde{\text{rec}}_{m-1}^q \right) \\
&+ \sum_{q=-1}^{2} \mu^q \Delta \tilde{\text{exp}}_m^q + \sum_{q=-1}^{2} \nu^q \left(\Delta \tilde{\text{exp}}_m^q - \Delta \tilde{\text{exp}}_{m-1}^q \right) \\
&+ \sum_{q=-1}^{2} \sum_{j=0}^{1} \pi^{j,q} \tilde{\text{heb}}_m^{j,q} + \sum_{q=-1}^{2} \sum_{j=0}^{1} \rho^{j,q} \left(\tilde{\text{heb}}_m^{j,q} - \tilde{\text{heb}}_{m-1}^{j,q} \right) + \epsilon_m^{RR+},
\end{aligned} \tag{1}
$$

where i_m is the intended federal funds rate in month m, and $\Delta \tilde{y}_m^q$, $\tilde{\pi}_m^q$, $\tilde{u}_m^q$, $\Delta \tilde{\text{rec}}_m^q$, $\Delta \tilde{\text{exp}}_m^q$ and $\tilde{\text{heb}}_m^{j,q}$ are the forecasts for real output growth, inflation, unemployment, receipts growth, expenditures growth and the high-employment budget surplus as a share of output in the previous ($q = -1$), current ($q = 0$), and subsequent ($q = 1, 2$) quarters. The high-employment budget surplus ($\tilde{\text{heb}}_m^{j,q}$) is a forecast of the federal surplus conditional on low unemployment—an acyclical measure of the fiscal stance. In September 1983, the Greenbook forecast switched from forecasting the surplus conditional on the unemployment rate being 5–5.1% to its being 6%. So, we interact $\tilde{\text{heb}}_m^{j,q}$ with pre- and post-September 1983 dummies, corresponding to the two variables: $\tilde{\text{heb}}_m^{0,q}$ and $\tilde{\text{heb}}_m^{1,q}$. The residuals obtained after running this regression, $\hat{\epsilon}_m^{RR+}$, are our measure of monetary shocks. We denote these shocks Romer–Romer+ since they are an extension of the initial Romer–Romer methodology.

Second, we use the shocks of Aruoba and Drechsel (2024). They extend the Romer–Romer methodology to include 132 forecast time series and 296 sentiment indicators. The forecasts used by Aruoba–Drechsel also originate from the Greenbook (renamed Tealbook in 2010) and include the original Romer–Romer variables as well as federal government spending. The sentiment indicators are based on a machine learning analysis of the various books prepared by the staff. Then, they estimate

$$
\Delta i_m = \alpha + \Gamma \left(i_{m-1}, X_m^{AD}, Z_m^{AD} \right) + \epsilon_m^{AD}, \tag{2}
$$

where $\Gamma(.)$ is a linear-quadratic function and X_t^{AD} and Z_t^{AD} are the vectors of forecasts and sentiment indicators. Since the right-hand side features 3,226 variables, equation (2) is estimated with a ridge regression. The estimated residual, $\hat{\epsilon}_m^{AD}$, is the monetary shock.

Third, we use the shocks of Bauer and Swanson (2023b). They build on a literature that uses asset price changes around FOMC announcements as a monetary shock. Market surprises, the reasoning goes, capture an exogenous component of monetary policy if the endogeneity of monetary policy is correctly anticipated by market participants. Bauer and Swanson (2023a) argue that these surprises are contaminated by a "response to news" effect: financial markets don't know how the Fed will respond to important macroeconomic news.

To solve this problem, Bauer and Swanson (2023b) propose regressing the monthly sum of the surprises on 6 control variables, which are available before the announcement and capture important economic news: non-farm payroll surprises, employment growth, implied skewness in the 10-year Treasury yield, and 3-month changes in the S&P 500, yield curve slope, and commodity price index. Thus, they run the regression:

$$\text{mps}_t = \alpha + \beta' X_{t-}^{BS} + \epsilon_t^{BS}. \tag{3}$$

Following Nakamura and Steinsson (2018a), mps_t is the first principal component of the changes in the first four quarterly Eurodollar futures contracts. X_{t-}^{BS} is the vector of control variables just described. Once again, $\hat{\epsilon}_t^{BS}$ is the monetary shock.

Each series of shocks, which we respectively denote Romer–Romer+, Aruoba–Drechsel, and Bauer–Swanson, has its pros and cons. The Aruoba–Drechsel and Bauer–Swanson shocks are less likely to be contaminated by endogeneity but they tend to exhibit lower statistical power at quarterly frequency, which is the frequency at which most fiscal data and shocks are available. Given this limitation, we begin with the Romer–Romer+ shocks in a quarterly vector autoregression (VAR) and then demonstrate the robustness of our findings using a mixed-frequency VAR (MF-VAR), which combines monthly Aruoba–Drechsel or Bauer–Swanson shocks, monthly macroeconomic variables, and quarterly fiscal variables. In the counterfactual exercises, we use the impulse response functions (IRFs) to the Romer–Romer+ shocks since most fiscal shocks are quarterly.

In figures A.1–A.2 of the appendix, we also present results for the shock series of Jarociński and Karadi (2020) and an extension of those of Miranda-Agrippino and Ricco (2021). We explain these shock series in appendix I.

2.2 Variables of Interest

We study the response of 5 fiscal variables (spending, tax receipts, transfers, interest payments, and debt, all in real terms) and 3 macroeconomic variables (GDP, inflation, and the nominal interest rate). The first three fiscal variables are fiscal instruments: they can be directly or indirectly controlled by the government. The other two endogenously depend on past values of debt, the interest rate, and inflation. The fiscal variables are at the federal level. State and local policy would also be interesting to study, but the narrative shocks that we use to construct the counterfactual have only been developed at the federal level. So, it is more realistic to focus on the latter.

All data series but debt and the nominal interest rate are from the National Income and Product Accounts (NIPA) tables. Our quarterly debt variable is net debt (total liabilities minus total financial assets) from the Financial Accounts of the United States (formerly known as the Flow of Funds Accounts). The interest rate is from the Federal Reserve Economic Data (FRED). We deflate the nominal series for spending, taxes, transfers, interest payments, debt, and GDP with the GDP deflator.[1]

To make the data stationary, we de-trend fiscal variables and GDP with the Gordon and Krenn (2010) procedure: (i) regress real GDP on a quadratic trend, (ii) divide real variables by this quadratic trend. Thus, fiscal variables are expressed in percentage of trend GDP. Another possible choice would be to cast the model in log-level. Compared to the logarithmic transformation, the Gordon-Krenn procedure has an interesting advantage: it preserves the budget constraint of the government.[2] We show in appendix C that the path of net debt from the Financial Accounts can be deduced approximately from the linearized budget constraint:

$$d_t \approx c_d d_{t-1} - c_\pi \pi_t + \underbrace{gs_t - tx_t + tr_t + int_t}_{\text{deficit}} + \text{deterministic terms}, \tag{4}$$

where d_t, gs_t tx_t, tr_t, and int_t are real debt, government spending, tax receipts, transfers and interest payments (all divided by trend GDP), and π_t is inflation. The coefficients c_d and c_π depend on the steady state values of inflation, the growth rate of trend GDP, and debt. Analytical expressions are given in appendix C; their 1947–2019 sample analogs are: $c_d \approx 0.98$ and $c_\pi \approx 0.58$. From equation (4), we can deduce a response of debt that is consistent with a well-defined budget constraint for the government. We report this construction in the body of the paper.[3] Similarly, we can obtain the response of the deficit without including it in the VAR, simply using the terms above the bracket in equation (4).

For mixed-frequency specifications, our list of variables requires some adjustments. First, we estimate monthly real GDP and GDP deflator series with the methodology proposed by Stock and Watson (2010) and followed by Jarociński and Karadi (2020). They use a Kalman filter to apportion quarterly GDP and GDP deflator data into monthly estimates thanks to a set of monthly indicators closely linked to economic activity and prices. Second, following Aruoba and Drechsel (2024) or Bauer and Swanson (2023b), we add the excess bond

[1]We provide an exhaustive list of our sources and definitions in appendix G.

[2]Ramey (2016) also argues in favor of this transformation, albeit on slightly different grounds: when computing a fiscal multiplier, estimates obtained with log-transformed data require a rescaling by the steady state spending to GDP ratio. Such rescaling is unnecessary with the Gordon-Krenn procedure, since it preserves relative levels.

[3]The values of c_d and c_π vary slightly across specifications because they are based on specification-specific sample averages. Since debt is one of the endogenous variables in our VAR, we could also report the response that is directly estimated. We report this other response and study the differences in appendix D. We also experiment with a constrained VAR, in which the two responses are the same.

premium of Gilchrist and Zakrajšek (2011) to the MF-VAR for its property as a predictor of the business cycle.[4] Third, we replace quarterly net debt from the Financial Accounts with monthly federal debt held by the public from US Treasury publications (Payne et al., 2025). Including a monthly fiscal variable in the MF-VAR is informative about monthly fluctuations in other fiscal variables, which improves the performance of the Kalman filter and state smoother that underlie the MF-VAR. The drawback of this debt concept, as we explain in appendix C, is that it does not exactly match the national accounts (figure C.1).[5] Fourth, we replace the 3-month Treasury bill rate with the 1-year Treasury bond rate. Since Gertler and Karadi (2015), the literature on high-frequency identification in VARs has traditionally preferred medium-term interest rates to capture the forward guidance component of monetary policy.

Our VAR sample for the Romer–Romer+ shocks spans 1947–2007. The start date ensures consistency with the sample for the narrative fiscal shocks used in the counterfactual scenarios; the endpoint reflects the onset of the zero lower bound. The samples for the Aruoba–Drechsel and Bauer–Swanson shocks begin in 1959 (when monthly GDP and the GDP deflator become available) and end in 2007 (prior to the zero lower bound) and 2019 (prior to the COVID-19 recession), respectively. The Romer–Romer+ shock series begins in the second quarter of 1970, the Aruoba–Drechsel series in October 1982, and the Bauer–Swanson series in February 1988. As discussed in the next section, we treat the Romer–Romer+ and Aruoba–Drechsel series as endogenous variables in the VAR, so we set them to zero before they become available. By contrast, we use the Bauer–Swanson series as an external instrument and estimate the mapping from reduced-form to structural shocks over the corresponding subsample (February 1988–December 2019).

2.3 Specifications

Our first specification is a quarterly vector autoregression (VAR) with 8 endogenous variables: spending, taxes, transfers, interest payments, debt, GDP, inflation, and the 3-month T-bill rate. As we explain below, some specifications will feature a monetary or fiscal shock among the endogenous variables, a so-called internal instrumental variable. We also include

[4] See also Gertler and Karadi (2015) on the inclusion of the excess bond premium in a VAR with monetary shocks.

[5] For the main results (figures 1–3), we construct impulse responses of debt that are consistent across specifications by iterating equation (4). We compare the responses of the various debt concepts in figure 7. We show results for Aruoba–Drechsel and Bauer–Swanson where we use quarterly debt from the Financial Accounts instead of monthly debt in figures A.6–A.7. The results are very similar, differing mostly in the smoothness of the IRFs. Figures A.6–A.7, however, should be taken with caution because the convergence of our estimation algorithm is poor (section E.6).

a quadratic time trend as a control. Formally, the reduced-form VAR can be written as

$$y_{t+1} = \sum_{l=1}^{L} B(l)' y_{t+1-l} + B_c' c_{t+1} + u_{t+1}, \quad u_{t+1} \sim \mathcal{N}(0, \Sigma), \tag{5}$$

where y_{t+1} is the vector of endogenous variables at time $t+1$, c_{t+1} the exogenous controls, and u_{t+1} the error term. The main competitors of VARs are local projections, which were proposed by Jordà (2005). Recent contributions have shown that both identify the same IRF in population (Plagborg-Møller and Wolf, 2021) and that VARs have a better bias-variance trade-off (Li et al., 2022). The latter advantage leads us to choose the VAR. We estimate this VAR with Bayesian techniques and a flat prior distribution for (B, Σ) (Jeffreys, 1946). We report the IRF at the mode of the posterior distribution of the parameters, which, given the flat prior distribution, coincides with the IRF that would be obtained by ordinary least squares or maximum likelihood estimation.[6] We refer to this IRF as the modal IRF.

Our second specification is a mixed-frequency vector autoregression (MF-VAR) that incorporates monthly shocks, monthly macroeconomic variables, and quarterly fiscal variables. The MF-VAR, an application of state space techniques, is a convenient way to handle the various frequencies at which the data is available (Giordani et al., 2012, Schorfheide and Song, 2015). The model and the state vector that contains the endogenous variables are defined at monthly frequency. Some elements of this state vector are only observed quarterly, but we can estimate their posterior distribution through Gibbs sampling.[7] An additional benefit of the Gibbs-sampling estimation of a state space model is that it can readily handle missing data: the excess bond premium is only available from 1973 and we can simply consider its corresponding element in the state vector as unobserved until then. We use a Minnesota prior distribution implemented with dummy observations to obtain smoother monthly IRFs (Del Negro and Schorfheide, 2012). Given the high dimension of the missing data, the mode of the parameter distribution is challenging to compute, so we simply report the point-wise median IRF.

We include the Romer–Romer+ or Aruoba–Drechsel shock series in our VAR, ordered first among contemporaneous variables. We recover the structural shock with a Cholesky decomposition. That is, we assume that the narrative shock, once we control for past values of endogenous variables, is exogenous to current and future macroeconomic conditions. This procedure, sometimes known as VAR with internal instrument, is recommended by Plagborg-Møller and Wolf (2021), who show that it is asymptotically equivalent to a local projection with the shock on the right-hand side. On the other hand, we follow Bauer and Swanson

[6]We give explicit formulas for the prior and posterior distributions of each specification in appendix E.

[7]We explain the algorithm and report convergence criteria in appendix E.6.

(2023b) and use the Bauer–Swanson shocks as an external instrument in a structural VAR (SVAR-IV): the instrument serves to identify the first column of the matrix that maps the structural shocks to the reduced form shocks (Stock and Watson, 2012, Mertens and Ravn, 2013).[8] The identification assumption is that the Bauer–Swanson monetary shock is orthogonal to the non-monetary structural shocks. Plagborg-Møller and Wolf show that this procedure is equivalent to a local projection under "(partial) invertibility, that is, the ability to recover the structural shock of interest [...] as a function of only current and past macro aggregates." Bauer and Swanson note that, although it is less conservative than the internal instrument, the SVAR-IV tends to produce tighter estimates.

We summarize our specification choices in table B.2.

3 Reality: Results

3.1 Main Results

We show the response of our variables to a Romer–Romer+ monetary shock in figure 1. We scale our IRFs such that the point estimate of the response of the nominal interest rate is 1 on impact.

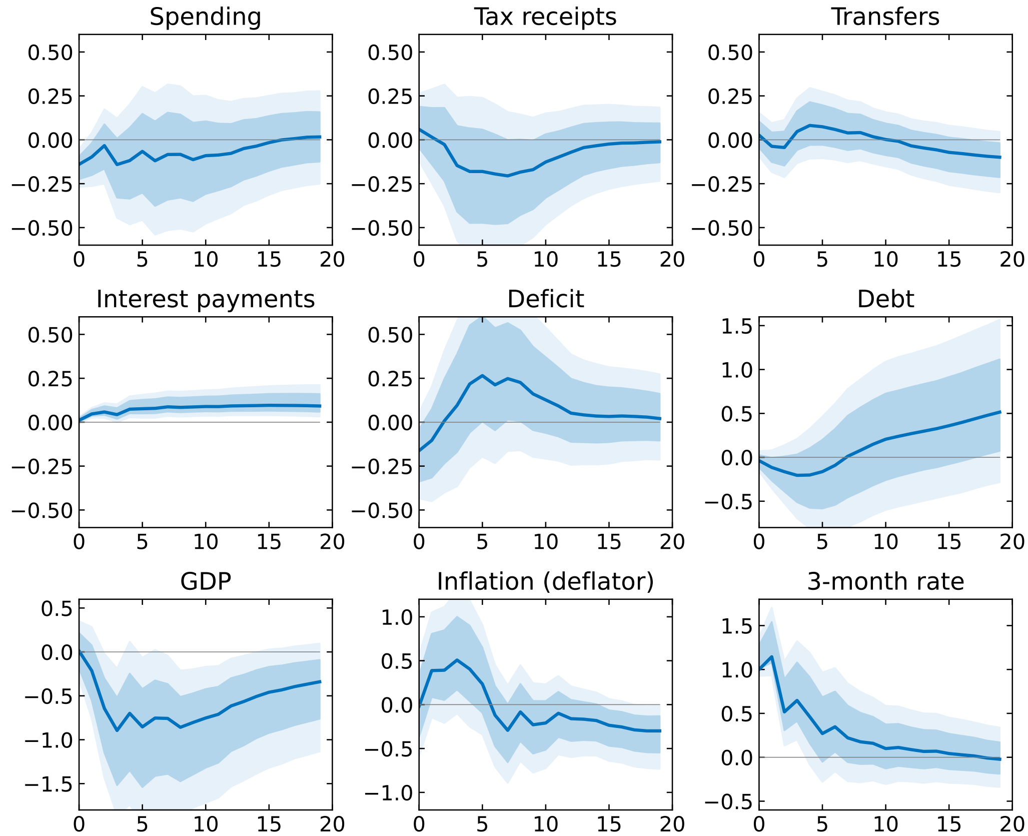
To give context, we first discuss the bottom row, which contains the well-known response of macroeconomic variables to a Romer–Romer monetary shock (Ramey, 2016, Nakamura and Steinsson, 2018b). The interest rate jumps on impact. It keeps increasing for one quarter and then slowly reverts towards 0. Real GDP falls in the year that follows the shock, stays low for another year, and starts recovering around the 10th quarter. Inflation falls after a year and stays persistently low for at least 4 years. The magnitude implies that for a 100-basis point rise in the 3-month nominal rate, GDP falls by about 0.75% (compared to trend) at the trough and inflation by 0.25 percentage points after a year.

The bottom row provides context, the top and middle ones are results. Spending and transfers are flat. Tax receipts fall slightly and interest payments increase. The combined effect of these changes is to increase the deficit. After a year, real debt builds up. This build-up in real debt is the result of an increase in the deficit and the fall in the price level entailed by the monetary shock. A 100-basis point rise in the 3-month nominal rate increases real debt by about 0.5 percentage points of trend GDP after 5 years.[9]

[8] We explain the SVAR-IV in section E.4.

[9] To be consistent with a theoretical model, the response of debt that we show here is the response derived from the linearized budget constraint, equation (C.3). We explain this construction in detail in appendix D and show that the implied response is qualitatively similar to the response of debt measured in the Financial Accounts (figure D.1). See also appendix C for a detailed explanation of the various debt concepts.

Figure 1: Response to Romer–Romer+ monetary shock

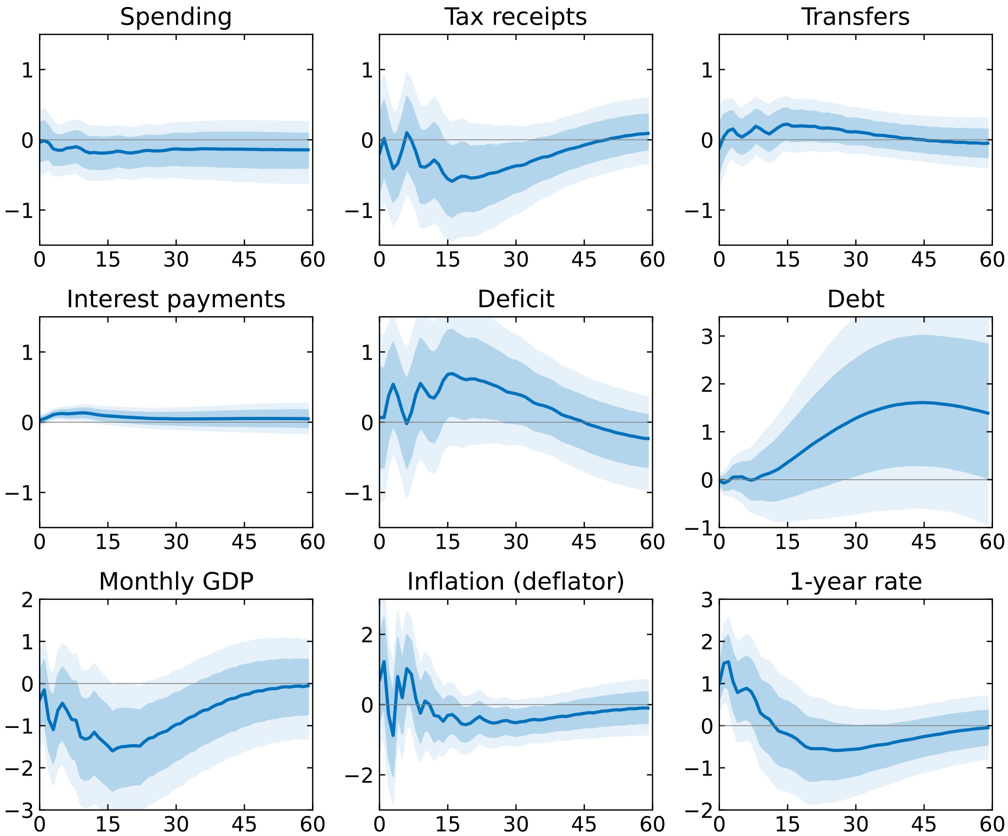


Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of the nominal interest rate is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Aruoba–Drechsel shocks deliver a similar message (figure 2). They cause a modest fall in tax receipts, which coincides with the fall in GDP, and a rise in interest payments. Spending and transfers seem mostly flat. These responses imply an increase in the deficit. Deflation kicks in after 2 years. The cumulated deficit and Fisher effect feed an increase in real debt. The similarity between the IRFs to Romer–Romer+ and Aruoba–Drechsel shocks shouldn't be a surprise: the two series rely on very similar identification strategies, although Aruoba and Drechsel use a more extensive set of controls.

Bauer–Swanson shocks imply consistent responses, but their IRFs are more front-loaded (figure 3). GDP falls vigorously and so do tax receipts. Inflation falls immediately and, after some turbulence, remains low for a year and a half. Spending barely responds. Transfers rise modestly, in line with the fall in GDP. The rise in interest payments is more modest

Figure 2: Response to Aruoba–Drechsel monetary shock



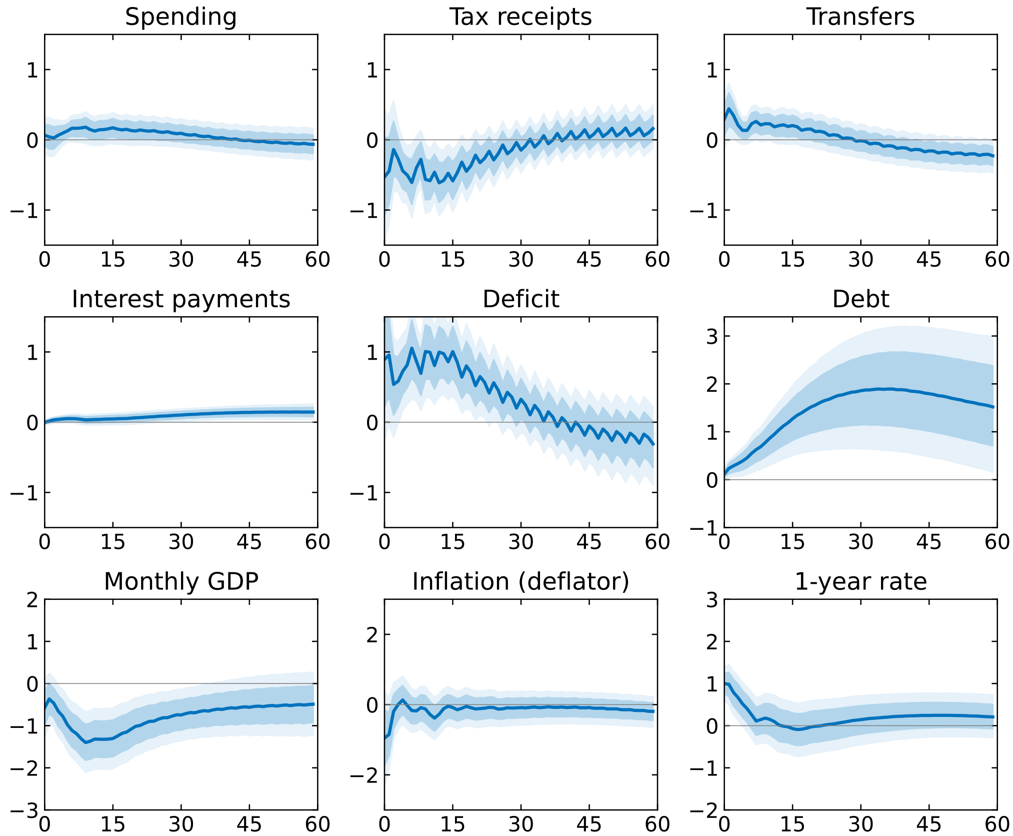
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of the nominal interest rate is 1 at time 0. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

than with Romer–Romer+ or Aruoba–Drechsel shocks: this behavior is consistent with the response of the nominal interest rate, which dissipates quickly after the shock.

We stress that the differences between the responses to the Romer–Romer+ and Aruoba–Drechsel shocks on the one hand, and the Bauer–Swanson shocks on the other hand, are not specific to our framework. Romer and Romer-style monetary shocks tend to imply a more sluggish response of output and prices than shocks based on high-frequency identification.[10] Similarly, the short-lived rise of the nominal interest rate after a Bauer–Swanson shock is in line with Bauer and Swanson's findings (figure 3). Thus, these IRFs suggest that fiscal vari-

[10] See also Miranda-Agrippino and Ricco (2021, figure 3). Miranda-Agrippino and Ricco's strategy to deal with the potential endogeneity of high-frequency shocks is different, but Bauer and Swanson (2023b, figure 7) show that both series of shocks yield similar IRFs.

Figure 3: Response to Bauer–Swanson monetary shock



Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of the nominal interest rate is 1 at time 0. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

ables behave mechanically given the response of macroeconomic variables—an interpretation that we develop in the next section.

3.2 Interpretation

Our interpretation of these results is that the fiscal side of the government is mostly unresponsive to monetary shocks. It leaves spending and transfers unchanged, lets tax receipts fall endogenously as a result of the fall in output and interest payments rise as a consequence of the increase in the interest rate. Debt must adjust to clear the budget constraint. Therefore, an interest rate hike leaves the federal government more indebted.

3.2.1 Output-Driven vs. Legislated Changes in Tax Receipts

Is the fall in tax receipts driven by legislated tax changes? This fall could be driven by: (i) the contraction in output and tax collection falling for a given tax schedule, (ii) an activist response of Congress in the face of monetary shocks, or (iii) chance correlation. Numbers (i) and (ii), though they highlight different mechanisms, would be valid causal effects of monetary policy. Number (iii) is worrisome in this context: the biggest Romer–Romer+ monetary policy shocks occurred in the early Volcker era (Coibion, 2012); at about the same time, Ronald Reagan was presiding over one of the largest tax cuts in US history. Luckily, an informal piece of evidence suggests that the response of receipts is mostly due to mechanism (i): on figure 1, the response of receipts follows that of output in the same direction. Moreover, the fact that the fall in receipts dissipates after a few years doesn't seem consistent with a change in the tax schedule, which one would expect to last longer.

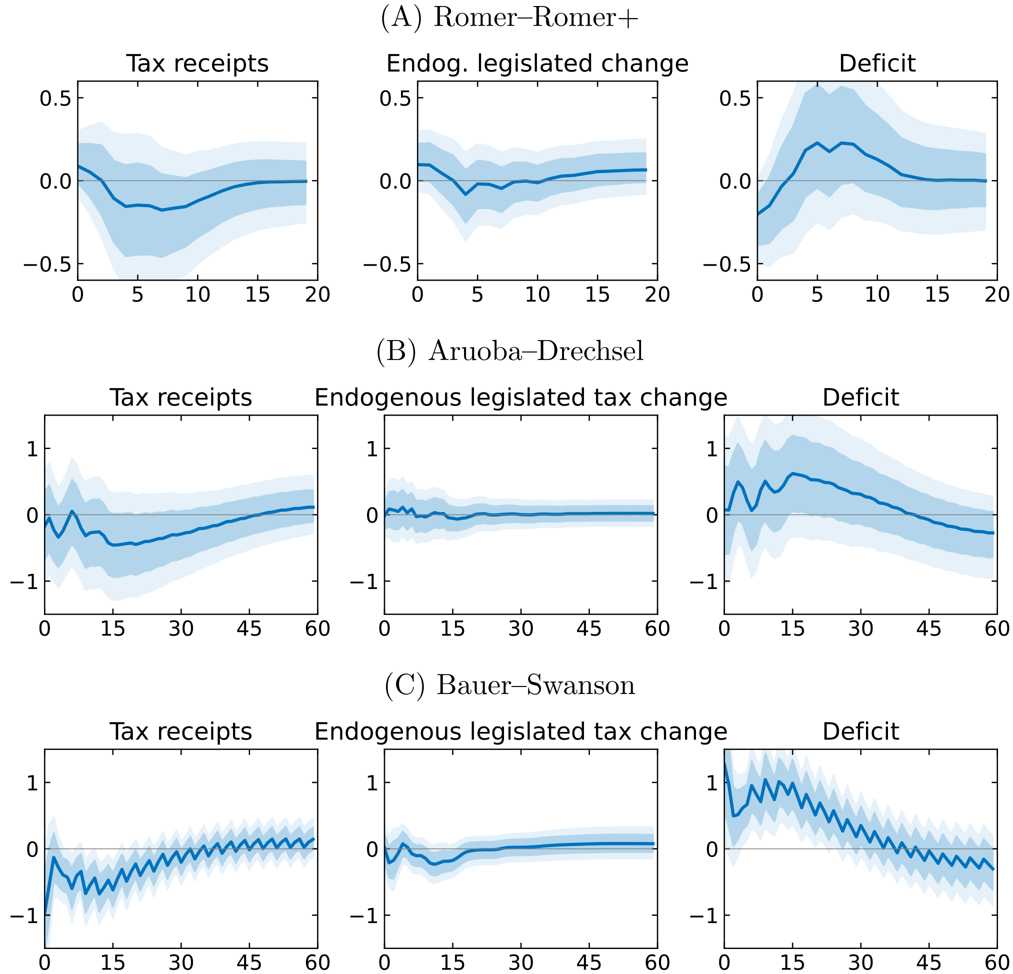
To investigate this question more formally, we use the database of legislated tax changes created by Romer and Romer (2010). They analyze the narrative record to quantify changes in the tax schedule, and classify them according to their underlying motivation. Thus, they distinguish four rationales that can drive a legislated tax change: finance extra spending, fight a recession, remedy an inherited deficit, and spur long-run growth. The first three categories may be endogenous to monetary policy.[11] The last category is exogenous to monetary policy, but it is a first order of concern for it includes the Reagan tax cuts of 1981.

Our strategy is to add the endogenous legislated changes to the VAR and include the exogenous legislated changes as controls. For the endogenous changes, we use the cumulative changes (expressed as a share of trend GDP). For the exogenous changes, we use the current value and a year of future values of the legislated change. The response of tax receipts and the deficit stay the same in response to the Romer–Romer+ shocks (figure 4, panel A). Since the Reagan tax cuts are part of the controls, this result addresses the concern that the response of tax receipts is due to chance correlation. As for the endogenous legislated changes, they do not decrease—they increase if anything. Therefore, the fall in tax receipts seems driven by their endogenous response to the fall in output, not by a legislated change in tax rates.

We repeat this exercise for the Aruoba–Drechsel and Bauer–Swanson shocks in panels B–C. We must omit the exogenous legislated tax changes as controls for two reasons: the database of legislated tax changes is quarterly while our VAR is monthly; the database stops

[11] Romer and Romer (2010) are interested in a different question: what are the effects of tax cuts on output? Hence their assessment of which tax changes are endogenous differs from ours. From their point of view, remedying an inherited deficit is exogenous since it is not driven by economic conditions. From our point of view, an inherited deficit can be the result of past monetary policy actions—the FOMC decides to generate less seigniorage for instance—, hence should be treated as potentially endogenous.

Figure 4: Interpretation: tax receipts



Note: Extensions of the results shown in figures 1–3. To the baseline VARs described in section 2, we add (i) endogenous legislated changes as an endogenous variable for the 3 shock series and (ii) exogenous legislated changes as controls for the Romer–Romer+ VAR. The solid line is the modal (Romer–Romer+)/median (Aruoba–Drechsel, Bauer–Swanson) IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.2.1 for more details.

in 2007 whereas the Bauer–Swanson sample ends in 2019. There is no simple way to handle sparsely observed control variables in a mixed-frequency VAR, so we do not include them.[12] The endogenous legislated changes are readily handled by our mixed-frequency VAR since they are part of the state space of the VAR model. In both cases, we see a similarly muted response of endogenous legislated tax changes, thus confirming that the fall in tax receipts is not driven by an endogenous response of marginal tax rates.

[12]The main rationale for including those exogenous changes as controls was the coincidence of the Reagan tax cuts (1981) and the Romer–Romer shocks of the early Volcker period (1979–83). This issue is less of a concern for the Aruoba–Drechsel and Bauer–Swanson shocks since the former are smaller than the Romer–Romer shocks in the early 1980s (Aruoba and Drechsel, 2024, figure 4) and the latter only start in 1988.

3.2.2 Unresponsive Transfers?

Is the lack of response of transfers plausible? It may seem surprising that transfers do not respond. Given the contraction in output, one would expect, for instance, a countercyclical response of unemployment benefits after the monetary contraction triggers an increase in unemployment. The answer to this puzzle is that most of the transfers provided by the federal government should not be expected to be countercyclical. In 2007, the three biggest categories, which accounted for 70% of federal transfers, were Social Security, Medicare and Medicaid.[13] There is little reason to expect that old-age, disability and health insurance payments go up in a recession. Programs such as unemployment insurance and food stamps are more likely to be countercyclical, but together they account for only 4% of the total. To check that our framework implies a plausible response of the countercyclical components of transfers, we add unemployment insurance and the unemployment rate to the VAR and plot the results in figure 5: consistent with the fall in GDP, both increase after a monetary shock, while total transfers remain unresponsive. This statement applies to the three shock series.

3.2.3 What Drives Deficit and Debt?

Is the response of real debt driven by the deficit or a Fisher effect? In figure 6, we plot the cumulative response of deficit and inflation, as well as the response of debt that is already plotted in figures 1–3, to the three shocks. The first two lines sum to the third one.[14] At least in the first 2 years that follow the shock, the cumulative deficit accounts for most of the response of debt in all three panels. The Fisher effect kicks in after 2–3 years, especially for the Romer–Romer+ and Aruoba–Drechsel shocks. In appendix figures A.3–A.5, we break this down even further by decomposing the cumulative response of deficit into its spending, receipts, transfers, and interest payments components. The rise in interest payments and the fall in tax receipts account for most of the deficit component of debt.

3.3 Debt: Additional Results

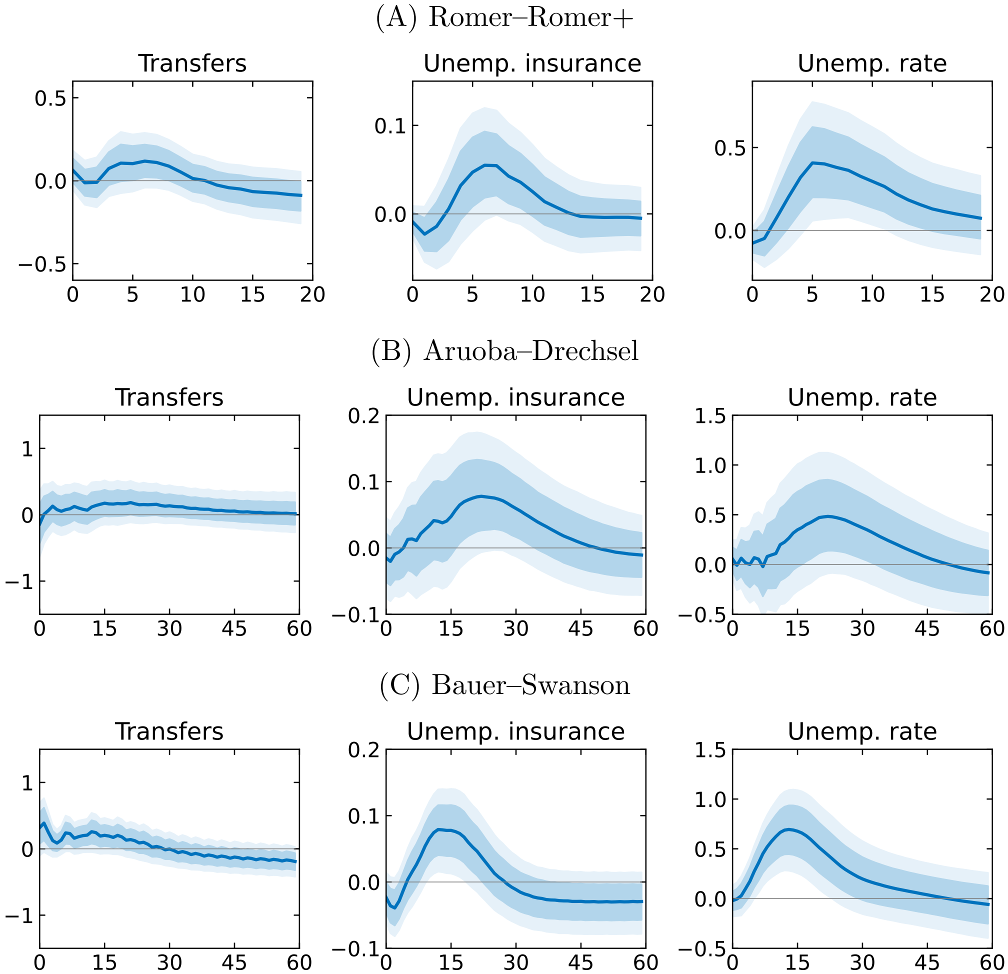
3.3.1 Concepts of Debt

We now turn our attention to other concepts of debt. As we explained in section 2.2, there are two potential concepts of federal debt: net debt from the national accounts and federal

[13] In appendix table B.1, we break federal transfers down by category.

[14] Compared to figures 1–3, those cumulative responses are divided by 4. Indeed, in keeping with the convention of the national accounts, our quarterly variables and inflation are annualized, so we need to re-scale them so that they're consistent with the response of debt.

Figure 5: Interpretation: transfers



Note: Extensions of the results shown in figures 1–3. To the baseline VARs described in section 2, we add unemployment insurance and the unemployment rate. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.2.2 for more details.

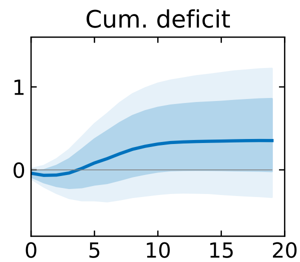
debt held by the public from the US Treasury.[15] In the results shown in figures 1–3, we constructed the debt response by iterating the linearized budget constraint (C.3). Thus, the response of debt that we show is consistent with national accounts. Another possibility is to study federal debt held by the public. The advantage is then that we can study the face and market values of that debt: by raising the interest rate, monetary policy should lower the price of government bonds.

To jointly study the response of face and market values of federal debt held by the public (FDHP) in our VARs, we include the lagged face value of FDHP and the current valuation term, which we construct by taking the difference between the market and face values of federal debt. We order the lagged face value first: in the notation of equation (5), the first

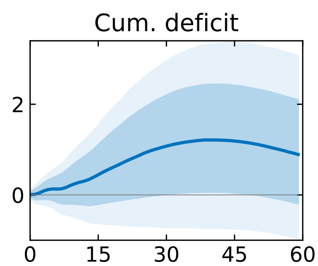
[15] We explain the differences between those sources in appendix C and compare them in figure C.1.

Figure 6: Interpretation: debt decomposition

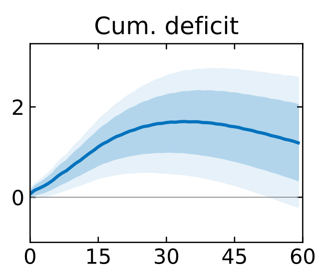
(A) Romer–Romer+



(B) Aruoba–Drechsel



(C) Bauer–Swanson



Note: Decomposition of the response of debt shown in figures 1–3 into its cumulated deficit and Fisher effect components. The solid line is the modal (Romer–Romer+)/median (Aruoba–Drechsel, Bauer–Swanson) IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.2.2 for more details.

component of y_t is d_{t-1}, followed by other endogenous variables at time t. Using the lagged (d_{t-1}), instead of current (d_t), face value of debt makes no conceptual difference when the monetary shock is an internal instrument (Romer–Romer+, Aruoba–Drechsel): d_t is allowed to respond to ϵ_t^{RR+} or ϵ_t^{AD} (d_{t-1} is not because it is ordered before ϵ_t^{RR+} or ϵ_t^{AD}.) On the other hand, when the monetary shock is an external instrument (Bauer–Swanson), d_t is only allowed to respond to the structural monetary shock *through other endogenous variables.*[16] We impose this restriction to avoid chance correlation between the Bauer–Swanson shock and debt issuance prompted by the financial crisis in the fall of 2008.[17] Since the face value

[16] Of course, d_{t-1} is not allowed to respond to the time-t structural monetary shock in any way; we explain implementation details in appendix E.4.

[17] We show results where we do not impose this restriction in figure A.11. A puzzling feature of this figure is the immediate jump in face and market values that follows a Bauer–Swanson shock. This jump is too

of debt is a stock variable that responds to endogenous variables in a mechanistic fashion, this restriction seems appropriate. The valuation term, of course, can freely respond to the monetary shock—we recover the response of the market value of debt by adding the responses of the face value and of the valuation term.

The results are shown in figure 7. The response of the face value is similar to the baseline (reproduced for convenience in the left column). Market value drops on impact, which is consistent with an increase in interest rates, but the difference with the face value dissipates after a year.

3.3.2 Long-Run Response of Debt

Is the debt increase permanent? Our IRFs seem to imply that debt is permanently higher after a monetary shock. We investigate this further in figure 8. The left plot extends the IRF of our original VAR (figure 1) to 10 years. It shows no sign of reversion. Plagborg-Møller and Wolf (2021) show that VARs and local projections are approximately asymptotically equivalent if the VAR includes at least as many lags as the horizon of the IRF. By extending our IRF over 10 years while keeping a year of lags, we are getting further from this theoretical benchmark. In the middle plot, we include 5 years of lags in the VAR. Since keeping as many variables in our VAR would imply more coefficients than observations, we keep only GDP, inflation, the interest rate, and quarterly net debt from the Financial Accounts. In the right-hand side plot, we estimate a local projection with an instrumental variable (Jordà, 2005, Jordà et al., 2015):

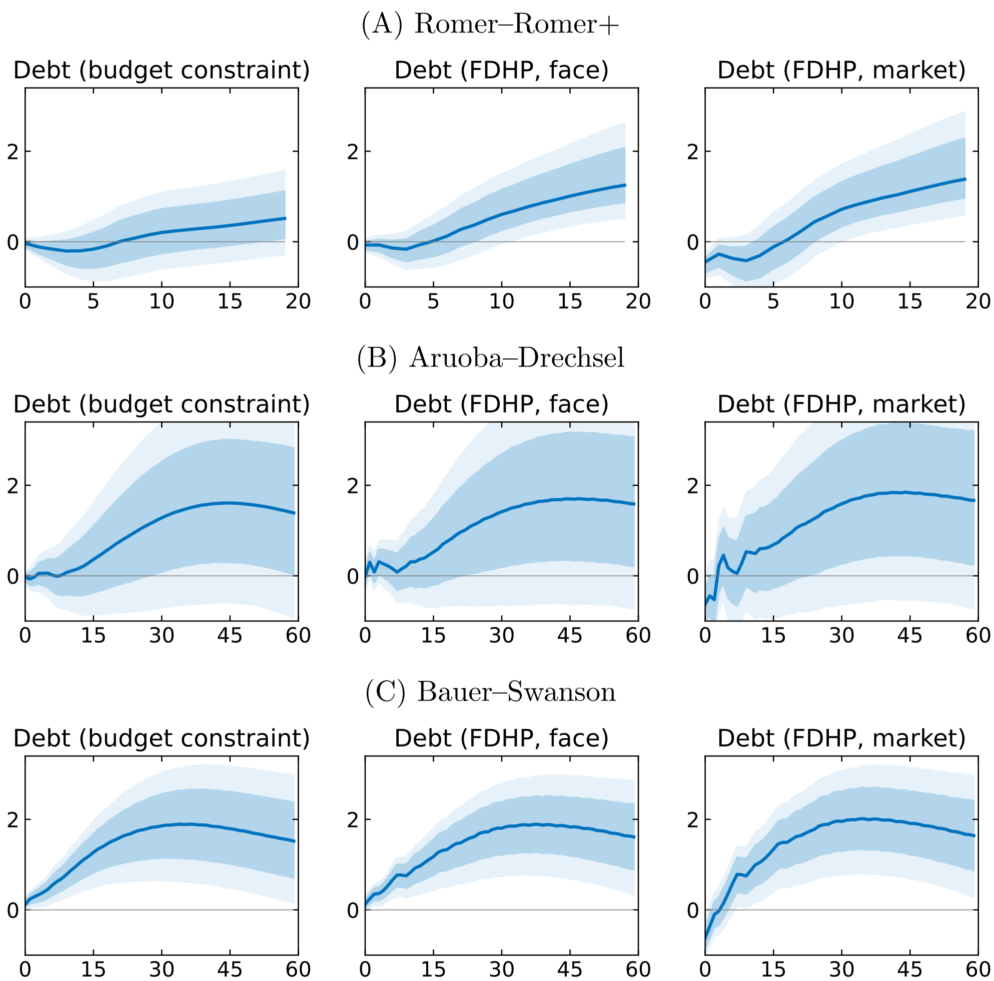
$$d_{t+h} = \alpha^h + \beta^h \Delta i_t^{3m} + \gamma^h d_{t-1} + \delta^{h'} X_{t-1} + e_t^h, \tag{6}$$

where Δi_t^{3m} (the change in the interest rate on 3-month Treasury bills) is instrumented with the Romer–Romer+ monetary shock and X_t is a vector of controls with a year of lags of debt, GDP, inflation, and the interest rate. The inclusion of at least a lag of the dependent variable (d_{t-1}) in the controls follows the recommendation of Montiel Olea and Plagborg-Møller (2021). The result is robust to our specification choice. So, we conclude that the increase in debt is permanent over an estimable horizon.[18]

large to be explained by the rise in the deficit or the fall in inflation (left-hand side column). Inspecting the results, we found that it was driven by chance correlation: according to Bauer and Swanson (2023b), there were contractionary monetary shocks in September and October 2008. These coincided with very large increases in the face value of federal debt (6% in September and 9% in October). These increases were driven by the Troubled Asset Relief Program (TARP), which was a response to the Global Financial Crisis.

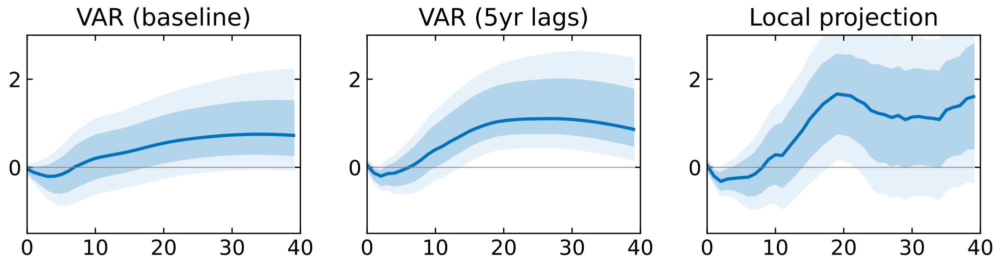
[18] In principle, we could extend our IRFs even further. We are reluctant to do so, however, since the sample of Romer–Romer+ shocks (1969–2007) is 39 years long hence contains less than 4 non-overlapping 10-year periods. With IRFs of 15 years, this number drops to 2; with 20 years or more, one cannot find 2 non-overlapping subsamples of the length of the IRF.

Figure 7: Debt concepts



Note: The left-hand side column is a reproduction of the baseline responses shown in figures 1–3. The middle and right-hand side columns contain the response of the face and market values of federal debt held by the public (FDHP) in a VAR that contains the lagged face value and a market valuation term (market value minus face value). The solid line is the modal (Romer–Romer+)/median (Aruoba–Drechsel, Bauer–Swanson) IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.3.1 for more details.

Figure 8: Long-Run Response of Debt to Romer–Romer+ shocks



Note: We study the response of debt in the long run. On the left, we extend the IRF of our baseline VAR to 10 years for Romer–Romer+ shocks. In the middle, we estimate a VAR with 5 years of lags and debt as the only fiscal variable. On the right, we estimate a local projection with heteroskedasticity autocorrelation consistent standard errors (Newey and West, 1987). The solid line is the modal IRF (VAR) or ordinary least squares estimate (local projection). Shaded areas represent 68% (dark) and 90% (light) credible (VAR) or confidence (local projection) intervals. See section 3.3.2 for more details.

We have also experimented with Aruoba–Drechsel and Bauer–Swanson shocks, but they deliver inconclusive results on the long-run response of debt. This is unsurprising since these series have shorter samples and smaller standard deviations. For instance, the Aruoba–Drechsel shock series lasts for less than 30 years, so they contain less than 3 non-overlapping 10-year periods. In the local projection, this translates into results that are very sensitive to small changes in the specification.

If debt is permanently higher, which fiscal variable adjusts to cover additional interest payments? We cannot answer this question because the adjustment is too small to detect. Indeed, consider the government budget constraint in steady state:

$$\frac{r^* - g^*}{1 + g^*} d^* = tx^* - gs^* - tr^*, \tag{7}$$

where r^* is the real interest rate, g^* the growth rate of trend GDP, d^* debt, tx^* taxes, gs^* government spending, and tr^* transfers (all expressed relative to trend GDP).[19] The left-hand side is the real service of debt adjusted for long-run GDP growth. The right-hand side is the primary surplus. In steady state, the government must run a surplus to finance a positive quantity of debt if $r^* > g^*$ and a deficit otherwise. A Romer–Romer+ shock that raises the interest rate by 100 basis points increases the debt-to-trend GDP by roughly 1 percentage point over the long run. The sample average for $(r_t - g_t)/(1 + g_t)$ is -0.6% in annual terms.[20] So, the required adjustment would be of -0.006 percentage points of trend

[19]See appendix C.4 for a derivation.

[20]For r_t we take the effective interest rate, which is interest payments divided by lagged debt and gross inflation. g_t is the growth rate of trend GDP described in section 2.2. It may seem surprising that $r_t - g_t$ is negative in our sample. Yet, this fact echoes the results of Blanchard (2019), who shows that the 1- and 10-year yields have been lower than nominal growth for most of postwar history, and Mehrotra and Sergeyev

GDP. This kind of precision is beyond the reach of time series econometrics, especially at long horizons. Even on impact, the width of our 68% credible interval for the deficit is larger than 0.2 percentage points of trend GDP (figure 1). Even if we assume $r^* = 4\%$ and $g^* = 0$, we obtain a required adjustment of 0.04 percentage points of trend GDP, a number that is five times smaller than the width of our 68% credible interval. Therefore, we must remain agnostic about the adjustment that takes place in the long run. Regardless, this adjustment might not even be quantitatively relevant: Auclert et al. (2020, figure 7) show that if the fiscal adjustment is delayed, the instrument used to balance the budget constraint (spending, taxes, or transfers) does not matter in their model; the instrument only matters if adjustment is immediate.

4 Counterfactual: Methodology

4.1 The McKay-Wolf (MKW) Method

In this section, we explain a methodology that was recently proposed by McKay and Wolf (2023). Readers who are familiar with it can jump to section 4.3.

McKay and Wolf's method answers a policy counterfactual question by means of time series regression. This method requires some structural assumptions but no commitment to a particular model. These assumptions are satisfied by the main models of the macroeconomics literature, such as the real business cycle model or the new Keynesian model, be it with representative or heterogeneous agents.

Consider a macroeconomic model that features non-policy variables, which can be observed or unobserved, and policy instruments. Those variables are linearized around their steady state value. The observed non-policy variables are collected in vector $\mathbf{x}$, the unobserved ones in $\mathbf{w}$, and the policy instruments in $\mathbf{z}$. The vectors $\mathbf{x}$, $\mathbf{w}$, and $\mathbf{z}$ feature the time path, from $t = 0$ to infinity, of each of the variables.[21] The model also features structural and policy shocks, collected in ε and ν, respectively.

The main theoretical assumption is that the equations that characterize the solution of the model can be separated into a non-policy block and a policy block. Formally, one must be able to write them in the following way:

$$\text{non-policy block:} \qquad \mathcal{H}_w\mathbf{w} + \mathcal{H}_\mathbf{x}\mathbf{x} + \mathcal{H}_\mathbf{z}\mathbf{z} + \mathcal{H}_\epsilon\varepsilon = \mathbf{0}, \quad \text{and} \tag{8}$$

(2021), who show that the median value of $r_t - g_t$ has been negative in the United States from 1870 or 1946 to 2016.

[21] In practice, an infinite time horizon is truncated at $t = T - 1$ for some finite T. Under such truncation, $\mathbf{x}$, $\mathbf{w}$, and $\mathbf{z}$ become $(n_\mathbf{x}T \times 1)$-, $(n_\mathbf{w}T \times 1)$-, and $(n_\mathbf{z}T \times 1)$-arrays, respectively, where $n_\mathbf{x}$, $n_\mathbf{w}$, and $n_\mathbf{z}$ are the number of observed non-policy variables, unobserved variables, and policy instruments, respectively.

$$\text{policy block:} \qquad \mathcal{A}_x \mathbf{x} + \mathcal{A}_{\mathbf{z}} \mathbf{z} + \nu = \mathbf{0}. \tag{9}$$

Crucially, the elements of the $\mathcal{H}$ matrices are not allowed to depend on the policy rules $\mathcal{A}$. The solution is assumed to be unique. We write it as

$$\begin{pmatrix} \mathbf{w} \\ \mathbf{x} \\ \mathbf{z} \end{pmatrix} = \underbrace{\begin{pmatrix} \Theta_{w,\varepsilon,\mathcal{A}} & \Theta_{w,\nu,\mathcal{A}} \\ \Theta_{x,\varepsilon,\mathcal{A}} & \Theta_{x,\nu,\mathcal{A}} \\ \Theta_{z,\varepsilon,\mathcal{A}} & \Theta_{z,\nu,\mathcal{A}} \end{pmatrix}}_{=\Theta} \times \begin{pmatrix} \varepsilon \\ \nu \end{pmatrix}. \tag{10}$$

Each row of Θ contains the IRF of the variables of the model to the shocks.

Most workhorse models of modern macroeconomics, once linearized, can be written in this way. Take, for instance, the representative agent New Keynesian (RANK) model (Galí, 2015, ch. 3), which we augment with government spending:

$$\begin{aligned}
\text{NK Phillips curve:} \quad & \pi_t = \beta E_t \pi_{t+1} + \kappa \hat{y}_t + \psi a_t, \\
\text{dynamic IS:} \quad & \hat{y}_t = -\frac{1}{\sigma} \left(\hat{\imath}_t - E_t \pi_{t+1} \right) + E_t \hat{y}_{t+1} + s_g \left(\hat{g}_t - E_t \hat{g}_{t+1} \right), \\
\text{Taylor rule:} \quad & \hat{\imath}_t = \phi_\pi \pi_t + \phi_y \hat{y}_t + v_t, \quad \text{and} \\
\text{government spending:} \quad & \hat{g}_t = \rho \hat{g}_{t-1} + w_t,
\end{aligned}$$

where π is inflation, $\hat{y}$ output, $\hat{g}$ government spending, $\hat{\imath}$ the nominal interest rate — the latter three expressed in deviation from steady state —, a, v and w are technology, monetary and fiscal shocks. Putting π and $\hat{y}$ in $\mathbf{x}$, $\hat{\imath}$ and $\hat{g}$ in $\mathbf{z}$, a in ε, and v and w in ν, one can write this model in the prescribed format. The rows of the matrices of equation (8), the non-policy block, would feature the Phillips curve and IS equations expressed from time $t = 0$ to infinity, while those of equation (9), the policy block, would have the Taylor and spending rules.

Assume now that we are interested in the effect of a structural shock path ε under some counterfactual policy rule:

$$\hat{\mathcal{A}}_x \mathbf{x} + \hat{\mathcal{A}}_{\mathbf{z}} \mathbf{z} = \mathbf{0}. \tag{9'}$$

In the context of the RANK model sketched above, an example would be that we want to study the effect of a natural rate shock under a Taylor rule with different coefficients, or a strict inflation targeting policy. As in McKay and Wolf (2023), $\mathbf{x}_{\mathcal{A}}(\varepsilon)$ and $\mathbf{x}_{\hat{\mathcal{A}}}(\varepsilon)$ respectively denote the path of the non-policy variables under the prevailing and counterfactual policy rules following shock ε; $\mathbf{z}_{\mathcal{A}}(\varepsilon)$ and $\mathbf{z}_{\hat{\mathcal{A}}}(\varepsilon)$ denote that of policy variables $\mathbf{z}$.

McKay and Wolf (2023) show that the counterfactual response of endogenous variables can be recovered from the IRFs under the *prevailing* policy rule:

$$\mathbf{x}_{\hat{\mathcal{A}}}(\varepsilon) = \mathbf{x}_{\mathcal{A}}(\varepsilon) + \mathbf{\Theta}_{\mathbf{x},\nu,\mathcal{A}} \times \hat{\nu}, \quad \text{and} \tag{11}$$

$$\mathbf{z}_{\hat{\mathcal{A}}}(\varepsilon) = \mathbf{z}_{\mathcal{A}}(\varepsilon) + \mathbf{\Theta}_{\mathbf{z},\nu,\mathcal{A}} \times \hat{\nu}, \tag{12}$$

where $\hat{\nu}$ is the unique solution of

$$\hat{\mathcal{A}}_x \left[\mathbf{x}_{\mathcal{A}}(\varepsilon) + \mathbf{\Theta}_{\mathbf{x},\nu,\mathcal{A}} \times \hat{\nu}\right] + \hat{\mathcal{A}}_z \left[\mathbf{z}_{\mathcal{A}}(\varepsilon) + \mathbf{\Theta}_{\mathbf{z},\nu,\mathcal{A}} \times \hat{\nu}\right] = \mathbf{0}. \tag{13}$$

The essence of this result is that a counterfactual policy rule has the same effect as an appropriate sequence of policy shocks that mimics said counterfactual rule.

This result, however, isn't directly implementable, as it requires knowledge of each element of the $\Theta_{x,\nu,\mathcal{A}}$ and $\Theta_{z,\nu,\mathcal{A}}$ matrices. Empirically, that means knowing the response of endogenous variables to each possible shock path! In reality, it is impossible to estimate the response of macroeconomic variables to more than a few of those. To circumvent this difficulty, McKay and Wolf (2023) propose to solve

$$\min_{\mathbf{s}} \left|\left|\hat{\mathcal{A}}_x \left(\mathbf{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \mathbf{\Omega}_{\mathbf{x},\mathcal{A}} \times \mathbf{s}\right) + \hat{\mathcal{A}}_z \left(\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A}} \times \mathbf{s}\right)\right|\right|. \tag{14}$$

Each column of $\Omega_{x,\mathcal{A}}$ and $\Omega_{z,\mathcal{A}}$ is the empirical estimate of the IRF of $\mathbf{x}$ and $\mathbf{z}$ to a shock path. The minimization problem consists in choosing loadings $\mathbf{s}$ to implement the counterfactual policy rule as well as possible. Note that solving equation (14) does not require specifying the full model, only knowledge of the IRF to the structural shocks ($\mathbf{x}_{\mathcal{A}}(\boldsymbol{\varepsilon})$, $\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$) and to a few policy shock paths ($\Omega_{x,\mathcal{A}}$, $\Omega_{z,\mathcal{A}}$).

Summary: the MKW method answers policy counterfactual questions based on time series regressions. It requires only minimal structural assumptions that embed most standard macroeconomic models. It is immune to the Lucas (1976) critique, because economic agents' expectations about a future policy change are already reflected in the impulse responses to a policy shock path.

4.2 Application to Fiscal-Monetary Interactions

We are interested in the effect of a monetary shock conditional on various fiscal rules. Note that this is not exactly what the method described above was originally meant to do: McKay and Wolf (2023) are concerned with the effect of a structural shock depending on policy rules; we are concerned with the effect of a shock to a policy instrument (the Federal Funds Rate target) depending on rules for other policy instruments (government spending, taxes, and transfers). It is, however, straightforward to extend the framework to the latter case: one can always write the monetary rule in the non-policy block. The monetary shock then plays the role of the structural shock and the instruments are the typical fiscal instruments of a macroeconomic model: government spending, taxes, and transfers.

We stress that our counterfactual experiments hold the monetary rule constant, but that the path of the nominal interest rate may change. Indeed, besides the monetary shock, monetary policy may endogenously respond to the change in macroeconomic conditions prompted by the change in the fiscal rule. If fiscal austerity lowers GDP and inflation, monetary policy may respond and the nominal interest rate may be lower than it would be under the prevailing rule given the monetary shock. Formally, equations (11–12) include endogenous response of all variables ($\Theta_{x,\mathcal{A}}$, $\Theta_{z,\mathcal{A}}$), including those of the policy rate of the central bank.

How does our approach relate to the theoretical literature on fiscal-monetary interactions? First, we are agnostic about whether monetary and fiscal policies are passive or active in the sense of Leeper (1991). Similarly, the economy could be described by a fiscal theory of the price level as much as by a more conventional New Keynesian framework.[22] On the other hand, we assume that there are no regime switches throughout the sample: the monetary rule must remain constant since we place it in the non-policy block.[23] Thus, we do not nest the models of Bianchi (2013) or Bianchi and Ilut (2017). We do nest, however, the latest generation of models of monetary-fiscal interactions, which feature "shock-specific" rules. This new class of models was recently introduced by Bianchi et al. (2023): they propose a framework where the central bank accommodates unfunded transfer shocks with inflation (Fiscally-led rule), while it does not accommodate other fiscal policy shocks by actively responding to inflation (Monetary-led rule), so that "Monetary-led and Fiscally-led rules coexist in [the] model, and the policy coordination is shock-specific" (p. 5). We formally show in appendix F that our approach can be mapped into this framework. Last, monetary-fiscal interactions take place without a regime change in HANK models, largely because of the breakdown of the Ricardian equivalence. Our framework nests these models.

[22]See, for example, the model of Cochrane (2023, chapter 5). The non-policy block is made of equations (A1.51, A1.52, A1.54, A1.56), the policy block equations (A1.53, A1.55, A1.57).

[23]Changes in the fiscal rules over time would be easily handled since they are in the policy block (McKay and Wolf, 2023, appendix A.4).

4.3 Implementation

4.3.1 Fiscal Shocks

We use a variety of fiscal policy shocks. The MKW procedure ideally demands an infinity of those. To approach that ideal, we have extensively surveyed the literature starting from Valerie Ramey's handbook chapter on "Macroeconomic Shocks and Their Propagation" (Ramey, 2016).

Most fiscal shocks fall under two umbrellas: narrative approach and structural identification. The narrative approach relies on a reading of the historical record: the researcher reads official documents to identify the rationale for policy changes. If this rationale is exogenous to the state of the economy, the change is retained as a valid policy shock, rejected otherwise. While there exists many variations, every series of narrative shock stems from a seminal paper: Ramey (2011) for spending, Romer and Romer (2010) for taxes, and Romer and Romer (2016) for transfers. Structural identification relies on a structural vector autoregression (SVAR). In the simplest case, researchers identify the fiscal shock with a Cholesky decomposition: they regress the fiscal variable on past values of several macroeconomic variables and assume that the residual doesn't respond contemporaneously to those variables. Therefore, said residual is a valid shock to infer the effect of fiscal policy. While the assumption that spending doesn't respond contemporaneously to GDP is plausible, it is clearly dubious for taxes: at given marginal tax rates, tax receipts should be positively correlated with GDP. Hence, more elaborate versions of this scheme control for the contemporaneous response of taxes to macroeconomic variables: this approach was pioneered by Blanchard and Perotti (2002) and perfected by Caldara and Kamps (2017).

For each fiscal instrument, we have one narrative and one structural shock series (table 1). We make a quick list here and describe their construction in detail in appendix I. We always take the narrative shocks from the reference papers cited above. The spending shocks are Ramey's original series. We use the Romer–Romer legislated tax changes motivated by long-run growth.[24] For transfers, Romer and Romer distinguish long-run from temporary changes in Social Security benefits. We only use the long-run changes since temporary changes have no distinguishable effect in quarterly data. For structurally identified shocks, we use the Cholesky identification for spending and the Caldara–Kamps identification for taxes and transfers.[25]

[24] See section 3.2.1 in the main text for details on their methodology. See also section I.3 in the appendix for why we use these shocks instead of later variations introduced by Mertens and Ravn (2012, 2013).

[25] In the baseline, we use a version of Caldara and Kamps's full fiscal rule that controls for the contemporaneous response of fiscal variables to our 3 macroeconomic variables (GDP, inflation, and the interest rate). In appendix B, we also show some results for the simple fiscal rule, which only allows for a contemporaneous response to GDP. See appendix I for more details on the Caldara–Kamps methodology.

Table 1: Fiscal shocks

Identification	Description	Source
	Spending	
Narrative	Future changes in military spending constructed by reading the specialized press	Ramey (2011)
Cholesky	Cholesky identification with spending ordered first	Blanchard and Perotti (2002)
	Taxes	
Narrative	Legislated tax changes motivated by long-run growth	Romer and Romer (2010)
Proxy	Reduced form shock for taxes purged from contemporaneous response to non-fiscal variables (GDP, inflation, nominal interest rate)	Caldara and Kamps (2017)
	Transfers	
Narrative	Long-run transfer legislated changes	Romer and Romer (2016)
Proxy	Reduced form shock for transfers purged from contemporaneous response to non-fiscal variables (GDP, inflation, nominal interest rate)	Our computation based on methodology of Caldara and Kamps (2017)

Note: See appendix I for more details on the construction of these shocks.

4.3.2 Specification

For each of the three fiscal instruments, we follow the procedure described in section 2: we estimate a VAR with the narrative fiscal shocks ordered first and the eight endogenous variables of interest (five fiscal and three macroeconomic variables).

To be internally consistent, we always re-estimate the structural shock within our VAR. For instance, Blanchard and Perotti (2002) identify spending shocks by running a VAR with spending, taxes, and GDP and running a Cholesky decomposition with spending ordered first. So, the identification assumption is that spending does not respond contemporaneously to other variables. Since we have more variables in our VAR, we adapt this scheme by ordering spending after the narrative shock, but before the other endogenous variables. Thus, our structural shock is not exactly the same as that of Blanchard and Perotti, although it is identified in the same spirit.

Following McKay and Wolf (2023), we fix the response to the monetary shock at the point estimate shown in figure 1 and use the joint distribution of the IRFs to the two fiscal shocks to construct the counterfactual scenario. Thus, the credible bands account for the joint uncertainty of the response to these shocks, but not for the uncertainty in the response to the initial monetary shock. This procedure also implies that we re-estimate the coefficients of the reduced-form VAR for the fiscal shocks, allowing them to differ from those of the monetary VAR. According to Plagborg-Møller and Wolf (2021), whose results we have already referred to in section 2.3, VARs and local projections estimate the same responses in population. Therefore, our monetary and fiscal impulse responses can be interpreted as arising from local projections, which allows us to combine them consistently.

5 Counterfactual: Results

5.1 Response to Fiscal Shocks

The response of macroeconomic variables to fiscal shocks is not our main focus, so we relegate those to the appendix, figures A.12–A.17. We only comment here on the sometimes counter-intuitive effects of those shocks on deficit and debt.

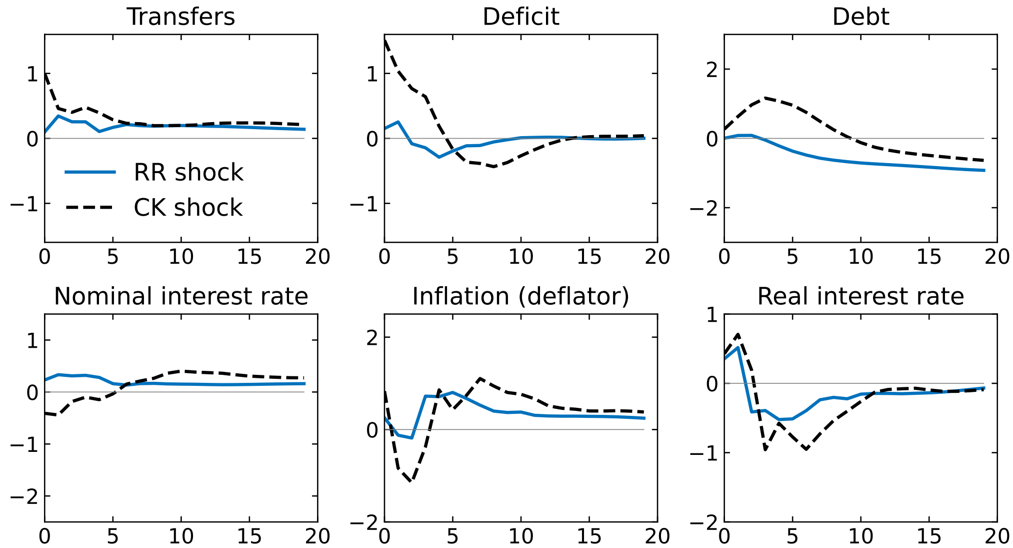
Spending tends to increase deficit and debt, but the short run effect can be different (figures A.12–A.13). For instance, the Ramey shocks only increase spending after a few quarters, but increase GDP immediately. Ramey argues that this feature reflects anticipation effects. The implication is that the deficit falls on impact as tax receipts endogenously rise. Moreover, inflation increases, which cuts the real value of debt. As a result, the immediate effect of a Ramey spending hike is to lower the deficit and cut real debt. The Blanchard-Perrotti shocks, since they are more front-loaded, imply more intuitive patterns of deficit and debt.

Tax increases lower the deficit in the short run, but the effect dissipates quickly, perhaps because of a strong Laffer curve effect (figures A.14–A.15). Moreover, tax increases tend to lower inflation, thus raising the real value of debt. On balance, the first effect (short-lived deficit versus increase in the real value of debt) dominates, so that the effect on debt is negative.

Transfer shocks have radical implications (figures 9 and A.16–A.17). Increases in transfers powerfully stimulate GDP, so that tax receipts increase enough to mute or even reverse the effect on the deficit. Besides, they're so inflationary that the real value of debt is lower after 5 years.

Those results echo two recent theoretical contributions. Angeletos et al. (2024) show that self-financed increases in transfers are possible in theory. In their model, the main mechanism is the non-Ricardian behavior of households. If their marginal propensity to consume (MPC) is sufficiently high, they spend the transfer, the economy is stimulated so tax receipts increase and the deficit shrinks: transfers pay for themselves (at least partially). Our results prove that this mechanism is credible empirically. In fact, the point estimates

Figure 9: Response of selected variables to transfer shocks



Note: Selected results from the VAR with transfer shocks (figures A.16–A.17). This VAR includes the Romer–Romer transfer shock, spending, tax receipts, transfers, interest payments, debt, GDP (all in real terms), inflation, and the 3-month nominal interest rate. RR and CK shocks stand for Romer–Romer and Caldara–Kamps transfer shocks, respectively.

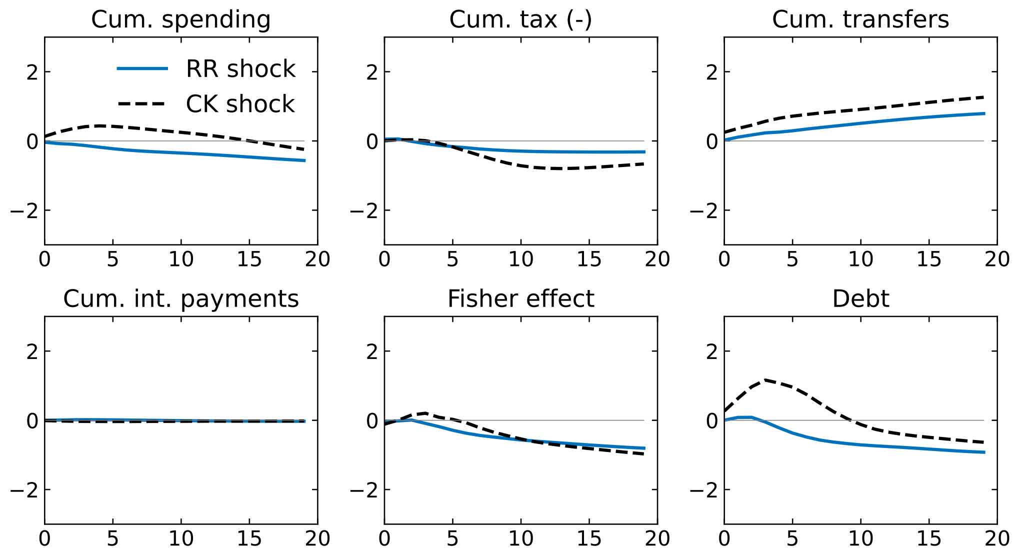
suggest that, not only are transfer hikes self-financed, they leave the government with less debt![26] One reason reality may overcome theory might be that the real interest rate falls within a year. In their model, the central bank is neither accommodating nor fighting the transfer shock. In our sample, it seems to have been accommodating it. This observation is consistent with the estimated model of Bianchi et al. (2023), who find that unfunded transfer shocks—shocks which are financed by inflation because the central bank behaves accommodatingly—were prevalent in post-WWII data. We find evidence that is consistent with this phenomenon in a purely empirical framework.

To better understand our transfer result, we decompose the response of debt in figure 10. Compared to figure 9, the response of the fiscal flow variables is shown in cumulative terms so that they sum to the response of debt. The increase in transfers is partly compensated by an increase in tax collection driven by the increase in output—we show the negative of the cumulative response of taxes since an increase in tax collection corresponds to a fall in debt. The cumulative response of interest payments is muted. Cumulative spending temporarily increases in response to the Caldara–Kamps shock, but dissipates after two years. More importantly, this additional expenditure is dwarfed by the Fisher effect. As

[26] We investigated the robustness of this result to adding another year of lags in the VAR, using Caldara and Kamps's simple fiscal rule, or putting taxes in front of transfers to identify the CK shock. See figure A.18. We cannot always reject 0, but the point estimate is always negative.

a result, increased tax collection and Fisher effect both contribute to the stabilization, or even fall, of debt. This figure is consistent with the two models that we have mentioned. The prominence of the Fisher effect evokes Bianchi et al.'s emphasis on the fiscal origins of inflation. The contribution of tax receipts echoes Angeletos et al.'s Keynesian cross logic.

Figure 10: Response of debt and GDP components to transfer shocks



Note: Selected results from the VAR with transfer shocks (figures A.16–A.17). This VAR includes the Romer–Romer transfer shock, spending, tax receipts, transfers, interest payments, debt, GDP (all in real terms), inflation, and the 3-month nominal interest rate. RR and CK shocks stand for Romer–Romer and Caldara–Kamps transfer shocks, respectively. The first five IRFs (cumulative spending, taxes, transfers, interest payments, and Fisher effect) sum to the sixth one (debt).

In appendix figure A.19, we add the two main components of GDP, consumption and investment, to our VAR. The increase in consumption mirrors the increase in GDP, with a sharp peak two quarters after a Romer–Romer shock and a more prolonged increase between the second and eighth quarters after a Caldara–Kamps shock. Investment increases with the same timing.

The concomitant increases of consumption or investment with that of output echo the general equilibrium effects of HANK models. In the models of Kaplan et al. (2018) and Auclert et al. (2024), because households have a high marginal propensity to consume, transfers stimulate consumption, which stimulates output; because output increases, households' income increases, which prompts more consumption, and so on. More recently, Winberry et al. (2025) show that a similar general equilibrium effect can arise in a New Keynesian model with heterogeneous firms, in which constrained firms translate output movements to investment demand. Our empirical result in figure A.19 that investment increases along with output in response to a transfer shock is also consistent with their suggested mechanism.

We emphasize the fiscal implications of fiscal shocks because they have received surprisingly little attention in the empirical literature. The papers that proposed these fiscal shocks tend to focus on their macroeconomic consequences, often neglecting their fiscal side effects. Exceptions are Mertens and Ravn (2013), who found a muted effect of tax changes on deficit and debt, and Ramey (2016), who noted that narrative tax changes are a weak instrument for tax receipts.

5.2 Counterfactual Policy: Debt Stabilization

We study the counterfactual response of the economy to a monetary shock, assuming that the federal government stabilizes real debt by changing one of its instruments: spending, taxes, or transfers.

We show the first scenario, a spending cut, in figure 11. The government sharply curtails spending. This spending cut, however, is partly self-defeating: it lowers GDP, which lowers tax receipts, which raises the deficit. So, the government must cut spending more than one-for-one to obtain the desired stabilization. (Since our fiscal variables are expressed as a fraction of trend GDP, the magnitudes on the y-axis can be directly compared across variables.) As a result, this scenario implies a much more pronounced fall in output than the prevailing rule, as well as more deflation in the short run.

The second scenario, a tax hike, delivers a similar message (figure 12). The government raises tax rates to run down the deficit and stabilize debt. GDP and inflation fall. This fall is less brutal, but more front-loaded than with spending.

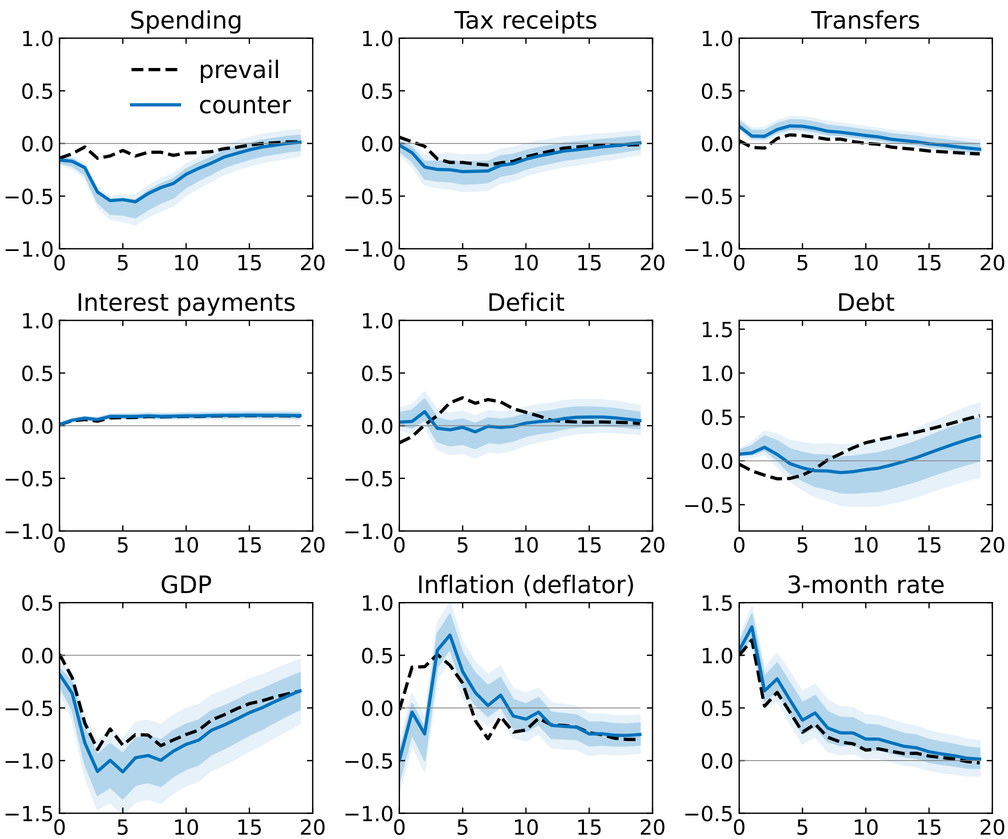
The third scenario, a change in transfers, delivers a surprising insight: the government can stabilize debt with a transfer hike (figure 13)! Transfers increase the deficit, but that effect is dampened by the increase in GDP, which stimulates tax receipts. Moreover, the transfers are very inflationary, which reduces real debt. The real interest rate falls compared to the prevailing rule.

We reiterate that these counterfactual experiments hold the monetary rule constant, but not necessarily the path of the nominal interest rate. That is why the path of the nominal interest rate under each counterfactual scenario is not exactly the same as under the prevailing rule.

5.3 Counterfactual Policy: Robustness

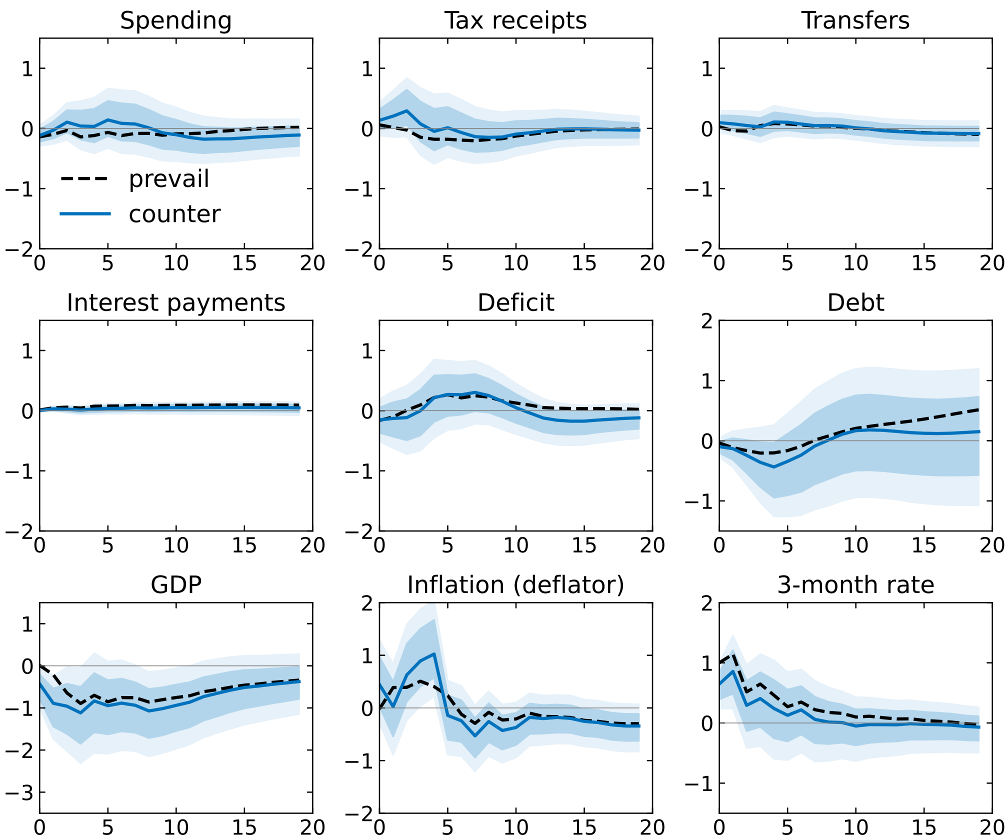
In appendix H, we present a different way to construct our credible bands. In their article, McKay and Wolf (2023) are silent about the construction of credible (or confidence) bands for their method. Should the loadings on the fiscal shocks, $\mathbf{s}$, remain constant or should they

Figure 11: Counterfactual—debt stabilization with spending



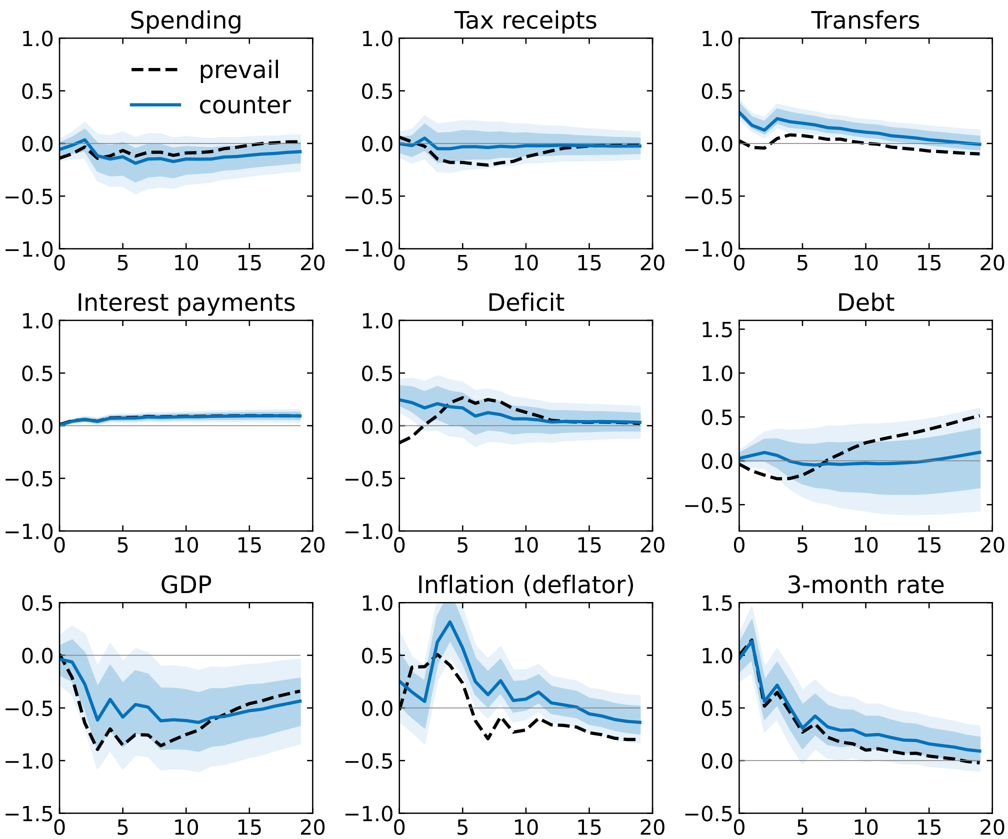
Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through spending. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

Figure 12: Counterfactual—debt stabilization with taxes



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through taxes. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

Figure 13: Counterfactual—debt stabilization with transfers



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through transfers. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

be updated for each draw of the distribution? The first method holds the loadings on the shocks constant, while the second one re-optimizes those for every draw. Studying the code of McKay and Wolf's replication package, we found that they use the latter. In our context, however, the former may be slightly more sensible since a given fiscal instrument is actually a collection of instruments.[27] A tax shock, for instance, is a collection of changes in many marginal tax rates across several types of taxes (individual income, corporate incomes, *etc.*). So, updating loadings across draws changes the policy mix across the distribution, thereby exacerbating the volatility of the underlying policy.

In the body of the text, we use the first method; in appendix H, we present the second one. Under the second method, the credible bands are tighter for debt by construction in the three counterfactual scenarios; they are similar for other variables in the spending and transfers scenarios. In the tax scenario, the distribution becomes asymmetric, with the modal IRF lying near the lower bound of the 68% credible band. We attribute this asymmetry to the similar shapes of the debt responses implied by the two shock series. If we use only one of the series (Romer–Romer or Caldara–Kamps), the credible bands tighten, although we do not perfectly stabilize debt in the short run.

In appendix figures A.26–A.28, we look at a different counterfactual rule: deficit stabilization. The main message is similar for spending and transfers. Stabilizing the deficit with spending requires a stronger contraction in output; stabilizing the deficit with transfers lessens the contraction in output. We fail to really change the deficit when it comes to taxes. Indeed, our tax shocks mostly change the deficit in the short run, while the monetary shock raises it in the medium run. This illustrates a limitation of the MKW method: with only two shocks, the counterfactual scenario is not perfectly enforced. This limitation is particularly pronounced when one tries to stabilize a flow instead of a stock. Our debt counterfactual fares better because, in each period, we use the whole cumulative path of past values of deficit and inflation.

5.4 Summary and Takeaway for HANK Models

We summarize our results in table 2 by computing the average response of the main variables over the first 3 years that follow the monetary shock.[28] In reality, GDP falls by 0.66% compared to trend for a shock that increases the real interest rate by 37 basis points. This implies an elasticity of $0.66/0.37 = 1.78$.[29] It may seem surprising that the average response

[27]We show how we can map unobserved fiscal instruments into our framework in section F.

[28]In tables B.3–B.5, we compute these averages for 1, 2, and 4 years after the shock.

[29]Using a VAR and the original version of the Romer–Romer shocks, Nakamura and Steinsson (2018b, table A.1) report an elasticity of 0.7.

of inflation is positive and insignificant: this result is due to the initial (insignificant) increase in inflation apparent on figure 1. Inflation only starts falling after a year. This is a well-known implication of the Romer–Romer shocks: the price level takes a long time to start falling. If we focus on the second and third years that follow the shock, the average response of inflation is negative (-11 basis points).

Table 2: Monetary policy and fiscal response—3-year average

		Counterfactual		
	Actual	Spending	Taxes	Transfers
	(1)	(2)	(3)	(4)
GDP	-0.66	-0.84	-0.91	-0.45
	(0.37)	(0.16)	(0.49)	(0.23)
	[-1.12,-0.40]	[-1.02,-0.72]	[-1.44,-0.50]	[-0.66,-0.21]
Inflation	0.07	0.07	0.07	0.28
	(0.15)	(0.07)	(0.20)	(0.09)
	[-0.06, 0.23]	[0.00, 0.13]	[-0.12, 0.26]	[0.22, 0.40]
Nominal interest rate	0.43	0.53	0.23	0.50
	(0.25)	(0.10)	(0.32)	(0.16)
	[0.27, 0.73]	[0.45, 0.65]	[-0.09, 0.52]	[0.37, 0.67]
Real interest rate	0.37	0.43	0.22	0.23
	(0.22)	(0.10)	(0.30)	(0.14)
	[0.22, 0.64]	[0.36, 0.54]	[-0.08, 0.49]	[0.09, 0.36]

Note: Average response over the first 3 years to a Romer–Romer+ monetary shock, depending on the fiscal response. Column (1) is the actual response described in section 3. Columns (2–4) are the counterfactual responses under the three scenarios described in section 5.2: debt stabilization through spending, taxes, or transfers. They respectively correspond to figures 11, 12, and 13. The number in parentheses is the standard error. The numbers between brackets are the bounds of the 68% credible interval.

As expected from section 5.2, GDP falls more if spending or taxes stabilize debt, less if transfers do. The response of GDP is 1.3 times larger with spending, 1.4 times with taxes. With transfers, the response is less pronounced: GDP falls by 0.45% compared to trend on average. So, in light of our results, it seems plausible that transfers can stabilize debt at no cost in terms of GDP. This statement does not imply that the outcome is achieved at no cost. The second row of the table makes clear one reason why transfers pay for themselves: inflation. We do not make any welfare statement here. Yet, the response of inflation to the transfer policy is a reminder that stabilizing debt and GDP is not necessarily optimal.

Are these results consistent with HANK models? Matching the many IRFs that we have estimated in a quantitative heterogeneous-agent model is beyond the scope of this paper. Yet, we can venture an informal comparison with two important models of that literature whose authors conduct similar experiments: Kaplan et al. (2018, table 8) and Auclert et al. (2020, figure 7). Interestingly, both models imply that letting debt adjust—what we call reality—is less destabilizing than using taxes or spending. This qualitative ranking is consistent with our estimates.[30] On the other hand, neither model predicts that transfers stabilize output compared to debt adjustment. This shouldn't come as a surprise: in both models, the government borrows in real debt, which kills the Fisherian channel of transfers. This channel is key to our counterfactual results.

We expect the framework of Bianchi et al. (2023) to generate predictions that are qualitatively similar to our results, should the fiscal authority stabilize debt with unfunded transfers after a monetary contraction—the authors do not explicitly conduct this experiment. In their model, an unfunded transfer shock is expansionary and lowers real debt (figure IV). By the logic of the MKW method, a systematic increase in unfunded transfers in response to a monetary shock should therefore stabilize debt while attenuating the contraction.

6 Conclusion

In this paper, we estimate the response of fiscal policy to monetary shocks and the response of the economy under counterfactual rules for fiscal policy. In reality, fiscal policy mostly relies on debt to deal with the fiscal consequences of a monetary policy action. This strategy dampens the effect of monetary policy on output compared to debt stabilization through spending or taxes, but it amplifies it compared to debt stabilization through transfers.

[30]The model of Kaplan et al. has no macroeconomic persistence or long-term debt, so any quantitative comparison would be heroic. The model of Auclert et al. achieves persistence through sticky expectations and long-term debt; moreover, it is estimated out of an IRF to a Romer–Romer monetary shock. Auclert et al. do not report the average response of output, but eyeballing figure 7 suggests that the response of output if spending (resp. taxes) is used to clear the budget constraint is twice (resp. 1.5 times) larger than if debt adjusts. These ratios would be smaller than we estimate, but potentially within the credible intervals (table 2).

References

Alves, F., G. Kaplan, B. Moll, and G. L. Violante (2020). A further look at the propagation of monetary policy shocks in HANK. *Journal of Money, Credit and Banking 52*(S2), 521–559.

Angeletos, G.-M., C. Lian, and C. K. Wolf (2024). Can deficits finance themselves? *Econometrica 92*(5), 1351–1390.

Aruoba, S. B. and T. Drechsel (2024). Identifying monetary policy shocks: A natural language approach. Working paper.

Ascari, G., P. Beck-Friis, A. Florio, and A. Gobbi (2023). Fiscal foresight and the effects of government spending: It's all in the monetary-fiscal mix. *Journal of Monetary Economics 134*, 1–15.

Auclert, A., M. Rognlie, and L. Straub (2020). Micro jumps, macro humps: Monetary policy and business cycles in an estimated HANK model. Working paper.

Auclert, A., M. Rognlie, and L. Straub (2024). The intertemporal Keynesian cross. *Journal of Political Economy 132*(12), 4068–4121.

Barnichon, R. and G. Mesters (2023). A sufficient statistics approach for macroeconomic policy. *American Economic Review 113*(11), 2809–2845.

Bauer, M. D. and E. T. Swanson (2023a). An alternative explanation for the "Fed information effect". *American Economic Review 113*(3), 664–700.

Bauer, M. D. and E. T. Swanson (2023b). A reassessment of monetary policy surprises and high-frequency identification. *NBER Macroeconomics Annual 37*, 87–155.

Bianchi, F. (2013). Regime switches, agents' beliefs, and post-world war ii u.s. macroeconomic dynamics. *Review of Economic Studies 80*(2), 463–490.

Bianchi, F., R. Faccini, and L. Melosi (2023). A fiscal theory of persistent inflation. *Quarterly Journal of Economics 138*(4), 2127–2179.

Bianchi, F. and C. Ilut (2017). Monetary/fiscal policy mix and agents' beliefs. *Review of Economic Dynamics 26*, 113–139.

Blanchard, O. (2019). Public debt and low interest rates. *American Economic Review 109*(4), 1197–1229.

Blanchard, O. and R. Perotti (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics 117*(4), 1329–1368.

Breitenlechner, M., M. Geiger, and M. Klein (2025). The fiscal channel of monetary policy.

Caldara, D. and C. Kamps (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *Review of Economic Studies 84*, 1015–1040.

Caramp, N. and D. Silva (2018). Fiscal origins of monetary paradoxes. Working paper.

Cochrane, J. (2023). *The Fiscal Theory of the Price Level*. Princeton: Princeton University Press.

Cochrane, J. H. (1999). A frictionless view of U.S. inflation. *NBER Macroeconomics Annual 13*, 323–384.

Coibion, O. (2012). Are the effects of monetary policy shocks big or small? *American Economic Journal: Macroeconomics 4*(2), 1–32.

Del Negro, M. and F. Schorfheide (2012). Bayesian macroeconometrics. In J. Geweke, G. Koop, and H. Van Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, Chapter 7, pp. 293–389. Oxford, UK: Oxford University Press.

Galí, J. (2015). *Monetary Policy, Inflation, and the Business Cycle: An Introduction to the New Keynesian Framework and Its Applications* (Second ed.). Princeton: Princeton University Press.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Gilchrist, S. and E. Zakrajšek (2011). Credit spreads and business cycle fluctuations. *American Economic Review 102*(4), 1692–1720.

Giordani, P., M. Pitt, and R. Kohn (2012). Bayesian inference for time series state space models. In J. Geweke, G. Koop, and H. Van Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, Chapter 3, pp. 60–124. Oxford, UK: Oxford University Press.

Gordon, R. J. and R. Krenn (2010). The end of the great depression 1939–41: Policy contributions and fiscal multipliers. Working paper.

Jarociński, M. and P. Karadi (2020). Deconstructing monetary policy surprises—The role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 186*(1007), 453–461.

Jordà, O. (2005). Estimation and inference of impulse responses by local projections. *American Economic Review 95*(1), 161–182.

Jordà, O., M. Schularick, and A. M. Taylor (2015). Betting the house. *Journal of International Economics 96*, S2–S18.

Kaplan, G., B. Moll, and G. L. Violante (2018). Monetary policy according to HANK. *American Economic Review 108*(3), 697–743.

Kurcz, F. (2025). Quantifying the fiscal channel of monetary policy.

Leeper, E. (1991). Equilibria under 'active' and 'passive' monetary and fiscal policies. *Journal of Monetary Economics 27*, 129–147.

Li, D., M. Plagborg-Møller, and C. K. Wolf (2022). Local projections vs. vars: Lessons from thousands of dgps. Working paper.

Lucas, Jr., R. E. (1976). Econometric policy evaluation: A critique. *Carnegie-Rochester Conference Series on Public Policy 1*, 19–46.

McKay, A. and C. K. Wolf (2023). What can time-series regressions tell us about policy counterfactuals? *Econometrica 91*(5), 1695–1725.

Mehrotra, N. R. and D. Sergeyev (2021). Debt sustainability in a low interest rate world. *Journal of Monetary Economics 124S*, S1–S18.

Mertens, K. and M. O. Ravn (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy 4*(2), 145–181.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review 103*(4), 1212–1247.

Miranda-Agrippino, S. and G. Ricco (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics 13*(2), 74–107.

Montiel Olea, J. L. and M. Plagborg-Møller (2021). Local projection inference is simpler and more robust than you think. *Econometrica 89*(4), 1789–1823.

Nakamura, E. and J. Steinsson (2018a). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Nakamura, E. and J. Steinsson (2018b). Identification in macroeconomics. *Journal of Economic Perspectives 32*(3), 59–86.

Newey, W. K. and K. D. West (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica 55*(3), 703–708.

Payne, J., B. Szőke, G. Hall, and T. J. Sargent (2025, 12). Costs of financing U.S. federal debt under a gold standard: 1791–1933. *Quarterly Journal of Economics 140*(1), 793–833.

Plagborg-Møller, M. and C. K. Wolf (2021). Local projections and VARs estimate the same impulse responses. *Econometrica 89*(2), 955–980.

Ramey, V. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics 126*(1), 1–50.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2A, pp. 71–162. Amsterdam: Elsevier.

Romer, C. D. and D. H. Romer (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review 94*(4), 1055–1084.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Romer, C. D. and D. H. Romer (2016). Transfer payments and the macroeconomy: The effects of Social Security benefit increases, 1952–1991. *American Economic Journal: Macroeconomics 8*(4), 1—42.

Schorfheide, F. and D. Song (2015). Real-time forecasting with a mixed-frequency VAR. *Journal of Business & Economic Statistics 33*(3), 366–380.

Sterk, V. and S. Tenreyro (2018). The transmission of monetary policy through redistributions and durable purchases. *Journal of Monetary Economics 99*, 124–137.

Stock, J. H. and M. W. Watson (2010). Distribution of quarterly values of GDP/GDI across months within the quarter. Research memorandum.

Stock, J. H. and M. W. Watson (2012). Disentangling the channels of the 2007-–09 recession. *Brookings Papers on Economic Activity 42*(1), 81–135.

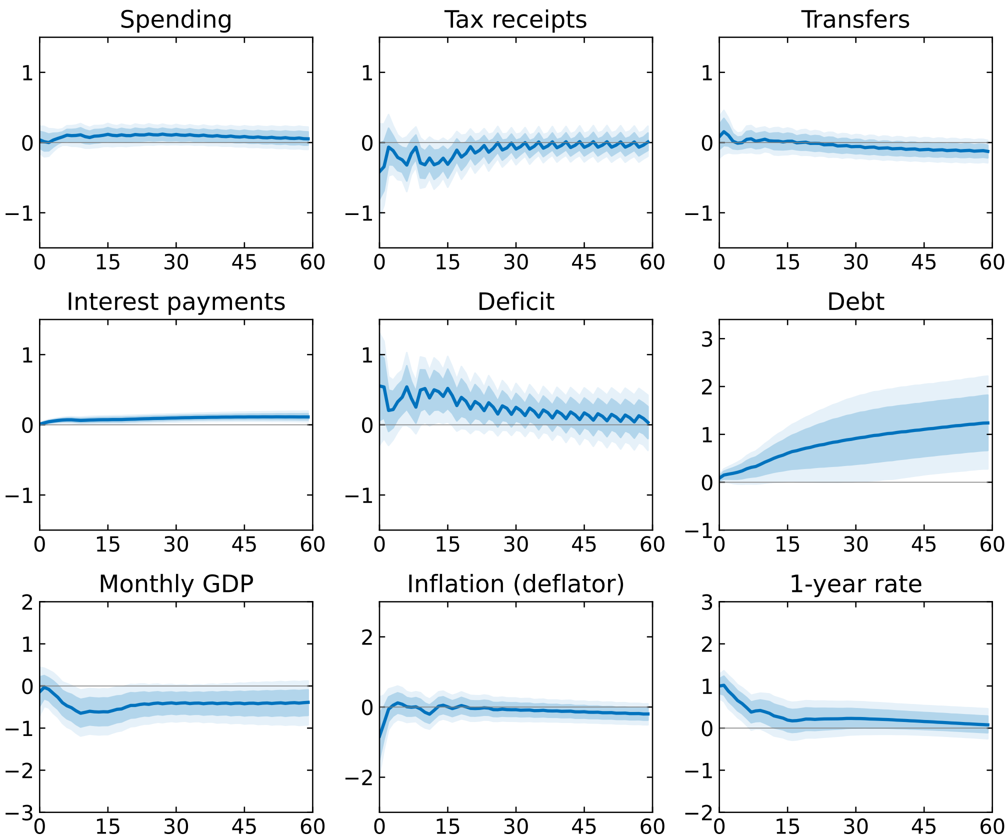
Winberry, T., A. Auclert, M. Rognlie, and L. Straub (2025). New Keynesian economics with household and firm heterogeneity. NBER Working Paper No. 34611.

Wolf, C. K. (2023). The missing intercept: A demand equivalence approach. *American Economic Review 113*(8), 2232–2269.

ONLINE APPENDIX

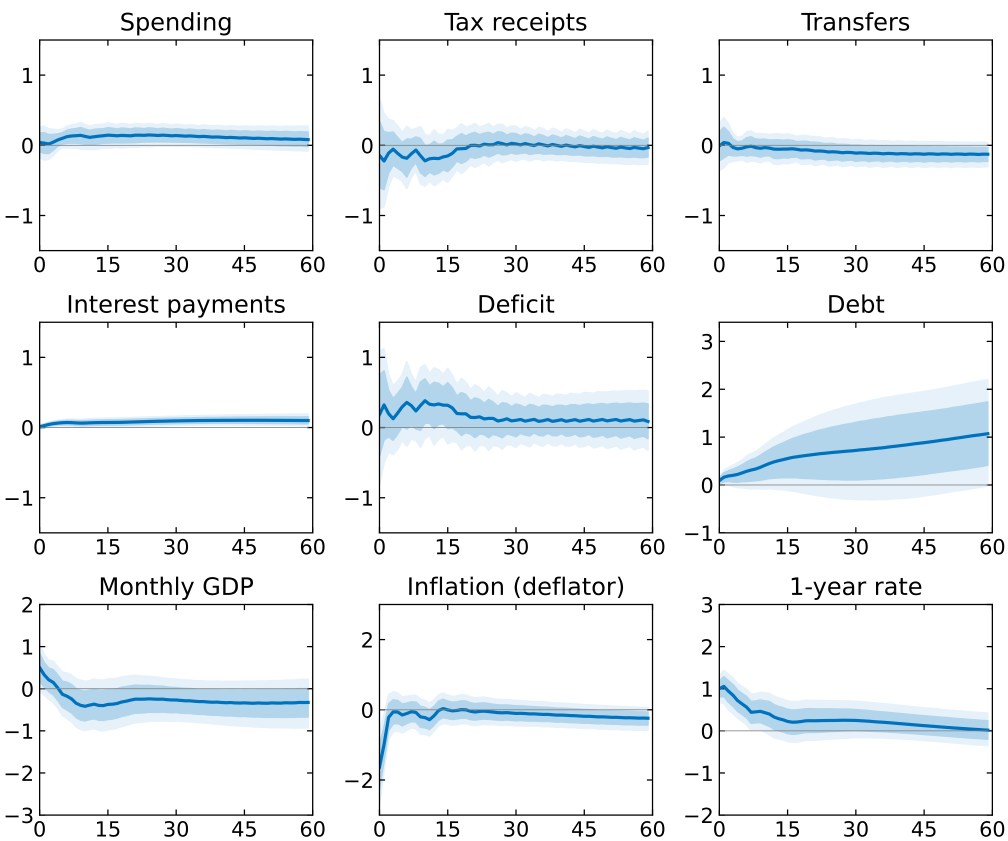
A Additional Figures

Figure A.1: Response to Jarociński–Karadi monetary shock



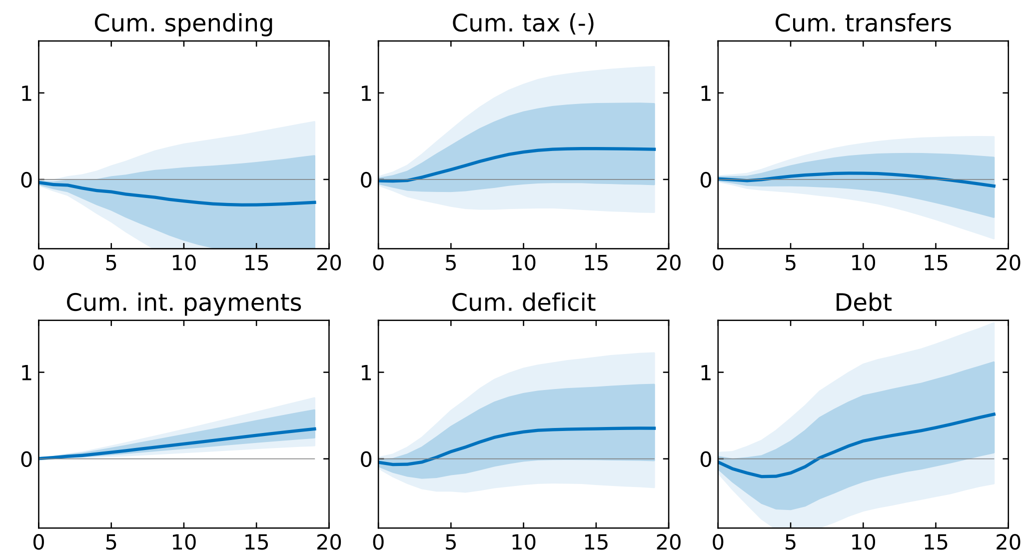
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of the nominal interest rate is 1 at time 0. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.2: Response to Miranda-Aggrippino–Ricco+ monetary shock



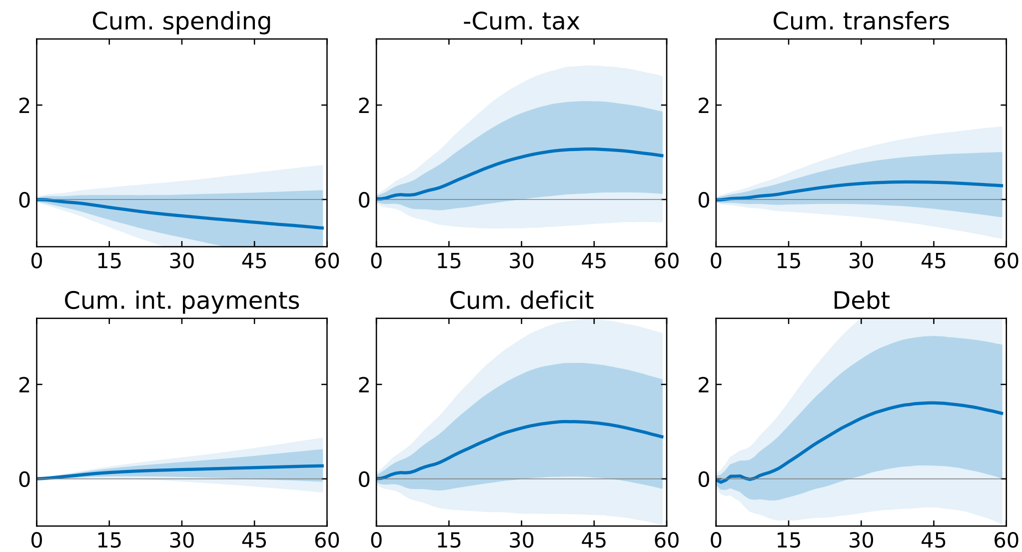
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of the nominal interest rate is 1 at time 0. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.3: Decomposition of response of deficit to Romer–Romer+ monetary shock



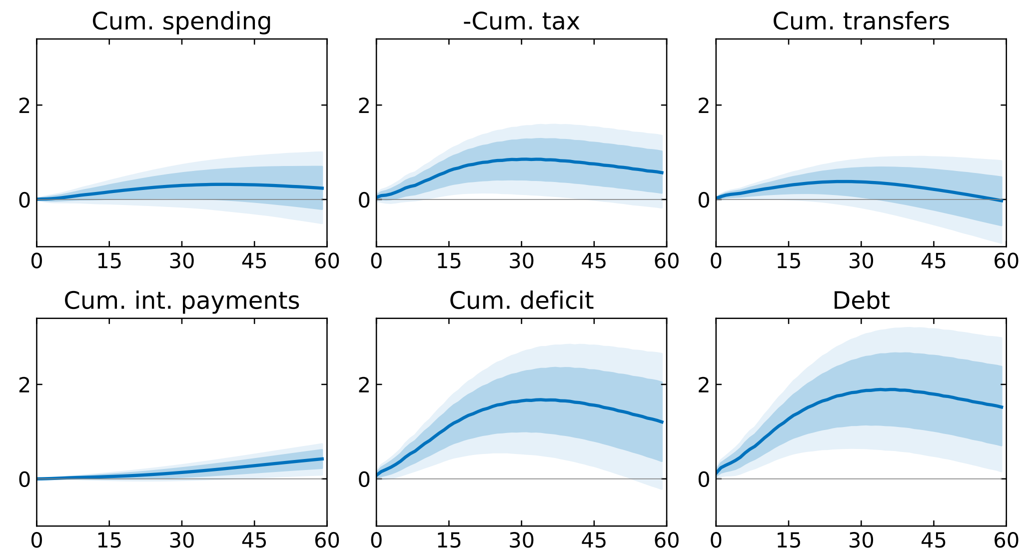
Note: Additional results from the baseline VAR (figure 1). The last chart is the same as the debt chart of figure 1. The other charts are the cumulative responses of spending, tax receipts, transfers, interest payments, and the deficit. The first four responses exactly sum to the fifth. These cumulative responses are deduced from equation (C.3). The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals.

Figure A.4: Decomposition of response of deficit to Aruoba–Drechsel monetary shock



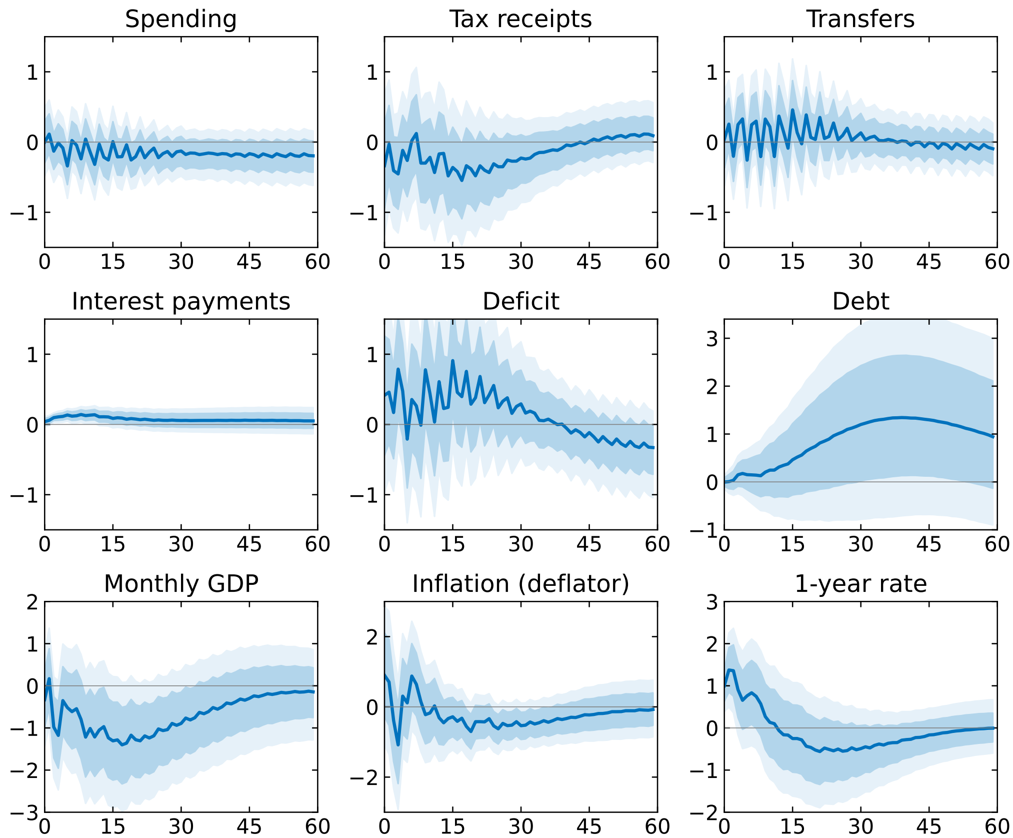
Note: Additional results from the VAR with Aruoba–Drechsel shocks (figure 2). The last chart is the same as the debt chart of figure 2. The other charts are the cumulative responses of spending, tax receipts, transfers, interest payments, and the deficit. The first four responses exactly sum to the fifth. These cumulative responses are deduced from equation (C.3). The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals.

Figure A.5: Decomposition of response of deficit to Bauer–Swanson monetary shock



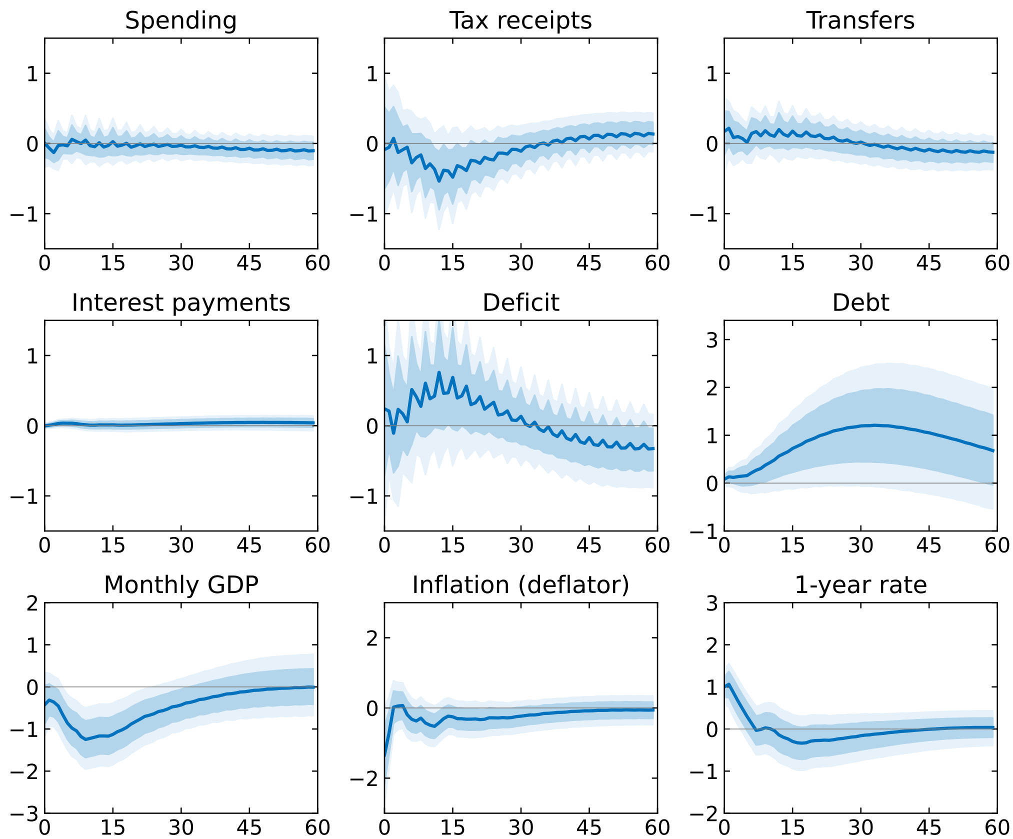
Note: Additional results from the VAR with Bauer–Swanson shocks (figure 3). The last chart is the same as the debt chart of figure 3. The other charts are the cumulative responses of spending, tax receipts, transfers, interest payments, and the deficit. The first four responses exactly sum to the fifth. These cumulative responses are deduced from equation (C.3). The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals.

Figure A.6: Aruoba–Drechsel with debt from the Financial Accounts



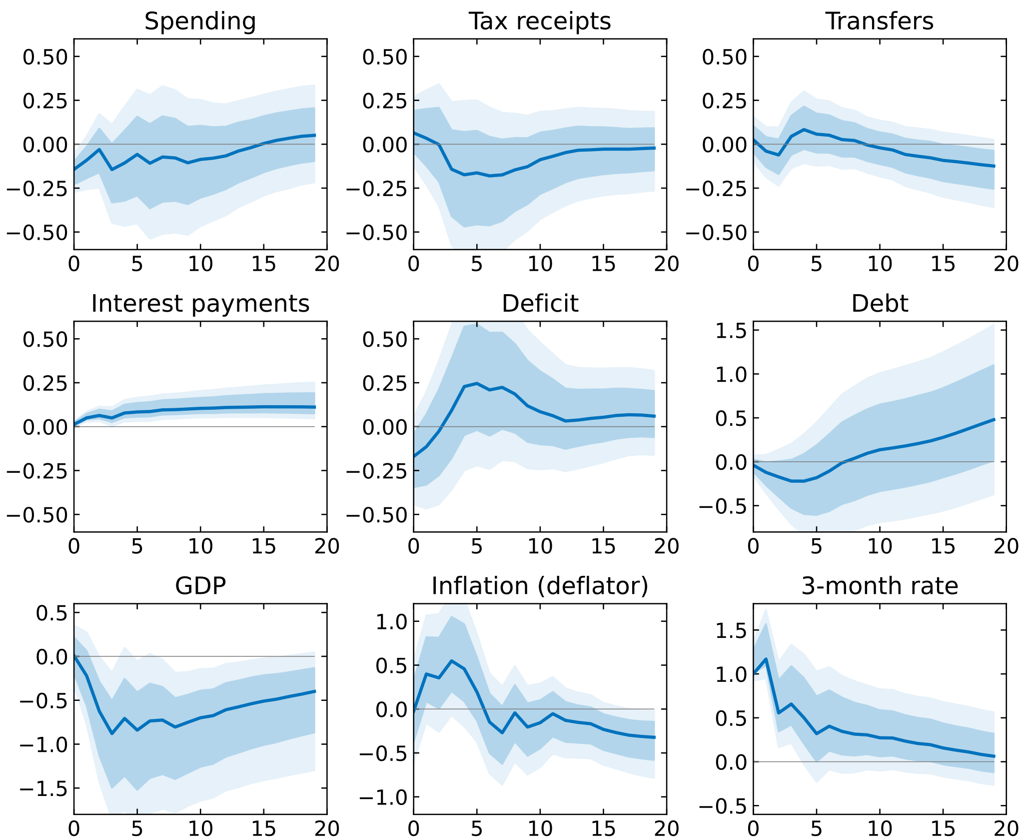
Note: Same VAR as figure 2 with the face value of debt replaced by net debt from the Financial Accounts. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.3.1 for details on these debt concepts. This figure should be taken with caution as the effective sample size is low (table E.1).

Figure A.7: Bauer–Swanson with debt from the Financial Accounts



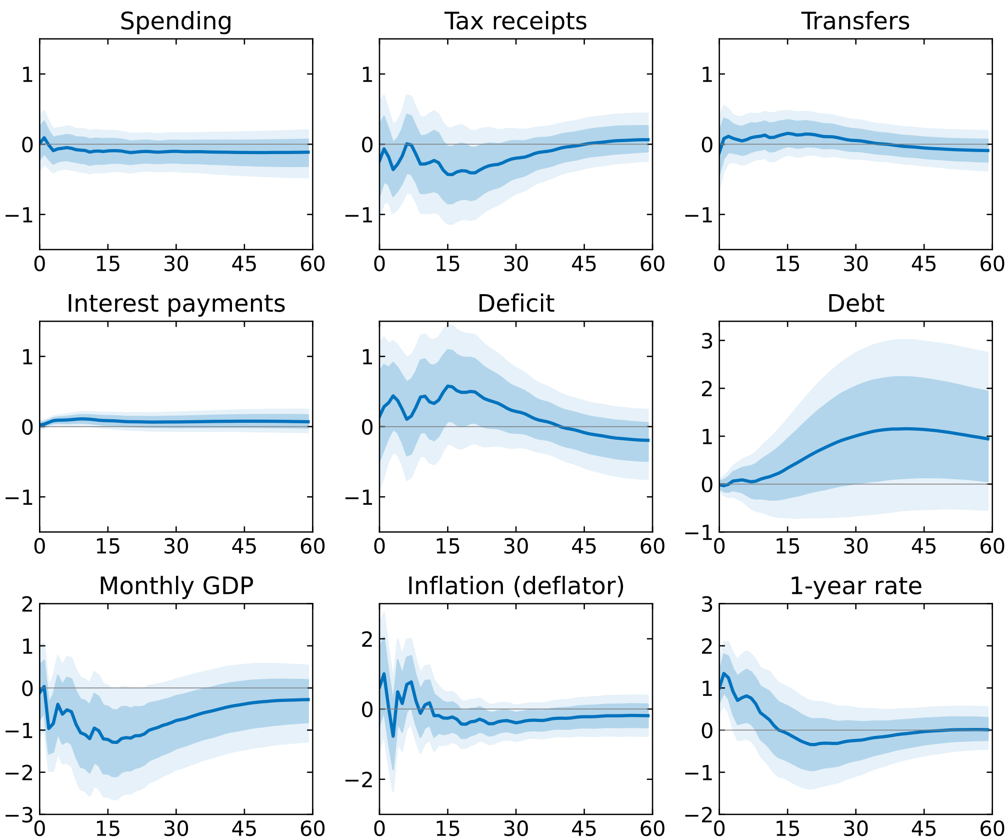
Note: Same VAR as figure 3 with the face value of debt replaced by net debt from the Financial Accounts. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section 3.3.1 for details on these debt concepts. This figure should be taken with caution as the effective sample size is low (table E.1).

Figure A.8: Constrained VAR (Romer–Romer+)



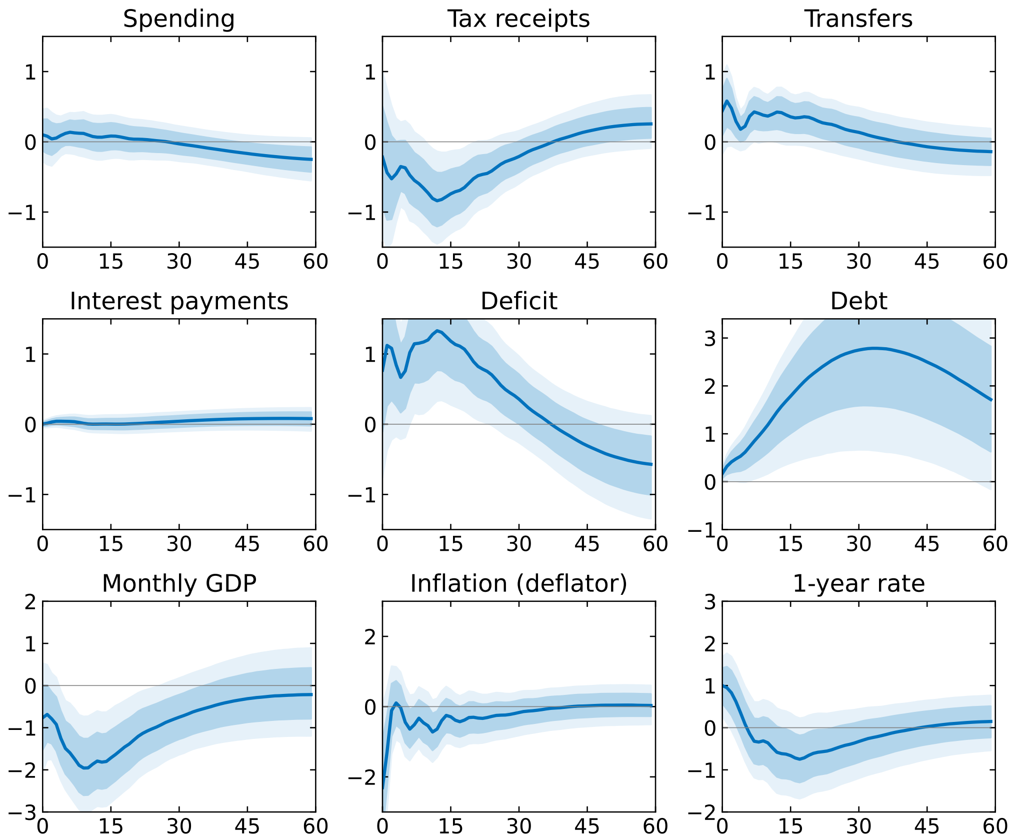
Note: Constrained VAR with the specification of figure 1. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section E.3 for the constrained VAR methodology.

Figure A.9: Constrained VAR (Aruoba–Drechsel)



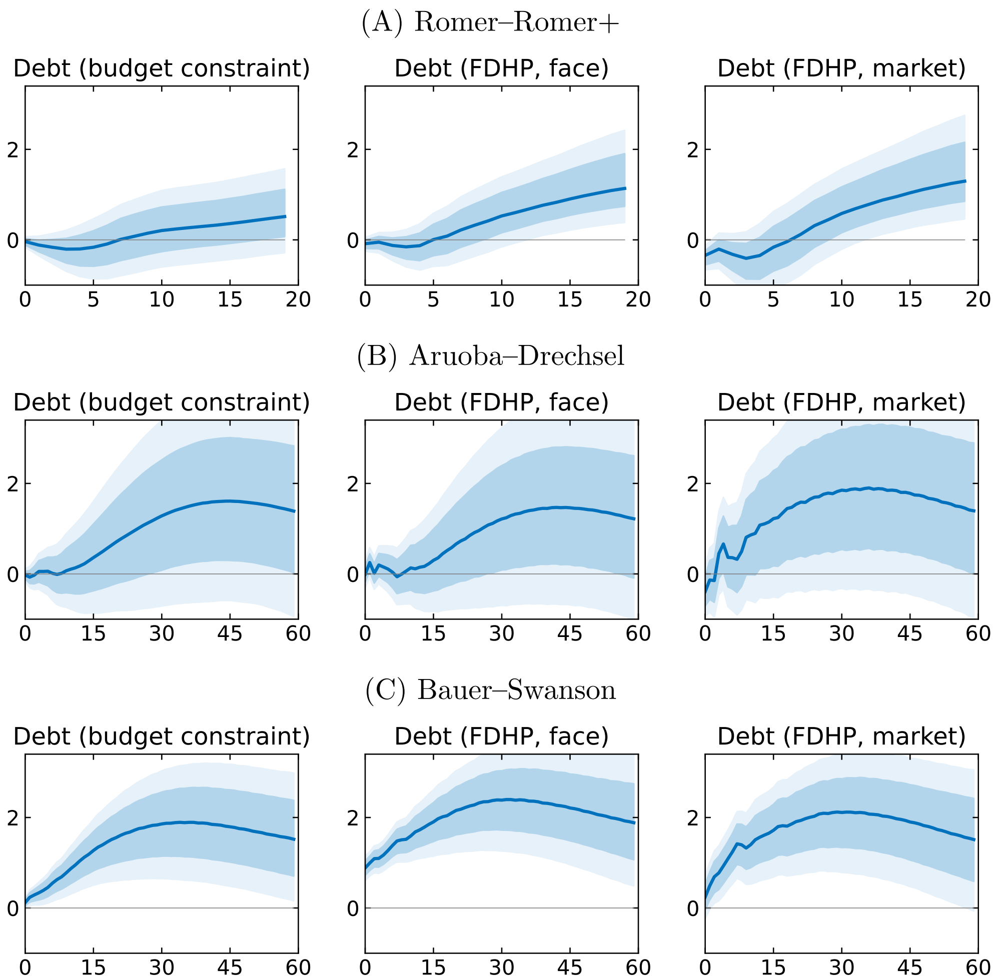
Note: Constrained VAR with the specification of figure A.6. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See section E.3 for the constrained VAR methodology.

Figure A.10: Constrained VAR (Bauer–Swanson)



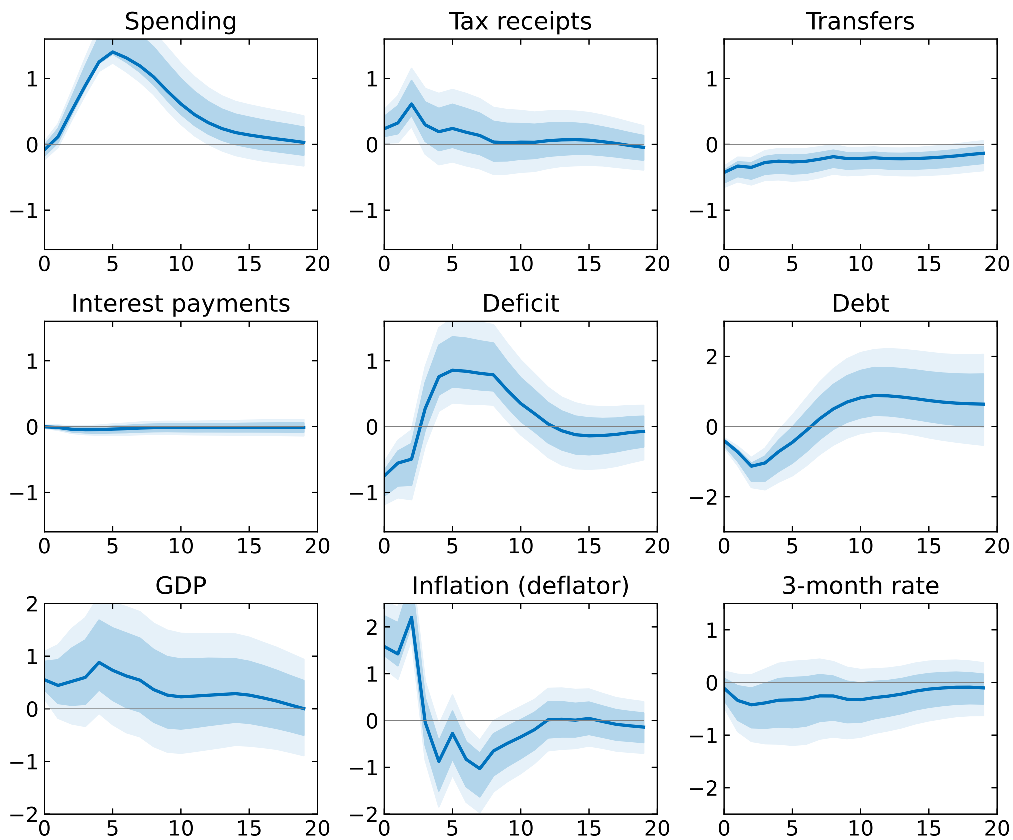
Note: Constrained VAR with the specification of figure A.7. See section E.3 for the constrained VAR methodology. The solid line is the median IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals.

Figure A.11: Debt concepts without timing assumption



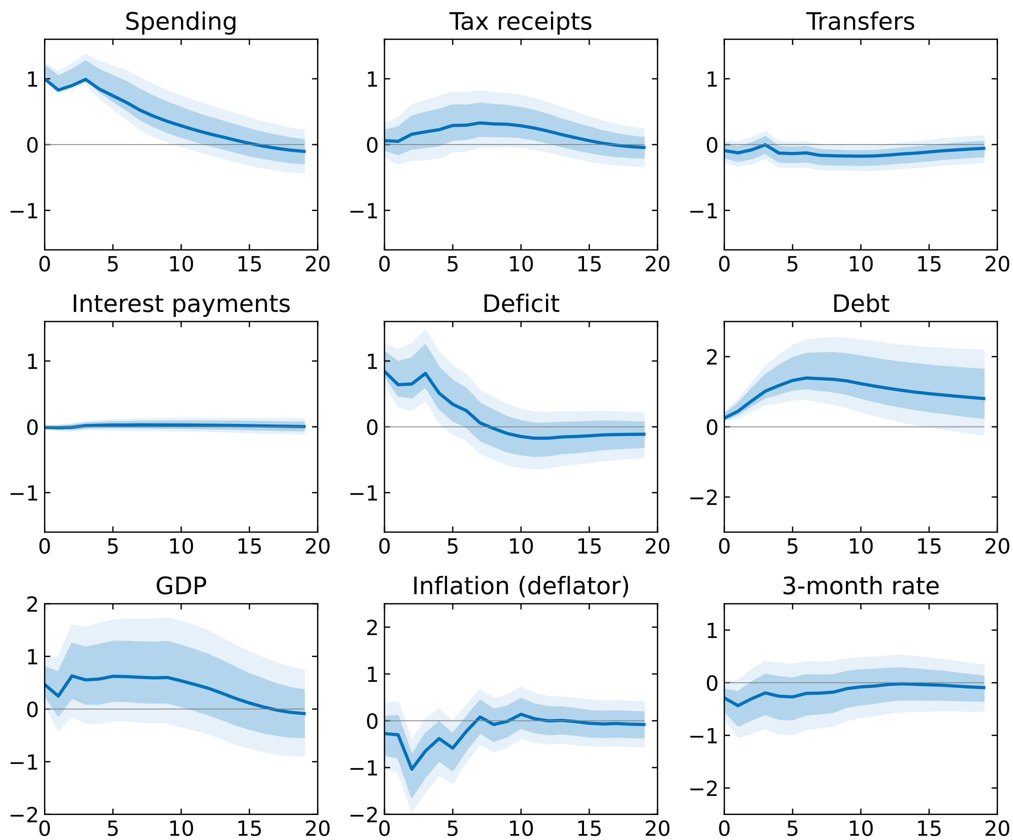
Note: The left-hand side column is a reproduction of the baseline responses shown in figures 1–3. The middle and right-hand side columns contain the response of the face and market values of federal debt held by the public (FDHP) in a VAR with market or face value as the debt concept. See section 3.3.1 for more details. The solid line is the modal (Romer–Romer+)/median (Aruoba–Drechsel, Bauer–Swanson) IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals.

Figure A.12: Response to Ramey spending shock



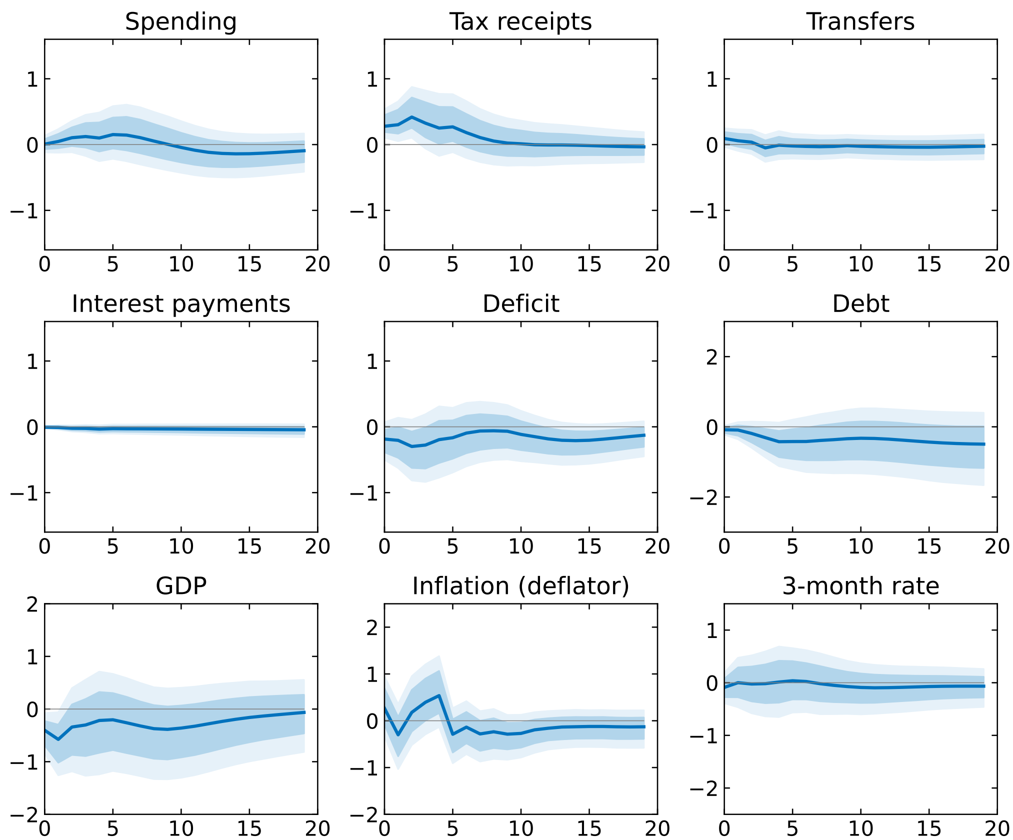
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of spending to the Blanchard–Perroti shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.13: Response to Blanchard–Perotti spending shock



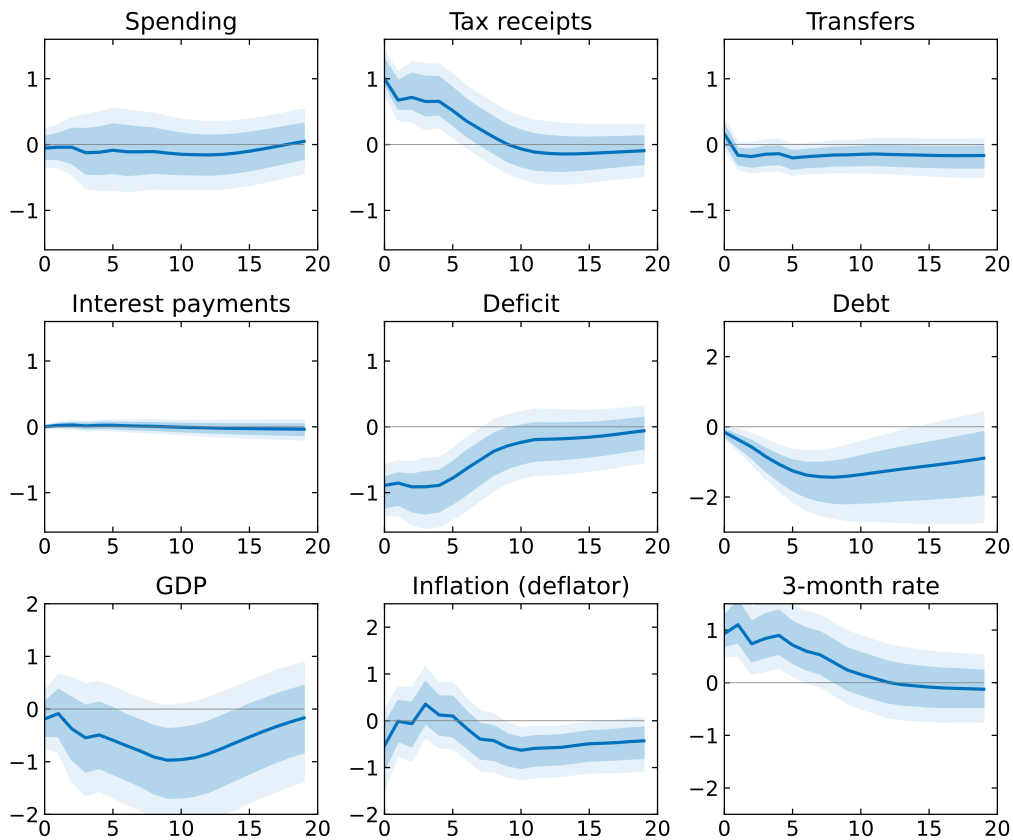
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of spending to the Blanchard–Perroti shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.14: Response to Romer–Romer tax shock



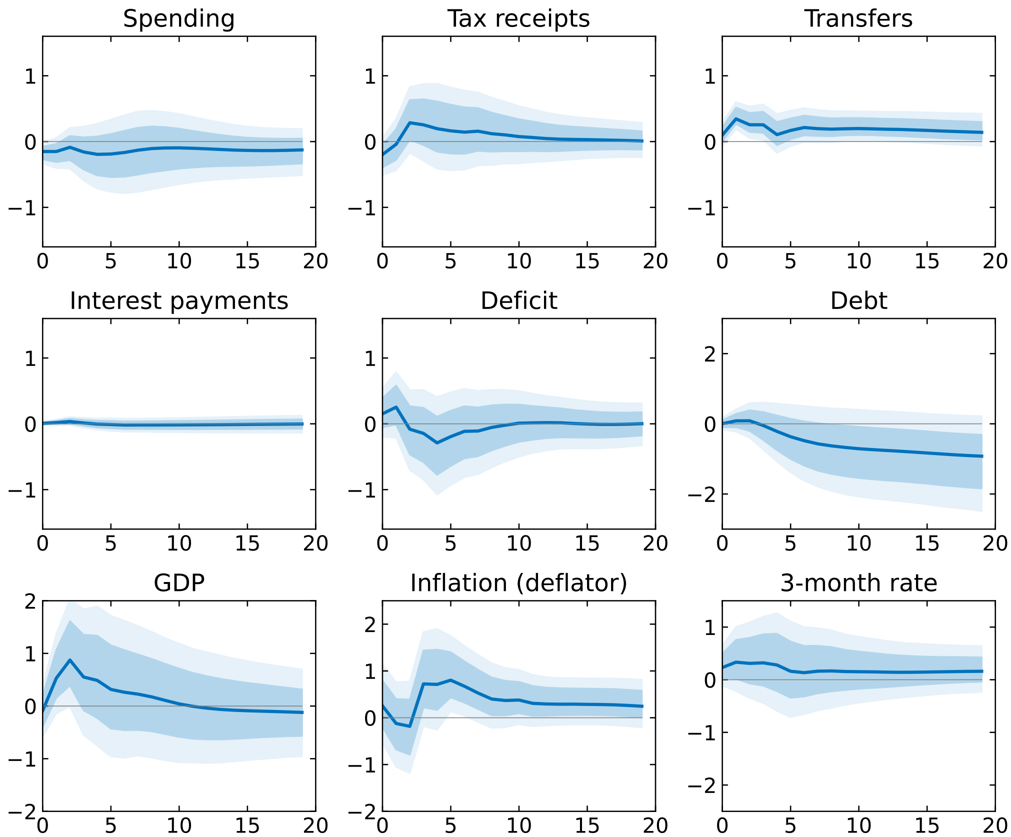
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of taxes to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.15: Response to Caldara–Kamps tax shock



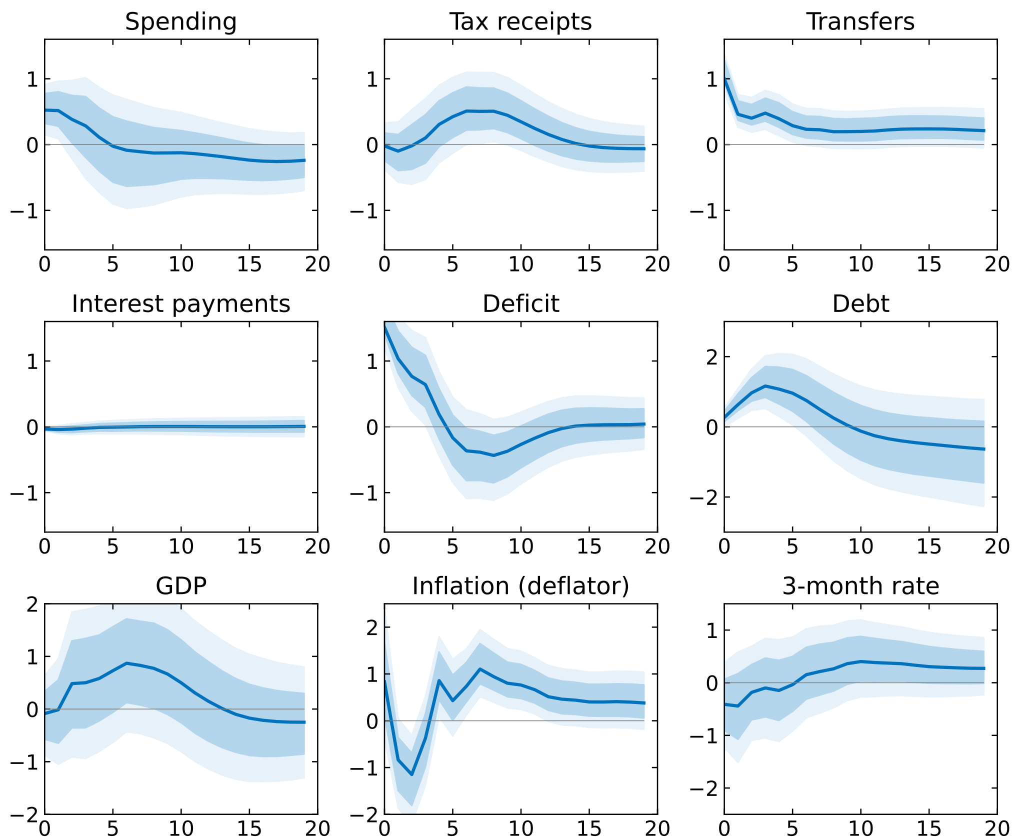
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of taxes to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.16: Response to Romer–Romer transfer shock



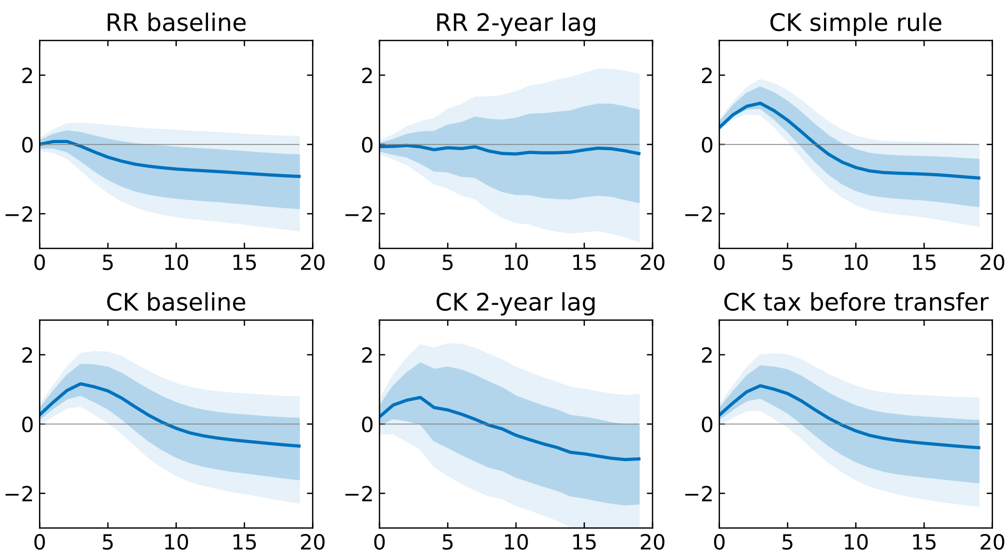
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of transfers to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.17: Response to Caldara–Kamps transfer shock



Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of transfers to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

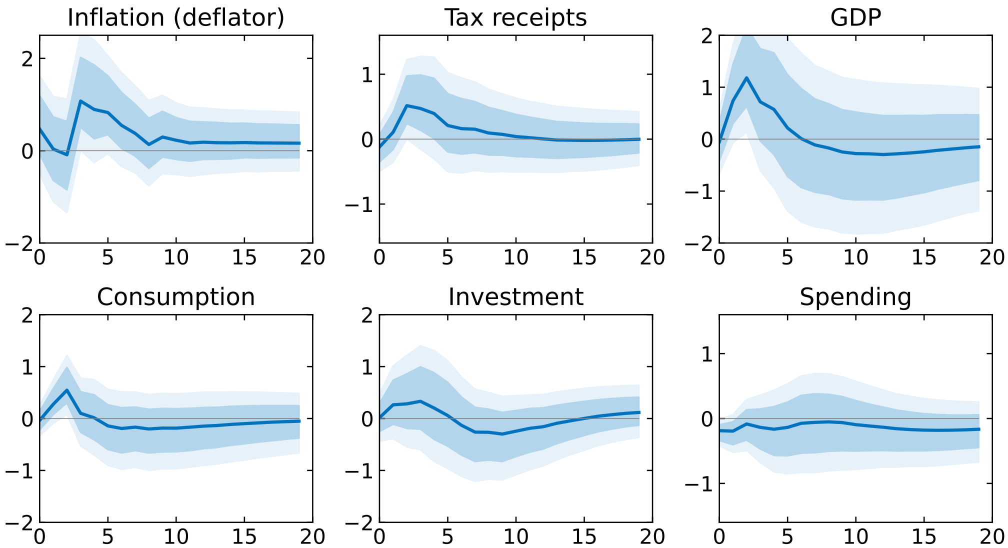
Figure A.18: Debt response to transfer shocks in various specifications



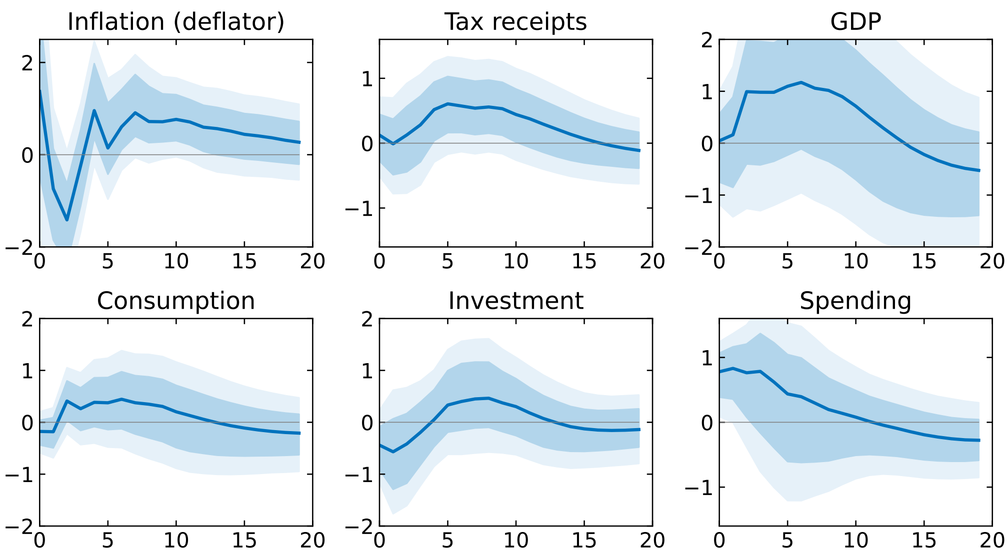
Note: Response of real debt in various specifications. “RR baseline” and “CK baseline” correspond to the Romer–Romer and Caldara–Kamps transfer shocks. They are the same as the response of debt in figures A.16 and A.17. In the second column, we include 2 years of lags in the VAR (instead of 1). In the third column, we use Caldara and Kamps’s simple fiscal rule and order taxes before transfers—the RR response is the same as in the baseline. See appendix I for details on the Caldara–Kamps identification.

Figure A.19: VAR with consumption and investment

(A) Romer–Romer transfer shock

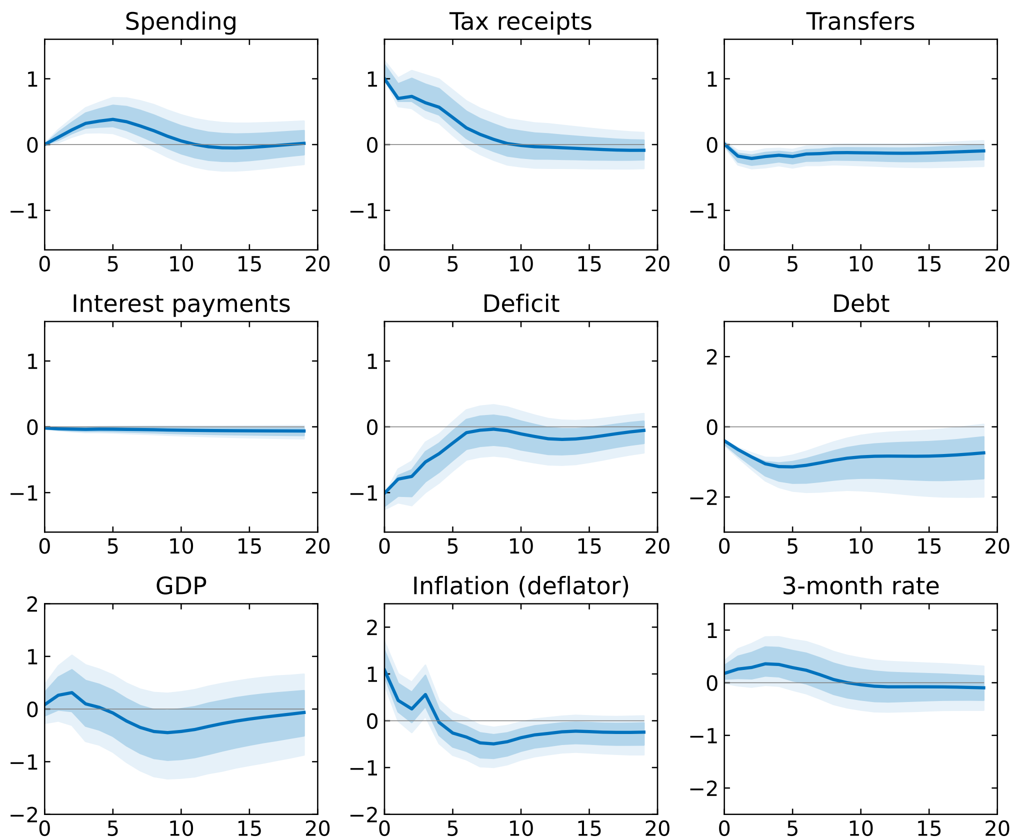


(B) Caldara–Kamps transfer shock



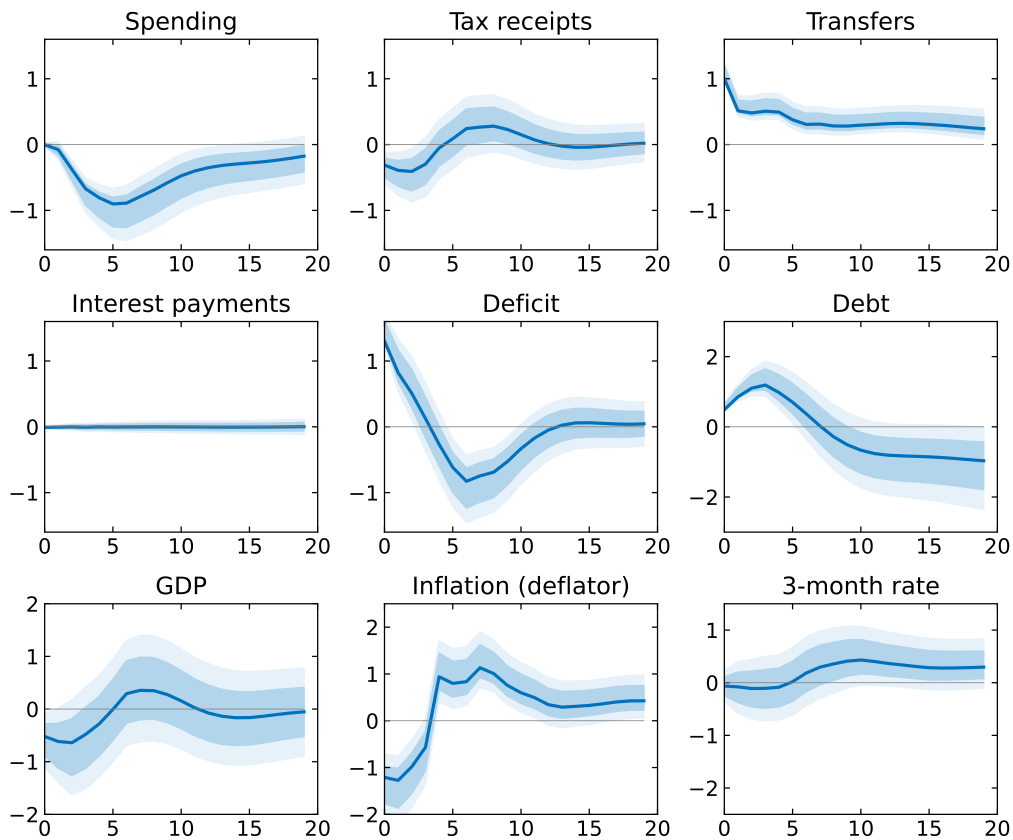
Note: Selected results from the VAR with consumption and investment. This VAR includes the Romer–Romer transfer shock, spending, tax receipts, transfers, interest payments, debt, GDP, consumption, investment (all in real terms), inflation, and the 3-month nominal interest rate.

Figure A.20: Response to Caldara–Kamps tax shock with simple fiscal rule



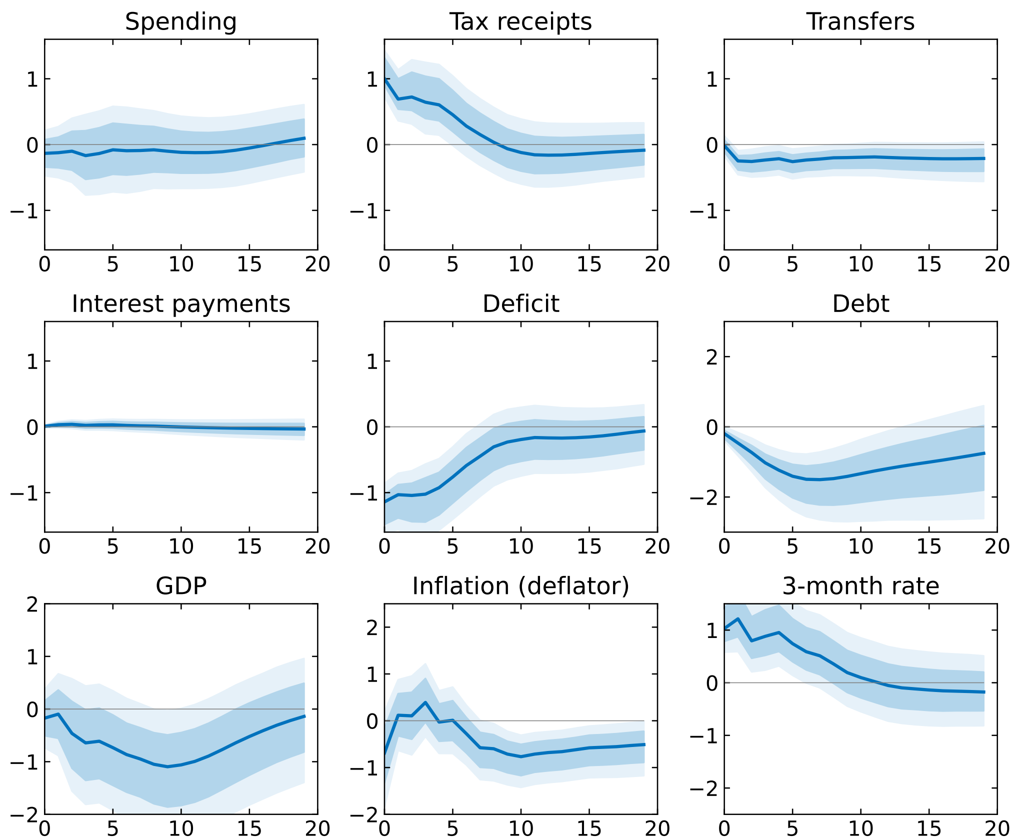
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of taxes to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.21: Response to Caldara–Kamps transfer shock with simple fiscal rule



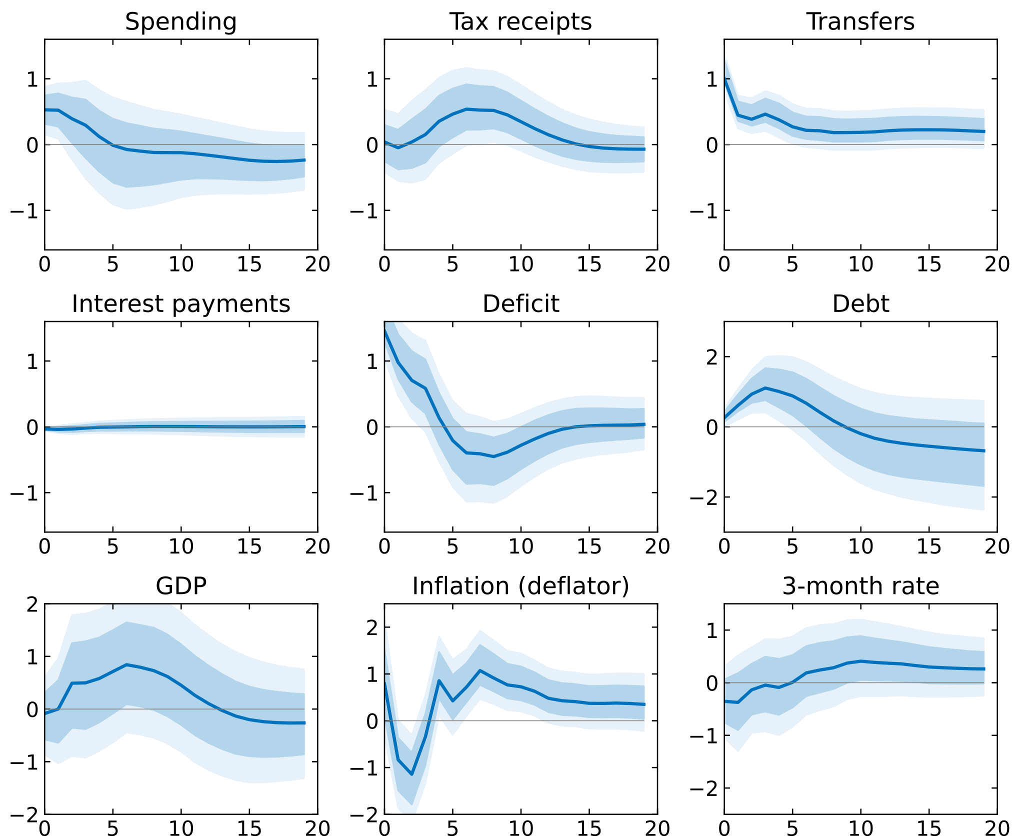
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of transfers to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.22: Response to Caldara–Kamps tax shock with taxes ordered before transfers



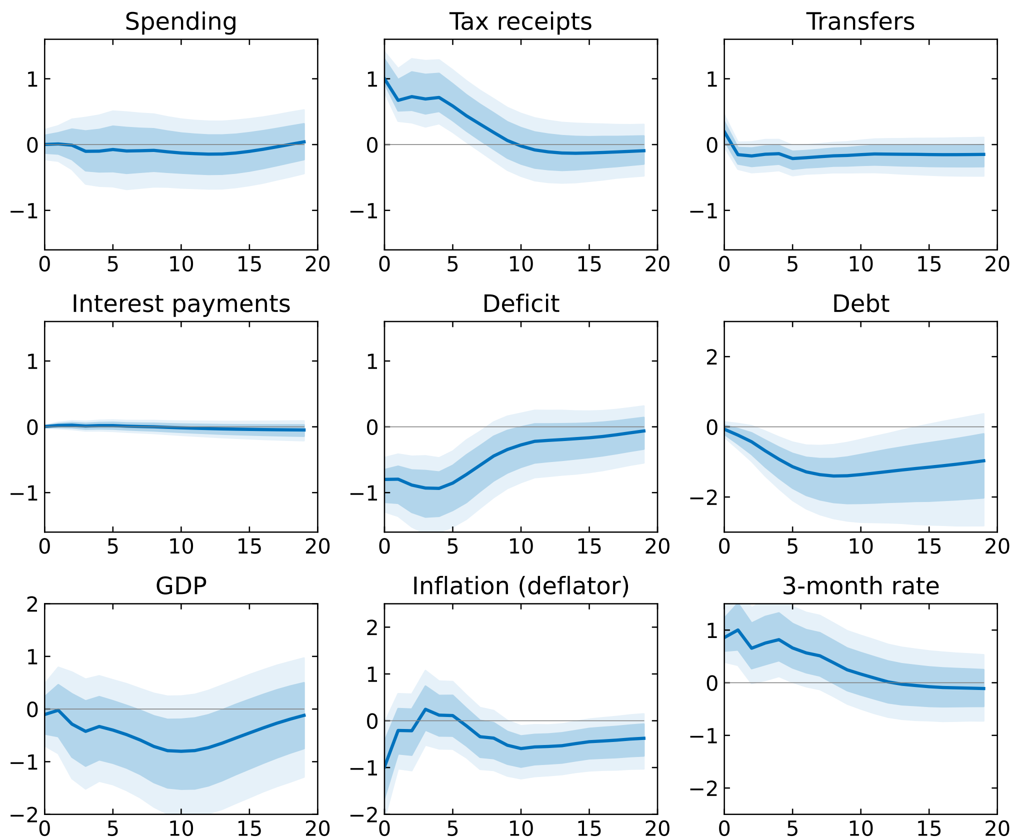
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of taxes to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.23: Response to Caldara–Kamps transfer shock with taxes ordered before transfers



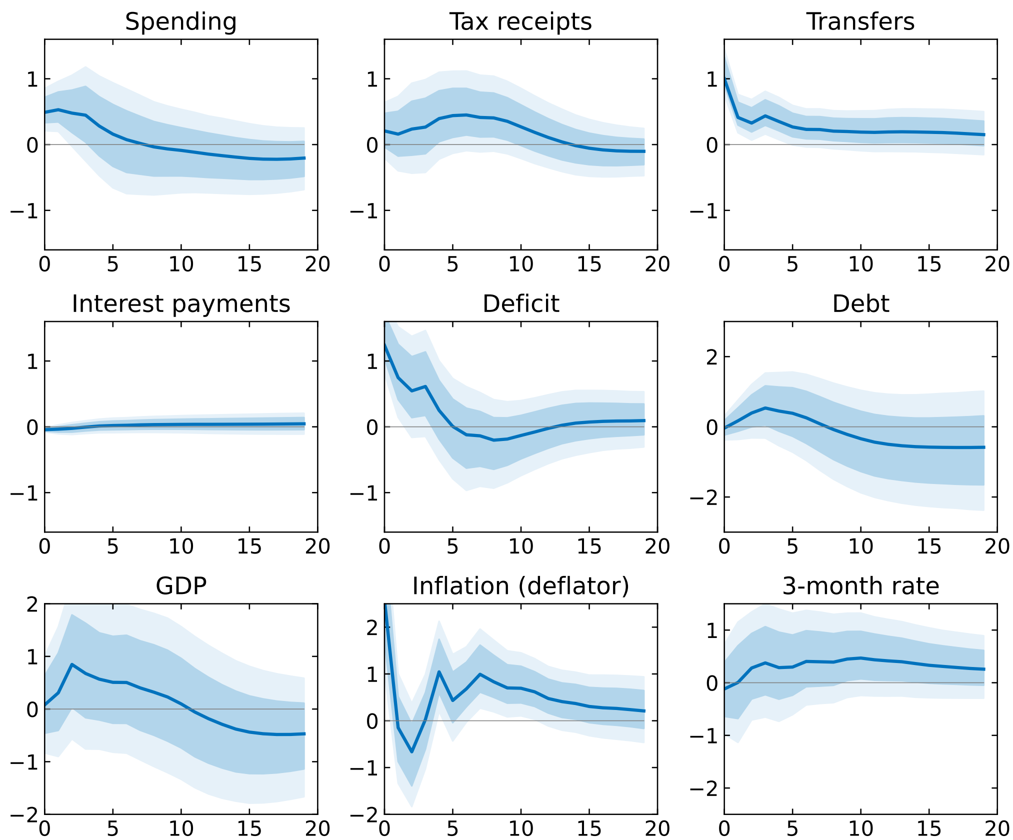
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of transfers to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.24: Response to Caldara–Kamps tax shock, using their vintage as instrument



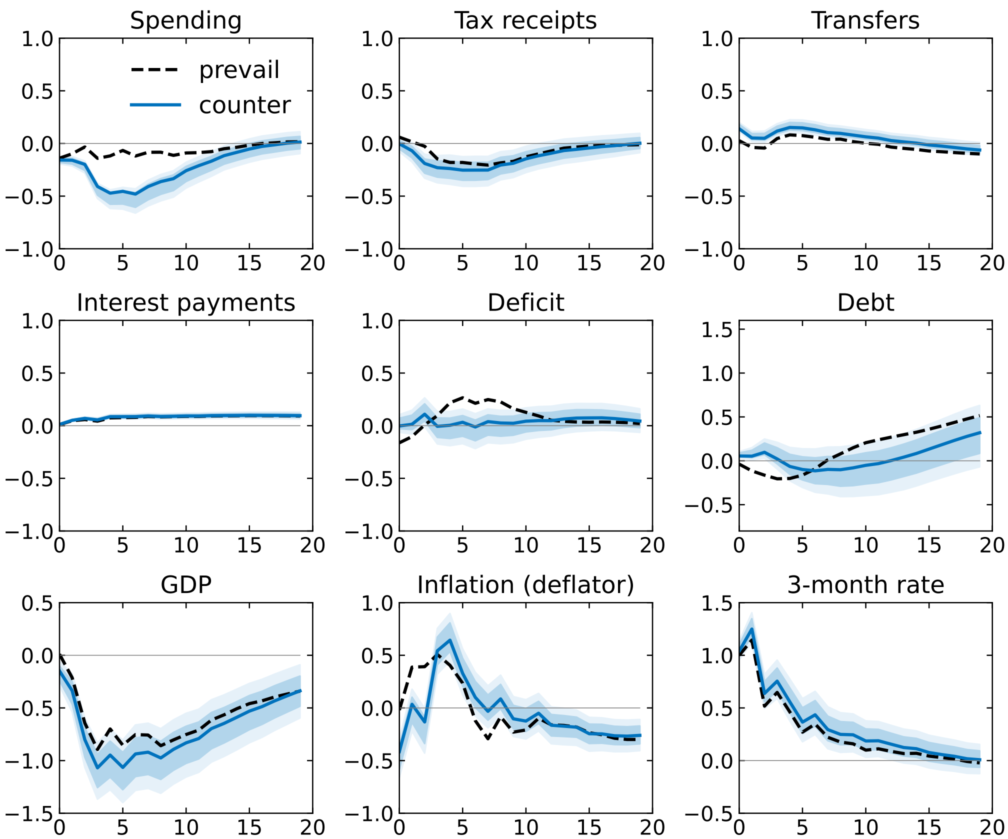
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of taxes to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.25: Response to Caldara–Kamps transfer shock, using their vintage as instrument



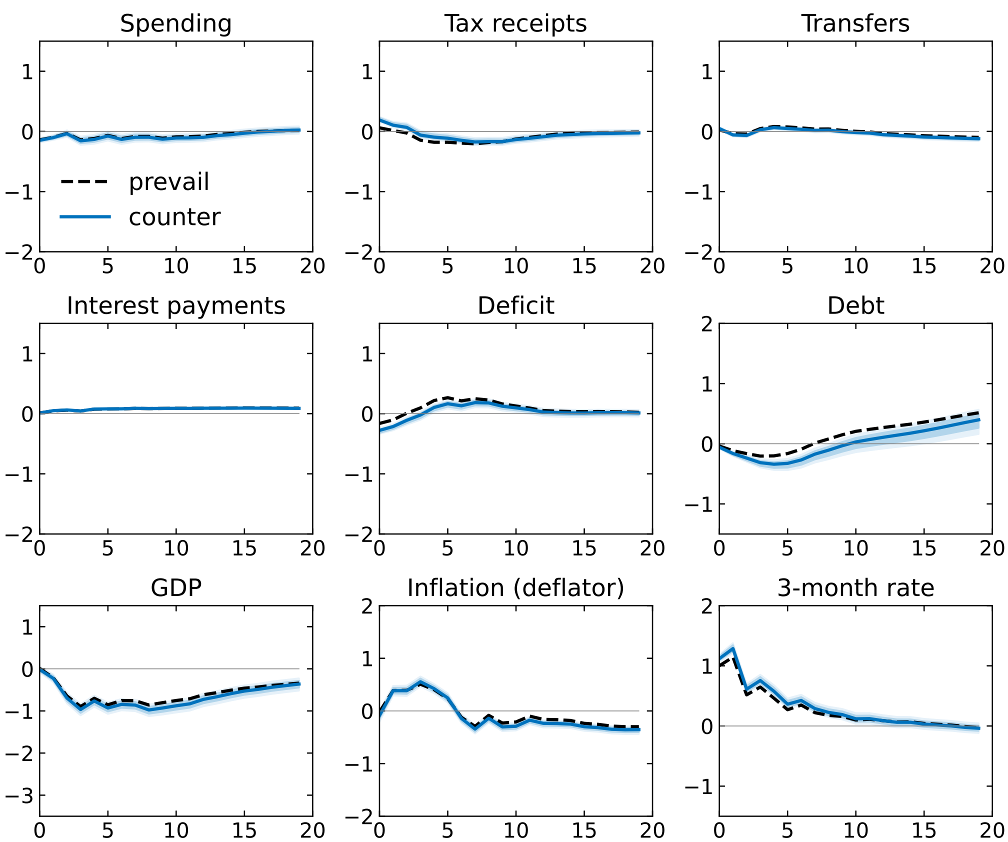
Note: VAR includes spending, tax receipts, transfers, interest payments, a concept of debt, GDP (all in real terms), inflation, and a nominal interest rate. Narrative shocks enter as an internal instrumental variable and high-frequency shocks as an external instrumental variable. The response of the deficit is computed from the responses of the flow fiscal variables. The response of debt is obtained by iterating the linearized budget constraint (4). The IRFs are scaled such that the point estimate of the response of transfers to the Caldara–Kamps shock is 1 at time 0. The solid line is the modal IRF. Shaded areas represent 68% (dark) and 90% (light) credible intervals. See sections 2 and 4.3 for more details on the methodology.

Figure A.26: Counterfactual—deficit stabilization with spending



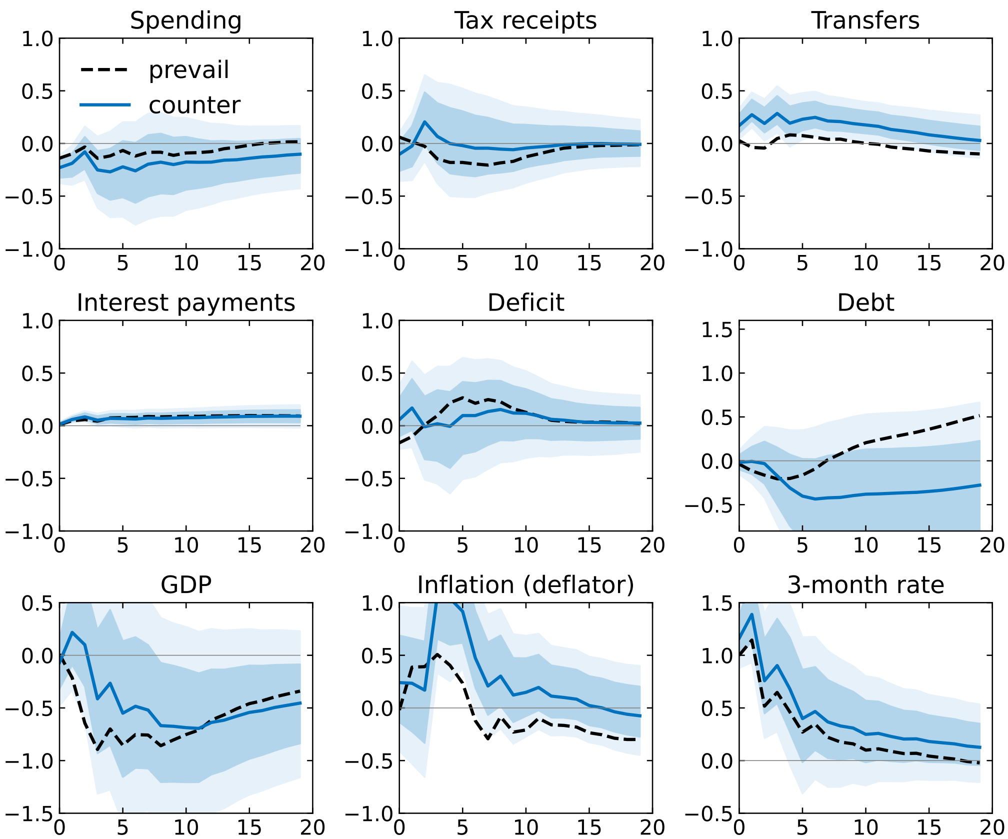
Note: Counterfactual response of the economy to a monetary shock if the government stabilizes deficit through spending. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

Figure A.27: Counterfactual—deficit stabilization with taxes



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes deficit through taxes. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

Figure A.28: Counterfactual—deficit stabilization with transfers



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes deficit through transfers. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1).

B Additional Tables

Table B.1: Federal transfers

	1947	1993	2007	2019
Total (billion dollars)	11	792	1,757	3,107
Social benefits	**73%**	**74%**	**72%**	**75%**
Social security	4%	38%	33%	33%
Medicare	0%	19%	24%	25%
Unemployment insurance	14%	4%	2%	1%
Railroad retirement	2%	1%	1%	0%
Pension benefit guaranty	0%	0%	0%	0%
Veterans' life insurance	3%	0%	0%	0%
Workers' compensation	0%	0%	0%	0%
Military medical insurance	0%	0%	0%	0%
Veterans' benefits	48%	2%	2%	4%
Supplemental Nutrition Assistance Program (SNAP)	0%	3%	2%	2%
Black lung benefits	0%	0%	0%	0%
Supplemental security income	0%	3%	2%	2%
Refundable tax credits	0%	1%	3%	5%
Other	3%	2%	2%	2%
To the rest of the world	0%	1%	1%	1%
Grants-in-aid to state and local governments	**8%**	**20%**	**20%**	**20%**
General public service	n.a.	0%	0%	0%
National defense	n.a.	0%	0%	0%
Public order and safety	n.a.	0%	0%	0%
Economic affairs	n.a.	1%	1%	0%
Housing and community services	n.a.	0%	1%	1%
Health	n.a.	10%	12%	14%
Recreation and culture	n.a.	0%	0%	0%
Education	n.a.	2%	2%	1%
Income security	n.a.	6%	4%	3%
Other transfers to the rest of the world	**17%**	**3%**	**2%**	**2%**
Capital transfer payments	**1%**	**3%**	**5%**	**3%**

Note: Based on NIPA tables 3.2, 3.12U and 3.24U. Detailed data on grants-in-aid is unavailable before 1993.

Table B.2: Main monetary specifications

Figure	Shock	Specific variable(s)	IV	Prior	Sample	Freq.
1, 6.a	Romer–Romer+	Net debt (Financial Accounts)	IIV	Flat	1947–2007	Q
2, 6.b	Aruoba–Drechsel	Monthly GDP, monthly inflation, and federal debt held by the public (face value), excess bond premium	IIV	MN	1959–2007	M
3, 6.c	Bauer–Swanson	Same as figure 2	XIV	MN	1959–2019	M
4.a		Figure 1 + endogenous legislated tax changes as endogenous variable + exogenous legislated tax changes as controls				
4.b–c		Figures 2–3 + endogenous legislated tax changes as endogenous variable				
5.a–c		Figures 1–3 + UI, UR				
7.a–c		Figures 1–3 with debt replaced by lagged federal debt held by the public (face value) and valuation component (market - face value)				
A.1	Jarociński–Karadi	Same as figure 2	XIV	MN	1959–2019	M
A.2	Miranda-Aggrippino–Ricco+	Same as figure 2	XIV	MN	1959–2019	M

Note: All specifications include at least 5 fiscal variables (quarterly spending, tax receipts, transfers, and interest payments from the NIPA tables and a concept of debt), 3 macroeconomic variables (GDP, inflation, and a nominal interest rate), and a quadratic time trend. Abbreviations: IV stands for instrumental variable, IIV for internal instrumental variable, XIV for external instrumental variable, MN for Minnesota prior, Freq. for frequency, Q and M for quarterly and monthly frequency. Cells left blank are the same as in the reference figure (1, 2, or 3).

Table B.3: Monetary policy and fiscal response—1-year average

		Counterfactual		
	Actual	Spending	Taxes	Transfers
	(1)	(2)	(3)	(4)
GDP	-0.44	-0.62	-0.85	-0.25
	(0.33)	(0.13)	(0.47)	(0.21)
	[-0.82,-0.16]	[-0.78,-0.52]	[-1.38,-0.46]	[-0.44,-0.03]
Inflation	0.32	-0.06	0.50	0.27
	(0.27)	(0.11)	(0.36)	(0.17)
	[0.10, 0.63]	[-0.23,-0.01]	[0.17, 0.88]	[0.11, 0.43]
Nominal interest rate	0.83	0.94	0.55	0.84
	(0.22)	(0.09)	(0.33)	(0.16)
	[0.73, 1.17]	[0.86, 1.05]	[0.21, 0.86]	[0.70, 1.02]
Real interest rate	0.40	0.70	-0.09	0.43
	(0.30)	(0.12)	(0.41)	(0.18)
	[0.17, 0.76]	[0.63, 0.87]	[-0.55, 0.26]	[0.26, 0.62]

Note: Average response over the first year to a Romer–Romer+ monetary shock, depending on the fiscal response. Column (1) is the actual response described in section 3. Columns (2–4) are the counterfactual responses under the three scenarios described in section 5.2: debt stabilization through spending, taxes, or transfers. They respectively correspond to figures 11, 12, and 13. The number in parentheses is the standard error. The numbers between brackets are the bounds of the 68% credible interval.

Table B.4: Monetary policy and fiscal response—2-year average

		Counterfactual		
	Actual	Spending	Taxes	Transfers
	(1)	(2)	(3)	(4)
GDP	-0.60	-0.81	-0.87	-0.37
	(0.39)	(0.16)	(0.51)	(0.23)
	[-1.07,-0.30]	[-1.00,-0.70]	[-1.43,-0.44]	[-0.58,-0.12]
Inflation	0.19	0.12	0.26	0.36
	(0.20)	(0.09)	(0.25)	(0.12)
	[0.03, 0.41]	[0.02, 0.19]	[0.03, 0.52]	[0.26, 0.50]
Nominal interest rate	0.58	0.68	0.36	0.62
	(0.26)	(0.11)	(0.35)	(0.17)
	[0.41, 0.92]	[0.60, 0.81]	[-0.00, 0.69]	[0.47, 0.80]
Real interest rate	0.40	0.49	0.18	0.26
	(0.23)	(0.10)	(0.31)	(0.15)
	[0.23, 0.68]	[0.40, 0.60]	[-0.14, 0.47]	[0.11, 0.39]

Note: Average response over the first 2 years to a Romer–Romer+ monetary shock, depending on the fiscal response. Column (1) is the actual response described in section 3. Columns (2–4) are the counterfactual responses under the three scenarios described in section 5.2: debt stabilization through spending, taxes, or transfers. They respectively correspond to figures 11, 12, and 13. The number in parentheses is the standard error. The numbers between brackets are the bounds of the 68% credible interval.

Table B.5: Monetary policy and fiscal response—4-year average

	Actual	Counterfactual Spending	Counterfactual Taxes	Counterfactual Transfers
	(1)	(2)	(3)	(4)
GDP	-0.63	-0.79	-0.84	-0.48
	(0.35)	(0.15)	(0.46)	(0.23)
	[-1.06,-0.39]	[-0.96,-0.67]	[-1.32,-0.46]	[-0.69,-0.26]
Inflation	0.01	0.01	0.00	0.22
	(0.14)	(0.06)	(0.18)	(0.08)
	[-0.12, 0.15]	[-0.06, 0.06]	[-0.17, 0.17]	[0.16, 0.32]
Nominal interest rate	0.34	0.43	0.17	0.42
	(0.22)	(0.09)	(0.28)	(0.14)
	[0.19, 0.61]	[0.36, 0.53]	[-0.11, 0.42]	[0.30, 0.57]
Real interest rate	0.35	0.41	0.21	0.23
	(0.22)	(0.09)	(0.29)	(0.14)
	[0.20, 0.61]	[0.34, 0.51]	[-0.07, 0.48]	[0.08, 0.35]

Note: Average response over the first 4 years to a Romer–Romer+ monetary shock, depending on the fiscal response. Column (1) is the actual response described in section 3. Columns (2–4) are the counterfactual responses under the three scenarios described in section 5.2: debt stabilization through spending, taxes, or transfers. They respectively correspond to figures 11, 12, and 13. The number in parentheses is the standard error. The numbers between brackets are the bounds of the 68% credible interval.

Table B.6: Monetary policy and fiscal response—5-year average

	Actual	Counterfactual Spending	Counterfactual Taxes	Counterfactual Transfers
	(1)	(2)	(3)	(4)
GDP	-0.58	-0.71	-0.75	-0.48
	(0.34)	(0.14)	(0.43)	(0.22)
	[-0.98,-0.37]	[-0.87,-0.60]	[-1.20,-0.41]	[-0.67,-0.27]
Inflation	-0.05	-0.05	-0.06	0.15
	(0.13)	(0.06)	(0.17)	(0.08)
	[-0.18, 0.07]	[-0.10, 0.01]	[-0.23, 0.09]	[0.10, 0.26]
Nominal interest rate	0.27	0.35	0.13	0.36
	(0.21)	(0.08)	(0.25)	(0.13)
	[0.13, 0.51]	[0.29, 0.44]	[-0.12, 0.34]	[0.25, 0.50]
Real interest rate	0.34	0.39	0.23	0.23
	(0.21)	(0.09)	(0.28)	(0.14)
	[0.20, 0.59]	[0.31, 0.48]	[-0.03, 0.48]	[0.09, 0.35]

Note: Average response over the first 5 years to a Romer–Romer+ monetary shock, depending on the fiscal response. Column (1) is the actual response described in section 3. Columns (2–4) are the counterfactual responses under the three scenarios described in section 5.2: debt stabilization through spending, taxes, or transfers. They respectively correspond to figures 11, 12, and 13. The number in parentheses is the standard error. The numbers between brackets are the bounds of the 68% credible interval.

C Government Accounting

C.1 Sources of Fiscal Data

The first potential source of fiscal data in the United States is national accounting. Flow variables can be found in the National Income and Product Accounts (NIPA) tables, which are published by the Bureau of Economic Analysis (BEA). Stock variables are available in the Financial Accounts of the United States (Z.1), formerly known as the Flow of Funds Accounts, which are published by the Federal Reserve. As we will explain in the next section, the NIPA tables and the Financial Accounts deliver a coherent view of the fiscal situation: taken together, the two sources are consistent with the budget constraint of the government. The national accounts have two advantages: (i) they allow us to properly distinguish government spending (consumption and investment) from transfers; (ii) they are consistent with macroeconomic aggregates, mainly GDP.

The second potential source is the US Treasury itself. Its various publications (*Monthly Treasury Statement*, *Monthly Statement of the Public Debt*, *Treasury Bulletin*) contain information on federal outlays, tax receipts, and debt. Payne et al. (2025) used these sources to construct series for federal debt valued at par (face value) or market price (market value) since independence.[1] The main drawback of these sources is that they do not systematically distinguish spending from transfers—both are lumped together as outlays. Because of this disadvantage, we rely primarily on national accounts. The Treasury data, however, is available at higher frequency (monthly) than the national accounts (quarterly), a feature that we use in our mixed-frequency VAR (section 2.3).

C.2 Budget Constraint: NIPA Tables and Financial Accounts

In nominal terms, the law of motion of government debt is given by the following identity:

$$D_t = D_{t-1} + GS_t - TX_t + TR_t + INT_t, \tag{C.1}$$

where D_t is debt, GS_t government spending, TX_t taxes, TR_t transfers, and INT_t interest payments. In practice, this is not perfectly accurate, as the NIPA tables, which are the source for the flow variables, and the Financial Accounts, which is the source for government debt, treat minor items differently.[2] Equation (C.2) is, however, a very good approximation, as figure C.1 (panel A, left) shows. We use as initial value that of debt at the beginning of the sample and construct a series by iterating equation (C.1): the new series is almost

[1] See Hall et al. (2018) for a detailed explanation of the data collection.

[2] See Federal Reserve Board (2025, pp. 10–11).

indistinguishable from net debt (total liabilities minus total financial assets) according to the Financial Accounts.

Equation (C.1) can be rewritten with deflated and de-trended quantities:

$$d_t = \frac{d_{t-1}}{(1+\pi_t)(1+g_t)} + gs_t - tx_t + tr_t + int_t, \tag{C.2}$$

where π_t is the inflation rate, g_t is the growth rate of the trend, and lowercase letters denote the deflated and de-trended variables. Equation (C.2) is exactly implied by equation (C.1), which, as we have seen, is a very good approximation of reality. Equation (C.2), however, is not linear anymore. Linearizing in level around a situation where $d_{t-1} = d^*$, $\pi_t = \pi^*$, and $g_t = g^*$, we obtain

$$d_t \approx c_d d_{t-1} - c_\pi(\pi_t - \pi^*) - c_g(g_t - g^*) + gs_t - tx_t + tr_t + int_t, \tag{C.3}$$

where

$$c_d = \frac{1}{(1+\pi^*)(1+g^*)}, \quad c_\pi = \frac{d^*}{(1+\pi^*)^2(1+g^*)}, \quad \text{and} \quad c_g = \frac{d^*}{(1+\pi^*)(1+g^*)^2}.$$

We repeat the comparison of debt from the Flow of Funds with those implied by equations (C.2–C.3). For equation (C.3), we use the 1947–2019 averages of debt, inflation, and the growth rate of trend GDP:

$$d^* \approx 0.58, \quad \pi^* \approx (1+3.1\%)^{1/4} - 1 \approx 0.77\%, \quad g^* \approx (1+3.2\%)^{1/4} - 1 \approx 0.79\%,$$
$$c_d \approx 0.98, \quad c_\pi \approx 0.58, \quad \text{and} \quad c_g \approx 0.58.$$

Figure C.1 (panel A, right) makes clear that even the linearized equation is a very accurate approximation of the debt to trend GDP ratio according to the Financial Accounts.

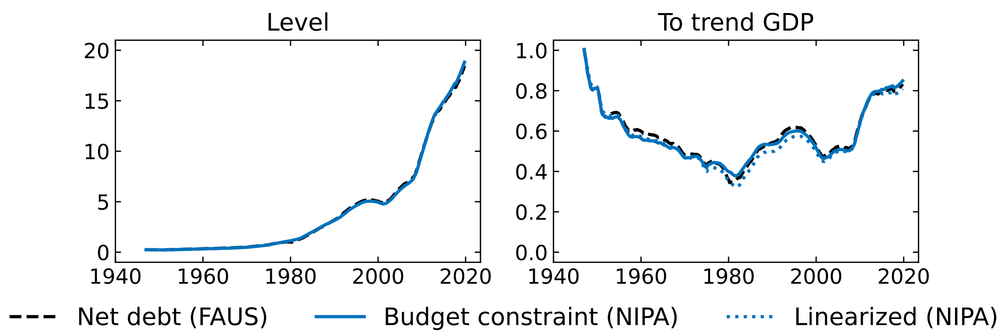
C.3 Debt Concepts

Our preferred concept of debt is net debt from the Financial Accounts of the United States (Z.1), which we define as total liabilities minus total financial assets. We deduct financial assets because they are the counterpart of a financial transaction, not of real economic activity. As we showed in the previous section, this concept of debt is consistent with the fiscal variables in the NIPA tables.

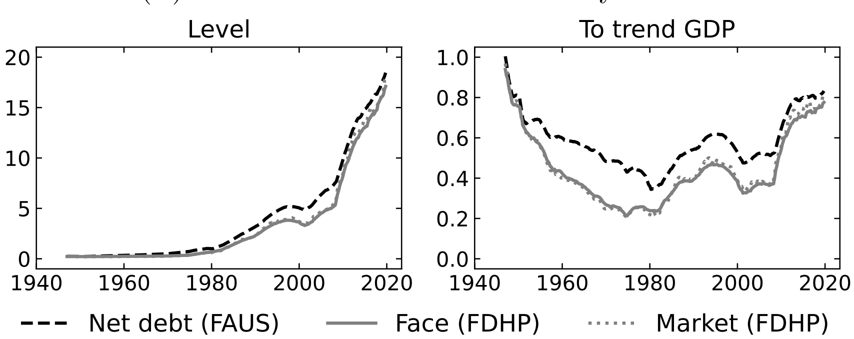
Another possible concept is debt held by the public (FDHP) as measured by the US Treasury, which is the sum of the value of all marketable and nonmarketable public debt

Figure C.1: Debt concepts

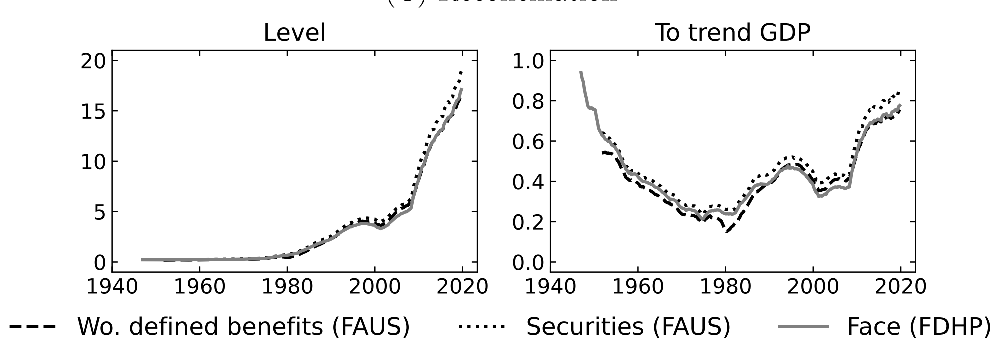
(A) National Accounts



(B) Financial Accounts and Treasury sources



(C) Reconciliation



Note: Various debt concept compared to the Flow of Funds. FAUS stands for Financial Accounts of the United States, FDHP for federal debt held by the public. The left-hand side plots ("level") represent nominal amounts in trillion dollars; the right-hand side plots ("to trend GDP") are ratios to the trend of real GDP. "Budget constraint" and "linearized" are obtained by iterating in equations (C.1–C.3) with data from the NIPA tables. "Wo. defined benefits" is debt according to the Financial Accounts minus the claims of defined benefit retirement funds of federal government employees. "Securities" is the sum of total marketable and nonmarketable securities according to the FAUS. Face and market values are the face and market values of the federal debt held by the public, abbreviated FDHP (Payne et al., 2025).

securities net of the holdings of government agencies and trust funds. The value of these securities can be expressed at par (face value) or at their market price (market value). The change in the face value is about equal to the sum of on- and off-budget deficits.[3] As we show in figure C.1 (panel B), face and market values imperfectly track our preferred concept of debt, hence national accounts, although they follow a similar trend. The discrepancy is particularly large from the 1960s to the 2000s.

To understand the differences between the two concepts, we consider the various components of financial assets and liabilities that underlie the Financial Accounts (table C.1). First, financial assets are much smaller than liabilities (less than 20%). Second, most of liabilities (around 90%) are marketable securities, nonmarketable securities, or the defined benefits of the pension funds of federal government employees. The first two components roughly correspond to debt held by the public. We check in figure C.1 (panel C) that the face value is very close to either our preferred concept of debt minus defined benefits (dashed black line) or the sum of marketable and nonmarketable securities according to the Financial Accounts (black dotted line).

We note that neither concept corresponds to "total public debt". Total public debt includes all government holdings, even those of government agencies and trust funds and the Federal Reserve. Net debt (FAUS) and federal debt held by the public do not include the holdings of government agencies and trust funds, but they include those of the Federal Reserve. Since we focus on monetary-fiscal interactions, the appropriate level of consolidation is the fiscal side of the federal government, which is why we include the holdings of the Federal Reserve in our debt concepts, but not other intragovernmental holdings.

C.4 Budget Constraint: Steady State

To express equation (C.2) in steady state, we introduce the effective nominal interest rate: $i_t = INT_t/D_{t-1} = int_t/d_{t-1} \times (1+\pi_t)(1+g_t)$. This nominal interest rate is tied to the ex-post effective real interest rate by the Fisher equation: $1 + i_t = (1+r_t)(1+\pi_t)$. So, equation (C.2) becomes

$$d_t = \frac{1+r_t}{1+g_t} d_{t-1} + gs_t - tx_t + tr_t. \tag{C.4}$$

[3] The sum of the deficits is equal to the change in public and agency securities minus investments of governments accounts and the change in cash and monetary assets and other adjustments. See table FF0-1 in the *Treasury Bulletin*. The off-budget deficit is the sum of the deficits of the Federal Old-Age and Survivors Insurance Trust Fund (Social Security retirement), the Federal Disability Insurance Trust Fund (Social Security disability), and the Postal Service Fund.

Table C.1: Federal government balance sheet

	1947	1970	2007	2019
Total financial assets	**52**	**97**	**810**	**3,134**
Monetary gold and SDRs	44%	12%	28%	14%
Checkable deposits	7%	10%	7%	13%
Total time and savings deposits	0%	0%	0%	0%
Other deposits	3%	6%	2%	1%
Other reserves and claims	0%	0%	0%	0%
Debt securities	0%	0%	1%	0%
Loans	24%	60%	37%	53%
Corporate equities	0%	0%	0%	1%
Other equity	2%	2%	6%	2%
Trade receivables	0%	7%	7%	2%
Total taxes receivable	20%	2%	10%	12%
Total liabilities	**291**	**636**	**8,079**	**21,590**
Special drawing rights (SDRs) allocations	0%	0%	0%	0%
Treasury currency	2%	1%	0%	0%
Postal savings system deposits	1%	0%	0%	0%
Debt securities	60%	39%	56%	77%
Including: marketable securities	*60%*	*37%*	*56%*	*77%*
Loans	17%	13%	19%	11%
Including: nonmarketable securities	*17%*	*13%*	*19%*	*11%*
Life insurance reserves	2%	1%	1%	0%
Trade payables	0%	1%	3%	2%
Total miscellaneous liabilities	18%	45%	21%	9%
Including: defined benefits	*10%*	*44%*	*19%*	*8%*
Net debt (preferred concept)	**239**	**539**	**7,269**	**18,456**

Source: Financial Accounts of the United States (Z.1), table L106. Note: Nominal amounts are in billion dollars. Net debt is total liabilities minus total financial assets.

Expressing this relationship in steady state, we obtain equation (7), which we reproduce here for convenience:

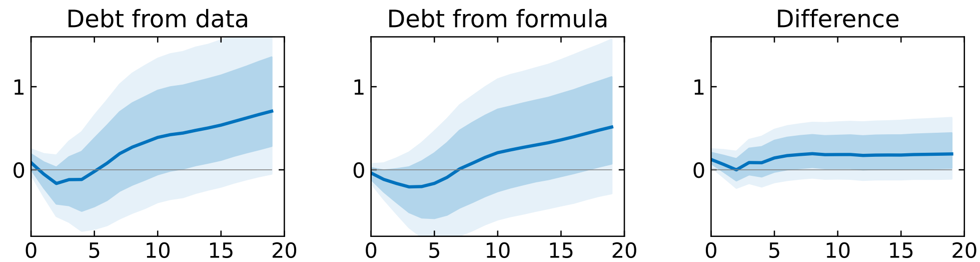
$$\frac{r^* - g^*}{1 + g^*} d^* = tx^* - gs^* - tr^*. \tag{7}$$

D VAR and Budget Constraint

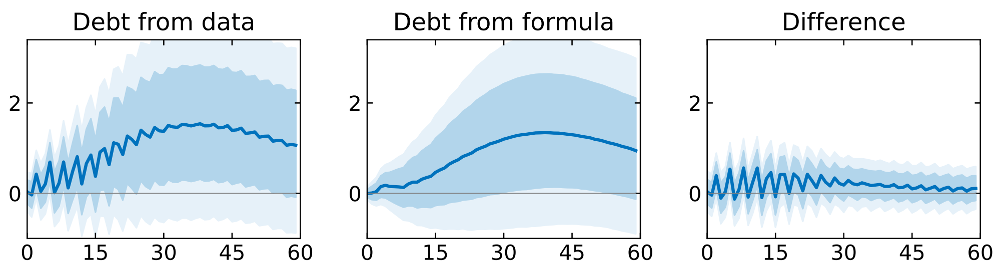
In the body of the paper, we construct the response of debt by iterating the budget constraint (4), using sample averages for the value of the coefficients. In figure D.1, we compare two responses. On the left, we show the response of debt from the Financial Accounts, in the middle that of debt from the formula, and on the right the difference between the two. Debt from the data responds even more strongly than debt from the formula in the Romer–Romer+ case.

Figure D.1: Financial Accounts and Iterated Budget Constraint

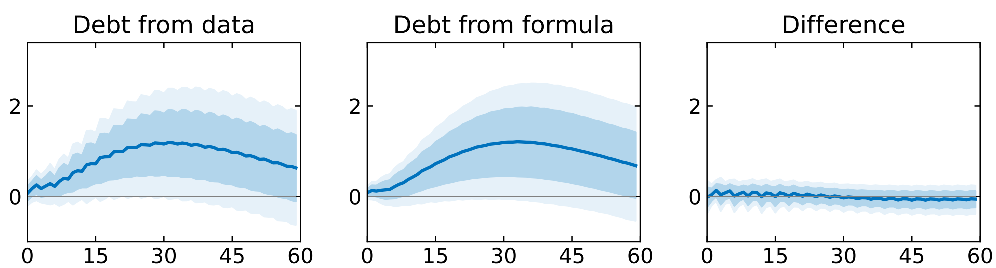
(A) Romer–Romer+



(B) Aruoba–Drechsel



(C) Bauer–Swanson



Note: The VARs underlying panels A, B, and C are those underlying figures 1, A.6, A.7 respectively. The left-hand side chart is the response of debt as it features in the VAR—it is the same as the debt chart of figures 1, A.6, A.7. The middle chart is the response of debt deduced from equation (4), using the sample averages for d^*, π^*, g^*. The right-hand side chart is the difference between the first and second charts. Shaded areas represent 68% (dark) and 90% (light) credible intervals. Panels B–C should be taken with caution since the effective sample size is low (table E.1).

We can explicitly incorporate the government budget constraint if we include lagged debt in our VAR and restrict some coefficients in the debt equation. To understand this approach, consider the linearized budget constraint (C.3), which we reproduce here for convenience:

$$d_t \approx c_d d_{t-1} - c_\pi(\pi_t - \pi^*) - c_g(g_t - g^*) + \underbrace{gs_t + tr_t - tx_t + int_t}_{\text{deficit}}, \tag{C.3}$$

where d_t, gs_t, tr_t, tx_t, and int_t are debt, spending, transfers, tax receipts, and interest payments (all relative to trend GDP), π_t is inflation, g_t is the deterministic growth rate of trend GDP, and c_d, c_π, c_g are constant numbers. If contemporaneous debt (d_t) is in the data vector (y_t), there is no simple way to map equation (C.3) into a line of the reduced-form VAR model, equation (5), because d_t appears on the right-hand side of the latter. If we use lagged debt instead, the problem disappears since d_t is in y_{t+1} and the other variables are in y_t. We can then restrict the coefficients on endogenous variables other than d_{t-1} and π_t: we set the coefficients on gs_t, tr_t, int_t to 1, that on tx_t to -1, and those on the other endogenous variables to 0.[4] g_t, the deterministic growth rate of trend GDP, is not part of our variables but, in practice, it is almost exactly an affine function of time.[5] So, we include a linear time trend in the debt equation to replace it. In summary, the equation for debt becomes

$$d_t = b^d_d d_{t-1} + b^d_\pi \pi_t + gs_t + tr_t - tx_t + int_t + b^d_t t + \bar{b}^d + u^d_t. \tag{D.1}$$

$\bar{b}^d$, b^d_t, b^d_d, b^d_π are estimated within our VAR. When we identify the structural shocks, we do not allow d_t (which is determined at time t) to respond to other components of y_{t+1} (which are determined at $t+1$)—we implement this by ordering it first. The approximation error at time t (u^d_t) is orthogonal to structural innovations in other endogenous variables at time $t+1$.

We show the results for all variables and our 3 shocks in figures A.8–A.10. Note that in these constrained VARs, we use quarterly debt from the Financial Accounts, as in Figures D.1, 1, A.6, and A.7. For Romer–Romer+, this exercise delivers results that are very similar to figure 1. Compared to figure D.1, the response of debt is closer to the response predicted by the formula than the one from the data. For Aruoba–Drechsel and Bauer–Swanson, the results are also close to figures 2–3.

[4] Antolín-Díaz et al. (2025) face a similar issue to enforce the Campbell and Shiller (1988) decomposition. Instead of including a lagged variable, they impose restrictions on Σ as well as on the matrices B_l. Drawing from a restricted posterior for Σ requires a specially designed importance sampling algorithm. Since we include lagged debt in our VAR, we can impose the budget constraint without any restriction on Σ, thus circumventing this difficulty. Drawing from a restricted posterior for the matrices B_l only entails a slight modification of the usual posterior distributions (propositions E.1 and E.2).

[5] In our sample, a regression of g_t on a constant and a linear time trend yields an R^2 of 1 up to the third digit.

E VAR: Full Description

E.1 VAR: Notations

The reduced-form VAR model is described in equation (5), which we reproduce here for convenience:

$$y_{t+1} = \sum_{l=1}^{L} B(l)' y_{t+1-l} + B_c' c_{t+1} + u_{t+1}, \quad u_{t+1} \sim \mathcal{N}(0, \Sigma), \tag{5}$$

where y_{t+1} is the vector of endogenous variables at time $t+1$, c_{t+1} the exogenous controls, and u_{t+1} the error term. The $B(l)$, $1 \leq l \leq L$, are (n, n) matrices where n is the number of endogenous variables. We can rewrite equation (5) by gathering the lagged values of the endogenous variables and the controls:

$$y_{t+1} = B' x_{t+1} + u_{t+1}, \quad u_{t+1} \sim \mathcal{N}(0, \Sigma),$$

where $x_{t+1} = \left(y_t', y_{t-1}', \ldots, y_{t+1-L}, c_{t+1}\right)'$ and $B = (B(1)', B(2)', \ldots, B(L)', B_c')'$. Stacking the observations y_t', x_t' vertically, we can write the seemingly unrelated regressions version of the VAR:

$$Y = XB + U, \quad U \sim \mathcal{MN}(0, I_T, \Sigma). \tag{E.1}$$

The dimensions of Y, X, U, B, and Σ are respectively (T, n), (T, k), (T, n), (k, n), and (n, n) where T is the number of observations, n the number of endogenous variables, k is the number of right-hand side variables. Moreover, we have: $k = n \times L + q$, where L is the number of lags and q the number of control variables. $\mathcal{MN}(0, I_T, \Sigma)$ is the matrix normal distribution with variance among rows I_T and variance among columns Σ.

E.2 Unconstrained VAR: Prior and Posterior Distributions

E.2.1 Quarterly Unconstrained VAR

For the quarterly VARs, we assume the following flat prior distribution (Jeffreys, 1946):

$$p(B, \Sigma) \propto |\Sigma|^{-(n+1)/2}. \tag{E.2}$$

This prior implies the usual normal-inverse Wishart posterior distribution:

$$\operatorname{vec}(B) \mid \Sigma, Y, X \sim \mathcal{N}\left(\operatorname{vec}(\tilde{B}), \tilde{\Omega}\right), \quad \text{and} \tag{E.3}$$

$$\Sigma \mid Y, X \sim \mathcal{W}^{-1}\left(\tilde{S}, \tilde{\nu}\right), \tag{E.4}$$

where

$$\begin{aligned}
\tilde{B} &= \left(X'X\right)^{-1}\left(X'Y\right), \\
\tilde{\Omega} &= \Sigma \otimes \left(X'X\right)^{-1}, \\
\tilde{S} &= Y'Y - Y'X(X'X)^{-1}X'Y, \quad \text{and} \\
\tilde{\nu} &= T - k.
\end{aligned}$$

Note that the mean, median, and mode of the posterior distribution of B coincide with the equation-by-equation ordinary least squares (OLS) estimator of B—a standard property of normal models in Bayesian econometrics. Since IRFs are non-linear functions of B, their mean and median depend on the whole distribution and may not exactly correspond to their OLS counterparts. The IRF at the mode the distribution of B, on the other hand, is the same as the IRF at the OLS estimator. In practice, these three IRFs are very close.

E.2.2 Monthly Unconstrained VAR

For the monthly VARs, we assume the following Minnesota prior distribution:

$$B_i' \sim \mathcal{N}\left(\underline{B}_i', \frac{1}{\kappa_i}\Sigma\right), \quad 1 \leq i \leq k, \quad \text{and} \tag{E.5}$$

$$\Sigma \sim \mathcal{W}^{-1}\left(\underline{S}, \underline{\nu}\right), \tag{E.6}$$

where B_i is row i of B. We implement this prior with dummy observations.[6] For each $i \in [\![1, k]\!]$, we can add the dummy observation $(\underline{Y}_i, \underline{X}_i) = (\sqrt{\kappa_i}\underline{B}_i, \sqrt{\kappa_i}e_i)$ where e_i is the row vector of dimension k whose element i is 1 and other elements are 0. The likelihood of that observation is proportional to the density of the prior distribution in equation (E.5):

$$|\Sigma|^{-1/2} \exp\left(-\frac{1}{2}\left(\underline{Y}_i - \underline{X}_i B\right)\Sigma^{-1}\left(\underline{Y}_i - \underline{X}_i B\right)'\right) \propto \mathcal{N}\left(\underline{B}_i, \frac{1}{\kappa_i}\Sigma; B_i\right).$$

[6] See Del Negro and Schorfheide (2012) for a handbook exposition of prior implementation with dummy observations.

Similarly, we implement equation (E.6) by adding $\underline{\underline{\nu}}$ times the following dummy observations: $\left(\underline{s}_j v_j, (0)_k\right)$, $1 \leq j \leq n$, where $\underline{s}_j$ is a scalar, v_j is a row vector of dimension n whose j^{th} element is 1 and other elements are 0 and $(0)_k$ is a row vector of dimension k whose elements are 0. The joint likelihood of those observations is proportional to the density of the prior distribution in equation (E.6):

$$\left(\prod_{j=1}^{n} |\Sigma|^{-1/2} \exp\left(-\frac{1}{2}\underline{s}_j^2 v_j \Sigma^{-1} v_j'\right)\right)^{\underline{\underline{\nu}}} = |\Sigma|^{-n\underline{\underline{\nu}}/2} \exp\left(-\frac{1}{2}\text{tr}\left(\Sigma^{-1}\underline{\underline{\nu}}\sum_{j=1}^{n}\underline{s}_j^2(v_j'v_j)\right)\right)$$
$$\propto \mathcal{W}^{-1}\left(\underline{S}, \underline{\nu}; \Sigma\right),$$

where

$$\underline{S} = \underline{\underline{\nu}} \begin{pmatrix} \underline{s}_1^2 & 0 & \vdots & 0 \\ 0 & \underline{s}_2^2 & \vdots & 0 \\ \dots & \dots & \vdots & \dots \\ 0 & 0 & \vdots & \underline{s}_n^2 \end{pmatrix}, \quad \text{and} \quad \underline{\nu} = n\underline{\underline{\nu}} - n - 1$$

This prior also implies a normal-inverse Wishart posterior distribution:

$$\text{vec}(B) \mid \Sigma, \bar{Y}, \bar{X} \sim \mathcal{N}\left(\text{vec}(\tilde{B}), \tilde{\Omega}\right), \quad \text{and} \tag{E.7}$$
$$\Sigma \mid \bar{Y}, \bar{X} \sim \mathcal{W}^{-1}\left(\bar{S}, \bar{\nu}\right), \tag{E.8}$$

where

$$\bar{B} = \left(\bar{X}'\bar{X}\right)^{-1}\left(\bar{X}'\bar{Y}\right),$$
$$\bar{\Omega} = \Sigma \otimes \left(\bar{X}'\bar{X}\right)^{-1},$$
$$\bar{S} = \bar{Y}'\bar{Y} - \bar{Y}'\bar{X}(\bar{X}'\bar{X})^{-1}\bar{X}'\bar{Y}, \quad \text{and}$$
$$\bar{\nu} = T + \underline{\nu}.$$

where $\bar{Y}$ and $\bar{X}$ are the stacked actual and dummy observations.

E.3 Constrained VAR: Prior and Posterior Distributions

E.3.1 Quarterly Constrained VAR

The vector y_t is partitioned into its debt component and its other variables: $y_t' = (d_{t-1}, y_t^{o\prime})$. We can do the same for B: $B = \begin{pmatrix} B^d & B^o \end{pmatrix}$. Superscript d denotes the first column, which contains the coefficients for the debt equation; superscript o is for the $n-1$ other equations. We can similarly partition Σ:

$$\Sigma = \begin{pmatrix} \Sigma_{dd} & \Sigma_{do} \\ \Sigma_{do}' & \Sigma_{oo} \end{pmatrix}.$$

We write the restriction as a subset of the rows of B^d, $\mathcal{R} \subset [1,k]$, being equal to a constant column vector c:

$$B^d_{\mathcal{R}} = c. \tag{E.9}$$

We denote the subset of unrestricted columns of B^d by $\mathcal{U}$. Note that we have

$$XB^d = X^{\mathcal{R}} B^d_{\mathcal{R}} + X^{\mathcal{U}} B^d_{\mathcal{U}} = X^{\mathcal{R}} c + X^{\mathcal{U}} B^d_{\mathcal{U}},$$

where $X^{\mathcal{R}}$ ($X^{\mathcal{U}}$) are the columns of X whose coefficients are restricted (unrestricted) in the debt equation.

Finally, we denote the vectorized matrix of unrestricted coefficient, $\text{vec}\left(B \setminus B^d_{\mathcal{R}}\right)$, by $\mathcal{B}$, and define the following matrix:

$$F := \begin{pmatrix} I_{\mathcal{U}} & \mathbf{0}_{(k,(n-1)k)} \\ \mathbf{0}_{((n-1)k,k-k_{\mathcal{R}})} & I_{(n-1)k} \end{pmatrix}.$$

F selects, in the debt equation, the variables whose coefficients are unrestricted:

$$(I_n \otimes X) F = \begin{pmatrix} X^{\mathcal{U}} & (0) \\ (0) & I_{n-1} \otimes X \end{pmatrix}.$$

For the quarterly constrained VAR, we still use the flat prior of equation (E.2). In a companion note, we show the following proposition.

Proposition E.1 (Jeffreys's prior) *With the flat prior (E.2), the posterior distributions of B and Σ are*

$$\begin{aligned}
&vec(B) \mid \Sigma, Y, X \sim \mathcal{N}\left(vec\left(\tilde{B}\right), \tilde{\Omega}\right), \\
&\Sigma_{dd}^{-1}\Sigma_{do} \mid \left(\Sigma^{-1}\right)_{oo}^{-1}, Y, X \sim \mathcal{MN}\left(\tilde{S}_{dd}^{-1}\tilde{S}_{do}, \tilde{S}_{dd}^{-1}, \left(\Sigma^{-1}\right)_{oo}^{-1}\right), \\
&\left(\Sigma^{-1}\right)_{oo}^{-1} \mid Y, X \sim \mathcal{W}^{-1}\left(\left(\tilde{S}^{-1}\right)_{oo}^{-1}, \tilde{\nu}\right), \quad \textit{and} \\
&\Sigma_{dd} \mid Y, X \sim \mathcal{W}^{-1}\left(\tilde{S}_{dd} + \tilde{R}_{dd}, \tilde{\nu} + k_{\mathcal{R}} - (n-1)\right),
\end{aligned}$$

where

$$\begin{aligned}
\tilde{\mathcal{B}} &= \left(F'\left(\Sigma^{-1}\otimes X'X\right)F\right)^{-1}\left(F'\left(\Sigma^{-1}\otimes X'\right)vec(Z)\right),\\
\tilde{\Omega} &= \left(F'\left(\Sigma^{-1}\otimes X'X\right)F\right)^{-1},\\
\tilde{S} &= Z'Z - Z'X(X'X)^{-1}X'Z,\\
\tilde{R}_{dd} &= Z^{d\prime}X\left((X'X)^{-1} - I_{\mathcal{U}}\left(I_{\mathcal{U}}'X'XI_{\mathcal{U}}\right)^{-1}I_{\mathcal{U}}'\right)X'Z^{d},\\
Z &= Y - X^{\mathcal{R}}c\begin{pmatrix}1 & (0)_{1,n-1}\end{pmatrix}, \quad \textit{and}\\
\tilde{\nu} &= T-k.
\end{aligned}$$

E.3.2 Quarterly Constrained VAR

For the monthly constrained VAR, we also use a Minnesota prior, but equation (E.5) needs to be modified since some of the elements of B are restricted. The constrained analog of equations (E.5–E.6) is

$$B_i' \sim \mathcal{N}\left(\underline{B}_i', \frac{1}{\kappa_i}\Sigma\right), \quad i\in\mathcal{U}, \tag{E.10}$$

$$B_i^{o\prime} \sim \mathcal{N}\left(\underline{B}_i^{o\prime}, \frac{1}{\kappa_i}\left(\Sigma^{-1}\right)_{oo}^{-1}\right), \quad i\in\mathcal{R}, \quad \text{and} \tag{E.11}$$

$$\Sigma \sim \mathcal{W}^{-1}\left(\underline{S}, \underline{\nu}\right), \tag{E.12}$$

where $\left(\Sigma^{-1}\right)_{oo}$ denotes the bottom right quadrant of the inverse of Σ. We again implement this prior with dummy observations.

Proposition E.2 (Minnesota prior) *With the Minnesota prior (E.10–E.12), the posterior distributions of B and Σ are*

$$\begin{aligned}
&\mathcal{B}\mid\Sigma,\bar{Y},\bar{X} \sim \mathcal{N}\left(\bar{\mathcal{B}},\bar{\Omega}\right),\\
&\Sigma_{dd}^{-1}\Sigma_{do}\mid\left(\Sigma^{-1}\right)_{oo}^{-1},\bar{Y},\bar{X} \sim \mathcal{MN}\left(\bar{S}_{dd}^{-1}\bar{S}_{do}, \bar{S}_{dd}^{-1}, \left(\Sigma^{-1}\right)_{oo}^{-1}\right),\\
&\left(\Sigma^{-1}\right)_{oo}^{-1}\mid\bar{Y},\bar{X} \sim \mathcal{W}^{-1}\left(\left(\bar{S}^{-1}\right)_{oo}^{-1}, \bar{\nu}\right), \quad \textit{and}\\
&\Sigma_{dd}\mid\bar{Y},\bar{X} \sim \mathcal{W}^{-1}\left(\bar{S}_{dd}+\bar{R}_{dd}, \bar{\nu}-(n-1)\right),
\end{aligned}$$

where

$$\begin{aligned}
\bar{\mathcal{B}} &= \left(F'\left(\Sigma^{-1}\otimes\bar{X}'\bar{X}\right)F\right)^{-1}\left(F'\left(\Sigma^{-1}\otimes\bar{X}'\right)vec(\bar{Z})\right),\\
\bar{\Omega} &= \left(F'\left(\Sigma^{-1}\otimes\bar{X}'\bar{X}\right)F\right)^{-1},
\end{aligned}$$

$$
\begin{aligned}
&\bar{S} = \bar{Z}'\bar{Z} - \bar{Z}'\bar{X}(\bar{X}'\bar{X})^{-1}\bar{X}'\bar{Z} + \underline{S}, \\
&\bar{R}_{dd} = \bar{Z}^{d\prime}\bar{X}\left((\bar{X}'\bar{X})^{-1} - I_{\mathcal{U}}\left(I_{\mathcal{U}}'\bar{X}'\bar{X}I_{\mathcal{U}}\right)^{-1} I_{\mathcal{U}}'\right)\bar{X}'\bar{Z}^{d}, \\
&\bar{Z} = \bar{Y} - \bar{X}^{\mathcal{R}}cM^{d}, \quad \textit{and} \\
&\bar{\nu} = T + \underline{\nu}.
\end{aligned}
$$

Propositions E.1 and E.2, which are proved in a companion note, are generalizations of equations (E.3–E.4) and (E.7–E.8). For instance, for proposition E.1, without restrictions we have: $\mathcal{R} = \emptyset$, so $Y = Z$ and $I_{\mathcal{U}} = I$. The posterior distribution for Σ becomes

$$
\begin{aligned}
&\Sigma_{dd}^{-1}\Sigma_{do} \mid \left(\Sigma^{-1}\right)_{oo}^{-1}, Y, X \sim \mathcal{MN}\left(\tilde{S}_{dd}^{-1}\tilde{S}_{do}, \tilde{S}_{dd}^{-1}, \left(\Sigma^{-1}\right)_{oo}^{-1}\right), \\
&\left(\Sigma^{-1}\right)_{oo}^{-1} \mid Y, X \sim \mathcal{W}^{-1}\left(\left(\tilde{S}^{-1}\right)_{oo}^{-1}, \tilde{\nu}\right), \quad \text{and} \\
&\Sigma_{dd} \mid Y, X \sim \mathcal{W}^{-1}\left(\tilde{S}_{dd}, \tilde{\nu} - (n-1)\right),
\end{aligned}
$$

where $\tilde{S} = Y'Y - Y'X(X'X)^{-1}X'Y$. These equations are simply the marginal distributions of the elements of Σ when it follows an inverse-Wishart distribution with scale matrix $\tilde{S}$ and degrees of freedom $\tilde{\nu}$ (Gupta and Nagar, 2018, theorem 3.3.9).

E.4 Structural VAR with External Instrument (SVAR-IV)

The SVAR-IV methodology was pioneered by Stock and Watson (2012) and Mertens and Ravn (2013).

In a structural VAR (SVAR), the assumption is that the dynamics of the economy can be summarized as:

$$
A(0)'y_{t+1} = \sum_{l=1}^{L} A(l)'y_{t+1-l} + A_c'c_{t+1} + e_{t+1}, \quad e_{t+1} \sim \mathcal{N}(0, I). \tag{E.13}
$$

If $A(0)'$ is invertible, we can recover equation (5) with $B(l)' = (A(0)')^{-1}A(l)'$ and $B_c' = (A(0)')^{-1}A_c'$.

In a structural VAR, the reduced-form shocks, u_t, are assumed to be a linear combination of structural shocks, e_t:

$$
u_t = (A(0)')^{-1}e_t.
$$

We can use this expression to rewrite non-monetary reduced form shocks, u_t^{-m}, in terms of the reduced-form monetary shock, u_t^m, and the non-monetary structural shocks, e_t^{-m}:

$$u_t^{-m} = \psi_m u_t^m + \omega_{-m} e_t^{-m}. \tag{E.14}$$

We can instrument u_t^m with the Bauer–Swanson (or Jarociński–Karadi, Miranda-Agrippino–Ricco+) monetary shock, which identifies ψ_m if the identified monetary shock is orthogonal to structural non-monetary shocks. Plagborg-Møller and Wolf (2021) show that the SVAR-IV is asymptotically equivalent to a local projection under partial invertibility, that is, there must exist a vector b that can be recovered from the reduced-form parameters and such that: $e_t^m = b' u_t$.

We focus on the case of a monetary shock in the aforementioned explanation, but the technique can be applied to any shock. There can be several instrumented reduced-form shocks as long as we have the appropriate instruments. The Caldara–Kamps tax and transfer shocks of section I are an example. If the VAR contains a pre-determined variable or a narrative shock as in section 3.3.1, we can order it first and assume: $u_t^1 = e_t^1$. Equation (E.14) then becomes:

$$u_t^{-1,-m} = \psi_1 u_t^1 + \psi_m u_t^m + \omega_{-1,-m} e_t^{-1,-m}. \tag{E.15}$$

E.5 Minnesota Prior Distribution: Parametric Choices

As usual with a Minnesota prior distribution, we center the distribution around a random walk. For each variable, the mean of the coefficient is 1 for the first own lag of the variable and 0 for the rest. Formally, we have

$$\underline{B}_i^j = \begin{cases} 1, & 1 \leq i = j \leq n, \quad \text{and} \\ 0, & \text{otherwise,} \end{cases}$$

where $\underline{B}_i^j$ is the prior mean for row i, column j of the B matrix. For the variance, we have

$$\kappa_i = \begin{cases} \left(\lambda_1 \times (l(i))^{\lambda_2} \times \hat{\sigma}_{\tilde{\imath}(i)}\right)^2, & 1 \leq i \leq n \times L, \quad \text{and} \\ \left(\lambda_0 \times \lambda_1\right)^2, & n \times L < i \leq k, \end{cases}$$

where $\tilde{\imath}(i) \in [\![1, n]\!]$ and $l(i) \in [\![1, L]\!]$ respectively are the variable and lag to which row i corresponds and $\hat{\sigma}_{\tilde{\imath}(i)}$ is the standard deviation of the residuals of an OLS estimation of the equation for variable $\tilde{\imath}(i)$ in the VAR system (5). For λ, we choose

$$\lambda_0 = \left(\frac{1}{10^5}\right)^{\frac{1}{2}}, \quad \lambda_1 = 5, \quad \text{and} \quad \lambda_2 = 1.$$

For the prior distribution of the covariance matrix, we use: $\underline{s}_j = \hat{\sigma}_j$, $1 \leq j \leq n$, and $\underline{\underline{\nu}} = 3$.

E.6 Estimation Algorithm: Details and Performance

When no data is missing, which is the case for most quarterly specifications, the posterior distribution is known and the distribution can be sampled independently. We take 20,000 draws.

When data is missing, we use Gibbs sampling: (i) given parameters, we infer the posterior distribution of the missing data—we obtain mean and variance by Kalman filtering and state smoothing and then we draw from the distribution using simulation smoothing; (ii) given observed and missing data, we draw from the posterior distribution of the parameters. See Durbin and Koopman (2012, chapter 4) for a textbook treatment of simulation smoothing and Giordani et al. (2012) for a handbook exposition of missing data in VARs. We take 400,000 draws, burn the first 80,000 draws, and keep every 16^{th} draw of the remaining 320,000, which leaves us with 20,000 draws.

By construction, the draws obtained by Gibbs sampling are not independent. We compute the effective sample size of our chain to check that we have achieved a reasonable sampling of the posterior distribution. We follow Vehtari et al. (2021) and compute it with rank-normalized draws. We report effective sample sizes in table E.1. Apart from unconstrained monthly specifications in which we use quarterly debt from the Financial Accounts, (figures A.6–A.7), we always achieve an effective sample size above 400. That's why we take figures A.6–A.7 with caution and prefer specifications with monthly federal debt in the hands of the public (figures 2–3) or a constrained VAR (figures A.9–A.10).

F Unobserved Fiscal Instruments

F.1 General Case

In macroeconomic models, tax receipts and transfer payments are often endogenous objects whose rule is not directly controlled by the government. For instance, tax receipts depend on economic activity as well as a collection of tax rates. If unemployment insurance is modeled, transfer payments depend on the unemployment rate as well as the level of unemployment

Table E.1: Effective Sample Sizes

Figure	ESS	Figure	ESS
1	18,078	A.9	3,379
2	1,240	A.10	4,177
3	570	A.11a(2)	17,721
4a	18,656	A.11a(3)	18,151
4b	556	A.11b(3)	421
4c	752	A.11c(3)	1,576
5a	10,933	A.12, A.13	16,666
5b	4,437	A.14, A.15	18,231
5c	926	A.16, A.17	18,113
7a(2,3)	17,063	A.18(2,5)	18,024
7b(2,3)	2,633	A.18(3), A.21	17,882
7c(2,3)	1,320	A.18(6), A.23	17,611
8(2)	18,199	A.19	17,286
A.1	1,205	A.20	18,271
A.2	2,330	A.22	17,453
A.6	102	A.24	17,833
A.7	73	A.25	18,113
A.8	17,685		

Note: We compute the effective sample size (ESS) of each impulse response coefficient to a structural shock after rank-normalization, as suggested by Vehtari et al. (2021). For each VAR model, we report the minimum ESS across the impulse response coefficients up to five years of all the variables in the model. Each VAR model is indicated by the figure(s) in which the impulse responses from the model are presented. When a VAR model is used only in a subset of plots within a figure, we indicate the plot numbers in the parentheses.

benefits. In these examples, the government directly controls the rule for tax rates and unemployment benefits, not those for tax receipts and expenses on unemployment insurance. Interpreted literally, the MKW would require observing every tax rate or transfer benefit as well as several exogenous shock paths for each of those. In practice, we only observe tax receipts and transfers payments, and we only have two shock paths for each of those.

How can we map our exercise to such a model? Consider a theoretical environment with a collection of (perhaps unobserved) instruments, z_i, $1 \leq i \leq n$. We can think of those as tax rates on various types and brackets of income, or transfer benefits for various situations (poverty, unemployment, retirement...). Suppose, on the other hand, that we observe another variable, X, which we seek to stabilize. In our counterfactual exercises, this variable is debt or deficit. It belongs to the non-policy block since it moves endogenously in response to the fiscal instruments and economic fluctuations. The MKW minimization problem, which we reproduce here for convenience, is the following:

$$\min_{\mathbf{s}} \left|\left| \hat{\mathcal{A}}_x \left(\mathbf{x}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \mathbf{\Omega}_{\mathbf{x},\mathcal{A}} \times \mathbf{s} \right) + \hat{\mathcal{A}}_z \left(\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A}} \times \mathbf{s} \right) \right|\right| . \tag{14}$$

Each row of the $\hat{\mathcal{A}}_x$ and $\hat{\mathcal{A}}_z$ matrices defines the counterfactual policy rule of one instrument for one time period.

Since equation (14) contains the empirical response of the instruments to the monetary ($\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$) and policy ($\Omega_{z,\mathcal{A}}$) shocks, not observing the instruments would seem damning. We can, however, work around this problem by enforcing a policy rule that exhibits a weighted average of the unobserved instrument combinations that actually happen in response to the non-policy and policy shocks. Formally, let us choose (without loss of generality) the stabilization of X as the policy rule for the first instrument, z_1: $X_t = 0$ for all t. As no instrument appears in this equation, the corresponding rows of the $\hat{\mathcal{A}}_z$ matrix are made of zeros. Hence, we don't need to observe $\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$ and $\Omega_{z,\mathcal{A}} \times \mathbf{s}$ to know the value of the corresponding rows of $\hat{\mathcal{A}}_z \left(\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon}) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A}} \times \mathbf{s} \right)$: that value is simply 0. For the other instruments, we choose the following policy rule:

$$z_{it} = \frac{\mathbf{z}_{\mathcal{A},\mathbf{z_{it}}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A},\mathbf{z_{it}}} \times \mathbf{s}}{\mathbf{z}_{\mathcal{A},\mathbf{z_{1t}}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A},\mathbf{z_{1t}}} \times \mathbf{s}} \times z_{1,t}, \qquad i > 1, \tag{F.1}$$

where the z_{it} subscript denotes the line that corresponds to the instrument z_i at time t in $\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$ and $\Omega_{z,\mathcal{A}}$. The corresponding row of $\hat{\mathcal{A}}_z$ (which we denote by $\hat{\mathcal{A}}_{z_{it}}$) contains $\frac{\mathbf{z}_{\mathcal{A},\mathbf{z_{it}}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A},\mathbf{z_{it}}} \times \mathbf{s}}{\mathbf{z}_{\mathcal{A},\mathbf{z_{1t}}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A},\mathbf{z_{1t}}} \times \mathbf{s}}$ in the column for z_{1t}, -1 in the column for z_{it}, and 0 in the other columns. So, it is trivial to check that the matrix product between this row and the vector $\mathbf{z}_{\mathcal{A}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A}} \times \mathbf{s}$ is 0: $\hat{\mathcal{A}}_{z_{it}}(\mathbf{z}_{\mathcal{A}}(\varepsilon) + \mathbf{\Omega}_{\mathbf{z},\mathcal{A}} \times \mathbf{s}) = \mathbf{0}$. Since equation (F.1) does not feature any non-policy variable, the corresponding row of $\hat{\mathcal{A}}_x$ only contains zeros. Therefore, the corresponding row of $\hat{\mathcal{A}}_x\left(\mathbf{x}_{\mathcal{A}}(\boldsymbol{\varepsilon})+\boldsymbol{\Omega}_{\mathbf{x},\mathcal{A}}\times\mathbf{s}\right)+\hat{\mathcal{A}}_z\left(\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})+\boldsymbol{\Omega}_{\mathbf{z},\mathcal{A}}\times\mathbf{s}\right)$ is also 0 and irrelevant to the minimization problem. As a result, the only rows that matter for the minimization problem are those that correspond to the policy rule for z_{1t} and where the unobserved instruments don't appear, so that we don't need to know $\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$ and $\Omega_{z,\mathcal{A}}\times\mathbf{s}$.

Once again equation (F.1) is not just a mathematical trick. Implicitly, this policy rule enforces a weighted average of the instrument combinations that happen in response to the non-policy shock ($\mathbf{z}_{\mathcal{A}}(\boldsymbol{\varepsilon})$) and to the policy shocks ($\Omega_{z,\mathcal{A}}\times\mathbf{s}$). We don't observe those combinations, but they are the empirically relevant ones.

F.2 Application to Bianchi et al. (2023)

As we explained in sections 4.3 and 5.1, Bianchi et al. (2023) propose a model with "shock-specific" monetary rules. The central bank responds actively to funded shocks, while it lets inflation rise in response to unfunded shocks. Their model features a "shadow economy," which registers the path that variables would take if there were only unfunded shocks. This speaks directly to the unobserved instruments we've described in section F.1: in the data, we don't observe the shadow economy (by definition); neither do we observe which shocks are funded or unfunded.

Formally, we can map their model into the MKW framework by collecting $\hat{g}$, $\hat{z}^b$, $\hat{z}$, $\hat{\tau}_L$, $\hat{\tau}_K$, and their shadow economy counterparts in the vector of policy variables ($\mathbf{z}$), ζ_g, ζ_z, ζ^M, ζ^F in the vector of policy shocks ($\boldsymbol{\nu}$), the remaining endogenous variables in the vectors of unobserved ($\mathbf{w}$) and observed ($\mathbf{x}$) policy variables, and the remaining shocks in the vector of non-policy shocks ($\boldsymbol{\varepsilon}$). The non-policy block is made of equations (66–83, 89) and their shadow counterparts, while the policy block features equations (84–88, 91–95).[7] Then, the fact that the distinction between funded and unfunded fiscal instruments is unobservable is addressed as in appendix F.1, by enforcing the fiscal rule to be a weighted average of the instrument combinations that happen in response to the non-policy and policy shocks.

In light of the Bianchi-Faccini-Melosi model, the transfer results of section 5.1 suggest that transfer shocks are mostly unfunded. If so, our transfer scenario relies on unfunded transfers: the fiscal authority fights the monetary contraction by increasing unfunded expenses. This forces the central bank to let inflation rise in response, thus undoing the monetary contraction.

[7]Since, in our application of the MKW method, the structural shock is a monetary policy shock, the Taylor rule (89) belongs to the non-policy block (see section 4.3). This implies that we cannot handle the effective lower bound as Bianchi et al. do. Indeed, under realistic "informational requirements," the non-policy block must be linear. See McKay and Wolf (2023, appendices A.8–9).

G Sources

The sources are summarized in table G.1. The definition of the fiscal flow variables is:

$$\begin{aligned} \text{spending} &= \text{consumption expenditures (line 25)} + \text{subsidies (line 36)} \\ &\quad + \text{gross investment (line 45)} - \text{capital consumption (line 48)}, \\ \text{transfers} &= \text{current transfer payments (line 26)} \\ &\quad + \text{capital transfer payments (line 46)}, \quad \text{and} \\ \text{tax receipts} &= \text{total receipts (line 40)}. \end{aligned}$$

("Line" refers to the line in NIPA table 3.2.)

Table G.1: Data sources

Variable	Definition	Source
FOMC Forecasts	See section 2.1	Croushore and van Norden (2018)
Fiscal shocks	See section 4.3.1	Ramey (2016) and own computations
GDP	GDP in constant 2012 prices	NIPA table 1.1.6
Price index	GDP deflator	NIPA table 1.1.4
Nominal interest rates	3-month and 1-year rates on Treasury securities	Federal Reserve Board downloaded from Fred
Excess bond premium	Gilchrist and Zakrajšek (2011)	Federal Reserve Board
Fiscal flow variables	See appendix G	NIPA table 3.2
Net debt (preferred concept)	Total liabilities (FGTLBLQ027S) minus total financial assets (FGTFASQ027S)	Financial Accounts of the United States (Z.1) downloaded from FRED
Federal debt held by the public	Debt held by private investors and the Federal Reserve	Payne et al. (2025)

Note: Sources for the data used in the paper.

H Credible Bands in the MKW Method

In the main text, we constructed credible bands by minimizing equation (14) at the modal IRF. Holding the loadings $\mathbf{s}$ constant, we then compute the counterfactual response under this policy for each draw of the posterior distribution of the IRFs to spending, tax, or transfer shocks. As we mentioned in the text, McKay and Wolf (2023) are silent about the construction of confidence of credible (or confidence) bands, but their code indicates that they minimize equation (14) for each draw of the IRFs to the fiscal shocks. We follow that strategy in this appendix.

We show the results obtained by recomputing the loadings in figures H.1–H.3. Perhaps surprisingly, recomputing the loadings does not necessarily lead to wider credible bands. In fact, by construction, the bands are narrower for debt since, for each draw, we recompute the loadings that stabilize debt as well as possible. On the other hand, for the stabilizing fiscal instrument, the bands tend to be wider since the loadings on that instrument are more volatile. For macroeconomic variables, the bands are of similar width in the spending and transfers counterfactual scenarios.

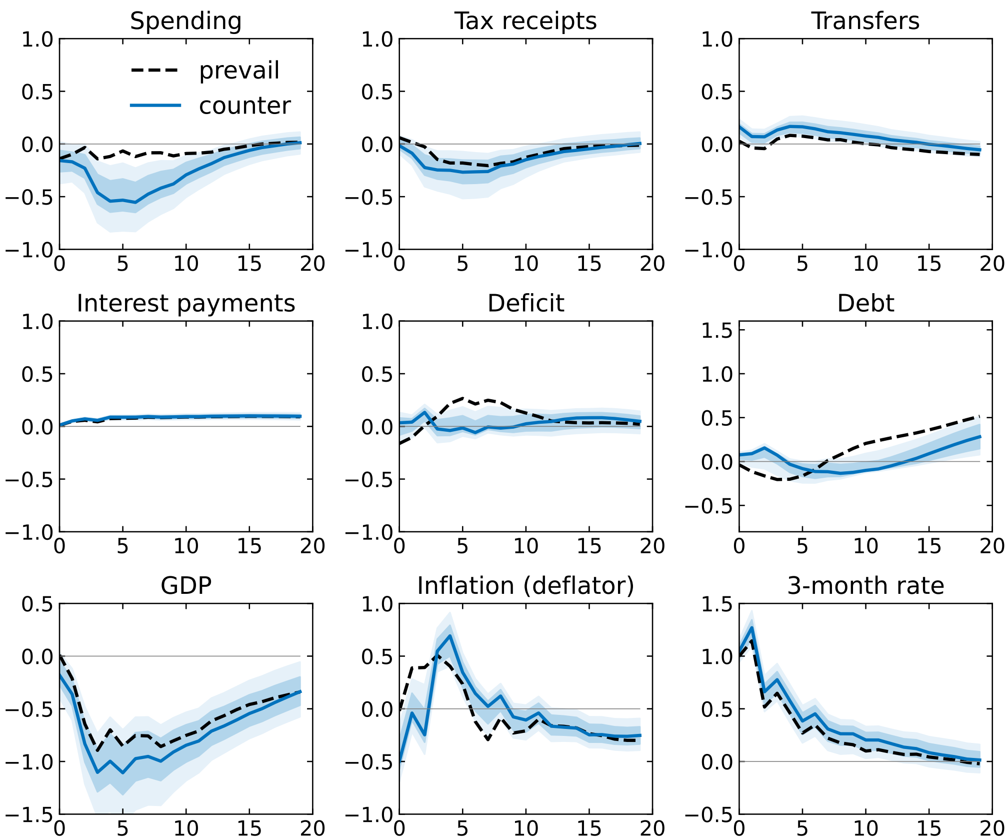
A notable exception is the tax counterfactual scenario (figure H.2): the response under the prevailing rule lies within the 68% confidence band for taxes and GDP. Moreover, the modal IRF lies on the edge of the 68% interval. Inspecting the mechanism, we found that this imprecision stemmed from the similar shapes of the debt responses across the two shocks. At the modal IRFs, the loading on the Romer–Romer shock is positive while that on the Caldara-Kamps shock is negative: the solution of the minimization problem raises taxes with the Romer–Romer shock while cutting them with the Caldara-Kamps shock in order to stabilize debt. When we recompute the loadings, they become very volatile. As we show in table H.1, the loadings have the same sign as the modal ones (positive on Romer-Romer, negative on Caldara–Kamps) with a probability of 0.46; both have inverted signs (negative on Romer-Romer, positive on Caldara–Kamps) with a probability of 0.34. Such sign flips for both loadings only happen with a probability of less than 2% in the spending or transfer counterfactual scenario (table H.1).

Table H.1: MKW loadings—probability distribution

		Structural shock	
		Positive	Negative
Narrative shock	Spending $(-0.31, -0.04)$		
	Positive	0.02	0.01
	Negative	0.33	0.64
	Taxes $(1.21, -0.26)$		
	Positive	0.14	0.46
	Negative	0.34	0.06
	Transfers $(0.29, 0.24)$		
	Positive	0.75	0.07
	Negative	0.16	0.02

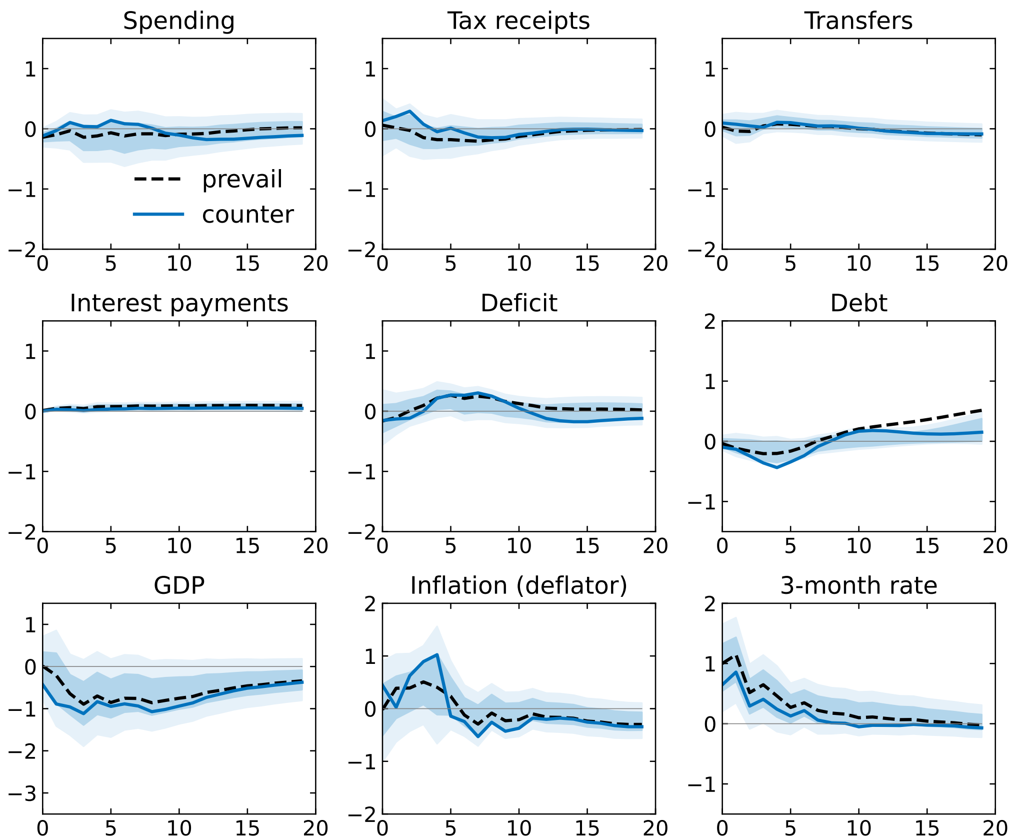
Note: Probability of the sign combination of the loadings on the narrative (Ramey, Romer–Romer) and structural shocks (Blanchard–Perotti, Caldara–Kamps). The numbers in parenthesis are the loadings at the modal IRFs.

Figure H.1: Counterfactual—debt stabilization with spending (updated loadings)



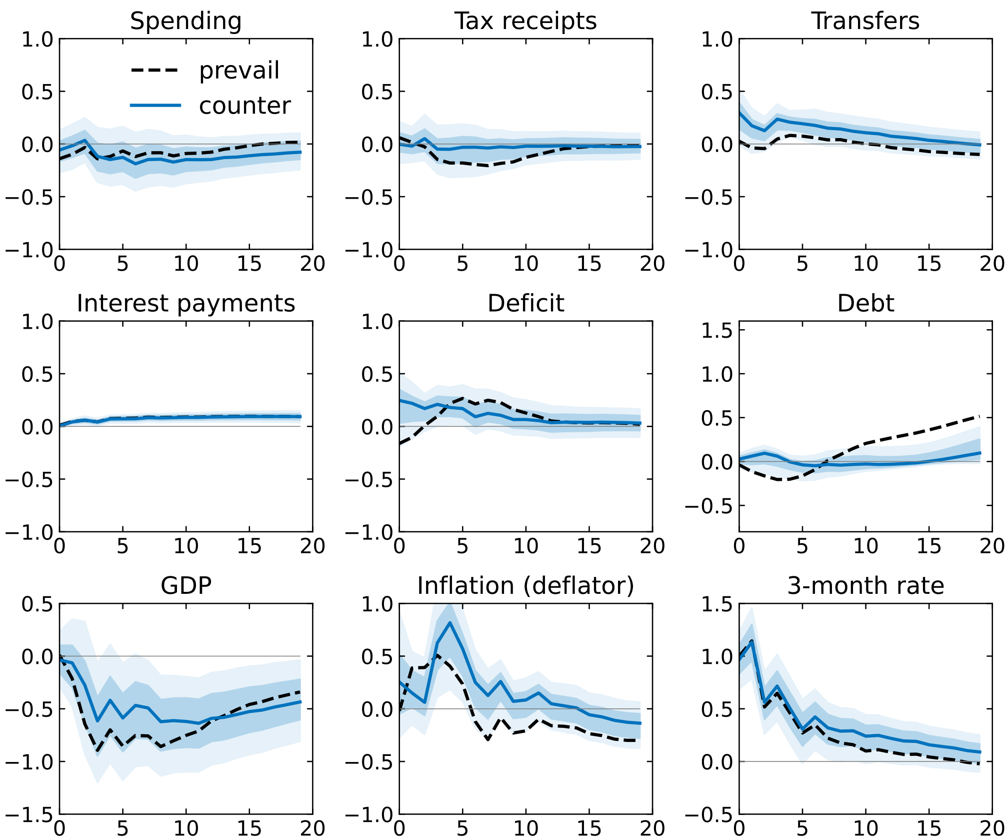
Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through spending. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1). Compared to figures 11–13, we update loadings for each draw of the posterior distribution. See section H for a detailed discussion.

Figure H.2: Counterfactual—debt stabilization with taxes (updated loadings)



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through taxes. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1). Compared to figures 11–13, we update loadings for each draw of the posterior distribution. See section H for a detailed discussion.

Figure H.3: Counterfactual—debt stabilization with transfers (updated loadings)



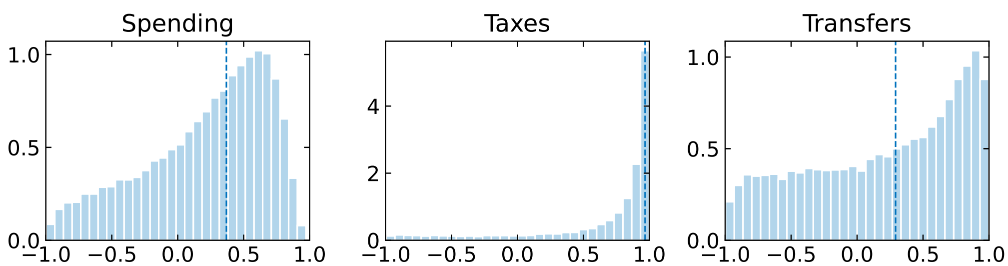
Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through transfers. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual scenario is constructed with the MKW method (section 4.1). Compared to figures 11–13, we update loadings for each draw of the posterior distribution. See section H for a detailed discussion.

To formally quantify the similarity of the shapes of the IRFs, we compute their cosine similarity. Cosine similarity is a multi-dimensional generalization of the cosine function. It is commonly used in text analysis to assess the similarity of two documents (Salton and McGill, 1983, pp. 120–3). It is, for example, at the heart of large language models (Krantz and Jonker, 2025). Cosine similarity is the ratio of the dot product of the two IRFs to the product of their Euclidian norms:

$$\text{cosine similarity} = \frac{\Omega^1_{d,\mathcal{A}} \cdot \Omega^2_{d,\mathcal{A}}}{\left|\left|\Omega^1_{d,\mathcal{A}}\right|\right| \times \left|\left|\Omega^2_{d,\mathcal{A}}\right|\right|},$$

where $\Omega^1_{d,\mathcal{A}}$ and $\Omega^2_{d,\mathcal{A}}$ are the debt IRF for the first and second shocks. With two time periods $(t = 0, 1)$, this ratio would be equal to the cosine of the angle between the vectors pointing from the origin to $(\Omega^1_{d,\mathcal{A},t=0}, \Omega^1_{d,\mathcal{A},t=1})$ and to $(\Omega^2_{d,\mathcal{A},t=0}, \Omega^2_{d,\mathcal{A},t=1})$. Its value would be 1 if the vectors point in the same direction, 0 if they are orthogonal, and -1 if they point in opposite direction. We would expect the MKW method to fare better if the measure is close to 0. We plot the posterior distribution of the cosine similarity in figure H.4. The draws are spread across the $(-1, 1)$ interval for the spending and transfers IRFs. On the other hand, they are concentrated near 1 for the tax IRFs, thus confirming our visual statement that the two IRFs have similar shapes.

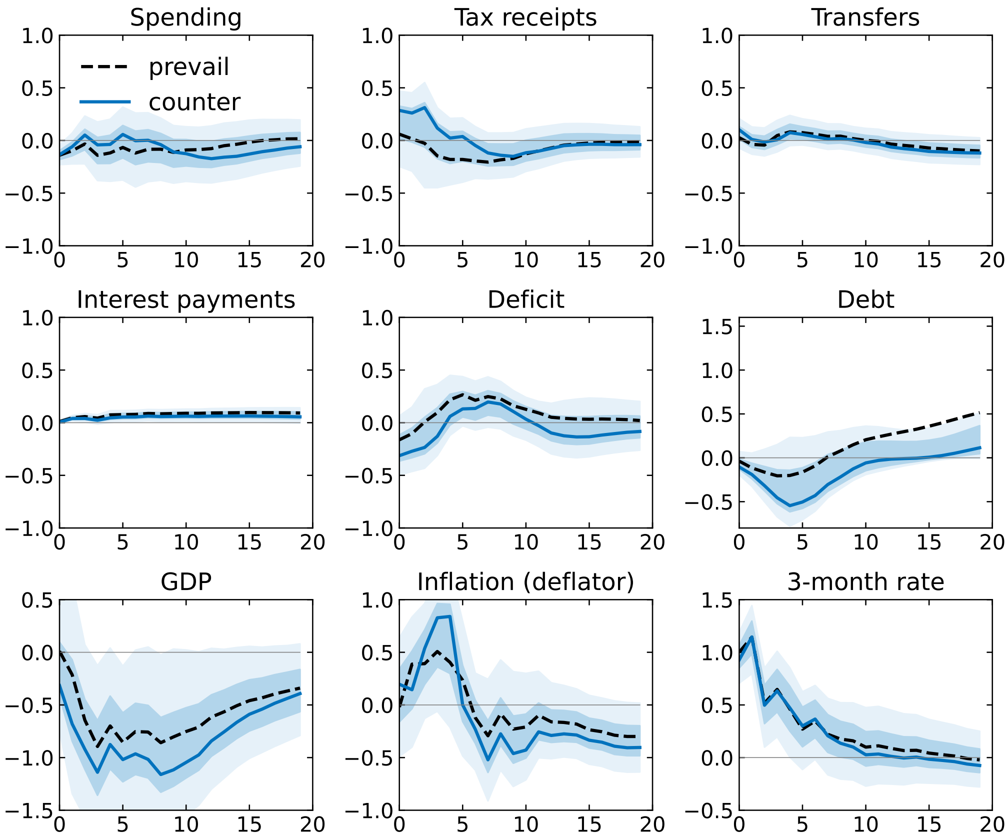
Figure H.4: Cosine similarity of the IRFs for debt



Note: Posterior density of the cosine similarity of the debt IRFs for spending, tax, and transfer shocks. The vertical dashed line is the cosine similarity at the modal IRFs.

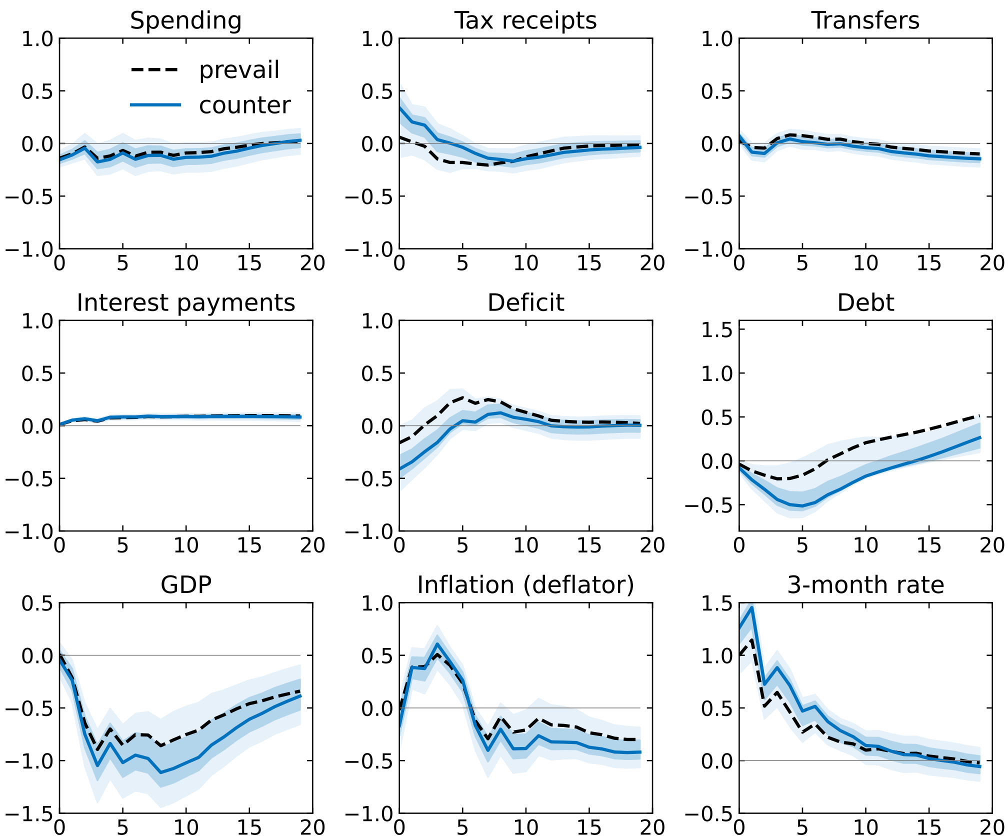
To validate this explanation, we can estimate the same counterfactual scenario (debt stabilization through taxes) with a single shock series (Romer-Romer *or* Caldara–Kamps). This exercise is inherently more limited: with one less shock path, the approximation of the counterfactual rule can only be worse. In particular, we are unlikely to be able to stabilize the whole time path of debt. To handle this difficulty, we adopt a quadratic weighting scheme in equation (14). Debt being a stock, we put more weight on the long-run response of debt than on short run deviations from the rule. Formally, we weight the deviation from the rule at time t by a factor of $(t + 1)^2$. We show the results in figures H.5–H.6. Taxes increase and GDP falls, although debt is imperfectly stabilized.

Figure H.5: counterfactual—debt stabilization with Romer–Romer tax shocks only



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through taxes. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual is constructed with the MKW method (section 4.1). Loadings are updated for each draw of the posterior distribution. Compared to figures H.2, we only use the Romer–Romer shock and weight each deviation from the rule by the square of its time horizon. See section H for a detailed discussion.

Figure H.6: counterfactual—debt stabilization with Caldara–Kamps tax shocks only



Note: Counterfactual response of the economy to a monetary shock if the government stabilizes debt through taxes. The dashed black line is the actual response under the prevailing rule. The blue line and shaded areas are the counterfactual scenario and its 68% (dark) and 90% (light) credible intervals. This counterfactual is constructed with the method of MKW (section 4.1). Loadings are updated for each draw of the posterior distribution. Compared to figures H.2, we only use the Caldara–Kamps shock and weight each deviation from the rule by the square of its time horizon. See section H for a detailed discussion.

I Detailed Description of the Shocks

I.1 Monetary Shocks

Romer–Romer+ monetary shocks: see section 2.1.

Aruoba–Drechsel monetary shocks: see section 2.1.

Bauer–Swanson monetary shocks: see section 2.1.

Jarocinski-Karadi monetary shocks: Jarociński and Karadi (2020) are concerned with the information effect of monetary policy (Romer and Romer, 2000, Nakamura and Steinsson, 2018). If the Federal Reserve has superior information about the state of the economy, asset price changes around monetary policy decisions not only reflect true monetary shocks, they also convey information about current and future economic fundamentals. Jarociński and Karadi decompose each surprise into its effect on the first principal component of the first four federal funds futures and the S&P 500 index. They propose two approaches. In the first one, they impose sign restrictions in a structural VAR: true monetary shocks imply a negative co-movement between federal funds futures and stocks, while central bank information effect imply a positive one. In the second one ("poor man's"), monetary shocks are simply asset price changes on FOMC days when federal funds futures and stocks moved in different directions. We follow the poor man's approach here.

Miranda-Agrippino–Ricco+ monetary shocks: Miranda-Agrippino and Ricco (2021) tackle the information effect by running a Romer–Romer regression with the change in the fourth federal funds future on the left-hand side. We follow their procedure, the only difference being the addition of fiscal forecasts. We start by running

$$\begin{aligned}\Delta FF4_m = \alpha_0 &+ \sum_{q=-1}^{3} \gamma^q \Delta\tilde{y}_m^q + \sum_{q=-1}^{2} \zeta^q \left(\Delta\tilde{y}_m^q - \Delta\tilde{y}_{m-1}^q\right)\\ &+ \sum_{q=-1}^{3} \eta^q \tilde{\pi}_m^q + \sum_{q=-1}^{2} \theta^q \left(\tilde{\pi}_m^q - \tilde{\pi}_{m-1}^q\right)\\ &+ \iota\tilde{u}_m^0 + \sum_{q=-1}^{2} \tau^q \left(\tilde{u}_m^q - \tilde{u}_{m-1}^q\right)\\ &+ \sum_{q=-1}^{3} \kappa^q \Delta\tilde{\text{rec}}_m^q + \sum_{q=-1}^{2} \lambda^q \left(\Delta\tilde{\text{rec}}_m^q - \Delta\tilde{\text{rec}}_{m-1}^q\right)\\ &+ \sum_{q=-1}^{3} \mu^q \Delta\tilde{\text{exp}}_m^q + \sum_{q=-1}^{2} \nu^q \left(\Delta\tilde{\text{exp}}_m^q - \Delta\tilde{\text{exp}}_{m-1}^q\right)\end{aligned} \tag{I.1}$$

$$+\sum_{q=-1}^{3}\pi^{1,q}\tilde{\text{heb}}_m^{1,q}+\sum_{q=-1}^{2}\rho^{1,q}\left(\tilde{\text{heb}}_m^{1,q}-\tilde{\text{heb}}_{m-1}^{1,q}\right)+\xi_m^{MAR+},$$

where $\Delta FF4_m$ is the change in the high-frequency change in the fourth Federal Funds rate future at meeting m and other notations are those of equation (1). Then, we sum the estimated residuals within months and run a regression of this new object, $\bar{\xi}_t^{MAR+}$, on 12 lags of itself to purge serial correlation from the shocks:

$$\bar{\xi}_t^{MAR+}=\phi_0+\sum_{l=1}^{12}\phi_l\bar{\xi}_{t-l}^{MAR+}+\epsilon_t^{MAR+}. \tag{I.2}$$

In equation (I.2), months without a meeting are filled with 0 on the right-hand side but do not enter as separate observations. ϵ_t^{MAR+} is set to 0 in those months. The sample for the estimation of the shocks is 1990–2009. The stopping point is determined by the onset of the zero lower bound. Once again, this procedure is exactly the one followed by Miranda-Agrippino and Ricco. The only difference is the addition of fiscal forecasts in equation (I.1).

I.2 Spending Shocks

Ramey spending shocks: Ramey (2011) reads periodicals, mainly *Business Week*, to measure the public's expectations of future military spending. The shock is an estimate of the change in the present discounted value of future spending.

Blanchard–Perroti spending shocks: Blanchard and Perotti (2002) run a VAR with spending, taxes, and GDP. To identify the spending shock, they assume that spending doesn't respond contemporaneously to taxes or GDP, a so-called Cholesky decomposition of the covariance matrix of the reduced-form shocks. We reproduce their approach in spirit by ordering spending first after the Ramey shocks in our 9-variable VAR. Hence, our version of the Blanchard–Perroti spending shocks controls for contemporaneous values of the Ramey ones.

I.3 Tax Shocks

Romer–Romer tax shocks: Romer and Romer (2010) read presidential speeches and Congressional reports to assess the motivation of changes in the tax code. They distinguish four motivations: (i) finance extra spending, (ii) fight a recession, (iii) remedy an inherited deficit, and (iv) spur long-run growth. In their exercise, the last two categories are treated as exogenous. Since we are interested in the effect on fiscal variables, (iii) is problematic, so we only use (iv).

Mertens and Ravn revisit these shocks in a series of papers to study anticipations (Mertens and Ravn, 2012), personal versus corporate income taxes (Mertens and Ravn, 2013), or the implied tax multiplier (Mertens and Ravn, 2014). While the distinction between anticipated and unanticipated tax shocks seems promising in the context of the MKW method–it implies different paths for the fiscal instrument—the anticipated tax changes vary so much in their anticipation horizon that trying to estimate their effect based on their announcement date has little statistical power.[8] We also experimented with personal and corporate tax changes, but those have similar implications for deficit and debt hence did not materially affect our ability to implement the desired counterfactual policy. Finally, the Mertens-Ravn "exogenous" series lump together motivations (iii) and (iv), which, once again, is problematic in our context since motivation (iii) (remedying an inherited deficit) is endogenous to fiscal variables. These experiments and limitation led us to settle for the subset of the original Romer–Romer tax shocks described in the previous paragraph.

Caldara–Kamps tax shocks: Caldara and Kamps (2017) run a VAR with some fiscal and macroeconomic variables. Denoting the vector of reduced-form shocks by u_t, they assume an invertible mapping $A(0)'$ between u_t and the vector of structural shocks e_t:

$$u_t = (A(0)')^{-1} e_t. \tag{CK5}$$

(CK5) stands for Caldara and Kamps's equation (5). Just as in section E.4, this equation can be rearranged into:

$$u_t^p = \psi_0 u_t^{np} + \omega_p e_t^p, \tag{CK8}$$

where superscripts p and np respectively denote the policy and non-policy variables. (Here, "non-policy" here should be interpreted as non-fiscal policy.) Equation (CK8) is a fiscal rule that ties innovations in a policy variable (e.g. taxes) to innovations in non-policy variables (e.g. GDP) and structural policy shocks (e.g. exogenous change in tax rates). To identify the structural policy shock e_t^p, Caldara and Kamps propose regressing u_t^p on instrumental variables for the non-policy variables, thus identifying the systematic component of the fiscal rule. Their non-policy variables are GDP, inflation, and the nominal interest rate.

They consider two cases. In the "simple fiscal rule" case, they assume that taxes do not respond contemporaneously to inflation and the interest rate, hence the corresponding rows of ψ_0 are set equal to 0. In the "general fiscal rule" case, they allow taxes to respond to

[8] This limitation led Mertens and Ravn to date the shocks at implementation and study the behavior of the economy before said implementation.

all non-policy variables. As instrumental variables, they use the Fernald (2014) measure of total factor productivity adjusted for factor utilization for GDP, the Hamilton (2003) series of oil price shocks for inflation, and the Romer and Romer (2004) series of monetary policy shocks for the interest rate. In the baseline, we assume the "full fiscal rule." We replace the Hamilton oil price shocks with the more recent series of Känzig (2021)—Hamilton's series are a weak instrument for oil prices (Stock and Watson, 2012)—and the Romer–Romer monetary shocks with those of Gertler and Karadi (2015) to avoid using the Romer–Romer series twice. At quarterly frequency, the Gertler–Karadi series are a stronger instrument than the refined monetary shocks explained in section I.1. We discuss the Fernald, Känzig, and Gertler–Karadi in detail in section I.5. In figures A.20 and A.21, we use the "simple fiscal rule" instead. Finally, the policy shocks, e_t^p, are identified by orthogonality. The ordering of the non-policy variables does not matter, but the one of the policy variables does. In the baseline, we put spending first, transfers second, and taxes third. That is, we assume that spending does not contemporaneously respond to transfers and taxes, and transfers do not contemporaneously respond to taxes. We swap the ordering of transfers and taxes in figures A.22 and A.23. Note that, like with the Blanchard–Perroti spending shocks, we re-estimate those shocks within our VAR.

I.4 Transfer Shocks

Romer–Romer transfer shocks: Romer and Romer (2016) study Social Security increases from 1952 to 1991. Like with tax changes, they read administrative documents to understand what motivated these increases. They argue that most of the increases "occurred somewhat randomly". Until 1974, Social Security payments weren't automatically adjusted for inflation. Until the early 1990s, "substantial variation in inflation and occasional bursts of retroactive payments resulting from idiosyncratic factors, as well as a legislated change in the timing of cost-of-living adjustments, led to irregular and variable benefit changes" (p. 1). They find a few changes made for countercyclical purposes, which they exclude. They distinguish between permanent and temporary changes. We only use the permanent changes as the temporary ones have no power in quarterly data.

Caldara–Kamps transfer shocks: see the discussion of the Caldara–Kamps tax shocks in section I.3. Caldara and Kamps did not estimate the effect of transfer shocks. We apply the methodology that they proposed for taxes to transfers.

I.5 Instruments for Caldara–Kamps Identification

Fernald TFP shocks: Fernald (2014) estimates total factor productivity, adjusted for utilization. While replicating Caldara and Kamps's (2017) approach, we noticed that the version of the Fernald (2014) series that is in their replication package differs from what can be downloaded from John Fernald's website.[9] The Fernald series are updated several times a year and they underwent an important revision in 2014, sometimes threatening the results of earlier papers (Cascaldi-Garcia, 2017). We were unable to understand which version Caldara and Kamps use or whether they applied a transformation. In our baseline, we use the December 2013 vintage from Fernald's website. Subsequent vintages are too weak an instrument to produce meaningful results (see below).[10] In figures A.24–A.25, we show the results if we use the version of the Fernald series that is in Caldara and Kamps's replication package.

Kanzig oil supply news shocks: Känzig (2021) uses the change in the price of oil futures around announcements of the Organization of the Petroleum Exporting Countries (OPEC).

Gertler–Karadi monetary shocks: Gertler and Karadi (2015) use the change in the price of federal funds rate futures around FOMC meetings as a measure of monetary shocks. Early contributors to that identification scheme were: Bagliano and Favero (1999), Cochrane and Piazzesi (2002), Faust et al. (2004), Barakchian and Crowe (2013).

Instrument relevance: We assess the relevance of these instruments in tables I.1.1 – I.1.3. Because the instruments are available over different time periods, we estimate separate first-stage regressions for each instrument, restricting the sample to the period in which the instrument is observable.[11] The Fernald TFP shocks, Kanzig oil shocks, and Gertler–Karadi monetary shocks are available over 1948Q2–2007Q4, 1983Q3–2007Q4, and 1988Q4–2007Q4, respectively. Like Caldara and Kamps (2017, table 2), we find that oil shocks are a weak instrument for inflation (table I.1.2, column 2), while the TFP and monetary shocks are strong instruments for GDP and the nominal interest rate, respectively (table I.1.1, column 1 and table I.1.3, column 3). As noted above, we prefer the 2013 vintage of Fernald's TFP shocks to more recent vintages, as it is a substantially stronger instrument (table I.1.1, column 1 and 2). We also report the same regression with the version of Fernald's shock from

[9] See `www.frbsf.org/economic-research/indicators-data/total-factor-productivity-tfp/` for the latest version or `www.johnfernald.net/TFP` for a version history.

[10] Kurmann and Sims (2021) explain that the most consequential innovation in the 2014 revision is a change in the method for de-trending industry hours per worker, upon which the measure for factor utilization is based. As a result of this change, the high-frequency fluctuations become unfiltered, potentially including cyclical measurement errors.

[11] In the same spirit, Caldara and Kamps run separate first-stage regressions for the monetary shock instrument and other instruments.

Caldara and Kamps's online appendix (table I.1.1, column 3). We also check the Hamilton oil shocks used by Caldara and Kamps and find that they are even weaker than the Kanzig oil shocks for inflation (table I.1.2, column 3).

Table I.1.1: Instrument relevance for Caldara–Kamps identification—GDP

	(1) GDP	(2) GDP	(3) GDP	(4) GDP	(5) GDP
			Tax		
Fernald TFP 13	0.0013				
	(0.0001)				
Fernald TFP 23		0.0004			
		(0.0001)			
Fernald TFP CK			0.0008		
			(0.0001)		
Kanzig Oil				-0.0003	
				(0.0003)	
Gertler-Karadi					0.0030
					(0.0050)
F-statistic	92.62	9.06	38.88	1.54	0.35
			Transfers		
Fernald TFP 13	0.0012				
	(0.0001)				
Fernald TFP 23		0.0004			
		(0.0001)			
Fernald TFP CK			0.0007		
			(0.0001)		
Kanzig Oil				-0.0004	
				(0.0003)	
Gertler-Karadi					0.0031
					(0.0049)
F-statistic	81.82	7.31	34.06	1.93	0.40

Note: Relevance of various instruments for Caldara–Kamps identification. See section I for more details. The tax and transfer VARs differ only because they don't feature the same narrative shock. The number in parenthesis is the standard error.

Table I.1.2: Instrument relevance for Caldara–Kamps identification—inflation

	(1) Inflation	(2) Inflation	(3) Inflation	(4) Inflation
		Tax		
Fernald TFP 13	-0.0007			
	(0.0003)			
Kanzig Oil		0.0004		
		(0.0005)		
Hamilton Oil			0.0001	
			(0.0002)	
Gertler-Karadi				-0.0093
				(0.0068)
F-statistic	6.64	0.87	0.57	1.84
		Transfers		
Fernald TFP 13	-0.0005			
	(0.0003)			
Kanzig Oil		0.0004		
		(0.0004)		
Hamilton Oil			0.0001	
			(0.0002)	
Gertler-Karadi				-0.0116
				(0.0068)
F-statistic	4.23	0.63	0.32	2.85

Note: Relevance of various instruments for Caldara–Kamps identification. See section I for more details. The tax and transfer VARs differ only because they don't feature the same narrative shock. The number in parenthesis is the standard error.

Table I.1.3: Instrument relevance for Caldara–Kamps identification—interest rate

	(1) Nominal i.r.	(2) Nominal i.r.	(3) Nominal i.r.
	Tax		
Fernald TFP 13	0.0001		
	(0.0001)		
Kanzig Oil		-0.0001	
		(0.0002)	
Gertler-Karadi			0.0155
			(0.0032)
F-statistic	1.46	0.05	23.96
	Transfers		
Fernald TFP 13	0.0001		
	(0.0001)		
Kanzig Oil		-0.0001	
		(0.0002)	
Gertler-Karadi			0.0156
			(0.0032)
F-statistic	1.36	0.08	23.90

Note: Relevance of various instruments for Caldara–Kamps identification. See section I for more details. The tax and transfer VARs differ only because they don't feature the same narrative shock. The number in parenthesis is the standard error.

References

Antolín-Díaz, J., I. Petrella, and J. Rubio-Ramírez (2025). Dividend momentum and stock return predictability: A Bayesian approach. *Review of Financial Studies*. Forthcoming.

Bagliano, F. C. and C. A. Favero (1999). Information from financial markets and VAR measures of monetary policy. *European Economic Review 43*, 825–837.

Barakchian, S. M. and C. Crowe (2013). Monetary policy matters: Evidence from new shocks data. *Journal of Monetary Economics 60*, 950–966.

Bianchi, F., R. Faccini, and L. Melosi (2023). A fiscal theory of persistent inflation. *Quarterly Journal of Economics 138*(4), 2127–2179.

Blanchard, O. and R. Perotti (2002). An empirical characterization of the dynamic effects of changes in government spending and taxes on output. *Quarterly Journal of Economics 117*(4), 1329–1368.

Caldara, D. and C. Kamps (2017). The analytics of SVARs: A unified framework to measure fiscal multipliers. *Review of Economic Studies 84*, 1015–1040.

Campbell, J. Y. and R. J. Shiller (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies 1*(3), 195–228.

Cascaldi-Garcia, D. (2017). News shocks and the slope of the term structure of interest rates: Comment. *American Economic Review 107*(10), 3243–3249.

Cochrane, J. H. and M. Piazzesi (2002). The Fed and interest rates: A high-frequency identification. *American Economic Review 92*(2), 90–95.

Croushore, D. and S. van Norden (2018). Fiscal forecasts at the FOMC: Evidence from the Greenbooks. *Review of Economics and Statistics 100*(5), 933–945.

Del Negro, M. and F. Schorfheide (2012). Bayesian macroeconometrics. In J. Geweke, G. Koop, and H. Van Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, Chapter 7, pp. 293–389. Oxford, UK: Oxford University Press.

Durbin, J. and J. Koopman (2012). *Time Series Analysis by State Space Methods* (Second ed.). Oxford, UK: Oxford University Press.

Faust, J., E. T. Swanson, and J. H. Wright (2004). Identifying VARS based on high frequency futures data. *Journal of Monetary Economics 51*, 1107–1131.

Federal Reserve Board (2025). Z.1: Financial accounts of the United States – All table descriptions. Technical report. `https://www.federalreserve.gov/apps/fof/Guide/z1_tables_description.pdf`.

Fernald, J. (2014). A quarterly, utilization-adjusted series on total factor productivity. FRB San Francisco Working Paper No. 2012-19.

Gertler, M. and P. Karadi (2015). Monetary policy surprises, credit costs, and economic activity. *American Economic Journal: Macroeconomics 7*(1), 44–76.

Gilchrist, S. and E. Zakrajšek (2011). Credit spreads and business cycle fluctuations. *American Economic Review 102*(4), 1692–1720.

Giordani, P., M. Pitt, and R. Kohn (2012). Bayesian inference for time series state space models. In J. Geweke, G. Koop, and H. Van Dijk (Eds.), *The Oxford Handbook of Bayesian Econometrics*, Chapter 3, pp. 60–124. Oxford, UK: Oxford University Press.

Gupta, A. K. and D. K. Nagar (2018). *Matrix Variate Distributions*. Taylor & Francis.

Hall, G., J. Payne, and T. Sargent (2018). Us federal debt 1776–1960: Quantities and prices. Working paper.

Hamilton, J. D. (2003). What is an oil shock? *Journal of Econometrics 113*(2), 363–398.

Jarociński, M. and P. Karadi (2020). Deconstructing monetary policy surprises—The role of information shocks. *American Economic Journal: Macroeconomics 12*(2), 1–43.

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences 186*(1007), 453–461.

Krantz, T. and A. Jonker (2025). What is cosine similarity? `https://www.ibm.com/think/topics/cosine-similarity`. IBM Think article.

Kurmann, A. and E. Sims (2021). Revisions in utilization-adjusted tfp and robust identification of news shocks. *Review of Economics and Statistics 103*(2), 216–235.

Känzig, D. R. (2021). The macroeconomic effects of oil supply news: Evidence from opec announcements. *American Economic Review 111*(4), 1092–1125.

McKay, A. and C. K. Wolf (2023). What can time-series regressions tell us about policy counterfactuals? *Econometrica 91*(5), 1695–1725.

Mertens, K. and M. O. Ravn (2012). Empirical evidence on the aggregate effects of anticipated and unanticipated US tax policy shocks. *American Economic Journal: Economic Policy 4*(2), 145–181.

Mertens, K. and M. O. Ravn (2013). The dynamic effects of personal and corporate income tax changes in the United States. *American Economic Review 103*(4), 1212–1247.

Mertens, K. and M. O. Ravn (2014). A reconciliation of SVAR and narrative estimates of tax multipliers. *Journal of Monetary Economics 68*, S1–S19.

Miranda-Agrippino, S. and G. Ricco (2021). The transmission of monetary policy shocks. *American Economic Journal: Macroeconomics 13*(2), 74–107.

Nakamura, E. and J. Steinsson (2018). High-frequency identification of monetary non-neutrality: The information effect. *Quarterly Journal of Economics 133*(3), 1283–1330.

Payne, J., B. Szőke, G. Hall, and T. J. Sargent (2025, 12). Costs of financing U.S. federal debt under a gold standard: 1791–1933. *Quarterly Journal of Economics 140*(1), 793–833.

Plagborg-Møller, M. and C. K. Wolf (2021). Local projections and VARs estimate the same impulse responses. *Econometrica 89*(2), 955–980.

Ramey, V. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics 126*(1), 1–50.

Ramey, V. A. (2016). Macroeconomic shocks and their propagation. In J. B. Taylor and H. Uhlig (Eds.), *Handbook of Macroeconomics*, Volume 2A, pp. 71–162. Amsterdam: Elsevier.

Romer, C. D. and D. H. Romer (2000). Federal reserve information and the behavior of interest rates. *American Economic Review 90*(3), 429–457.

Romer, C. D. and D. H. Romer (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review 94*(4), 1055–1084.

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Romer, C. D. and D. H. Romer (2016). Transfer payments and the macroeconomy: The effects of Social Security benefit increases, 1952–1991. *American Economic Journal: Macroeconomics 8*(4), 1—42.

Salton, G. and M. J. McGill (1983). *Introduction to Modern Information Retrieval*. New York: McGraw-Hill.

Stock, J. H. and M. W. Watson (2012). Disentangling the channels of the 2007-–09 recession. *Brookings Papers on Economic Activity 42*(1), 81–135.

Vehtari, A., A. Gelman, D. Simpson, B. Carpenter, and P.-C. Bürkner (2021). Rank-normalization, folding, and localization: An improved $\hat{R}$ for assessing convergence of MCMC (with discussion). *Bayesian Analysis 16*(2), 667–718.