

INTERNATIONAL MONETARY FUND

Public Debt Dynamics and the Impact of Fiscal Policy

Nikhil Patel and Adrian Peralta-Alva

WP/24/87

***IMF Working Papers* describe research in progress by the author(s) and are published to elicit comments and to encourage debate.**

The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

**2024
APR**



WORKING PAPER

IMF Working Paper
Research Department

Public Debt Dynamics and the Impact of Fiscal Policy
Prepared by Nikhil Patel and Adrian Peralta-Alva*

Authorized for distribution by Prachi Mishra
April 2024

IMF Working Papers describe research in progress by the author(s) and are published to elicit comments and to encourage debate. The views expressed in IMF Working Papers are those of the author(s) and do not necessarily represent the views of the IMF, its Executive Board, or IMF management.

ABSTRACT: Public debt-to-GDP ratios have undergone substantial fluctuations over both the short and long term. Most recently, global debt-to-GDP ratios peaked at 100% on average in 2020 due to COVID-19, retracting substantially by 2022. To understand what drives these movements, we propose a structural approach to debt decompositions based on a SVAR identified with narrative sign restrictions. We find that GDP growth shocks and the corresponding comovements of macroeconomic variables are the key drivers of debt to GDP, accounting for 40% of the observed yearly variation in 17 advanced economies since the 1980s. Discretionary fiscal policy changes, in turn, account for less than 20% of the observed changes. The analysis also finds the primary balance multiplier on GDP to be very small. We reconcile our results with the literature, underscoring the importance of accurate shock identification and accounting for cross-country heterogeneity.

RECOMMENDED CITATION: Patel, Nikhil and Adrian Peralta-Alva (2024) "Public Debt Dynamics And the Impact of Fiscal Policy" IMF Working Paper WP/24/87, International Monetary Fund, Washington, D.C.

JEL Classification Numbers:	C30, E32, E60, H60.
Keywords:	Public Debt; Fiscal Consolidation; Structural Vector Autoregression.
Author's E-Mail Address:	npatel@imf.org ; aperalta-alva@imf.org

* We are grateful to Julien Acalin, Marina Azzimonti, Filippo Ferroni, Carlos Goncalves, Luciana Juvenal, Daniel Leigh, Hanno Lustig, Prachi Mishra, Emi Nakamura, Ivan Petrella, Francisco Roch, Juan Rubio Ramirez, Martin Uribe and seminar participants at the Society of Economic Dynamics annual meetings, Midwest Econometrics Group, Midwest Macro Meetings, IMF and the SNDE and Canadian Economic Association annual meetings for helpful comments and discussions and to Swapnil Aggarwal, Sergio Garcia Monroy, Carlos Angulo Saborio and Chenxu Fu for excellent research assistance.

Contents

1	Introduction	2
2	Empirical Approach	6
2.1	Structural Shocks and Historical Decompositions	6
2.2	Identification Using Narrative Sign Restrictions	7
3	Data and Estimation	7
3.1	Model Specification	7
3.2	Estimation	9
4	Results	10
4.1	Structural Debt Decompositions	10
4.2	Dynamic Effects of Structural Shocks	12
4.3	The Value of Narrative Restrictions	13
5	Towards a Reconciliation with Contractionary Austerity Results in the Literature	14
5.1	The Problem of Foresight and Invertibility	15
5.2	The Role of Heterogeneity	17
5.3	Proxy SVARs and the Tax Multiplier	18
6	Robustness: Priors, Global Shocks, Unexpected Inflation, and Predictability of our Narrative Data	18
7	Conclusion	19

1 Introduction

Public debt-to-GDP ratios have undergone substantial fluctuations over both the short and long term. The recent COVID-19 pandemic led to a dramatic escalation, pushing the global average ratio to nearly 100% in 2020—a 15-percentage-point increase from the previous year (Figure 1). Remarkably, the subsequent years, 2021 and 2022, saw these ratios contract, reversing approximately half of the 2020 surge. This contraction was facilitated by economic recovery, fiscal policy normalization, and inflation-induced debt deflation.

This recent volatility is part of a broader historical pattern. For instance, the 1980s saw debt-to-GDP ratios rise substantially due to policy-driven welfare expansions, tax reforms, and fiscal stimuli. Conversely, the 1990s experienced declines as a result of targeted fiscal reforms and favorable economic conditions. The 2008-2009 financial crisis brought another surge, highlighting the perpetual struggle between fiscal stimulus and sustainability.

A substantial body of empirical literature has focused on fiscal multipliers. These studies have introduced advanced methods to address endogeneity and identification issues, aiming to understand the impact of fiscal policy on key macroeconomic outcomes. Our paper is, however, the first to use those methods to focus on debt to GDP. This is significant as the connection between debt-to-GDP and macroeconomic outcomes is not straightforward. The standard debt accumulation equation linking changes in debt to the primary balance, interest rates, and GDP growth, does not typically hold in the data. Large residuals frequently arise, prompted by financial operations below the line, exchange rate changes and other valuation effects. Furthermore, debt ratios do not merely respond to GDP fluctuations. In fact, debt displays higher volatility compared to GDP.

Our study aims to identify the determinants of changes in public debt-to-GDP ratios in advanced economies since the 1980s. Similar in spirit to the pioneering work of [Mertens and Ravn \(2013\)](#), our approach exploits the benefits of both, SVARs and the information in narrative data. Our core SVAR is based on sign restrictions for growth and fiscal shocks following the methodology of [Mountford and Uhlig \(2009\)](#). The setup is then extended with narrative sign restrictions on the primary balance shock, in line with the approach of [Antolín-Díaz and Rubio-Ramírez \(2018\)](#). Narrative shocks are derived from a specific subset of the literature that identifies shifts in the primary balance unrelated to the business cycle, such as [Guajardo et al. \(2014\)](#), [Gunter et al. \(2021\)](#), and [Carrière-Swallow et al. \(2021\)](#).

The Structural Vector Autoregression (SVAR) we employ includes six variables: GDP growth, government revenues, primary balance, debt to GDP, inflation, and the effective interest rate on debt. The SVAR accounts for three distinct shocks: a demand-driven GDP growth shock, a supply-driven GDP growth shock, and a primary balance shock orthogonal to the others. The latter encapsulates discretionary primary balance consolidations or expansions.

The structural shocks we consider account jointly for around 60% of the variations in the debt-to-GDP ratio. We find that GDP growth shocks, and the average comovements of GDP with other relevant macro and fiscal variables in response to these shocks are the major contributors. GDP shocks explain roughly 40% of the yearly variation in debt to GDP ratios for the median advanced economy. Discretionary policy shocks on the other hand explain less than a fifth. In the baseline model, we do not explicitly incorporate monetary policy

shocks. Unexpected inflation has been important in driving debt to GDP in certain periods and countries including the US.¹ For the period we consider this has been less relevant. More importantly, an extension of the model to explicitly allow for monetary policy shocks reveals that our main results remain unchanged.

Our findings also diverge from prior research as we find a small impact of discretionary fiscal consolidations on GDP growth (i.e primary balance multiplier). We argue there are two reasons for this. Firstly, compared to other approaches, our SVAR equipped with narrative sign restrictions more effectively addresses the foresight problem discussed by [Ramey \(2016\)](#). Secondly, our methodology offers a more flexible framework to account for variations across countries.

Empirical research on fiscal policy often faces the 'foresight problem'. This arises when private agents anticipate policy changes or react in advance to what an econometrician would recover as exogenous shocks in the future. Such anticipated shocks can lead to non-invertible representations in the SVAR, rendering the shocks non-fundamental. [Beaudry et al. \(2019\)](#) introduce a diagnostic for assessing the significance of non-fundamentality. Based on this diagnostic, our methodology and identification approach seems to outperform other techniques, like direct use of narrative shocks or an SVAR identified solely through sign restrictions. We further demonstrate that when we (i) use the structurally-identified SVAR shocks from our method into a standard local projection approach, or (ii) employ our SVAR shocks as instrumental variables for narrative shocks in a local projection method, a null multiplier is observed, consistent with our baseline estimates. This, given the conditions for equivalence of SVARs and local projections, shows that the distinct outcomes of our paper stem from our differing approaches to identification.

We further assess the validity of our methodology following [Cochrane \(2004\)](#)'s approach, and provide evidence of the orthogonality of our VAR shocks relative to macroeconomic forecasts. External validation of the performance of our method is obtained through two additional experiments. First, we apply our methodology to the study of the impact of tax increases in the US., reproducing the high negative multiplier found by [Mertens and Ravn \(2013\)](#). Second, we find that periods of exogenous fiscal expansions identified by our method align with those of [Ben Zeev et al. \(2023\)](#) that are based on military spending news, even though we do not include any narrative restrictions on fiscal expansions in our estimation as cross-country narrative shocks are only available for fiscal consolidations.

Our method also deals directly with cross-country heterogeneity, deviating from previous studies that predominantly relied on panel regression models and potentially overlooking nuanced country-specific differences. Applying local projections to individual countries (with narrative data directly), we find a distribution (across countries) of primary balance multipliers centered around zero, consistent with our baseline SVAR estimates. Yet, further scrutiny shows that the real divergence between our results and those in existing studies primarily hinges on the better performance of our identification approach to deal with the problem of foresight. When we again employ the SVAR-derived shocks in country-specific local projections we find estimates for the multiplier that are highly correlated with ours.

Our research provides insights that challenge and refine the prevailing understanding on the drivers of public debt-to-GDP ratios. We find that the empirical reaction of the

¹See for instance [Acalin and Ball \(2023\)](#) and [Hall and Sargent \(2011\)](#)

macroeconomy and policy variables to GDP shocks play a foundational role in accounting for observed changes in debt to GDP. Given debt to GDP ratios are so high, our analysis suggests a careful look at whether the current policy calibrations (and implicit reactions of policy to shocks) are adequate is granted. Additionally, our study emphasizes the limited impact of fiscal consolidations on these ratios, pointing instead to the significance of debt reduction during periods of growth. A detailed analysis on policies to reduce public debt to GDP ratios is considered by [Ando et al. \(2023\)](#).

Another important dimension of our findings is the 'zero multiplier' effect. This highlights the importance of methodologies that effectively tackle challenges like endogeneity and the foresight problem. By addressing these challenges more robustly, our study provides more nuanced quantitative outcomes. The implications extend beyond just understanding public debt dynamics; they call for a re-evaluation of current fiscal policy strategies and underline the need for methodologies that accurately capture the intricacies of fiscal and economic interactions.

The remainder of this paper is organized as follows. The remainder of this section discusses the related literature. Section 2 presents the methodology. Section 3 discusses the data and estimation setup. Section 4 presents the main results, which are followed by a few diagnostics and extensions in Sections 5 and 6. Section 7 concludes with a summary of the main findings and highlights avenues for future research.

Literature Review

This paper is linked to three strands of the literature. First, to studies on the drivers of sovereign debt. Second, to SVAR approaches analyzing the effect of macroeconomic shocks. And, third, to the literature on fiscal multipliers and macroeconomic effects of fiscal consolidations.

The literature on drivers of sovereign debt has focused largely on proximate drivers that emerge from a mechanical decomposition of changes in debt to GDP ratios into various subcomponents such as interest expenses, primary balance and GDP growth, based on identities. Recent examples of this approach can be found for instance in [Cochrane \(2019\)](#) and [Hall and Sargent \(2011\)](#) for the US and [Das and Ghate \(2022\)](#), for India.

While these estimates provide an accurate accounting of the proximate drivers, they are silent on the fundamental primitive shocks that ultimately drive debt, as well as on causality. For instance, a mechanical decomposition may reveal that debt fluctuations are driven to a large extent by the primary balance, but may fail to capture that the primary balance itself is driven by more fundamental business cycle or commodity price shocks. Our approach is geared towards addressing these issues.

The second point of intersection is with the literature on structural vector auto regressions for identifying and studying the impact of macroeconomic shocks. In this context, our work is most closely related to [Mountford and Uhlig \(2009\)](#) who study the impact of discretionary fiscal actions on GDP using sign restrictions. The authors find, focusing only on US. data, a sizable multiplier-especially when the consolidation is driven by higher taxes. But the paper is silent about the effects of fiscal shocks on the debt-to-GDP ratio. We build on their framework by combining it with the narrative sign restrictions approach of [Antolín-Díaz and Rubio-Ramírez \(2018\)](#). In particular, we constrain the sign of the primary balance shock in

our setup to be consistent with the narrative evidence documented in the literature. [Cherif and Hasanov \(2018\)](#) also examine the impact of fiscal consolidations on debt ratios using a VAR with a non-linear debt accumulation equation and a different identification strategy, finding a moderate effect on the debt ratio similar to our results.

Finally, our paper links with a vast literature dealing with the macroeconomic impact of fiscal consolidations and expansions focusing mostly on the impact on GDP. Contrary to the new Keynesian result of fiscal consolidations being contractionary, a sizable literature has documented significant instances of expansionary fiscal austerity. For instance [Alesina and Ardagna \(2010\)](#) uncover several episodes in which spending cuts adopted to reduce deficits were associated with economic expansions rather than recessions. [Giavazzi and Pagano \(1990\)](#) uncovered similar evidence for expansionary austerity looking at episodes in Denmark and Ireland, and [Alesina and Perotti \(1997\)](#) and [Alesina and Ardagna \(1998\)](#) reach similar conclusions for a broader set of countries.

These findings on expansionary austerity were challenged by [Guajardo et al. \(2014\)](#). They argued that even the cyclically adjusted primary balance and related measures typically used in the literature are not immune from capturing fluctuations in fiscal deficits that can be attributed to responses to macroeconomic conditions. Instead, they build on the seminal work of [Romer and Romer \(2010\)](#) for the US by compiling a narrative database covering several countries to identify instances of primary balance adjustments motivated solely by a desire to reduce the budget deficit and ensure long-term public financial sustainability. They draw on a rich and diverse set of documents to collect this narrative information, including country specific (central bank and fiscal budget records) as well as multilateral sources (including OECD economic surveys and IMF staff reports). They conclude that in contrast to the literature emphasising expansionary austerity, the impact of fiscal austerity is highly contractionary.

[Jordà and Taylor \(2016\)](#) illustrate that the narrative shocks used by [Guajardo et al. \(2014\)](#) are predictable, and propose a propensity score matching approach to isolate the impact of exogenous fiscal consolidations. Upon doing so, they still find fiscal consolidations to be contractionary, but mostly in downturns. This approach however relies on leveraging a full panel data set and assuming homogeneity across countries, which, as we show, can still result in substantial biases.

The closest paper in spirit to ours is [Mertens and Ravn \(2013\)](#), who develop an estimation strategy mixing SVARs and narrative information. Their identification is based on assuming that narrative measures of tax changes correlate with latent shocks but are orthogonal to other structural shocks. The paper studies the United States only, and the focus is the impact of specific tax instruments (personal income taxes, and corporate income taxes) on GDP. Personal income taxes are found to have important effects on GDP (a large negative multiplier). Our approach also mixes SVARs with narrative information, but we employ narrative sign restrictions for identification. This approach allows us to cover many countries. In addition, historical decompositions can be readily computed, which help us determine the contributions of the different shocks to fluctuations in debt to GDP. Importantly, when we apply our narrative-sign-restrictions based to estimate the tax multiplier in the US, we also find a negative multiplier similar to [Mertens and Ravn \(2013\)](#). This offers further assurances on the reliability of our methods.

The literature on the fiscal multiplier has been mostly silent on the effects on debt to GDP

ratios, and as the previous discussion illustrated, the issue of the impact of fiscal consolidation on output is somewhat unsettled. Our paper provides a novel perspective through the use of narrative sign restrictions on both issues.

In concluding the review of relevant literature, it is pertinent to observe that some recent studies, such as [Barnichon et al. \(2022\)](#), have focused on the influence of individual fiscal policy instruments - such as taxes or spending, and the respective timing and orientation of these measures - on GDP. Contrary to these recent approaches, we center on the primary balance. Our focus is dictated by the primary balance's direct impact on the debt-to-GDP ratio, which forms the crux of our research.

2 Empirical Approach

This section provides a brief summary of how the narrative sign restriction approach can be used to compute historical decompositions and impulse responses analysed in the subsequent sections.² The structural vector autoregression has the general form

$$y'_t A_0 = \sum_{l=1}^p y'_{t-l} A_l + c + \epsilon'_t, 0 < t < T \quad (2.1)$$

where y_t is an n by 1 vector of variables, ϵ_t is an n by 1 vector of structural shocks, A_l is an $n \times n$ matrix of parameters for lags $0 \leq l \leq p$. A_0 is an invertible matrix, c is a $1 \times n$ vector of parameters, p is the lag length and T is the sample size. The vector ϵ_t , conditional on past information and initial conditions y_0, \dots, y_{1-p} is Gaussian with mean zero and covariance matrix I_n , the identity matrix. The model described in equation (2.1) can be written as

$$y'_t A_0 = x'_t A_+ + \epsilon'_t, 0 < t < T \quad (2.2)$$

where $A'_+ = [A'_1, A'_2, \dots, A'_p, c']_{m \times n}$ and $x'_t = [y'_{t-1}, \dots, y'_{t-p}, 1]_{m \times 1}$ for $1 \leq t \leq T$, where $m = np + 1$. the reduced form implied by equation (2.2) is given by:

$$y'_t = x'_t B + u'_t, 0 < t < T \quad (2.3)$$

where $B = A_+ A_0^{-1}$, and $E[u_t u'_t] = \Sigma = [A_0 A'_0]^{-1}$. The matrices B and Σ are reduced form parameters, while A_0 and A_+ are structural parameters. Let $\Theta = (A_0, A_+)$ collect the value of the structural parameters.

2.1 Structural Shocks and Historical Decompositions

Given a value Θ of structural parameters and the data, the structural shocks at time t are defined by

$$\epsilon'_t(\Theta) = y'_t A_0 - x'_t A_+, \text{ for } 1 \leq t \leq T \quad (2.4)$$

²Additional details including a complete description of the methodology are available in [Antolín-Díaz and Rubio-Ramírez \(2018\)](#) from which this section draws.

The historical decomposition computes the contribution of the structural shocks to the observed change in the variables between two periods. Formally, the contribution of the j th structural shock to the observed change in the i th variable between periods t and $t + h$ is given by:

$$H_{i,j,t,t+h}(\Theta) = \sum_{l=0}^h e'_{i,n} L_l(\Theta) e_{j,n} e'_{j,n} \epsilon_{t+h-l}(\Theta) \quad (2.5)$$

where $e_{j,n}$ is the j th column of I_n

2.2 Identification Using Narrative Sign Restrictions

Sign restrictions have emerged as a popular alternative to traditional VAR identification schemes in recent times, since they allow researchers to identify shocks by flexibly imposing minimal restrictions on impulse responses that are grounded in theory. Formally, consider any continuous function $F(\Theta)$ from the structural parameters to the space of r byn matrices, where r is a natural number. Sign restrictions will take the form

$$S_j F(\theta) e_{j,n} > 0 \quad (2.6)$$

for $1 \leq j \leq n$ where S_j is an s_j byn matrix of full row rank, with $0 \leq s_j$. The value of s_j indicates a number of sign restrictions being used to identify the j th. the precise nature of the sign restrictions is contained in S_j and $F(\Theta)$. In our analysis, we employ a version of the narrative sign identification that restricts structural parameters so that the structural shocks are of a particular sign for some dates. Other types of restrictions; including those on magnitudes and relative importance of a particular shock are also possible through the use of this narrative approach ([Antolín-Díaz and Rubio-Ramírez \(2018\)](#)).

3 Data and Estimation

The analysis is based on a panel of yearly data covering seventeen advanced economies from 1981-2019 for which we have narrative information. Annual frequency maximizes the number of countries and years with available data. For the same reason, the level of aggregation for fiscal variables employed is general government as opposed to central.

For years before 2011, the primary source of data is the Historical Public Finance Database (HPFD). For years after 2011 where HPFD is not available, we use the Global Debt Database (GDD) and the IMF's World Economic Outlook (WEO) database.

3.1 Model Specification

Table 1 summarizes the variables used as well as the sign restrictions that are imposed. The SVAR consists of the following six variables that figure prominently in standard debt decomposition identities: (1) Real GDP growth; (2) Growth rate of real government revenues; (3) change in primary balance to GDP; (4) change in debt to GDP; (5) nominal effective interest rate on public debt; and (6) GDP deflator inflation.

The backbone of the identification is a set of sign restrictions that are used to identify three shocks—two GDP shocks (supply and demand) and a discretionary primary balance shock orthogonal to the growth shocks. The identification setup is then sharpened by using narrative restrictions to discipline the third shock for which narrative information is available in the existing literature.

More specifically, we identify two types of GDP growth shocks in the spirit of [Mountford and Uhlig \(2009\)](#) using classical sign restrictions. A demand shock is identified as one which leads output and inflation to move in the same direction on impact, with government revenues reacting procyclically. A supply shock embeds the same procyclical response of revenues, but imposes output and inflation to move in opposite direction.

While similar in spirit, there are some differences from the sign restriction setup of [Mountford and Uhlig \(2009\)](#). First, while they consider only a single shock to GDP, we focus on identifying two separate shocks (supply and demand) in order to quantify and compare their relative contributions to gyrations in growth and debt to GDP.

Second, although we exploit the procyclicality of revenues, our focus is on the primary balance to GDP ratio. This is in line with our emphasis on uncovering the drivers of debt, instead of comparing the relative importance of different fiscal instruments. That said, when drawing comparisons to the literature, we also consider the impact specifically of tax changes using narrative data for the US. Third, while [Mountford and Uhlig \(2009\)](#) work with quarterly data and impose the sign restrictions for four quarters, we impose them for a single period given the annual frequency of our data.

We identify a shock which moves the primary balance in a manner orthogonal to the growth shocks identified above. While this can be done with sign restrictions alone as in [Mountford and Uhlig \(2009\)](#), we augment the identification of this shock by complementing the sign restriction with narrative information from the literature. To do this, we incorporate qualitative information from narrative shocks identified by a series of papers that are precisely aimed at capturing movements in the primary balance unrelated to business cycles. We do this using updated measures (up till 2019) based on the methodology in [Guajardo et al. \(2014\)](#)

Figure 2 illustrates the universe of narrative shocks available in the dataset. In our estimation we use a subset of these as narrative restrictions in the VAR, by only introducing the restriction in the first year in cases where there is a sequence of consecutive years with narrative shocks. For instance, the narrative database codes positive consolidation shocks for Austria in the years 2000 and 2001. In our VAR, we only impose the narrative restriction, i.e. constrain the shock to be positive in year 2000, and leave it unconstrained in 2001. In addition, in line with the commonly established narrative, we impose a negative demand shock in 2008 as an additional narrative restriction.

There are two key reasons for adapting this approach. First, given the rich set of sign restrictions with two additional shocks already imposed in the identification, the scope for incorporating additional restrictions becomes limited. The larger the number of narrative restrictions imposed, the more difficult it is to find a set of candidate solutions that satisfy all the restrictions (both sign and narrative), a problem that is particularly acute when a large number of narrative restrictions are imposed. The literature has therefore only used a handful of restrictions under this methodology, and this often proves to be ample. For instance, [Antolín-Díaz and Rubio-Ramírez \(2018\)](#) show that even a small number of narrative

shocks can make significant improvements to the VAR identification, and in fact use just a single one in their VAR studying the effects of monetary policy.³

Second, as shown by [Alesina et al. \(2015\)](#) using an earlier vintage of the same narrative database that we use, fiscal adjustments are best viewed as multi-year plans instead of a sequence of shocks in consecutive years. Their analysis further suggests that information beyond the first year of a fiscal plan is more likely to confound the estimation of exogenous and unexpected shifts in fiscal stance, either because it is pre-announced (for instance in the first year of the plan), but also because there are often frequent adjustments during fiscal plans that significantly alter and even reverse the sign of fiscal adjustments beyond the first year of the program. Reducing the influence of these confounding effects is also an important reason why we focus on the first year whenever we encounter a multi-year sequence in the underlying narrative database.

Examining the narrative fiscal shocks as fiscal treatments (i.e a 0/1 outcome), [Jordà and Taylor \(2016\)](#) argue that they are predictable to a large extent based on lagged values of macroeconomic variables that are typically not included in estimations in the literature. If this predictability is pronounced, then it biases the estimates obtained. It is therefore important to examine the extent to which the narrative restrictions that we incorporate in our VARs suffers from this predictability problem. To do so, we regress the actual magnitudes of the narrative shocks (instead of a discrete 0/1 variable as in their case) on the same set of variables considered in [Jordà and Taylor \(2016\)](#), which includes lagged values of the output gap, the level of debt to GDP, GDP growth and the narrative shocks, as well as country fixed effects.

Similar to [Jordà and Taylor \(2016\)](#), this exercise indeed suggests that the narrative shocks are to a significant extent predictable, with an R square around 0.4. That said, this overall predictability masks the fact that for almost all (95 percent) of the narrative shocks that we incorporate as narrative restrictions in the VAR, this regression still yields a positive residual, suggesting that these narrative shocks, even after accounting for the impact of macroeconomic variables, do carry an unexplained fiscal consolidation component that is not predictable by lagged macroeconomic variables. Moreover, it is interesting to note that this number is much lower (around 70 percent) if we consider all narrative shocks instead of just the first year in a multi-year episode, which provides further justification for our approach of using the latter, as consolidations beyond the first year of an episode are more likely to be predictable and already anticipated by agents.

3.2 Estimation

We estimate the VAR country by country with two lags using Bayesian Methods. For the reduced form VAR, we impose a Minnesota Prior. Given the small sample size, the shrinkage features offered by the Minnesota prior are particularly helpful in the present context. We follow the version of the Minnesota prior discussed in [Sims and Zha \(1998\)](#) and [Del Negro \(2011\)](#) which entails a random covariance matrix for the reduced form VAR innovations, and choose the shrinkage parameters to maximize marginal data density—see for instance [Canova](#)

³See also [Giacomini et al. \(2022\)](#) on the value of even a small number of narrative restrictions in sign-identified VARs.

(2007) and [Giannone et al. \(2015\)](#).

For identifying the structural parameters by imposing narrative sign restrictions, we use the approach of [Antolín-Díaz and Rubio-Ramírez \(2018\)](#), which in turn builds on [Rubio-Ramírez et al. \(2010\)](#), who propose using the Haar prior as a computationally efficient Bayesian approach to implementing sign restrictions. The Haar prior treats all realizations of the rotation matrices used to implement sign restrictions as equally likely and implies that the structural impulse response prior is invariant to post-multiplication by orthogonal matrices. A few recent studies have questioned the appropriateness of this prior in some stylized settings given that the rotation matrix itself does not enter the likelihood and hence the prior cannot be overruled by the data (see for instance [Baumeister and Hamilton \(2015\)](#) and [Baumeister and Hamilton \(2018\)](#) for an illustration of this point). However, as argued by [Inoue and Kilian \(2021\)](#), the influence of the Haar prior tends to be negligible in structural VAR models that are tightly identified by sign restrictions, and more so when augmented with identification structures that entail multiple shocks as well as narrative restrictions as is the case in this paper. Moreover, as shown by [Arias et al. \(2023\)](#), uniform prior over the set of orthogonal matrices is not only sufficient but also necessary to have a uniform joint prior and posterior distributions over the identified set for the vector of impulse responses when identifying multiple shocks as we do, and is more suited than alternatives such as [Giacomini and Kitagawa \(2021\)](#) when conducting inference with respect to multiple shocks.

4 Results

This section summarizes the results. To answer the first question on drivers of debt, we focus on historical decompositions, and to answer the second one on the impact of fiscal consolidations unrelated to business cycles, we then turn to impulse responses from the estimated SVAR.⁴

We estimate the model country-by country and summarize medians across 17 advanced economies. The natural alternative to this approach would have been to estimate a panel VAR. We opt for the country-by country approach over a panel VAR because the latter imposes dynamic homogeneity across countries, which may not be justifiable given the large heterogeneity in debt dynamics across countries already documented in the literature. Indeed, as we show in section 5 the data offers strong evidence in favor of heterogeneity across countries and ignoring this feature can bias the inference.

4.1 Structural Debt Decompositions

SVAR-based historical decompositions give the year-by-year contribution of each one of the shocks to the observed changes in each of the variables in the system, with a residual component capturing whatever the specified shocks cannot explain, along with a purely deterministic component. From this point of view, our analysis offers a structural version of standard debt decompositions with the clear advantage that in our case shocks are structurally identified, and orthogonal. Figure 3 shows an example of the historical decomposition of debt

⁴All estimations are performed using the Empirical Macro Toolbox developed by [Ferroni and Canova \(2021\)](#).

to GDP for one country (USA) computed from the VAR, and Table 2 displays the median historical decomposition of debt to GDP ratios across countries in our sample. The numbers reported take the median of absolute contributions of the different shocks across countries and time.

The table reveals that growth shocks account for about 40% of the fluctuations in debt for the median country in the sample. Within growth shocks, the share of demand shocks is somewhat higher by about 7%. On the other hand, at 16% for the median country, the discretionary primary balance shock account for less one-fifth of debt fluctuations.

In the subsequent analysis, we further scrutinize the data through various lenses to examine cross-country and temporal variations. As disclosed in Table 2, the findings indicate that the aggregate contribution, as well as individual shock contributions to the changes in the debt-to-GDP ratio, exhibit remarkable uniformity across nations. Italy emerges as a notable exception, wherein the cumulative effect of the identified shocks accounts for less than 50% of the fluctuations in the debt ratio. Moreover, Italy's discretionary fiscal policy contributes minimally compared to other nations.

In an analysis partitioned by periods of rising and falling debt ratios, distinct patterns emerge. Table 3 demonstrates that the share of primary balance consolidation shocks increases from an average of 14% during debt ratio ascension to 18% in periods of decline. This variation is corroborated by Figure 4, which reveals significant cross-country heterogeneity. Ireland serves as a particularly illuminating case. The influence of discretionary fiscal consolidation nearly doubles when the debt ratio is in decline, compared to periods of rising debt.⁵ Moreover, Figure 5 displays large residuals for the years 2014 and 2015, when debt-to-GDP declined substantially. These residuals are not a flaw but a validation: they affirm that the model is well-specified in that it does not fit non-economic fluctuations, which are outside its scope. Such anomalies, specifically in 2014 and 2015, were largely a result of 'below-the-line' fiscal operations and statistical recalibrations that inflated GDP by over 20%. The model's inability to capture these distortions underscores its appropriate focus on economic variables.⁶

Lastly, we assess whether periods of atypical fluctuations in debt ratios exhibit distinct contributing factors. To this end, we isolate for each country a subset of five years featuring the most significant increases and decreases in debt ratios. This subset is subsequently compared to the average values presented in prior tables. Table 3 confirms that, even when evaluated against this criterion, the influential factors for debt ratios exhibit limited variability on average.

Overall, our results suggest that the contribution of the identified shocks to fluctuations in debt ratios is fairly stable across countries, albeit with some notable differences overall as in the case of Italy, or in the relative importance of discretionary shocks in debt ratio decreases as opposed to increases. (Table 4) for large increases and decreases

⁵For a comprehensive examination of the Irish context, see Figure 5. This figure highlights that significant reductions in debt, such as those observed in the late 1980s, mid-1990s, and 2014–2015, are primarily driven by substantial primary balance shocks. Conversely, episodes of considerable debt accumulation, notably around the Global Financial Crisis and early 1980s, are largely attributable to growth factors with minimal discretionary impact.

⁶For an expanded discussion on 'below-the-line' operations and the 2015 GDP inflation, the reader is directed to the IMF Article IV Staff Reports for Ireland [International Monetary Fund \(2015\)](#) and [International Monetary Fund \(2017\)](#).

Historical decompositions focus on levels of the variable under consideration. A natural alternative is to look at what explains the second moments of the debt to GDP ratio. In a SVAR this could be considered by looking at the Forecast Error Variance. Table 5 summarizes the results, and a symmetric message emerges. The structural shocks considered here are relevant drivers of the debt to GDP, with growth shocks being the most important force driving the volatility of debt to GDP.

4.2 Dynamic Effects of Structural Shocks

This section summarizes the median impulse response over six years of all variables to the model’s one standard deviation orthogonal structural shocks. The shaded areas represent the 68 percent highest posterior density credible sets computed using inverse variance weights for the IRFs when the sign restrictions and narrative sign restrictions are satisfied.⁷

The demand shock (Figure 6) shows that there is a strong counter-cyclical response of the primary balance, and, importantly, that the overall effect on debt to GDP is also countercyclical. The latter is a result, since no conditions are imposed on the level of debt. The primary balance response is front loaded (the strongest response is on impact) and is relatively long lasting. Since the SVAR shocks are symmetric, these results indicate that countries reduce their debts while in economic expansions, and increase them in contractions. The Figure also illustrates that the effective interest rate increases in response to a demand shock.

Supply shocks (Figure 7) have qualitatively similar effects on the variables of the model (with a front-loaded primary balance response, and a countercyclical debt reaction). Unlike the demand shock, the supply shock has an insignificant impact on the effective interest rate on debt. Monetary policy could be one factor behind this difference. In the case of a demand shock, the output-inflation tradeoff is more favorable than in the case of a supply shock when output and inflation move in the opposite direction. Central banks with dual mandates on output and inflation would therefore unambiguously prefer to raise rates in response to a demand shock, whereas their response to a supply shock would be contingent on how they tackle the output-inflation tradeoff.

Discretionary primary balance consolidations (primary balance shocks not related to the business cycles, through the narrative information, and sign restriction identification) also tend to be front-loaded and durable (Figure 8). We find that these shocks also reduce debt ratios on average, although not by much.⁸

The figure also reveals that Primary balance consolidations de-linked from the growth shocks have a minimal effect on GDP. As discussed in the introduction, the growth effects of fiscal consolidations remains an active area of debate in the literature.

Our identification approach is novel for the analysis of macro-fiscal issues, and allows us to leverage narrative information while preserving the advantages of shock exogeneity from the SVARs. While both [Guajardo et al. \(2014\)](#) and [Jordà and Taylor \(2016\)](#) aim to capture the impact of fiscal consolidations that are unrelated to macroeconomic conditions, they note

⁷The use of 68% confidence intervals is relatively standard in the literature- see for instance [Sims and Zha \(1999\)](#), [Murphy \(2015\)](#). Similar to [Di Pace et al. \(2023\)](#), we use inverse variance weights to aggregate country level impulse responses

⁸See [Bi et al. \(2013\)](#) for a characterization of conditions under which consolidations can fail to reduce debt.

that narrative data is not immune from potential biases. For instance, if countries postpone fiscal consolidation until the economy recovers, or strengthen it when growth unexpectedly slows to keep deficits in check, then the narrative shocks are confounded by business cycle influences, leading to biased results. Similarly, they note how if policymakers officially motivate fiscal consolidation in terms of ensuring long-term debt sustainability, while in fact being motivated by cyclical considerations, this may again bias the results obtained from the pure narrative approach. Our method is geared towards addressing concerns of this type. By explicitly identifying two GDP growth shocks, it aims to purge the discretionary shocks from macroeconomic influences in a precise manner. For this reason, we restrict attention to the qualitative information contained in the narrative shocks, allowing them to be arbitrary close to zero even in periods in which narrative restrictions on signs of shocks are imposed, rather than the precise quantitative magnitudes.

Finally, our results also reveal that instead of declining on the back of a fiscal consolidation, inflation moves very little and in fact ends up rising, albeit by a small amount. This is in contrast to predictions based on the aggregate demand channel of fiscal consolidations, which would typically entail a decline in inflation. That said, the moderate rise in inflation is consistent with the influence of several channels which might be at play. First, as shown by the rise in revenues, the average consolidation recovered in our analysis does entail a significant increase in taxes, which contribute directly to inflation via an increase in prices. Second, consolidations may also push up inflation if they are associated with a depreciation of the exchange rate that feeds into inflation. Third, our result of a moderate increase in inflation and the very low impact on output are consistent with models that incorporate good-specific habits (see for instance [Ravn et al. \(2006\)](#)). [Ravn et al. \(2012\)](#) show that the time varying markups implied by the deep habits feature can help match the response of the real exchange rate in response to fiscal spending shocks, and [Jacob \(2015\)](#) shows that this deep habits feature when combined with reasonable levels of price stickiness can substantially decrease the impact of fiscal shocks on output and inflation, even to the extent of changing the sign of the response of the latter, consistent with our results.

4.3 The Value of Narrative Restrictions

To evaluate the contribution of narrative restrictions, we estimate our SVAR with only sign restrictions and contrast the results with our baseline. The main takeaway is that, consistent with [Antolín-Díaz and Rubio-Ramírez \(2018\)](#), we find that narrative information, even when sparse, proves to be quite valuable in the VAR model.

Figure 9 displays the posterior distribution of the primary balance shock during narrative periods, both with and without the imposition of narrative restrictions. The first panel shows the comparison for the entire sample, whereas the second panel zooms in on one specific episode—the 1990 consolidation in the US. From the narrative data, we know that all these shock draws should be positive. However, the figure shows that identification based purely on sign restrictions also includes several negative values that should not exist, illustrating the sharpening effect of narrative information. In fact, for the US in 1990, the mean of the shock without narrative restrictions is even slightly negative (-0.07), whereas the mean with narrative restriction is 0.3, very close to the actual value of the shock in the narrative database of [Guajardo et al. \(2014\)](#) which is 0.36, even though we only use the sign and not

the magnitude as part of the narrative restriction.

Consistent with the literature, we also find that the introduction of narrative restrictions changes both the precision, and in some cases the qualitative inference drawn from the data. On precision, Table 6 shows the difference between the width of confidence intervals between IRFs computed with and without narrative restrictions, expressed as percentage differences so that negative values indicate smaller confidence intervals for the VAR with narrative restrictions. The numbers therefore reveal that narrative information tightens the Highest Posterior Density (HPD) credible sets for most impulse responses, with the impact being especially noticeable for the primary balance shock.

More strikingly, Figure 10 shows that the introduction of narrative information leads to significant changes in key conclusions with respect to the impact of a primary balance shock on GDP and the debt ratio. In particular, with narrative sign restrictions, as discussed above, we find the impact on GDP and debt ratios to be negative but small. In contrast, the VAR without narrative restrictions actually yields a positive (albeit still small) impact on GDP, and a much sharper negative impact on debt ratios.

Together, the analyses presented in Figure 9, Table 6 and Figure 10, underscore the value of incorporating narrative restrictions into the VAR estimation process. Narrative restrictions not only sharpen the identification of shocks but also enhance the accuracy of HPD credible sets, ultimately influencing key conclusions.

5 Towards a Reconciliation with Contractionary Austerity Results in the Literature

Our results consistently suggest the response of consolidations on GDP to be minimal. This is in stark contrast to studies that have found a strong negative impact on GDP using versions of the same narrative fiscal dataset that we use, such as [Guajardo et al. \(2014\)](#) and [Carrière-Swallow et al. \(2021\)](#).

There are at least two potential sources for this discrepancy. The first is that our approach, by explicitly modeling GDP shocks, and combining the benefits of SVARs with narrative information, deals with the problems of foresight and endogeneity better. The second possible source of the discrepancy is heterogeneity across countries. Conventional panel regression techniques often employed in previous research typically impose homogeneity across countries. This is true even when the panel introduces both country and time fixed effects. This homogeneity assumption can introduce substantial biases to the estimates, especially in dynamic settings, as has been underscored in the literature (see for instance [Pesaran and Smith \(1995\)](#)).

Finally, we perform two additional external validation tests. First, we apply our approach to study the tax multiplier in the US. [Mertens and Ravn \(2013\)](#) pioneered an approach which also aims to leverage narrative information while preserving the benefits of SVARs, and find a large negative tax multiplier for the US. Our identification method, based on narrative sign restrictions, when applied to the same data to identify tax revenue shocks, also finds a large negative multiplier. Secondly, we find that fiscal expansion shocks identified by our method for the US agrees with military spending shocks of [Ben Zeev et al. \(2023\)](#) despite not

incorporating such narrative data in the estimation.

5.1 The Problem of Foresight and Invertibility

The survey article [Ramey \(2016\)](#) summarizes a key issue in empirical work and the identification of macroeconomic policy shocks. In the context of fiscal policy, it is likely that changes in policies are anticipated by private and public agents in advance of their occurrence. The presence of such foresight can hamper the recovery of fiscal shocks using regressions on contemporaneous and past macroeconomic variables. Under these conditions the VAR representation becomes non-fundamental and the true shocks cannot be recovered.

As noted by ? and [Chahrour and Jurado \(2022\)](#) among others, invertibility is a property of the disturbances of a structural model. As such, it can only be tested with reference to a given (set of) candidate model(s), and there is no precise testing approach that is entirely data driven. We nevertheless provide some empirical evidence suggesting that the consequences of this problem are limited in our setting, especially in comparison to other approaches used in the literature.

First, drawing from [Beaudry et al. \(2019\)](#), we leverage a diagnostic designed to assess the quantitative significance of non-fundamentalness. This involves computing the R-squared statistic from a regression of candidate structural shocks on observables not explicitly considered in the empirical model. The goodness of fit measures derived from this model offer insight into potential biases due to non-invertibility. Notably, simulations from these authors indicate that an R-squared value below 0.25 suggests that, although non-invertibility might exist theoretically, its practical implications are not significant.

Following their approach, Table 7 shows the R square statistics when we regress different measures of fiscal shocks on certain observables that are not included in our VAR but could in theory be important determinants of shocks that we wish to identify. The first column in the table shows the R square statistics when we use contemporaneous and lagged macroeconomic variables. We include the levels of the primary balance to GDP, real government revenues, real GDP, the output gap and the VIX as a proxy for global financial conditions. In the second column, in addition to the contemporaneous and lagged macro variables from column 1, we add one and five year ahead forecasts of debt to GDP, GDP growth, and primary balance to GDP.

The table shows that across both specifications, regressions of narrative shocks on macro variables and forecasts yields fairly large R square statistics in excess of 0.25. On the other hand, for the SVAR shocks, the explanatory power in the regression is much lower, with R squares well below 0.10.

Although these exercises do not provide a decisive test of invertibility or lack thereof, they do offer suggestive evidence that the quantitative importance of this problem is likely to be limited in our baseline VARs that combine narrative and sign restrictions, relative to both SVARs that use only sign restrictions and local projections that use the narrative shocks directly.

As a further check on the validity of the shocks that we use to draw key inferences in the paper, we appeal to the insight of [Cochrane \(2004\)](#), who argued that in the case of monetary policy shocks, to measure the effects of monetary policy on output, it is enough that the shock is orthogonal to output forecasts. It does not have to be orthogonal to other variables,

and may even be predictable from contemporaneous and past macro variables, and need not necessarily be a shock to the information set of agents. We check for this condition for fiscal shocks with respect to forecasts of two key variables on which we conduct inference in our analysis—namely GDP growth and changes in debt to GDP. In Table 8, we regress one and five year ahead forecasts for these variables on VAR (sign-narrative) shocks as well as the narrative shocks directly. The results show that while the VAR shocks are orthogonal to the forecasts of GDP growth and changes in debt to GDP with no significant predictive power, the same is not true for narrative shocks, which turn out to be systematically correlated with forecasts for changes in debt to GDP.

To further examine the consequences of these issues, we estimate impulse responses using local projection methods common in the literature (e.g. [Carrière-Swallow et al. \(2021\)](#)) for different estimators and shocks. As established by [Plagborg-Møller and Wolf \(2021\)](#), given the same identification method, the impulse responses from SVARs and those of local projections converge to the same (as the sample size grows). From this perspective, there are three reasons why our results may differ. First, the different asymptotic properties of the methods (due to sample size). Second, our identification is based on narrative sign restrictions and does not use the narrative data directly. And third, we estimate the model country by country country level instead of using panel methods such as [Jordà and Taylor \(2016\)](#) that impose homogeneity across countries. To consider each of these differences, we first estimate the dynamic response of GDP growth to a primary balance shock using a fixed effects estimator that controls for the lagged dependent variable as well as country and time fixed effects. We consider two versions of what is a primary balance shock. The first uses the narrative data directly, and the second uses the shocks from our SVAR analysis which are less affected by the problem of foresight, as discussed earlier. The specification of the estimation is as follows:

$$\Delta y_{i,t+h} = \alpha^h \Delta y_{i,t-1} + \beta^h (Shock_{it}) + \alpha_i^h + \delta_t^h + \epsilon_{it}^h \quad (5.1)$$

where $\Delta y_{i,t}$ denotes real GDP growth in country i in year t , $Shock_{it}$ is the primary balance shock, and α_i^h and δ_t^h are country and time fixed effects respectively.⁹

The results are summarized in Figure 11. The top-left panel reproduces the results of using the narrative shocks directly using a two-way fixed effects estimator. Note that, consistent with the previous literature, we recover the sharp and persistent negative decline in GDP growth in response to a narrative fiscal consolidations shock from the local projection method.

For comparison, the top-right panel shows the outcome when we replace narrative shocks with our structurally identified VAR shocks with sign restrictions only, and the bottom-left panel shows the estimates with identified shocks from the VAR with narrative-sign restrictions.¹⁰

When standard methods are "fed" with the exogenous shocks estimated from our SVAR, the zero fiscal multiplier of our baseline analysis also emerges, suggesting potential biases from relying solely on narrative data. The zero fiscal multiplier result is obtained despite the fact that unlike the baseline VAR with sign-narrative restrictions, this approach controls for common global factors via time fixed effects, which serves as an additional robustness

⁹We use the most parsimonious specification for illustrative purposes, but the results are robust to the inclusion of additional control variables

¹⁰We use the median of the extracted shocks for each country and year from the SVAR.

check on the VAR results. Differences in results with local projections in the literature are those due to identification methods, and not to sample size.

Note that in line with studies that have shown a positive impact of fiscal shocks identified via cyclically adjusted primary balance, point estimates in the top-right panel are higher than ones obtained by imposing narrative sign restrictions in the bottom-left panel.

As our final robustness check, the bottom-right panel of Figure 11 shows the response of GDP growth using narrative shocks as in the top-left panel, but instrumented using the sign-narrative VAR shocks used in the bottom-left panel. This approach addresses the endogeneity concerns of using the narrative shocks directly, while at the same time accounting for uncertainty and measurement error in the VAR extracted shocks, which by their very nature are set-identified. The VAR shocks turn out to be strong instruments for the narrative shocks, with the first stage regression yielding an F statistic of 15. Moreover, estimated response of GDP growth using this IV approach is almost identical to our baseline.

One potential pitfall with our identification is that the narrative restrictions that we use correspond solely to fiscal consolidations, as opposed to expansions. Yet, the linear SVAR that we employ implicitly assumes a symmetry of responses between expansions and recessions. This may in principle sound problematic, for example in the case of the US where certain episodes of increased military spending (and thus exogenous primary balance expansions) have been important drivers of debt. Our method could in principle fail to identify them and instead wrongly attribute their influence to either supply or demand shocks. At least for the US, we find this potential problem does not materialize. Indeed, we find that in both years (1999 and 2003) where our sample overlaps with the news shocks of [Ben Zeev et al. \(2023\)](#) our primary balance shock is correctly identified: agreeing with the sign of the spending news shock, which is positive in those years. Furthermore, [Ben Zeev et al. \(2023\)](#) document that positive and negative shocks identified via their military news measure do not have a significantly different effect.

5.2 The Role of Heterogeneity

We now shift our focus to a second key divergence between our study and prior literature: the issue of inherent cross-country heterogeneity. Our methodological approach mitigates this limitation by adopting a country-by-country SVAR framework. This individualized treatment allows us to accommodate the inherent differences between countries, offering a more nuanced and accurate picture of the impact of discretionary primary balance consolidations.

To empirically investigate the role of heterogeneity, we estimate local projections for each country individually over a three year horizon, examining the distribution of primary balance multipliers that emerge. The resulting distribution in the left panel in Figure 12 illustrates that when heterogeneity is accounted for in country-by-country local projections, the GDP response to consolidations is significantly attenuated, and is centered around zero. Interestingly, this distribution is qualitatively similar to the cross-country distribution obtained from our baseline SVAR estimates, shown for comparison in the right panel of the Figure. This finding underscores the potential biases introduced by homogeneity assumptions prevalent in prior studies ([Favero et al. \(2011\)](#)).

Further analysis reveals, nevertheless, that exogeneity, rather than heterogeneity, seems to be the key issue. To see this, the left panel of Figure 13 shows that the correlation between

the coefficients estimated country by country using local projections on purely narrative shocks (and hence allowing for country heterogeneity) and the corresponding estimates from the sign-narrative VAR is essentially zero. On the other hand, the right panel shows that when we replace narrative shocks with our VAR derived sign narrative shocks in the same local projection estimator, the correlation with the VAR estimates is much higher. These results further corroborate the findings in Figure 11, which shows that even upon imposing country homogeneity in a two-way fixed effect regression, replacing purely narrative shocks with the VAR-derived shocks essentially recovers the VAR result of insignificant impact of consolidations on GDP.

5.3 Proxy SVARs and the Tax Multiplier

We conclude this section by relating our work to a paper close in spirit to ours. [Mertens and Ravn \(2013\)](#) (MR) study the tax multiplier in the United States. To deal with endogeneity issues, the authors develop a proxy VAR approach using narrative shocks as external instruments. This combines the benefits of the SVARs (exogenous shocks), while extracting useful information from narrative data. How would our method perform when applied to the estimate the tax multiplier in the US.? This test serves as an additional external validation of our methods.

We retain the two growth shocks as in our core analysis, but replace the primary balance shock with shocks to government revenues. Figure 14 shows the response to a government revenue shock for the US using our data sample from 1981-2019, and our narrative information. In this case we also find that GDP growth declines in response to a revenue increase shock. The response, however, has a wide confidence band and is not statistically significant.

To sharpen our estimates, we expand the sample and draw from the narrative data of [Mertens and Ravn \(2013\)](#) and estimate the VAR on their sample for comparison (1950-2006). We focus on tax increases (to remain close to our earlier approach). As shown in Figure 15 our approach now finds a statistically significant negative multiplier, consistent with the authors' original findings, offering reassurances about the validity of our methods. In terms of magnitude, our multiplier peaks at about 3, which is somewhat larger than the peak multiplier of 2.5 in [Mertens and Ravn \(2013\)](#), but smaller than the revenue multiplier in [Mountford and Uhlig \(2009\)](#) which peaks at 5.

6 Robustness: Priors, Global Shocks, Unexpected Inflation, and Predictability of our Narrative Data

Finally, we consider the robustness of some of our main results to different priors and to the inclusion of global shocks. We also discuss the issue of predictability of the narrative information.

Given our short time sample, in our baseline estimations we preferred to use Minnesota priors to discipline the estimates of the reduced form VAR by imposing a shrinkage structure. However, the second column in Table 9 shows that our main result on the relative importance of growth versus discretionary primary balance shocks in the historical decomposition of debt

to GDP is robust to a more uninformative Jeffrey’s prior. The ratio of the contribution of growth to discretionary shocks remains around 2.5 under either prior.

Our approach so far has been agnostic on whether the identified shocks originate within or beyond a country’s borders. To explicitly allow for the fact that common global factors could play an important role in driving fluctuations in debt ratios and other macroeconomic aggregates, we augment the baseline SVAR to include the 10 year nominal yield on US government bonds as an exogenous variable. Column 3 in Table 9 shows that this variable can indeed account for a large fraction of the movements in debt ratios. But the new estimates further strengthen our main finding that luck (growth shocks) is the main driver of debt ratios, in contrast to discretionary fiscal policy.

The last column in 9 shows that the relative contribution of growth vs discretionary fiscal shocks remains largely unchanged if we augment the baseline VAR specification by identifying an additional monetary policy shock via impact sign restrictions on the interest rate and inflation in the spirit of [Mountford and Uhlig \(2009\)](#).

[Jordà and Taylor \(2016\)](#) document, using a standard regression approach, that the narrative information we employ is predictable. Specifically, past values of debt can forecast the magnitude of the narrative shocks. This is a problem because the VAR representation is not fundamental if anticipated shocks are present (VAR innovations would not be just contemporaneous rotations of structural shocks but moving averages, functions of current and past structural shocks). Our strategy attenuates the problem because we do not use the actual values of the narrative shocks, but only use the information of the presence of a shock in the narrative sign restrictions. To further explore the issue, we reproduce in our data the regression in [Jordà and Taylor \(2016\)](#). Then, we filter our narrative information based on residual of that regression and use only the non predictable part of the shock (for example if a positive consolidation is expected we only keep it if the actual consolidation is larger than what was expected from the regression). The results are essentially the same as in the baseline.

7 Conclusion

This paper explores the dynamics of sovereign debt in advanced economies using a structural approach to debt decompositions. Its basis is a Structural Vector Autoregression (SVAR) framework augmented with narrative sign restrictions. The empirical analysis reveals that exogenous shocks, specifically demand and supply growth shocks, account for approximately forty percent of the variations in the debt-to-GDP ratio. In contrast, discretionary policy shocks contribute less than a fifth of these variations. Notably, the analysis provides evidence that discretionary primary balance consolidations exert a modest influence on the debt-to-GDP ratio and have, on average, an insignificant impact on GDP.

Our paper discusses two significant sources for the discrepancy between our result of a zero primary balance multiplier and previous literature that has found a strong negative impact of consolidations on GDP. First, we emphasize the difficulties in dealing with the problem of foresight. Our SVAR analysis uses narrative information while also isolating primary balance movements that are unrelated to business cycle and growth shocks. Available tests of the possible effect of non-fundamentalness suggest our method is indeed better in

handling possible biases. Second, we highlight that our SVAR estimation, conducted on a country-by-country basis, inherently accommodates cross-country heterogeneity. This approach effectively sidesteps the restrictive assumption of coefficient homogeneity across units imposed by panel regressions that is known to lead to biased estimates.

It is worth noting that while our methodology aims to rigorously address issues related to exogeneity and cross-country heterogeneity, alternative methods, such as those by [Jordà and Taylor \(2016\)](#), offer other ways to tackle exogeneity. These methods, however, come with their own limitations, specifically the need for extensive panel data and the associated challenges such as the imposition of cross-country homogeneity. Therefore, while our approach offers important insights, future work may benefit from exploring the trade-offs involved in different methodological choices.

External validation further substantiates the robustness of our methodology. We apply our framework to estimate the tax multiplier in the United States, finding values quantitatively in line with the pioneering work of [Mertens and Ravn \(2013\)](#). Further, our methodology accurately identifies the magnitude of the 1990 US. consolidation shock in [Guajardo et al. \(2014\)](#), despite relying only on the sign from the narrative shock.

Looking forward, this research paves the way for several fruitful avenues. Future work could extend the analysis to emerging economies, providing a broader perspective on global sovereign debt dynamics. Additionally, researchers might investigate potential non-linearities in debt dynamics as well as the impact of consolidations on GDP growth, which we on average find to be indistinguishable from zero. Lastly, our analysis focused on primary balance to GDP ratios given the main objective of uncovering drivers of public debt and our use of narrative shocks that are specifically designed to capture exogenous changes in the primary balance to GDP ratio. A promising avenue for future work would entail extending the analysis for select countries with available data to zoom in on different components of the primary balance, namely revenues and expenditures, and their subcomponents such as different types of taxes as well as expenditure components (government consumption vs investment).

References

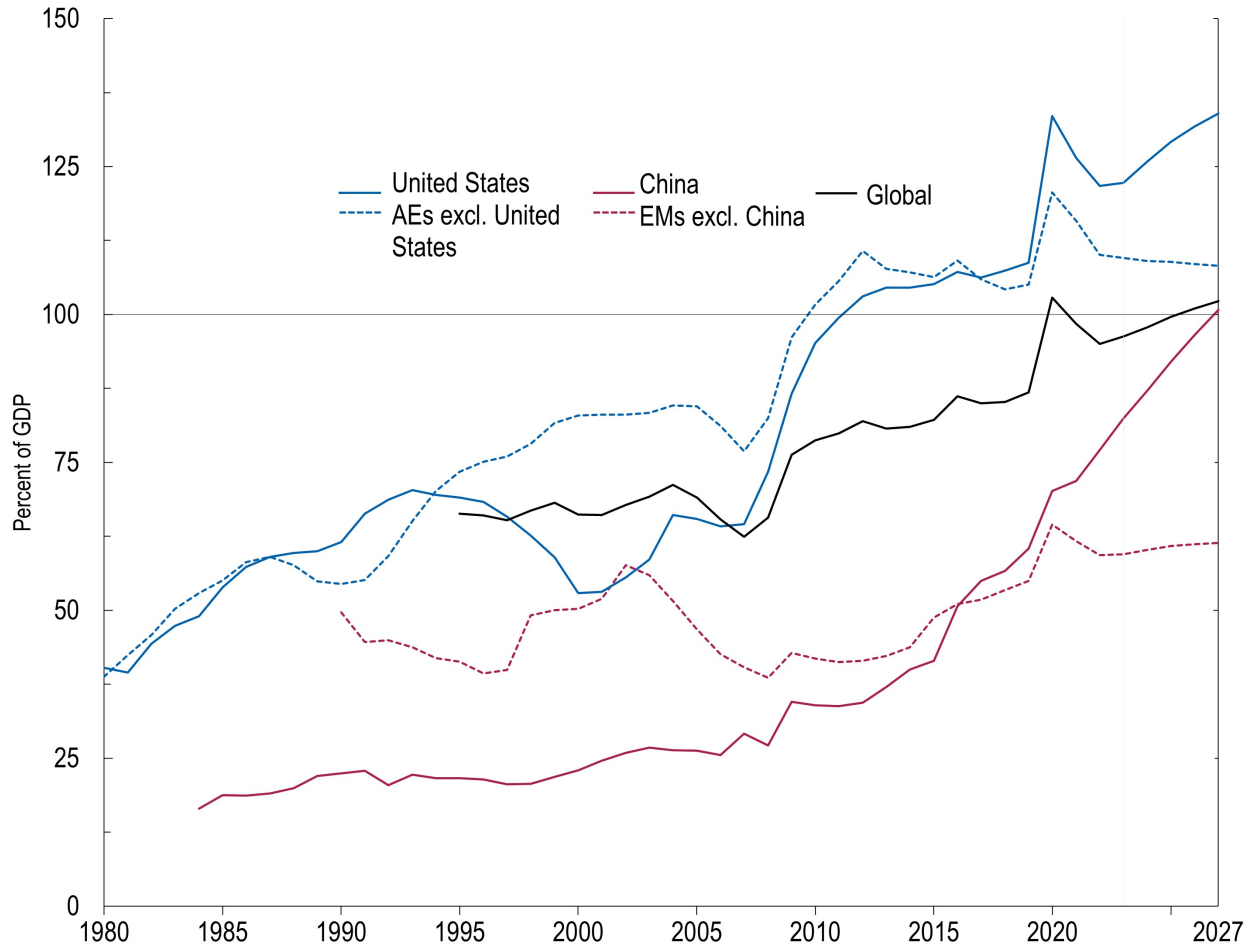
- [Acalin, Julien and Laurence M Ball](#), “Did the US really grow out of its World War II debt?,” 2023.
- [Alesina, Alberto and Roberto Perotti](#), “Fiscal adjustments in OECD countries: composition and macroeconomic effects,” *Staff Papers*, 1997, *44* (2), 210–248.
- [and Silvia Ardagna](#), “Tales of fiscal adjustment,” *Economic policy*, 1998, *13* (27), 488–545.
- [and —](#), “Large changes in fiscal policy: taxes versus spending,” *Tax policy and the economy*, 2010, *24* (1), 35–68.
- [, Carlo Favero, and Francesco Giavazzi](#), “The output effect of fiscal consolidation plans,” *Journal of International Economics*, 2015, *96*, S19–S42.

- Ando, Sakai, Prachi Mishra, Nikhil Patel, Adrian Peralta-Alva, and Andrea Presbitero**, “Fiscal Consolidation and Public Debt,” *CEPR Working Paper DR18548*, 2023.
- Antolín-Díaz, Juan and Juan F Rubio-Ramírez**, “Narrative sign restrictions for SVARs,” *American Economic Review*, 2018, *108* (10), 2802–29.
- Arias, Jonas, Juan F Rubio-Ramirez, and Daniel F Waggoner**, “Uniform priors for impulse responses,” 2023.
- Barnichon, Regis, Davide Debortoli, and Christian Matthes**, “Understanding the Size of the Government Spending Multiplier: Its in the Sign,” *The Review of Economic Studies*, 2022, *89* (1), 87–117.
- Baumeister, Christiane and James D Hamilton**, “Sign restrictions, structural vector autoregressions, and useful prior information,” *Econometrica*, 2015, *83* (5), 1963–1999.
- and —, “Inference in structural vector autoregressions when the identifying assumptions are not fully believed: Re-evaluating the role of monetary policy in economic fluctuations,” *Journal of Monetary Economics*, 2018, *100*, 48–65.
- Beaudry, Paul, Patrick Feve, Alain Guay, and Franck Portier**, “When is nonfundamentalness in SVARs a real problem?,” *Review of Economic Dynamics*, 2019, *34*, 221–243.
- Ben Zeev, Nadav, Valerie A Ramey, and Sarah Zubairy**, “Do government spending multipliers depend on the sign of the shock?,” in “AEA Papers and Proceedings,” Vol. 113 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2023, pp. 382–387.
- Bi, Huixin, Eric M Leeper, and Campbell Leith**, “Uncertain fiscal consolidations,” *The Economic Journal*, 2013, *123* (566), F31–F63.
- Canova, Fabio**, *Methods for applied macroeconomic research*, Vol. 13, Princeton university press, 2007.
- Carrière-Swallow, Yan, Antonio C David, and Daniel Leigh**, “Macroeconomic Effects of Fiscal Consolidation in Emerging Economies: New Narrative Evidence from Latin America and the Caribbean,” *Journal of Money, Credit and Banking*, 2021, *53* (6), 1313–1335.
- Chahrour, Ryan and Kyle Jurado**, “Recoverability and expectations-driven fluctuations,” *The Review of Economic Studies*, 2022, *89* (1), 214–239.
- Cherif, Reda and Fuad Hasanov**, “Public debt dynamics: the effects of austerity, inflation, and growth shocks,” *Empirical Economics*, 2018, *54*, 1087–1105.
- Cochrane, John**, “Comments on "A new measure of monetary shocks: Derivation and implications" By Christina Romer and David Romer. July 17 2004, presented at NBER EFG meeting,” *Handbook of macroeconomics*, 2004.

- Cochrane, John H**, “The value of government debt,” Technical Report, National Bureau of Economic Research 2019.
- Das, Piyali and Chetan Ghate**, “Debt decomposition and the role of inflation: A security level analysis for India,” *Economic Modelling*, 2022, p. 105855.
- Del Negro, Marco**, “Bayesian macroeconometrics,” 2011.
- Di Pace, Federico, Luciana Juvenal, and Ivan Petrella**, “Terms-of-trade shocks are not all alike,” 2023.
- Favero, Carlo, Francesco Giavazzi, and Jacopo Peregò**, “Country heterogeneity and the international evidence on the effects of fiscal policy,” *IMF Economic Review*, 2011, 59 (4), 652–682.
- Ferroni, Filippo and Fabio Canova**, “A Hitchhiker’s Guide to Empirical Macro Models,” 2021.
- Giacomini, Raffaella and Toru Kitagawa**, “Robust Bayesian inference for set-identified models,” *Econometrica*, 2021, 89 (4), 1519–1556.
- , – , and **Matthew Read**, “Narrative restrictions and proxies,” *Journal of Business & Economic Statistics*, 2022, 40 (4), 1415–1425.
- Giannone, Domenico, Michele Lenza, and Giorgio E Primiceri**, “Prior selection for vector autoregressions,” *Review of Economics and Statistics*, 2015, 97 (2), 436–451.
- Giavazzi, Francesco and Marco Pagano**, “Can severe fiscal contractions be expansionary? Tales of two small European countries,” *NBER macroeconomics annual*, 1990, 5, 75–111.
- Guajardo, Jaime, Daniel Leigh, and Andrea Pescatori**, “Expansionary austerity? International evidence,” *Journal of the European Economic Association*, 2014, 12 (4), 949–968.
- Gunter, Samara, Daniel Riera-Crichton, Carlos A Vegh, and Guillermo Vuletin**, “Non-linear effects of tax changes on output: The role of the initial level of taxation,” *Journal of International Economics*, 2021, 131, 103450.
- Hall, George J and Thomas J Sargent**, “Interest rate risk and other determinants of post-WWII US government debt/GDP dynamics,” *American Economic Journal: Macroeconomics*, 2011, 3 (3), 192–214.
- Inoue, Atsushi and Lutz Kilian**, “The role of the prior in estimating VAR models with sign restrictions,” *Center for Financial Studies Working Paper*, 2021, (660).
- International Monetary Fund**, “Ireland 2015 Article IV Consultation Staff Report,” 2015.
- , “Ireland 2017 Article IV Consultation Staff Report,” 2017.

- Jacob, Punnoose**, “Deep habits, price rigidities, and the consumption response to government spending,” *Journal of Money, Credit and Banking*, 2015, 47 (2-3), 481–510.
- Jordà, Òscar and Alan M Taylor**, “The time for austerity: estimating the average treatment effect of fiscal policy,” *The Economic Journal*, 2016, 126 (590), 219–255.
- Mertens, Karel and Morten O Ravn**, “The dynamic effects of personal and corporate income tax changes in the United States,” *American Economic Review*, 2013, 103 (4), 1212–1247.
- Mountford, Andrew and Harald Uhlig**, “What are the effects of fiscal policy shocks?,” *Journal of applied econometrics*, 2009, 24 (6), 960–992.
- Murphy, Ryan H**, “Unconventional Confidence Bands in the Literature on the Government Spending Multiplier,” *Econ Journal Watch*, 2015, 12 (1), 72–83.
- Pesaran, M Hashem and Ron Smith**, “Estimating long-run relationships from dynamic heterogeneous panels,” *Journal of econometrics*, 1995, 68 (1), 79–113.
- Plagborg-Møller, Mikkel and Christian K. Wolf**, “Local Projections and VARS Estimate the Same Impulse Responses,” *Econometrica*, 2021, 89, 955–980.
- Ramey, Valerie A**, “Macroeconomic shocks and their propagation,” *Handbook of macroeconomics*, 2016, 2, 71–162.
- Ravn, Morten O, Stephanie Schmitt-Grohé, and Martín Uribe**, “Consumption, government spending, and the real exchange rate,” *Journal of Monetary Economics*, 2012, 59 (3), 215–234.
- Ravn, Morten, Stephanie Schmitt-Grohé, and Martin Uribe**, “Deep habits,” *The Review of Economic Studies*, 2006, 73 (1), 195–218.
- Romer, Christina D and David H Romer**, “The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks,” *The American Economic Review*, 2010, 100 (3), 763–801.
- Rubio-Ramirez, Juan F, Daniel F Waggoner, and Tao Zha**, “Structural vector autoregressions: Theory of identification and algorithms for inference,” *The Review of Economic Studies*, 2010, 77 (2), 665–696.
- Sims, Christopher A and Tao Zha**, “Bayesian methods for dynamic multivariate models,” *International Economic Review*, 1998, pp. 949–968.
- **and —**, “Error bands for impulse responses,” *Econometrica*, 1999, 67 (5), 1113–1155.

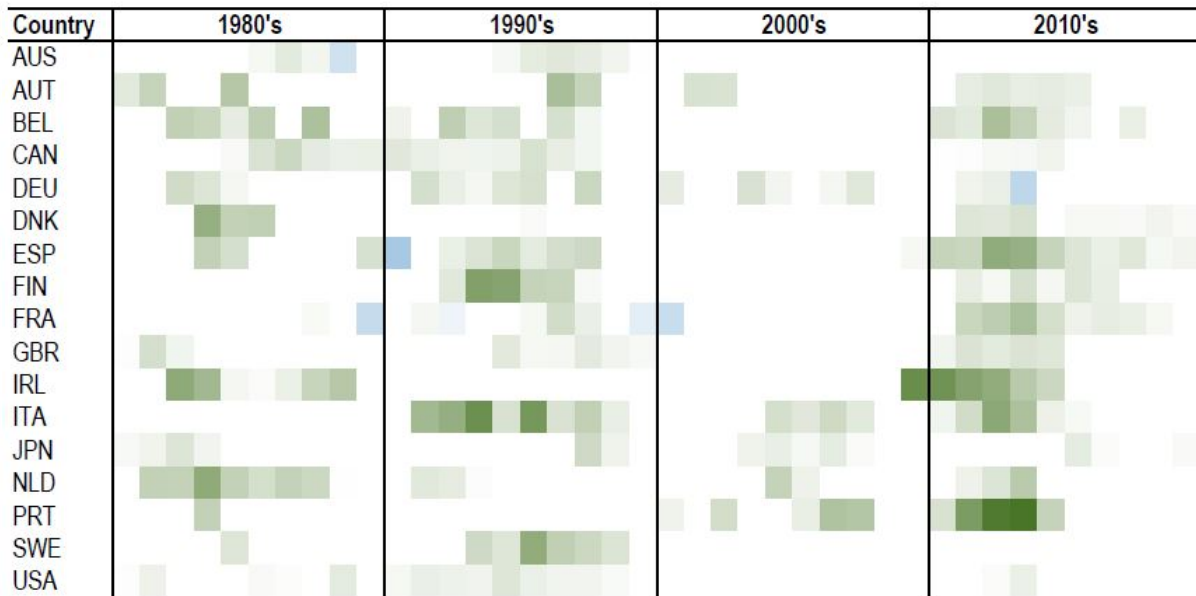
Figure 1 – Evolution of Debt to GDP Ratios



Notes:

Source: IMF World Economic Outlook (WEO) Database

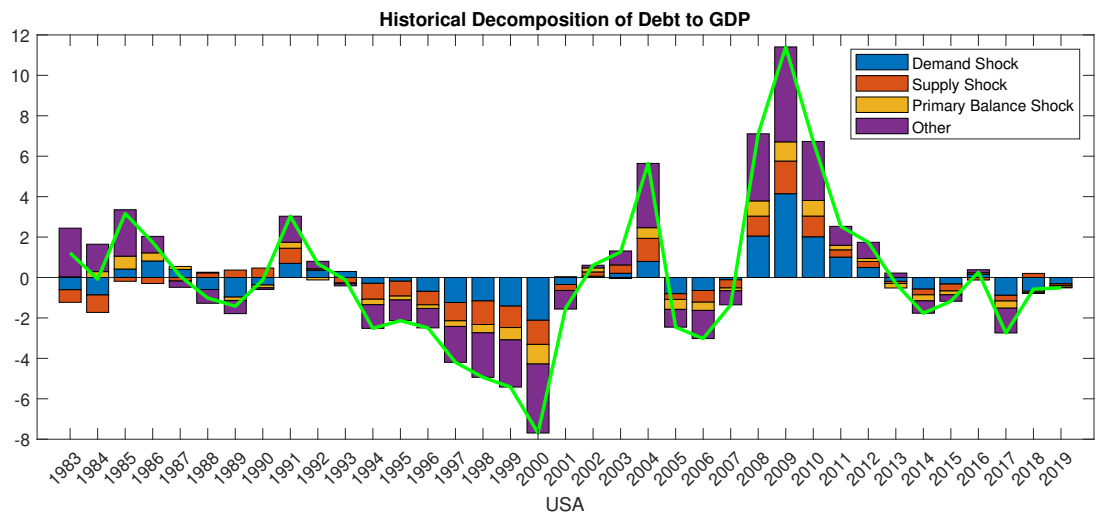
Figure 2 – Narrative Shocks in Advanced Economies



Notes: Green and blue denote positive and negative narrative primary balance shocks.

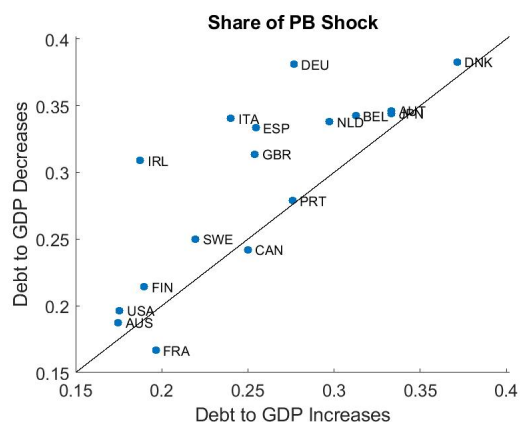
Sources: [Guajardo et al. \(2014\)](#)

Figure 3 – Historical Decomposition of Debt to GDP year-on year Changes in the US



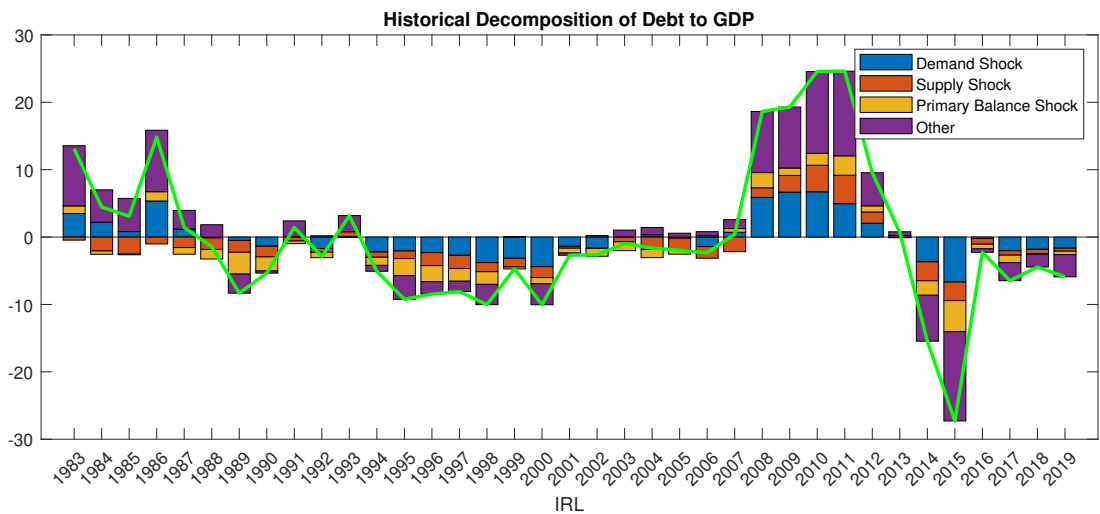
Notes: Based on the baseline VAR with narrative sign restrictions described in the text. The "other" category combines unidentified VAR shocks as well as the initial condition.

Figure 4 – Contribution of Primary Balance Shock to the Total Contribution of Identified Shocks to Debt Ratio Increases and Decreases



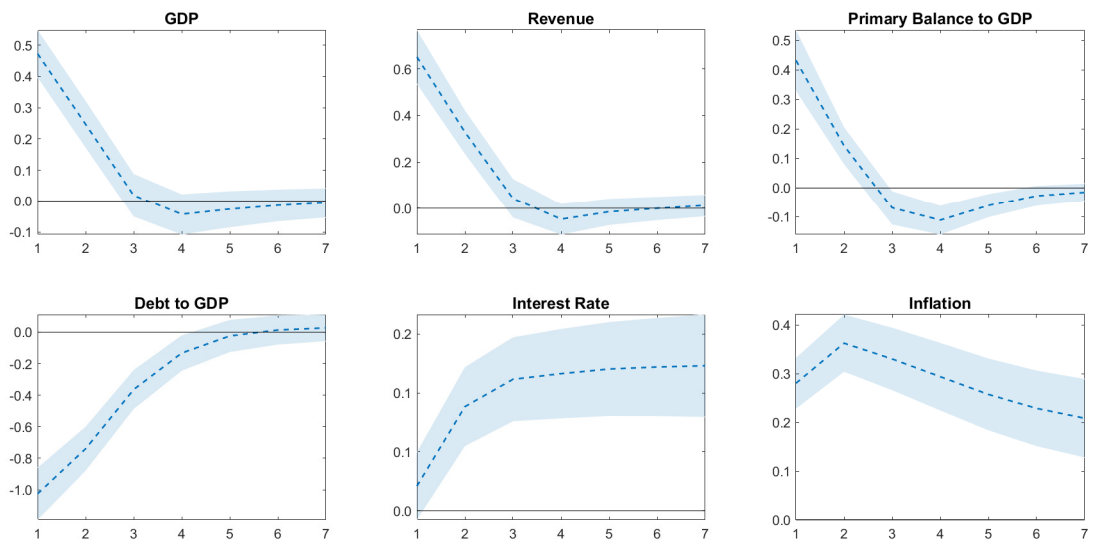
Notes: Notes: Median (over time) of posterior means across countries.

Figure 5 – Historical Decomposition of Debt to GDP year-on year Changes in Ireland



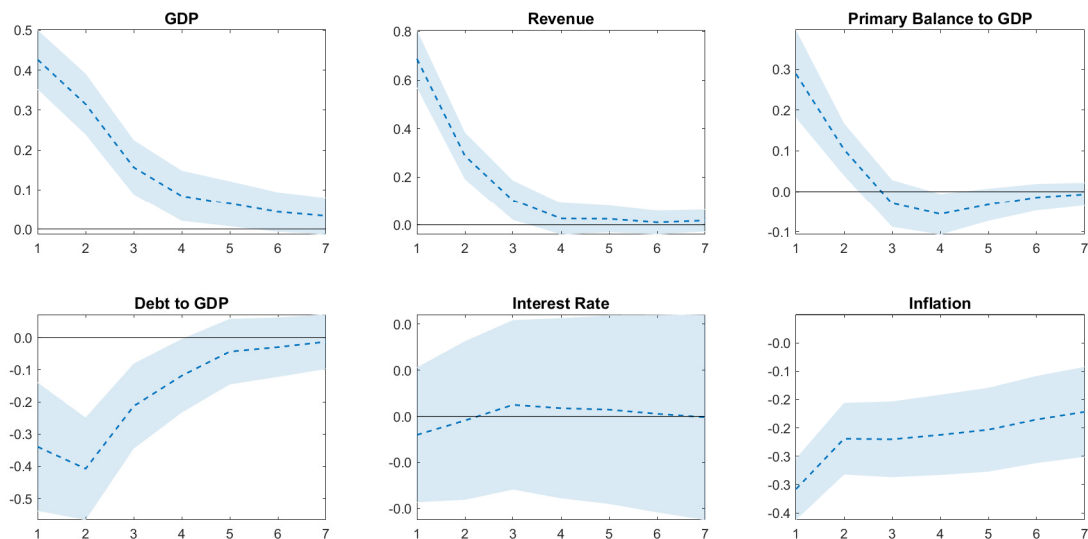
Notes: Based on the baseline VAR with narrative sign restrictions described in the text. The "other" category combines unidentified VAR shocks as well as the initial condition.

Figure 6 – Impulse Response: Demand Shock



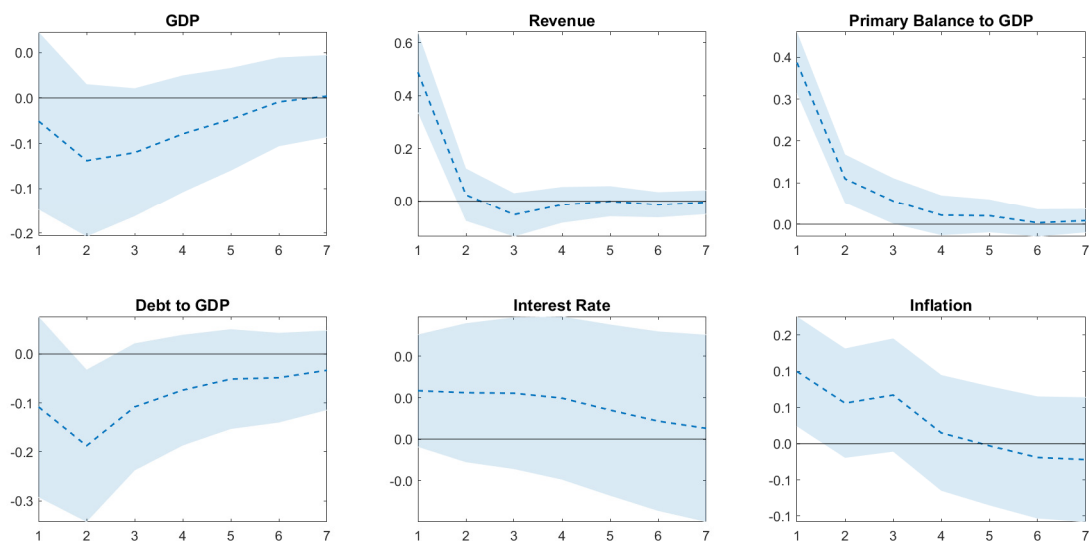
Notes: Medians and 68th percentile confidence intervals. Impulse responses are weighted using inverse variance weights. Based on an annual sample of 17 advanced economies from 1981-2019.

Figure 7 – Impulse Response: Supply Shock



Notes: Medians and 68th percentile confidence intervals. Impulse responses are weighted using inverse variance weights. Based on an annual sample of 17 advanced economies from 1981-2019.

Figure 8 – Impulse Response: Primary Balance Shock



Notes: Medians and 68th percentile confidence intervals. Impulse responses are weighted using inverse variance weights. Based on an annual sample of 17 advanced economies from 1981-2019.

Figure 9 – Primary Balance Shocks in Narrative Periods: With and Without Narrative Restrictions

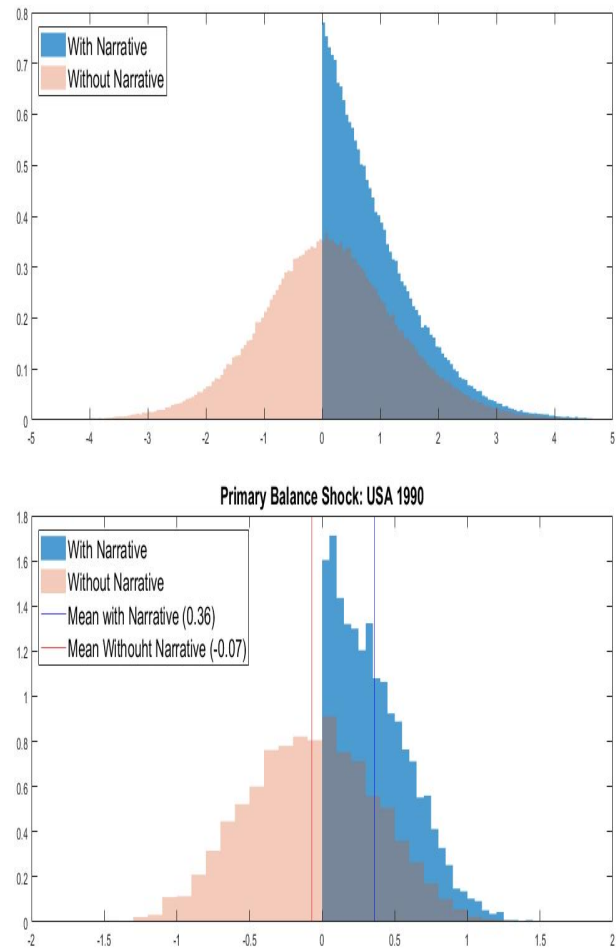


Figure 10 – The Impacts of Narrative Information

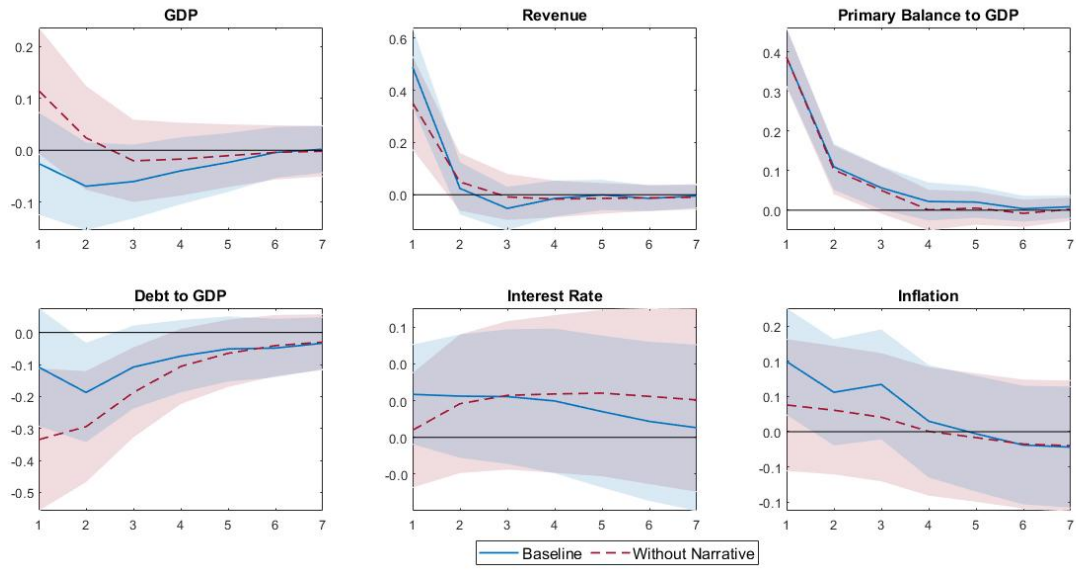
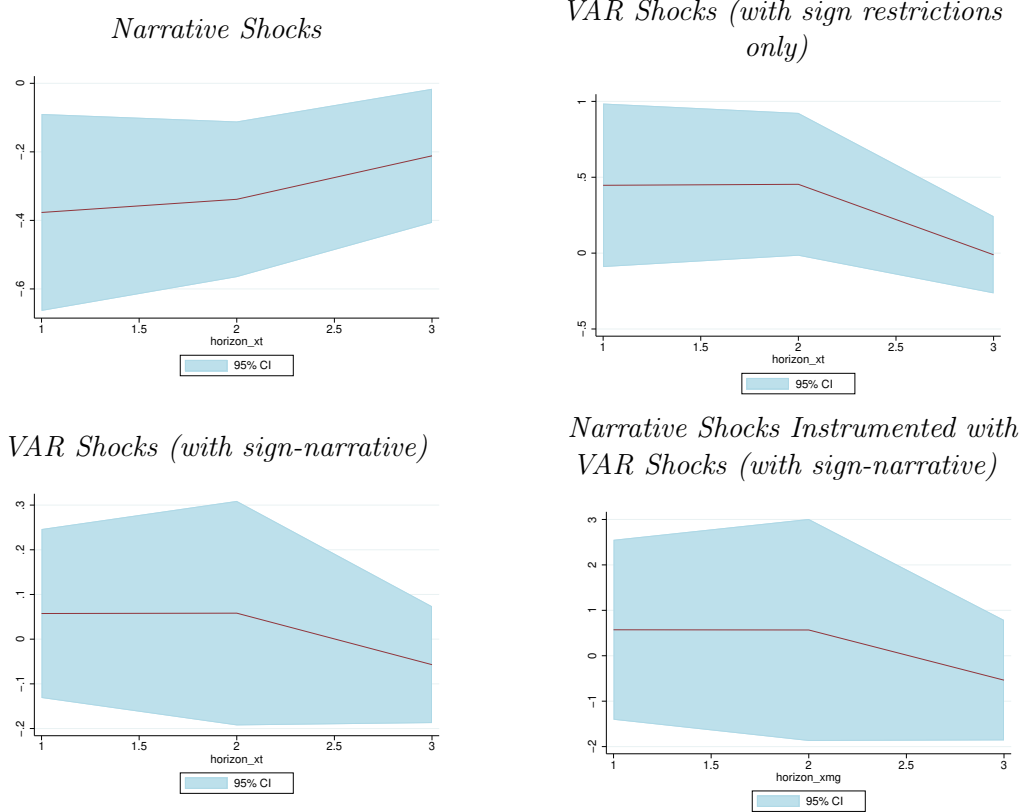


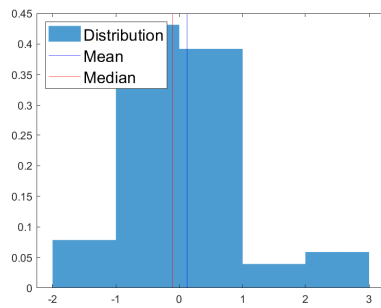
Figure 11 – Panel Fixed Effects Local Projections: Narrative Shocks vs VAR Shocks



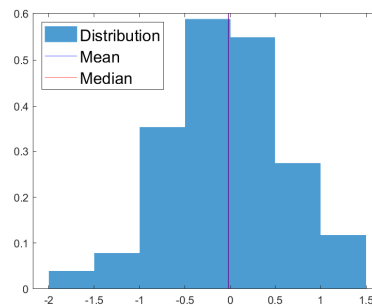
Notes: Impulse response of GDP growth to primary balance consolidation shocks. The regression controls for lagged dependent variable and country and time fixed effects. Standard errors clustered by country. Sample of 17 advanced economies from 1983-2019.

Figure 12 – Cross-Country Heterogeneity in Impact of Consolidations on Growth

Local Projections using Narrative Shocks



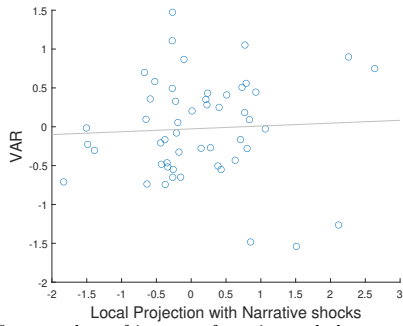
VAR with Sign-Narrative Restrictions



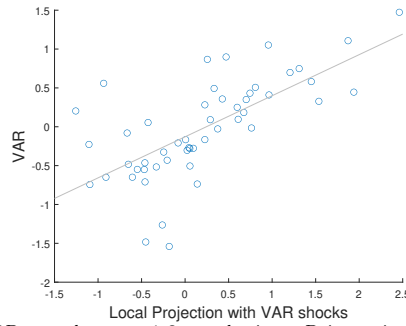
Notes: Histograms of the impact of a primary balance consolidation on GDP growth over a 1-3 year horizon. Point estimates for local projections. Medians for VARs.

Figure 13 – Impact of Consolidations on Growth at the Country Level: Correlations across methods

*VAR vs Local Projections using
Narrative Shocks
(Correlation <0.1)*

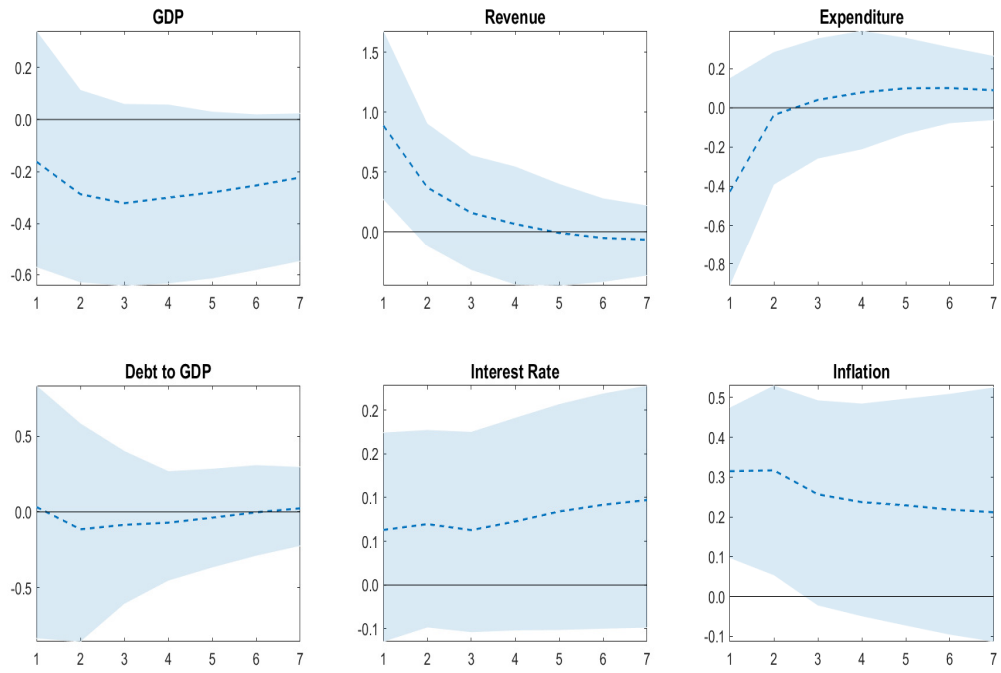


*VAR vs Local Projections Using VAR
Shocks
(Correlation =0.7)*



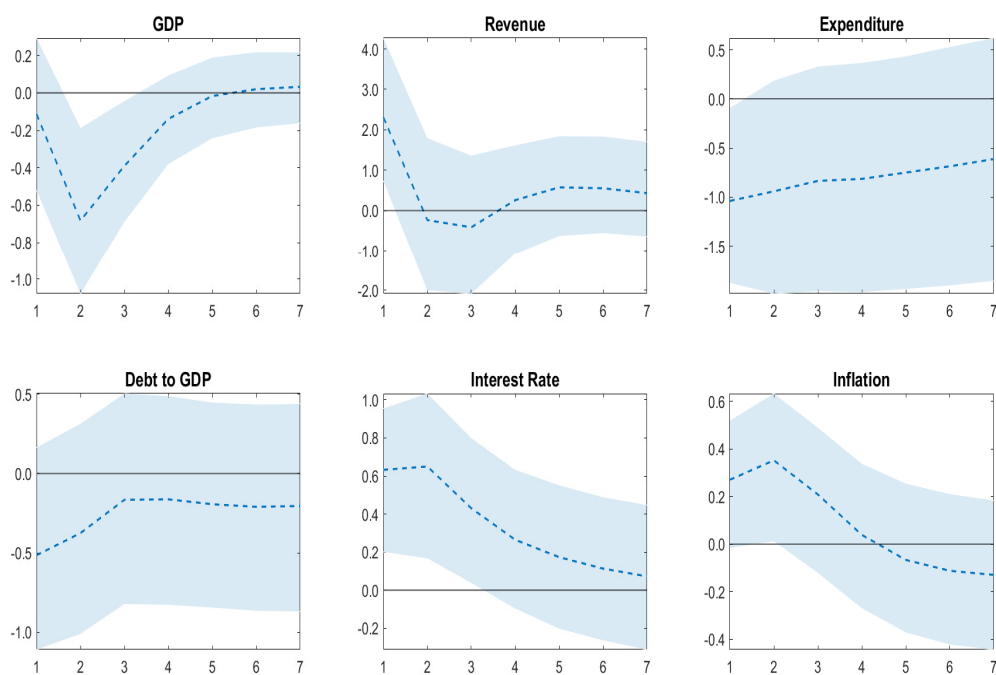
Notes: Scatter plots of impact of a primary balance consolidation on GDP growth over a 1-3 year horizon. Point estimates for local projections. Medians for VARs.

Figure 14 – Impulse Response to a Government Revenue Shock for the US (1981-2019)



Notes: Median and 16-84th percentile confidence intervals.

Figure 15 – Impulse Response to a Government Revenue Shock for the US (1952-2006) using *Mertens and Ravn (2013)* Shocks as Narrative Restrictions



Notes: Median and 16-84th percentile confidence intervals. Tax increase shocks from *Mertens and Ravn (2013)* database imposed as narrative restrictions.

Table 1 – *Sign-restrictions-imposed*

Sign restrictions imposed on the VAR impulse responses

	GDP Growth	Revenue Growth	Primary Balance/GDP	Interest Rate	Inflation	Debt to GDP
Growth Shocks						
Demand Shock	+	+			+	
Supply Shock	+	+			-	
Discretionary PB Shock			+			

Notes: The table denotes the impact sign restrictions imposed on the impulse response of the identifies shocks (along the row) on variables (along the column) in the VAR. Empty cells imply that no sign restriction is imposed for the specific shock/variable pair. In the VAR, the primary balance to GDP and debt to GDP ratios are included in first difference to reduce the influence of deterministic trends in these variables dominating the historical decomposition

Table 2 – Summary of Historical Decomposition

	Median	'DEU'	'DNK'	'ESP'	'FIN'	'FRA'	'GBR'	'IRL'	'ITA'	'NLD'	'PRT'	'USA'
Growth Shocks	41	36	38	42	45	44	42	43	35	37	43	46
Demand	24	19	19	25	28	25	29	26	23	22	22	27
Supply	17	17	19	17	17	19	13	17	12	15	21	19
Primary Balance Shock	16	17	23	17	11	11	16	15	13	18	16	11
Total: Identified Shocks	57	53	61	59	56	55	58	58	48	53	59	57

Notes: Means based on 5000 posterior draws for each country based on a SVAR with narrative sign restrictions estimated using Minnesota Prior.

Table 3 – Contribution of Identified Shocks to Debt Ratio Increases vs Decreases

	Overall	Debt to GDP Increase	Debt to GDP decrease
Growth Shocks	41	40	43
Demand Shock	24	25	25
Supply Shock	17	15	18
Primary Balance Shock	16	14	18
Total: Identified Shocks	57	54	61

Notes: Medians across countries reported separately for years with debt ratio increases vs decreases based on a SVAR with narrative sign restrictions estimated using Minnesota Prior.

Table 4 – Contribution of Identified Shocks: Sample Average vs Period with Large Changes in Debt Ratios

	Overall	Years with Large Changes in Debt to GDP
Growth Shocks	41	39
Demand Shock	24	24
Supply Shock	17	15
Primary Balance Shock	16	15
Total: Identified Shocks	57	54

Notes: Medians across countries reported separately for the overall sample in the first column and conditional on years with 5 largest increases and 5 largest decreases in the debt ratio within a country.

Table 5 – Historical decomposition vs Forecast Error Variance

	Forecast Error Variance Decomposition	Historical Decomposition
Growth Shocks	38	41
Demand Shock	22	24
Supply Shock	16	17
Primary Balance Shock	15	16
Total: Identified Shocks	53	57

Table 6 – Narrative Information Tightens HPD Credible Sets

	Demand Shock	Supply Shock	Primary Balance Shock
GDP Growth	0.01	-1.32	-3.49
Revenue Growth	-0.65	-1.78	-1.62
Primary Balance	-12.62	-7.96	-5.19
Debt	-0.77	-2.55	-2.16
Interest Rate	0.22	-2.09	-1.26
Inflation	-0.1	-2.69	-1.14

Notes: % difference between HPD sets for impulse responses in the VAR with narrative-sign restrictions relative to only sign restrictions, cumulated over the impulse response horizon of 6 years. A negative denotes smaller HPD sets for impulse response in the VAR with narrative-sign restrictions relative to only sign restrictions, and a cross mark indicates the opposite.

Table 7 – *R squared diagnostic for Non-Fundamentalness*

	Contemporaneous and lagged controls	Contemporaneous and lagged controls, plus forecasts
Narrative Shocks	0.471	0.582
SVAR (Sign-narrative)	0.104	0.156
SVAR (sign-only)	0.239	0.424
Observations	476	407

Notes: The table shows the coefficient of determination (R squared) based on panel regressions of shocks specified in the first column on a list of explanatory variables as well as country fixed effects based on a sample of 17 countries from 1990-2019. Explanatory variables include two lags of the levels of the primary balance to GDP, real government revenues, real GDP, the output gap and the VIX. The column with forecasts includes, in addition to the previous explanatory variables, one and five year ahead forecasts of debt to GDP, GDP growth and primary balance to GDP.

Table 8 – Orthogonality of Shocks with respect to Macroeconomic Forecasts

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	$\Delta GDP_{t-1,t}^f$		$\Delta GDP_{t-1,t+4}^f$		$\Delta \left(\frac{Debt}{GDP} \right)_{t-1,t}^f$		$\Delta \left(\frac{Debt}{GDP} \right)_{t-1,t+4}^f$	
VAR shocks	-0.044 (0.041)		-0.045 (0.040)		0.51 (0.42)		1.17 (0.69)	
Narrative Shocks		-0.056 (0.058)		-0.068 (0.059)		1.04** (0.40)		3.38*** (1.07)
Constant	0.086*** (0.0011)	0.11*** (0.021)	0.22*** (0.0011)	0.25*** (0.021)	0.040*** (0.011)	-0.33** (0.15)	-4.97*** (0.018)	-6.18*** (0.39)
Observations	408	408	408	408	407	407	407	407
R-squared	0.001	0.002	0.001	0.003	0.004	0.030	0.005	0.070

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.

Notes: Based on a sample of 17 countries from 1995-2019. All regressions include country fixed effects. Standard errors clustered by country. $X_{t,-1t+h}^f$ denotes the h period year ahead forecast of variable X made in period $t - 1$.

Table 9 – Departures from the Baseline Specification

	Baseline	Jeffrey's Prior	VAR-X	Monetary Policy Shock
Growth Shocks	41	35	18	39
Demand Shock	24	21	11	22
Supply Shock	17	14	7	17
Primary Balance Shock	16	14	7	14
US (10 year) Nominal Rate	0	0	40	
Monetary Policy Shock				14
Residual	44	50	35	33
Growth Shocks/Primary Balance Shock	2.56	2.50	2.57	2.78

Notes: “Baseline” denotes the baseline specification discussed in earlier sections, where the six variable VAR is estimated with two lags using the Minnesota prior. The second column uses the Jeffrey’s prior instead of Minnesota. The third column adds US nominal 10 year yields as an exogenous variable to the VAR in addition to the 6 endogenous variables. The fourth column adds an additional monetary shock, which is identified with inflation and the effective interest rate moving in opposite directions on impact.



PUBLICATIONS

Public Debt Dynamics and the Impact of Fiscal Policy
Working Paper No. WP/2024/087