

Dybowski, T. Philipp

Conference Paper

Tracing the Role of Foresight on the Effects of U.S. Tax Policy: Evidence from a Time-Varying SVAR

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Public debt and fiscal foresight, No. F03-V1

Provided in Cooperation with:

Verein für Socialpolitik / German Economic Association

Suggested Citation: Dybowski, T. Philipp (2015) : Tracing the Role of Foresight on the Effects of U.S. Tax Policy: Evidence from a Time-Varying SVAR, Beiträge zur Jahrestagung des Vereins für Socialpolitik 2015: Ökonomische Entwicklung - Theorie und Politik - Session: Public debt and fiscal foresight, No. F03-V1, ZBW - Deutsche Zentralbibliothek für Wirtschaftswissenschaften, Leibniz-Informationszentrum Wirtschaft

This Version is available at:

<https://hdl.handle.net/10419/113049>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

Tracing the Role of Foresight on the Effects of U.S. Tax Policy: Evidence from a Time-Varying SVAR*

T. Philipp Dybowski [†]

August 26, 2015

Abstract

Recent empirical research emphasizes the importance of foresight for tax policy analyses. According to Leeper et al. (2013a), failing to model foresight adequately can lead to biased inference in empirical models. The authors reveal this bias by augmenting the SVAR model of Blanchard and Perotti (2002) with a measure for tax foresight. I extend this finding by transforming their model into a time-varying SVAR with stochastic volatility. This approach allows to study the time-variations of tax foresight, and resulting effects on tax policy shocks over time. Two findings stand out: First, both anticipated and unanticipated tax shocks show considerable movements over time. The magnitude of these shocks suggests that some tax reforms are more anticipated than others. Second, I find that the bias in the tax shocks is more pronounced during the 1980s and 1990s, suggesting that tax reforms in these decades were to a higher degree anticipated than in the 1960s and 1970s. The results compare well to other studies and find support in anecdotal evidence on documented U.S. tax reforms.

JEL Classification: C11, C32, D84, E62, H20, H30

Keywords: Fiscal multiplier, tax expectations, identified VARs, Bayesian estimation, news

*I thank Bernd Kempa, Jana Riedel, Philipp Adämmer and Samad Sarferaz for helpful comments on a drafted version of this paper.

[†]Institute of International Economics, University of Münster, Universitätsstr. 14-16, 48143 Münster, Germany, Phone: +49 251 8328663, E-mail: philipp.dybowski@wiwi.uni-muenster.de.

1 Introduction

Recent empirical research on fiscal policy has been focused on the incorporation of fiscal foresight to the analysis of policy reactions. The necessity of including such information stems mainly from the insight that anticipated policy changes evoke economic reactions that differ from unanticipated changes. Economic theory attributes this difference to forward-looking agents who adjust their behavior upon the arrival of fiscal news. An unanticipated policy change affects individuals' behavior only when the new policy is implemented, whereas an anticipated policy change may cause reactions well ahead of the actual implementation (see, e.g., Hall, 1971; Judd, 1985; Auerbach, 1989; House and Shapiro, 2006).

Aside from a theoretical justification, recent econometric studies find that failing to model foresight adequately can lead to biased inference in empirical models. Yang (2005) shows within a simulated model how estimates of tax effects are biased if foresight is ignored. Leeper et al. (2012) find the information content of fiscal policy news to be a time-varying process, and show within a DSGE framework the consequences of ignoring this time-variation. Leeper et al. (2013a) find that the presence of foresight misaligns the econometrician's information set with that of economic agents to the extent that the econometrician has a smaller subset of information available to her. In such cases, the true structural shocks can not be recovered. Although these three studies underline the importance for empirical models to account for foresight, the consequences for specific applications in fiscal policy research are just starting to be quantified.

The empirical literature on the effects of fiscal policy is divided into two approaches. The narrative approach, which achieves identification by relying on rich data sets that contain information on legislated tax changes (see, e.g., Ramey and Shapiro, 1998; Edelberg et al., 1999; Burnside et al., 2004; Eichenbaum and Fisher, 2005), and the conventional SVAR approach, which identifies relevant policy shocks by imposing certain restrictions on the covariance or parameter matrices (see, e.g., Fatás and Mihov, 2001; Perotti, 2005; Galí et al., 2007). Within both approaches the importance to account for policy foresight has

been acknowledged and attempts to disentangle anticipated from unanticipated policy shocks are either based on timing rules or certain model restrictions (see, e.g., Ramey, 2011; Mertens and Ravn, 2012; Blanchard and Perotti, 2002; Mountford and Uhlig, 2009).

Leeper et al. (2013a) criticize these attempts to account for foresight as tax reforms are too complex and heterogeneous to be classified by simple timing rules or restrictions. Yang (2007) lends support to this notion by documenting detailed time lines for postwar U.S. tax reforms, underlining the uniqueness of each tax reform. Fortune (1996) and Leeper et al. (2012) show that fiscal news contain more concrete information at some times and less in others. The latter argue that information flows surrounding tax reforms can be divided into two kinds of lags. An ‘inside lag’, constituting the time between a first informal policy proposal until it is passed, and an ‘outside lag’, stretching from when legislation is enacted until it takes effect. During these lags media reports or political campaigns can lead to the build up of expectations, which are not captured by simple timing rules or restrictions.

Instead, Leeper et al. (2013a) (in the following ‘LWY’) use the spread between tax-exempt U.S. municipal bonds and taxable U.S. Treasury bonds to gauge the effects of tax foresight. This so-called implicit tax rate resembles a weighted average of discounted future tax rates and thus provides a flexible measure of tax foresight. Accordingly, there is no need to define a period of foresight ex-ante, or to specify a functional form for information flows. The authors argue that it is especially suitable to recover the true structural shocks of the model. To underscore their argument, LWY re-evaluate the SVAR of Blanchard and Perotti (2002) (in the following ‘BP’) by adding the implicit tax rate. Results show an upward bias in BP’s output multipliers for tax shocks, which LWY attribute to the omission of tax foresight. However, as their reassessment of BP’s findings is based on a time-invariant SVAR, no inference on time-varying effects of foresight can be made.

This paper seeks to address this gap by estimating two versions of time-varying parameter VAR (TVP-VAR) models with stochastic volatility, using both the exact same fiscal data as in BP, and the implicit tax rate data as in LWY. Within this framework it is possible to

study the time variations of anticipated and unanticipated tax shocks, as well as the bias that results from omitting tax news from an econometric model. This is of particular interest as the degree of foresight that surrounds tax changes varies over time (Leeper et al., 2012). If news contain less relevant information on future tax changes in some periods, and more in others, this should be reflected in the variation of the anticipated and the unanticipated tax shock, as well as the bias.

I find positive effects on output after an anticipated tax increase. The unanticipated tax increase causes the expected negative responses in output. This has been shown in earlier studies by Mertens and Ravn (2011), Mertens and Ravn (2012), and Leeper et al. (2013a). More interestingly, I find that the responses to the anticipated and the unanticipated tax shocks vary considerably over time. Between the early 1970s and the early 1990s the response to an anticipated tax shock roughly doubles. In contrast, the unanticipated tax shock shows less of a trend, but more fluctuations between 1960 to 1997. Most notable is a distinct negative peak in 1975 and a turning point at the beginning of the 1980s, after which the impact of the unanticipated tax shock first declines and then fluctuates for the rest of the sample.

A comparison of the time-varying tax revenue shocks in the BP and LWY specifications reveals the bias and its dynamics over time. I find this bias to be largely close to zero during the 1960s until 1975, and comparatively high going into the 1980s and 1990s. These findings suggest that tax reforms during the 1980s and 1990s were accompanied by a higher degree of tax foresight. This adds to findings from Fortune (1996) and Leeper et al. (2012), and is supported by anecdotal evidence on documented U.S. tax reforms.

The remainder of this paper is structured as follows. Section 2.1 describes the data and explains the construction of the implicit tax rate. Section 2.2 describes the model setup and the identification scheme. Section 2.3 explains the estimation procedure before giving a short discussion of the priors in Section 2.4. Section 3 presents and analyzes the key findings. Section 4 concludes.

2 Empirical analysis

2.1 Data and the implicit tax rate

As argued in previous studies, the spread between U.S. municipal bonds and U.S. Treasury bonds contains information about expected tax changes (Poterba, 1989; Fortune, 1996; Park, 1997; Kueng, 2011). The reason lies in their differential tax treatment. Municipal bonds, in contrast to U.S. Treasury bonds, are exempt from federal taxes. Given that asset markets are efficient, an investor is indifferent between buying a tax-exempt municipal bond and a taxable Treasury bond if the yield on a municipal bond \mathcal{Y}^M equals the after-tax yield of a government bond \mathcal{Y} . Assuming equal maturities m for both bonds, callability, market risk, and credit risk amongst others, an implicit tax rate can be defined as $\tau_{t,m}^I = 1 - \mathcal{Y}^M/\mathcal{Y}_m$, which contains information about expected future tax rates. Leeper et al. (2012, 2013a) derive this relationship as follows: A newly issued tax-exempt municipal bond with maturity m , a par value of \$1, and per-period coupons C_m^M , sells at par if

$$1 = \frac{C_m^M}{\sum_{t=1}^m (1 + R_t^r)^t} + \frac{1}{(1 + R_m^r)^m}, \quad (1)$$

where R_t^r is the after-tax nominal interest rate that applies to after-tax payments made in period t .

Under no arbitrage conditions, a taxable bond with equal maturity, a par value of \$1, and coupon payments C_m , sells at par if

$$1 = \frac{\sum_{t=1}^m C_m(1 - \tau_t^e)}{\sum_{t=1}^m (1 + R_t^r)^t} + \frac{1}{(1 + R_m^r)^m}, \quad (2)$$

where τ_t^e denotes the future expected tax rate in period t . Solving (1) and (2) as a function of the relative coupon spread C_m^M/C_m , and recognizing that the yield-to-maturity equals the

coupon payments for bonds that sell at par, yields

$$\frac{y_m^{\mathcal{M}}}{y_m} = \sum_{t=1}^m \omega_t (1 - \tau_t^e), \quad (3)$$

where $\omega_t = \delta_t / \sum_{t=1}^m \delta_t$ and $\delta_t = (1 + R_t^r)^{-t}$. Since the weights ω_t sum to unity (3) simplifies to $\tau_{t,m}^I = \sum_{t=1}^m \omega_t \tau_t^e$, matching the above definition of the implicit tax rate. Thus, the implicit tax rate at time t represents a weighted average of discounted expected future tax rates over the maturity of the bonds. As such, it should respond to any news concerning anticipated future tax changes that occur during the periods until maturity. Leeper et al. (2012) find the information content of the implicit tax rate with respect to tax policy to be a highly time-varying process. This justifies my approach to utilize a time-varying SVAR model as to explicitly take into account such time variation.

The data set for the empirical analysis comprises quarterly data from 1960:I–1997:IV. Data on net taxes nt_t , government spending g_t , output x_t , as well as the elasticity of net taxes to output a_{1t}^* are exactly as in BP. All variables are seasonally adjusted and in logarithms of real per capita values.¹ Data for the implicit tax rate (in the remainder of this paper denoted by s_t) are as in LWY, and contain data on prime-grade U.S. municipal bonds and U.S. Treasury bonds with maturity lengths of 5 years.²

2.2 Model setup and identification

I employ a time-varying parameters VAR(p) model with stochastic volatility, as proposed by Primiceri (2005) and Del Negro and Primiceri (2013), and a modified identification scheme similar to Pereira and Lopes (2014). This identification scheme translates the analysis of the time-invariant BP and LWY models into a time-varying setup.

¹For detailed descriptions concerning the construction of these time series see Blanchard and Perotti (2002), Section III and the Appendix A.1.

²I thank Todd B. Walker for providing the data for the implicit tax rate. For a detailed description of the data and its sources see Leeper et al. (2013b), Section S.1.3.2.

The time-varying VAR(p) model has the reduced form:

$$\mathbf{y}_t = \boldsymbol{\beta}_{0,t} + \sum_{j=1}^p \boldsymbol{\beta}_{j,t} \mathbf{y}_{t-j} + \mathbf{u}_t, \quad t = 1, \dots, T, \quad (4)$$

where \mathbf{y}_t is an $n \times 1$ vector of endogenous variables, $\boldsymbol{\beta}_{0,t}$ is an $n \times 1$ vector of time-varying constants, $\boldsymbol{\beta}_{j,t}$ are p time-varying $n \times n$ parameter matrices, and \mathbf{u}_t is an $n \times 1$ vector of Gaussian reduced-form errors. By defining $\boldsymbol{\Theta}_t = (\boldsymbol{\beta}_{0,t}, \boldsymbol{\beta}_{1,t}, \dots, \boldsymbol{\beta}_{p,t})$ and $\mathbf{Z}_{t-1} = (\mathbf{1}, \mathbf{y}_{t-1}, \dots, \mathbf{y}_{t-p})'$ the model can be expressed in matrix form

$$\mathbf{y}_t = \boldsymbol{\Theta}_t \mathbf{Z}_{t-1} + \mathbf{u}_t, \quad (5)$$

which can be re-written such that

$$\mathbf{y}_t = (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n) \boldsymbol{\theta}_t + \mathbf{u}_t, \quad (6)$$

with $\boldsymbol{\theta}_t = \text{vec}(\boldsymbol{\Theta}_t)$.

The reduced-form innovations are structured as follows

$$\mathbf{A}_t \mathbf{u}_t = \mathbf{B}_t \mathbf{e}_t, \quad (7)$$

where \mathbf{A}_t and \mathbf{B}_t are the $n \times n$ matrices of contemporaneous coefficients, and \mathbf{e}_t is an $n \times 1$ vector of Gaussian structural innovations. To complete the model, the structural innovations are specified as

$$\mathbf{e}_t = \mathbf{D}_t \boldsymbol{\varepsilon}_t, \quad (8)$$

where \mathbf{D}_t is an $n \times n$ diagonal matrix containing the standard deviations of the orthogonal shocks, and $\boldsymbol{\varepsilon}_t$ is an $n \times 1$ Gaussian vector with $E(\boldsymbol{\varepsilon}_t) = 0$ and $E(\boldsymbol{\varepsilon}_t \boldsymbol{\varepsilon}_t') = \mathbf{I}_n$.

Under these assumptions the reduced-form covariance matrix can be recovered by

$$\Sigma_t = \text{Var}(\mathbf{u}_t) = \mathbf{A}_t^{-1} \mathbf{B}_t \mathbf{D}_t \mathbf{D}_t' \mathbf{B}_t' (\mathbf{A}_t^{-1})'. \quad (9)$$

Given the model setup, the BP and the LWY models apply the following identification schemes. The BP model forms the baseline specification, with net taxes nt_t , government expenditure g_t , and output x_t . In the LWY specification, the model is augmented by the implicit tax rate s_t , to account for tax foresight. The ordering of the variables is as follows: $\mathbf{y}_t \equiv [nt_t, g_t, x_t]'$ for the BP specification and $\mathbf{y}_t \equiv [nt_t, g_t, x_t, s_t]'$ for the model proposed by LWY.

BP achieve identification in their time-invariant SVAR by utilizing information regarding features of the U.S. spending and tax/transfer systems, where the elasticity of net taxes to output a_{13}^* is calibrated outside of the VAR.³ Together with zero restrictions on mutual feedback between net taxes and government expenditure, as well automatic feedback from output to government expenditure within one quarter, matrices \mathbf{A} and \mathbf{B} become

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & -a_{13}^* \\ 0 & 1 & 0 \\ -c_{31} & -c_{32} & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & 0 & 0 \\ b_{21} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (10)$$

This leaves parameters b_{21} , c_{31} and c_{32} to be estimated within the VAR. As u_t^{nt} appears as endogenous regressor in the output equation, BP estimate c_{31} and c_{32} by 2SLS estimation, using the residuals of the tax and government spending equations as instruments for the endogenous regressors in the output equation.

In a time-varying context, Pereira and Lopes (2014) show that an equivalent identification scheme to that in (10) exists. Matrices \mathbf{A}_t and \mathbf{B}_t then become

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 & -a_{13,t}^* \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} 1 & 0 & 0 \\ \beta_{21,t} & 1 & 0 \\ \beta_{31,t} & \beta_{32,t} & 1 \end{pmatrix}, \quad (11)$$

³See Appendix A.2. of Blanchard and Perotti (2002) for more details.

and as such yield the same impulse-responses as in (10), but within the time-varying framework.⁴ Structural shocks estimated from (11) fully coincide with those from (10), except for a scale factor of $1/(1 - c_{31}a_{13}^*)$ for the output innovation in the third equation. Since $a_{13,t}^*$ is again calibrated outside the VAR, the only parameters to be estimated in (11) are contained in the \mathbf{B}_t matrix. Due to its lower triangular structure, parameters in \mathbf{B}_t are predetermined in the time-varying estimation procedure, making an estimation by 2SLS redundant.

The identification in the LWY model follows the same logic. Augmenting the system with the implicit tax rate s_t in a fourth equation transforms the time-varying matrices \mathbf{A}_t and \mathbf{B}_t to

$$\mathbf{A}_t = \begin{pmatrix} 1 & 0 & -a_{13,t}^* & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{B}_t = \begin{pmatrix} 1 & 0 & 0 & 0 \\ \beta_{21,t} & 1 & 0 & 0 \\ \beta_{31,t} & \beta_{32,t} & 1 & 0 \\ \beta_{41,t} & \beta_{42,t} & \beta_{43,t} & 1 \end{pmatrix}. \quad (12)$$

Again, all parameters to be estimated are contained in \mathbf{B}_t , and the structural shocks fully coincide with the time-invariant specification in Leeper et al. (2013a), except for the scaling factor $1/(1 - c_{31}a_{13}^*)$ for the output innovation in the third equation.

2.3 Estimation procedure

To obtain impulse response functions for the above models, the time-varying covariance matrix $\mathbf{\Sigma}_t$ has to be estimated. This is done in three blocks: The first block consists of the time-varying coefficient states in vector $\boldsymbol{\theta}_t$ of Eq. (6). The second block embodies the non-zero and non-unity covariance states in \mathbf{B}_t of Eq. (11) for the BP model, and Eq. (12) for the LWY model. Matrix \mathbf{A}_t does not contain any unknown parameters and therefore does not have to be estimated. The third block comprises the volatility states, contained on the main diagonal of \mathbf{D}_t .

Following the vector notation in Eq. (6), the free parameters of \mathbf{B}_t and \mathbf{D}_t are also stacked in vectors, row by row, where $\mathbf{b}_t = [\beta_{21,t}, (\beta_{31,t} \beta_{32,t}), \dots, (\beta_{n1,t} \beta_{n(n-1),t})]'$ and $\mathbf{d}_t = \text{diag}(\mathbf{D}_t)$. The coefficient and covariance states are assumed to follow random walks, and the volatility

⁴See Appendix C of Pereira and Lopes (2014).

states evolve as geometric random walks,

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\nu}_t, \quad (13)$$

$$\mathbf{b}_t = \mathbf{b}_{t-1} + \boldsymbol{\zeta}_t, \quad (14)$$

$$\log \mathbf{d}_t = \log \mathbf{d}_{t-1} + \boldsymbol{\eta}_t. \quad (15)$$

For estimation purposes it is assumed that all the innovations are jointly normally distributed such that

$$\begin{pmatrix} \boldsymbol{\varepsilon}_t \\ \boldsymbol{\nu}_t \\ \boldsymbol{\zeta}_t \\ \boldsymbol{\eta}_t \end{pmatrix} \sim N(\mathbf{0}, \mathbf{V}) \quad \text{with} \quad \mathbf{V} = \begin{pmatrix} \mathbf{I}_n & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{Q} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{S} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{W} \end{pmatrix}, \quad (16)$$

where the \mathbf{S} matrix is restricted to be block-diagonal.⁵ All other covariance matrices are left unrestricted.

The model can be written in state-space form, and is estimated using Bayesian simulation methods. Equations (6), (7) and (8) constitute the measurement equations for each block, linking the time-varying parameters to the data. These time-varying parameters resemble unobserved state variables, whose laws of motion are defined by the transition equations (13), (14), and (15). Sequentially for each block, the Gibbs sampler is applied to draw the state vectors from posterior Gaussian distributions. The means and variances of these posterior distributions are obtained by a forward run of the Kalman filter, followed by a backward recursion of the simulation smoother (see Carter and Kohn, 1994 and Kim et al., 1998). After initializing \mathbf{b}^T , \mathbf{d}^T , \mathbf{s}^T , \mathbf{V} , the Gibbs sampler iterates over the following steps to generate the histories of the parameters, e.g. $\boldsymbol{\theta}^T = [\boldsymbol{\theta}_1 \boldsymbol{\theta}_2 \dots \boldsymbol{\theta}_T]$.⁶

I Initialize \mathbf{b}^T , \mathbf{d}^T , \mathbf{s}^T , \mathbf{V} .

II Sample $\boldsymbol{\theta}^T$ from the conditional posterior $p(\boldsymbol{\theta}^T | \mathbf{y}^T, \mathbf{b}^T, \mathbf{d}^T, \mathbf{V})$.

III Sample \mathbf{b}^T from the conditional posterior $p(\mathbf{b}^T | \mathbf{y}^T, \boldsymbol{\theta}^T, \mathbf{d}^T, \mathbf{V})$.

⁵See Primiceri (2005) for more details on these assumptions and a discussion about relaxing them.

⁶See the Appendix for a more detailed description of the Gibbs sampling procedure.

IV Sample \mathbf{s}^T from the conditional posterior $p(\mathbf{s}^T | \mathbf{y}^T, \boldsymbol{\theta}^T, \mathbf{b}^T, \mathbf{V})$.

V Sample \mathbf{d}^T from the conditional posterior $p(\mathbf{d}^T | \mathbf{y}^T, \boldsymbol{\theta}^T, \mathbf{b}^T, \mathbf{s}^T, \mathbf{V})$.

VI Sample \mathbf{V} , by sampling \mathbf{Q} , \mathbf{W} and \mathbf{S} from $p(\mathbf{Q}, \mathbf{W}, \mathbf{S} | \mathbf{y}^T, \boldsymbol{\Theta}^T, \mathbf{B}^T, \mathbf{D}^T)$
 $= p(\mathbf{Q} | \mathbf{y}^T, \boldsymbol{\Theta}^T, \mathbf{B}^T, \mathbf{D}^T) \cdot p(\mathbf{W} | \mathbf{y}^T, \boldsymbol{\Theta}^T, \mathbf{B}^T, \mathbf{D}^T) \cdot$
 $p(\mathbf{S}_1 | \mathbf{y}^T, \boldsymbol{\Theta}^T, \mathbf{B}^T, \mathbf{D}^T) \cdot \dots \cdot p(\mathbf{S}_{n-1} | \mathbf{y}^T, \boldsymbol{\Theta}^T, \mathbf{B}^T, \mathbf{D}^T).$

VII Go back to II.

2.4 Choice of Priors

I follow conventional choices from the TVP-VAR literature in calibrating the priors for the model. I choose mainly the same priors as Primiceri (2005) and Pereira and Lopes (2014). However, I calibrate the priors on the basis of the entire sample from 1960:I to 1997:IV instead of specifying a pre-sample. This approach is suggested by Canova and Ciccarelli (2006) and Canova (2007), and also adapted by Kirchner et al. (2010), in cases where the overall sample is comparatively short. Applying an initial training sample of 40 periods as in Primiceri (2005) would mean to sacrifice about 25% of observations. In light of the high number of parameters to be estimated and the already limited sample size, a further truncation seems undesirable. Moreover, the scope of this analysis is to obtain a close time-varying match to the original BP and LWY models, which are based on a sample from 1960:I to 1997:IV.

Prior distributions for the initial states and the hyperparameters are calibrated according to estimates of the identified time-invariant VARs, as proposed in BP and LWY. The resulting point estimates $(\hat{\boldsymbol{\theta}}^{OLS}, \hat{\mathbf{b}}^{OLS}, \log \hat{\mathbf{d}}^{OLS})$ form the means for the Gaussian initial states. Covariance matrices for $\boldsymbol{\theta}_0$ and \mathbf{b}_0 are set to four times multiples of their asymptotic

covariances and for \mathbf{d}_0 to an identity matrix:

$$\begin{aligned}\boldsymbol{\theta}_0 &\sim N(\hat{\boldsymbol{\theta}}^{OLS}, 4 \cdot V(\hat{\boldsymbol{\theta}}^{OLS})), \\ \mathbf{b}_0 &\sim N(\hat{\mathbf{b}}^{OLS}, 4 \cdot V(\hat{\mathbf{b}}^{OLS})), \\ \log \mathbf{d}_0 &\sim N(\log \hat{\mathbf{d}}^{OLS}, \mathbf{I}_n).\end{aligned}\tag{17}$$

The hyperparameters $(\mathbf{Q}, \mathbf{S}, \mathbf{W})$ have conjugate inverse-Wishart priors. Their scaling parameters consist of the respective estimated coefficient covariance matrices over the training sample, multiplied by a scaling factor and their respective degrees of freedom. The scaling factors (k_Q, k_S, k_W) express the prior beliefs about the degree of time variation:⁷

$$\begin{aligned}\mathbf{Q} &\sim IW(k_Q^2 \cdot T_0 \cdot V(\hat{\boldsymbol{\theta}}^{OLS}), T_0), \\ \mathbf{W} &\sim IW(k_W^2 \cdot (n+1) \cdot \mathbf{I}_n, (n+1)), \\ \mathbf{S}_1 &\sim IW(k_S^2 \cdot 2 \cdot V(\hat{\mathbf{b}}_1^{OLS}), 2), \\ \mathbf{S}_2 &\sim IW(k_S^2 \cdot 3 \cdot V(\hat{\mathbf{b}}_2^{OLS}), 3), \\ \mathbf{S}_3 &\sim IW(k_S^2 \cdot 4 \cdot V(\hat{\mathbf{b}}_3^{OLS}), 4),\end{aligned}\tag{18}$$

where $k_Q = 0.01$, $k_S = 0.1$ and $k_W = 0.1$. For the BP specification $n = 3$, but set to $n = 4$ for the LWY model, as it includes the implicit tax rate as fourth variable. Note also that \mathbf{S}_3 only exists in the time-varying LWY model. Given the high dimensionality of parameters to estimate, I follow the literature and choose a parsimonious lag length of $p = 2$ in both specifications. The above procedure is applied to create 100.000 draws from the posterior, while discarding the first 80.000 draws and keeping every 10th draw from the remaining 20.000 to eliminate possible autocorrelation of the draws. Inference is made from posterior means.

I report transformed impulse response functions to represent (non-cumulative) multipliers in the vein of Blanchard and Perotti (2002). As such, each multiplier has the interpretation

⁷See Primiceri (2005) and Pereira and Lopes (2014) for a more detailed discussion of the influence of the scaling parameters on the time variation of the states.

of a dollar response to a dollar shock in the tax revenue variable. I follow Leeper et al. (2013b) and apply the tax revenue data to give the impulse responses for the implicit tax rate the interpretation of a dollar shock to anticipated tax revenue.

3 Empirical Results

The discussion of my findings is structured as follows: In section 3.1 I present the time-varying output multipliers in the LWY model, and discuss their dynamics to anticipated and unanticipated tax shocks. In section 3.2 I examine to what extent the omission of tax foresight leads to a bias in the time-varying SVAR, and discuss how the dynamics of the bias relate to the degree of foresight present over different decades.

3.1 Time-Varying Anticipated and Unanticipated Tax Shocks

I begin the analysis by looking at time variations in the anticipated and unanticipated tax shock between the three decades of the 1970s, 1980s and 1990s. This segmentation is motivated by studies of Fortune (1996) and Leeper et al. (2012), who find that the information content of news regarding tax changes varies considerably across these decades. Figure 1 shows aggregated impulse responses, computed as the average impulse responses over every quarter in the respective decade. Results compare well to the time-invariant results reported in FIGURE S.1. in Leeper et al. (2013b) as they show the expected shapes and magnitudes. Of key interest are the responses of output to both a positive implicit tax shock (panel (a)) and tax revenue shock (panel (c)), as these have the interpretations of an anticipated and unanticipated tax shock, respectively. Two quarters after impact the implicit tax shock has a positive effect on output. This reflects anticipation effects of fiscal foresight that lasts over several quarters, and states that anticipated increases in tax rates are expansionary in the short to medium run. The unanticipated tax shock has the expected negative impact on output over all horizons. Panels (b) and (d) report responses of tax revenues to the implicit tax rate and vice-versa. They underpin the plausibility of the identification by showing a positive

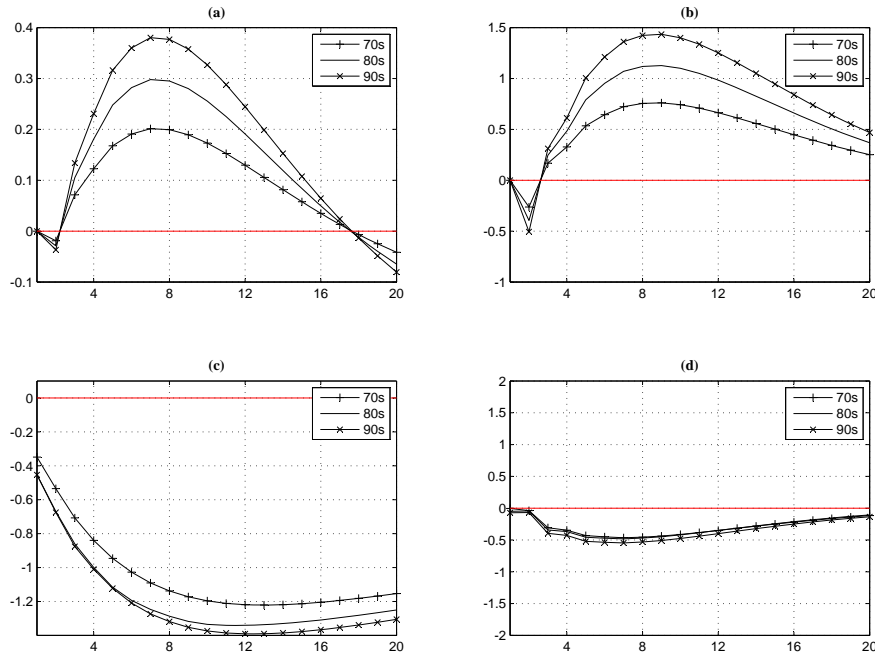


Figure 1: Average Mean Impulse Responses for the 1970s, 1980s, and 1990s, (a) Implicit Tax rate on Output, (b) Implicit Tax Rate on Tax Revenues, (c) Tax Revenues on Output, and (d) Tax Revenues on Implicit Tax Rate.

effect of tax revenues to an implicit tax shock, and an effect of tax revenue to the implicit tax shock not significantly different from zero.⁸ Figures 2 and 3 further corroborate the above findings, as they show clear signs of shifts in the multipliers. For both the anticipated and the unanticipated tax shock the average effect has grown stronger over the decades. Whereas the peak impact of an anticipated tax shock, reached after 7 quarters, is 0.20 in the 1970s, it rises to 0.29 in the 1980s, and to 0.37 in the 1990s. This is almost a doubling in magnitude between the 1970s and 1990s. A similar evolution, yet less pronounced, is observed for the unanticipated tax shock. In the 1970s the biggest impact is reached after 13 quarters, with -1.22. Then, during the 1980s the impact grows in severity reaching its maximum negative impact at -1.35, and -1.39 after a further slight increase going into the 1990s. Figure 2 also reveals that responses of output to the anticipated tax shock are significant between 4 and 10 quarters over all three decades. The unanticipated tax shock, reported in Figure 3, has a significant and negative impact on output over almost the entire horizon. In the 1980s, and

⁸To conserve space, plots with confidence bands are not reported here, but are available upon request.

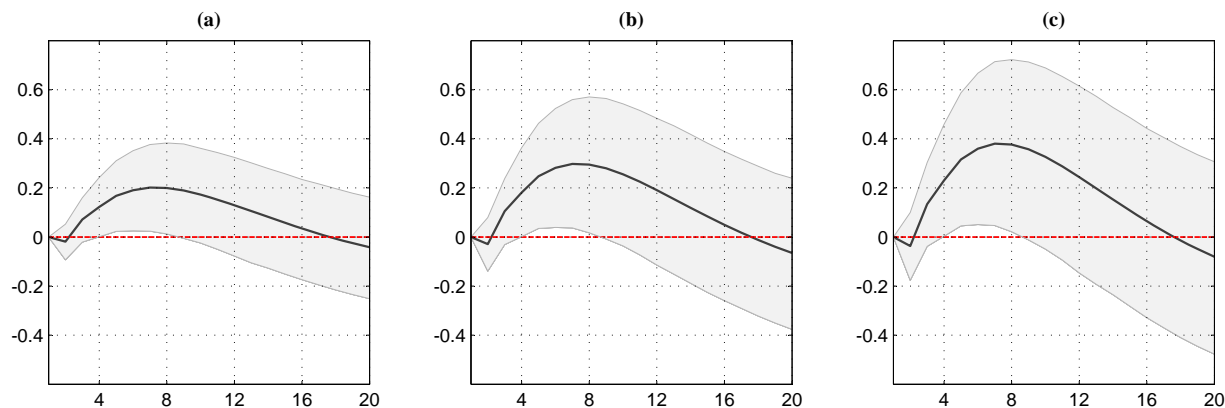


Figure 2: Average Mean Impulse Responses of Output to Implicit Tax Shock for the (a) 1970s, (b) 1980s, and (c) 1990s, with 32-th and 68-th percentiles

more profoundly in the 1990s, the unanticipated tax shock loses its significance on impact, and only becomes significant after 1 quarter and 3 quarters, respectively.

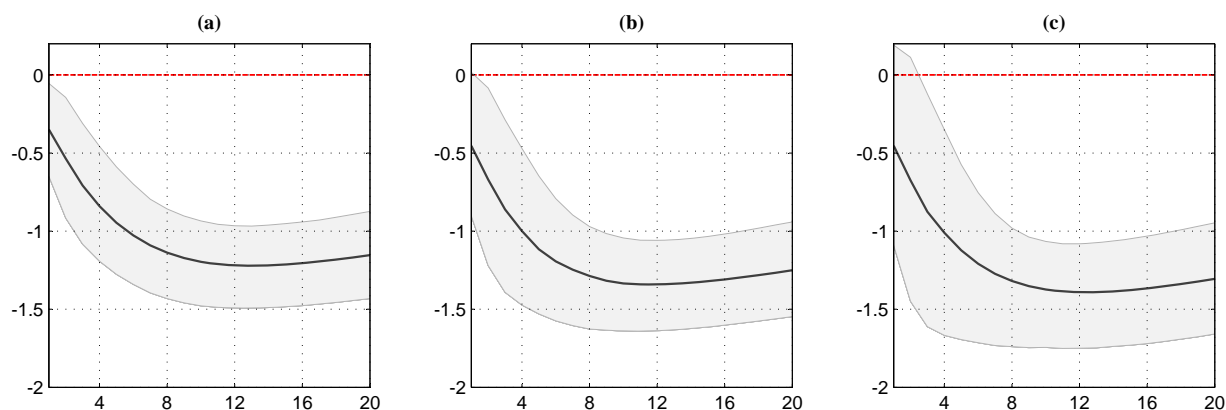


Figure 3: Average Mean Impulse Responses of Output to Tax Revenues Shock for the (a) 1970s, (b) 1980s, and (c) 1990s, with 32-th and 68-th percentiles

Since the above impulse responses report decade averages, they only reveal shifts between these time periods, but not variations within them. Figures 4 and 5 report the anticipated and unanticipated output multipliers for the entire sample from 1960 to 1997. This refined resolution helps to track down the role of tax foresight as accurately as possible and to link specific shifts with anecdotal evidence from documented U.S tax reforms.

Figure 4 reveals that the anticipated tax shock starts out with little time variation until 1975. At horizons 4, 8 and 12, the anticipated shock gains in importance going into the 1980s,

reaching a plateau around the years 1983–1984.⁹ A final surge takes place from around 1986 to roughly 1992, before declining somewhat towards the end of the sample in 1997.

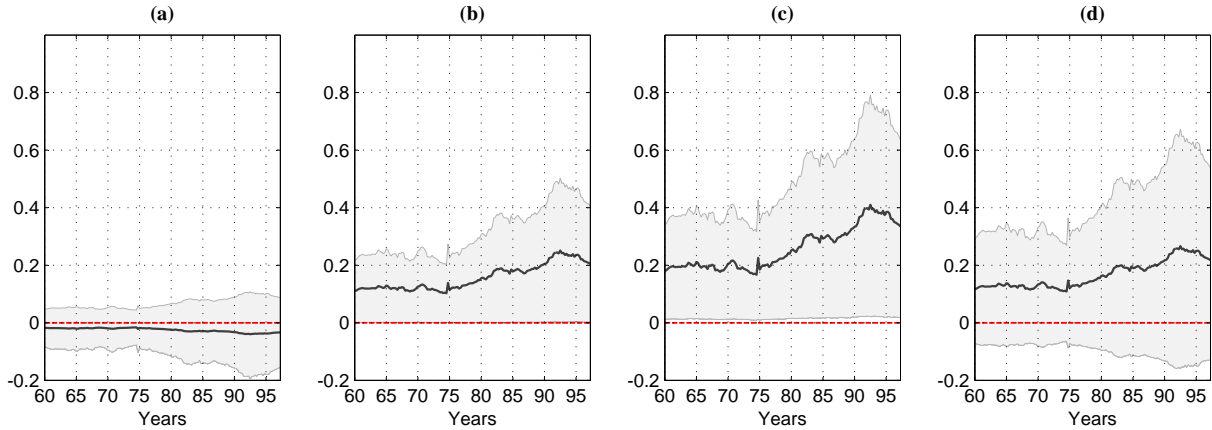


Figure 4: Mean Impulse Responses of Output to Implicit Tax Shock over the entire sample, (a) after 1 quarter, (b) after 4 quarters, (c) after 8 quarters, and (d) after 12 quarters, with 32-th and 68-th percentiles

Figure 5 shows the dynamics of the unanticipated tax shock to output over the entire sample. Most apparent is the spike in 1975, which is most profound between 4 and 12 quarters. Following this isolated spike, the response to the unanticipated shock gradually gains in magnitude over the late 1970s. The early 1980s seem to resemble a turning point of this development, as the response markedly decreases towards the end of the decade. Yet another spike in magnitude can be observed around 1991, before decreasing considerably towards the end of the sample.

Anecdotal evidence sheds light on some of the dynamics behind these shifts. According to BP, a large, isolated, well-identified tax cut temporarily took place in 1975:II, namely the Tax Reduction Act of 1975. U.S. President Gerald Ford signed this law on March 29, 1975, only two months after having called for an Economic Summit Conference to discuss policy measures to overcome an economic downturn. In face of this sudden policy change, the tax cut in 1975:II can be regarded as an unanticipated tax shock.¹⁰ The output multipliers in Figure 4 and Figure 5 support this notion. In 1975:II, a small jump is recorded for the

⁹The anticipated tax shock is not recorded at horizon 0, as the identification scheme restricts the contemporaneous reaction of output to the implicit tax rate to zero.

¹⁰This argument is consistent with timing rules from the narrative approach. For example, Mertens and Ravn (2012) classify a tax shock as unanticipated if its enactment lies within 90 days after it becomes law.

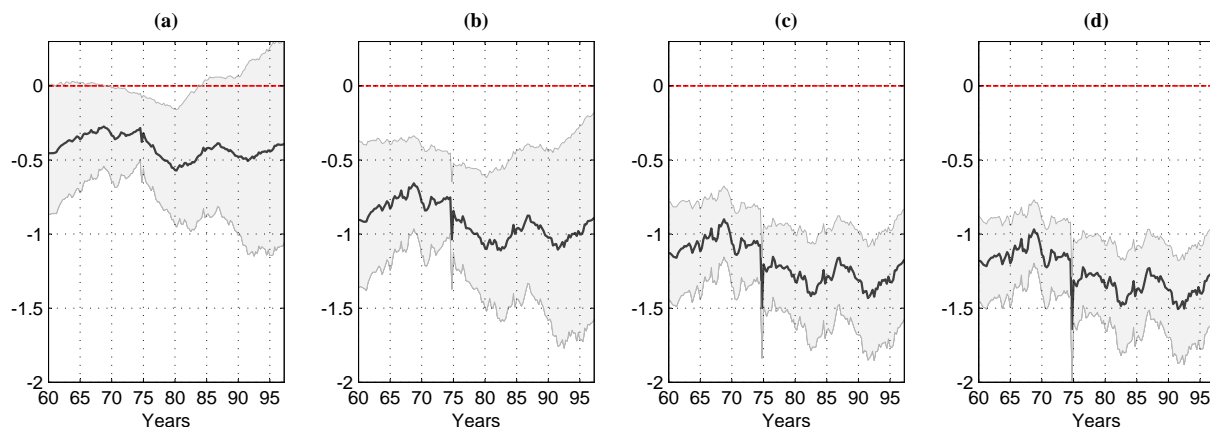


Figure 5: Mean Impulse Responses of Output to Tax Revenues Shock over the entire sample, (a) on impact, (b) after 4 quarters, (c) after 8 quarters, and (d) after 12 quarters, with 32-th and 68-th percentiles

anticipated tax shock, yet compared to the rise in magnitude during subsequent decades, this spike is quite modest. In contrast, the negative effect of an unanticipated tax shock on output plunges in 1975:II. This reveals that the effectiveness of the unanticipated tax shock temporarily increased substantially in 1975. In the following years until around 1980, both the anticipated and the unanticipated shock grow in magnitude.

The negative downward trend of the unanticipated shock comes to a halt in the early 1980s. A plausible explanation provides Ronald Reagan's presidential campaign during 1979 and 1980, which was based on announcements concerning economic reforms, including tax policy changes. Various announcements and campaign speeches helped to build up expectations of tax cuts. After his inauguration in 1981, Reagan officially called for tax relief and enacted such with the Economic Recovery Tax Act of 1981, which sought to phase-in several measures until 1984. This was followed by the announcement in early 1984 to pursue plans of structural reform – later enacted as the Tax Reform Act of 1986. These examples suggest that tax policy between 1980 and 1986 was to a higher extent accompanied by foresight. Again, Figures 4 and 5 support this notion. From early 1980, when Reagan's prospects of winning office rose, until late 1986, when the last tax reform was enacted, the magnitude of the unanticipated tax shock on output declined. Meanwhile, the effect of the anticipated tax shock grew stronger during that time, with an additional surge between 1981 and 1984.

Overall, the analysis reveals the time variation in the anticipated and the unanticipated tax shock, and anecdotal evidence indicates the importance of tax foresight on these dynamics.

3.2 Tracing the bias

To shed light on the time-varying presence of tax foresight, it is necessary to study the bias that results from the time-varying tax multipliers in the BP model along with its dynamics. Table 1 reports tax multipliers for the time-varying BP and LWY models, as well as the bias during the 1970s, 1980s, and 1990s, defined as the difference between corresponding multipliers of both models.

Table 1: Output Multipliers for an Anticipated Tax Shock (A), Unanticipated Tax Revenue Shock in the LWY model (U), and Tax Revenue Shock in the BP model (BP)

	0 qtr	4 qtr	8 qtr	12 qtr	20 qtr	Avg qtr
Average for 1970s						
BP	-0.348*	-0.813*	-1.086*	-1.153*	-1.108*	-0.902
LWY (U)	-0.378*	-0.860*	-1.146*	-1.223*	-1.157*	-0.953
LWY (A)	0.000	0.123*	0.200*	0.130	-0.042	0.082
Diffs. (BP - U)	0.030	0.047	0.060	0.069	0.049	0.051
Average for 1980s						
BP	-0.416*	-0.917*	-1.189*	-1.250*	-1.194*	-0.993
LWY (U)	-0.472*	-1.004*	-1.288*	-1.352*	-1.268*	-1.077
LWY (A)	0.000	0.176*	0.287*	0.186	-0.063	0.117
Diffs. (BP - U)	0.056	0.087	0.099	0.102	0.074	0.084
Average for 1990s						
BP	-0.388	-0.906*	-1.214*	-1.286*	-1.226*	-1.004
LWY (U)	-0.452	-1.006*	-1.318*	-1.390*	-1.303*	-1.094
LWY (A)	0.000	0.227*	0.370*	0.240	-0.079	0.152
Diffs. (BP - U)	0.064	0.100	0.104	0.104	0.077	0.090
Avg. bias	0.050	0.078	0.088	0.092	0.067	0.075

The table reports mean output multipliers averaged over the 1970s, 1980s, and 1990s. Diffs. (BP - U) gives the difference of the output multipliers to tax revenue shocks from the BP and the LWY specifications. Avg qtr denotes the average effect over all horizons. Avg bias denotes the average bias over the three decades. An asterisk (*) indicates that the interval between the 32-th/68-th percentiles does not include zero.

The results indicate an upward bias in the BP tax multipliers for all quarters over all three decades (see Diffs. (BP-U)). On average, the impact effect is biased by 0.05 during the 1970s, 1980s, and 1990s. This average rises to 0.078 after 4 quarters, when the anticipated

tax shock first becomes statistically significant (compare Figure 2). It rises further to 0.088 and 0.092 at the 8 and 12 quarter horizons, respectively. This confirms the time-invariant findings of LWY that the BP tax revenue shock understates the negative effect on output.

In addition, my results allow for an assessment of the time variations of the bias. Comparing the three decades, the 1970s show the lowest average bias over all quarters with 0.051. A sharp increase is documented for the 1980s, as the average bias over all quarters is 0.084. For the 1990s the bias rises yet again, although more modestly, to 0.09 on average over 20 quarters.

To reveal the full dynamics of the bias, Figure 6 plots the difference of tax multipliers from the two models (a) on impact and (b) after 4 quarters. For the latter case, this corresponds to the quarter in which the response of output to the implicit tax rate first becomes statistically significant. Two results stand out: First, the bias in the tax multiplier of the BP model varies substantially over time, thereby justifying my time-varying SVAR approach. Second, the notion from Table 1 is refined in the sense that the bias increases markedly from 1975 onwards, rendering the 1980s and 1990s as decades with a sustainedly higher bias.

Figure 6, in combination with findings from Yang (2005), provide an understanding of the bias' time variations and also provides an interpretation of the bias itself. Yang (2005) shows that the conventional econometric model, disregarding tax foresight, yields unbiased estimates if tax disturbances are unanticipated. Thus, a conventional VAR yields estimates close to the correct theoretical responses in the absence of foresight. Results are different, however, when tax foresight builds up due to publicly communicated tax changes. Then, conventional VARs yield biased estimates. If one acknowledges that the correct theoretical responses are only recovered by a VAR that incorporates potential future tax information, then the difference of impulse responses (i.e. the bias) of a conventional VAR and the correctly specified VAR should reveal periods of greater or lesser tax foresight. Hence time variations of the bias, as depicted in Figure 6, implicitly reflect different degrees of foresight

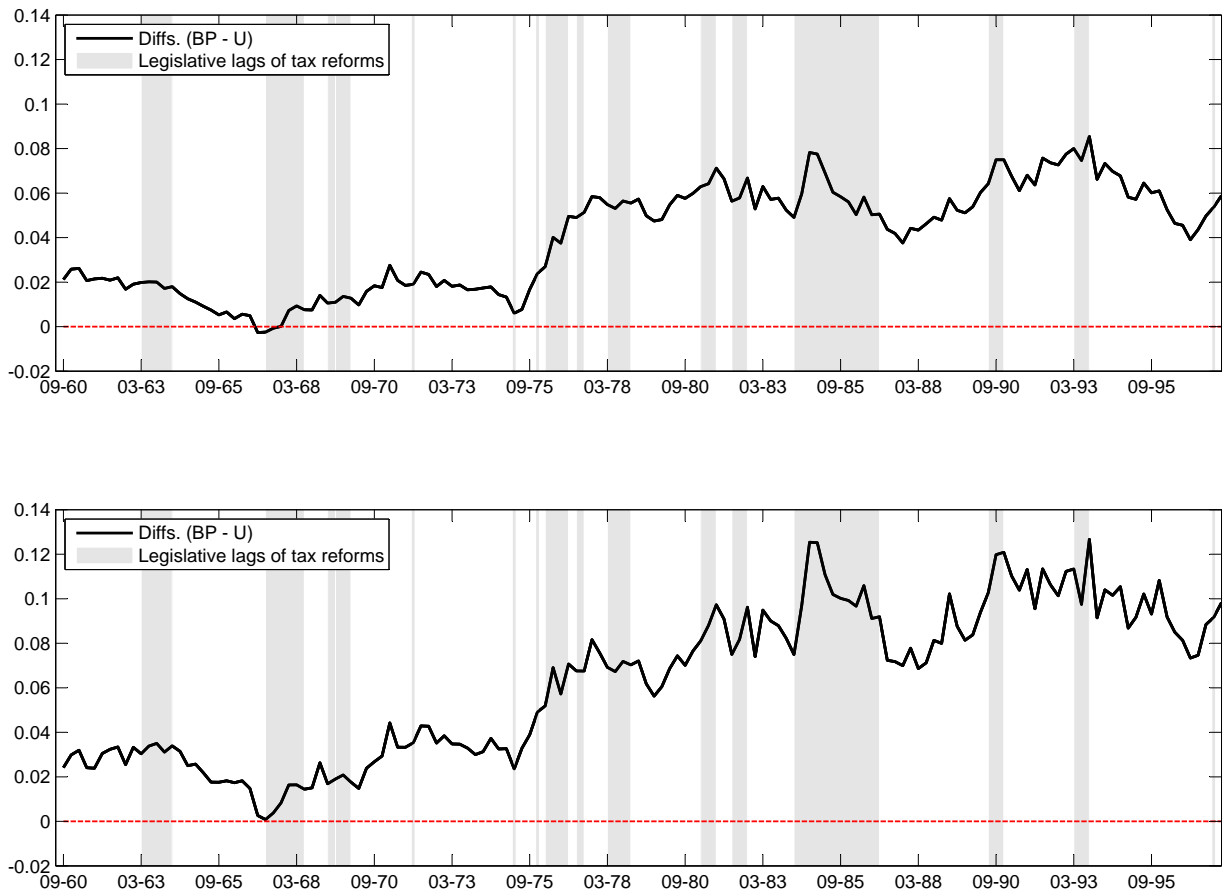


Figure 6: Differences of Output Multipliers for a Tax Revenue Shock (BP) and an Unanticipated Tax Revenue Shock (LWY(U)) on (a) impact and (b) after 4 quarters. Shaded areas indicate legislation lags of U.S. tax reforms as documented in Yang (2007).

between 1960 and 1997.

With this in mind, two features of the dynamics depicted in Figure 6 are noteworthy and warrant a discussion. First, tax foresight temporarily spiked during several tax reforms. Second, tax foresight was on average lower during the 1960s and 1970s, but higher during the 1980s and 1990s.

Leeper et al. (2012) argue that, in principle, a spike in expectations should occur when a tax policy change is announced, followed by a drop in expectations as soon as the new legislation is implemented. This follows from the rationale that political debates on tax reforms attract media attention, in turn building up expectations of future tax changes until they have been implemented. The difficulty lies in the timing of these periods. As the legislative lag, other than the inside-outside lags, can be pinned down by documented dates, it is used as proxy for the length of potential tax policy foresight. The shaded areas in Figure 6 indicate the legislative lags of U.S. tax reforms, documented in Yang (2007). For several tax reforms (e.g. in 1968, 1976, 1981, 1982, 1986, 1990 and 1993) the bias increases markedly during the legislative lag and drops towards the end of that period, or shortly thereafter. The clearest example for this dynamic is the Tax Reform Act of 1986. The degree of foresight peaks after President Reagan first signaled the tax reform in his State of the Union address on January 25, 1984, and deteriorates towards the implementation of the legislation on October 22, 1986.

The second interesting observation of Figure 6 is the difference in levels of tax foresight: The first quarter of 1975 marks a turning point for the presence of foresight. Whereas the bias is largely close to zero for years prior to 1975, it rises substantially towards the beginning of the 1980s and remains on a high level going into the 1990s. Yang (2007) points out several features of U.S. tax reforms during the time in question that help provide an explanation for the sustainingly higher bias. Accordingly, tax reforms in the 1970s were primarily motivated by efforts to stimulate the economy short term. As such, these tax

reforms were a discretionary countercyclical tool to fight recessions.¹¹ In contrast, tax policy in the 1980s and 1990s was mostly concerned with fighting the ever growing budget deficits that resulted after the tax cut in 1981.¹² As budget deficits are more persistent than the cyclically changing state of the economy, it is likely that economic agents in the 1980s had more tax foresight than in the 1970s. It is easier to expect tax reforms that address an omnipresent deficit than sudden shocks to the economy, such as the recessions in the 1970s. A second argument is given by phase-in schedules and so-called sunset (or temporary) provisions defined in tax law. Yang (2007) reveals that especially the structural tax reforms of the 1980s and 1990s made frequent use of such phase-in provisions. For example, the Economic Recovery Tax Act of 1981 and the Tax Reform Act of 1986 specified provisions that sought to phase-in measures over several years ahead. Given this embodiment of tax reforms in the 1980s and 1990s it is plausible to observe a higher degree of foresight during that time.

4 Conclusions

This paper contributes to the recent findings of Leeper et al. (2012, 2013a) that foresight is a time-varying process and that estimated tax policy shocks are biased, if they fail to account for tax foresight adequately. I extend LWY's re-evaluation of BP's prominent SVAR approach by transforming these models into time-varying SVARs with stochastic volatility. This provides insights on the time variations of anticipated and unanticipated tax policy shocks, the time shifts of the bias in the time-varying BP model, and on the degree of tax foresight during different time periods. Three findings stand out:

First, I find the expected negative effect of an unanticipated tax shock on output, and

¹¹Tax reforms associated with the rationale of stimulating the economy were: (a) the Tax Reduction Act of 1975, (b) the Revenue Adjustment Act of 1975, (c) the Tax Reduction and Simplification Act of 1977, and (d) the Revenue Act of 1978.

¹²Tax reforms motivated by raising revenues and/or reducing deficits were: (a) the Tax Equity and Fiscal Responsibility Act of 1982, (b) the Deficit Reduction Act of 1984, (c) the Omnibus Budget Reconciliation Act of 1990, and (d) the Omnibus Budget Reconciliation Act of 1993.

an expansive reaction of output to an anticipated tax shock. Second, both the anticipated and the unanticipated tax shocks show considerable movements over time. After two stable decades in the 1960s and 1970s, the anticipated tax shock rises in magnitude going into the 1980s, and again going into the early 1990s, before declining towards the end of the decade. Overall the magnitude almost doubles between the mid 1970s and the early 1990s. The unanticipated tax shock shows more fluctuations over the entire sample from 1960 to 1997. Following a temporary negative spike in 1975, the severity of the unanticipated shock rises towards the end of the decade, but declines for most parts of the 1980s. Following a temporary surge on the verge of the 1990s, severity again declines towards the end of the sample. Third, I find that the bias in BP's tax shock is largely close to zero during the 1960s and 1970s, increasing sustainedly during the late 1970s, and high during the 1980s and 1990s. This suggests that a higher degree of tax foresight was present in the 1980s and 1990s compared to the 1960s and 1970s. Anecdotal evidence of the legislative lag and motivation of documented U.S. tax reforms support my findings.

References

- Auerbach, A. J. (1989). Tax Reform and Adjustment Costs: The Impact on Investment and Market Value. *International Economic Review*, 30(4):pp. 939–962.
- Blanchard, O. and Perotti, R. (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.
- Burnside, C., Eichenbaum, M., and Fisher, J. D. (2004). Fiscal shocks and their consequences. *Journal of Economic Theory*, 115(1):89–117.
- Canova, F. (2007). *Methods for Applied Macroeconomic Research*. Volume 13. Princeton University Press.
- Canova, F. and Ciccarelli, M. (2006). Estimating multi-country VAR models. Working Paper Series 0603, European Central Bank.
- Carter, C. K. and Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81(3):541–553.
- Del Negro, M. and Primiceri, G. (2013). Time-varying structural vector autoregressions and monetary policy: a corrigendum. *Federal Reserve Bank of New York Staff Reports*, 619.
- Edelberg, W., Eichenbaum, M., and Fisher, J. D. (1999). Understanding the Effects of a Shock to Government Purchases. *Review of Economic Dynamics*, 2(1):166–206.
- Eichenbaum, M. and Fisher, J. D. M. (2005). Fiscal Policy in the Aftermath of 9/11. *Journal of Money, Credit and Banking*, 37(1):pp. 1–22.
- Fatás, A. and Mihov, I. (2001). The effects of fiscal policy on consumption and employment: theory and evidence. *CEPR Discussion Paper No. 2760*.

- Fortune, P. (1996). Do municipal bond yields forecast tax policy? *New England Economic Review*, (Sep):29–48.
- Fuller, W. A. (1996). *Introduction to Statistical Time Series*. Volume 230. John Wiley & Sons Inc.
- Galí, J., Lpez-Salido, J. D., and Valls, J. (2007). Understanding the Effects of Government Spending on Consumption. *Journal of the European Economic Association*, 5(1):pp. 227–270.
- Hall, R. E. (1971). The Dynamic Effects of Fiscal Policy in an Economy with Foresight. *The Review of Economic Studies*, 38(2):pp. 229–244.
- House, C. L. and Shapiro, M. D. (2006). Phased-In Tax Cuts and Economic Activity. *The American Economic Review*, 96(5):pp. 1835–1849.
- Judd, K. L. (1985). Short-Run Analysis of Fiscal Policy in a Simple Perfect Foresight Model. *Journal of Political Economy*, 93(2):pp. 298–319.
- Kim, S., Shephard, N., and Chib, S. (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models. *The Review of Economic Studies*, 65(3):361–393.
- Kirchner, M., Cimadomo, J., and Hauptmeier, S. (2010). Transmission of Government Spending Shocks in the Euro Area: Time Variation and Driving Forces. *ECB Working Paper No.1219*.
- Kueng, L. (2011). Tax News: Identifying the Household Consumption Response to Tax Expectations using Bond Prices. *Manuscript, University of California, Berkley*.
- Leeper, E. M., Richter, A. W., and Walker, T. B. (2012). Quantitative effects of fiscal foresight. *American Economic Journal: Economic Policy*, 4(2):115–144.
- Leeper, E. M., Walker, T. B., and Yang, S.-C. S. (2013a). Fiscal Foresight and Information Flows. *Econometrica*, 81(3):1115–1145.

- Leeper, E. M., Walker, T. B., and Yang, S.-C. S. (2013b). Supplement to “Fiscal Foresight And Information Flows”. *Econometrica*, 81(3):1115–1145.
- Mertens, K. and Ravn, M. O. (2011). Understanding the aggregate effects of anticipated and unanticipated tax policy shocks. *Review of Economic Dynamics*, 14(1):27–54. Special issue: Sources of Business Cycles.
- Mertens, K. and Ravn, M. O. (2012). Empirical Evidence on the Aggregate Effects of Anticipated and Unanticipated US Tax Policy Shocks. *American Economic Journal: Economic Policy*, 4(2):145–81.
- Mountford, A. and Uhlig, H. (2009). What are the effects of fiscal policy shocks? *Journal of Applied Econometrics*, 24(6):960–992.
- Park, S. (1997). The Relationship between Government Financial Condition and Expected Tax Rates Reflected in Municipal Bond Yields. *National Tax Journal*, 50(1):23–38.
- Pereira, M. C. and Lopes, A. S. (2014). Time-varying fiscal policy in the U.S. *Studies in Nonlinear Dynamics and Econometrics*, 18(2):157–184.
- Perotti, R. (2005). Estimating the Effects of Fiscal Policy in OECD Countries. CEPR Discussion Papers 4842, C.E.P.R. Discussion Papers.
- Poterba, J. M. (1989). Tax reform and the market for tax-exempt debt. *Regional Science and Urban Economics*, 19(3):537–562.
- Primiceri, G. E. (2005). Time varying structural vector autoregressions and monetary policy. *The Review of Economic Studies*, 72(3):821–852.
- Ramey, V. A. (2011). Identifying Government Spending Shocks: It’s all in the Timing. *The Quarterly Journal of Economics*.

- Ramey, V. A. and Shapiro, M. D. (1998). Costly capital reallocation and the effects of government spending. *Carnegie-Rochester Conference Series on Public Policy*, 48(0):145–194.
- Rathke, A. and Sarferaz, S. (2014). Malthus and the Industrial Revolution: Evidence from a Time-Varying VAR. *CEPr Working Papers*, No. 4667.
- Yang, S.-C. S. (2005). Quantifying tax effects under policy foresight. *Journal of Monetary Economics*, 52(8):1557–1568.
- Yang, S.-C. S. (2007). A Chronology Of Postwar U.S. Federal Income Tax Policy. *Indiana University Bloomington Economics Department Center for Applied Economics and Policy Research (CAEPR) Working Paper*, 21.

Appendix

A The Gibbs sampling algorithm

A.1 Gibbs sampling overview

This section gives a brief overview over the Gibbs sampler for the estimated TVP-VAR as proposed in Primiceri (2005) and Del Negro and Primiceri (2013).¹³ After initializing \mathbf{b}^T , \mathbf{d}^T , \mathbf{s}^T , \mathbf{V} , the Gibbs sampler is performed by iterating over the following steps:

Sampling $\boldsymbol{\theta}^T$: The measurement equation (6) and the transition equation (13) form a linear Gaussian state-space system. Given the data, histories of covariance and volatility states (which constitute a history of $\boldsymbol{\Sigma}_t$) and the covariance of innovations \mathbf{Q} , the algorithm of Carter and Kohn (1994) can be applied to generate a history of $\boldsymbol{\theta}$'s.

Sampling \mathbf{b}^T : The measurement equation (7) and the transition equation (14) form a linear Gaussian state-space system. Due to the lower triangular structure of \mathbf{B}_t the algorithm can be applied equation-by-equation, from top to bottom. A history of \mathbf{b} 's can be generated, given the data, histories of coefficient and volatility states, and the covariance of innovations \mathbf{S} .

Sampling \mathbf{d}^T : The measurement equation based on a linearized version of (8) and the transition equation (15) form a linear state-space system. Applying the approximation method proposed by Kim et al. (1998), renders the system Gaussian and the algorithm of Carter and Kohn (1994) can be applied. A history of \mathbf{d} 's can be generated, given the data, histories of coefficient and covariance states, and the covariance of innovations \mathbf{W} .

Sampling \mathbf{V} : Given the histories of $\boldsymbol{\Theta}$'s, \mathbf{B} 's and \mathbf{D} 's all innovations in the transition equations are observable. Thus, the covariances in $(\mathbf{Q}, \mathbf{S}, \mathbf{W})$ can be generated.

¹³In the description of the algorithm I follow the outlines of Rathke and Sarferaz (2014).

A.2 Step II - Drawing VAR parameters $\boldsymbol{\theta}^T$

Conditional on \mathbf{b}^T , \mathbf{D}^T and \mathbf{V} , the measurement equation of (6) together with the transition equation of (13) form a linear normal state-space system

$$\begin{aligned}\mathbf{y}_t &= (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n)\boldsymbol{\theta}_t + \mathbf{u}_t, \\ \boldsymbol{\theta}_t &= \boldsymbol{\theta}_{t-1} + \boldsymbol{\nu}_t.\end{aligned}$$

Therefore, the algorithm of Carter and Kohn (1994) can be applied to draw a history of $\boldsymbol{\theta}^T$.

A.3 Step III - Drawing covariance states \mathbf{b}^T

Taking $\boldsymbol{\theta}^T$ as given, it is possible to observe \mathbf{u}_t by rewriting (6) as

$$\mathbf{u}_t = \mathbf{y}_t - (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n)\boldsymbol{\theta}_t.$$

Using (8), we can express (7) as follows

$$\mathbf{A}_t\mathbf{u}_t = \mathbf{B}_t\mathbf{D}_t\boldsymbol{\varepsilon}_t. \tag{19}$$

Recall that \mathbf{A}_t contains only known elements, and that it is the estimation of \mathbf{B}_t that is of interest. Therefore, we can rewrite (19) as

$$\hat{\mathbf{y}}_t = \mathbf{B}_t\mathbf{D}_t\boldsymbol{\varepsilon}_t, \tag{20}$$

where $\hat{\mathbf{y}}_t = \mathbf{A}_t\mathbf{u}_t$.¹⁴ Due to the identification structures imposed in (11) and (12), \mathbf{B}_t is a lower triangular matrix with ones on the main diagonal. Therefore, (20) can be written as

$$\hat{\mathbf{y}}_t = \mathbf{L}_t\mathbf{b}_t + \mathbf{D}_t\boldsymbol{\varepsilon}_t, \tag{21}$$

¹⁴The conversion of (19) to (20) shows the equivalence to the covariance states estimation in Primiceri (2005), as long as all elements of \mathbf{A}_t are known.

where \mathbf{L}_t has the following structure

$$\mathbf{L}_t = \begin{pmatrix} 0 & \cdots & \cdots & 0 \\ \hat{y}_{1,t} & 0 & \cdots & \vdots \\ 0 & (\hat{y}_{1,t} \ \hat{y}_{2,t}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & (\hat{y}_{1,t} \ \hat{y}_{2,t} \ \cdots \ \hat{y}_{n,t}) \end{pmatrix}.^{15} \quad (22)$$

Together with the transition equation $\mathbf{b}_t = \mathbf{b}_{t-1} + \boldsymbol{\zeta}_t$ from (14) this completes a Gaussian state-space system. However, some dependent variables in the measurement equation also appear on the R.H.S of the equation, rendering the system nonlinear. The additional assumption of \mathbf{S} being block diagonal solves this problem, allowing the algorithm of Carter and Kohn (1994) to be applied equation by equation, starting with the second equation. Note that in this case not only does $\hat{y}_{i,t}$ not show up on the R.H.S. of the i -th equation, but due to the triangular structure, the vectors $\hat{y}_{i,t}$ in matrix \mathbf{L}_t can be treated as predetermined in the same equation.¹⁶

A.4 Steps IV and V - Drawing volatility states \mathbf{D}^T

With $\boldsymbol{\theta}_t$ and \mathbf{B}_t given, and all elements of \mathbf{A}_t known, we can combine equations (7) and (8) to observe \mathbf{e}_t ,

$$\mathbf{e}_t = \mathbf{B}_t^{-1} \mathbf{A}_t (\mathbf{y}_t - (\mathbf{Z}'_{t-1} \otimes \mathbf{I}_n) \boldsymbol{\theta}_t) = \mathbf{D}_t \boldsymbol{\varepsilon}_t. \quad (23)$$

Note that this system of measurement equations is nonlinear. Squaring and taking logarithms of every element converts the system to be linear,

$$\log(y_{i,t}^*)^2 = 2 \log d_{i,t} + \log \varepsilon_{i,t}^2, \quad i = 1, \dots, n, \quad (24)$$

¹⁵In contrast to Primiceri (2005), who estimates the parameters of \mathbf{A}_t , we estimate \mathbf{B}_t on the R.H.S of (20). Therefore, the \mathbf{L}_t matrix contains positive values of $\hat{y}_{i,t}$.

¹⁶See the Appendix of Primiceri (2005) for a more detailed discussion of this estimation block.

where we define $y_{i,t}^* = e_{i,t}$. Since $(y_{i,t}^*)^2$ can take on very small values, an offset constant is used to make the estimation procedure more robust. The approximated linear state-space form is

$$\mathbf{y}_t^{**} = 2\mathbf{h}_t + \boldsymbol{\vartheta}_t \quad (25)$$

$$\mathbf{h}_t = \mathbf{h}_{t-1} + \boldsymbol{\eta}_t, \quad (26)$$

where $y_{i,t}^{**} = \log[(y_{i,t}^*)^2 + \bar{c}]$, $h_{i,t} = \log d_{i,t}$, and $\vartheta_{i,t} = \log \varepsilon_{i,t}^2$. \bar{c} denotes an offset constant set to 0.001 as introduced by Fuller (1996, pp. 494-7). Note that this linear state-space system is not Gaussian, as the innovations of the measurement equation are distributed as $\log \chi^2(1)$. To transform the system into a Gaussian one, each element of $\boldsymbol{\vartheta}_t$ is approximated by a mixture of normal densities as described in Kim et al. (1998). Since the covariance matrix of $\boldsymbol{\varepsilon}_t$ is an identity matrix, the covariance matrix of $\boldsymbol{\vartheta}_t$ is also diagonal, making the approximation possible for each element of $\boldsymbol{\vartheta}_t$. Kim et al. (1998) match a number of moments of the $\log \chi^2(1)$ distribution with a mixture of seven normal densities. The constants to match the moments are the component probabilities q_j , means $m_j - 1.2704$, and variances v_j^2 , $j = 1, \dots, 7$, as reported in Table 2. The approximation of each element of ϑ_t is

$$f(\vartheta_{i,t}) \approx \sum_{j=1}^7 q_j f_N((\vartheta_{i,t} | m_j - 1.2704, v_j^2)). \quad (27)$$

An alternative formulation of (27) is

$$\vartheta_{i,t} | s_{i,t} = j \sim N(m_j - 1.2704, v_j^2), \quad (28)$$

$$\Pr(s_{i,t} = j) = q_j, \quad (29)$$

where $\mathbf{s}^T = [\mathbf{s}_1, \dots, \mathbf{s}_T]$ is a matrix of unobserved states $s_{i,t} \in 1, \dots, 7$ that indicate which member of the normal distribution mixture is used for the approximation of each $\vartheta_{i,t}$. Conditional on \mathbf{y}^{**T} and \mathbf{h}^T , it is possible to sample each $s_{i,t}$ independently from the probability

mass function defined by

$$\Pr(s_{i,t} = j | y_{i,t}^{**}, h_{i,t}) \propto q_j f_N(y_{i,t}^{**} | 2h_{i,t} + m_j - 1.2704, v_j^2), \quad (30)$$

with $j = 1, \dots, 7$, $i = 1, \dots, n$, and $t = 1, \dots, T$. Conditional on $\mathbf{y}^T, \boldsymbol{\theta}^T, \mathbf{b}^T, \mathbf{s}^T$, and \mathbf{V} , and the normal approximation as in Kim et al. (1998), the measurement equation of (25) and the transition equation of (27) form a linear Gaussian state-space system, to which the algorithm of Carter and Kohn (1994) can be applied.

Table 2: Selection of the mixing distribution to be $\log \chi^2(1)$

ω	$q_j = \Pr(\omega = j)$	m_j	v_j^2
1	0.00730	-10.12999	5.79596
2	0.10556	-3.97281	2.61369
3	0.00002	-8.56686	5.17950
4	0.04395	2.77786	0.16735
5	0.34001	0.61942	0.64009
6	0.24566	1.79518	0.34023
7	0.25750	-1.08819	1.26261

Source: Kim et al. (1998).

A.5 Step VI - Sampling Covariances \mathbf{V}

The prior and conditional posterior distributions of \mathbf{Q}, \mathbf{W} and the blocks on the diagonal of \mathbf{S} are inverse-Wishart. Conditional on $\boldsymbol{\theta}^T, \mathbf{b}^T$ and \mathbf{h}^T the innovations $\boldsymbol{\nu}_t, \boldsymbol{\zeta}_t$ and $\boldsymbol{\eta}_t$ in the transition equations are observable, allowing to draw the hyperparameters from these inverse-Wishart posteriors.