

A narrative approach to a fiscal DSGE model

THORSTEN DRAUTZBURG
Federal Reserve Bank of Philadelphia

Structural DSGE models are used for analyzing both policy and the sources of business cycles. Conclusions based on full structural models are, however, potentially affected by misspecification. A competing method is to use partially identified SVARs based on narrative shocks. This paper asks whether both approaches agree. Specifically, I use narrative data in a DSGE-SVAR that partially identify policy shocks in the VAR and assess the fit of the DSGE model relative to this narrative benchmark. In developing this narrative DSGE-SVAR, I develop a tractable Bayesian approach to proxy VARs and show that such an approach is valid for models with a certain class of Taylor rules. Estimating a DSGE-SVAR based on a standard DSGE model with fiscal rules and narrative data, I find that the DSGE model identification is at odds with the narrative information as measured by the marginal likelihood. I trace this discrepancy to differences in impulse responses, identified historical shocks and policy rules. The results indicate monetary accommodation of fiscal shocks.

KEYWORDS. Fiscal policy, monetary policy, DSGE model, Bayesian estimation, narrative shocks, Bayesian VAR.

JEL CLASSIFICATION. C32, E32, E52, E62.

1. INTRODUCTION

Dynamic Stochastic General Equilibrium (DSGE) models are a widespread research and policy tool, as showcased in the influential work of [Christiano, Eichenbaum, and Evans \(2005\)](#) and [Smets and Wouters \(2007\)](#). But as structural models, DSGE models could be misspecified and their counterfactual predictions therefore misleading. [Faust \(2009\)](#), for example, argued that microfoundations of DSGE models are weak and that their quantitative success could mask misspecification. [Sims \(2005\)](#) cautioned that DSGE models should not displace alternative identification schemes. Narrative approaches (e.g., [Romer and Romer \(2004\)](#)), are one such alternative. They analyze the effects of the same

Thorsten Drautzburg: tdrautzburg@gmail.com

I am grateful to Pooyan Amir-Ahmadi, Jesus Fernandez-Villaverde, Pablo Guerron-Quintana, my discussant Ed Herbst, Jim Nason, Barbara Rossi, Juan Rubio-Ramirez, Frank Schorfheide, Tom Stark, Keith Sill, Harald Uhlig, Mark Watson, Jonathan Wright and audiences at the Atlanta Fed, the 2014 EEA, the ECB, the 2015 ESWC, Emory, Goethe, Madison, the 2015 NBER-NSF SBIES and NBER SI, the Philadelphia Fed, the Riksbank, the 2014 SED, the Fall 2015 System Macro Committee, and Wharton for comments and suggestions. I would also like to thank Valerie Ramey, Karel Mertens and Morten Ravn for making their code and data publicly available and Nick Zarra for excellent research assistance. Any errors are mine. The views expressed herein are my own. They do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia, the Federal Reserve System or its Board of Governors.

© 2020 The Author. Licensed under the [Creative Commons Attribution-NonCommercial License 4.0](#). Available at <http://qeconomics.org>. <https://doi.org/10.3982/QE1083>

shocks as the typical DSGE model. Narrative methods rely on fewer structural assumptions and incorporate additional information relative to standard macroeconomic time series. I provide a framework for incorporating this information in DSGE model estimation and quantifying misspecification compared with narrative studies.

Existing approaches to DSGE model misspecification focus on statistical specification. As surveyed by [Fernández-Villaverde, Rubio-Ramírez, and Schorfheide \(2016\)](#), there are various approaches for addressing model misspecification. For example, impulse-response function (IRF) matching as in [Christiano, Eichenbaum, and Evans \(2005\)](#) side-steps fully parametrizing all shock processes of the DSGE model, which may introduce misspecification. However, such limited-information approaches give up on tracking a large number of macroeconomic indicators and shocks, which is one of the main successes of DSGE models such as [Smets and Wouters \(2007\)](#). The DSGE-VAR framework of [Del Negro and Schorfheide \(2004\)](#) is a full information framework that quantifies model misspecification relative to a reduced-form vector-autoregression (VAR).¹ [Del Negro, Schorfheide, Smets, and Wouters \(2007\)](#) used this framework to show that DSGE models are subject to a small degree of misspecification. But since their benchmark model is a reduced-form VAR, their metric for comparing models is unaffected by shock identification.

This paper introduces a structural VAR (SVAR), identified using narrative shock proxies, to assess DSGE-model misspecification. It extends the DSGE-VAR framework for assessing misspecification in reduced-form VARs to a DSGE-SVAR, partially identified with external instruments. The DSGE-SVAR allows us to quantify the misspecification of the DSGE identification scheme by comparing marginal likelihoods: If a given DSGE model and the narrative SVAR agree, including information from the model increases precision, and hence the marginal likelihood. As in [Park \(2011\)](#), I distinguish the fit of the DSGE dynamics from the fit of the DSGE identification. The innovation of this paper is to incorporate external data to build a DSGE-SVAR that partially identifies the model. I show that the DSGE-SVAR can imply substantive differences in IRFs compared to the traditional DSGE-VAR.

The DSGE-SVAR builds on a simple implementation of a Bayesian proxy SVAR. This simple proxy SVAR implementation is an additional contribution of this paper. Similar to [Caldara and Herbst \(2019\)](#), I use a Bayesian approach to implement the idea of the proxy-VAR first proposed by [Stock and Watson \(2012\)](#) and [Mertens and Ravn \(2013\)](#) in frequentist settings. Both [Caldara and Herbst \(2019\)](#) and my approach simply append a measurement equation to the VAR equation, but they use a structural VAR as the starting point. In contrast, I use a reduced form VAR as the statistical model and inference with a flat prior that uses a textbook Gibbs sampler. The structural implications follow from closed-form expressions based on the reduced form, similar to a standard VAR such as [Uhlig \(2005\)](#). I elicit the VAR prior indirectly via the structural parameters of a DSGE model, but add priors for the measurement equation similar to [Caldara and Herbst \(2019\)](#). They use a Minnesota prior and priors for the coefficient and

¹[Waggoner and Zha \(2012\)](#) confronted DSGE model misspecification differently. Instead of estimating a constant mixture between the DSGE model and the VAR, they allow for Markov switching between the models.

signal-to-noise ratio of the measurement equation. With a proper prior, the models both here and in [Caldara and Herbst \(2019\)](#) require a Metropolis-within-Gibbs sampler. Unlike their setting, I allow for possibly missing instruments, an important issue in applied work.

My result that the proxy SVAR identification is valid for a class of DSGE models is the second building block for my DSGE-SVAR: I provide conditions under which the instrument-identified VAR in [Mertens and Ravn \(2013\)](#) correctly identifies shocks and policy rule coefficients in models with standard Taylor-type policy rules, such as [Leeper, Plante, and Traum \(2010\)](#) and [Fernandez-Villaverde, Guerron-Quintana, Kuester, and Rubio-Ramirez \(2015\)](#). This property of the narrative VAR contrasts with traditional VARs that identify shocks through contemporaneous zero restrictions. DSGE models that match the VAR then need to assume that economic agents only react to policy shocks with a delay. The narrative VAR approach is valid if the data are generated from a widely used class of DSGE models, without restricting the timing. The key condition for my result is that the information set in the VAR captures the variables policy-makers pay attention to. This theoretical result mirrors the empirical result in [Caldara and Herbst \(2019\)](#) that credit spreads may be an important policy variable.

My application contributes to the literature on fiscal and monetary policy DSGE models, which is important from a substantive point of view: With monetary policy constrained by the zero lower bound (ZLB), “stimulating” fiscal policy has gained a lot of attention and influential papers such as [Christiano, Eichenbaum, and Rebelo \(2011\)](#) have used quantitative DSGE models for the analysis of fiscal policies. Since the fiscal building blocks of DSGE models are less well studied than, say, the Taylor rule for monetary policy (e.g., [Clarida, Galí, and Gertler \(2000\)](#)), assessing the fiscal policy implications of these models is warranted. I use narrative measures for government spending, tax rates and monetary policy shocks, building on the work in [Ramey \(2011\)](#), [Mertens and Ravn \(2013\)](#) and [Romer and Romer \(2004\)](#), to estimate a standard medium-scale DSGE-SVAR with Taylor rules for fiscal and monetary policy. Even though my results show that fiscal DSGE models can match VAR estimates along some dimensions, my overall results caution against the use of standard DSGE models for fiscal policy analysis. However, this finding is conditional on the assumption that the proxy variables are good instruments.

Specifically, I find that the best-fitting model puts significant weight on DSGE model dynamics, but the data prefer the weakest prior on the DSGE model covariance structure. This indicates that the DSGE model does not agree with the narrative SVAR identification. These results are robust to estimating the model with news shocks and expectations data, and for other model variants. To provide evidence why the data dislikes the DSGE model specification, I analyze the implied impulse-response functions, historical shocks and policy rules. I find that monetary policy accommodates fiscal shocks in the impulse-response function analysis. While the corresponding pure DSGE can match these qualitative findings through the indirect fiscal effects on inflation and output, my analysis of policy rules reveals a systematic direct response. I also find responses to monetary shocks that resemble a price-puzzle that the pure DSGE model cannot match. These responses to monetary policy shocks are robust to different specifications such

as controlling for expectations. Still, the pure DSGE model matches the historical monetary policy shocks best, but struggles to match historical tax shocks.

Showing the model fit along several dimensions highlights the advantage of using likelihood-based inference: The VAR has implications for several structural objects that are all informative about the success of the DSGE model that I analyze. In contrast, limited-information methods such as impulse-response function matching try to maximize the fit along only one of the structural dimensions.

For the estimation, I use previously unused data on expectations for the identification. Using the historic, publicly available Greenbook records, I digitize quarterly expectations data on the different components of fiscal policy along with the economy as a whole. I use this data to update the instrument for short-term government spending in Ramey (2011), that is, defense spending surprises. Extending the sample on proxy variables sharpens inference significantly. I also use the data to control for fiscal foresight in a model extension featuring news shocks. Fiscal foresight does affect some estimates, but leaves the main conclusions unchanged.

Since it provides an intuitive way to incorporate prior information, the narrative DSGE-VAR framework can also be of interest for future narrative studies. In my application, the posterior distributions over the effects of fiscal shocks remain wide when only few proxy measures are available. Even though here I could add government spending shock proxies through extra data work, this is not possible for other applications, such as taxes. In such cases, incorporating prior information sharpens the posterior by shrinking the estimates toward theory-consistent policy rules. This is similar to the work by Arias, Caldara, and Rubio-Ramirez (2015), who document that even weak structural priors over policy rules help to identify shocks via sign restrictions.

This paper is structured as follows: Section 2 gives an overview of the methodological approach. Section 3 describes methods in detail and may be skipped by some practitioners. Sections 4 and 5 describe the empirical specification and the empirical results. Appendices A and B can be found in the Online Supplemental Material (Drautzburg (2020)). Appendix A contains the proofs and Appendix B describes the data and computational details. Online Appendix C, available on the author's website, also shows additional empirical results.² In what follows, I use regular font letters for scalar variables, bold lower case letters for vectors and bold upper case letters for matrices.

2. NARRATIVE DSGE-SVARs: THE IDEA

The DSGE-VAR approach in Del Negro and Schorfheide (2004) estimates a weighted average of a reduced-form VAR and a fully structural DSGE model. For estimating VARs, this approach is useful because it allows to elicit priors for the many reduced-form VAR parameters via priors for relatively few structural parameters of a DSGE model. Since the structural parameters typically have a straightforward interpretation, eliciting priors is simpler in the DSGE model.

²https://sites.google.com/site/tdrautzburg/research/NarrativeDSGE_web.pdf.

The DSGE-VAR also allows to quantify whether the economy is well described by the structural DSGE model. Since VARs typically summarize the dynamics of macroeconomic movements well, comparing the DSGE model fit to that of the VAR yields a meaningful measure of misspecification, based on the reduced-form statistical properties. Formally, the larger the weight the best-fitting DSGE-VAR puts on the DSGE model, the lower the concern about misspecification.

This paper takes the approach to measuring DSGE model misspecification further. First, it replaces the reduced-form VAR with a structural proxy-VAR. This SVAR itself allows causal inference on impulse-response functions, historical structural shocks and of policy rules, when applied to policy shocks. It thus goes beyond assessing the purely statistical fit and also speaks to the identifying assumptions about structural relationships in the DSGE model.

If the dynamic relationships between variables are pinned down, identification in the proxy VAR is governed by the contemporaneous covariance between forecast errors and proxy variables. I therefore measure DSGE model misspecification along two dimensions: (1) the fit of macroeconomic dynamics, and (2) the contemporaneous covariance structure that implies the causal relationships. The weight on the DSGE model is my summary measure of the model fit and corresponds to the prior precision. An intuitive measure of this precision is how many observations T_0^B and T_0^V it would take to arrive at that prior precision for the model dynamics and the covariance structure.

2.1 Formal setting

Formally, write the VAR with p lags in the (zero-mean) observables \mathbf{y}_t and its companion form as

$$\mathbf{y}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{A}\boldsymbol{\varepsilon}_t, \quad (2.1a)$$

$$\mathbf{x}_t = \begin{bmatrix} \mathbf{B} & \\ \mathbf{I}_{(p-1)m} & \mathbf{0} \end{bmatrix} \mathbf{x}_{t-1} + \begin{bmatrix} \mathbf{A} \\ \mathbf{0} \end{bmatrix} \boldsymbol{\varepsilon}_t, \quad \mathbf{x}_t \equiv \begin{bmatrix} \mathbf{y}'_t & \mathbf{y}'_{t-1} & \cdots & \mathbf{y}'_{t-(p-1)} \end{bmatrix}', \quad (2.1b)$$

where the lag coefficients \mathbf{B} and the covariance matrix $\boldsymbol{\Sigma} \equiv \mathbf{A}\mathbf{A}'$ are identified from the data, but not \mathbf{A} . $\boldsymbol{\varepsilon}_t$ are structural shocks of the same dimension as \mathbf{y}_t . The corresponding linearized DSGE model has the following state-space representation:

$$\mathbf{y}_t = \mathbf{B}^*(\boldsymbol{\theta})\mathbf{x}_{t-1}^* + \mathbf{A}^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*, \quad (2.2a)$$

$$\mathbf{x}_t^* = \mathbf{D}^*(\boldsymbol{\theta})\mathbf{x}_{t-1}^* + \mathbf{C}^*(\boldsymbol{\theta})\boldsymbol{\varepsilon}_t^*, \quad (2.2b)$$

where $\mathbb{E}[\boldsymbol{\varepsilon}_t^*(\boldsymbol{\varepsilon}_t^*)'] = \mathbf{I}_m$. m is the dimensionality of \mathbf{y}_t . $\boldsymbol{\theta}$ is a vector of DSGE model parameters. The structural shocks here are $\boldsymbol{\varepsilon}_t^*$. Let $\boldsymbol{\Sigma}^* \equiv \mathbf{A}^*(\mathbf{A}^*)'$ be the covariance matrix of \mathbf{y}_t .

By itself, the VAR in equations (2.1) is observationally equivalent to VARs with other \mathbf{A} matrices. We can, therefore, compare the DSGE model in equations (2.2) only to reduced form moments—VAR dynamics and forecast-error covariances. To build a DSGE-SVAR,

I add a measurement equation to both models. Assuming that the VAR spans the true shocks ε_t^* , this observation equation is given by

$$\mathbf{z}_t = \begin{bmatrix} \mathbf{G} & \mathbf{0} \end{bmatrix} \varepsilon_t + \text{noise}_t. \quad (2.3)$$

The zero restriction in the observation equation for the instruments breaks the observational equivalence between various structural VARs with the same reduced form presentation. Adding this equation to the VAR and the DSGE model allows me to evaluate whether the DSGE model is able to match the static correlations between the shocks. These correlations are determined by the restriction that, except for noise, $\mathbf{z}_t = [\mathbf{G}, \mathbf{0}] \varepsilon_t$, and $\mathbf{v}_t = \mathbf{A} \varepsilon_t$. In addition to the covariance matrix Σ of the forecast errors \mathbf{v}_t , the covariance Γ between instruments and forecast errors now enters the likelihood of the model. Consequently, it also influences the optimal weight put on the DSGE model and the VAR.

Formally, under normality, the likelihood of equation (2.1a) is unchanged when we consider some $(\tilde{\mathbf{A}}, \tilde{\varepsilon})$ instead of $(\mathbf{A}, \varepsilon)$ for any $\tilde{\mathbf{A}} \equiv \mathbf{A}\mathbf{Q}$ and $\tilde{\varepsilon} \equiv \mathbf{Q}'\varepsilon$, where \mathbf{Q} is an orthonormal rotation matrix: $\tilde{\mathbf{A}}(\tilde{\mathbf{A}})' = \mathbf{A}\mathbf{A}'$ and $\tilde{\varepsilon} \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is equal in distribution to ε . In contrast, the zero restriction in (2.3) implies that Γ varies with orthonormal rotations of the shocks, unlike Σ . It then becomes possible to evaluate the structural fit of the DSGE model along with its statistical fit, which is my main focus.

DSGE-VARs also help to estimate the dynamic coefficients \mathbf{B} , and I also measure the DSGE-model statistical fit along this “dynamic” dimension. \mathbf{B} has a number of entries equal to $m^2 \times p$, where m is the number of variables and p the number of lags. Even in medium-sized VARs \mathbf{B} is large, and prior information on it from the DSGE model may well help to improve the fit.

With the narrative instruments, one can assess model misspecification beyond the marginal likelihood of the DSGE-SVAR. The estimable covariance Γ of instruments and forecast errors and the structure that equation (2.3) imposes on Γ allows one to compare the VAR and the DSGE model along several dimensions, such as structural IRFs, implied policy rules and historical shock paths. While analyzing these structural implications typically requires extra assumptions beyond the structure in (2.3), the data on the instruments allows one to proceed without assuming that the same mapping from the reduced-form DSGE covariance matrix Σ^* to \mathbf{A}^* also applies to (Σ, \mathbf{A}) in the VAR, as in [Del Negro and Schorfheide \(2004\)](#).

Traditional narrative analyses, such as [Romer and Romer \(2004\)](#), are limited to analyzing impulse-response-functions to a single shock. [Antolin-Diaz and Rubio-Ramirez \(2016\)](#) also analyzed impulse-response functions in VARs, but they only look at a few, large realizations of a shock. Compared to my approach, the merit of these limited information approaches is that they get by with partially specified models. In contrast, I embed the assumption that narrative measures such as [Romer and Romer \(2004\)](#) are correct and incorporate them into a larger structural model to assess DSGE models as a competing approach. DSGE models identify shocks through fully parametrizing the economy, including shock processes. Similar to DSGE models, my approach is a full information framework that, nevertheless, gets by with fewer identifying assumptions than a full-fledged DSGE model.

2.2 Example of identification through narrative proxies

To illustrate the identification through narrative proxies and contrast it with the identification in a fully structural model, consider the small fiscal SVAR in Blanchard and Perotti (2002) and Mertens and Ravn (2014). Blanchard and Perotti (2002) identified the model by calibrating a number of structural parameters, akin to the parametric identification of DSGE models. In contrast, Mertens and Ravn (2014) used shock proxies for the identification.

The VAR includes government spending g_t , tax revenues r_t , and output y_t :

$$\begin{bmatrix} g_t \\ r_t \\ y_t \end{bmatrix} = \mathbf{B} \begin{bmatrix} g_{t-1} \\ r_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} v_{g,t} \\ y_{r,t} \\ v_{y,t} \end{bmatrix}, \quad \begin{bmatrix} v_{g,t} \\ v_{r,t} \\ v_{y,t} \end{bmatrix} \equiv \mathbf{A} \begin{bmatrix} \epsilon_{g,t} \\ \epsilon_{r,t} \\ \epsilon_{y,t} \end{bmatrix}.$$

While \mathbf{B} and \mathbf{A} could be linked in DSGE models through their dependence on common parameters θ , I ignore \mathbf{B} in this example to focus on shock identification. It is useful to follow Blanchard and Perotti (2002) and Mertens and Ravn (2014) to rewrite $\mathbf{v}_t = \mathbf{A}\epsilon_t$ as

$$\begin{bmatrix} v_{g,t} \\ v_{r,t} \\ v_{y,t} \end{bmatrix} \equiv \begin{bmatrix} 0 & 0 & \eta_{gy} \\ 0 & 0 & \eta_{ry} \\ \kappa_{yg} & \kappa_{yr} & 0 \end{bmatrix} \begin{bmatrix} v_{g,t} \\ v_{r,t} \\ v_{y,t} \end{bmatrix} + \begin{bmatrix} \sigma_g & \gamma_{gr} & 0 \\ \gamma_{rg} & \sigma_r & 0 \\ 0 & 0 & \sigma_y \end{bmatrix} \begin{bmatrix} \epsilon_{g,t} \\ \epsilon_{r,t} \\ \epsilon_{y,t} \end{bmatrix}. \quad (2.4)$$

Blanchard and Perotti (2002) calibrated several of the parameters in their analysis of this SVAR, similar to how DSGE models use information outside the model to calibrate parameters or to set priors. The standard DSGE-VAR approach would, for a given parameter draw $(\hat{\theta}, \hat{\mathbf{A}}^*)$, back out the VAR counterpart $\hat{\mathbf{A}}$ as $\text{chol}(\Sigma(\hat{\theta})) \times \mathbf{Q}(\mathbf{A}^*(\hat{\theta}))$, where $\mathbf{Q}(\mathbf{A}^*(\hat{\theta})) = \text{chol}(\Sigma^*(\hat{\theta}))^{-1}(\mathbf{A}^*(\hat{\theta}))$.

In contrast, I follow the recent SVAR literature, using structural modeling and external instruments for identification. I show below that a class of policy rules in DSGE models implies zero restrictions on γ_o . Unlike Blanchard and Perotti (2002), who use institutional reasoning to calibrate η_{ry} and set $\eta_{gy} = 0$, I follow Mertens and Ravn (2014) to identify these parameters from external instruments: When two external instruments \mathbf{z}_t are available that jointly span the fiscal shocks, we can recover $\kappa = [\kappa_{yg}, \kappa_{yr}]$ from the data. The identifying assumption is that the instruments are orthogonal to ϵ_{yt} . Intuitively, one can then use the two instruments to identify the 1×2 vector κ , which, in turn, allows to back out ϵ_{yt} up to scale. Knowing ϵ_{yt} identifies $\eta = [\eta_{gy}, \eta_{ry}]$.

The DSGE-SVAR assesses which weight on the VAR and the DSGE model best fits the macro data $[g_t, r_t, y_t]$ as well as the instruments \mathbf{z}_t . For given dynamics, the model fit depends on the variances and covariances of the forecast errors of $[g_t, r_t, y_t]$ and \mathbf{z}_t , which are the identifying relationships in the data. I also assess how the DSGE model dynamics fit the data.

3. NARRATIVE (DSGE-)SVARS: IDENTIFICATION AND ESTIMATION

In this section, I first discuss identification and Bayesian estimation for the instrument-identified SVAR. The second part of this section discusses the link between SVAR and

DSGE models and outlines how to use a DSGE model for eliciting priors and to quantify its possible misspecification.

3.1 Narrative BVAR

Spelling out the distributional assumptions, the VAR with proxy variables is summarized by

$$\mathbf{y}_t = \boldsymbol{\mu}_y + \mathbf{B}\mathbf{x}_{t-1} + \mathbf{v}_t, \quad (3.1a)$$

$$\mathbf{v}_t = \mathbf{A}\boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \mathbf{I}_m), \quad (3.1b)$$

$$\mathbf{z}_t = \boldsymbol{\mu}_z + \begin{bmatrix} \mathbf{G} & \mathbf{0} \end{bmatrix} \boldsymbol{\varepsilon}_t + \mathbf{u}_t, \quad \mathbf{u}_t \stackrel{\text{i.i.d.}}{\sim} \mathcal{N}(\mathbf{0}, \boldsymbol{\Omega}). \quad (3.1c)$$

Here, \mathbf{y}_t is the $m \times 1$ data vector, \mathbf{x}_{t-1} stacks p lags of the data and deterministic terms, \mathbf{B} contains the lag coefficient matrices of the equivalent VAR(p) model, as well as constants and trend terms, \mathbf{v}_t is the vector of forecast errors and \mathbf{z}_t contains m_z narrative shock measures. $\boldsymbol{\varepsilon}_t$ are the structural shocks. \mathbf{u}_t is measurement error in the narrative shocks.

The standard assumptions of instrument validity and relevance correspond to two assumptions on $\boldsymbol{\Gamma}$, the covariance between instruments and forecast errors.

ASSUMPTION 1. *For some invertible square matrix G , the covariance matrix $\boldsymbol{\Gamma}$ can be written as*

$$\boldsymbol{\Gamma} \equiv \text{Cov}[\mathbf{z}_t, \mathbf{v}_t] = \begin{bmatrix} \mathbf{G} & \mathbf{0} \end{bmatrix} \mathbf{A}'. \quad (3.2)$$

The assumption that \mathbf{G} is invertible follows [Mertens and Ravn \(2013\)](#) and corresponds to the assumption that the instruments are relevant. The zero restriction is the standard exclusion restriction. It could be violated if the instruments load on shocks other than the ones intended, or if the VAR information set is insufficient. The VAR information set could be insufficient, for example, in the presence of news shocks, when the VAR only spans the shocks if the proxies are included in the VAR.

Stacking the equations in (3.1), the model can be written compactly as

$$\begin{bmatrix} \mathbf{y}_t \\ \mathbf{z}_t \end{bmatrix} \Big| \mathbf{y}^{t-1} \sim \mathcal{N} \left(\begin{bmatrix} \boldsymbol{\mu}_y + \mathbf{B}\mathbf{y}_{t-1} \\ \boldsymbol{\mu}_z \end{bmatrix}, \mathbf{V} \right), \quad \mathbf{V} \equiv \begin{bmatrix} \mathbf{A}\mathbf{A}' & \boldsymbol{\Gamma}' \\ \boldsymbol{\Gamma} & \tilde{\boldsymbol{\Omega}} \end{bmatrix} \quad (3.3)$$

where $\tilde{\boldsymbol{\Omega}} = \boldsymbol{\Omega} + [\mathbf{G}, \mathbf{0}][\mathbf{G}, \mathbf{0}]'$ is the covariance matrix of the narrative instruments.

The components of the variance-covariance matrix \mathbf{V} for the VAR stacked with the instruments embodies the restrictions that (partially) identify the shocks.

3.1.1 Identification given parameters [Stock and Watson \(2012\)](#) and [Mertens and Ravn \(2013\)](#) have shown how to identify shocks in a VAR with external instruments. This section largely follows [Mertens and Ravn \(2013\)](#). It considers the case of as many instruments as shocks to be identified, with $m_z \leq m$.

The following lemma summarizes results from the literature. Its notation uses partitions of \mathbf{A} and $\mathbf{\Gamma}$: $\mathbf{A} = [\mathbf{A}^{[1]}, \mathbf{A}^{[2]}] = \begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix}$, $\mathbf{A}^{[1]} = [\mathbf{A}'_{11}, \mathbf{A}'_{21}]'$ with $\mathbf{A}_{11} (m_z \times m_z)$ being invertible and where \mathbf{A}_{21} is $((m - m_z) \times m_z)$. Similarly, partition $\mathbf{\Gamma} = [\mathbf{\Gamma}_1, \mathbf{\Gamma}_2]$ with $\mathbf{\Gamma}_1$ being $m_z \times m_z$.

LEMMA 1 (Stock and Watson (2012), Mertens and Ravn (2013)). *Under Assumption 1, the impact of shocks with narrative instruments is generally identified up to an $m_z \times m_z$ scale matrix \mathbf{A}_{11} whose outer product $\mathbf{A}_{11}\mathbf{A}'_{11}$ is known given $\mathbf{\Sigma}$ and $\mathbf{\kappa}' = \mathbf{\Gamma}_1^{-1}\mathbf{\Gamma}_2$, requiring an extra $\frac{(m_z-1)m_z}{2}$ identifying restrictions and the impulse vector is given by $[\mathbf{I}, \mathbf{\kappa}']'\mathbf{A}_{11}$.*

PROOF. See Appendix A.1. □

Thus, for $m_z > 1$, the extra data by itself only identifies a set of IRFs, although the statistical fit of the model is invariant to the particular choice of \mathbf{A}_{11} . To uniquely characterize impulse responses or historical shocks, however, I need further assumptions to partially identify specific responses. Before discussing identification, it is worth pointing out that identification is not affected by instruments that are missing at random conditional on time $t - 1$ information, as the following corollary shows.

COROLLARY 1. *Assume that $\tilde{z}_{jt} = z_{jt}$ with probability $\phi_j \in (0, 1]$ and equal to its sample mean otherwise. Define $\tilde{\mathbf{\Gamma}} = \text{Cov}[\tilde{\mathbf{z}}_t, \mathbf{v}_t]$ and $\tilde{\mathbf{\kappa}}' = \tilde{\mathbf{\Gamma}}_1^{-1}\tilde{\mathbf{\Gamma}}_2$. Then $\tilde{\mathbf{\kappa}} = \mathbf{\kappa}$ and identification is unaffected by $\phi_j < 1$.*

PROOF. $\tilde{\mathbf{\Gamma}} = \sum_{j=1}^{m_z} \phi_j \text{Cov}[z_{j,t}, \mathbf{v}_t] = \sum_{j=1}^{m_z} \phi_j \mathbf{G}_j \circ \mathbf{A}' = [\text{diag}([\phi_j]_{j=1}^{m_z})\mathbf{G}, \mathbf{0}]\mathbf{A}'$. Then $\tilde{\mathbf{\kappa}}' = \tilde{\mathbf{\Gamma}}_1^{-1}\tilde{\mathbf{\Gamma}}_2 = (\mathbf{A}'_{11})^{-1}\mathbf{G}^{-1} \text{diag}([\phi_j]_{j=1}^{m_z})^{-1} \text{diag}([\phi_j]_{j=1}^{m_z})\mathbf{G}\mathbf{A}'_{12} = (\mathbf{A}'_{11})^{-1}\mathbf{A}'_{12} = \mathbf{\Gamma}_1^{-1}\mathbf{\Gamma}_2 = \mathbf{\kappa}$. Because $\mathbf{A}^{[1]} = [\mathbf{I}_{m_z}, \mathbf{\kappa}']'\mathbf{A}_{11}$ and $\mathbf{A}_{11}\mathbf{A}'_{11}$ depends on $\mathbf{\Sigma}$ and $\mathbf{\kappa}$ only, identification is unaffected. □

Intuitively, the scale of the covariance between instruments and forecasts errors does not affect identification; only the relative effect of instruments on the various forecasts errors does.

To achieve point-identification of the instrument-identified shocks, a simple Cholesky-type assumption is appropriate when identifying a certain class of VAR or DSGE models: I show below that this choice is correct for policy instruments under commonly made assumptions. I therefore use this particular choice as my baseline. The previous working paper version (Drautzburg (2016)) also provides an alternative statistical factorization for comparison.

The baseline factorization is a triangular decomposition in a population two-stage least squares (2SLS) representation of the previous problem, following Mertens and Ravn (2013). This 2SLS procedure uses the instruments \mathbf{z}_t to purge the forecast error variance to the first m_z variables in \mathbf{y}_t from shocks other than $\boldsymbol{\varepsilon}_t^{[1]}$, given $\mathbf{\Gamma}$, $\mathbf{\Sigma}$.

Mertens and Ravn (2013) called the resulting residual variance-covariance matrix $\mathbf{S}_1\mathbf{S}'_1$ and proposed either an upper or a lower triangular (Cholesky) decomposition of $\mathbf{S}_1\mathbf{S}'_1$. Generalizing the motivating example (2.4), \mathbf{v}_t can be rewritten as follows:

$$\begin{bmatrix} \mathbf{v}_{1,t} \\ \mathbf{v}_{2,t} \end{bmatrix} = \begin{bmatrix} \mathbf{0} & \boldsymbol{\eta} \\ \boldsymbol{\kappa} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{v}_{1,t} \\ \mathbf{v}_{2,t} \end{bmatrix} + \begin{bmatrix} \mathbf{S}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_{1,t} \\ \boldsymbol{\varepsilon}_{2,t} \end{bmatrix}, \quad (3.4)$$

where $\boldsymbol{\eta} = \mathbf{A}_{12}\mathbf{A}_{22}^{-1}$ and $\boldsymbol{\kappa} = \mathbf{A}_{21}\mathbf{A}_{11}^{-1}$ are functions of $\boldsymbol{\Sigma}$, $\boldsymbol{\Gamma}$, given in Appendix A.2. Using that $\mathbf{v}_t = \mathbf{A}\boldsymbol{\varepsilon}_t$ and simple substitution allows me to rewrite this system as

$$\begin{bmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\varepsilon}_{1,t} \\ \boldsymbol{\varepsilon}_{2,t} \end{bmatrix} = \begin{bmatrix} \mathbf{v}_{1,t} \\ \mathbf{v}_{2,t} \end{bmatrix} = \begin{bmatrix} (\mathbf{I} - \boldsymbol{\eta}\boldsymbol{\kappa})^{-1} \\ (\mathbf{I} - \boldsymbol{\kappa}\boldsymbol{\eta})^{-1}\boldsymbol{\kappa} \end{bmatrix} \mathbf{S}_1\boldsymbol{\varepsilon}_{1,t} + \begin{bmatrix} (\mathbf{I} - \boldsymbol{\eta}\boldsymbol{\kappa})^{-1}\boldsymbol{\eta} \\ (\mathbf{I} - \boldsymbol{\kappa}\boldsymbol{\eta})^{-1} \end{bmatrix} \mathbf{S}_2\boldsymbol{\varepsilon}_{2,t},$$

This shows how identifying \mathbf{S}_1 identifies $\mathbf{A}^{[1]}$ up to a triangular factorization as

$$\mathbf{A}^{[1]} = \begin{bmatrix} \mathbf{A}_{11} \\ \mathbf{A}_{21} \end{bmatrix} = \begin{bmatrix} (\mathbf{I} - \boldsymbol{\eta}\boldsymbol{\kappa})^{-1} \\ (\mathbf{I} - \boldsymbol{\kappa}\boldsymbol{\eta})^{-1}\boldsymbol{\kappa} \end{bmatrix} \text{chol}(\mathbf{S}_1\mathbf{S}'_1). \quad (3.5)$$

3.1.2 Identification of policy rules using instruments When identifying shocks to a certain class of policy rules, the lower Cholesky decomposition of $\mathbf{S}_1\mathbf{S}'_1$ identifies the true impact matrix $\mathbf{A}^{[1]}$. Generally, this is only true up to an orthonormal rotation. But even if one can show that two particular factorizations do not affect the results substantially, this does not generally mean that the results are robust to different identifying assumptions (cf. Watson (1994, footnote 42)). It is thus reassuring that, for what I call “observable Taylor-type rules” policy rules, a much sharper result holds since the narrative VAR correctly recovers \mathbf{A}^1 .

DEFINITION 1. An observable Taylor-type rule in economy (2.2) for variable $y_{p,t}$ is of the form:

$$y_{p,t} = \sum_{i=m_p+1}^m \eta_{p,i}y_{i,t} + \boldsymbol{\lambda}_p\mathbf{x}_{t-1} + \sigma_p\boldsymbol{\varepsilon}_{p,t},$$

where $\boldsymbol{\varepsilon}_{p,t} \subset \boldsymbol{\varepsilon}_t$ is *i.i.d.* and $y_{i,t} \subset \mathbf{y}_t$, $i = 1, \dots, n_p$.

While the policy rules are part of a structural model, whether they are observable depends on the VAR specification. For example, the canonical Taylor rule for monetary policy based on current inflation and an output measure is a useful clarifying example. It is an observable rule according to Definition 1 only if the output measure is included in the VAR-observables. An output gap constructed relative to a measure of output outside the VAR, such as a flexible price output, generally violates the assumption because monetary policy then also reflects other policy shocks contemporaneously. In the Blanchard and Perotti (2002) example (2.4), the policy rules are observable when $\gamma_{rg} = \gamma_{gr} = 0$. Policy rules with expectation measures would be observable, if the correct expectation measures are included. For example, if the monetary policy rule is correctly specified with

Greenbook expectation measures, as estimated in Coibion and Gorodnichenko (2012), then my approach accommodates this rule if Greenbook expectations are included in the VAR.

The identification problem simplifies if all shocks spanned by the instruments affect observable policy rules—or all but one shock affect observable policy rules. Formally, the following proposition shows that in a model with observable policy rules, \mathbf{S}_1 has a special structure that allows me to identify it uniquely using $\mathbf{\Gamma}$, $\mathbf{\Sigma}$, up to a normalization. Equivalently, when the VAR approximates the dynamics of the underlying DSGE model well, a notion Appendix A.4 makes precise, and Assumption 1 holds in the structural model, then the narrative VAR recovers the actual policy rules based on the procedure in (3.5). Trivially, when there is a single policy rule to be identified, as in Del Negro and Schorfheide (2009), the proposition also applies.

PROPOSITION 1. *Let $\mathbf{\Sigma} = \mathbf{A}\mathbf{A}'$ and order the policy variables such that the $m_p = m_z$ or $m_p = m_z - 1$ observable Taylor rules are ordered first and $\mathbf{\Gamma} = [\mathbf{G}, \mathbf{0}]\mathbf{A}$. Then $\mathbf{A}^{[1]}$ defined in (3.5) satisfies $\mathbf{A}^{[1]} = \mathbf{A}[\mathbf{I}_{m_z}, \mathbf{0}_{(m-m_z) \times (m-m_z)}]'$ up to a normalization of signs on the diagonal if*

(a) m_z instruments jointly identify shocks to $m_p = m_z$ observable Taylor rules w.r.t. the economy (2.2), or

(b) m_z instruments jointly identify shocks to $m_p = m_z - 1$ observable Taylor rules w.r.t. the economy (2.2) and $\eta_{p,m_z} = 0$, $p = 1, \dots, m_p$.

PROOF. See Appendix A.2. □

While the proof proceeds by Gauss–Jordan elimination, the intuition in case (a) can be understood using partitioned regression logic: $\mathbf{S}_1\mathbf{S}_1'$ is the residual variance of the first m_p forecast errors after accounting for the forecast error variance due to the last $m - m_p$ observed variables. Including the nonpolicy variables that enter the Taylor rule directly among the observables controls perfectly for the systematic part of the policy rules and leaves only the covariance matrix induced by policy shocks. With observable Taylor rules, this covariance matrix is diagonal. The VAR shocks can then be written as in (3.4) with a diagonal \mathbf{S}_1 matrix. If the last shock does not follow an observable Taylor rule, \mathbf{S}_1 is still lower-triangular. In either case, the Cholesky factorization in (3.5) works.

Formally, the Cholesky factorization of $\mathbf{S}_1\mathbf{S}_1'$ proposed by Mertens and Ravn (2013) imposes the $\frac{m_z(m_z-1)}{2}$ zero restrictions needed for exact identification. The structure imposed by having observable Taylor rules rationalizes these restrictions in a class of models. In fact, the mechanics of the proof would carry through if the block of policy rules had a Cholesky structure, confirming that what is needed for identification via instrumental variables in the model is precisely the existence of $\frac{m_z(m_z-1)}{2}$ restrictions.³ More

³In the notation of Appendix A.2, \mathbf{D}_0 need only be lower triangular to achieve identification, not diagonal as implied in Proposition 1. Indeed, there are extra restrictions coming from the assumption of observable Taylor rules rather than allowing for a block-diagonal structure in the Taylor rules. Rather than imposing these overidentifying restrictions, I back out the implied Taylor-type policy rules and analyze the distribution of implied coefficient loadings.

generally, identification requires restrictions on the contemporaneous interaction between policy instruments that need not have the form of observable Taylor rules.

3.1.3 Posterior uncertainty Now, I consider the case when the priors and posteriors over Γ and other parameters are nondegenerate.⁴ I allow for the following conjugate form of reduced-form priors: $\beta \sim \mathcal{N}(\bar{\beta}_0, \mathbf{N}_0)$ and $\mathbf{V}^{-1} \sim \mathcal{W}((\nu_0 \mathbf{S}_0)^{-1}, \nu_0)$. Unlike the common Normal-Wishart prior (e.g., Uhlig (1994)), the priors over β and \mathbf{V}^{-1} here are independent, for reasons I discuss below.

The proxy VAR in equation (3.1) is, formally, a seemingly unrelated regression (SUR). Inference is, therefore, that of a standard SUR model (e.g., Rossi, Allenby, and McCulloch (2005, Chapter 3.5)). In the special case in which the control variables for \mathbf{z}_t coincide with the variables used in the VAR, the SUR model collapses to a standard hierarchical Normal-Wishart posterior.

To derive the posterior, first stack the vectorized model (3.3):

$$\mathbf{y}_{\text{SUR}} = \mathbf{X}_{\text{SUR}} \beta_{\text{SUR}} + \mathbf{v}_{\text{SUR}}, \quad \mathbf{v}_{\text{SUR}} \sim \mathcal{N}(\mathbf{0}, \mathbf{V} \otimes \mathbf{I}_T), \quad (3.6)$$

where $\mathbf{y}_{\text{SUR}} = [\mathbf{y}'_T, \dots, \mathbf{y}'_1, \mathbf{z}'_T, \dots, \mathbf{z}'_1]'$ and \mathbf{v}_{SUR} is defined analogously. In addition, I use the following definitions:

$$\mathbf{V} = \begin{bmatrix} \mathbf{A}\mathbf{A}' & \Gamma' \\ \Gamma & \tilde{\Omega} \end{bmatrix}, \quad \beta_{\text{SUR}} = \begin{bmatrix} \text{vec}(\mathbf{B}) \\ \text{vec}(\boldsymbol{\mu}_z) \end{bmatrix},$$

$$\mathbf{X}_{\text{SUR}} = \begin{bmatrix} \mathbf{I}_{m_y} \otimes \mathbf{X}_y & \mathbf{0}_{T(m_y p + 1) \times T m_z} \\ \mathbf{0}_{T m_z \times T m_y} & \mathbf{I}_{m_z} \otimes \mathbf{X}_z \end{bmatrix}, \quad \mathbf{X}_y = [\mathbf{Y}_{-1} \quad \dots \quad \mathbf{Y}_{-p} \quad \mathbf{1}_T], \quad \mathbf{X}_z = [\mathbf{1}_T].$$

Using these definitions to transform the model makes the errors independently normally distributed. The transformation takes advantage of the block-diagonal structure of the covariance matrix: $\tilde{\mathbf{v}} = \tilde{\mathbf{Y}} - \tilde{\mathbf{X}}\beta \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Standard conditional Normal-Wishart posterior distributions arise from the transformed model. For the transformation, it is convenient to define \mathbf{U} as the Cholesky decomposition of \mathbf{V} such that $\mathbf{U}'\mathbf{U} = \mathbf{V}$:

$$\tilde{\mathbf{X}} = ((\mathbf{U}^{-1})' \otimes \mathbf{I}_T) \begin{bmatrix} \mathbf{I}_{m_y} \otimes \mathbf{X}_y & \mathbf{0}_{T(m_y p + 1) \times T m_z} \\ \mathbf{0}_{T m_z \times T m_y} & \mathbf{I}_{m_z} \otimes \mathbf{X}_z \end{bmatrix}, \quad \tilde{\mathbf{Y}} = ((\mathbf{U}^{-1})' \otimes \mathbf{I}_T) \mathbf{y}_{\text{SUR}},$$

$$\mathbf{N}_{XX}(\mathbf{V}) = \tilde{\mathbf{X}}' \tilde{\mathbf{X}}, \quad \mathbf{N}_{XY}(\mathbf{V}) = \tilde{\mathbf{X}}' \tilde{\mathbf{Y}},$$

$$\mathbf{S}_T(\beta) = \frac{1}{\nu_0 + T} \begin{bmatrix} (\mathbf{Y} - \mathbf{X}\mathbf{B})' \\ (\mathbf{Z} - \mathbf{1}_T \boldsymbol{\mu}'_z)' \end{bmatrix} \left[(\mathbf{Y} - \mathbf{X}\mathbf{B}) \quad (\mathbf{Z} - \mathbf{1}_T \boldsymbol{\mu}'_z) \right] + \frac{\nu_0}{\nu_0 + T} \mathbf{S}_0.$$

With these definitions, the following lemma holds (Rossi, Allenby, and McCulloch (2005, Chapter 3.5), or Appendix A.3).

LEMMA 2. *The conditional likelihoods are, respectively, conditionally conjugate with Normal and Wishart priors. Given independent priors $\beta \sim \mathcal{N}(\bar{\beta}_0, \mathbf{N}_0)$ and $\mathbf{V}^{-1} \sim$*

⁴I abstract from potentially weak instruments, an issue surveyed by Lopes and Polson (2014). I show that in my application, the instruments are highly correlated with the identified shocks.

Algorithm 1 SUR-VAR.

1. Initialize $\mathbf{V}^{(0)} = \mathbf{S}_T(\bar{\boldsymbol{\beta}}_T)$.
2. Repeat for $i = 1, \dots, n_G$:
 - (a) Draw $\boldsymbol{\beta}^{(i)} | \mathbf{V}^{(i-1)}$ from (3.7a).
 - (b) Draw $\mathbf{V}^{(i)} | \boldsymbol{\beta}^{(i)}$ from (3.7b).

$\mathcal{W}((\nu_0 \mathbf{S}_0)^{-1}, \nu_0)$ and defining $\bar{\boldsymbol{\beta}}_T(\mathbf{V}) = (\mathbf{N}_{XX}(\mathbf{V}) + \mathbf{N}_0)^{-1}(\mathbf{N}_{XY}(\mathbf{V}) + \mathbf{N}_0 \bar{\boldsymbol{\beta}}_0)$, the conditional posterior distributions are given by

$$\boldsymbol{\beta} | \mathbf{V}, \mathbf{y}^T \sim \mathcal{N}(\bar{\boldsymbol{\beta}}_T(\mathbf{V}), (\mathbf{N}_{XX}(\mathbf{V}) + \mathbf{N}_0)^{-1}), \quad (3.7a)$$

$$\mathbf{V}^{-1} | \boldsymbol{\beta}, \mathbf{y}^T \sim \mathcal{W}(\mathbf{S}_T(\boldsymbol{\beta})^{-1} / (\nu_0 + T), \nu_0 + T). \quad (3.7b)$$

In general, no closed-form posterior is available. The exception occurs when SUR collapses to ordinary least squares (OLS): If $\mathbf{X}_z = \mathbf{X}_y$, then $\tilde{\mathbf{X}} = \mathbf{I}_{m_c+m_z} \otimes \mathbf{X}_y$ and $\mathbf{N}_{XX} = (\mathbf{V}^{-1} \otimes \mathbf{X}_y' \mathbf{X}_y)$ and analogously for \mathbf{N}_{XY} . In this special case, closed forms are available for the marginal distribution of \mathbf{V} , allowing me to draw directly from the posterior. Generally, however, the block closed-form structure gives rise to a natural Gibbs sampler as in Algorithm 1.

3.1.4 Instruments with missing values While Corollary 1 shows that missing instruments do not affect shock identification, a point-mass of instruments at their sample mean will typically affect the model fit. Missing data on instruments is common in macro time series: High frequency identification of monetary policy shocks using fed funds futures data (e.g., Kuttner (2001)) is restricted to the introduction of this financial instrument in the early 1990s. Mertens and Ravn (2013) include only 13 nonmissing instruments for the U.S. personal income tax rate since 1947. The Bayesian approach provides a natural way to avoid dealing with mass points. It only requires joint normality of the instruments and structural shocks when the shock is actually observed. For the other periods, I impute the missing data under the same distributional assumption.

Let $\iota_{V,t}$ denote the index of the rows of the covariance matrix \mathbf{V} corresponding to the missing instruments in period t . Let $\iota_{V,t}^c$ denote the complementary index. Finally, let $\iota_{z,t}$ index the rows of missing instruments in the m_z row vector \mathbf{z}_t with complement $\iota_{z,t}^c$. Joint normality then implies that the missing instruments have the following distribution:

$$\hat{\mathbf{z}}_{\iota_{z,t},t} \left| \begin{bmatrix} \mathbf{y}_t - \mathbf{B}\mathbf{x}_{t-1} - \boldsymbol{\mu}_y \\ \mathbf{z}_{\iota_{z,t}^c,t} \end{bmatrix}, \mathbf{V} \right. \\ \left. \stackrel{\text{i.i.d.}}{\sim} \mathcal{N} \left(\boldsymbol{\mu}_{\iota_{z,t},z} + \begin{bmatrix} \mathbf{y}_t - \mathbf{B}\mathbf{x}_{t-1} - \boldsymbol{\mu}_y \\ \mathbf{z}_{\iota_{z,t}^c,t} - \boldsymbol{\mu}_{\iota_{z,t}^c,z} \end{bmatrix} \boldsymbol{\beta}(\iota_{z,t}), \mathbf{V}_{\iota_{V,t},\iota_{V,t}} - \boldsymbol{\beta}(\iota_{z,t}) \mathbf{V}_{\iota_{V,t}^c,\iota_{V,t}^c} \boldsymbol{\beta}(\iota_{z,t})' \right), \quad (3.8)$$

where $\boldsymbol{\beta}(\iota_{z,t}) = (\mathbf{V}_{\iota_{V,t}^c,\iota_{V,t}^c})^{-1} \mathbf{V}_{\iota_{V,t}^c,\iota_{V,t}}$ is the population OLS regression coefficient of the missing data on the observed data. The Gibbs sampler easily accommodates this imputation, leading to Algorithm 2.

Algorithm 2 SUR-VAR with missing values.

-
1. Initialize $\mathbf{V}^{(0)} = \mathbf{S}_T(\bar{\boldsymbol{\beta}}_T)$ and $\mathbf{z}_{\iota_z, t, t}^{(0)} = \mathbf{0}_{\dim(\iota_z, t) \times 1} \forall t$.
 2. Repeat for $i = 1, \dots, n_G$:
 - (a) Draw $\boldsymbol{\beta}^{(i)} | \mathbf{V}^{(i-1)}, \{\hat{\mathbf{z}}_{\iota_z, t, t}^{(i-1)}\}_{t=1}^T$ from (3.7a).
 - (b) Draw $\mathbf{V}^{(i)} | \boldsymbol{\beta}^{(i)}, \{\hat{\mathbf{z}}_{\iota_z, t, t}^{(i-1)}\}_{t=1}^T$ from (3.7b) replacing $\mathbf{z}_{\iota_z, t, t} = \hat{\mathbf{z}}_{\iota_z, t, t}^{(i-1)} \forall t$.
 - (c) Draw $\{\hat{\mathbf{z}}_{\iota_z, t, t}^{(i)}\}_{t=1}^T | \boldsymbol{\beta}^{(i)}, \mathbf{V}^{(i)}$ from (3.8).
-

3.2 Narrative DSGE-SVAR

Building on the results on identification and estimation in the narrative SVAR, I show how to use the DSGE model to elicit a prior using dummy variables and how to measure model misspecification, building on [Del Negro and Schorfheide \(2004\)](#). Throughout, I assume that the VAR is sufficiently informative so that $\boldsymbol{\varepsilon}_t^* = \boldsymbol{\varepsilon}_t$ in a VAR(∞). The VAR is sufficiently informative under the ABCD condition in [Fernandez-Villaverde, Rubio-Ramirez, Sargent, and Watson \(2007\)](#). My estimation only considers parameters that guarantee this condition. I also verify numerically that this condition holds approximately with a finite lag order.

3.2.1 Prior elicitation A natural way to elicit a prior over the parameters of the VAR, tracing back to [Theil and Goldberger \(1961\)](#), is through dummy observations. [Del Negro and Schorfheide \(2004\)](#) used this approach to elicit a prior for the VAR based on a prior over structural parameters of a DSGE model. Here, I adapt their approach for use within the SUR framework. The likelihood function of the SUR model is not conjugate with a closed-form joint prior over $(\boldsymbol{\beta}, \mathbf{V}^{-1})$, unless the right-hand side variables in the instrument equations are the same as those in the VAR equations itself. Consequently, the DSGE model implied prior also fails to generate unconditionally conjugate posteriors. I thus consider independent priors for the dynamics and the covariance matrix. In this approach, the prior is available in closed form, but the conditional prior variance of $\boldsymbol{\beta}$ is necessarily independent of \mathbf{V}^{-1} , unlike in the standard DSGE-VAR model. Even in the more general SUR case, however, using dummy variables still generates conditional posteriors in closed form and conditional priors with the same intuitive interpretation as in [Del Negro and Schorfheide \(2004\)](#).

To implement the prior, I follow [Del Negro and Schorfheide \(2004\)](#): I generate the prior for $\mathbf{B}, \mathbf{V}^{-1}$ by integrating out the disturbances to avoid unnecessary sampling error. The prior for the VAR coefficients is centered at the coefficients of a population regression of the VAR parameters in data simulated from the DSGE model. The prior for the variance covariance matrix is centered on the population covariance of the DSGE model forecast errors for the VAR variables and builds in the identifying assumptions on the narrative instruments. $\bar{\mathbf{V}}_0(\boldsymbol{\theta})$ satisfies Assumption 1 with $\mathbf{G} = \text{diag}([c_1, \dots, c_{m_z}])$. I assume measurement error that is independent across instruments. The variance of the measurement error is parametrized to be ω_z^2 times the univariate shock variance.

Together, c_i and ω_i map into a signal-to-noise ratio for each proxy variable. Formally, the prior is centered at

$$\bar{\mathbf{B}}_0^y(\boldsymbol{\theta}) = \mathbb{E}^{\text{DSGE}}[\mathbf{X}_0 \mathbf{X}_0' | \boldsymbol{\theta}]^{-1} \mathbb{E}^{\text{DSGE}}[\mathbf{X}_0 \mathbf{Y}_0' | \boldsymbol{\theta}] \Leftrightarrow \bar{\boldsymbol{\beta}}_0^y(\boldsymbol{\theta}) = \text{vec}(\bar{\mathbf{B}}_0^y(\boldsymbol{\theta})), \quad (3.9a)$$

$$\bar{\mathbf{V}}_0(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{A}(\boldsymbol{\theta})^* (\mathbf{A}(\boldsymbol{\theta})^*)' & (\mathbf{A}(\boldsymbol{\theta})^*) [\text{diag}([c_1, \dots, c_{m_z}], \mathbf{0})]' \\ [\text{diag}([c_1, \dots, c_{m_z}], \mathbf{0})] (\mathbf{A}(\boldsymbol{\theta})^*)' & \text{diag}([\omega_i^2 + c_i^2] [\mathbf{A}_1(\boldsymbol{\theta})^* (\mathbf{A}_1(\boldsymbol{\theta})^*)']_{ii})_{i=1}^{m_z} \end{bmatrix}, \quad (3.9b)$$

where $\mathbb{E}^{\text{DSGE}}[\cdot | \boldsymbol{\theta}]$ denotes the unconditional expectation based on the linear DSGE model (2.2) when the coefficient matrices are generated by the structural parameters $\boldsymbol{\theta}$. Here, $\mathbf{A}_1(\boldsymbol{\theta})^* \equiv [\mathbf{I}_{m_z}, \mathbf{0}] \mathbf{A}(\boldsymbol{\theta})^*$. The specific assumptions on $\bar{\mathbf{V}}_0(\boldsymbol{\theta})$ can be relaxed. However, note that the specific form of \mathbf{G} does not matter for shock identification, as long as it is invertible. This is similar to the logic of Corollary 1. But the assumptions on \mathbf{G} may affect the model and I report a robustness check with an alternative parametrization.

The prior incorporates the Normal likelihood over the dummy observations and Jeffrey's prior over \mathbf{V}^{-1} along with scale factors chosen to make the prior information equivalent to T_0^B observations about \mathbf{B} and T_0^V observations on \mathbf{V}^{-1} . Per Lemma 2, the corresponding marginal priors are

$$\boldsymbol{\beta} | \mathbf{V}^{-1}, \boldsymbol{\theta} \sim \mathcal{N}(\bar{\boldsymbol{\beta}}_0(\boldsymbol{\theta}), (T_0^B \times \tilde{\mathbf{X}}_0' (\mathbf{V}^{-1} \otimes \mathbf{I}) \tilde{\mathbf{X}}_0)^{-1}),$$

$$\tilde{\mathbf{X}}_0 = \text{diag}([\mathbf{X}_0^y, \dots, \mathbf{X}_0^y, \mathbf{X}_0^z, \dots, \mathbf{X}_0^z]),$$

$$\mathbf{V}^{-1} | \boldsymbol{\beta}, \boldsymbol{\theta} \sim \mathcal{W}_{m+m_z}(\text{SSR}_0(\boldsymbol{\beta}, \boldsymbol{\theta})^{-1}, T_0^V),$$

$$\text{SSR}_0(\boldsymbol{\beta}, \boldsymbol{\theta}) = T_0^V \times \bar{\mathbf{V}}_0(\boldsymbol{\theta}) + T_0^B ([\mathbf{Y}_0(\boldsymbol{\theta}), \mathbf{Z}_0(\boldsymbol{\theta})] - \mathbf{X}_0 \mathbf{B}(\boldsymbol{\beta})) ([\mathbf{Y}_0(\boldsymbol{\theta}), \mathbf{Z}_0(\boldsymbol{\theta})] - \mathbf{X}_0 \mathbf{B}(\boldsymbol{\beta}))'.$$

Appendix A.3 provides details on the prior densities and the corresponding dummy variables.

Estimating the DSGE-VAR requires estimating $\boldsymbol{\theta}$, and thus an extra step. For given $\boldsymbol{\theta}$, Lemma 2 characterizes the posterior for $\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}$, as before. To estimate $\boldsymbol{\theta}$, I add a Metropolis-within-Gibbs step to the previous Gibbs sampler 2 to simulate from $\boldsymbol{\theta} | \boldsymbol{\beta}, \mathbf{V}^{-1}$; see Algorithm 3, which also accommodates missing values for the instruments. The resulting distribution of the structural parameters of the DSGE model—a byproduct of the DSGE-VAR estimation—may be of interest in itself.

When the distribution of both \mathbf{B} and \mathbf{V}^{-1} is nondegenerate, the conditional posterior for $\boldsymbol{\theta}$ simplifies, as in Geweke (2005, p. 77):

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{B}, \mathbf{V}^{-1}, \mathbf{Y}, \mathbf{Z}) &= \frac{f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}, \mathbf{V}^{-1}) \pi(\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}, \mathbf{V}^{-1}) \pi(\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} = \frac{p(\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int \pi(\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\ &\propto \pi(\mathbf{B}, \mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \end{aligned} \quad (3.10)$$

However, in the limit of $T_0^B \rightarrow \infty$, the prior for $\text{vec}(\mathbf{B})$ is a point-mass at $\bar{\boldsymbol{\beta}}_0(\boldsymbol{\theta})$, and $\mathbf{V} = \bar{\mathbf{V}}_0(\boldsymbol{\theta})$ when $T_0^V \rightarrow \infty$. $f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}, \mathbf{V}^{-1})$ then becomes a function of $\boldsymbol{\theta}$ and the posterior

Algorithm 3 DSGE-VAR with missing data.

-
- (1) Initialize VAR parameters: Set $\mathbf{B}_{(0)}, \mathbf{V}_{(0)}^{-1}$ to OLS estimates.
 - (2) Initialize structural parameters: $\boldsymbol{\theta}_0 = \int_{\boldsymbol{\theta}} \boldsymbol{\theta} \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}$.
 - (3) Initialize missing instruments: $\mathbf{Z}_{\iota_{Z,t},t}^{(0)} = \mathbf{0}_{\dim(\iota_{Z,t}) \times 1} \forall t$.
 - (4) Metropolis–Hastings within Gibbs:
 - (a) Draw a candidate $\boldsymbol{\theta}_c$ from $\boldsymbol{\theta}_c \sim F_{\boldsymbol{\theta}}(\cdot | \boldsymbol{\theta}_{(d-1)})$. Assign unstable draws or draws violating the ABCD condition (Fernandez-Villaverde et al. (2007)) zero density.
 - (b) With probability $\alpha_{d-1,i}(\boldsymbol{\theta}_c)$, set $\boldsymbol{\theta}_{(d)} = \boldsymbol{\theta}_c$, otherwise, set $\boldsymbol{\theta}_{(d)} = \boldsymbol{\theta}_{(d-1)}$.

$$\alpha_{d-1,i}(\boldsymbol{\theta}_c) = \min \left\{ 1, \frac{\pi(\mathbf{B}_{(d-1)}, \mathbf{V}_{(d-1)}^{-1} | \boldsymbol{\theta}_c) \pi(\boldsymbol{\theta}_c)}{\pi(\mathbf{B}_{(d-1)}, \mathbf{V}_{(d-1)}^{-1} | \boldsymbol{\theta}_{(d-1)}) \pi(\boldsymbol{\theta}_{(d-1)})} \frac{f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_{(d-1)} | \boldsymbol{\theta}_c)}{f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_c | \boldsymbol{\theta}_{(d-1)})} \right\}. \quad (3.11)$$

- (c) Draw $\mathbf{B}_{(d)} | \boldsymbol{\theta}_{(d)}, \mathbf{V}_{(d-1)}^{-1}, \{\hat{\mathbf{Z}}_{\iota_{Z,t},t}^{(d-1)}\}_{t=1}^T$ according to (3.7a) and including dummy observations.
 - (d) Draw $\mathbf{V}_{(d)}^{-1} | \mathbf{B}_{(d)}, \boldsymbol{\theta}_{(d)}, \{\hat{\mathbf{Z}}_{\iota_{Z,t},t}^{(d-1)}\}_{t=1}^T$ according to (3.7b) and including dummy observations.
 - (e) Draw $\{\hat{\mathbf{Z}}_{\iota_{Z,t},t}^{(d)}\}_{t=1}^T | \mathbf{B}_{(d)}, \boldsymbol{\theta}_{(d)}, \mathbf{V}_{(d)}^{-1}$ according to (3.8).
 - (f) If $d < D$, increase d by one and go back to (a), or else exit.
-

for $\boldsymbol{\theta}$ becomes

$$\pi(\boldsymbol{\theta} | \mathbf{B}, \mathbf{V}^{-1}, \mathbf{Y}, \mathbf{Z}) = \frac{f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}(\boldsymbol{\theta})^{-1}) \pi(\boldsymbol{\theta})}{\int f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}(\boldsymbol{\theta})^{-1}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \propto f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}(\boldsymbol{\theta})^{-1}) \pi(\boldsymbol{\theta}), \quad (3.12)$$

that is, the posterior is proportional to the VAR approximation to the DSGE likelihood function for the data times the prior. If only $T_0^B \rightarrow \infty$, then an intermediate case arises:

$$\begin{aligned} \pi(\boldsymbol{\theta} | \mathbf{B}, \mathbf{V}^{-1}, \mathbf{Y}, \mathbf{Z}) &= \frac{f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}^{-1}) \pi(\mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta})}{\int f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}^{-1}) \pi(\mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}) d\boldsymbol{\theta}} \\ &\propto f(\mathbf{Y}, \mathbf{Z} | \mathbf{B}(\boldsymbol{\theta}), \mathbf{V}^{-1}) \pi(\mathbf{V}^{-1} | \boldsymbol{\theta}) \pi(\boldsymbol{\theta}). \end{aligned} \quad (3.13)$$

In these limiting cases, the kernel from (3.12) or (3.13) take the place of the kernel from (3.10) in Algorithm 3.

3.2.2 Marginal likelihood The joint distribution of the data $[\mathbf{Y}, \mathbf{Z}]$, the missing instruments $\hat{\mathbf{Z}} \equiv \{\hat{\mathbf{z}}_{\iota_{Z,t}}\}_t$, the VAR parameters $\boldsymbol{\beta}, \mathbf{V}^{-1}$ and the DSGE model parameters $\boldsymbol{\theta}$ is given by

$$p(\mathbf{Y}, \mathbf{Z}, \hat{\mathbf{Z}}, \boldsymbol{\beta}, \mathbf{V}^{-1}, \boldsymbol{\theta}) = p(\mathbf{Y}, \mathbf{Z}, \hat{\mathbf{Z}} | \boldsymbol{\beta}, \mathbf{V}^{-1}) p(\boldsymbol{\beta}, \mathbf{V}^{-1} | \boldsymbol{\theta}) p(\boldsymbol{\theta}),$$

where equation (A.12) in Appendix A.5 spells out the individual components.

Integrating out the missing observations, the VAR parameters, and the DSGE parameters gives the marginal likelihood:

$$p(\mathbf{Y}, \mathbf{Z}) = \int \int \int \int p(\mathbf{Y}, \mathbf{Z}, \hat{\mathbf{Z}}, \boldsymbol{\beta}, \mathbf{V}^{-1}, \boldsymbol{\theta}) d\boldsymbol{\beta} d\mathbf{V}^{-1} d\hat{\mathbf{Z}} d\boldsymbol{\theta}. \quad (3.14)$$

Because the prior is a function of T_0^V , T_0^B , the marginal likelihood is, implicitly, indexed by these prior hyperparameters. I now discuss how to interpret the marginal likelihood as a function of the prior DSGE model weights. Next, I summarize the computation of the marginal likelihood.

Interpretation Del Negro and Schorfheide (2004) introduced the prior DSGE model weight with the highest marginal likelihood as a diagnostic of misspecification: If the DSGE prior generates observations with properties like the data, these are informative and improve the model fit by reducing the prior probability of the ill-fitting parameters. Del Negro et al. (2007) showed in an AR(1) case with known variance that when DSGE model and sample moments differ, there can be an interior optimum for the prior weight. This trades off shrinkage with bias. In other cases, they find that the best fitting prior weight can diverge, so that a high weight on the pure DSGE model or the (almost) flat prior VAR can emerge as optimal. Thus, the analysis of the prior weight that maximizes the marginal likelihood is meaningful and has a clear interpretation.

My model differs from Del Negro et al. (2007) in two dimensions: First, because of the extra information via instruments, my prior is not conjugate. Second, I allow for the weight on the covariance matrix and the dynamics to differ. I therefore characterize the behavior of the marginal likelihood in terms of the prior weight in the case of my model in Appendix A.6. For the case of known model dynamics, I characterize the marginal likelihood analytically and prove a lemma characterizing the slope of the marginal likelihood in the scalar case. The analytical results imply that the marginal likelihood is strictly increasing in T_0^V when the prior variance fits well enough—and vice versa, when the DSGE model fits the sample variance (sufficiently) poorly. The Appendix verifies the same in a numerical example for the empirically relevant matrix case. Together, the analytical and numerical results show that when the DSGE model prior fits the sample moments well, the marginal likelihood is strictly increasing in the number of dummy observations.

Computation I combine the methods of Chib (1995) and Geweke (1999) to compute the likelihood: Specifically, I compute the marginal likelihood conditional on a specific $\boldsymbol{\theta}_d$ —the three inner integrals in (3.14)—using the method of Chib (1995) for models with fully conditional posteriors. Combined with the prior, this conditional marginal likelihood gives the kernel of the $\boldsymbol{\theta}$ posterior that I use in the Geweke (1999) algorithm. While relatively few draws—2000 draws after 1000 burn-in draws—are accurate to ± 0.1 log points given $\boldsymbol{\theta}_d$, the repeated approximation takes time. Since the posterior draws $\{\boldsymbol{\theta}_d\}_d$ are autocorrelated, I subsample every j th draw to yield a more efficient sample of 1000 posterior draws that economizes on computing time. See Appendix A.7 for details.

4. EMPIRICAL SPECIFICATION

4.1 *Data and sample period*

The estimation uses seven macroeconomic indicators: Government spending, the average labor tax rate, and the effective Federal Funds Rate (FFR) are the fiscal and monetary policy instruments. The other variables are real GDP, real investment (including consumer durables), the debt-to-GDP ratio and GDP inflation. GDP and its components are in per capita terms. All fiscal variables aggregate the federal government with state and local governments.

To include periods of significant variation in fiscal policy, the sample starts in 1947:Q1. This period includes Korean War expenditures as well as episodes of declining and rising debt-to-GDP ratios. [Bohn \(1991\)](#) argued that long samples are important to capture slow moving debt dynamics. I stop the estimation in 2007:Q4, before the zero lower bound became binding. Following [Ramey \(2011\)](#), I allow for a quadratic trend. [Francis and Ramey \(2009\)](#) showed that a quadratic trend can capture unmodeled factors such as demographics.⁵

I use newly digitized Greenbook data as a proxy for defense spending and to compute an updated series on monetary policy shocks. Specifically, I follow [Ramey \(2011\)](#) and use one-quarter ahead defense spending forecast errors as a proxy for exogenous government purchases. [Ramey \(2011\)](#) used the forecast errors from the Survey of Professional Forecasters (SPF), but the SPF only covers defense until 1982. In contrast, Greenbooks reported defense spending forecasts continuously since 1969.⁶

In an extension, I also use the newly digitized Greenbook data to control for fiscal foresight. Specifically, I use Greenbook forecasts on federal government purchases and revenue four quarters out as observables in the VAR to control for the possibility of news shocks. One can show that including such forecasts makes an otherwise nonfundamental VAR fundamental in a version of the work-horse New Keynesian model with news shocks. Empirically, the announcements and implementation dates in [Yang \(2007\)](#) suggest that U.S. tax reforms are typically implemented within a year. Thus, expectations of policy four quarters ahead should capture fiscal foresight.

The other shock proxies are from the literature and cover taxes and monetary policy. The tax instruments are constructed by [Mertens and Ravn \(2013\)](#) building on [Romer and Romer \(2010\)](#). [Romer and Romer \(2010\)](#) proceeded in two steps. They first identify legislated tax changes. Each of these tax changes is associated with a predicted change in tax liabilities, which is used to quantify the size of the tax change. In a second step, [Romer and Romer \(2010\)](#) consulted documents issued by the legislative and the executive branches to classify the tax changes as endogenous or as exogenous. Exogenous changes are not motivated by the current state of the business cycle, but rather by long-run growth objectives or inherited deficit concerns. [Mertens and Ravn \(2013\)](#) additionally consulted legislative and executive documents to parse out tax rate changes antici-

⁵I detrend prior to estimation to match the detrended data with my stationary DSGE model.

⁶Restricted to the same sample period at the SPF forecast errors, the posterior uncertainty is larger with the Greenbook forecasts. This is intuitive because forecast averaging typically improves forecasts. In the full sample, the extra data availability serves to reduce the uncertainty.

pated by more than 90 days and to divide the tax changes into those affecting personal income and corporate income.⁷

The measure of monetary policy shocks is constructed using a mix of the approach for the two fiscal shock proxies. Romer and Romer (2004) first extracted the desired change in the Federal Funds Rate target from documents, and then projected it on the previous interest rate level, and staff nowcasts and forecasts of the unemployment rate, output growth and the inflation rate. The residual is the monetary policy shock. To the extent that monetary policy only considers current and future inflation and economic activity, the residual is the monetary policy shock. I update the Romer and Romer (2004) series of monetary policy shocks beyond 1996. Appendix B.1 describes construction of the macroeconomic data and the proxy variables.

4.2 DSGE model specification

In this section, I outline the empirical specification of the generic DSGE model in (2.2). The model is based on the standard medium-scale New Keynesian model as exemplified by Christiano, Eichenbaum, and Evans (2005) and follows closely Smets and Wouters (2007). There is monopolistic competition in intermediate goods markets and the labor market with Calvo frictions in price and wage adjustment, partial price and wage indexation and real frictions such as investment adjustment cost and habit formation. I add labor, capital and consumption taxes as in Drautzburg and Uhlig (2015) and fiscal rules as in Leeper, Plante, and Traum (2010) and Fernandez-Villaverde et al. (2015). Here, I only discuss the specification of fiscal and monetary policy. The remaining model equations are detailed in Appendix A.8.

The monetary authority sets interest rates according to the following standard Taylor rule:

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r)(\eta_{r,\pi} \hat{\pi}_t + \eta_{r,y} \hat{y}_t) + \xi_t^r, \quad (4.1)$$

where ρ_r controls the degree of interest rate smoothing and $\eta_{r,x}$ denotes the reaction of the interest rate to deviations of variable x from its trend. As a robustness-check, I have also estimated the model with a Taylor rule that includes the the level and the change of the output gap (i.e., the deviation of output from output in a frictionless world). ξ_t^r follows an AR(1) process.

The fiscal rules allow for both stabilization of output and the debt burden as well as smoothing of the different fiscal instruments:⁸

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + (1 - \rho_g) \left(-\eta_{g,y} \hat{y}_t - \eta_{g,b} \frac{\bar{b}}{\gamma \bar{y}} \hat{b}_t \right) + \xi_t^g, \quad (4.2a)$$

⁷I do not model shocks to corporate income taxes. Since personal and corporate income shock proxies are correlated, this may bias my results. However, dropping the tax shock altogether in a variant of my model with news shocks produces results similar to my baseline estimates.

⁸Leeper, Plante, and Traum (2010) used no lag in debt and GDP, while Fernandez-Villaverde et al. (2015) used a one-quarter lag.

$$\hat{s}_t = \rho_s \hat{s}_{t-1} + (1 - \rho_s) \left(-\eta_{s,y} \hat{y}_t - \eta_{s,b} \frac{\bar{b}}{\gamma \bar{y}} \hat{b}_t \right) + \xi_t^s, \quad (4.2b)$$

$$\frac{\bar{w}\bar{n}}{\bar{y}} d\tau_t^n = \rho_\tau \frac{\bar{w}\bar{n}}{\bar{y}} d\tau_{t-1}^n + (1 - \rho_\tau) \left(\eta_{\tau^n, y} \hat{y}_t + \eta_{\tau^n, b} \frac{\bar{b}}{\gamma \bar{y}} \hat{b}_t \right) + \xi_t^{\tau, n}. \quad (4.2c)$$

The disturbances ξ_t^o follow exogenous AR(1) processes: $\xi_t^o = \rho_o \xi_{t-1}^o + \epsilon_t^o$.⁹ The sign of the coefficients in the expenditure components g_t and s_t is flipped so that positive estimates always imply consolidation in good times ($\eta_{o,y} > 0$) or when debt is high ($\eta_{o,b} > 0$).

Following [Christiano, Trabandt, and Walentin \(2011\)](#), I include a cost channel of monetary policy: Firms have to borrow at the nominal interest rate to pay the wage bill at the beginning of the period. This allows monetary policy to cause inflation in the short run.

The debate about variable selection in singular DSGE models is still ongoing; see [Guerron-Quintana \(2010\)](#), [Canova, Ferroni, and Matthes \(2013\)](#) and the comment by [Iskrev \(2014\)](#). When fitting the model to the data, I consider as many structural shocks as observables in the observation equation of the DSGE model (2.2a) and the VAR (2.1a). Including the policy variables naturally suggests to include the corresponding policy shocks. Additionally, I include a total factor productivity (TFP) and investment specific technology shock to explain GDP and investment, a price markup shock to contribute to inflation and a shock to lump-sum transfers to add variation to the debt-to-GDP ratio.

In the robustness check that allows for fiscal foresight, I model fiscal policy as the sum of a contemporaneous and an anticipated component. The contemporaneous component is given by (4.2) as before. Analogous rules determine the anticipated component two quarters out in terms of current observables. In addition, I allow for a productivity news shock, also revealed two quarters in advance.

Backing out the DSGE model implied historical shocks generally requires Kalman smoothing. Here, I exploit that under the invertibility Assumption A.2, Lemma 2 implies that the data eventually fully reveal the hidden state variables of the DSGE model. The contemporaneous state uncertainty matrix is then zero. I initialize the Kalman filter with this matrix so that the Kalman smoother coincides with the Kalman filter. Given parameters, I use Dynare ([Adjemian, Bastani, Juillard, Mihoubi, Perendia, Ratto, and Villemot \(2011\)](#)) to solve the model.

To limit the dimensionality of the estimation problem, I calibrate a number of structural parameters and focus on the estimation of policy rules and shock processes (Table 1). These parameters largely correspond to the prior mean in [Smets and Wouters \(2007\)](#). Average tax rates are calibrated as in [Drautzburg and Uhlig \(2015\)](#).

Priors for policy rules and shock processes are standard: I follow [Smets and Wouters \(2007\)](#) for common parameters and choose similar priors for new parameters of the fiscal rules. To parametrize the observation equations for the instruments, I assume that

⁹Not only the fiscal policy shocks but all shocks in my specification follow univariate AR(1) processes, unlike [Smets and Wouters \(2007\)](#) who allowed some shocks to follow ARMA(1, 1) processes. Ruling out MA(1) components helps to guarantee that a VAR can approximate the DSGE model dynamics as discussed by [Fernandez-Villaverde et al. \(2007\)](#).

TABLE 1. Calibrated parameters.

Parameter	Value
Elasticity of substitution (inverse) σ	1.500
Discount rate (quarterly)	0.5%
Capital share α	0.300
Depreciation rate δ	0.020
Net TFP growth (quarterly)	0.4%
Steady state gross wage markup	1.150
Kimball parameters	10.000
Steady state government spending	0.200
Steady state consumption tax rate	0.073
Steady state capital tax rate	0.293
Steady state labor tax rate	0.165

both the matrix of loadings \mathbf{G} and the covariance matrix of measurement errors are diagonal, as in (3.9b). My prior for the loadings c_i is that they are centered around unity, given the appropriate scaling of the narrative variables. Both the loadings and the relative standard errors of the measurement error have inverse gamma priors, so that the instruments are relevant for all parameter draws. The prior for the relative standard deviation is a relatively tight prior with a mean of 1.0 and a standard deviation of 0.5. This prior is intended to make the instruments informative to allow them to influence other parameter estimates: Overall the prior mean implies a signal-to-noise ratio of two in terms of standard deviations. Table C.1 in Appendix C lists all estimated parameters alongside their prior distributions.

4.3 DSGE-SVAR model specification

I use a VAR with $p = 4$ lags throughout this paper. The baseline model has seven variables and all three narrative instruments listed earlier. In the model with news shocks, I add expectations of government spending and overall output four quarters ahead as observables, along with shocks to future spending and productivity. When estimating the news shock model, I give up on identifying tax shocks because the short sample period leaves only a small number of nonzero proxies. The specification of the lags follows Ramey (2011), but I also find that the finite lag approximation to the underlying DSGE model does well empirically: Figure B.2 for the baseline model and Figure B.3 in the Appendix show that the VAR dynamics match that of the underlying estimated DSGE model well. In the reported results, I only demean the proxy variables by setting \mathbf{X}_t^z to a constant. In unreported results, I have checked the robustness of the flat prior results for controlling for defense spending excess returns (Fisher and Peters (2010)) and commodity prices.

To calibrate the Gibbs sampler, I discard the first 50,000 draws as a burn-in period and keep every 20th draw thereafter until accumulating 5000 draws. With a low prior weight on the DSGE model, this generates negligible autocorrelations of model sum-

mary statistics. The sampler is less efficient with a stronger DSGE model prior but performs well with the above sample size for moderate weights on the DSGE prior. Appendix B.3 also presents evidence on the convergence of the parameter estimates based on the Brooks and Gelman (1998) diagnostic and compares IRFs for chains started at different seeds and with longer lengths.

5. RESULTS

I now discuss the fit of the narrative DSGE-SVAR along its dynamic and its identifying dimension, measured as the tightness of priors over coefficients \mathbf{B} , and the covariance matrix \mathbf{V} . Then I analyze the implied impulse-response functions, historical shocks and policy functions in the DSGE-SVAR. Last, I analyze the estimated DSGE model parameters.

5.1 Marginal likelihood

Figure 1(a) shows the marginal likelihood for the baseline model as a function of the strength of the prior for DSGE model dynamics, T_0^B , for different values of T_0^V . The marginal likelihood initially increases in T_0^B , then flattens and eventually falls for all values of T_0^V . There is thus clear evidence for an interior peak in terms of the model dynamics. However, the data clearly prefer the weakest weight for the prior on the DSGE model identification. The best-fitting model features $T_0^B = 4T$ and $T_0^V = \frac{1}{5}T$, that is, a weight of four sample sizes for the model coefficient \mathbf{B} and of one-fifth for the covariance matrix \mathbf{V} .¹⁰ Thus, a pure DSGE model is misspecified, but less so for model dynamics than for shock-identification.

The qualitative finding that DSGE model dynamics agree more with the data than the shock identification is robust to fiscal foresight. To do so, I consider a variant of the model with spending and technology news shocks and observed expectations on output and government spending. Figure 1(b) shows that the results mirror Figure 1(a). Qualitatively, the results are unchanged: The best-fitting model puts an intermediate weight on the DSGE-model prior for model dynamics, but at each point the data prefer a weaker prior on the DSGE-model covariance. The fact that both the likelihood and the relative weight on model dynamics are higher than in the baseline case likely reflects the shorter estimation sample in this extended model.

Other model variants also imply that the data prefer the weakest prior for the DSGE-model covariance \mathbf{V} . Appendix C (Figure C.8(a)) shows that the marginal likelihood is decreasing in the weight on \mathbf{V} for five alternative variants. I consider three variations on policy rules: A Smets and Wouters (2007)-type monetary policy rule that considers the output gap, the baseline monetary policy rule extended to react to fiscal policy and the baseline spending rule, extended to react to real interest rates. Only the rule with the output gap improves the fit of the best-fitting model by about 10 log points, but the

¹⁰For the best-fitting model, I have recomputed the likelihood of the best-fitting model using different seeds and doubling the number of simulations for the two additional seeds. The three different estimates lie 2 to 6 (natural) log-points from each other. The plots shows results averaged across chains.

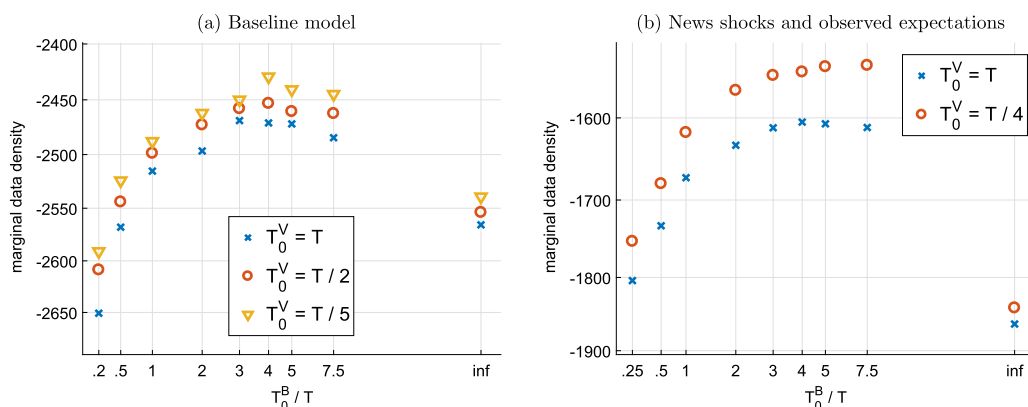


FIGURE 1. Marginal likelihood for varying DSGE model weights: baseline model and model with news shocks. The marginal narrative DSGE-SVAR likelihood peaks with interior weight on the DSGE model dynamics, but with a small weight on the covariance structure: For the baseline model in panel (a), $T_0^B = 4T$ and $T_0^V = \frac{1}{5}T$ (i.e., adding four full samples worth of observations on DSGE model dynamics but only one-fifth of a sample worth on shock and their covariance), yields the highest data density. Panel (b) shows the fit for a variant of the model with news shocks and observed expectations. With a shorter sample period, the best-fitting model in panel (b) is $T_0^B = 7.5T$ and $T_0^V = \frac{1}{4}T$.

slope in terms of the prior weight is as negative for all three variations as in the baseline model. A model version without the cost channel of monetary policy also improves the fit, but the slope is more negative than in the baseline. A version with an ad hoc investment friction, which makes the investment-specific shock correlated with current output, improves the fit the most and leads to a less negative slope. This change relaxes the identifying assumptions of the DSGE model and is thus further evidence that identifying restrictions of the DSGE model are at odds with the data.

In addition, I have also checked that my results are not driven by the assumptions about the measurement equation: Appendix C (Figure C.8(b)) shows a similar shape of the marginal likelihood when doubling (reducing) the prior for the relative standard deviation of the measurement noise from 0.5 to 1.0 (to 0.1). Allowing for a nondiagonal \mathbf{G} matrix also yields similar results. Together, my findings suggest that the empirical findings here extend beyond the specific model put forward.

To better understand the model fit along the two dimensions of shock identification and dynamics, I proceed by presenting the IRFs that shape the model dynamics in the baseline model. Subsequently, I turn to the policy shocks directly to gain intuition on the identification. Policy rule estimates and, more generally, DSGE model parameter estimates provide an additional way to understand the results summarized by the likelihood. For presenting the historical shocks and the estimated IRFs, I focus on the best-fitting DSGE-VAR model with a weak prior to over \mathbf{V} and an informative prior over \mathbf{B} , that is, the model with $T_0^V = \frac{1}{5}T$ and $T_0^B = 4 \times T$.

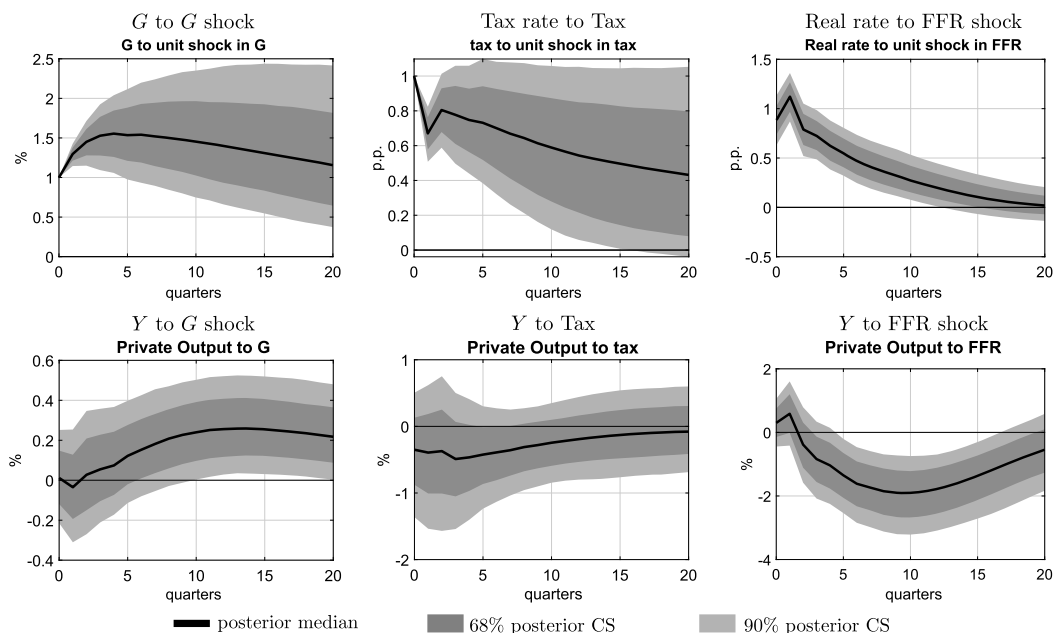


FIGURE 2. Policy shocks and output responses in best-fitting DSGE-SVAR ($T_0^B = 4T$ and $T_0^V = \frac{1}{5}T$). There is a buildup in government spending in response to a government spending shock, causing a significant and lasting increase in output. A tax shock raises tax rates persistently but implies a smooth decline to zero. With a one quarter lag, output drops persistently. A FFR shock increases the real rate and causes a significant output drop after two quarters, with no significant response on impact.

5.2 IRFs

5.2.1 Benchmark DSGE-SVAR estimates Figure 2 shows responses of private output to the three policy shocks identified in the DSGE-SVAR along with the policy instruments themselves: Shown as lines and shaded areas, respectively, are the pointwise posterior median and the 68% and 90% credible sets. It is instructive to report results for private output. Private output is overall GDP minus government consumption and investment, computed as $\hat{y}_t - 0.2\hat{g}_t$. The results confirm intuition: Spending increases are expansionary, tax and interest rate increases are contractionary. However, the responses differ in their size, timing and estimated precision.

Start with the government spending shock. A 1% shock leads to an additional buildup in government spending that declines but persists for more than 5 years. This spending increase causes private sector output to rise persistently, but with a 1 year delay. Consequently, the impact multiplier is centered around 1, but the credible set rises to a range between 1.25 and 2.0 after 5 years, as Figure 3 shows. The DSGE-SVAR estimate is consistent with other estimates of multipliers that use variation in defense spending: Ramey (2011, p. 31) estimates a 5-year cumulative multiplier of 1.2. Amir-Ahmadi and Drautzburg (2017, Figure 4.13) identified an interval of 1.0 to 3.0 consistent with macro sign restrictions and industry-level heterogeneity restrictions at the 5-year horizon.

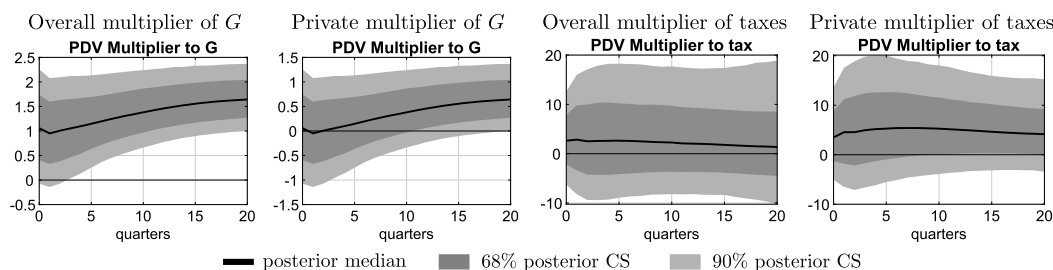


FIGURE 3. Present-discounted value multipliers for overall and private output for fiscal shocks in best-fitting DSGE-VAR ($T_0^B = 4T$ and $T_0^V = \frac{1}{3}T$). PDV multipliers are defined as the ratio of the discounted sum of GDP changes to the discounted sum of government spending, using a discount factor of 0.99 per quarter. I use a share of G in GDP of 20% and of labor taxes in GDP of 10%. For the numerator, I only consider first round effects: For shocks to G , I only consider sum the increase in G on expenditure. For shocks to taxes, I use (minus) the increase in tax rates, times the average size of the tax base.

Tax shocks have a half-life of slightly less than 5 years, but decline smoothly. The estimate of their effect on private output is noisy, but I find that an increase in tax rates by one percentage point leads, with a one year delay, to a decrease in private sector activity, which lasts for about 1 year—significant with a 68% probability. The traditional output multiplier, however, is insignificant, because overall output may rise with 68% probability due to off-setting effects on government spending. The multiplier on private output, in contrast, is weakly positive with 68% probability starting at the six quarter horizon (Figure 3). The wide credible sets reflect the sparse data on tax shock proxies.

Shocks to monetary policy, in contrast to fiscal shocks, have a half-life of only about 1 year. The rise in the nominal interest rate causes the real rate to rise and leads, with a delay of half a year, to a hump-shaped drop in private output that peaks about 2.5 years after the shock. The effects revert to zero after 5 years.

Using the narrative DSGE-SVAR rather than a DSGE-VAR approach to identify IRFs matters. Figure 4 contrasts the two approaches for the six responses from Figure 2. For the three policy instruments in the top panel, the DSGE-VAR yields responses that lie largely on top of the DSGE-SVAR policy instrument responses. But there are noticeable differences for the responses of private output: The DSGE-VAR response of private output to GDP is significantly positive on impact, as opposed to the insignificant DSGE-SVAR impact response. The DSGE-VAR also shifts the approach to a monetary policy shock considerably. It shows an immediate drop in output, eliminating the surprising pattern of an initial increase in output that only leads to a drop in output with a delay.

These results are robust. Robustness checks (available upon request) show that the the results are robust across sample periods, to using the original Ramey (2011) and Romer and Romer (2004) instruments and to controlling for expectations. Importantly, qualitatively similar responses also hold when using an upper triangular factorization of $S_1 S_1'$, or when factoring the shocks by adopting the procedure in Uhlig (2003) to factor $S_1 S_1'$. But the assumption that policy is described by observable Taylor rules substantially sharpens the inference.

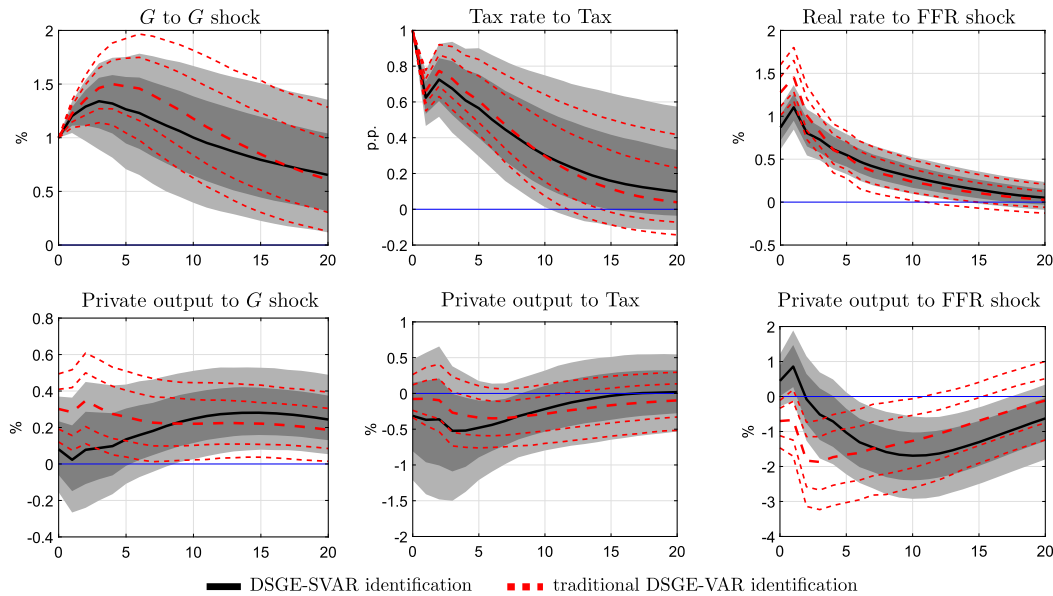


FIGURE 4. Policy shocks and output responses in best-fitting DSGE-VAR ($T_0^B = 4T$ and $T_0^V = \frac{1}{5}T$) compared with standard DSGE-VAR identification. Shown are the pointwise posterior median, and the 68% and 90% credible sets. The traditional DSGE-VAR identification uses $\mathbf{A} = \text{chol}(\boldsymbol{\Sigma}(\hat{\boldsymbol{\theta}})) \times \mathbf{Q}(\mathbf{A}^*(\boldsymbol{\theta}))$, where $\mathbf{Q}(\mathbf{A}^*(\boldsymbol{\theta})) = \text{chol}(\boldsymbol{\Sigma}^*(\boldsymbol{\theta}))^{-1}(\mathbf{A}^*(\boldsymbol{\theta}))$.

5.2.2 DSGE-SVAR versus the DSGE model To understand the shortcomings of the DSGE model that lead the data to reject the identifying restrictions implied by the DSGE model and to prefer only a modest weight on the model dynamics, I compare the estimates DSGE-SVAR with the corresponding DSGE model. Impulse-response functions reflect both identification and model dynamics and this section discusses the shortcomings of the DSGE model relative to the DSGE-SVAR impulse responses.¹¹ In particular, the pure DSGE model roughly matches the responses of most variables to fiscal shocks that include significant responses of monetary policy to these fiscal shocks. But the DSGE-SVAR produces responses to monetary policy shocks reminiscent of the price puzzle, which the pure DSGE model fails to match. It also cannot replicate the response of government spending to the other identified shocks.

For the government spending shock in Figure 5(a),¹² the first-order mismatch is that the DSGE model overstates the government spending buildup, and thus the size of the wealth effect. Otherwise, the DSGE model largely succeeds in matching the DSGE-SVAR responses after taking parameter uncertainty into account. For example, the DSGE model matches the zero impact response of private output that turns positive with a delay. The pure DSGE model also matches the implied spending multipliers; see Appendix C. The second striking fact about the identified responses is that both inflation

¹¹The results are based on the state-space representation of the DSGE model. But this is inconsequential, since the VAR approximation and the state-space representation yield virtually identical IRFs. See Figure B.3 in the Appendix.

¹²The responses of debt to all shocks are shown in Figure C.9 in Appendix C.

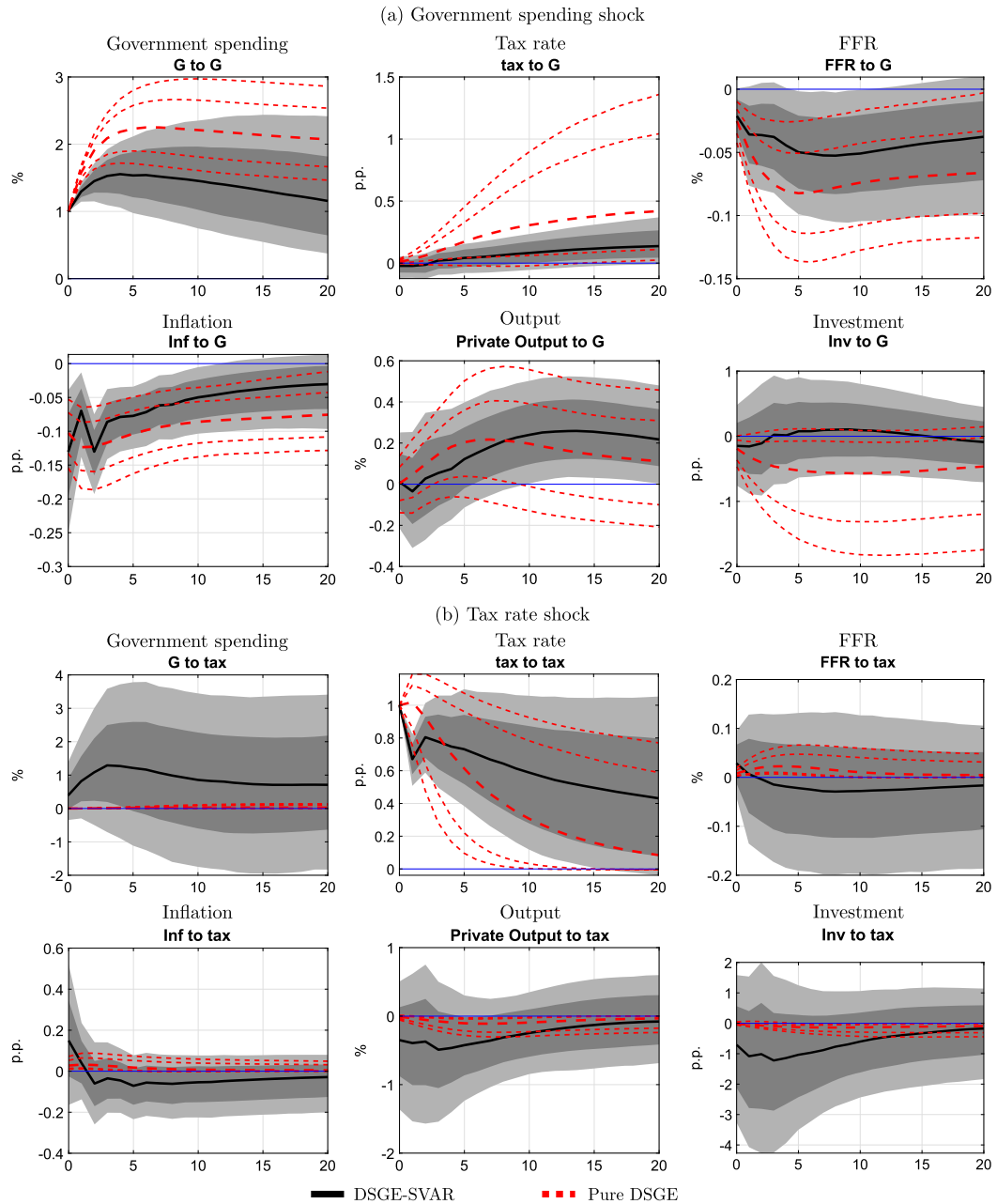


FIGURE 5. Full set of responses to fiscal policy shocks with best-fitting prior ($T_0^B = 4T$, $T_0^V = \frac{1}{5}T$). Panel (a): The pure DSGE model and the best-fitting DSGE-SVAR largely agree on the response to a government spending shock, except for the size of the build-up in government spending. Importantly, the pure DSGE model replicates the drop in the funds rate and the inflation rate following a government spending shock. Panel (b): The estimates of the response to taxes are noisy. Given the uncertainty, the pure DSGE model and the best-fitting DSGE-VAR largely agree on the response to a tax spending shock, except for the flat response of government spending in the DSGE model.

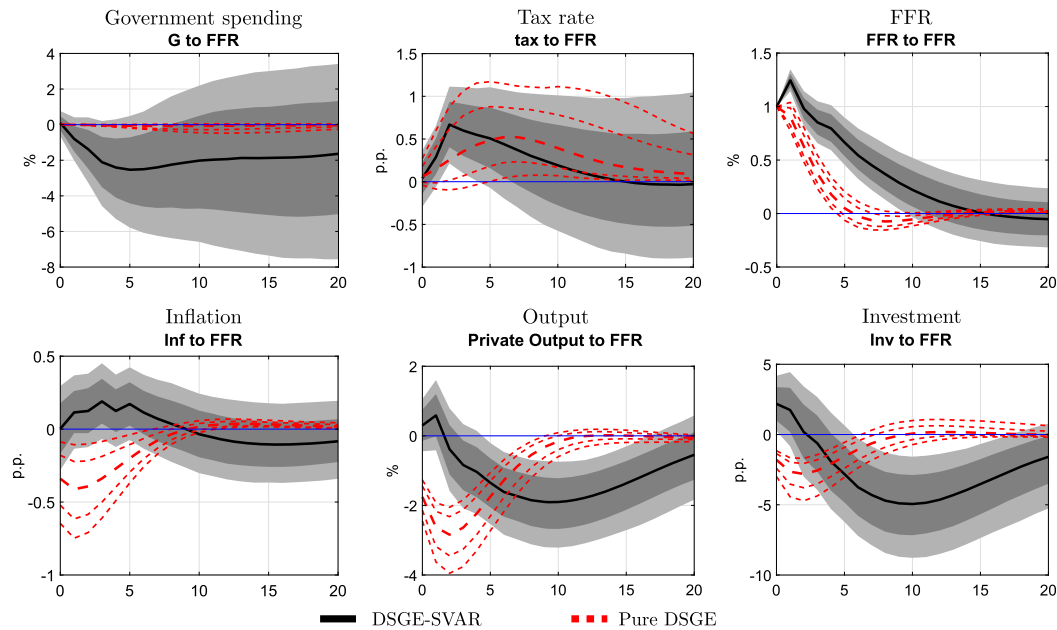


FIGURE 6. Full set of responses to FFR shock with best-fitting prior ($T_0^B = 4T$, $T_0^V = \frac{1}{5}T$). The estimates of the response to monetary policy shocks are the most precise. The estimates show that, despite including a cost-channel of monetary policy, the pure DSGE model cannot fit the initial responses of private output and inflation, as well as the delayed decline in government spending.

and the FFR drop both in the DSGE-SVAR and the pure DSGE model. This is possible even though monetary policy in the model reacts only to output and inflation.

Figure 5(b) shows the responses to a tax rate shock. The DSGE model matches the dynamics of the tax rate and most DSGE-SVAR responses to it, but cannot replicate the estimated initial increase in government spending after a tax hike. The pure DSGE model otherwise matches the qualitative features of the DSGE-SVAR responses, aided by the large uncertainty surrounding the tax shock estimates. The pure DSGE model also matches the pattern of the implied private output multiplier that is initially close to zero and becomes significantly positive at longer horizons; see Figure C.10 in Appendix C. The estimates also indicate, albeit noisily, monetary accommodation of the contractionary tax shock. In the pure DSGE model this response is driven by the underlying reaction to output and inflation.

The responses to the monetary policy shock in Figure 6 are the most precisely estimated responses and show the largest discrepancies between the DSGE-SVAR and the pure DSGE model. The latter struggles to explain the initial responses of output, investment and inflation—reminiscent of the price puzzle. This is despite including a cost channel of monetary policy as in Christiano, Trabandt, and Walentin (2011).¹³ The DSGE

¹³This finding may hint at another dimension of misspecification of monetary policy shock measures. For example, Caldara and Herbst (2019) argue that monetary policy also reacts to credit spread shocks.

model also fails to match the sizable fiscal contraction at the 1-year horizon that the DSGE-VAR implies.

5.3 Historical shocks

In addition to impulse-responses, historical shocks are an important implication of DSGE models. Similar to the overlap of impulse-response functions, I now consider the correlation of the identified shocks in the DSGE-SVAR and the pure DSGE model as a measure of misspecification. While Rudebusch (1998) and Sims (1998) have debated the validity of such a comparison across VARs, here the comparison is straightforward because the models share the same information set.

For the identified policy shocks, the shocks line up well across models and the disagreement is attributable in roughly equal parts to differences in dynamics and covariances. Table 2(a) shows the correlations for varying DSGE-model weights for each of the three shocks. The leftmost block shows the median correlation for each of the shocks for the best-fitting model, along with the posterior 90% credible set. The middle block takes the weight on the model dynamics to infinity, that is, it imposes the VAR-approximation to the DSGE model dynamics. The right block also imposes the DSGE model covariance structure. Moving from left to right, the average correlation between the three shocks rises from 0.68 to 0.86 when the model dynamics are imposed. Mechanically, the correlation becomes perfect, up to approximation error, when also imposing the DSGE model covariance structure. The two drops in the correlations of the identified shocks—by an

TABLE 2. Historical shocks in various DSGE-SVARs: correlation with DSGE model and nonzero instruments.

Shock	$T_0^V = \frac{1}{5}T, T_0^B = 4 \times T$		$T_0^V = \frac{1}{5}T, T_0^B \nearrow \infty$		$T_0^V \nearrow \infty, T_0^B \nearrow \infty$	
	Median	(90% band)	Median	(90% band)	Median	(90% band)
(a) DSGE-SVAR vs. DSGE						
<i>G</i>	0.76	(0.58, 0.85)	0.87	(0.61, 0.96)	1.00	(1.00, 1.00)
Tax	0.51	(0.23, 0.76)	0.76	(0.41, 0.91)	0.99	(0.99, 1.00)
FFR	0.77	(0.66, 0.85)	0.94	(0.88, 0.98)	1.00	(1.00, 1.00)
Average	0.68		0.86		1.00	
(b) "First stage": DSGE-SVAR vs. IV						
<i>G</i>	0.48	(0.40, 0.54)	0.50	(0.34, 0.56)	0.51	(0.50, 0.52)
Tax	0.54	(0.37, 0.68)	0.51	(0.28, 0.66)	0.36	(0.33, 0.42)
FFR	0.62	(0.59, 0.65)	0.58	(0.57, 0.59)	0.55	(0.54, 0.56)
Average	0.55		0.53		0.47	

Note: The table shows the posterior median correlations (and posterior credible sets) between the identified DSGE-SVAR shocks and the corresponding structural shocks from the DSGE model, as well as the correlation of the structural shocks with the nonmissing instrumental variables (IV). Underlying are draws from the joint posterior of the structural parameters θ and the corresponding VAR parameters.

In results available upon request, I verified that controlling for BAA bond spreads and a number of other variables does not change the result in the flat prior VAR. Figure C.12 shows that also the results of the DSGE-SVAR are robust to controlling for expectations.

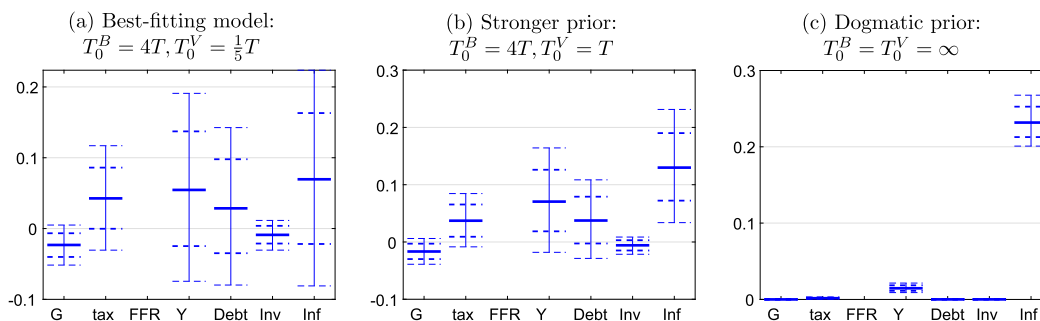


FIGURE 7. Monetary policy rule estimates. Shown are the posterior median and 68% and 90% credible set of the monetary policy rule coefficients implied by the partially identified VAR with DSGE model prior. With the best-fitting DSGE-VAR prior, the model is inconclusive about monetary tightening in response to higher inflation but finds evidence of accommodating fiscal policy. Only priors putting more weight on the DSGE-model identification find a significant reaction to output and inflation.

average of 0.14 when relaxing the prior on the DSGE-model covariance structure and by an average of 0.18 when relaxing the prior on the dynamics—show that shock identification requires getting the forecast errors right and correctly mapping the forecast errors to structural shocks.

In terms of individual shocks, the DSGE-SVAR and the DSGE model disagree the most on the tax shock. For the best-fitting model, the right block of Table 2(a) shows that the monetary policy shock paths and the fiscal policy shock paths line up about equally well with correlations of 0.76 and 0.77. For tax shocks, in contrast, the discrepancy is more pronounced with a correlation of only 0.51 across models, although the posterior uncertainty is substantial. Also, the correlations between the (nonzero) instruments and the DSGE-SVAR variables reflect the challenge the DSGE model faces in matching the tax shocks. The best-fitting DSGE-VAR has correlations of the identified shocks with the instruments of 0.48 for the spending shock, 0.54 for the tax shock and 0.62 for the monetary policy shock. However, particularly as the weight on the identification scheme increases, the correlations approach that of the pure DSGE model. The pure DSGE model has a correlation of the tax shock proxies with the identified shocks of only 0.36, much lower than the values of 0.51 and 0.55.

5.4 Policy rules

Partial identification of policy shocks implies identification of the underlying policy rules. Arias, Caldara, and Rubio-Ramirez (2015) exploited this to identify shocks with sign restrictions, and it underlies the link between the narrative identification and the DSGE model in Proposition 1. It is therefore instructive to back out the underlying policy rule coefficients to understand the mechanics of the DSGE-SVAR model. These coefficients correspond to the coefficients $(1 - \rho_o) \times \eta_o$ in the policy rules (4.1) and (4.2).

The raw data only provide noisy information about policy rules: As I document in Appendix C (Table C.3), the coefficients of all three policy rules have wide posterior

confidence bands, which are often insignificant for the best-fitting model. Figure 7 illustrates the estimates for the monetary policy rule. Both the 68% and the 90% credible sets for the responses of monetary policy to inflation and output include zero. A stronger DSGE model prior over \mathbf{V} helps by shrinking the imprecise estimates of the policy rule coefficients toward those implied by the DSGE model prior. This shrinkage renders the inflation and output coefficients positive. With the DSGE model prior, the policy rule estimates imply that the Fed increases interest rates in response to higher inflation or output growth. With a dogmatic prior on the DSGE model, we recover that monetary policy only responds to inflation and output. However, the data clearly prefer the model with the diffuse policy rule estimates.

For both fiscal policy shocks, the estimated DSGE-SVAR impulse responses found that monetary policy accommodates: Lower interest rates follow expansionary G shocks and higher interest rates the contractionary tax shocks. The pure DSGE model could match this because fiscal shocks affect inflation via marginal costs—indirectly via the wealth effects and the pass-through of higher taxes. My policy rule estimate, however, indicates that there is also a direct element of accommodation in the monetary policy rule.

Both reduced-form correlations and the institutional setting make the estimated direct fiscal-monetary policy plausible. For example, Calomiris (2012) characterized the Federal Reserve as a “highly politicized entity” that depends on the federal government (p. 68)—the fiscal authority. Romer and Romer (2014) provided qualitative evidence that the Federal Reserve has indeed considered fiscal policy in its monetary policy decisions. They document staff presentations to the FOMC suggesting monetary accommodation of the 1964 and 1972 tax cuts (p. 38f) as well as monetary easing in response to the 1990 budget agreement. Estimating monetary policy rule directly in the reduced form also yields a qualitatively similar result:

$$\text{FFR}_t = \underset{(-2.32)}{-2.90} + \underset{(18.24)}{0.89} \text{FFR}_{t-1} + \underset{(1.96)}{0.17} \pi_t + \underset{(1.60)}{0.06} \Delta \ln Y_t + \underset{(-3.24)}{-0.12} \Delta \ln \text{Debt}_t + \underset{(-2.40)}{-1.83} \ln \frac{G_t}{Y_t}, \quad (5.1)$$

where $N = 213$, $R^2 = 0.90$ and the parenthesis show t -statistics based on Newey–West robust standard errors estimated with four lags. This reduced form estimate is robust across specification and sample periods (results available upon request). Similarly, reduced form spending rules also give a role to the real short-term interest rate.

While the model misspecification manifests itself in the policy block, it is not confined to it: Extending the policy rules does not fix the misspecification. I reestimated the DSGE-SVAR with a monetary policy rule that responds to debt levels or spending and, separately, with a fiscal policy rule that responds to the real rate. Both modifications reduce the discrepancies of the investment IRF to both spending and tax shocks. However, both specifications produce responses of private output to government spending and taxes in the DSGE model which are, for the first few quarters, more at odds with the SVAR than the original specification. And the policy rule estimates in the DSGE-SVAR change little. As a consequence, it is unsurprising that these model variations do not change the marginal data density much, either.

5.5 Parameter estimates

As a byproduct of the estimation of the DSGE-SVAR, the estimator also yields estimates of the structural DSGE model parameters. For brevity, I present the full estimates only in Table C.1 in the Appendix. One feature that emerges is that the DSGE-model parameters feature only a low persistence of most structural shocks, but a high degree of smoothing in the policy rules. This helps to explain the inability of the pure DSGE model to match the responses of government spending to the other policy shocks. With a stronger prior on \mathbf{V} , the parameter estimates also feature higher real and nominal frictions in the form of adjustment and fixed costs, and the Calvo price stickiness.¹⁴ With few exceptions, the posterior differs significantly from the prior and the policy rule estimates are economically meaningful. The exception is the wage indexation parameter whose posterior is close to its prior. This is plausible because I do not use wages in the estimation.

I now turn to the parameters that are specific to my narrative DSGE-VAR: the parameters of the observation equation for the narrative instruments in the DSGE model. Figure 8 shows the signal to noise ratio estimated for each of the three narrative shocks. For the best-fitting model with $T_0^B = 4T$ in panel (a), the posterior median signal-to-noise ratio is about 0.5 for all three instruments. With the stronger prior on \mathbf{V} , the distribution of the estimated signal-to-noise ratios seems to shift somewhat down in panel (b) compared to panel (a). Intuitively, because the data prefers the looser prior on the DSGE model identification, it selects lower signal-to-noise ratios when forced to tighten the prior. The differences concern, however, mostly the center of the distribution, while the credible sets overlap.

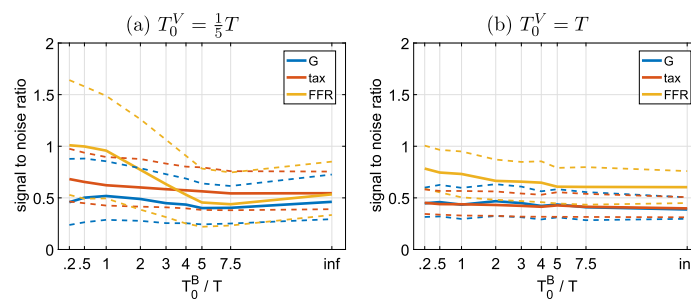


FIGURE 8. Signal-to-noise ratio of instruments with varying DSGE model weight. The plot shows the signal-to-noise ratio as a function of the prior weight given to the DSGE model. The ratio is defined as the standard deviation of the instrument attributable to the structural shock divided by the standard deviation of the measurement noise. The prior signal-to-noise ratio in the DSGE model is 2. The plot implies that the data is informative about the signal-to-noise ratio that falls initially quickly and then stabilizes around 0.4 for the two fiscal instruments and falls toward 0.75 for the monetary policy shock.

¹⁴Figure B.7 in the Appendix uses the Brooks and Gelman (1998) diagnostic to show that the posterior simulator for θ has converged reasonably well.

6. CONCLUSION

A key question for academics and practitioners using quantitative DSGE models is whether these models agree with methods that “get by with weak identification” (Sims (2005, p. 2)). This paper develops a DSGE-SVAR to assess the potential misspecification of DSGE models with respect to the identification of shocks through external instruments, extending earlier work (Del Negro et al. (2007)) that does not assess the DSGE model identification.

Estimating my medium-scale DSGE-SVAR model with fiscal Taylor rules reveals systematic model misspecification. The prior over DSGE model dynamics improves the overall statistical fit as measured by the marginal data density. However, I show that the fit worsens when putting a stronger prior on the DSGE model covariance structure, which embeds the identifying restrictions. This finding carries over from the baseline model with contemporaneous shocks to an extended model with news shocks and observed expectations, as well other model variants.

Looking at impulse responses shows that the best-fitting DSGE-VAR and the corresponding pure DSGE model largely agree on the dynamics following fiscal shocks, but the pure DSGE model cannot capture the interaction between policy tools. The model has trouble matching the dynamics following a monetary policy shock. Analyzing historical shocks shows that the DSGE model struggles particularly to explain the tax shock proxy. My estimates of policy rules show that systematic monetary policy broadly moves with fiscal policy, reinforcing spending or tax increases. However, simply reestimating the model with richer policy rules does not change the main findings. Overall, my results suggest that more work is needed to reconcile medium-scale DSGE models for fiscal policy with narrative methods.

REFERENCES

- Adjemian, S., H. Bastani, M. Juillard, F. Mihoubi, G. Perendia, M. Ratto, and S. Villemot (2011), “Dynare: Reference manual, version 4.” Dynare Working Paper 1, CEPREMAP. [820]
- Amir-Ahmadi, P. and T. Drautzburg (2017), “Identification through heterogeneity.” Working Paper 17-11, Federal Reserve Bank of Philadelphia. [824]
- Antolin-Diaz, J. and J. F. Rubio-Ramirez (2016), “Narrative sign restrictions for SVARs.” Working Paper 2016-16, Federal Reserve Bank of Atlanta. [806]
- Arias, J. E., D. Caldara, and J. Rubio-Ramirez (2015), “The systematic component of monetary policy in SVARs: An agnostic identification procedure.” Unpublished manuscript, Duke University. [804, 830]
- 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. [807, 810]
- Bohn, H. (1991), “Budget balance through revenue or spending adjustments?: Some historical evidence for the United States.” *Journal of Monetary Economics*, 27 (3), 333–359. [818]

- Brooks, S. P. and A. Gelman (1998), “General methods for monitoring convergence of iterative simulations.” *Journal of Computational and Graphical Statistics*, 7, 434–455. [822, 832]
- Caldara, D. and E. Herbst (2019), “Monetary policy, real activity, and credit spreads: Evidence from Bayesian proxy SVARs.” *American Economic Journal: Macroeconomics*, 11 (1), 157–192. [802, 803, 828]
- Calomiris, C. W. (2012), “Macroeconomic effects of Federal Reserve forward guidance: Comment.” *Brookings Papers on Economic Activity*, Spring, 55–63. [831]
- Canova, F., F. Ferroni, and C. Matthes (2013), “Choosing the variables to estimate singular DSGE models.” CEPR Discussion Paper 9381. [820]
- Chib, S. (1995), “Marginal likelihood from the Gibbs output.” *Journal of the American Statistical Association*, 90, 1313–1321. [817]
- Christiano, L., M. Eichenbaum, and S. Rebelo (2011), “When is the government spending multiplier large?” *Journal of Political Economy*, 119 (1), 78–121. [803]
- Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005), “Nominal rigidities and the dynamic effects of a shock to monetary policy.” *Journal of Political Economy*, 113 (1), 1–45. [801, 802, 819]
- Christiano, L. J., M. Trabandt, and K. Walentin (2011), “DSGE models for monetary policy analysis.” In *Handbook of Monetary Economics*, Vol. 3 (B. M. Friedman and M. Woodford, eds.), Chapter 7, 285–367, Elsevier, Amsterdam. [820, 828]
- Clarida, R., J. Galí, and M. Gertler (2000), “Monetary policy rules and macroeconomic stability: Evidence and some theory.” *Quarterly Journal of Economics*, 115 (1), 147–180. [803]
- Coibion, O. and Y. Gorodnichenko (2012), “Why are target interest rate changes so persistent?” *American Economic Journal: Macroeconomics*, 4 (4), 126–162. [811]
- Del Negro, M. and F. Schorfheide (2004), “Priors from general equilibrium models for vars.” *International Economic Review*, 45 (2), 643–673. [802, 804, 806, 814, 817]
- Del Negro, M. and F. Schorfheide (2009), “Monetary policy analysis with potentially misspecified models.” *American Economic Review*, 99 (4), 1415–1450. [811]
- Del Negro, M., F. Schorfheide, F. Smets, and R. Wouters (2007), “On the fit of new Keynesian models.” *Journal of Business & Economic Statistics*, 25, 123–143. [802, 817, 833]
- Drautzburg, T. (2016), “A narrative approach to a fiscal DSGE model.” Working Paper 16-11, Federal Reserve Bank of Philadelphia. [809]
- Drautzburg, T. (2020), “Supplement to ‘A narrative approach to a fiscal DSGE model.’” *Quantitative Economics Supplemental Material*, 11, <https://doi.org/10.3982/QE1083>. [804]

Drautzburg, T. and H. Uhlig (2015), “Fiscal stimulus and distortionary taxation.” *Review of Economic Dynamics*, 18 (4), 894–920. [819, 820]

Faust, J. (2009), “The new macro models: Washing our hands and watching for icebergs.” *Sveriges Riksbank Economic Review*, 1, 45–68. [801]

Fernandez-Villaverde, J., P. A. Guerron-Quintana, K. Kuester, and J. Rubio-Ramirez (2015), “Fiscal volatility shocks and economic activity.” *American Economic Review*, 105 (11), 3352–3384. [803, 819]

Fernandez-Villaverde, J., J. F. Rubio-Ramirez, T. J. Sargent, and M. W. Watson (2007), “ABCs (and Ds) of understanding VARs.” *American Economic Review*, 97 (3), 1021–1026. [814, 816, 820]

Fernández-Villaverde, J., J. F. Rubio-Ramírez, and F. Schorfheide (2016), “Solution and estimation methods for DSGE models.” In *Handbook of Macroeconomics*, Vol. 2 (J. B. Taylor and H. Uhlig, eds.), Chapter 9, 527–724, Elsevier, Amsterdam. [802]

Fisher, J. D. M. and R. Peters (2010), “Using stock returns to identify government spending shocks.” *Economic Journal*, 120 (544), 414–436. [821]

Francis, N. and V. A. Ramey (2009), “Measures of per capita hours and their implications for the technology-hours debate.” *Journal of Money, Credit and Banking*, 41 (6), 1071–1097. [818]

Geweke, J. (1999), “Using simulation methods for Bayesian econometric models: Inference, development, and communication.” *Econometric Reviews*, 18 (1), 1–73. [817]

Geweke, J. (2005), *Contemporary Bayesian Econometrics and Statistics*. Wiley Series in Probability and Statistics. Wiley, New York. [815]

Guerron-Quintana, P. A. (2010), “What you match does matter: The effects of data on DSGE estimation.” *Journal of Applied Econometrics*, 25 (5), 774–804. [820]

Iskrev, N. (2014), “Choosing the variables to estimate singular DSGE models: Comment.” Dynare Working Paper 41, CEPREMAP. [820]

Kuttner, K. N. (2001), “Monetary policy surprises and interest rates: Evidence from the Fed funds futures market.” *Journal of Monetary Economics*, 47 (3), 523–544. [813]

Leeper, E. M., M. Plante, and N. Traum (2010), “Dynamics of fiscal financing in the United States.” *Journal of Econometrics*, 156 (2), 304–321. [803, 819]

Lopes, H. F. and N. G. Polson (2014), “Bayesian instrumental variables: Priors and likelihoods.” *Econometric Reviews*, 33 (1–4), 100–121. [812]

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. [802, 803, 808, 809, 810, 811, 813, 818]

Mertens, K. and M. O. Ravn (2014), “A reconciliation of SVAR and narrative estimates of tax multipliers.” *Journal of Monetary Economics*, 68 (Supplement), 1–19. [807]

- Park, W. Y. (2011), "Evaluation of DSGE models: With an application to a two-country DSGE model." Unpublished manuscript, Princeton University. [802]
- Ramey, V. A. (2011), "Identifying government spending shocks: It's all in the timing." *Quarterly Journal of Economics*, 126 (1), 1–50. [803, 804, 818, 821, 824, 825]
- Romer, C. D. and D. H. Romer (2004), "A new measure of monetary shocks: Derivation and implications." *American Economic Review*, 94 (4), 1055–1084. [801, 803, 806, 819, 825]
- 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. [818]
- Romer, C. D. and D. H. Romer (2014), "Transfer payments and the macroeconomy: The effects of social security benefit changes, 1952–1991." Working Paper 20087, National Bureau of Economic Research. [831]
- Rossi, P. E., G. M. Allenby, and R. McCulloch (2005), *Bayesian Statistics and Marketing*. Wiley Series in Probability and Statistics. Wiley, New York. [812]
- Rudebusch, G. D. (1998), "Do measures of monetary policy in a VAR make sense?" *International Economic Review*, 39 (4), 907–931. [829]
- Sims, C. A. (1998), "Comment on Glenn Rudebusch's "Do measures of monetary policy in a VAR make sense?"". *International Economic Review*, 39 (4), 933–941. [829]
- Sims, C. A. (2005), "The state of macroeconomic policy modeling: Where do we go from here?" [801, 833]
- Smets, F. and R. Wouters (2007), "Shocks and frictions in US business cycles: A Bayesian DSGE approach." *American Economic Review*, 97 (3), 586–606. [801, 802, 819, 820, 822]
- Stock, J. H. and M. W. Watson (2012), "Disentangling the channels of the 2007–2009 recession." *Brookings Papers on Economic Activity*, Spring, 81–141. [802, 808, 809]
- Theil, H. and A. S. Goldberger (1961), "On pure and mixed statistical estimation in economics." *International Economic Review*, 2 (1), 65–78. [814]
- Uhlig, H. (1994), "What macroeconomists should know about unit roots: A Bayesian perspective." *Econometric Theory*, 10 (3/4), 645–671. [812]
- Uhlig, H. (2003), "What moves GNP?" Draft, Humboldt Universität zu Berlin. [825]
- Uhlig, H. (2005), "What are the effects of monetary policy on output? Results from an agnostic identification procedure." *Journal of Monetary Economics*, 52 (2), 381–419. [802]
- Waggoner, D. F. and T. Zha (2012), "Confronting model misspecification in macroeconomics." *Journal of Econometrics*, 171 (2), 167–184. [802]
- Watson, M. W. (1994), "Vector autoregressions and cointegration." In *Handbook of Econometrics*, Vol. 4, (R. F. Engle and D. McFadden, eds.) Chapter 47, 2843–2915, Elsevier, Amsterdam. [810]

Yang, S.-C. S. (2007), "A chronology of federal income tax policy: 1947–2009." Working Paper 2007-021, CAEPR. [818]

Co-editor Frank Schorfheide handled this manuscript.

Manuscript received 1 March, 2018; final version accepted 2 June, 2019; available online 11 December, 2019.