A Narrative Approach to a Fiscal DSGE Model

This version: July 16, 2014
First version: February 2014
Thorsten Drautzburg

– PRELIMINARY AND INCOMPLETE –

Abstract

Structural DSGE models are used both for analyzing 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 VARs using, for example, narrative shock measures. This paper proposes to use narrative shock measures to aid in the estimation of DSGE models. Here I adapt the existing methods for shock identification with external “narrative” instruments for Bayesian VARs. I find that VAR-identified government spending shocks, tax shocks, and monetary policy shocks line up well with those based on DSGE models. A standard DSGE with simple fiscal rules fails, however, to capture policy interactions and the response of private sector variables to government spending. Estimating the loadings of narrative shocks in structural shocks in the fiscal DSGE model, I find empirical support for the assumption that narrative shocks only reflect the structural shock they are supposed to measure. Methodologically, I show that the proposed Bayesian SUR-type estimator also captures the uncertainty about the instrument quality and has, with enough data, good sampling properties in the present application.

Keywords: Fiscal policy; DSGE model; Bayesian estimation; narrative shocks; Bayesian VAR
1 Introduction

With monetary policy constrained by the Zero Lower Bound, “stimulating” fiscal policy has gained a lot of attention. Quantitative DSGE models such as Christiano et al. (2005) have been used to analyze the effect of discretionary fiscal policy interventions (Christiano et al., 2009). But are the fiscal policy implications of these models consistent with alternative identification schemes? To answer this question, this paper uses narrative measures of policy shocks to compare histories and responses VAR-identified shocks to the histories and responses from a standard fiscal DSGE model.

Systematically assessing the structural “fit” of DSGE models complements the existing literature which has documented statistical fit. For example, Smets and Wouters (2007) showed that an estimated medium-scale DSGE model could beat a VAR in terms of the forecasting performance. Negro et al. (2013) argue that with small modifications, i.e. adding a parsimoniously modeled financial sector and observed inflation expectations, an otherwise standard medium-scale DSGE model also fits well the US experience since the Great Recession. However, an adequate structure is a required maintained assumption in this literature. Since the fiscal building blocks of DSGE models are less well studied than, say, the Taylor rule for monetary policy (e.g. Clarida et al., 2000), assessing the fiscal policy implications of these models is warranted. Indeed, fully structural and partial identification of fiscal shocks can lead to widely different conclusions about the drivers of business cycles: Rossi and Zubairy (2011) document that when applying a Blanchard and Perotti (2002)-type identification of government spending shocks in a VAR, the fraction of the variance of GDP driven by those shocks rises significantly with the forecast horizon, whereas the DSGE-model based variance decomposition in Smets and Wouters (2007) implies opposite pattern.  

Methodologically, this paper complements the recent advances VAR-based shock identification using narrative measures as instruments (Stock and Watson, 2012; Mertens and Ravn, 2013) in the frequentist framework by proposing a simple Bayesian analogue. To achieve identification with multiple narratively identified shocks, I

1Table 2 in Rossi and Zubairy (2011) implies that the fraction of the variance of GDP driven by government spending shocks rises with the forecast horizon from below 5% at four quarters to 35% at 40 quarters whereas (Smets and Wouters, 2007, Figure 1) implies that the output variance explained by GDP falls from roughly 30% at one quarter to about 15% at four quarters and less than 5% at 40 quarters. Note that throughout this paper, I focus on discretionary fiscal policy, as opposed to the effect of policy rules.
maximize the impact of a given shock to the forecast error variance attributable to the narratively identified shocks. This is a conditional version of the procedure used in Uhlig (2003). The implications for inference of the proposed framework are best understood in analogy to 2 stage least squares reasoning. It allows for natural inference about both the second stage coefficients on structural shocks as well as the first stage covariance between instruments and shocks. As a baseline, I specify a flat prior over the first stage covariances and still find in a Monte Carlo study that, with realistic sample sizes for US post-WWII data, the implied estimator has good sampling properties. When only few observations are available, I show that the proposed estimator with an improper prior has (Bayesian) confidence intervals with an actual size exceeding the nominal size. In comparison, frequentist estimators ignoring first stage uncertainty seem to produce confidence intervals understating the actual size.

Since forming priors on the many structural parameters in the VAR is hard, a natural extension, which I investigate in ongoing work, is to elicit priors over the structural parameters using priors over DSGE models. This has been proposed by Del Negro and Schorfheide (2004) and can be extended to the present application by simply adding observation equations for the narrative shocks to the DSGE model and specifying priors over the measurement error of the narrative shocks.

In this paper, I focus on surprise government spending, income taxes, and monetary policy shocks using instruments from Ramey (2011), Mertens and Ravn (2013), and Romer and Romer (2004), respectively. I use the standard medium-scale DSGE model from Christiano et al. (2005) and Smets and Wouters (2007), but add linear fiscal policy rules as in Leeper et al. (2010) and Fernandez-Villaverde et al. (2011). According to the preliminary results, I find that the implied shock histories of the VAR and the DSGE model roughly coincide for all three shocks. Also the response to a tax shock is broadly consistent. However, the VAR evidence points to a stimulating effect of government spending on private sector activity and fiscal-monetary policy interactions not predicted by the DSGE model. If robust, this would suggest a role for productive government spending as in Drautzburg and Uhlig (2011) or consumption complementarities as in Coenen et al. (2012).

A different interpretation of the differences in the responses and policy interaction is that narrative shocks reflect shocks other than the ones assumed in the model. The DSGE model allows to test whether a given narrative shock indeed loads up on more than one structural shock. Preliminary findings provide some evidence in
favor of the correct structural interpretation of the narrative shock measures.

Under the maintained assumption that the narrative shock measures correctly reflect just one structural shock, partial identification of impulse-response-functions in the VAR is of interest also when estimating DSGE models under partial information by IRF-matching as in Christiano et al. (2005) or Altig et al. (2005). Since no stark zero-restrictions are needed, the proposed identification scheme may provide a useful set of moments which can be more naturally matched by DSGE models.

In assessing the structural fit of DSGE models, this paper relates to implicit or explicit discomfort with structural identification using DSGE models in the literature. Structural DSGE models serve as a tool to assess the quantitative plausibility of “stories” about the causes of economic fluctuations and consequences of different policies. Sims (2005) cautions that economists may be set back “[i]f Bayesian DSGE’s displace methods that try to get by with weak identification and in the process reinforce the excess weight we give to story-spinning” (p. 2). Mian and Sufi (2012) only use micro data to identify demand shocks and then resort to partial model-specification to assess the macro implications. Ferreira (2013), in his analysis of risk shocks, also shies away of using an estimated DSGE model to back out shocks and instead uses only model-derived sign restrictions for fear of the DSGE model misspecification. Drautzburg (2013) identifies shocks from cross-sectional data relying only on part of the structural model and then uses them as an input into a dynamic structural model.

This paper is structured as follows: First, I describe estimation and shock identification in the VAR. Second, I lay out the DSGE model used for comparison. The third section describes the data and the empirical results, focusing on the implications for the history of shocks and IRFs. The fourth and final section evaluates whether, seen through the lenses of the DSGE model, the narrative shocks correctly reflect only a given structural shock.

2 BVAR estimation with narrative instruments

In this section, I follow the exposition in Mertens and Ravn (2013) to derive identifying restrictions based on narrative shock measures. I first discuss population results without parameter uncertainty to fix ideas. Then I proceed to discuss the sampling scheme and provide a simple Monte Carlo study of its sampling properties.

2Section ?? in the appendix follows the derivation in Stock and Watson (2012).
Finally, I discuss how to obtain responses of variables omitted from the core VAR to the identified shocks.

2.1 Narrative BVAR

I use the following notation for the data generating process:

\[ y_t = \mu_y + B y_{t-1} + v_t \quad (2.1a) \]

\[ v_t = A\epsilon_t^{str}, \epsilon_t^{str} \sim N(0, I_m) \quad (2.1b) \]

\[ z_t = \mu_z + F v_t + \Omega^{-1/2} u_t, u_t \sim N(0, I_k) \quad (2.1c) \]

Here, \( Y_t \) is the observed data, \( B \) is a matrix containing the (possible stacked) lag coefficient matrices of the equivalent VAR(p) model as well as constants and trend terms, \( v_t \) is the \( m \)-dimensional vector of forecast errors, and \( z_t \) contains \( k \) narrative shock measures.

Note that, knowing \( B \), \( v_t \) is data. We can thus also observe \( \text{Var}[v_t] = AA' \equiv \Sigma \).

Note that \( A \) is identified only up to an orthonormal rotation: \( \bar{A}Q(\bar{A}Q)' = AA' \) for \( \bar{A} = \text{cho}(\Sigma) \) and \( QQ' = I \).

The observation equation for the narrative shocks (2.1c) can alternatively be written as:

\[ z_t = \begin{bmatrix} G & 0 \end{bmatrix} \epsilon_t^{str} + \Omega^{-1/2} u_t = \begin{bmatrix} G & 0 \end{bmatrix} A^{-1} A\epsilon_t^{str} + \Omega^{-1/2} u_t \equiv F \]  

By imposing zero restrictions on \( G \), knowledge of \( F \) and \( AA' \) identifies the shocks which are not included in 2.2. However, only the covariance matrix \( FA \) is needed for identification. It is therefore convenient for inference and identification to introduce a shorthand for the covariance matrix between the instruments and the forecast errors:

\[ \Gamma \equiv \text{Cov}[z_t, v_t] = FA \]  

The model in (2.1) can then be written compactly as:

\[ \begin{bmatrix} y_t \\ z_t \end{bmatrix} Y_{t-1} \sim \mathcal{N} \left( \begin{bmatrix} \mu_y + B y_{t-1} \\ \mu_z \end{bmatrix}, \begin{bmatrix} AA' & \Gamma' \\ \Gamma & \tilde{\Omega} \end{bmatrix} \right), \quad (2.4) \]
where \( \tilde{\Omega} = \Omega + FAA'F' \) is the covariance matrix of the narrative instruments.

### 2.1.1 Identification given parameters

This section largely follows Mertens and Ravn (2013). It considers the case of as many instruments as shocks to be identified, with \( k \leq m \).

Partition \( A = [\alpha^{[1]}, \alpha^{[2]}] = \begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{bmatrix} \), \( \alpha^{[1]} = [\alpha'_1, \alpha'_2]' \) with both \( \alpha_{11}(k \times k) \) being invertible and \( \alpha_{21}((m - k) \times k) \).

Using the definitions of \( \Gamma \) and the forecast errors gives from (2.3):

\[
\Gamma \equiv \text{Cov}[z_t, v_t] = \text{Cov}[z_t, A\epsilon^{\text{str}}_t] = \begin{bmatrix} G & 0 \end{bmatrix} A' = G\alpha'_1 = [Ga'_{11}, Ga'_{21}] \tag{2.5}
\]

If the narrative shocks together identify an equal number of structural shocks, \( G \) is of full rank. Since also \( \alpha_{11} \) is invertible:

\[
G\alpha'_{11} = \Gamma_1, \tag{2.6a}
\]

\[
\alpha'_{21} = G^{-1}\Gamma_2 = \alpha'_{11}(\Gamma_1^{-1}\Gamma_2), \equiv \alpha'_{11}\kappa' \tag{2.6b}
\]

where \( \Gamma \) is a known reduced form parameter matrix that can be estimated.

Hence the (structural) impulse-vector to shocks 1, \ldots, \( k \) satisfies:

\[
\alpha^{[1]} = \begin{bmatrix} \alpha_{11} \\ \alpha_{21} \end{bmatrix} = \begin{bmatrix} I_k \\ (\Gamma_1^{-1}\Gamma_2)' \end{bmatrix} \alpha_{11}, \tag{2.7}
\]

where \( \Gamma_2 \) is \( k \times (m - k) \) and \( \Gamma_1 \) is \( k \times k \). This \( m \times k \) dimensional vector is a known function of the \( k^2 \) parameters in \( \alpha_{11} \). It therefore restricts \( (m - k)k \) elements of \( A \).

An alternative way of stating the under-identification of VARs without external instruments or other restrictions on the coefficients is to note that \( A \) is only identified up to multiplication by an orthonormal rotation matrix \( Q' \). Since \( A \) is assumed to be of full rank and, hence, \( \Sigma \) is positive definite, the Choleski factorization of \( \Sigma \) exists: \( \tilde{A}\tilde{A}' = \Sigma \). Since \( A \) and \( \tilde{A} \) are of full rank, there is a unique \( Q \) such that \( A = \tilde{A}Q \).

Here, \( Q \) and therefore \( A \) is typically only partially identified. Order these columns first as above. Then:

\[
[\alpha^{[1]}, \alpha^{[2]}] = \tilde{A}Q = [\tilde{A}q^{[1]}, \tilde{A}q^{[2]}]
\]
Thus:

\[ \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} = \alpha^{[1]} = \tilde{A}q^{[1]} \quad \Leftrightarrow \quad q^{[1]} = \tilde{A}^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} \]

\( \alpha_{11} \) is partially pinned down from the requirement that \( Q \) be a rotation matrix.

In particular:

\[
I_k = (q^{[1]})'q^{[1]} = \alpha'_{11} \begin{bmatrix} I_k \\ \kappa \end{bmatrix}'(\tilde{A}')^{-1}\tilde{A}^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} = \alpha'_{11} \begin{bmatrix} I_k \\ \kappa \end{bmatrix}'\Sigma^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} \quad (2.8)
\]

\[ \Leftrightarrow (\alpha'_{11})^{-1}(\alpha_{11})^{-1} = \begin{bmatrix} I_k \\ \kappa \end{bmatrix}'\Sigma^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \Leftrightarrow \alpha_{11} \alpha'_{11} = \left( \begin{bmatrix} I_k \\ \kappa \end{bmatrix}'\Sigma^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \right)^{-1} \quad (2.9) \]

This requires \( \frac{k(k-1)}{2} \) additional restrictions to identify \( \alpha_{11} \) completely.

The following Lemma summarizes the above results:

**Lemma 1** (Stock and Watson, 2012; Mertens and Ravn, 2013) The impact of shocks with narrative instruments is generally identified up to a \( k \times k \) scale matrix \( \alpha_{11} \) whose outer product \( \alpha_{11}\alpha'_{11} \) is known, requiring an extra \( \frac{(k-1)k}{2} \) identifying restrictions and the impulse vector is given by (2.7). Proof: See Appendix A.1.

A simple way to identify \( \alpha_{11} \) would be to require it to be lower diagonal. This imposes that the one-period ahead forecast error variance of the first variable in the VAR that is jointly attributed to the \( k < m \) identified shocks is explained exclusively by the first shock. To avoid the stark zero restrictions the Choleski-factorization imposes, I adapt a more general identification scheme described in Uhlig (2003) and used by Barsky and Sims (2009). I choose \( \alpha_{11} \) to maximize the forecast error variance attributed to the first shock over the horizon \( \{\bar{h}, \ldots, \bar{h}\} \). Following the steps in Uhlig (2003), this amounts to solving the following principal components problem:

\[
\max_{\lambda_{\ell}} S_{\bar{h}}^{\alpha_{11}} = \lambda_{\ell} \bar{q}_{\ell}, \quad \ell \in \{1, \ldots, k\}. \quad (2.10a)
\]

\[
S = \sum_{h=0}^{\bar{h}} (\bar{h} + 1 - \max\{h, h\}) \left( B^h \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \tilde{\alpha}_{11} \right)' e_{1} e_{1}' \left( B^h \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \tilde{\alpha}_{11} \right), \quad (2.10b)
\]

\(^3\)Here I used that the variance whose forecast error variance is of interest is among the first \( k \) variables. Otherwise, redefine \( \alpha_{11} \) to include \( \kappa \) or reorder variables as long as \( \alpha_{11} \) is guaranteed to be invertible.
where $\tilde{\alpha}_{11}$ is the Choleski factorization associated with (2.9) and $e_1$ is a selection vector with zeros in all except the first position. Then the desired $\alpha_{11}$ is given by:

$$\alpha_{11} = \tilde{\alpha}_{11}q_t^\alpha,$$

where the eigenvectors $\tilde{q}_t^\alpha$ can be normalized to form an orthogonal matrix because $S$ is symmetric.

Identified structural shocks can now be easily obtained from the data. Recall that $Q = [q^{[1]}, q_{\perp}^{[1]}]$ with $q_{\perp}^{[1]}$ such that $(q_\perp^{[1]})'q^{[1]} = 0$ and $(q_\perp^{[1]})'q_\perp^{[1]} = I$. Then use that $\epsilon_t = A^{-1}u_t = (A')^{-1}u_t$ and that $Q^{-1} = Q'$ to get

$$\epsilon_t = Q'A^{-1}u_t = \begin{bmatrix} (q^{[1]})' \\ (q_\perp^{[1]})' \end{bmatrix} A^{-1}u_t = \begin{bmatrix} \alpha_{11}'[I_k \ k'] (A')^{-1} \end{bmatrix} A^{-1}u_t \Rightarrow \epsilon_{11} = \alpha_{11}'[I_k \ k'] \Sigma^{-1}v_t$$

The next section considers how inference is affected when the covariance matrix underlying the identification scheme has to be estimated.

### 2.1.2 Posterior uncertainty

Here, I consider the case when the posterior over $\Gamma$ is non-degenerate. I abstract from potential weak instruments, which are discussed in, for example, Kleibergen and Zivot (2003) and surveyed in Lopes and Polson (2014).

Inference is analogous to inference in a SUR model (e.g. Rossi et al., 2005, ch. 3.5). In the special case in which the control variables for $Z_t$ coincide with the variables used in the VAR, the SUR model collapses to a standard scheme hierarchical Normal-Wishart posterior. Stack the vectorized model (2.4) as follows:

$$Y_{SUR} = X_{SUR}\beta_{SUR} + v_{SUR}, \quad v_{SUR} \sim N(0, V \otimes I_T), \quad (2.11)$$

\[\text{Note that, by construction, this gives orthogonal historical shocks: } \begin{bmatrix} \kappa \\ \kappa' \end{bmatrix} \Sigma^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} = \begin{bmatrix} \kappa \\ \kappa' \end{bmatrix} \Sigma^{-1} \begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11} = I_k \text{ from (2.8). The responses to other shocks are partially identified, up to } (m - k) \times (m - k) \text{ matrix } \lambda:\]

$$q_\perp^{[1]} = \tilde{A}' \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix} \lambda, \quad \lambda \lambda' = \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix} \Sigma \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix}^{-1},$$

where the restrictions on $\lambda$ follow from the orthogonality condition $(q_\perp^{[1]})'q_\perp^{[1]} = I_{m-k}$.
using the following definitions:

\[
V = \begin{bmatrix}
AA' & \Gamma' \\
\Gamma & \tilde{\Omega}
\end{bmatrix}
\]

\[
\beta_{SUR} = \begin{bmatrix}
\text{vec}(B) \\
\text{vec}(\mu_z)
\end{bmatrix}
\]

\[
\beta_{SUR} = \begin{bmatrix}
\text{vec}(B) \\
\text{vec}(\mu_z')
\end{bmatrix}
\]

\[
Y_{SUR} = \begin{bmatrix}
y_{1,1, \ldots, y_{1,T}}, \ldots, [y_{m_c,1, \ldots, y_{m_c,T}], [z_{1,1, \ldots, z_{1,T}], \ldots, [z_{m_z,1, \ldots, z_{m_z,T}]'}
\end{bmatrix}
\]

\[
X_{SUR} = \begin{bmatrix}
I_{m_y} \otimes X_y' & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z'
\end{bmatrix}
\]

\[
X_y = \begin{bmatrix}
Y_{-1} & \ldots & Y_{-p} & 1_T
\end{bmatrix}
\]

\[
X_z = \begin{bmatrix}
1_T
\end{bmatrix}
\]

\[
v_{SUR} = \begin{bmatrix}
v_{1,1, \ldots, v_{1,T}}, \ldots, [v_{m_c,1, \ldots, v_{m_c,T}], [v_{1,1, \ldots, v_{1,T}], \ldots, [v_{m_z,1, \ldots, v_{m_z,T}]'}
\end{bmatrix}
\]

The model is transformed to make the transformed errors independently normally distributed, taking advantage of the block-diagonal structure of the covariance matrix. Standard conditional Normal-Wishart posterior distributions arise from the transformed model. For the transformation it is convenient to define \( U \) as the Choleski decomposition of \( V \) such that

\[
\tilde{X} = ((U^{-1})' \otimes I_T) \begin{bmatrix}
I_{m_y} \otimes X_y' & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z'
\end{bmatrix} (V^{-1} \otimes I_T) \begin{bmatrix}
I_{m_y} \otimes X_y & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z
\end{bmatrix}
\]

\[
N_{XX}(V) = \tilde{X}' \tilde{X} = \begin{bmatrix}
I_{m_y} \otimes X_y' & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z'
\end{bmatrix} (V^{-1} \otimes I_T) \begin{bmatrix}
I_{m_y} \otimes X_y & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z'
\end{bmatrix}
\]

\[
N_{XY}(V) = \begin{bmatrix}
I_{m_y} \otimes X_y' & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes X_z'
\end{bmatrix} (V^{-1} \otimes I_T) \begin{bmatrix}
I_{m_y} \otimes Y & 0_{T(m_yp+1) \times Tm_z} \\
0_{Tm_z \times Tm_y} & I_{m_z} \otimes Z
\end{bmatrix}
\]

\[
S_T(\beta) = \frac{1}{\nu_0 + T} \begin{bmatrix}
(Y - XB)' \\
(Z - 1_T \mu_z')
\end{bmatrix} \begin{bmatrix}
(Y - XB) & (Z - 1_T \mu_z')
\end{bmatrix} + \frac{\nu_0}{\nu_0 + T} S_0.
\]

Given the above definitions, the following Lemma holds (Rossi et al., 2005, ch. 3.5):

**Lemma 2** The conditional likelihoods are, respectively, conditionally conjugate with
Normal and Wishart priors. The conditional posterior distributions are given by:

\[ \bar{\beta}_T(V) = (N_{XX}(V) + N_0)^{-1}(N_{XY}(V) + N_0\bar{\beta}_0) \]
\[ \beta|V,Y^T \sim \mathcal{N}(\bar{\beta}_T(V), (N_{XX}(V) + N_0)^{-1}) , \]
\[ V^{-1}|\beta,Y^T \sim \mathcal{W}(S_T(\bar{\beta})^{-1}/(\nu_0 + T), \nu_0 + T). \] (2.12a)

If \( X_z = X_y \), then \( \tilde{X} = I_{m_z+m_y} \otimes X_y \) and \( N_{XX} = (V^{-1} \otimes X_y'X_y) \) and analogous for \( N_{XY} \). In this special case, closed forms are available for the marginal distribution of \( V \), allowing to draw directly from the posterior.

I use the following algorithm for the Gibbs sampler:

1. Initialize \( V^{(0)} = S_T(\bar{\beta}_T) \).

2. Repeat for \( i = 1, \ldots n_G \):
   (a) Draw \( \beta^{(i)}|V^{(i-1)} \) from (2.12a).
   (b) Draw \( V^{(i)}|\beta^{(i)} \) from (2.12b).

To evaluate the joint posterior, I discard the first 1,000 draws and then keep every 20th draw. Figure 9 in the appendix compares the densities of the resulting log-likelihood in the first and the second half of the remaining sample from the posterior. Consistent with convergence of the sampler, the two distributions line up closely.\(^5\)

### 2.2 Simulated sampling properties

One criterion to judge the proposed prior and Bayesian inference scheme is to investigate its frequentist properties in a Monte Carlo study. I simulate 100 datasets based on the actual point estimates for the reduced-form VAR as well as the instrument-based inference about a structural impulse-response vector. I initialize the VAR at the zero vector and drop the first 100 observations and keep the last \( T = 236 \) observations as the basis for estimating the reduced-form VAR. Different scenarios

\(^5\)For computational purposes, I avoid computing the inverse in (2.12a) directly by using a QR-factorization. Since I only consider priors implemented using dummy observations and an otherwise flat prior: \( N_0 = 0 \) and \( N_T(V) = N_{XX}(V) \). Then factor \( \tilde{X} = Q_X R_X \), where \( Q_X \) is an orthogonal matrix. It follows that \( N_{XX}(V) = R_X' Q_X Q_X R_X = R_X' R_X \) and therefore \( N_{XX}(V)^{-1} = (R_X R_X)^{-1} = (R_X)^{-1} (R_X^')^{-1} \), so that \( R_X^{'-1} \mathcal{N}(\tilde{X}(R_X^')^{-1} N_{XY}(V), I) \) gives a draw from the posterior.
for instrument availability are considered for the structural inference: The fraction of missing observations varies from about 10% to 90% of the observations.

The data generating process is given by the maximum likelihood point estimate of a VAR in the core variables and its covariance with the Romer and Romer (2004) monetary policy shocks as the single observed instrument. I choose these shocks as a starting point for the analysis because they are available for about half (47%) of the sample. To vary the fraction of missing observations, I keep the point estimates constant, but vary the fraction of observations for the instrument set to zero.

For each dataset \( m \), I compute the pointwise posterior confidence interval using my Bayesian procedure, with and without conditioning on the observed covariance between instruments and VAR forecast errors. Similarly, I use the wild bootstrap proposed in Mertens and Ravn (2013) to conduct frequentist inference. Each procedure yields an estimate of the true IRF \( \{ I_{h,m} \}_{h=0}^H \) for each point as the maximum likelihood estimate and a pointwise confidence interval for horizons up to \( H \):

\[
\{ \hat{C}(\alpha)_{h,m} \}_{h=0}^H = \{ [c(\alpha)_{h,m}, \tilde{c}(\alpha)_{h,m}] \}_{h=0}^H.
\]

The superscript \( j \) indexes the different methods. \( \alpha \) is the nominal size of the confidence interval: A fraction \( \alpha/2 \) of draws from the posterior or the bootstrapped distribution for \( I_{h,m} \) lies under \( c(\alpha)_{h,m} \).

To assess the actual size of the confidence intervals of method \( j \), I compute for each \((h, m)\) whether the truth at horizon \( h \), \( I_{h,m} \) lies outside the pointwise confidence interval: \( 1 \{ I_{h,m} \notin \hat{C}(\alpha)_{h,m} \} \). The actual size is then estimated as:

\[
\hat{\alpha}^j_h = \frac{1}{100} \sum_{m=1}^{100} 1 \{ I_{h,m} \notin \hat{C}(\alpha)_{h,m} \}
\]

In Figure 1 below, I plot the deviation of the actual size for model \( j \) at horizon \( h \) from the nominal size: \( \hat{\alpha}^j_h - \alpha \) for \( \alpha = 0.32 \). The procedures that ignore “first stage” uncertainty about the covariance between the forecast errors and the external instruments understate the size of the confidence intervals substantially at short horizons (“Bayes – certain”, BaC and “Bootstrap – certain”, BoC), while the Bayesian procedure which allows for uncertainty about the covariance matrix errs on the conservative side (“Bayes–uncertain”, BaU). For the latter scheme, the actual size is typically only zero to five percentage points above the nominal size when half of the instruments are nonzero.\(^7\) However, when only one in ten observations

\(^6\)Appendix A.2 describes the algorithm.

\(^7\)The actual size depends slightly on the variable under consideration.
for the instrument is available, the actual size exceeds the nominal size by around ten percentage points. Overall, the Monte Carlo study may suggest that the proposed Bayesian procedure has good sampling properties and accounts properly for the uncertain covariance between instruments and forecast errors in the context of the present application.

2.3 Inferring the response of additional variables

In the empirical application to tax shocks, I face the challenge that only few (namely 13) non-zero observations on the tax shocks are available. Without additional (prior) information, inference on the effect of tax shocks becomes infeasible.\(^8\) When extending the analysis to include additional variables of interest, I therefore maintain the assumption in (2.4) that the variables in the above VAR are influenced by a number of shocks equal to the number of variables only. Following Uhlig (2003), I call the above VAR the “core” and the extra parameters the “periphery”.

In the language of Jarocinski and Mackowiak (2013), I am imposing Granger-causal priority of the core variables over the peripheral variables. This is akin to the forecasting of variables outside the DSGE model using an estimated DSGE model in Schorfheide et al. (2010). The key premise is that the core VAR allows to identify the structural shocks and the peripheral model then allows to trace the effect of these shocks on additional variables not needed for the shock identification.\(^9\) Implicitly, a similar assumption seems to be underlying the rotation of additional variables in and out of a smaller VAR in Ramey (2011).

Zha (1999) shows that such a block-recursive system can easily be implemented by including current values of the core variables as (exogenous) regressors to a separate periphery-VAR, for which standard inference applies:

\[
Y_{p,t} = B_{p,p} Y_{p,t-1} + B_{p,c} Y_{c,t-1} + B_{p,0} Y_{c,t} + A_{p,c} \epsilon_t.
\] (2.13)

\(^8\)Inference becomes infeasible insofar as the Bayesian confidence intervals blow up. This is consistent with the classical properties I find for my Bayesian estimator in the Monte Carlo study of section 2.2: As the number of instruments approaches the number of shocks, the actual size of the confidence intervals becomes increasingly larger than their nominal size.

\(^9\)If one were willing to impose a proper prior on the parameters of the model, one could use the elegant result in Jarocinski and Mackowiak (2013) to compute the posterior probability of the core-periphery model relative to an unrestricted model. However, the typical Minnesota prior used in Jarocinski and Mackowiak (2013) implies strong evidence for the more parsimonious model. However, the prior does affect structural responses significantly.
Figure 1: Monte Carlo analysis of actual minus nominal size $\hat{\alpha}^j_h - \alpha$
Conditional on current core variables, the peripheral variables in (2.13) are independent of the shocks in (2.4). Given $Y^T$, the posterior is therefore independent and parameters are drawn according to the following hierarchical procedure (e.g. Uhlig, 1994):

\[\Sigma_p^{-1} \equiv (A_p'(A_p')^{-1})^{-1} \sim W_{n\nu_0,\nu_T}(S_{p,T}^{-1},\nu_0 + \nu_T),\] (2.14a)

\[N_{p,T} = N_{p,0} + X'_pX_p,\]

\[S_{p,T} = \frac{1}{\nu_T}((\hat{B}_p - \bar{B}_{p,0})'N_{p,0}N_{p,T}^{-1}X'_pX_p(\hat{B}_p - \bar{B}_{p,0}) + \frac{\nu_0}{\nu_T}S_0 + \frac{\nu_T - \nu_0}{\nu_T}\hat{\Sigma}_p)
\]

\[B_p|\Sigma_p \sim N(B_{p,T}, \Sigma_p \otimes N_{p,T}^{-1}).\] (2.14b)

Here, $\hat{B}_p = (X'_pX_p)^{-1}X'_pY_p$, $\bar{B}_{p,T} = N_{p,T}^{-1}(N_{p,0}\bar{B}_{p,0} + X'_pX_p\hat{B}_p)$, $\hat{\Sigma}_p = T^{-1}Y'_p(I - X_p(X'_pX_p)^{-1}X'_p)Y_p$. In the computations I use the flat prior suggested by Uhlig (1994) with $\nu_0 = 0$, $N_0 = 0$.

As in Uhlig (2003), the response of peripheral variables to identified shocks is $B_{c,0}'\hat{A}q_{IV}$ on impact. In general, for the core and periphery, the model can be stacked to embody exclusion restrictions and extra lags to yield the response at time $h$ as

\[\begin{bmatrix} I_{m+m_p} & 0_{(m+m_p)\times(p-1)(m+m_p)} \end{bmatrix} B^h \begin{bmatrix} \hat{A}q_{IV} \\ 0_{(p-1)(m+m_p)\times m} \end{bmatrix}.\]

### 3 DSGE model estimation

In this section I outline the DSGE model used for comparison with the VAR. The model follows the well-known Smets and Wouters (2007) model with monopolistic competition in intermediate goods markets and the labor market and Calvo frictions to price and wage adjustment, partial price and wage indexation, and real frictions such as investment adjustment cost and habit formation. I ass labor, capital, and consumption taxes as in Drautzburg and Uhlig (2011) and fiscal rules as in Leeper et al. (2010). Here I only discuss the specification of fiscal and monetary policy. The remaining model equations are detailed in Section A.3 in the appendix.

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

\[\hat{r}_t = \rho_r\hat{r}_{t-1} + (1 - \rho_r)(\psi_{r,\pi}\hat{\pi}_t + \psi_{r,y}\tilde{y}_t + \psi_{r,\Delta y}\Delta\tilde{y}_t) + \epsilon'_t,\] (3.1)

where $\rho_r$ controls the degree of interest rate smoothing and $\psi_{r,x}$ denotes the reaction...
of the interest rate to deviations of variable \( x \) from its trend. \( \tilde{y} \) denotes the output gap, i.e. the deviation of output from output in a frictionless world.\(^{10}\)

The fiscal rules are:\(^{11}\)

\[
\hat{g}_t = -\psi_{g,y} \tilde{y}_t - \psi_{g,b} \tilde{b}_t + \xi_t^g \\
\hat{s}_t = -\psi_{s,y} \tilde{y}_t - \psi_{s,b} \tilde{b}_t + \xi_t^s \\
\frac{\tilde{w}n}{\tilde{y}} d\tau^n_t = \psi_{\tau^n,y} \tilde{y}_t + \psi_{\tau^n,b} \tilde{b}_t + \xi_t^{\tau,n} \\
\frac{(\tilde{r}^k - \delta) \tilde{k}}{\tilde{y}} d\tau^k_t = \psi_{\tau^k,y} \tilde{y}_t + \psi_{\tau^k,b} \tilde{b}_t + \xi_t^{\tau,k}
\]

The disturbances \( \xi_t^i \) follow exogenous AR(1) processes: \( \xi_t^i = \rho_i \xi_{t-1}^i + \epsilon_t^i \). Note that the sign of the coefficients in the expenditure components \( g_t, s_t \) are flipped so that positive estimates always imply consolidation in good times (\( \psi_{o,g} > 0 \)) or when debt is high (\( \psi_{o,b} > 0 \)).

The consolidated government budget constraint is:\(^{12}\)

\[
\frac{\tilde{b}}{\tilde{r}} (\hat{b}_t - \hat{r}) + \frac{\tilde{w}n}{\tilde{y}} (d\tau^n_t + \tilde{\tau}^n (\tilde{w}_t + \tilde{n}_t)) + \frac{\tilde{c}}{\tilde{y}} \tilde{\tau}^c \hat{c}_t + \frac{(\tilde{r}^k - \delta) \tilde{k}}{\tilde{y}} (d\tau^k_t + \tilde{\tau}^k (\tilde{r}_t^k \tilde{r}^k - \delta) + \tilde{k}_{t-1}^p) \\
= \hat{g}_t + \hat{s}_t + \frac{\tilde{b}}{\gamma \tilde{y}} (\hat{b}_{t-1} - \hat{\pi}_t)
\]

4 Results

4.1 Data and specifications

I use an updated version of the Smets and Wouters (2007) dataset for the estimation of both the DSGE and VAR model. I follow their variable definitions with two exceptions: I include consumer durables among the investment goods and use total hours worked from Francis and Ramey (2009) rather than private hours. They show that government hours differ in the early post-war period systematically from private hours worked. This can be particularly relevant when examining the effect

\(^{10}\)Money supply is assumed to adjust to implement the interest rate and fiscal transfers are adjusted to accommodate monetary policy.

\(^{11}\)Note that Leeper et al. (2010) assume there is no lag in the right hand side variables, while Fernandez-Villaverde et al. (2011) use a one quarter lag.

\(^{12}\)Seigniorage revenue for the government enters negatively in the lump-sum transfer to households \( \hat{s}_t \).
of fiscal spending on the economy.\textsuperscript{13} I also extend the sample period to 1947:Q1 to 2007:Q4, stopping before the ZLB became binding.\textsuperscript{14}

Sources for narrative shock measures are Romer and Romer (2004) for monetary policy shocks, the Survey of Professional Forecasters’ real defense spending forecast errors from Ramey (2011), and tax shock instruments are from Mertens and Ravn (2013). These are, broadly speaking, the subset of the shocks in Romer and Romer (2010) which are not considered to be motivated by economic conditions.

While it is arguably important to include a long sample covering WWII and the Korean War to have important instances of exogenous variation in government spending (cf. Ramey, 2011), I consider subsamples of the data before and after 1982 in robustness exercises.\textsuperscript{15}

I follow Leeper et al. (2010) in constructing time series for taxes, except for including state and local tax revenue in the calculation of revenue, tax, and debt data, similar to Fernandez-Villaverde et al. (2011). The significant share of state and local governments in both government debt and spending motivates this broader definition: Using Flow of Fund (FoF) data to initialize the time series of government debt at par value, I calculate debt levels using the cumulative net borrowing of all three levels of government. Municipal debt as a share of total government debt has increased from little over 5% in 1945 to around 35% in 1980 before falling to about 20% in 2012 according to FoF data.\textsuperscript{16} The share of state and local governments in government consumption and investment has been above 50% since 1972. Details are given in Appendix A.4.

\textsuperscript{13}I normalize all measures by the civilian non-institutional population as in Smets and Wouters (2007), but unlike Francis and Ramey (2009).

\textsuperscript{14}A significant downside of this strategy is that one period of significant fiscal changes during peacetime is omitted.

\textsuperscript{15}Bohn (1991) uses both annual data from 1792 to 1988 and quarterly data from 1954 to 1988. He argues that long samples are preferable because debt is often slow moving and because it contains more episodes such as wars which might otherwise appear special. However, he finds that the results for the shorter quarterly sample are qualitatively unchanged.

4.1.1 VAR specification

Six variables form the core VAR: Government spending, the average personal income tax rate, real per capita output and debt, the federal funds rate, and the inflation rate. The periphery consists of five variables: real per capita consumption, investment, government revenue, and total hours worked, as well as real wages.

A key input into formal and informal model selection criteria is the forecasting performance. Table 1 shows that, with the exception of inflation and hours, the peripheral variables improve forecasts very little: the R-square barely changes when they are included. For inflation, adding five more variables would increase the forecasting power from a little less than 74% to little less then 79%. Hours are hardest to predict without the peripheral variables and the R-squared would rise 25 percentage points. However, hours are a peripheral variable themselves in my baseline specification.\(^{17}\)

As instruments, I use real defense spending forecast errors from Ramey (2011) as instruments for government spending shocks, personal income tax rate shocks from Mertens and Ravn (2013) as instruments for tax shocks, and monetary policy surprises from Romer and Romer (2004) as instruments for monetary policy shocks. This leaves an extra three degrees of freedom, which here I resolve by imposing that monetary policy does not affect fiscal policy contemporaneously and tax shocks not affect spending contemporaneously.

Since instrument data is scarce, I follow Mertens and Ravn (2013) and set missing observations for instruments to zero – only five tax instrument observation overlap with the spending instrument. Setting missing observations to zero weakens the sample correlation from -0.5 to approximately zero so that the interpretation of spending and tax shocks may have to be made with caution.

In line with the benchmark DSGE model, I allow for a linear, deterministic trend in all variables (without imposing a common trend, however). I also consider

<table>
<thead>
<tr>
<th>Core + Periphery</th>
<th>Core</th>
<th>Periphery</th>
<th>Core only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>G    Tax</td>
<td>Y    Debt</td>
<td>FFR   Inf</td>
</tr>
<tr>
<td>Core only</td>
<td>99.86 95.20 99.95 99.98 95.44 78.53</td>
<td>99.76 99.99 99.74 98.65 99.97</td>
<td></td>
</tr>
<tr>
<td>Core only</td>
<td>99.81 94.46 99.94 99.97 94.85 73.60</td>
<td>99.58 99.96 99.22 73.68 99.72</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: R-squared of core model vs full model

\(^{17}\)Including hours instead of output among the core variables widens the confidence intervals somewhat and may switch implication inference about output and hours in response to tax shocks.
a quadratic trend and no trend, as in Ramey (2011) and Mertens and Ravn (2013), respectively.

4.1.2 DSGE specification

Currently, estimation of the DSGE model uses nine variables and does mirror the distinction between the core and the periphery, but allows all shocks to influence all variables. In a future revision, this will be addressed by introducing measurement error in the variables corresponding to the periphery, rather than interpreting every shock as structural. The variables used in the estimation are the six core variables except tax rates plus private consumption and investment, labor tax revenue, and real wage growth. Priors, when available, follow Smets and Wouters (2007). The exception is the labor supply elasticity, which I fix at unity prior to estimation.\footnote{Otherwise the estimation would imply an implausibly low labor supply elasticity.} I estimate the DSGE model using ?.

4.2 VAR results

Broadly speaking, the estimated output response to government spending, tax, and monetary policy shocks are unsurprising: Output rises in response to a government spending shock, and falls in response to positive tax shocks and unexpected monetary tightening. Other results are, however, worth noting. In response to monetary policy shocks, there is an apparent price puzzle as inflation rises and real economic activities starts falling only with a delay. There are significant interactions between the policy instruments: Government consumption falls significantly in response to contractionary monetary or tax policy shocks. Some private sector activity increases in response to government spending shocks, in particular private consumption. While the precise policy interactions depend on the identification scheme, the price puzzle and response of private activity appears to be a robust feature.

Is the estimated fiscal-monetary policy interaction plausible? A recent paper by Romer and Romer (2014) provides evidence that the Federal Reserve has indeed considered fiscal policy in its monetary policy decisions. Romer and Romer (2014) document staff presentations to the Federal Open Market Committee (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. The tax-inflation
The blue lines represent the median and 60% pointwise Bayesian confidence intervals when shocks are identified using narrative instruments. The green lines represent a standard Choleski identification ordering government spending first, taxes second, and the FFR fifth, after output and debt, but before inflation. The red lines represent the median and 60% pointwise Bayesian confidence interval for the estimated DSGE model.

Figure 2: VAR result comparison – aggregate / private
The blue lines represent the median and 60% pointwise Bayesian confidence intervals when shocks are identified using narrative instruments. The green lines represent a standard Choleski identification ordering government spending first, taxes second, and the FFR fifth, after output and debt, but before inflation. The red lines represent the median and 60% pointwise Bayesian confidence interval for the estimated DSGE model.

Figure 3: VAR result comparison – fiscal
The blue lines represent the median and 60% pointwise Bayesian confidence intervals when shocks are identified using narrative instruments. The green lines represent a standard Choleski identification ordering government spending first, taxes second, and the FFR fifth, after output and debt, but before inflation. The red lines represent the median and 60% pointwise Bayesian confidence interval for the estimated DSGE model.

Figure 4: VAR result comparison – prices
nexus is also reflected in staff presentations and in the comments by at least one FOMC: According to Romer and Romer (2014) the FRB staff saw social security tax increases in 1966, 1973, and 1981 as exerting inflationary pressure (p. 40). In general, Romer and Romer (2014) conclude that monetary policy did not counteract expansionary tax policy – but may have tightened policy in response to fiscal transfer programs (p. 41f.).

4.2.1 Comparison with Choleski-identification

How does the narrative identification differ from standard Choleski identification? First, by estimating the identifying covariance matrix rather than imposing restrictions, it produces substantially wider confidence intervals: Giving up the prior information about the contemporaneous effect of shocks comes at the price of a loss of precision.

Second, from a substantive point of view, the Choleski identification of the response to government spending shocks is largely consistent with the narrative identification. However, some significant differences arise for the response to tax shocks and monetary policy shocks: The Choleski-identification fails to capture the large contemporaneous output and investment drop in response to surprise tax increases. Similarly, the short-run output and investment responses to monetary policy shocks are, almost by assumption, significantly lower than implied by the narrative identification (Figure 2).

Interestingly, however, the price puzzle arises under both the Choleski and the narrative identification approaches. As Figure 3 shows, taxes are estimated to increase in responses to surprise monetary tightening under either identification approach and inflation, in Figure 4, mimics the pattern of taxes. The fiscal-monetary policy interaction is therefore not just a result of the particular approach to identification I take in this study.

4.2.2 First stage uncertainty

Figure 5 also illustrates, for the example of inflation responses to monetary policy shocks, the loss of precision that comes with estimating the covariance between forecast errors and instruments. The blue lines in the figure represent the pointwise 68% confidence interval coming from the baseline model, while the red line only accounts for uncertainty about the regression coefficients and neglects “first stage
uncertainty”, as discussed in Section 2.2. Particular in the very short run, ignoring first stage uncertainty shrinks the confidence interval by roughly two thirds. Intuitively, over longer horizons uncertainty about dynamics becomes more important and the ignoring first stage uncertainty matters less. These findings are in line with the simulated sampling properties of the different estimators in Figure 1.\(^{19}\)

### 4.2.3 Commodity prices and the price puzzle

Hanson (2004) reviews the literature on the price puzzle and finds that the various “fixes” such as including commodity prices in the VAR have no obvious justification in terms of increased forecast performance. Nevertheless, I include his preferred measure of commodity prices as an exogenous regressor in both the VAR regressors \(X_y\) and for the instrument regressors \(X_z\). Figure 5 shows the results. Comparing the right and the left panel reveals that the differences are minuscule.

The blue lines represent the median and 68% pointwise Bayesian confidence intervals when shocks are identified using narrative instruments and uncertainty about the covariance matrix is taken into account. The red line is the analogue object, but only allows for uncertainty in the regressors with regards to shock identification. The left column is the baseline VAR, while the right column includes commodity price inflation as an additional exogenous regressor.

**Figure 5:** Robustness of inflation response to monetary policy shocks

\(^{19}\)Figures 10 to 12 in the appendix show the complete set of responses.
4.2.4 Short vs medium-term identification

Using medium-run identification to identify the instrument-identified shocks separately has only a limited effect on the results. Figure 6 compares the response of the shocked variable themselves, output, and inflation to the three identified shocks under both identification schemes. Unsurprisingly, identifying the government spending shock by maximizing its medium-run variance contribution implies a more persistent effect on government spending. In line with larger wealth effects, output rises more when the effect on government spending is more pronounced. The responses to a monetary policy shock hardly change, while the median responses to tax shocks are weaker, but also more precisely estimated.

4.2.5 Theoretical variance decomposition

The three identified shocks together explain about 40 to 50% of the variance in government spending, taxes, and the federal funds rate, as the upper panel in Table 2 shows.\(^{20}\) Of the explained variance, the majority of the variation comes from the “own” shock, i.e. the tax shock explains more of the variation in taxes than the government spending or the monetary policy shock. However, there is also significant posterior uncertainty surrounding these posterior medians.

(a) Short-run (Choleski) factorization of \(\alpha_{11}'\alpha_{11}\)

(b) Medium-run factorization of \(\alpha_{11}'\alpha_{11}'\)

<table>
<thead>
<tr>
<th></th>
<th>G</th>
<th>tax</th>
<th>Y</th>
<th>Debt</th>
<th>FFR</th>
<th>Inf</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0.27 (0.13, 0.44)</td>
<td>0.08 (0.03, 0.16)</td>
<td>0.19 (0.08, 0.33)</td>
<td>0.11 (0.03, 0.29)</td>
<td>0.08 (0.02, 0.22)</td>
<td>0.20 (0.09, 0.35)</td>
</tr>
<tr>
<td>tax</td>
<td>0.13 (0.05, 0.27)</td>
<td>0.19 (0.08, 0.36)</td>
<td>0.14 (0.05, 0.27)</td>
<td>0.06 (0.01, 0.19)</td>
<td>0.06 (0.02, 0.19)</td>
<td>0.11 (0.04, 0.21)</td>
</tr>
<tr>
<td>FFR</td>
<td>0.09 (0.03, 0.20)</td>
<td>0.14 (0.06, 0.27)</td>
<td>0.10 (0.04, 0.21)</td>
<td>0.08 (0.03, 0.22)</td>
<td>0.34 (0.11, 0.53)</td>
<td>0.10 (0.04, 0.20)</td>
</tr>
</tbody>
</table>

Posterior median with 68% posterior confidence intervals in parenthesis.

Table 2: Theoretical unconditional variance decomposition for core variables

One obvious implication of these results is that specifications treating movements in government spending as exogenous miss important feedback mechanisms: While government spending shocks account for a large fraction of the variation in government spending, the posterior median leaves 73% of the variance unaccounted.

\(^{20}\)The remainder is attributable to the remaining shocks which are not separately identified.
The blue lines represent the median and 60% pointwise Bayesian confidence intervals when shocks are identified using narrative instruments and the benchmark short-run factorization. The red lines represent the counterpart with medium-run factorization.

Figure 6: Short-run vs medium run factorization
for. This lends support to the proposed fiscal extension of the canonical DSGE model in Section 3.

Policy interactions are also reflected in the theoretical variance decomposition: 14% of the variance in tax rates is attributed to the identified monetary policy shock, almost as much as by the tax shock itself. Tax shocks are half as important as government spending shocks themselves for variation in government spending. Fiscal shocks do, however, play only a limited role in explaining interest rate movements. When using the medium run-factorization (see the lower panel in Table 2) to identify shocks, however, the influence of monetary policy shocks on government spending and interest rates falls, but the effect of government spending shocks on interest rates rises from 8% to 16%.

Government spending is also identified as an important source of variation in inflation rates, independent of the factorization scheme. Tax shocks are less important for inflation than government spending shocks, but as important as monetary policy shocks.

4.3 Narrative DSGE model results

4.3.1 Impulse-responses

To compare the impulse-response functions coming from the DSGE model to those of the VAR, I consider again Figures 2 to 4. If the pointwise distribution of the DSGE responses differs only insignificantly from the pointwise distribution implied by the VAR I conclude that the models are consistent.

For tax shocks, the responses are mostly consistent across the two models: Output, consumption, and investment fall significantly. Compared to the VAR the DSGE model overstates, however, the decrease in debt and the increase in the tax rate required to generate an increase in revenue of one percent on impact. Unlike the VAR, it also predicts increased government spending for several years and a rise in real wages.

The estimated responses to government spending and monetary policy shocks differ substantially: In response to monetary policy shocks real activity falls only with a lag of several quarters in the VAR, but immediately in the DSGE model. Inflation and taxes rise slightly in the VAR estimates while both fall significantly according to the DSGE model. The previous discussion suggests that this may be
due to the response of taxes to monetary tightening in the VAR and their inflationary effect, which is absent in the estimated DSGE model.

While the expansion in output and hours worked is consistent across the DSGE and the VAR model, the responses of private activity and monetary policy differ significantly: The VAR model predicts a slow increase in private consumption and investment following a surprise increase in government spending, while the DSGE model predicts crowding out. Moreover, the VAR predicts accommodating monetary policy and falling inflation, while the DSGE model suggests the opposite. Taking the VAR estimates at face value, they may be interpreted as support for productive government investment as in Drautzburg and Uhlig (2011). Productive government investment may explain the lack of crowding out and falling prices, as it increases private sector productivity.

4.3.2 Historical shocks

Overall, both models have similar implications for the historical fiscal and monetary policy shocks. Figure 7 compares the time series of identified fiscal shocks and the monetary shock in the VAR to the identified shocks in the estimated DSGE model. The median of the time series is highly positively correlated, with correlation of 0.63 for tax shocks, 0.74 for government spending shocks and 0.75 for monetary policy shock.

A high correlation between the implied structural shocks is a necessary condition for consistency of the VAR results and DSGE results. However, implicit in the VAR analysis has been the assumption that the instruments are not related to non-policy shocks. I know assess this assumption.

5 Narrative shocks and structural shocks

For most of this paper, I have worked under the assumption that narrative shocks are noisy measures of a particular structural shock. This has allowed inference about the response to these structural shocks. However, what if narrative shock measures are better interpreted as reflecting several structural shocks? For example, the precursor to the monetary policy shock measure used in Romer and Romer (2010), namely Romer and Romer (1989) has been criticized in the discussion for not controlling for the endogeneity of shocks. For example, Dotsey and Reid (1992) argue that the Romer and Romer (1989) shocks reflect supply shocks.
Government spending shock – correlation of medians = 0.74

Tax shock – correlation of medians = 0.63

Monetary policy shock – correlation of medians = 0.75

Figure 7: Identified shocks comparison
To formally test whether the narrative shock measures I used correspond to the structural shocks they are assumed to correspond to, I append an extra set of measurement equations to the DSGE model. As in the VAR in (2.2), the instruments are interpreted as a noisy measure of an underlying structural shock. Here, however, I allow the narrative instrument to load on more than just one shock. Comparing the fit of the model with and without restrictions provides evidence for or against the assumption of a meaningful map between structural and narrative shocks.

The measurement equations linking the model to the structural shocks are as follows:

\begin{align}
  z_t^\tau &= c_{\tau} \varepsilon_t^\tau + \sum_{j \neq \tau} \varepsilon_t^j + \omega_{u}^\tau u_t^\tau, \\
  z_t^m &= c_{m} \varepsilon_t^m + \sum_{j \neq m} \varepsilon_t^j + \omega_{u}^m u_t^m, \\
  z_t^g &= c_{g} \mathbb{E}_{t-1} [\varepsilon_t^g - \varepsilon_{t-1}^g] + \sum_{j \neq g} \varepsilon_t^j + \omega_{u}^g u_t^g, \\
  z_t^a &= c_{a} (\varepsilon_t^a - \varepsilon_{t-1}^a) + \sum_{j \neq a} \varepsilon_t^j + \omega_{u}^a u_t^a,
\end{align}

where $\varepsilon_t^m$ represents the structural shock associated with instrument $m$ and $j \neq m$ indexes the remaining structural shocks in the model. $u_t^i \sim \mathcal{N}(0,1)$ is iid measurement error. $\xi_t^g$ is an exogenous AR(1) process with innovations given by $\varepsilon_t^g$. The superscripts $\tau, m, g$ represent, respectively, the tax shock, the monetary policy shock, and the government spending shock. In addition to the shocks used in the estimation above, I also include the quarterly measure of TFP growth by Fernald (2012), which I denote by $z_t^a$. The prior is that, with appropriate re-scaling, the coefficient on the structural shock itself, $c_{i}$, is centered around unity, while all other shock loadings $c_{j}$ are centered around zero.

Including the narrative shocks in the equation one at a time, I compute the Bayes factor in favor of the restricted model implicit in the VAR analysis vs. the unrestricted model. Table 3 shows the results. For the narrative policy instruments used in the VAR, the data provides more support for the restricted model consistent with the VAR assumptions than for the unrestricted model. In particular, based on the modified harmonic mean approximation to the posterior data density, the Bayes factor is between +2.3 and +4.1 log points.

In contrast, the data provides strong evidence against the interpretation of mea-
sured TFP growth as reflecting only structural TFP shocks in the model. Several shocks have a positive loading on TFP growth: the labor tax shock and the price markup shock both load significantly negative, while the wage markup shock, the government spending shock, and the investment specific shock load positively on TFP. If one trusts the model, this would provide evidence for mis-measurement of the TFP time series. Alternatively, trusting the measured series, one might want to force the model to fit the observed TFP series, possibly by including TFP as an observable (e.g. Schmitt-Grohe and Uribe, 2012) or by allowing for a more general variance-covariance structure of shocks as in Curdia and Reis (2010).

### Table 3: Assessing the restriction of narrative instruments to a single shock

<table>
<thead>
<tr>
<th>Posterior data density</th>
<th>Government Spending</th>
<th>Taxes</th>
<th>Monetary Policy</th>
<th>TFP growth</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>MC approximation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>restricted</td>
<td>-2037.8</td>
<td>-1951.0</td>
<td>-1903.0</td>
<td>-2587.9</td>
</tr>
<tr>
<td>unrestricted</td>
<td>-2040.6</td>
<td>-1953.3</td>
<td>-1907.1</td>
<td>-2574.6</td>
</tr>
<tr>
<td>difference</td>
<td>2.8</td>
<td>2.3</td>
<td>4.1</td>
<td>-13.3</td>
</tr>
<tr>
<td><strong>Laplace approximation</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>restricted</td>
<td>-2033.1</td>
<td>-1946.8</td>
<td>-1895.6</td>
<td>-2585.2</td>
</tr>
<tr>
<td>unrestricted</td>
<td>-2033.4</td>
<td>-1947.8</td>
<td>-1902.1</td>
<td>-2562.2</td>
</tr>
<tr>
<td>difference</td>
<td>0.3</td>
<td>1.0</td>
<td>6.5</td>
<td>-23.0</td>
</tr>
</tbody>
</table>

6 Conclusion

A key problem in macroeconomics is to translate forecast errors into structural economic shocks. This paper examines whether the inference on shocks from a standard fully structural DSGE model lines up with the conclusions based on methods with much weaker identifying assumptions, such as narrative methods. I find that VAR-identified government spending shocks, tax shocks, and monetary policy shocks line up well with those based on DSGE models. Also estimating the loadings of narrative shocks on structural shocks in a standard DSGE DSGE model augmented with simple fiscal rules lends empirical support to the assumption that narrative shocks only reflect the structural shock they are supposed to measure. The extended DSGE fails, however, to capture policy interactions and the response of private sector variables to government spending. A potential fix to this failure is the introduction of productive government capital as in Drautzburg and Uhlig (2011).

Methodologically, I show how to estimate a narrative VAR in a Bayesian framework. The proposed Bayesian estimator also captures the uncertainty about the instrument quality and has, with enough data, accurate frequentist sampling properties in my application.
Since narrative data are scarce, sample splits or extensions of the core VAR are infeasible in the present setup which uses a flat prior for the VAR. A natural next step is to elicit priors for the VAR from DSGE models, as in Del Negro and Schorfheide (2004). Adding assumptions about measurement error, the DSGE model readily implies a prior for the narrative VAR.
References


Francis, Neville and Valerie A. Ramey, “Measures of per Capita Hours and Their Implications for the Technology-Hours Debate,” Journal of Money, Credit and Banking, 09 2009, 41 (6), 1071–1097.


A Appendix

A.1 Narrative Shock identification

Note that (2.6a) does not restrict $\alpha_{11}$ for any value of $G$, assuming that $\Gamma$ is of maximal column rank $k \leq m$. The LHS depends on $k^2$ different parameters of $\alpha_{11}$, and $\Gamma_1$ has also $k^2$ elements.

$\Sigma$ provides an extra $m \frac{m+1}{2}$ equations, which can be used to solve for the remaining $k \leq m$ elements in $\alpha_{11}$ plus the $m \times (m - k)$ elements in $\alpha_2$. Thus, in general, an extra $\frac{m(m - 1)}{2} - k(m - k) \geq 0$ restrictions are needed.\(^{21}\) For $m = 2$ shocks and $k = 1$ instruments, the model is exactly identified.

Define $\kappa = (\Gamma_1^{-1}\Gamma_2)'$, (A.1)

so that $\alpha_{21} = \kappa\alpha_{11}$. Then:

$$
\Sigma = \begin{bmatrix}
\alpha_{11}\alpha'_{11} + \alpha_{12}\alpha'_{12} & \alpha_{11}\alpha'_{11}\kappa' + \alpha_{12}\alpha'_{22} \\
\kappa\alpha_{11}\alpha'_{11} + \alpha_{22}\alpha'_{12} & \kappa\alpha_{11}\alpha'_{11}\kappa' + \alpha_{22}\alpha'_{22}
\end{bmatrix}
$$

The covariance restriction identifies the impulse response (or component of the forecast error) up to a $k \times k$ square scale matrix $\alpha_{11}$:

$$
u_t = A\epsilon_t = \begin{bmatrix} \alpha[1] & \alpha[2] \end{bmatrix} \epsilon_t = \alpha[1]\epsilon_t[1] + \alpha[2]\epsilon_t[2] = \begin{bmatrix} I_k & \alpha_{11} \end{bmatrix} \epsilon_t + \begin{bmatrix} \alpha_{12} \\
\alpha_{22}
\end{bmatrix} \epsilon_t[2]
$$

Given that $\epsilon_t[1] \perp \epsilon_t[2]$ it follows that:

$$
\begin{align*}
\text{Var}[u_t|\epsilon_t[1]] &= \alpha[2](\alpha[2])' = \begin{bmatrix} \alpha_{12}\alpha'_{12} & \alpha_{12}\alpha'_{22} \\
\alpha_{22}\alpha'_{12} & \alpha_{22}\alpha'_{22}
\end{bmatrix} \\
\text{Var}[u_t|\epsilon_t[2]] &= \alpha[1](\alpha[1])' = \begin{bmatrix} \alpha_{11}\alpha'_{11} & \alpha_{11}\alpha_{11}\kappa' \\
\kappa\alpha_{11}\alpha'_{11} & \kappa\alpha_{11}\alpha'_{11}\kappa'
\end{bmatrix} \\
\Sigma &= \text{Var}[u_t] = \text{Var}[u_t|\epsilon_t[1]] + \text{Var}[u_t|\epsilon_t[2]] = \begin{bmatrix} \Sigma_{12}\Sigma'_{12} & \Sigma_{12}\Sigma'_{22} \\
\Sigma_{22}\Sigma'_{12} & \Sigma_{22}\Sigma'_{22}
\end{bmatrix}
\end{align*}
$$

Note that:

$$
\begin{bmatrix} I_k \\
\kappa
\end{bmatrix} \alpha_{11} \epsilon_t[1]
$$

\(^{21}\) $\frac{m(m-1)}{2} - k(m-k) = \frac{1}{2}((m - (k + 0.5))^2 + k^2 - (0.25 + k)) = \frac{1}{2}m(m - 2k - 1) + k^2 \geq 0$ for $m \geq 1, m \geq k \geq 0$. 

Any vector in the nullspace of \([I_k \ \kappa']\) satisfies the orthogonality condition. Note that \(\left\{ \begin{bmatrix} I_k \\ \kappa \\ \kappa' \\ -I_{m-k} \end{bmatrix} \right\}\) is an orthogonal basis for \(\mathbb{R}^m\).

Define

\[ Z \equiv [Z^{[1]} \ Z^{[2]}] = \begin{bmatrix} I_k & \kappa' \\ \kappa & -I_{m-k} \end{bmatrix} \quad (A.2) \]

Note that \(Z^{[2]}\) spans the Nullspace of \(\alpha^{[1]}\). Hence, \((Z^{[2]})'v_t\) projects \(v_t\) into the Nullspace of the instrument-identified shocks \(\epsilon_t^{[1]}\).

\[
(Z^{[2]})'v_t = (Z^{[2]})'A\epsilon_t = (Z^{[2]})'[\alpha^{[1]} \ \alpha^{[2]}] \epsilon_t = (Z^{[2]})'[Z^{[1]}|\tilde{Z}|\alpha_{11} \ \alpha^{[2]}] \epsilon_t = 0 \quad (Z^{[2]})'\alpha^{[2]} \perp \epsilon_t^{[1]}
\]

Note that \((Z^{[2]})'\alpha^{[2]}\) is of full rank and I can therefore equivalently consider \(\epsilon_t^{[2]}\) or \((Z^{[2]})'v_t\). Thus, the expectation of \(v_t\) given \(\epsilon_t^{[2]}\) is given by:

\[
\mathbb{E}[v_t|\epsilon_t^{[2]}] = \text{Cov}[v_t, (Z^{[2]})'v_t] \text{Var}[(Z^{[2]})'v_t]^{-1}(Z^{[2]})'(Z^{[2]})'v_t,
\]

\[
v_t - \mathbb{E}[v_t|\epsilon_t^{[2]}] = (I - \text{Cov}[v_t, (Z^{[2]})'v_t]) \text{Var}[(Z^{[2]})'v_t]^{-1}(Z^{[2]})'v_t,
\]

\[
\text{Cov}[v_t, (Z^{[2]})'v_t] = \Sigma Z^{[2]} = \Sigma \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix}
\]

\[
\text{Var}[v_t|\epsilon_t^{[2]}] = \mathbb{E}[(I - \text{Cov}[v_t, (Z^{[2]})'v_t]) \text{Var}[(Z^{[2]})'v_t]^{-1}(Z^{[2]})'v_t v_t']
\]

\[
= \mathbb{E}[v_t v_t'] - \text{Cov}[v_t, (Z^{[2]})'v_t] \text{Var}[(Z^{[2]})'v_t]^{-1} \mathbb{E}[(Z^{[2]})'v_t v_t']
\]

\[
= \Sigma - \text{Cov}[v_t, (Z^{[2]})'v_t] \text{Var}[(Z^{[2]})'v_t]^{-1} \text{Cov}[v_t, (Z^{[2]})'v_t]
\]

\[
= \Sigma - \Sigma \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix} \left( \begin{bmatrix} \kappa & -I_{m-k} \end{bmatrix} \Sigma \begin{bmatrix} \kappa' \\ -I_{m-k} \end{bmatrix} \right)^{-1} \begin{bmatrix} \kappa & -I_{m-k} \end{bmatrix} \Sigma
\]

\[
= \begin{bmatrix} \alpha_{11} \alpha_{11}' & \alpha_{11} \alpha_{11} \kappa' \\ \kappa \alpha_{11} \alpha_{11}' & \kappa \alpha_{11} \alpha_{11}' \kappa' \end{bmatrix} \quad (A.3)
\]

This gives a solution for \(\alpha_{11} \alpha_{11}'\) in terms of observables: \(\Sigma\) and \(\kappa = \Gamma^{-1}_1 \times \Gamma_2\).

In general, \(\alpha_{11}\) itself is unidentified: An additional \(\frac{(k-1)k}{2}\) restrictions are needed to pin down its \(k^2\) elements from the \(\frac{(k+1)k}{2}\) independent elements in \(\alpha_{11} \alpha_{11}'\). Given \(\alpha_{11}\), the impact-response to a unit shock is given by:

\[
\begin{bmatrix} I_k \\ \kappa \end{bmatrix} \alpha_{11}
\]

Note that this leaves \(m(m-k)\) elements in \(\alpha^{[2]}\) unrestricted, for which there exist
\[ \frac{m(m+1)}{2} \] equations in \( \Sigma - \text{Var}[v_t|\epsilon^1_t] \), requiring an extra \[ \frac{m}{2}(m - (1 + 2k)) \] restrictions. (Note: if \( k \geq \frac{1}{2}(m - 1) \), this implies over-identification.)

Intuitively: \( \alpha^{[1]} \) has \( km \) parameters. There are \( (m - k) \times k \) covariances with other structural shocks \( (\Sigma_{12}) \) which help to identify the model, leaving \( k^2 \) parameters undetermined. An extra \[ \frac{k(k+1)}{2} \] restrictions comes from the Riccati equation via the residual variance, leaving \( \frac{k(k-1)}{2} \) parameters to be determined. (This should also work when identifying one shock at a time with instruments.) The problem with multiple instrumented shocks is that even if I know that each instruments is only relevant for one specific structural shock, they are silent on the covariance between shocks – the covariance between shocks could be contained in the first stage residual.

When each instruments exactly identifies one shock, then each instrument delivers \( (n - 1) \) identifying restrictions: Each instrument then identifies one column of \( A \) up to scale. To see this, consider (2.5) for the case of \( G = \text{diag}(g_1, \ldots, g_k, 0, \ldots, 0) \) and let \( \alpha^{[i]} \) now denote the \( i \)'th column of \( A \):

\[
\text{Cov}[z_t^{[i]}, v_t] = e_i G A' = g_i e_i e_i' A' = g_i (\alpha^{[i]})'.
\]  

(A.4)

Since the LHS is observable, (??) identifies \( \alpha^{[i]} \) up to scale, imposing \( n-1 \) restrictions on \( A \) for each instrument. In general, this leads to overidentifying restrictions on \( \Sigma \).

The (normalized) identified shocks are given by \( Fu \), where \( F = \Gamma \Sigma^{-1} \).

### A.2 Frequentist inference

Following Mertens and Ravn (2013), the bootstrap procedure I consider is characterized as follows:

1. For \( t = 1, \ldots, T \), draw \( \{e^b_t, \ldots, e^b_T\} \), where \( e^b_t \sim \text{iid} \) with \( \text{Pr}(e^b_t = 1) = \text{Pr}(e^b_t = -1) = 0.5 \).

2. Construct the artificial data for \( Y^b_t \). In a VAR of lag length \( p \), build \( Y^b_t, t > p \) as:
   - For \( t = 1, \ldots, p \) set \( Y^b_t = Y_t \).
   - For \( t = p + 1, \ldots, T \) construct recursively \( Y^b_t = \sum_{j=1}^{p} \hat{B}_j Y^b_{t-j} + e^b_t \hat{u}_t \).

3. Construct the artificial data for the narrative instrument:
   - For \( t = 1, \ldots, T \) construct recursively \( z^b_t = e^b_t z^b_t \).
A.3 DSGE model equations

A.3.1 Households

The law of motion for capital:

\[ \dot{k}_t^p = (1 - \bar{x} k^p) \dot{k}_{t-1}^p + \bar{x} k^p (\dot{x}_t + \dot{q}_{t+s}) \]  

(A.5)

Household wage setting:

\[ \dot{w}_t = \frac{\dot{w}_{t-1}}{1 + \bar{\gamma}} + \frac{\bar{\beta} \gamma E_t[\dot{w}_{t+1}]}{1 + \bar{\beta} \gamma} + \frac{(1 - \bar{\beta} \zeta w)(1 - \zeta w)}{(1 + \bar{\beta} \gamma) \zeta_w} A_w \left( \frac{\dot{c}_t - (h/\gamma) \dot{c}_{t-1}}{1 - h/\gamma} + \nu \dot{n}_t - \ddot{w}_t + \frac{d \tau^p_t}{1 - \tau^p} + \frac{d \tau^c_t}{1 + \tau^c} \right) 
- \frac{1 + \bar{\beta} \mu w}{1 + \bar{\beta} \gamma} \dot{\pi}_t + \frac{\bar{\gamma}_c}{1 + \bar{\beta} \gamma} E_t[\dot{\pi}_{t+1}] + \frac{\dot{\epsilon}_t^{\lambda w}}{1 + \bar{\beta} \gamma} \]  

(A.6)

Household consumption Euler equation:

\[ E_t[\dot{\xi}_{t+1} - \dot{\xi}_t] + E_t[d \tau^p_{t+1} - d \tau^c_t] = \frac{1}{1 - h/\gamma} E_t(\sigma - 1) \frac{1}{1 + \bar{\lambda}_w} \frac{1 - \tau^p}{\bar{\epsilon}} [\dot{n}_{t+1} - \dot{n}_t] - \sigma \left[ \dot{c}_{t+1} - \left( \frac{1}{\gamma} \right) c_t + \frac{h}{\gamma} \dot{c}_{t+1} \right] \]  

(A.7)

Other FOC (before rescaling of \( \dot{q}_t^h \)):

\[ E_t[\dot{Q}_{t+1} - \dot{Q}_t] = -\dot{q}_t^h - \dot{R}_t + E_t[\dot{\pi}_{t+1}], \]  

(A.8)

\[ \dot{Q}_t = -\dot{q}_t^h - (\dot{R}_t - E_t[\pi_{t+1}]) + \frac{1}{\dot{r}^k(1 - \tau^k) + \delta \tau^k + 1 - \delta} \times \left( \dot{r}^k (1 - \tau^k) d \tau^k_{t+1} \right) + (1 - \delta) E_t(\dot{Q}_{t+1}) \]  

(A.9)

\[ \dot{x}_t = \frac{1}{1 + \bar{\beta} \gamma} \left[ \dot{x}_{t-1} + \bar{\beta} \gamma E_t(\dot{x}_{t+1}) + \frac{1}{\gamma^2 S^{\gamma}(\gamma)} (\dot{Q}_t + \dot{q}_t^e) \right], \]  

(A.10)

\[ \dot{u}_t = \frac{d'(1)}{a''(1)} \dot{\tau}_t = \frac{1 - \psi_u}{\psi_u} \dot{\tau}_t. \]  

(A.11)
### A.3.2 Production side and price setting

The linearized aggregate production function is:

\[
\hat{y}_t = \bar{y} + \Phi \left( \hat{\alpha}_t + \zeta \hat{k}_{t-1} + \alpha (1 - \zeta) \hat{n}_t + (1 - \alpha) (1 - \zeta) \hat{n}_t \right),
\]

where \( \Phi \) are fixed costs. Fixed costs, in steady state, equal the profits made by intermediate producers.

The capital-labor ratio:

\[
\hat{k}_t = \hat{n}_t + \hat{w}_t - \hat{r}_k.
\]

Price setting:

\[
\hat{\pi}_t = \iota p_1 + \iota \pi_1 \beta \gamma \hat{\pi}_t - 1 + (1 - \zeta) \bar{\beta} \gamma_{1 + \iota \bar{p}} E_t \hat{\pi}_{t+1} + \bar{\beta} \gamma_{1 + \iota \bar{p}} E_t \hat{\pi}_{t+1}.
\]

### A.3.3 Market clearing

Goods market clearing requires:

\[
\hat{y}_t = \bar{y} \hat{c}_t + \bar{x} \hat{x}_t + \bar{x} g \hat{x}_g + \hat{g}_t + \bar{r} \hat{k}_t - \hat{c}_t.
\]

### A.3.4 Observation equations

The observation equations are given by (??) as well as the following seven observation equations from Smets and Wouters (2007) and three additional equations (A.18) on fiscal variables:

\[
\Delta \ln \hat{y}_t = \hat{g}_t - \hat{g}_{t+1} + (\gamma_g - 1),
\]

\[
\Delta \ln \hat{x}_t = \hat{x}_t - \hat{x}_{t+1} + (\gamma_x - 1),
\]

\[
\Delta \ln \hat{w}_t = \hat{w}_t - \hat{w}_{t+1} + (\gamma_w - 1),
\]

\[
\Delta \ln \hat{c}_t = \hat{c}_t - \hat{c}_{t+1} + (\gamma - 1),
\]

\[
\hat{n}_t = \bar{n}_t + \hat{n}_t,
\]

\[
\iota \hat{R}_t = \bar{R}_t + (\beta - 1) - 1,
\]

By allowing for different trends in the non-stationary observables I treat the data symmetrically in the VAR and the DSGE model.
I use the deviation of debt to GDP and revenue to GDP, detrended prior to the estimation, as observables:

\[ b_{t}^{obs} = \frac{\bar{b}}{\bar{y}}(\hat{b} - \hat{y}) + \bar{b}^{obs} \]  
(A.18a)

\[ rev_{t}^{n,obs} = \frac{\bar{w}}{\bar{c}} \left( \frac{d\tau_{t}}{\tau_{t}} + \bar{w} + \bar{n} - \bar{y}_{t} \right) + \bar{rev}_{t}^{n,obs} \]  
(A.18b)

\[ rev_{t}^{k,obs} = \frac{\bar{k}}{\bar{y}} (\bar{r} - \delta) \left( \frac{d\tau_{t}}{\tau_{t}} + \frac{\bar{r}}{\bar{k}} \tau_{t} + \bar{k}_{t}^{p} + \bar{k}_{t-1} - \bar{y}_{t} \right) + \bar{rev}_{t}^{k,obs} \]  
(A.18c)

### A.4 Data construction

I follow Smets and Wouters (2007) in constructing the variables of the baseline model, except for allocating durable consumption goods to investment rather than consumption expenditure. Specifically:

\[ y_{t} = \frac{(\text{nominal GDP: NIPA Table 1.1.5Q, Line 1})_{t}}{(\text{Population above 16: FRED CNP16OV})_{t} \times (\text{GDP deflator: NIPA Table 1.1.9Q, Line 1})_{t}} \]

\[ c_{t} = \frac{(\text{nominal PCE on nondurables and services: NIPA Table 1.1.5Q, Lines 5+6})_{t}}{(\text{Population above 16: FRED CNP16OV})_{t} \times (\text{GDP deflator: NIPA Table 1.1.9Q, Line 1})_{t}} \]

\[ i_{t} = \frac{(\text{Durables PCE and fixed investment: NIPA Table 1.1.5Q, Lines 4 + 8})_{t}}{(\text{Population above 16: FRED CNP16OV})_{t} \times (\text{GDP deflator: NIPA Table 1.1.9Q, Line 1})_{t}} \]

\[ \pi_{t} = \Delta \ln(\text{GDP deflator: NIPA Table 1.1.9Q, Line 1})_{t} \]

\[ r_{t} = \begin{cases} \frac{1}{4} (\text{Effective Federal Funds Rate: FRED FEDFUNDS})_{t} & t \geq (1954:Q3) \\ \frac{1}{12} (\text{3-Month Treasury Bill: FRED TB3MS})_{t} & \text{(else.)} \end{cases} \]

\[ n_{t} = \frac{(\text{Nonfarm business hours worked: BLS PRS85006033})_{t}}{(\text{Population above 16: FRED CNP16OV})_{t}} \]

\[ w_{t} = \frac{(\text{Nonfarm business hourly compensation: BLS PRS85006103})_{t}}{(\text{GDP deflator: NIPA Table 1.1.9Q, Line 1})_{t}} \]

When using an alternative definition of hours worked from Francis and Ramey (2009), I compute:

\[ n_{t}^{FR} = \frac{(\text{Total hours worked: Francis and Ramey (2009)})_{t}}{(\text{Population above 16: FRED CNP16OV})_{t}} \]

Fiscal data is computed following Leeper et al. (2010), except for adding state and local governments (superscript “s&l”) to the federal government account (superscript “f”), similar to Fernandez-Villaverde et al. (2011). Since in the real world
\[
\tau_i^c = \frac{(\text{production & imports taxes: Table 3.2, Line 4})_t + (\text{Sales taxes})_{skl}^t}{((\text{Durables PCE})_t + c_t) \times (\text{GDP deflator})_t - (\text{production & imports taxes})_{skl}^t} - (\text{Sales taxes})_{skl}^t
\]

\[
\tau_i^p = \frac{(\text{Proprietors' income})_t + (\text{wage income})_t + (\text{wage supplements})_t + (\text{capital income})_t}{(\text{Personal current taxes})_t}
\]

\[
\tau_i^n = \frac{\tau_i^p \left(\frac{1}{2}(\text{Proprietors' income})_t + (\text{wage income})_t + (\text{wage supplements})_t\right) + (\text{wage taxes})_t^f}{(\text{wage income})_t + (\text{wage supplements})_t + (\text{wage taxes})_t^f + \frac{1}{2}(\text{Proprietors' income})_t}
\]

\[
\tau_i^k = \frac{\tau_i^p (\text{capital income})_t + (\text{corporate taxes})_t^f + (\text{corporate taxes})_{skl}^t}{(\text{Capital income})_t + (\text{Property taxes})_{skl}^t}
\]

where the following NIPA sources were used:

- (Federal) production & imports taxes: Table 3.2Q, Line 4
- (State and local) sales taxes: Table 3.3Q, Line 7
- (Federal) personal current taxes: Table 3.2Q, Line 3
- (State and local) personal current taxes: Table 3.3Q, Line 3
- (Federal) taxes on corporate income minus profits of Federal Reserve banks: Table 3.2Q, Line 7 – Line 8.
- (State and local) taxes on corporate income: Table 3.3Q, Line 10.
- (Federal) wage tax (employer contributions for government social insurance): Table 1.12Q, Line 8.
- Proprietors' income: Table 1.12Q, Line 9
- Wage income (wages and salaries): Table 1.12Q, Line 3.
- Wage supplements (employer contributions for employee pension and insurance): Table 1.12Q, Line 7.
- Capital income = sum of rental income of persons with CCAdj (Line 12), corporate profits (Line 13), net interest and miscellaneous payments (Line 18, all Table 1.12Q)
Note that the tax base for consumption taxes includes consumer durables, but to be consistent with the tax base in the model, the tax revenue is computed with the narrower tax base excluding consumer durables.

\[
(\text{rev})^c_t = \tau^c_t \times (c_t - (\text{Taxes on production and imports})^f_t - (\text{Sales taxes})^{s&l}_t) \times \left((\text{Population above 16})^t \times (\text{GDP deflator})^t\right)
\]

\[
(\text{rev})^n_t = \tau^n_t \times ((\text{wage income})^t + (\text{wage supplements})^t + (\text{wage taxes})^f_t + \frac{1}{2}(\text{Proprietors’ income})^t)\]

\[
(\text{rev})^k_t = \tau^k_t \times ((\text{Capital income})^t + (\text{Property taxes})^{s&l}_t)
\]

I construct government debt as the cumulative net borrowing of the consolidated NIPA government sector and adjust the level of debt to match the value of consolidated government FoF debt at par value in 1950:Q1. A minor complication arises as federal net purchases of nonproduced assets (NIPA Table 3.2Q, Line 43) is missing prior to 1959Q3. Since these purchases typically amount to less than 1% of federal government expenditures with a minimum of -1.1%, a maximum of 0.76%, and a median of 0.4% from 1959:Q3 to 1969:Q3, two alternative treatments of the missing data leads to virtually unchanged implications for government debt. First, I impute the data by imposing that the ratio of net purchases of nonproduced assets to the remaining federal expenditure is the same for all quarters from 1959:Q3 to 1969:Q4. Second, I treat the missing data as zero.

In 2012 the FoF data on long term municipal debt was revised up. The revision covers all quarters since 2004, but not before, implying a jump in the debt time series. I splice together a new smooth series from the data before and after 2004 by imposing that the growth of municipal debt from 2003:Q4 to 2004:Q1 was the same before and after the revision. This shifts up the municipal and consolidated debt levels prior to 2004. The revision in 2004 amounts to $840bn, or 6.8% of GDP.

The above data is combined with data from the web appendices of Romer and Romer (2004), Fernald (2012), Ramey (2011), and Mertens and Ravn (2013) on narrative shock measures. I standardize the different narrative shock measures to have unit standard deviation.

Figure 8 shows that there are significant differences in the total hours measure from Francis and Ramey (2009) and used in Ramey (2011), arising both from the hours time series itself (blue solid line vs. red dotted line) and from the difference between population measures (dashed vs. solid lines). These different measures of the same concept in the model imply different estimates of the fixed cost and persistence of government spending in the model (not shown) and consequently also in the implied IRFs. Figure ?? illustrates this. Using the Francis and Ramey (2009) hours measure with the Smets and Wouters (2007) population measure lowers the

---

22http://www.bondbuyer.com/issues/121_84/holders-municipal-debt-1039214-1.html

"Data Show Changes in Muni Buying Patterns" by Robert Slavin, 05/01/2012 (retrieved 01/24/2014).
Figure 8: Comparison of hours worked measures

prediction of the DSGE model for the response of hours to a government spending shock.
A summary measure of the multivariate posterior distribution is consistent with convergence of the Gibbs sampler: The plot shows the density estimate of \( \frac{\ell(\theta^{(i)}|\cdot) - \mu^G_1}{\sigma^G_1} \), where \( \ell \) is the log-Likelihood of the data and \( \theta^{(i)} \) a draw from the Gibbs sampler. \( \sigma^G_1 \) denotes the standard deviation of \( \ell(\theta^{(i)}) \) in the 1st half of the draws from the Gibbs sampler and \( \mu^G_1 \) the corresponding mean.

Figure 9: Gibbs-Sampler of baseline model: Normalized Distribution of Log-Likelihood
Figure 10: Core-Periphery VAR results – aggregate / private sector quantities
Figure 11: Core-Periphery VAR results – government quantities
Figure 12: Core-Periphery VAR results – prices
Note: – lin. trend 1

Figure 13: Core-Periphery VAR results – aggregate / private sector quantities – medium-run factorization
Figure 14: Core-Periphery VAR results – government quantities – medium-run factorization
Figure 15: Core-Periphery VAR results – prices – medium-run factorization