Changing impact of shocks: a time-varying proxy SVAR approach

Haroon Mumtaz* Katerina Petrova†

February 2019

Abstract

In this paper we extend the Bayesian Proxy VAR to incorporate time variation in the parameters. A Gibbs sampling algorithm is provided to approximate the posterior distributions of the model's parameters. Using the proposed algorithm, we estimate the time-varying effects of taxation shocks in the US and the UK and find evidence for a decline in the impact of this shock on output growth.

Key words: time-varying parameters, stochastic volatility, proxy VAR, tax shocks

JEL codes: C2,C11, E3

1 Introduction

Structural changes in macroeconomic dynamics and in the propagation of economic shocks have been documented in numerous recent studies. In their seminal paper, Cogley and Sargent (2005) introduce a vector autoregression (VAR) with time-varying parameters and stochastic volatility (TVP-VAR) and provide evidence supporting shifts in the persistence and volatility of key US macroeconomic variables. Their model was extended by Primiceri (2005) who also allowed the

*Queen Mary University London. Email: H.Mumtaz@qmul.ac.uk

†University of St. Andrews, Email: Katerina.Petrova@st-andrews.ac.uk

contemporaneous coefficients of the VAR model to change over time thus allowing these models to be used for structural analysis. For example, using this model Benati and Mumtaz (2007) identify demand and supply shocks via sign restrictions to investigate the causes of the 'Great Moderation' in the US. Similarly, Baumeister and Peersman (2013) identify oil supply and demand shocks using sign restrictions and show that the price elasticity of oil demand in the US has declined over time. Gali and Gambetti (2009) employ TVP-VARs with long-run restrictions to investigate changes in US macroeconomic dynamics. Canova and Forero (2015) provide a general algorithm to estimate TVP-VARs when the shocks are identified via non-recursive identification schemes.

In parallel to this literature on TVP-VARs, the methods used for shock identification in VAR models have also seen rapid development. An approach that has gained popularity in recent empirical applications is the identification of shocks by using external instruments. This 'proxy SVAR' model was introduced by Mertens and Ravn (2013) and Stock and Watson (2008). This approach differs from standard identification methods because the contemporaneous impulse response is estimated using an instrumental variable procedure where the instrument is an exogenous proxy for the shock of interest, usually constructed via a narrative approach. This reliance on external information reduces the number of (possibly controversial) restrictions needed to identify the contemporaneous impulse response. As the proxy is used as an instrument and not an endogenous variable directly in the VAR, the effects of measurement error in the proxy can be alleviated. This combination of SVAR methods with a narrative approach to estimating causal effects is the key reason behind the increased popularity of proxy SVAR models in applied work.

In this paper, we propose an algorithm for the estimation of proxy SVAR models where the parameters of the model are allowed to vary over time. Therefore, we extend the range of methods for time-varying SVARs described in Canova and Forero (2015). We provide a Gibbs sampling algorithm to approximate the posterior distributions of the parameters. Our estimation procedure extends the Bayesian MCMC algorithm of Caldara and Herbst (2019) for fixed coefficient proxy SVAR models, by casting it in state-space form, and applying a 'state of the art' particle Gibbs algorithm to filter through the parameter time variation in the resulting nonlinear state space. We design a small simulation exercise to study the finite sample properties of our proposed algorithm showing that it displays a satisfactory performance.

Using the proposed model, we investigate the impact of tax shocks on US and UK GDP growth and consider whether these effects have changed over time. To identify the tax shocks, we use the narrative measures proposed by Mertens and Ravn (2012) and Cloyne (2013) as instruments for the US and the UK, respectively. Our results suggest that the response of US GDP to tax shocks declined sharply during the 1960s and has remained more stable, thereafter. In contrast, the response of UK GDP to this shock has declined persistently over the sample period and the tax shock has become less important in terms of contribution to the forecast error variance.

The remainder of the paper is organised as follows. Section 2 presents the proxy SVAR model with time-varying parameters. Section 3 describes the Gibbs algorithm and provides a small Monte Carlo exercise to evaluate the performance of the proposed algorithm. Finally, Section 4 contains our empirical analysis, and Section 5 includes some concluding remarks.

2 The time-varying proxy SVAR

We consider the following Gaussian VAR model with time-varying parameters:

$$Y_t = B_t X_t + u_t, \; u_t = \Sigma_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I_N) \tag{1}$$

where Y_t is a $N \times 1$ vector of endogenous variables, $X_t = [Y_{t-1}', .., Y_{t-P}', 1]'$ is $(NP+1) \times 1$ vector of regressors in each equation and B_t denotes the $N \times (NP+1)$ matrix of coefficients

$B_t = [B_{1t}, ..., B_{Pt}, c_t]$. This VAR model features time-varying autoregressive coefficients B_t, where we follow the literature in assuming that the evolution of the $N(NP+1) \times 1$ vector $b_t := vec(B_t')$ is described through a random walk transition equation:

$$b_t = b_{t-1} + Q_b^{1/2} \eta_t^b, \quad \eta_t^b \sim \mathcal{N}\left(0, I_{N(NP+1)}\right), \tag{2}$$

The time-varying covariance matrix of the reduced form residuals u_t can be written as:

$$\Sigma_t = (A_t q)(A_t q)' \tag{3}$$

where A_t is a lower triangular matrix with time-varying elements, and q is an element of the family of orthogonal matrices of size N, satisfying $q'q = I_N$. By considering all possible values of q, the matrix $A_t q$ spans the space of all possible contemporaneous matrices; a result which follows from the QR decomposition of a square matrix.

Next, we decompose A_t as:

$$A_t = \tilde{A}_t H_t^{1/2} \tag{4}$$

where $\tilde{A}_t$ is a lower triangular matrix with diagonal elements equal to one and H_t is diagonal. Following Primiceri (2005), we model the evolution of the unrestricted elements of $\tilde{A}_t$ (where we donote by the vector a_t the $N(N-1)/2 \times 1$ elements below the subdiagonal of $\tilde{A}_t$) as random walks:

$$\alpha_t = \alpha_{t-1} + Q_a^{1/2} \eta_t^a, \quad \eta_t^a \sim \mathcal{N}\left(0, I_{N(N-1)/2}\right). \tag{5}$$

Finally, we assume that the diagonal elements of H_t (summarised in an $N \times 1$ vector h_t) evolve as geometric random walks:

$$\ln h_t = \ln h_{t-1} + Q_h^{1/2} \eta_t^h, \quad \eta_t^h \sim \mathcal{N}(0, I_N). \tag{6}$$

2.1 Identification of shocks

The structural shocks of the VAR model ε_t are defined as

$$\varepsilon_t = A_{0,t}^{-1} u_t, \tag{7}$$

where $A_{0,t} = A_t q$. We assume, without loss of generality, that we are interested in identifying the first shock ε_{1t} in the $N \times 1$ vector of shocks $\varepsilon_t = [\varepsilon_{1t}, \varepsilon_{\cdot t}]$, where $\varepsilon_{\cdot t}$ contains the remaining $N-1$ elements in ε_t. To do this, we employ an instrument m_t described by the following equation:

$$m_t = \beta \varepsilon_{1t} + \sigma v_t, \quad v_t \sim \mathcal{N}(0, 1) \tag{8}$$

where $\mathbb{E}(v_t \varepsilon_t) = 0$. The instrument is assumed to be relevant ($\beta \neq 0$) and uncorrelated with other structural shocks ($\mathbb{E}(m_t \varepsilon_{\cdot t}) = 0$). As discussed in Caldara and Herbst (2019), the relevance of the instrument can be judged by calculating the reliability statistic of Mertens and Ravn (2013) which is defined as the squared correlation between m_t and ε_{1t}:

$$\rho^2 = \frac{\beta^2}{\beta^2 + \sigma^2} \tag{9}$$

While we write equation 8 as a fixed coefficient model in the benchmark case, our procedure easily allows the extension to a time-varying β and σ^2. For example one can assume that:

$$\beta_t = \beta_{t-1} + \sigma^{\beta} n_t^{\beta}, \quad n_t^{\beta} \sim \mathcal{N}(0,1) \tag{10}$$

$$\ln \sigma_t^2 = \ln \sigma_{t-1}^2 + \sigma^{\sigma} n_t^{\sigma}, \quad n_t^{\sigma} \sim \mathcal{N}(0,1) \tag{11}$$

This extended formulation would also allow the resulting reliability statistic to change over time.

2.1.1 The likelihood function and the role of the instrument

The covariance between the reduced form residuals and the instrument can be defined by:

$$\begin{pmatrix} u_t \\ m_t \end{pmatrix} | L_t \sim \mathcal{N}\left(0, L_t L_t'\right), \quad L_t = \begin{pmatrix} A_t q & 0 \\ \bar{b} & \sigma \end{pmatrix} \tag{12}$$

where $\bar{b}$ is a $1 \times N$ vector $\bar{b} = \begin{bmatrix} \beta & 0_{1\times(N-1)} \end{bmatrix}$, since

$$\begin{pmatrix} u_t \\ m_t \end{pmatrix} = L_t \begin{bmatrix} \varepsilon_t \\ v_t \end{bmatrix}.$$

To consider the role of the instrument we follow Caldara and Herbst (2019) and factor the likelihood of the model as:

$$p(Y_t, m_t|\Xi) = p(Y_t|\Xi)\, p(m_t|Y_t, \Xi) \tag{13}$$

where Ξ denotes all parameters and states of the model. Given the conditional normality assumption in equation 12, the conditional density $p(m_t|Y_t, \Xi)$ is also normal with mean $\mu_t = \beta q_1' A_t^{-1} u_t$ and variance $s = \sigma^2$, where q_1 is the first column of q.

As discussed in Caldara and Herbst (2019), μ_t can be interpreted as a linear combination of the orthogonalised residuals $A_t^{-1}u_t$. In the context of estimation, this means that draws from the posterior distribution that result in the difference between the proxy and this linear combination becoming smaller are given larger weight.

The key contribution of the current paper relative to Caldara and Herbst (2019) is that we embed the dependence on m_t within a state-space model. In our model, time-variation in the reduced-form of the model is enabled through A_t and B_t. However, the draws of q_1 ensure that the contemporaneous impact matrix $A_{0,t}$ accounts for the conditional likelihood of m_t.

The proposed model is related to recent contribution by Paul (2017) who incorporates proxies as exogenous variables in a TVP-VAR. He shows that this VARX approach leads to a consistent estimator of the relative or normalised impulse response. While the approach in Paul (2017) is attractive due to its simplicity, our model offers two advantages. First, since we use the proxy as an instrument, we can estimate reliability statistics and provide evidence on instrument relevance. Second, as we describe below, our procedure can easily accommodate missing values in the instrument series, and thus deal with an issue that is common in the existing literature.

It is also interesting to note that the literature on fixed coefficient Bayesian proxy SVARs has also considered alternative specifications for the model. In particular, Drautzburg (2016) and Rogers *et al.* (2016) link the instrument to the reduced form residuals and use $cov\left(u_t, m_t\right)$ to back out the implied normalised impulse vector. However, in a time-varying parameter model, this formulation can be problematic as it is difficult to separate changes in the transmission mechanism and possible shifts in instrument reliability. Therefore, we use the specification proposed in Caldara and Herbst (2019) as our starting point.

3 Estimation

The model is estimated using a Metropolis-within-Gibbs algorithm. In this section we describe the priors and provide a sketch of the algorithm. The detailed description of the conditional posteriors can be found in the technical Appendix.

3.1 Priors and starting values

Following Cogley and Sargent (2005) and Primiceri (2005), we set the prior for Q_b using a training of sample of T_0 observations. Denote the OLS estimates of the VAR coefficients and coefficient covariance as B_{OLS} and V_{OLS}. The prior for Q_b is inverse-Wishart $\mathcal{IW}\left(T_0 \times V_{OLS} \times \kappa_b, T_0\right)$ where the scaling parameter is $\kappa_b = 3.5 \times 10^{-4}$ in our empirical analysis below. We set the initial value $b_{0|0} \sim \mathcal{N}(B_{OLS}, V_{OLS})$. The prior for Q_a is inverse Wishart: $\mathcal{IW}\left(I_{N(N-1)/2} \times \kappa_a, T_{a,0}\right)$ where, in our empirical analysis, $\kappa_a = 1 \times 10^{-4}$. The prior for Q_h is also inverse Wishart: $\mathcal{IW}\left(I_N \times \kappa_h, T_{h,0}\right)$ where $\kappa_h = 1 \times 10^{-4}$. To obtain an initial draw for these state variables we run the algorithm for a standard TVP-VAR for a limited number of iterations and use the last draw as the initial value to be input into the MCMC algorithm described below.

Following Caldara and Herbst (2019), the prior for q is uniform and as described in Rubio-Ramirez *et al.* (2010) can be sampled from by taking the QR decomposition of a $N \times N$ matrix from the standard normal distribution. When fixed parameters for equation 8 are used, we employ a standard Normal-inverse Gamma prior: $p\left(\beta\right) \sim \mathcal{N}(\beta_0, V_\beta)$ and $p\left(\sigma^2\right) \tilde{} \mathcal{IG}^*\left(\sigma_0, v_0\right)$ where $\mathcal{IG}^*$ is an inverse gamma density, re-parameterised in terms of the mean σ_0 and variance v_0.

3.2 Gibbs sampling algorithm

The Gibbs sampling algorithm cycles through the following conditional posterior distributions:

Step 1. $p\left(A_t | \Xi_{-A_t}, Y_{1:T}, m_{1:T}\right)$. Note that Ξ_{-A_t} denotes the set of all parameters other than A_t.

Conditional on Ξ_{-A_t}, the state-space model can be written as:

$$\begin{aligned} Y_t &= B_t X_t + u_t \text{ observation} \\ u_t &= A_t q \varepsilon_t \text{ observation} \\ m_t &= \beta \varepsilon_{1t} + \sigma v_t \text{ observation} \\ \alpha_t &= \alpha_{t-1} + Q_a^{1/2} n_t^a \text{ transition} \\ \ln h_t &= \ln h_{t-1} + Q_h^{1/2} n_t^h \text{ transition} \end{aligned} \tag{14}$$

The state-space in (14) is nonstandard due to non-linearity of the first observation equation and because of the relationship between the instrument and ε_{1t}. Therefore, to sample $[\alpha_t', \ln h_t']'$ we employ a particle-Gibbs sampler. In a seminal contribution, Andrieu *et al.* (2010) show how a version of the particle filter, conditioned on a fixed trajectory for one of the particles can be used to produce draws that result in a Markov kernel with a target distribution that is invariant. We employ a version of the sampler introduced in Lindsten *et al.* (2014) who propose the addition of a step that involves sampling the 'ancestors' or indices associated with the particle that is been conditioned on. They show that this leads to a considerable improvement in the mixing of the algorithm even with a few particles.

Step 2. $p\left(b_t|\Xi_{-b_t}, Y_{1:T}, m_{1:T}\right)$. Given Ξ_{-b_t} containing all parameters except b_t, the state-space of the model can be written as:

$$\begin{pmatrix} Y_t \\ m_t \end{pmatrix} = \begin{pmatrix} \left(I_N \otimes X_t'\right) b_t \\ 0 \end{pmatrix} + \begin{pmatrix} u_t \\ m_t \end{pmatrix} \text{ observation}$$

$$b_t = b_{t-1} + Q_b \eta_t^b \text{ transition}$$

where the conditional covariance matrix of the observation equation residuals is:

$$cov\left(\begin{array}{c} u_t \\ m_t \end{array}\middle| \Xi_{-b_t}\right) = \begin{pmatrix} A_t A_t' & A_t q_1' \beta \\ \beta q_1 A_t' & \beta^2 + \sigma^2 \end{pmatrix}$$

This system is conditionally linear and Gaussian and we can use the Carter and Kohn (1994) algorithm to draw b_t from its posterior distribution.

Step 3. $p\left(q_1|\Xi_{-q_1}, Y_{1:T}, m_{1:T}\right)$. Following Caldara and Herbst (2019), we use an independence Metropolis step to sample q_1.

Step 4. $p\left(\beta, \sigma|\Xi_{-[\beta,\sigma]}, Y_{1:T}, m_{1:T}\right)$. The structural shock of interest ε_{1t} can be calculated as $\varepsilon_{1t} = A_t q_1 u_t$. First draw $p\left(\sigma^2|\Xi_{-[\beta,\sigma]}, Y_{1:T}, m_{1:T}\right)$. Assuming an inverse-Gamma prior, this conditional posterior is also inverse-Gamma and can be easily sampled from. If σ^2 is allowed to vary over time, the a Metropolis algorithm or a particle Gibbs step can be employed to sample from the conditional posterior. Moreover, conditioning on σ^2 equation 8 is a standard linear regression and conditional posterior which is Gaussian: $p\left(\beta|\Xi_{-[\beta,\sigma]}, \sigma, Y_{1:T}, m_{1:T}\right) \sim \mathcal{N}(\tilde{\beta}, \tilde{V})$. Note that if β were allowed to be time-varying, the Carter and Kohn (1994) algorithm can be used to draw from its conditional posterior.

Step 5. Draw from $p\left(Q_b|\Xi_{-Q_b}, Y_{1:T}, m_{1:T}\right), p\left(Q_a|\Xi_{-Qa}, Y_{1:T}, m_{1:T}\right)$ and $p\left(Q_h|\Xi_{-Qh}, Y_{1:T}, m_{1:T}\right)$. Assuming an inverse-Wishart prior, these conditional posterior distributions are also inverse-Wishart and these draws are standard.

Step 6. $p\left(m_{-t}|\Xi_{m_{-t}}, Y_{1:T}, m_{1:T}\right)$. This step is implemented if the instrument contains missing observations (denoted by m_{-t}). As in step 2, the model can then be written in the form of a conditionally linear, Gaussian state-space model with m_{-t} treated as a latent state. The Carter and Kohn (1994) algorithm can be used to draw the conditional posterior distribution.

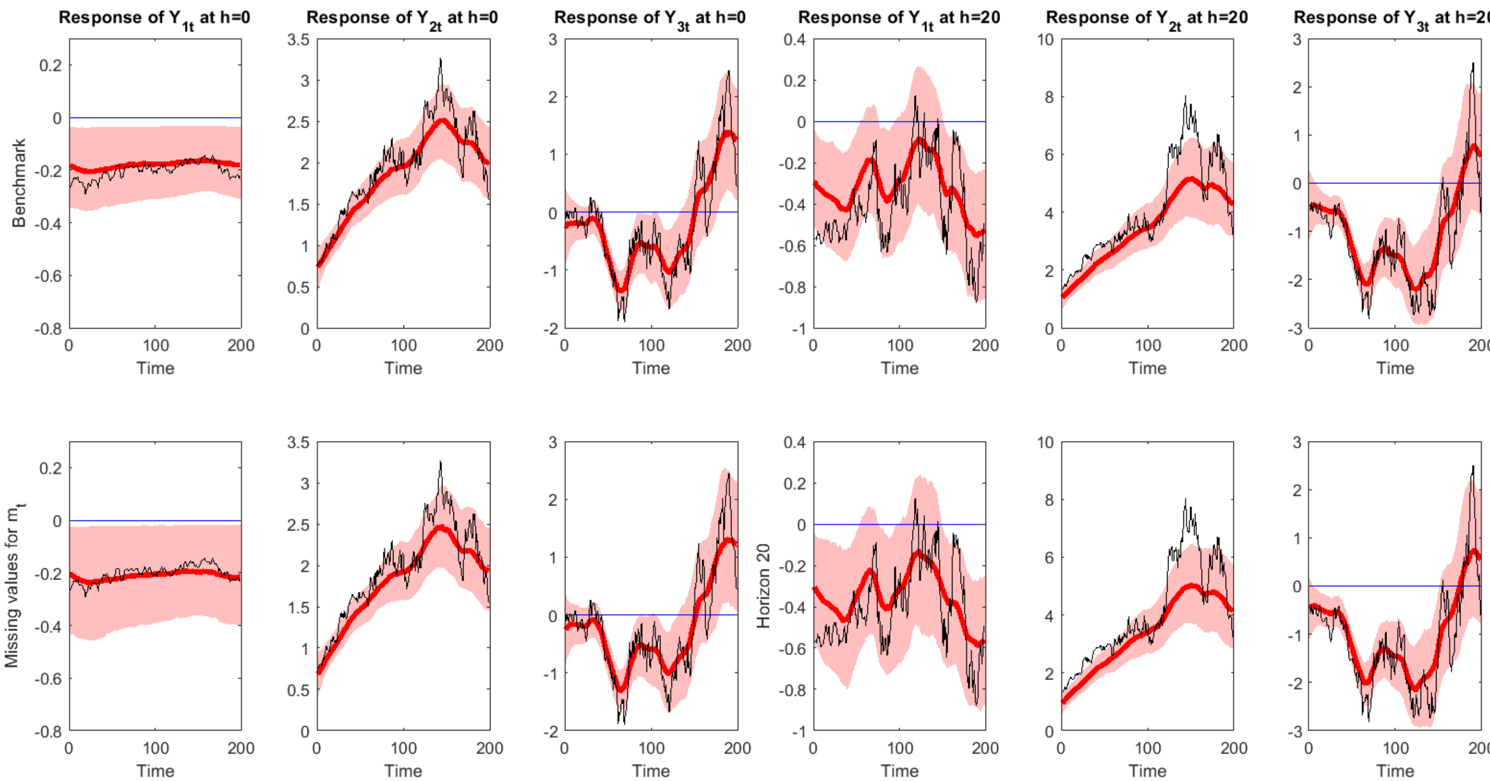

Figure 1: Impulse response to a 1 standard deviation shock to the first equation. The thick (red) line and shaded area represents the median and 1 standard deviation error band while the thin (black) line shows the true response. The impulse response horizon is denoted by h.

3.3 Estimation using simulated data

To test the algorithm we conduct a simple simulation experiment. We generate data from the following data generating process (DGP):

$$Y_t = B_t Y_{t-1} + (A_t q)\,\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I_3) \tag{15}$$

where Y_t is 3×1 vector, $q'q = I_3$ and $A_t = \tilde{A}_t H_t^{1/2}$. The time-varying parameters evolve as:

$$\begin{aligned}
vec\left(B_t'\right) &= vec\left(B_{t-1}'\right) + Q_b^{1/2}\eta_t^b,\ Q_b = I_{N^2} \times 0.0001,\ \eta_t^b \sim \mathcal{N}(0, I_{N^2}) \\
\alpha_t &= \alpha_{t-1} + Q_a^{1/2}\eta_t^a,\ Q_a = I_{N(N-1)/2} \times 0.01,\ \eta_t^a \sim \mathcal{N}(0, I_{N(N-1)/2}) \\
\ln h_t &= \ln h_{t-1} + Q_h^{1/2}\eta_t^h,\ Q^h = I_N \times 0.01,\ \eta_t^h \sim \mathcal{N}(0, I_N)
\end{aligned}$$

where α_t denote the non-zero and non-one elements of the lower triangular matrix $\tilde{A}_t$ and h_t is a vector containing the diagonal elements of the diagonal matrix H_t. The instrument is generated via:

$$m_t = 0.2\varepsilon_{1t} + 0.1^{1/2} v_t, \quad v_t \sim \mathcal{N}(0, 1)$$

where ε_{1t} is the first element of ε_t, the shock of interest. We generate 320 observations, from which we discard the first 100. We use 20 as a training sample, leaving 200 observations for estimation. We consider two cases for m_t. In the benchmark case, it is assumed that there are no missing observations for the instrument. We consider an alternative scenario where the first 80 observations for m_t are assumed to be missing.

The estimation of the model uses 5000 Gibbs iterations with a burn-in of 3000 iterations. The particle Gibbs step uses 10 particles. We repeat the simulation experiment 500 times with the state-variables kept constant in each iteration.

Figure 1 presents the true impulse response of all three variables to the structural shock ε_{1t} at horizon 0 and 20 and the estimated values. Consider the top panel which displays the results from the benchmark experiment. Although the estimated time-varying response is smoother than the true response, the estimates tracks the main structural shifts fairly well. The bottom panel of the figure shows that this conclusion does not change substantially when the DGP incorporates missing values in the instrument.

Overall, these simulation results provide some evidence that the MCMC algorithm proposed in this paper displays a satisfactory finite-sample performance.

4 The time-varying impact of taxation shocks in the United States and the United Kingdom

The impact of tax shocks on output has received considerable attention in the recent macroeconomic literature. As pointed out in Mertens and Ravn (2014), estimates of the US tax multiplier differ substantially with studies using narrative measures of tax shocks reporting estimates that are larger than those obtained using SVARs with zero restrictions. Mertens and Ravn (2014) propose a proxy SVAR to estimate the impact of taxation shocks. Their instrument is a refined version of the narrative tax measure of Romer and Romer (2010), the construction of which is described in Mertens and Ravn (2012). Romer and Romer (2010) build their shock measure by purging legislated tax changes from movements that are endogenous and driven by policy makers' concerns about growth. However, Mertens and Ravn (2012) argue that the Romer and Romer (2010) tax shock may not satisfy exogeneity as the proxy does not account for implementation lags. Instead Mertens and Ravn (2012) propose a proxy based on exogenous tax changes where legislation and implementation are less than a quarter apart. Using this measure, Mertens and Ravn (2014) estimate tax multipliers that lie towards the upper end of the range of estimates.

For the UK, Cloyne (2013) uses official budget sources to obtain data on tax changes. He then uses the methodology of Romer and Romer (2010) to isolate tax changes that are exogenous. Using this proxy for tax shocks, Cloyne (2013) estimates that a one percentage point decrease in the tax to GDP ratio increase GDP by 0.6 percent on impact.

While a number of studies have attempted to pin down the average estimate of taxation shocks, there is little existing evidence regarding changes in the transmission of this shock across time. An exception is Perotti (2005) who uses sub-sample estimates of the Blanchard and Perotti (2002) SVAR and finds some evidence of a decline in the impact of taxation shocks in the US. In contrast, Mertens and Ravn (2014) show that this sub-sample evidence for the US is much weaker when their proxy VAR is used to estimate the effects of tax shocks. For the UK, Cloyne (2013) presents estimates for sub-samples before and after 1979 and provides evidence that is consistent with a decline in the tax multiplier.

In this section, we re-consider the possibility of changes in the effects of tax shocks across time. In particular, we use the proposed proxy SVAR to estimate the time-varying response of output to tax shocks in the US and UK, respectively. For the US, the tax shock is identified by using the narrative shock measure of Mertens and Ravn (2012), while for the UK the proxy of Cloyne (2013) is used to estimate the model.

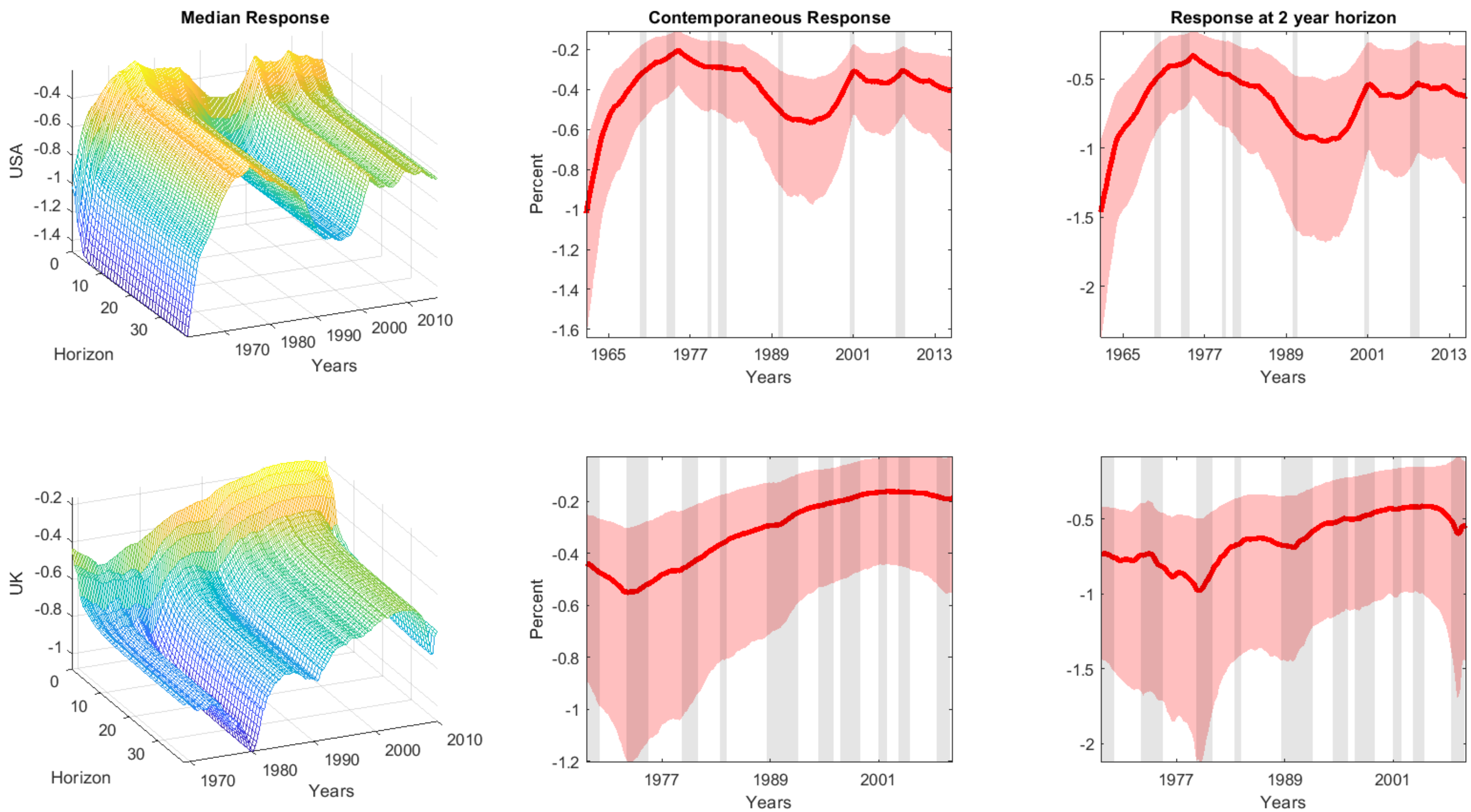

Figure 2: Cumulated response of y_t^{US} (top panel) and y_t^{UK} bottom panel to a shock normalised to increase taxes by one unit on impact at each point in time. The shaded area is the 68 percent error band while the vertical bars represent recession periods for the US and the UK as identified by the NBER and OECD, respectively.

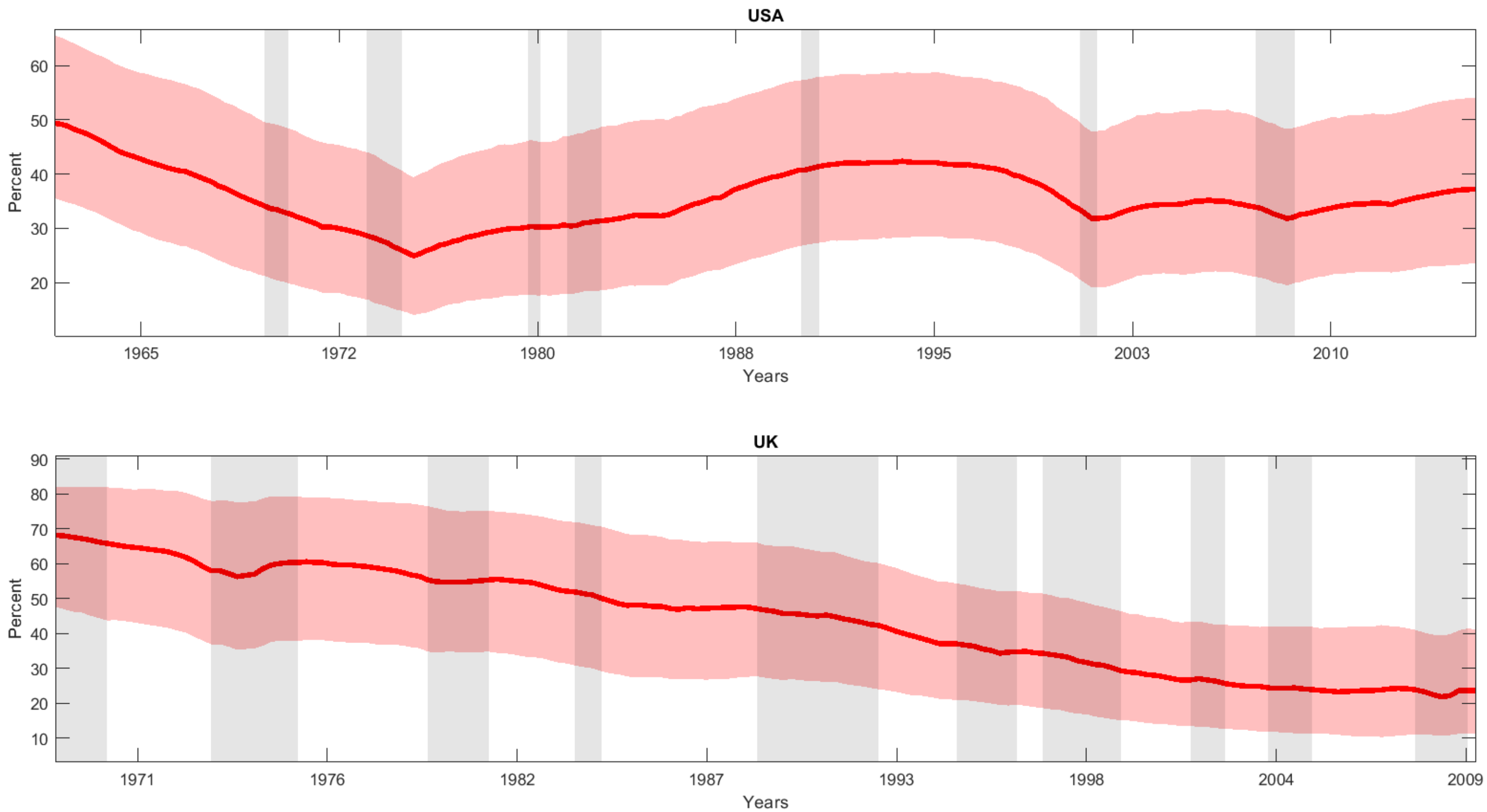

Figure 3: Contribution of the tax shock to the forecast error variance of y_t^{US} (top panel) and y_t^{UK} (bottom panel) at the two year horizon. The shaded area is the 68 percent error band while the vertical bars represent recession periods for the US and the UK as identified by the NBER and OECD, respectively.

4.1 Empirical model, data and priors

For both the US and UK, we estimate a TVP proxy SVAR(4) model using the MCMC algorithm proposed in Section 3. The model is defined as:

$$Y_t = B_t X_t + u_t \tag{16}$$

where X_t contains four lags and an intercept $X_t = [Y'_{t-1}, ..., Y'_{t-4}, 1]'$ and $var\left(u_t|A_t\right) = A_t A_t{}'$. The elements of the coefficient matrices B_t and A_t evolve over time as random walks.

For the US, the endogenous variables we include are (i) real per-capita federal government revenue $\left(T_t^{US}\right)$, (ii) real per-capita federal government spending $\left(G_t^{US}\right)$, and (iii) real per-capita GDP $\left(y_t^{US}\right)$. We follow Mertens and Ravn (2014) in defining the fiscal variables. As described in the technical Appendix, government spending is defined as the sum of Federal government consumption and investment. Taxes are calculated as current receipts of the Federal government plus contributions for social insurance less corporate income taxes from Federal Reserve banks. Both variables are deflated by the GDP deflator and divided by total population. The variables enter the model in log differences and the sample period runs from 1948Q2 to 2016Q2. Note that the tax shock proxy is only available from 1950Q1 to 2006Q4 and the missing observations are estimated as additional states in the model.

Following Cloyne (2013), the model for the UK is estimated using (i) Total tax receipts $\left(T_t^{UK}\right)$, (ii) Total government expenditure $\left(G_t^{UK}\right)$ and (iii) real GDP $\left(y_t^{UK}\right)$. Tax receipts are defined as the sum of tax and national insurance receipts while government expenditure is the sum of total managed expenditure less debt interest deflated by the GDP deflator. As in the case of the US, we transform the variables by taking log differences. The sample period runs from 1955Q2 to 2009Q4.

As described in Section 3.1, the prior for Q^b is set using OLS estimates of the VAR over a training sample of $T_0 = 50$ observations.[1] The degrees of freedom for the inverse Wishart priors for Q^a and Q^h are set to reflect the prior belief that changes in A_t are gradual. In particular, $T_{a,0}$ and $T_{h,0}$ are set to 25 for the US and 15 for the UK. The priors for the instrument equation parameters β and σ^2 are set to reflect the strong belief that the instruments are relevant and imply that $\rho \approx 0.3$. The basis of this prior belief is that fact that, both, Mertens and Ravn (2014) and Cloyne (2013) present extensive evidence to support the claim of instrument relevance.

The MCMC algorithm uses 100,000 replications with a burn-in of 50,000. Every 10th remaining draw is used for inference. The particle Gibbs step employs 10 particles. The technical appendix presents some evidence for convergence of the algorithm.

4.1.1 Empirical results

Figure 2 displays the cumulated response of y_t^{US} and y_t^{UK} to a taxation shock that increases taxes by one percent at each point in time. The top panel of figure shows that there is some evidence to indicate time-variation in the response of US GDP. However, this evidence is mainly limited to the beginning of the estimation sample. Over the second half of the 1960s, the two year ahead response declined by about one percent. Over the remaining sample, this response fluctuated around -0.5 percent to -1 percent.

For the UK, the estimated responses support the sub-sample results reported by Cloyne (2013). The (posterior median) response to the shock was relatively large before the 1980s with GDP growth falling by 0.8 percent on average at the two year horizon. After 1980, the average decline in y_t^{UK} at this horizon was estimated to be about 0.5 percent. As discussed in Cloyne (2013), this decline in the impact of fiscal shocks was possibly driven by an increase in the reaction of monetary policy to inflationary shocks.

[1] We impose the restriction that the VAR models are stable at each point in time. This condition is imposed via rejection sampling.

Figure 3 displays the contribution of the tax shock to the forecast error variance (FEV) of GDP growth over a two year horizon. The top panel shows that the contribution of the shock was substantial during the 1960s but began declining towards the end of this decade. This contribution has fluctuated around 30 percent over the remaining sample period. The contribution of the tax shock to UK GDP growth has declined persistently over the sample. This shock explained greater that 50 percent of the FEV of GDP growth before the 1980s. In contrast, this contribution averaged about 25 percent in the post-2000 period.[2]

5 Conclusion

This paper proposes a time-varying proxy SVAR model that can be used to estimate changes in the transmission of shocks. We provide a Gibbs sampling algorithm to approximate the posterior distribution of the parameters. Using a simple simulation experiment, we show that the algorithm displays reasonable performance. The proposed model is used to estimate the time-varying response to tax shocks in the US and the UK. Our results suggest that there has been a steady decline in the response of UK GDP growth to this shock with a corresponding fall in the contribution of this disturbance to UK GDP volatility. For the US, we find that the response to tax shocks displays time-variation in the early part of the sample, with a sharp decline during the early and mid-1960s.

In future work, it would be interesting to consider the possibility of changes in the transmission of monetary policy using the proposed model. Existing work on this issue has largely used TVP-VARs where the monetary policy shock is identified using zero or sign restrictions. The model that we propose in this paper could be used to investigate if identification via external instruments delivers different conclusions regarding changes in the impact of monetary disturbances.

[2] The technical Appendix contains some additional results suggesting that while the response of government spending to tax shock displays some time variation for the US, the estimated response is fairly constant over time for the UK.

References

Andrieu, Christophe, Arnaud Doucet and Roman Holenstein, 2010, Particle Markov chain Monte Carlo methods, *Journal of the Royal Statistical Society Series B* **72**(3), 269–342.

Baumeister, Christiane and Gert Peersman, 2013, Time-Varying Effects of Oil Supply Shocks on the US Economy, *American Economic Journal: Macroeconomics* **5**(4), 1–28.

Benati, Luca and Haroon Mumtaz, 2007, U.S. evolving macroeconomic dynamics - a structural investigation, *Working Paper Series 746*, European Central Bank.

Blanchard, Olivier and Roberto Perotti, 2002, An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output, *The Quarterly Journal of Economics* **117**(4), 1329–1368.

Caldara, Dario and Edward Herbst, 2019, Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs, *American Economic Journal: Macroeconomics* **forthcoming**.

Canova, Fabio and Fernando J. Pérez Forero, 2015, Estimating overidentified, nonrecursive, time varying coefficients structural vector autoregressions, *Quantitative Economics* **6**(2), 359–384.

Carter, C and P Kohn, 1994, On Gibbs sampling for state space models, *Biometrika* **81**, 541–53.

Cloyne, James, 2013, Discretionary Tax Changes and the Macroeconomy: New Narrative Evidence from the United Kingdom, *The American Economic Review* **103**(4), 1507–1528.

Cogley, Timothy and Thomas J. Sargent, 2005, Drift and Volatilities: Monetary Policies and Outcomes in the Post WWII U.S, *Review of Economic Dynamics* **8**(2), 262–302.

Drautzburg, Thorsten, 2016, A narrative approach to a fiscal DSGE model, *Working Papers 16-11*, Federal Reserve Bank of Philadelphia.

Gali, Jordi and Luca Gambetti, 2009, On the Sources of the Great Moderation, *American Economic Journal: Macroeconomics* **1**(1), 26–57.

Lindsten, Fredrik, Michael I. Jordan and Thomas B. Schön, 2014, Particle Gibbs with Ancestor Sampling, *Journal of Machine Learning Research* **15**, 2145–2184.

Mertens, Karel and Morten O. Ravn, 2012, Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks, *American Economic Journal: Economic Policy* **4**(2), 145–81.

Mertens, Karel and Morten O. Ravn, 2013, The Dynamic Effects of Personal and Corporate Income Tax Changes in the United States, *American Economic Review* **103**(4), 1212–47.

Mertens, Karel and Morten O. Ravn, 2014, A reconciliation of SVAR and narrative estimates of tax multipliers, *Journal of Monetary Economics* **68**, S1 – S19.

Paul, Pascal, 2017, The Time-Varying Effect of Monetary Policy on Asset Prices, *Working Paper Series 2017-9*, Federal Reserve Bank of San Francisco.

Perotti, Roberto, 2005, Estimating the effects of fiscal policy in OECD countries, *Proceedings* .

Primiceri, G, 2005, Time varying structural vector autoregressions and monetary policy, *The Review of Economic Studies* **72**(3), 821–52.

Rogers, John H., Chiara Scotti and Jonathan H. Wright, 2016, Unconventional Monetary Policy and International Risk Premia, *International Finance Discussion Papers 1172*, Board of Governors of the Federal Reserve System (U.S.).

Romer, Christina D. and David 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.

Rubio-Ramirez, Juan F, Daniel Waggoner and Tao Zha, 2010, Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference, *Review of Economic Studies* **77**(2), 665–696.

Stock, James H. and Mark W. Watson, 2008, What's New in Econometrics- Time Series, *Lecture 7*, National Bureau of Economic Research, Inc.