Lütkepohl, Helmut; Shang, Fei; Uzeda, Luis; Woźniak, Tomasz

Working Paper

Partial identification of heteroskedastic structural VARs: Theory and Bayesian inference

DIW Discussion Papers, No. 2081

Provided in Cooperation with:
German Institute for Economic Research (DIW Berlin)

Suggested Citation: Lütkepohl, Helmut; Shang, Fei; Uzeda, Luis; Woźniak, Tomasz (2024) : Partial identification of heteroskedastic structural VARs: Theory and Bayesian inference, DIW Discussion Papers, No. 2081, Deutsches Institut für Wirtschaftsforschung (DIW), Berlin

This Version is available at:
https://hdl.handle.net/10419/295170

DIW BERLIN

2081

Discussion Papers

Deutsches Institut für Wirtschaftsforschung 2024

Partial Identification of Heteroskedastic Structural VARs: Theory and Bayesian Inference

Helmut Lütkepohl, Fei Shang, Luis Uzeda and Tomasz Woźniak

IMPRESSUM

DIW Berlin, 2024

DIW Berlin
German Institute for Economic Research
Mohrenstr. 58
10117 Berlin

Tel. +49 (30) 897 89-0
Fax +49 (30) 897 89-200
https://www.diw.de

ISSN electronic edition 1619-4535

Papers can be downloaded free of charge from the DIW Berlin website:
https://www.diw.de/discussionpapers

Discussion Papers of DIW Berlin are indexed in RePEc and SSRN:
https://ideas.repec.org/s/diw/diwwpp.html
https://www.ssrn.com/link/DIW-Berlin-German-Inst-Econ-Res.html

Partial Identification of Heteroskedastic Structural VARs: Theory and Bayesian Inference

Helmut Lütkepohl[a], Fei Shang[b], Luis Uzeda[c,*], Tomasz Woźniak[d,**]

[a] *Freie Universität Berlin & DIW Berlin*
[b] *South China University of Technology & Yuexiu Capital Holdings Group*
[c] *Bank of Canada*
[d] *University of Melbourne*

Abstract

We consider structural vector autoregressions identified through stochastic volatility. Our focus is on whether a particular structural shock is identified by heteroskedasticity without the need to impose any sign or exclusion restrictions. Three contributions emerge from our exercise: (i) a set of conditions under which the matrix containing structural parameters is partially or globally unique; (ii) a statistical procedure to assess the validity of the conditions mentioned above; and (iii) a shrinkage prior distribution for conditional variances centred on a hypothesis of homoskedasticity. Such a prior ensures that the evidence for identifying a structural shock comes only from the data and is not favoured by the prior. We illustrate our new methods using a U.S. fiscal structural model.

Keywords: Identification Through Heteroskedasticity, Stochastic Volatility, Non-centred Parameterisation, Shrinkage Prior, Normal Product Distribution, Tax Shocks
JEL classification: C11, C12, C32, E62

*The views expressed in this paper are those of the authors and do not necessarily reflect the position of the Bank of Canada.

**Corresponding author. *Address:* Department of Economics, University of Melbourne, 111 Barry Street, FBE Building, Level 4, 3053 Carlton, Victoria, Australia, *Email address:* tomasz.wozniak@unimelb.edu.au.

1. Introduction

This paper considers the partial identification of a structural shock in a multivariate setup that is in line with the definition by Rubio-Ramírez, Waggoner and Zha (2010). This definition states that a structural shock is identified when the parameters of its corresponding equation within a system are globally identified, that is, up to being sign–normalised as in Waggoner and Zha (2003b). Partial identification is essential in empirical analyses using Structural Vector Autoregressions (SVARs) and focusing on fewer shocks than there are variables in the model. For example, it is often the case that one is interested in the identification of a specific shock, such as a monetary or a fiscal policy shock. Moreover, partial identification becomes even more important in larger systems of variables that, on the one hand, improve the forecasting performance of the model, resulting in more realistic impulse responses but, on the other one, increase the number of shocks that are not necessarily of interest or difficult to interpret (see Carriero, Clark and Marcellino, 2019).

In our approach, the source of partial identification is conditional heteroskedasticity that offers the potential to identify all the parameters of a given equation up to a sign following the seminal developments proposed by Rigobon (2003). We choose a specific model for conditional variances, namely Stochastic Volatility (SV) as proposed by Cogley and Sargent (2005), and in line with the identification ideas put forth by recent studies, such as Lewis (2021) and Bertsche and Braun (2022). Not only does this choice offer a flexible approach to address identification, but it also has been shown to be a key extension of homoskedastic SVARs that leads to improved forecasting performance (see, e.g., Clark and Ravazzolo, 2015; Chan and Eisenstat, 2018).

The main contribution of this paper is a general condition for the partial identification of a structural shock via heteroskedasticity. This condition states that a structural shock is identified up to sign if the sequence of its conditional variances is distinct from and not proportional to those for all other shocks. Our condition covers both conditional and unconditional heteroskedasticity and most of the heteroskedastic models used in

empirical studies. It is expressed explicitly in terms of conditional variances, simplifying the proof and granting a straightforward interpretation. In these respects, it stands out from the existing conditions on which we provide more details in Section 2. We further show that a shock having such a unique sequence of (conditional) variances leads to globally identified impulse response functions.

Importantly, the conditions for partial identification that we derive can be verified. In this regard, a second contribution of this paper is the development of a Savage-Dickey Density Ratio (SDDR) for the hypothesis of homoskedasticity that extends the procedure by Chan (2018) for univariate SV models to multivariate structural models and generalising it. This is achieved by proposing a new prior distribution for the parameter that is associated with changes in the latent log volatilities. Notably, such a parameter plays a key role in our framework as it allows us to verify if a particular shock of interest can be identified through heteroskedasticity. For convenience, we refer to it as volatility of the log-volatility parameter hereafter. The prior for this parameter is centred at the hypothesis of homoskedasticity and follows a hierarchical structure determining the level of shrinkage and ensuring the verification to be based on less arbitrary choices. Given the flexibility of our setup, we provide conditions under which the analysis using the SDDR is feasible. Our verification procedure generalises that by Lütkepohl and Woźniak (2020) to SVARs with a more flexible process for conditional volatility, namely the SV model. Notably, our approach verifies the conditions assumed to hold for the identification test proposed by Lewis (2021).

A third contribution of this paper is the provision of a detailed analysis of the marginal prior distribution of the conditional variances implied by the normality of the SV process innovations, our new prior for the volatility of the log-volatility parameter, and the non-centred SV process parameterisation proposed by Kastner and Frühwirth-Schnatter (2014) that we adapt to the SVAR context. We show that the marginal prior for conditional variances is centred around the hypothesis of homoskedasticity and exhibits strong shrinkage towards it while maintaining heavy tails. These features are essential for SVAR models with identification via

heteroskedasticity to be verified for two reasons. First, they standardise the SVAR model facilitating the identification and estimation of conditional variances and the structural matrix. Secondly, this setup requires the evidence in favour of heteroskedasticity to come from the data. We also point out a problem of a standard prior setup for the SV model in heteroskedastic SVARs used by Cogley and Sargent (2005) and more recently by Chan, Koop and Yu (2024).

Estimation is conducted using a Gibbs sampler that implements optimal techniques such as the structural matrix sampler by Waggoner and Zha (2003a), autoregressive slope row-by-row sampling by Chan et al. (2024), auxiliary mixture by Omori, Chib, Shephard and Nakajima (2007), and the ancillarity-sufficiency interweaving strategy by Kastner and Frühwirth-Schnatter (2014) granting efficient simulation smoothing when the heteroskedasticity is uncertain. All estimation procedures are accessible via the **R** package **bsvars** by Woźniak (2024a,b) implementing our algorithms in compiled code using **C++**, which speeds up the computations by orders of magnitude. Additionally, the package offers a wide range of structural and predictive analyses using heteroskedastic SVARs.

We illustrate our methods by studying the sources of identification of unanticipated tax shocks using a structural model by Mertens and Ravn (2014). We find evidence for identification via heteroskedasticity in an extended sample covering the period until 2023. In a recent paper, Lewis (2021) also considers identification via time-varying volatility to estimate the effects of unanticipated tax changes for Mertens and Ravn's data. However, he uses a frequentist approach and, hence, his methods differ fundamentally from those developed in the following.

Our paper is closely related to a number of studies that pursue identification through heteroskedasticity using different techniques. For example, Lütkepohl and Milunovich (2016) investigate identification by testing a heteroskedastic rank defined as the number of independent heteroskedastic processes following a Generalised Autoregressive Conditional Heteroskedasticity (GARCH) model. Lanne and Luoto (2021) propose a test for the validity of moment conditions based on kurtosis of the structural shocks that are

in line with their non-normality or heteroskedasticity. Lewis (2021) proposes a test based on the autocorrelation of the reduced-form residuals that assumes non-proportional changes in volatilities of the structural shocks. Lütkepohl, Meitz, Netšunajev and Saikkonen (2021) propose a test for identification through heteroskedasticity for a two-regime volatility model when the timing of the change is known. Finally, Lütkepohl and Woźniak (2020) develop th SDDR to verify identification via Markov-switching heteroskedasticity in a model with an arbitrary number of regimes.

The remainder of this paper is as follows. Section 2 introduces the SVAR model and the identification results, with proofs relegated to the appendix. Section 3 parametrises the conditional heteroskedasticity using the SV model while the underlying priors for the volatility process are presented in Section 4. The SDDR for the verification of the identification via heteroskedasticity is introduced in Section 5. The priors and estimation of the remaining groups of parameters is scrutinised in Section 6 and the appendix, respectively. Finally, we illustrate our methodology by investigating identification of unanticipated tax shocks in Section 7.

Notation

We use the following notation: $\mathbf{y}$ denotes the available data, $\mathbf{I}_N$ is the identity matrix of order N, $\mathbf{0}_{N\times N}$ and $\boldsymbol{\iota}_N$ are a matrix of zeros and a vector of ones of the indicated dimensions, respectively, the operator diag$(\cdot)$ puts the vector provided as its argument on the main diagonal of a diagonal matrix of the order corresponding to the dimension of the vector, an indicator function $\mathcal{I}(\cdot)$ takes the value of 1 if the condition provided as the argument holds and 0 otherwise, $\otimes$ denotes the Kronecker product of matrices. $A \setminus B$ defines the set with all elements of the set A which are not in the set B. $\Gamma(\cdot)$ denotes the gamma function, and $K_n(\cdot)$ denotes the modified Bessel function of the second kind. The following notation is used for statistical distributions: $\mathcal{N}$ stands for a univariate normal and $\mathcal{N}_N$ for the N-variate normal distribution. $\mathcal{NP}$ stands for a univariate normal product while log $\mathcal{NP}$ for the univariate log normal product distribution (to be defined in Section 4). The gamma distribution is denoted by $\mathcal{G}$, the inverted gamma 2 by $\mathcal{IG}2$, and the uniform distribution

by $\mathcal{U}$. Unless specified otherwise, n goes from 1 to N, t goes from 1 to T, and s goes from 1 to S.

2. Partial identification in heteroskedastic SVARs

In this section we establish results for partial identification of structural parameters that are applicable within a broad class of heteroskedastic SVAR models. Some theoretical results in this section can be interpreted as general matrix results and thus extendable to different modeling frameworks. Similarly general identification results can be found in Lewis (2021). These, however, are based on moments formulated in such a way so as to become testable in a frequentist setup. In contrast, we state our conditions so as to facilitate Bayesian analysis. To make notation less cumbersome, Theorem 1, Corollary 1 and Corollary 2 are presented using SVAR-consistent notation.

2.1. Model

We begin by describing two common representations for SVAR models. Consider first the following reduced-form equation:

$$\mathbf{y}_t = \mathbf{A}_1\mathbf{y}_{t-1} + \cdots + \mathbf{A}_p\mathbf{y}_{t-p} + \mathbf{A}_d\mathbf{d}_t + \mathbf{u}_t, \tag{1}$$

where $\mathbf{y}_t$ is an N-dimensional vector of observable time series variables, $\mathbf{A}_i$, $i = 1, \ldots, p$, are $N \times N$ autoregressive coefficient matrices, $\mathbf{d}_t$ is a $d \times 1$ vector containing deterministic terms such as the intercept, trend variables, or dummies, $\mathbf{A}_d$ is the corresponding $N \times d$ matrix of coefficients, and $\mathbf{u}_t = (u_{1.t}, \ldots, u_{N.t})'$ is an N-dimensional, zero-mean, serially uncorrelated error term.

The structural equations introduce a linear relationship between the reduced-form innovations, $\mathbf{u}_t$, and the structural shocks, $\mathbf{w}_t$, via the $N \times N$ contemporaneous effects matrix $\mathbf{B}_0$,

$$\mathbf{B}_0\mathbf{u}_t = \mathbf{w}_t, \tag{2}$$

where the structural shocks are additionally contemporaneously uncorrelated. Depending on the model used, the time-varying covariances, $\mathbb{E}[\mathbf{w}_t\mathbf{w}_t']$, or conditional covariances, $\mathbb{E}[\mathbf{w}_t\mathbf{w}_t'|\mathbf{w}_{t-1}, \mathbf{w}_{t-2}, \dots]$, of $\mathbf{w}_t$ are denoted by $\mathbf{\Lambda}_t = \text{diag}\left(\sigma_{1.t}^2, \dots, \sigma_{N.t}^2\right)$, where the $\sigma_{n.t}^2$ are the unconditional or conditional variances. We also consider an alternative structural form of the model:

$$\mathbf{u}_t = \mathbf{B}\mathbf{w}_t, \tag{3}$$

where $\mathbf{B} = \mathbf{B}_0^{-1}$ is a nonsingular $N \times N$ structural matrix that represents the impact effects of the structural errors $\mathbf{w}_t$ on the observed variables. In other words, $\mathbf{B}$ represents the zero-horizon impulse responses.

Both structural-form equations (2) and (3) have the same reduced-from equation (1). The assumptions for $\mathbf{w}_t$ imply that the unconditional or conditional covariance matrices of the residuals $\mathbf{u}_t$, $\mathbb{E}\left[\mathbf{u}_t\mathbf{u}_t'\right]$ and $\mathbb{E}\left[\mathbf{u}_t\mathbf{u}_t'|\mathbf{u}_{t-1}, \mathbf{u}_{t-2}, \dots\right]$ respectively, may be time-varying and are denoted by $\mathbf{\Sigma}_t$.

2.2. Identification

It is well known that the structural matrices $\mathbf{B}_0$ or $\mathbf{B}$ are not identified without additional restrictions. Below, we state the general conditions for partial identification of some of the parameters of these matrices as the following matrix result.

Theorem 1. *Let $\mathbf{\Sigma}_t$, $t = 0, 1, \dots$, be a sequence of positive definite $N \times N$ matrices and $\mathbf{\Lambda}_t = \text{diag}\left(\sigma_{1.t}^2, \dots, \sigma_{N.t}^2\right)$ a sequence of $N \times N$ diagonal matrices with $\mathbf{\Lambda}_0 = \mathbf{I}_N$. Suppose there exists a nonsingular $N \times N$ matrix $\mathbf{B}$ such that*

$$\mathbf{\Sigma}_t = \mathbf{B}\mathbf{\Lambda}_t\mathbf{B}', \qquad t = 0, 1, \dots. \tag{4}$$

Let $\boldsymbol{\sigma}_n^2 = (1, \sigma_{n.1}^2, \sigma_{n.2}^2, \dots)$ be a possibly infinite dimensional vector. Then the n^{th} column of $\mathbf{B}$ is unique up to sign if $\boldsymbol{\sigma}_n^2 \neq \boldsymbol{\sigma}_i^2 \quad \forall i \in \{1, \dots, N\}\backslash\{n\}$.

Proof. The proof is given in Appendix A.1. It proceeds by showing that, under the conditions of the theorem, any other nonsingular $N \times N$ matrix $\mathbf{B}_*$ satisfying equation (4) has the same n^{th} column as $\mathbf{B}$ except that all signs may be reversed. □

Theorem 1 has obvious implications for models set up as in (3). It implies that, if the n^{th} component of $\mathbf{w}_t$ has a sequence of variances $\boldsymbol{\sigma}_n^2 = (1, \sigma_{n.1}^2, \sigma_{n.2}^2, \dots)$ which is different from the variance sequence of any other component of $\mathbf{w}_t$, the n^{th} column of $\mathbf{B}$ will be identified up to a sign. The theorem generalizes identification results for some special volatility models that have been used in the literature on identification through heteroskedasticity (see, e.g., Kilian and Lütkepohl, 2017, Chapter 14). For example, it is easy to see that identification results for volatility models based on a finite number of volatility regimes as considered by Rigobon (2003), Rigobon and Sack (2003), Lanne and Lütkepohl (2008), Lanne, Lütkepohl and Maciejowska (2010), Netšunajev (2013), Herwartz and Lütkepohl (2014), Woźniak and Droumaguet (2015), Lütkepohl and Velinov (2016), Lütkepohl and Netšunajev (2017) are special cases of Theorem 1.

As we will consider SV models in the following, it is important to mention that Theorem 1 applies for such models. In this context, SV models have also been proposed by Lewis (2021) and Bertsche and Braun (2022). In such cases, the conditional covariance matrices of the reduced form errors for model (3) are given by

$$\boldsymbol{\Sigma}_t = \mathbf{B}\boldsymbol{\Lambda}_t\mathbf{B}',$$

where $\boldsymbol{\Lambda}_t = \text{diag}\left(\sigma_{1.t}^2, \dots, \sigma_{N.t}^2\right)$ is a diagonal matrix. If the $\sigma_{n.t}^2$ vary stochastically, as in SV dynamics, they will not be proportional with probability 1 and, hence, satisfy the conditions for identification of Theorem 1. So if any one of the structural errors has changing conditional variances, it will be identified, even if all the other components have constant conditional variance. We will use that insight in our Bayesian analysis of the SV model. It may be worth noting, however, that Theorem 1 also implies that a single shock may be homoskedastic and still be identified in case all other shocks are heteroskedastic. This discussion also shows that Theorem 1 generalises results for full identification in Sentana and Fiorentini (2001), Lewis (2021), and Bertsche and Braun (2022) to the case of partial identification.

As we will consider the structural form setup (2), it is useful to know that Theorem 1

implies that a single row of $\mathbf{B}_0$ is identified if the corresponding error term has a variance vector which is different from the variance vectors of all other equations. More precisely, Theorem 1 implies the following result:

Corollary 1. *Let* $\boldsymbol{\Sigma}_t$, $\boldsymbol{\Lambda}_t$, $t = 0, 1, \ldots$, *and* $\boldsymbol{\sigma}_n^2 = (1, \sigma_{n.1}^2, \sigma_{n.2}^2, \ldots)$, $n = 1, \ldots, N$, *be as in Theorem 1. Suppose there exists a nonsingular* $N \times N$ *matrix* $\mathbf{B}_0$ *such that*

$$\boldsymbol{\Sigma}_t = \mathbf{B}_0^{-1} \boldsymbol{\Lambda}_t \mathbf{B}_0^{-1\prime}, \quad t = 0, 1, \ldots. \tag{5}$$

Then the n^{th} *row of* $\mathbf{B}_0$ *is unique up to sign if* $\boldsymbol{\sigma}_n^2 \neq \boldsymbol{\sigma}_i^2 \quad \forall i \in \{1, \ldots, N\} \backslash \{n\}$.

Proof. – See Appendix A.2. □

Note that the vectors $\boldsymbol{\sigma}_j^2$ contain value one for the variance in period 0. This specific parameterisation gives the elements of vector $\boldsymbol{\sigma}_j^2$ the interpretation of variances relative to the variances for $t = 0$. We are using relative variances in our theorem and corollary because it makes it easier to state the result, is in line with our standardisation of the structural model, and also leads directly to the verification procedure discussed in Section 5.

The corollary generalises Theorem 1 of Lütkepohl and Woźniak (2020). It provides a general result on identification of a single equation through heteroskedasticity. It shows that a structural shock and, hence, the corresponding structural equation is identified if the sequence of variances is distinct from the variance sequences of any of the other shocks.

2.3. Impulse responses

Structural impulse responses are computed from the reduced-form impulse responses $\boldsymbol{\Phi}_i$, $i = 0, 1, \ldots$, which are obtained by the following recursions from the p^{th}-order reduced-

form VAR slope coefficients:

$$\mathbf{\Phi}_i = \begin{cases} \mathbf{I}_N \text{ for } i = 0, \\ \sum_{j=1}^{i} \mathbf{A}_j \mathbf{\Phi}_{i-j} \text{ for } i = 1, 2, \ldots, \end{cases} \tag{6}$$

where $A_j = 0$ for $j > p$ (Lütkepohl, 2005, Section 2.1.2). The structural impulse responses are the elements of the matrices $\mathbf{\Theta}_i = \mathbf{\Phi}_i \mathbf{B}_0^{-1} = \mathbf{\Phi}_i \mathbf{B}$, $i = 0, 1, \ldots$. Thus, for computing them, the structural matrices $\mathbf{B}_0$ or $\mathbf{B}$ are needed. In particular, if just one shock is identified through heteroskedasticity, the following result formalises the implications of Corollary 1 for impulse response analysis:

Corollary 2. *If the n^{th} row of $\mathbf{B}_0$ is identified and, hence, unique in model (2), then the n^{th} column of $\mathbf{B}_0^{-1}$ is unique and the structural impulse responses can be obtained by right-multiplying the matrices $\mathbf{\Phi}_i$ by the n^{th} column of $\mathbf{B}_0^{-1}$.*

Proof. See Appendix A.3. □

We present theoretical results for impulse responses based on the SVAR in (2) since this is the specification used for the empirical application in this paper. However, it is straightforward to verify that – for the SVAR setup in (3) – if just one shock is uniquely identified and the corresponding column of the $\mathbf{B}$ matrix is given, then all the responses to the n^{th} shock are uniquely obtained by right-multiplying the $\mathbf{\Phi}_i$ matrices by the n^{th} column of $\mathbf{B}$.

3. SVAR model with stochastic volatility

Our empirical model is the SVAR model corresponding to equation (2), that is, the specification with the structural matrix $\mathbf{B}_0$. The model below supplements this specification with the assumption of conditionally normally distributed and heteroskedastic error terms with conditional variances following the SV process. Conditional heteroskedasticity has been proven an essential extension of SVARs facilitating the model identification (e.g., Bertsche and Braun (2022)).

We make an additional assumption for the structural shocks $w_{n.t}$ from equation (2) to be jointly conditionally normally distributed, given the past observations of vector $\mathbf{y}_t$, with zero means and conditional variances $\sigma^2_{n.t}$:

$$\mathbf{w}_t|\mathbf{y}_{t-1}, \mathbf{y}_{t-2}, \cdots \sim \mathcal{N}_N\left(\mathbf{0}_{(N\times 1)}, \text{ diag}\left(\sigma^2_{1.t}, \ldots, \sigma^2_{N.t}\right)\right). \tag{7}$$

The joint normality and a diagonal covariance matrix imply conditional independence between the structural shocks.

In our setting, the conditional variances are specified by an SV process that is expressed in its non-centred parametrisation where the structural shocks are decomposed into conditional standard deviations and a standardised normal error term:

$$w_{n.t} = \sqrt{\sigma^2_{n.t}}\epsilon_{n.t}, \tag{8}$$

$$\epsilon_{n.t} \sim \mathcal{N}(0, 1). \tag{9}$$

The conditional variances are defined as the exponent of the volatility of the log-volatility parameter, ω_n, and the latent log-volatility term, $h_{n.t}$:

$$\sigma^2_{n.t} = \exp\left\{\omega_n h_{n.t}\right\}, \tag{10}$$

where $h_{n.t}$ follows a zero-mean stationary autoregressive process of order one with autoregressive parameter $\rho_n \in (-1, 1)$ and standardised normal shocks, $v_{n.t}$:

$$h_{n.t} = \rho_n h_{n.t-1} + v_{n.t}, \tag{11}$$

$$v_{n.t} \sim \mathcal{N}(0, 1). \tag{12}$$

We further assume that $h_{n.0} = 0$ implying that $\sigma^2_{n.0} = 1$. This non-centred parameterisation of state-space models was introduced by Frühwirth-Schnatter and Wagner (2010) and adapted for the SV model by Kastner and Frühwirth-Schnatter (2014) and Chan (2018).

Such a model specification ensures that the specification of the conditional variances

is in line with that from Theorem 1. Note that, given equations (11) and (12), the initial condition is equal to the unconditional expected value of $h_{n.t}$ that is equal to 0. This, however, does not fix the unconditional expected value of the conditional variances $\sigma^2_{n.t}$ that depends on ω_n and ρ_n. Therefore, we will standardise the system in a different way. Our standardisation uses a non-centred SV specification and a prior that shrinks the model towards homoskedasticity, as explained in Section 4.

The non-centred specification implies a centred specification via the one-to-one relationships $\tilde{h}_{n.t} = \omega_n h_{n.t}$, $\omega_n = \pm\sqrt{\sigma^2_{v.n}}$, and $\tilde{v}_{n.t} = \omega_n v_{n.t}$. Then equations (10), (11), and (12) can be rewritten respectively as:

$$\sigma^2_{n.t} = \exp\left\{\tilde{h}_{n.t}\right\} \tag{13}$$

$$\tilde{h}_{n.t} = \rho_n \tilde{h}_{n.t-1} + \tilde{v}_{n.t} \tag{14}$$

$$\tilde{v}_{n.t} \sim \mathcal{N}\left(0, \sigma^2_{v.n}\right). \tag{15}$$

These one-to-one relationships might suggest that these two specifications are equivalent. However, complementing them with commonly used prior distributions for the parameters ω_n and $\sigma^2_{v.n}$ reveals essential differences. Our choice of the non-centred parameterisation of the SV process for SVARs identified via heteroskedasticity is motivated by (i) its potential to provide the standardisation of the model, (ii) the feasibility of specifying a prior distribution for the conditional variances that is centred at a point corresponding to homoskedasticity of the structural shocks, and (iii) the possibility to verify the homoskedasticity hypothesis for each of the shocks, which helps to assess partial identification of the system. The latter is essential to determine which shocks are identified through heteroskedasticity. All of these features can be understood by the analysis of the prior distribution for the parameters and the implied prior for the conditional variances that we analyse in Section 4.

4. The prior for stochastic volatility

We propose a novel prior specification for the SV process and its parameters that is suitable for the SVAR models in the context of identification through heteroskedasticity. It is based on the model equations (8)–(12). Our objectives for setting this prior are that (i) the implied prior for the conditional variances is centred around a homoskedastic process, (ii) it provides a controlled level of shrinkage towards the hypothesis of homoskedasticity of the structural shocks, (iii) it supports the standardisation of the system at a point where $\sigma_{n.t}^2 = 1$ for all n and t (and, hence, the shock is homoskedastic), (iv) it leads to an efficient Gibbs sampler for the estimation of the model, and (v) it facilitates a reliable verification of the partial identification through heteroskedasticity using the SDDR.

By centring at and shrinking the prior for the conditional variances towards the hypothesis of homoskedasticity we address the critique of Bayesian SVARs identified through heteroskedasticity by Lewis (2021) and Bertsche and Braun (2022). They argue that the conventional priors enable the identification via the implied non-normality of residuals irrespectively of whether the structural shocks are heteroskedastic or not. Note that this feature is shared by the corresponding frequentist models that assume the normality of the SV equation innovations. Our priors proposed in the following let the data decide on whether particular shocks are homo- or heteroskedastic. As a consequence, in our setup, it is the data that decides on the partial or full identification of the SVAR system via heteroskedasticity. At the same time, our priors facilitate the estimation of a parametric model when it is partially identified and perform a valid verification of the identification using SDDRs.

This section first states the prior for the latent process $h_{n.t}$ and the parameters of the SV process. Then, it analyses the properties of the implied priors for the conditional variances and their logarithms. Finally, it discusses how partial identification of the SVAR model can be verified in our setup.

4.1. Priors for the SV process parameters

In order to state the prior distribution for $h_{n.t}$ we define a $T \times 1$ vector $\mathbf{h}_n = \begin{pmatrix} h_{n.1} & \dots & h_{n.T} \end{pmatrix}'$, and a $T \times T$ matrix $\mathbf{H}_{\rho_n}$ with ones on the main diagonal, with $-\rho_n$ on the first subdiagonal, and with zeros elsewhere.

The prior for the latent process $\mathbf{h}_n$ is determined by equations (11) and (12) which imply the following conditionally multivariate normal distribution, given parameter ρ_n:

$$\mathbf{h}_n \mid \rho_n \sim \mathcal{N}_T \left(\mathbf{0}_{T \times 1}, \left(\mathbf{H}'_{\rho_n} \mathbf{H}_{\rho_n} \right)^{-1} \right). \tag{16}$$

The remaining parameters of the SV process include ω_n – the essential parameter of the non-centred parameterisation responsible for the variance of the log-conditional variances, ρ_n – the latent process' autoregressive parameter, and $\sigma^2_{\omega_n}$ – the prior variance of ω_n. These parameters follow a hierarchical prior structure, including normal, uniform, and truncated gamma distributions, respectively:

$$\omega_n \mid \sigma^2_{\omega_n} \sim \mathcal{N}\left(0, \sigma^2_{\omega_n}\right), \tag{17}$$

$$\rho_n \mid \sigma^2_{\omega_n} \sim \mathcal{U}\left(-\sqrt{1-\sigma^2_{\omega_n}}, \sqrt{1-\sigma^2_{\omega_n}}\right), \tag{18}$$

$$\sigma^2_{\omega_n} \mid \rho_n \sim \mathcal{G}\left(\underline{S}, \underline{A}\right) \mathcal{I}\left(0 < \sigma^2_{\omega_n} < 1 - \rho^2_n\right). \tag{19}$$

This original specification is our proposal for the prior distributions of the SV process for structural shocks of the SVAR models in the context of identification through heteroskedasticity. It is complemented by three restrictions:

$$|\rho_n| < 1, \tag{20}$$

$$\frac{\sigma^2_{\omega_n}}{1-\rho^2_n} \leq 1, \tag{21}$$

$$\underline{A} > 0.5. \tag{22}$$

Under the first restriction, the latent SV process has no unit roots and is stationary. Note

that restrictions (20) and (21) result in the bounds for ρ_n as presented in the uniform prior for ρ_n from expression (18). Similarly, the truncation of the gamma prior for $\sigma^2_{\omega_n}$ results from the restriction (21). The last restriction from inequality (22) determines the marginal prior for ω_n and makes it particularly suitable for our setup, as explained below. We show that under restrictions (20) and (21), all our objectives regarding centring and shrinking of the prior for the conditional variances are met, whereas, under the last one, the SDDR for the verification of the partial identification provides reliable results and is free of arbitrary choices.

4.2. Implied priors for conditional variances and their logarithms

The SV process equations (8)–(12), the prior distributions for their parameters (16)–(19), and the restrictions (20)–(22) imply prior distributions for the interpretable sequences of random variables like conditional variances and their logarithms. The following definitions will simplify the subsequent discussion of these priors.

Definition 1. (Normal product distribution)
Let x and y denote two independent zero-mean normally distributed random variables with variances σ^2_x and σ^2_y respectively. Then, a random variable $z = xy$ follows the normal product distribution with zero mean and variance $\sigma^2_z = \sigma^2_x\sigma^2_y$, denoted by $z \sim \mathcal{NP}(\sigma^2_z)$, with density function given by: $\frac{1}{\pi\sqrt{\sigma^2_z}} K_0\left(\frac{|z|}{\sqrt{\sigma^2_z}}\right)$. □

The normal product distribution is known in the statistical literature. We state it here to clarify our notation. However, the following distribution is new and its density function is obtained by change of variables.

Definition 2. (Log normal product distribution)
Let a random variable z follow the normal product distribution with variance σ^2_z. Then, a random variable $q = \exp(z)$ follows the log normal product distribution, denoted by $q \sim \log\mathcal{NP}(\sigma^2_z)$, with density given by: $\frac{1}{\pi\sqrt{\sigma^2_z}} \frac{1}{q} K_0\left(\frac{|\log q|}{\sqrt{\sigma^2_z}}\right)$. □

We are now ready to state the implied marginal priors for the volatilities in Proposition 1:

Proposition 1. ***(Marginal distributions of conditional volatilities)***
Given the prior specification from equations (8)–(12) *and* (16)–(22)*, the marginal priors for the latent process* $h_{n.t}$*, log-conditional variances* $\log \sigma^2_{n.t} = \omega_n h_{n.t}$*, and conditional variances* $\sigma^2_{n.t} = \exp(\omega_n h_{n.t})$ *are given by the following normal, normal product, and log normal product distributions:*

(a) $h_{n.t} \mid \rho_n \sim \mathcal{N}\left(0, \frac{1-\rho_n^{2t}}{1-\rho_n^2}\right)$,

(b) $\log \sigma^2_{n.t} \mid \rho_n, \sigma^2_{\omega_n} \sim \mathcal{NP}\left(\sigma^2_{\omega_n} \frac{1-\rho_n^{2t}}{1-\rho_n^2}\right)$,

(c) $\sigma^2_{n.t} \mid \rho_n, \sigma^2_{\omega_n} \sim \log \mathcal{NP}\left(\sigma^2_{\omega_n} \frac{1-\rho_n^{2t}}{1-\rho_n^2}\right)$,

with the corresponding limiting distributions:

(d) $\lim_{t\to\infty} h_{n.t} \mid \rho_n \sim \mathcal{N}\left(0, \frac{1}{1-\rho_n^2}\right)$,

(e) $\lim_{t\to\infty} \log \sigma^2_{n.t} \mid \rho_n, \sigma^2_{\omega_n} \sim \mathcal{NP}\left(\frac{\sigma^2_{\omega_n}}{1-\rho_n^2}\right)$,

(f) $\lim_{t\to\infty} \sigma^2_{n.t} \mid \rho_n, \sigma^2_{\omega_n} \sim \log \mathcal{NP}\left(\frac{\sigma^2_{\omega_n}}{1-\rho_n^2}\right)$. □

Proof. **(a)** The result is based on the properties of a normal compound distribution that facilitates the integration of $\int p(h_{n.t}, h_{n.t-1}, \ldots, h_{n.1}) d(h_{n.t-1}, \ldots, h_{n.1})$, where the joint distribution under the integral is constructed from the conditional distributions $h_{n.t} \mid h_{n.t-1}, \ldots, h_{n.1} \sim \mathcal{N}(\rho_n h_{n.t-1}, 1)$, and using $h_{n.0} = 0$. **(b)** The result is obtained directly by applying Definition 1, the result **(a)** and the prior in expression (17). Points **(c)**–**(f)** are obtained as a straightforward consequence of the first two results. □

4.3. Non-centred and centred SV representations

In order to understand our prior setup it is essential to understand the properties of the implied prior distributions and compare them to a conventional prior for SV processes. Our proposed normal prior for the parameter ω_n from equation (17) implies the following

gamma prior for the conditional variance of the SV process in its centred parameterisation $\sigma^2_{v.n} = \omega^2_n$ with expected value $E\left[\sigma^2_{v.n} \mid \sigma^2_{\omega_n}\right] = \sigma^2_{\omega_n}$:

$$\sigma^2_{v.n} \mid \sigma^2_{\omega_n} \sim \mathcal{G}\left(2\sigma^2_{\omega_n}, \frac{1}{2}\right). \tag{23}$$

However, the conventional conjugate prior for a variance parameter belongs to the family of inverted-gamma distributions. We state this conventional prior for the sake of comparison to our proposal. Let the conventional prior for $\sigma^2_{v.n}$ in the centred parameterisation have the expected value equal to $\sigma^2_{\omega_n}$ and be specified as:

$$\sigma^2_{v.n} \mid \sigma^2_{\omega_n} \sim \mathcal{IG}2\left(2\sigma^2_{\omega_n}, 4\right), \tag{24}$$

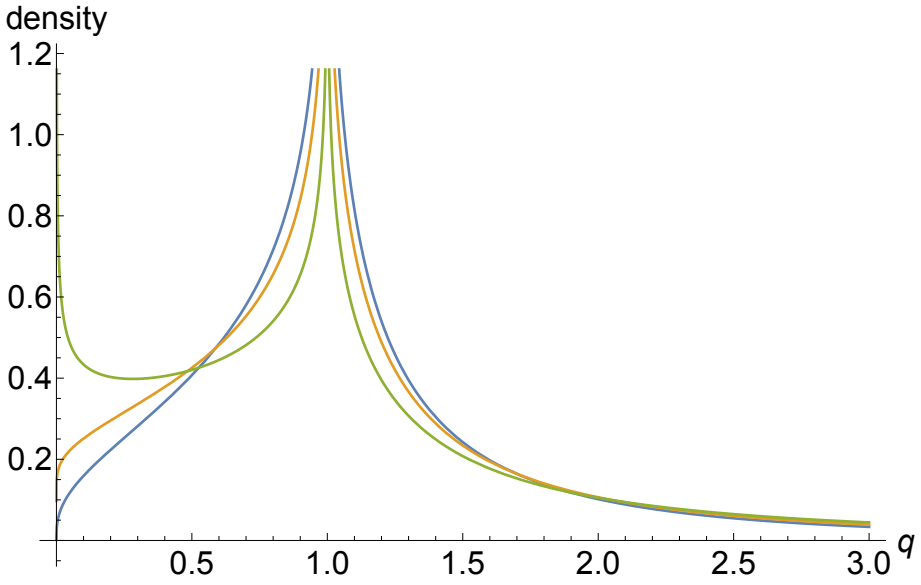
where $\mathcal{IG}2(\cdot,\cdot)$ denotes the inverted-gamma 2 distribution (see Bauwens, Lubrano and Richard, 1999, Appendix A). The alternative prior specifications from equations (23) and (24) illustrate how our approach is different from the conventional take on SV models. Our gamma prior for the volatility of the log-volatility parameter in equation (23) implies strong shrinkage towards the homoskedasticity (see Chan, 2018). Furthermore, the prior from equation (23) shows that our specification can be equivalently presented as the SV process in its centred parameterisation and with a gamma prior for the variance of the latent process.

4.4. Properties of the prior for conditional variances

At the beginning of the current section, we stated the objectives for our prior setup. Below we provide more insights into how these objectives are embedded in our priors and why that matters.

The first objective was to centre the prior for conditional variances around the hypothesis of homoskedasticity. As long as the prior expected values of the log-conditional variances following the normal product distribution are 0 for all t, that is, centred at the homoskedastic process, the expected values of the conditional variances following the log normal product distribution are not simply equal to 1, of

Figure 1: Densities of the log normal product for various values of the scale parameter.



Note: the blue, orange, and green lines correspond to the densities for the values of the scale parameter, σ_z^2, equal to 0.8, 1, and 1.5 respectively.

course. However, the log normal product distribution has a pole at value 1. Property 1 establishes when this distribution has a single mode at this point.

Property 1. (Single pole of log normal product distribution at point 1)
The log normal product distribution from Definition 2 has a single pole at point 1 iff $\sigma_z^2 \leq 1$. In this case, the value of the density function approaches 0 when the argument goes to 0. When $\sigma_z^2 > 1$, this density function differs from 0 and goes to infinity when q goes to 0.

We illustrate Property 1 in Figure 1. First, note that the condition for a single pole for the distribution of conditional variances for all t was stated in expression (21) and, consequently, defined our prior. Secondly, the single pole of the log normal product distribution implies a strong concentration of the prior probability mass at the point corresponding to the pole. This point is $\sigma_{n.t}^2 = 1$ implying homoskedasticity of the structural shocks. This extreme concentration of probability mass to the point of unbounded probability density function supports our claim that at the prior mode, the

SVAR model is not identified through heteroskedasticity. Still, the distribution is proper and allows a strong signal from the data to push the posterior probability mass towards heteroskedasticity. Finally, note that the inequality restriction is imposed on the parameter of the limiting log normal product distribution, $\sigma^2_{\omega_n}/(1-\rho^2_n)$, to ensure it holds for all t.

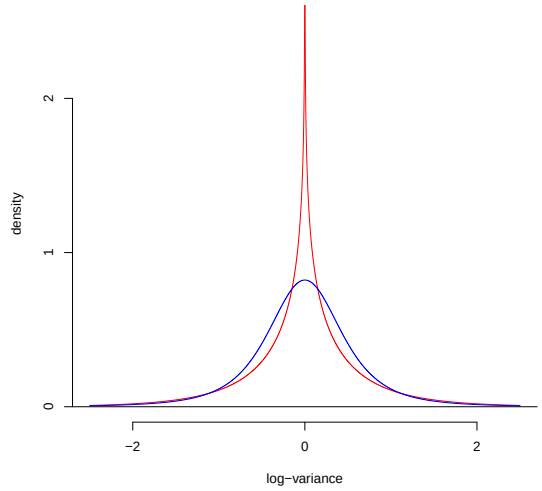
The second objective was to shrink the prior for conditional variances towards homoskedasticity. It is also obtained by imposing the restriction from Property 1 in inequality (21). This objective cannot be obtained without this restriction. The second pole at point 0 distributes probability mass over the interval from zero to 1 more evenly. In effect, the posterior estimates are either pushed towards very small values when they are shrunk towards 0 or take practically any value on the real scale. Additionally, our hierarchical prior exhibits the same shrinkage properties for the log-conditional volatilities as analysed in other contexts by Bitto and Frühwirth-Schnatter (2019) and Cadonna, Frühwirth-Schnatter and Knaus (2020). These properties include the extreme concentration of the prior mass at the point of homoskedasticity and heavy tails, allowing to accommodate the signals of heteroskedasticity from the data.

The latter remark leads to the third objective of specifying the prior so that the SVAR model is standardised. The conventional standardisations of the model in the frequentist approach rely either on setting the diagonal elements of the matrix B_0 to 1 or the unconditional expected value of the conditional variances, $E\left[\sigma^2_{n.t}\right]$, to value 1. Both of these solutions complicate the derivation of an efficient Bayesian estimation algorithm and might require cumbersome treatment of the local identification of the model. Our property relies on the appropriate centring and shrinking of the prior that results in standardised posterior output.

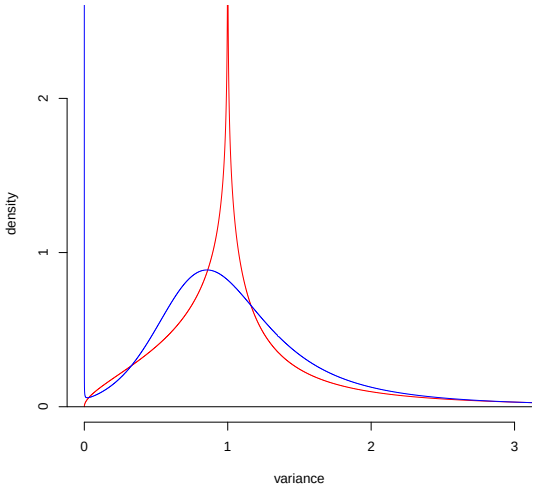
All of these features are clearly visible in Figure 2 that juxtaposes our prior distribution with a conventional setup. These figures present in red our prior distribution for the non-centred parameterisation with the restrictions (20)–(22), and in blue a conventional prior for the SV process in its centred parameterisation. In both setups here we apply a uniform prior for ρ_n on the interval $(-1, 1)$, and the unrestricted inverse gamma prior for $\sigma^2_{v.n}$ as in

Figure 2: Priors for log-variances and variances of Stochastic Volatility processes in their centred and non-centred parameterisations

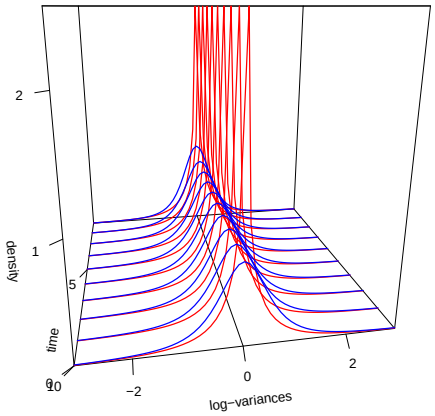
(a) marginal prior for a log-variance



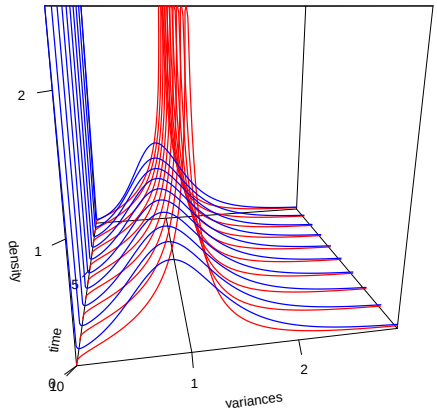
(b) marginal prior for a variance



(c) prior for log-variances over time



(d) prior for variances over time



Note: red corresponds to our non-centred and blue to the centred parameterisation of the Stochastic Volatility process.

expression (24).

Figure 2 (a) presents the marginal limiting prior for log-conditional variances for two cases: (i) our proposed prior is plotted in red where the hyper-parameters were integrated

out according to $\log \sigma^2_{n.t} = \int p\left(\log \sigma^2_{n.t} \mid \rho_n, \sigma^2_{\omega_n}\right) d\left(\rho_n, \sigma^2_{\omega_n}\right)$ with respect to the priors for ρ_n and $\sigma^2_{\omega_n}$ as in equations (18)–(21); (ii) a conventional prior is plotted in blue where the marginalisation is performed with respect to parameters ρ_n and $\sigma^2_{v.n}$ following the unrestricted prior described in the paragraph above. Our proposal is based on the normal product distribution, and the marginal prior inherits its properties, including the pole at zero, strong shrinkage towards the prior mean, and heavy tails. The conventional prior is Student-t distributed, not exhibiting that strong concentration of the prior probability mass around homoskedasticity.

Figure 2 (b) presents a similarly obtained marginal limiting prior distribution of the conditional variances and reveals major differences in the prior specifications. Firstly, note that our prior plotted in red has a single pole at value 1, and exhibits strong shrinkage towards this point. The conventional prior that is log t distributed has a mode at point 1 and a pole at point 0. The pole arises from the marginalisation of the log t distribution with respect to parameters ρ_n and $\sigma^2_{v.n}$ and is described in Property 2.

Property 2. (Prior for conditional variances in a centred SV model)
Consider a centred SV model from equations (13)–(15) with the inverted-gamma 2 prior for the volatility of the log-volatility parameter as in equation (24) that, for the sake of generality, is stated as:

$$\sigma^2_{v.n} \mid \sigma^2_{\omega_n} \sim \mathcal{IG}2\left(\sigma^2_{\omega_n}, \underline{\nu}\right). \tag{25}$$

Then,

(a) the log-volatilities, $\tilde{h}_{n.t}$, follow a Student-t marginal prior distribution (see Bauwens et al., 1999),

(b) the conditional variances, $\sigma^2_{n.t}$, follow a log t marginal prior distribution (see Hogg and Klugman, 1983),

(c) the prior distribution stated in **(b)** has a pole at point 0, unless $\underline{\nu}$ goes to infinity (see Callealta Barroso, García-Pérez and Prieto-Alaiz, 2020, for points (c)–(e)),

(d) the prior distribution stated in **(b)** has a second mode – a local maximum – at point $\exp\left\{-\frac{1}{2}\left[(\underline{\nu}+1)-\sqrt{(\underline{\nu}+1)^2-4\sigma^2_{\omega_n}}\right]\right\}$ iff $\sigma^2_{\omega_n} < \frac{(\underline{\nu}+1)^2}{4}$,

(e) the prior distribution stated in **(b)** has a median at point 1 only if $\underline{\nu}$ goes to infinity. □

Property 2 shows that the unrestricted centred SV parameterisation is highly problematic in SVAR applications. With unconstrained volatility of the log-volatility parameter, it does not ensure even the standardisation of the system about value $\sigma^2_{n.t} = 1$. Moreover, with any finite values of the shape hyper-parameter, $\underline{\nu}$, the pole at point 0 provides heavy local shrinkage towards a point in which the model is singular as it exhibits zero conditional variances of the structural shocks. In this context, our proposal satisfying all the stated objectives robustly leads to reliable posterior estimates and inferences.

Finally, Figures 2 **(c)** and 2 **(d)** present the shape of the priors for conditional variances and their logarithms in a dynamic setting for the first ten sample periods. These distributions inherit the properties from the marginal limiting distributions. All these arguments indicate the adequacy of our prior setup that includes an appropriate hierarchy and the restrictions that provide necessary features and interpretations in the context of priors for the volatility process of structural shocks for the SVAR model partially identified through heteroskedasticity.

5. Bayesian assessment of identification conditions

The assessment of the identification conditions is a challenging task for heteroskedastic SVARs because it often relies on the estimated parameters of a model that is potentially not identified (see the points made by Lewis (2021) and by Lütkepohl et al. (2021)). We follow Lütkepohl and Woźniak (2020) and complement the setup of our model with a Savage-Dickey Density Ratio by Verdinelli and Wasserman (1995) for the verification of the identification conditions from Theorem 1. More precisely, we verify the heteroskedasticity

of each structural shock by seeking evidence against the homoskedasticity restriction

$$\omega_n = 0. \tag{26}$$

If this condition holds, the latent volatility equation becomes deterministic with $h_{n.t} = 0$ and $\sigma^2_{n.t} = 1$ for all t, that is, the latent process is homoskedastic for a given n. This restriction is verified for each n, that is, for each structural shock. The results allow us to make statements regarding the identification of the SVAR model through heteroskedasticity and answer whether the system is fully or partially identified. Recall that, by Theorem 1, each shock with nonzero ω_n is identified through heteroskedasticity. The SDDR is particularly suitable for the verification of the identification conditions because it relies on the estimation outcome of one general model and does not require the estimation of restricted models.

In order to verify the restriction from equation (26) using the SDDR, we adapt the ideas presented by Frühwirth-Schnatter and Wagner (2010) for state space models and applied to SV models by Chan (2018). More specifically, we extend the heteroskedasticity verification procedure using SDDR by Chan (2018) for investigating the hypothesis (26) for the structural shocks of SVAR models. Several features make our procedure reliable. Firstly, it is embedded in the SV hierarchical prior structure that is suitable for the volatility process in the context of identification through heteroskedasticity. Secondly, we extend the normal prior distribution for ω_n, as in equation (17), proposed by Chan (2018) by a hierarchical prior for its variance $\sigma^2_{\omega_n}$ as in equation (19). In contrast, Chan (2018) simply fixes this prior variance. Our specification allows the subsequent estimation of $\sigma^2_{\omega_n}$ making the prior and the inference using the SDDR less dependent on arbitrary choices. Moreover, based on the results provided by Bitto and Frühwirth-Schnatter (2019) and Cadonna et al. (2020), the marginal prior for ω_n, based on a gamma scale mixture of normal distributions, combines the extreme shrinkage towards the hypothesis of homoskedasticity and heavy tails allowing the data to provide the signal of heteroskedasticity that informs posterior inference. An additional property that is necessary for the SDDR to provide reliable

outcomes is the bounded density function for the marginal prior for ω_n that we discuss below. Finally, our estimation algorithm for the SV parameters uses the ancillarity-sufficiency interweaving strategy by Kastner and Frühwirth-Schnatter (2014) that is the optimal solution making the MCMC algorithm efficient when it is unknown whether the process is homo- or heteroskedastic. Therefore, we provide the setup that is numerically efficient, sufficiently flexible to exploit the information from data, and suitable for partial identification verification of the SVARs identified through heteroskedasticity.

Our SDDR is defined by the ratio of the marginal posterior of ω_n ordinate to the marginal prior ordinate for this parameter both evaluated at point 0:

$$SDDR = \frac{p(\omega_n = 0|\mathbf{y})}{p(\omega_n = 0)}. \tag{27}$$

Small values of the SDDR provide evidence against the restriction $\omega_n = 0$. In turn, the posterior probability mass being more concentrated around the restriction than the prior probability mass is evidence in favor of the restriction. The analysis of the SDDR in equation (27) reveals that Property 3 establishing the bounded marginal prior for ω_n is essential. If this density is unbounded at point $\omega_n = 0$, then it is impossible for the data to provide evidence for heteroskedasticity, as the denominator of the SDDRs is equal to infinity and the SDDR itself goes to zero. Additionally, extending the prior by Chan (2018) allows our verification procedure to let the data speak about the heteroskedasticity of structural shocks.

In order to present the result of bounded marginal prior for ω_n, we first redefine the prior $\sigma^2_{\omega_n}$ from equation (19) as:

$$\sigma^2_{\omega_n} \sim \mathcal{G}\left(\underline{S}, \underline{A}\right) \tag{28}$$

and provide the density of the marginal prior for ω_n.

Proposition 2. ***(Density of marginal prior for*** ω_n***)***

The marginal prior density function for parameter ω_n *obtained by marginalising the joint prior*

distribution over $\sigma^2_{\omega_n}$, $p(\omega_n) = \int_0^\infty p\left(\omega_n \mid \sigma^2_{\omega_n}\right) p\left(\sigma^2_{\omega_n}\right) d\sigma^2_{\omega_n}$, *where the priors* $p\left(\omega_n \mid \sigma^2_{\omega_n}\right)$ *and* $p(\sigma^2_{\omega_n})$ *are given by expressions* (17) *and* (28) *respectively is given by:*

$$p(\omega_n) = \frac{|\omega_n|^{\underline{A}-\frac{1}{2}} K_{\underline{A}-\frac{1}{2}}\left(\sqrt{\frac{2}{\underline{S}}}|\omega_n|\right)}{\sqrt{\pi}\left(\sqrt{2}\right)^{\underline{A}-\frac{3}{2}} \Gamma\left(\underline{A}\right)\left(\sqrt{\underline{S}}\right)^{\underline{A}+\frac{1}{2}}}. \tag{29}$$

□

Proof. The integration proceeds by recognising the constant and kernel and applies to the latter which is facilitated using the normalising constant of the generalised inverse Gaussian distribution provided by Barndorff-Nielsen (1997). □

Below, we state that the existence of the upper bound for the density from equation (29) at point 0 depends on the value of hyper-parameter $\underline{A}$ by providing results on the following limit:

Property 3. (Upper bound of the marginal prior density for ω_n) (See Cadonna et al., 2020, Theorem 2).

$$\lim_{\omega_n \to 0} p(\omega_n) = \begin{cases} \infty & \text{for } 0 < \underline{A} \leq 0.5 \\ \dfrac{1}{\sqrt{2\pi\underline{S}}\left(\underline{A}^2 - \frac{1}{4}\right)} \dfrac{\Gamma\left(\underline{A} + \frac{3}{2}\right)}{\Gamma\left(\underline{A}\right)} & \text{for } \underline{A} > 0.5 \end{cases} \tag{30}$$

Therefore, the marginal prior density function for parameter ω_n is bounded from above iff the restriction, $\underline{A} > 0.5$, from equation (22) holds. Consequently, we set the hyper-parameter $\underline{A} = 1$ which reduces the gamma prior to an exponential distribution and the Bayesian Lasso prior considered by Belmonte, Koop and Korobilis (2014) for state-space models. However, other choices are possible and are best reviewed by Cadonna et al. (2020).

We further set the hyper-parameter $\underline{S} = 0.05$. This choice implies that practically the whole prior probability mass of $\sigma^2_{\omega_n}$ lays within the interval $(0, 1)$. Therefore, instead of imposing a strict restriction leading to a hard-to-sample-from truncated full conditional

posterior distribution for $\sigma^2_{\omega_n}$, we impose the restriction via an appropriate prior specification. As we explain in Appendix C, we first sample $\sigma^2_{\omega_n}$ from the generalised inverse Gaussian distribution whose value will be within the interval $(0, 1)$ with a probability close to one thanks to our prior setup. Subsequently, ρ_n is sampled from the truncated normal distribution given the most recent draw of $\sigma^2_{\omega_n}$ to fix the truncation set.

The SDDR can be easily computed as long as the densities of the full conditional posterior and the prior distributions are of a known analytical form. In Appendix C, we show that, given the data, the latent volatilities processes involved in our model and the parameters of the SVAR equation, the parameters ω_n can be independently sampled from the univariate normal full conditional posterior distributions with the mean $\overline{\omega}_n$ and variance $\overline{v}_{\omega_n}$ specified in equations (C.18)–(C.20). Then, the numerator of the SDDR can be computed using a sample of S draws from the posterior distribution by applying the marginal density ordinate estimator proposed by Gelfand and Smith (1990):

$$\widehat{p}(\omega_m = 0 \mid \mathbf{y}) = \frac{1}{S}\sum_{s=1}^{S} f_{\mathcal{N}}\left(0; \overline{\omega}_n^{(s)}, \overline{v}_{\omega_n}^{(s)}\right), \tag{31}$$

where $f_{\mathcal{N}}$ denotes the density function of a normal distribution, whereas $\overline{\omega}_n^{(s)}$ and $\overline{v}_{\omega_n}^{(s)}$ denote the values of the mean and variance in which the place of the parameters of the model are replaced by their s^{th} draws from the posterior.

6. Prior distribution for the SVAR parameters

Our objectives for setting the joint prior distribution for the structural matrix $\mathbf{B}_0$ and the autoregressive slope parameters collected in the matrix $\mathbf{A} = \begin{bmatrix}\mathbf{A}_1 & \dots & \mathbf{A}_p & \mathbf{A}_d\end{bmatrix}$ are that (i) it is conditionally conjugate, and thus, facilitates the derivation of an efficient Gibbs sampler for the estimation of the parameters, (ii) it is a reference prior that does not distort the shape of the likelihood function due to the local identification of the model as defined by Rubio-Ramírez et al. (2010), (iii) it can be interpreted as a Minnesota prior proposed by Doan, Litterman and Sims (1984), and (iv) it enjoys the flexibility of the hierarchical prior

specification thanks to which the essential hyper-parameters responsible for the level of shrinkage are estimated as argued by Giannone, Lenza and Primiceri (2015).

All these objectives are met when the prior for the structural matrix is set to the generalised-normal distribution proposed by Waggoner and Zha (2003a) and multivariate normal for the autoregressive parameters. Let $\mathbf{B}_{0.n}$ and $\mathbf{A}_n$ denote the nth row of the matrices $\mathbf{B}_0$ and $\mathbf{A}$, respectively. Then the prior distribution for matrix $\mathbf{B}_0$ is proportional to

$$p\left(\mathbf{B}_0 \mid \gamma_0\right) \propto \det\left(|\mathbf{B}_0|\right)^{\nu-N} \exp\left\{-\frac{1}{2}\sum_{n=1}^{N}\frac{1}{\gamma_{0.n}}\mathbf{B}_{0.n}\mathbf{B}_{0.n}'\right\}. \tag{32}$$

The parameters of this distribution are further assumed to be equation invariant. That feature makes this distribution the reference prior, which means that it is invariant to the rotations of the structural system up to permutation and sign change of its rows (see Woźniak and Droumaguet, 2015). The scale matrix of the distribution in (32) is set to $\gamma_{0.n}\mathbf{I}_N$, where $\gamma_{0.n}$ is a hyper-parameter, and the shape parameter is set to $\underline{\nu}_0 = N$, which makes the marginal prior distribution for the rows of $\mathbf{B}_0$ the N-variate normal distribution with the zero mean and covariance $\gamma_{0.n}\mathbf{I}_N$.

The prior distribution for each row of matrix $\mathbf{A}$ is multivariate normal, sharing features of the Minnesota prior. Therefore, the prior mean of $\mathbf{A}$ is equal to $\underline{\mathbf{A}} = \begin{bmatrix}\mathbf{D} & \mathbf{0}_{N\times(N(p-1)+d)}\end{bmatrix}$, where $\mathbf{D}$ is a diagonal matrix with zeros and ones on the diagonal depending on whether the corresponding variables in $\mathbf{y}_t$ are stationary or unit-root nonstationary. The matrix $\mathbf{D}$ is fixed at $\mathbf{I}_N$ if all variables in $\mathbf{y}_t$ are unit-root non-stationary or at $\mathbf{0}_{N\times N}$ if they are stationary. The covariances of the rows of $\mathbf{A}$ are given by diagonal matrices $\gamma_{A.n}\underline{\mathbf{\Omega}}$ with scalar hyper-parameters $\gamma_{A.n}$ and, where $\underline{\mathbf{\Omega}} = \operatorname{diag}\begin{pmatrix}\mathbf{p}^{-1\prime}\otimes\boldsymbol{\iota}_N' & 100\boldsymbol{\iota}_d'\end{pmatrix}$, and $\mathbf{p}^{-1}$ denotes a vector containing the reciprocal of integer values from 1 to p. This matrix provides the increasing level of shrinkage with increasing lag order of the autoregressive slope parameters, incorporating the ideas of the Minnesota prior of Doan et al. (1984). Furthermore, the prior variances of the parameters corresponding to the deterministic terms are equal to $100\gamma_A$, reflecting a

popular view that the shrinkage should be relatively weaker for these parameters.

Extending the prior by Giannone et al. (2015), the levels of shrinkage of the autoregressive and structural matrices follow a 3-level global-local hierarchical prior on the equation-specific shrinkage parameters $\gamma_{A.n}$ and $\gamma_{0.n}$:

$$\gamma_{0.n} \mid \underline{s}_{0.n} \sim \mathcal{IG}2\left(\underline{s}_{0.n}, \underline{\nu}_0\right), \qquad \underline{s}_{0.n} \mid \underline{s}_{\gamma_{0.n}} \sim \mathcal{G}\left(\underline{s}_{\gamma_0}, \underline{\nu}_{\gamma_0}\right), \qquad \underline{s}_{\gamma_0} \sim \mathcal{IG}2\left(\underline{s}_{s_0}, \underline{\nu}_{s_0}\right), \tag{33}$$

$$\gamma_{A.n} \mid \underline{s}_{A.n} \sim \mathcal{IG}2\left(\underline{s}_{A.n}, \underline{\nu}_A\right), \qquad \underline{s}_{A.n} \mid \underline{s}_{\gamma_{A.n}} \sim \mathcal{G}\left(\underline{s}_{\gamma_A}, \underline{\nu}_{\gamma_A}\right), \qquad \underline{s}_{\gamma_A} \sim \mathcal{IG}2\left(\underline{s}_{s_A}, \underline{\nu}_{s_A}\right). \tag{34}$$

We set $\underline{\nu}_0, \underline{\nu}_{\gamma_0}, \underline{s}_{s_0}$, and $\underline{\nu}_{s_0}$ to values 10, 10, 100, and 1 respectively to make the marginal prior for the elements of $\mathbf{B}_0$ quite dispersed, and $\underline{\nu}_A, \underline{\nu}_{\gamma_A}, \underline{s}_{s_A}$, and $\underline{\nu}_{s_A}$ all equal to 10, which facilitates relatively strong shrinkage for the autoregressive parameters in matrix $\mathbf{A}$ that gets updated, nevertheless. Providing sufficient flexibility on this 3-level hierarchical prior distribution was essential for a robust shape of the estimated impulse responses.

7. An empirical illustration: Verifying identification of tax shocks

When heteroskedasticity is used for identification in SVAR analysis, the shocks are distinguished by their variances or conditional variances. This approach provides distinct shocks without economic labels and requires some additional information to label the shocks. Such information is sometimes available in the form of specific shapes of the impulse responses associated with a shock or a specific sign pattern of the impact effects of the shocks.

To illustrate the methods developed in the previous sections, we will consider a fiscal SVAR model in which the unanticipated tax shock has been identified in different conventional ways. These alternative identification strategies include, for example, Blanchard and Perotti (2002) (henceforth BP) who use restrictions on the short-run effects of the shocks and the instantaneous interactions of the variables to identify their shocks, and by Mountford and Uhlig (2009) using sign restrictions. Moreover, Mertens and Ravn (2014) (henceforth MR), as revised by Ramey (2016), use an external instrument, a narrative measure of the tax shock proposed by Romer and Romer (2010).

Finally, Lewis (2021) (henceforth LE) uses heteroskedasticity and, hence, an approach in that respect similar to ours. He uses quite different frequentist estimation and inference methods, however, and he also needs more restrictive assumptions regarding the heteroskedasticity of the shocks. Specifically, he assumes non-proportional changes in the conditional variances of the structural shocks. We use the MR model as our benchmark to illustrate the use of our methodology for identifying the tax shock via heteroskedasticity, and the narrative measure by Romer and Romer (2010) to ensure a correct labelling of the shocks.

7.1. A simple fiscal SVAR

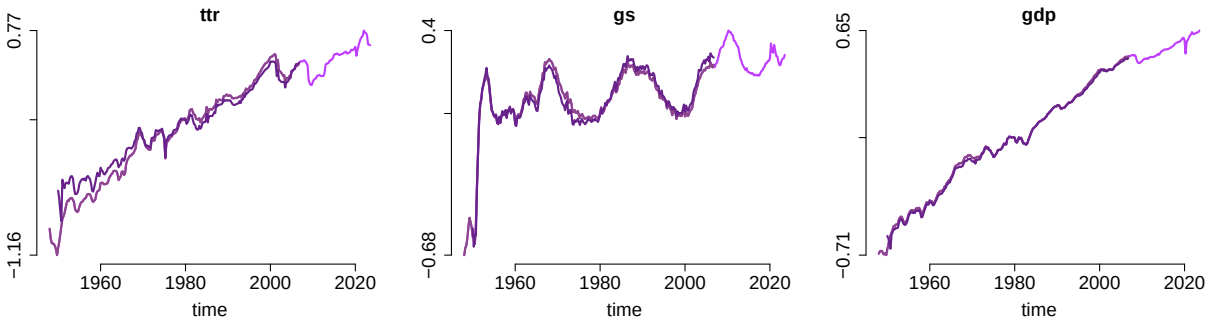
MR specify a three variable fiscal system including total tax revenue, denoted by *ttr*, government spendings, *gs*, and gross domestic product, *gdp*, and they express all the quarterly variables in real, log, per person terms. We will also consider these three variables and investigate whether the tax shock can be identified by our methodology.

In order to investigate identification through heteroskedasticity in this fiscal system, we use three alternative samples of different length and partly different values even for overlapping periods. They are plotted in Figure 3, where it can be seen that the series are different but similar in overlapping periods. The shortest sample, hereafter MR-sample, uses the data from MR and LE that is downloaded directly from Karl Merten's website[1]. Following the data construction described by MR, total tax revenue, government spendings, and gross domestic product, as well as the GDP deflator are taken from NIPA Tables number 3.2, 3.9.5, 1.1.5, and 1.1.9, respectively, provided by the U.S. Bureau of Economic Analysis (2024c,d,a,b), and the population variable is provided by Francis and Ramey (2009). This data spans the period 1950Q1 to 2006Q4.

We extend the sample to the latest available observations in 2023Q3 with modifications in the population variable that is replaced by one matching Francis and Ramey's definition and provided by the U.S. Bureau of Labor Statistics (2024). Based on these variables we form two samples, both of which contain longer time series than MR and start in

[1]The spreadsheet is available at: https://karelmertenscom.files.wordpress.com/2017/09/jme2014_data.xls

Figure 3: Data plots of the three samples used for estimation.



Note: The figure plots three series for three samples: the 2023-sample plotted in light pink includes observations from 1948Q1 to 2023Q3 ($T = 303$), the 2006-sample plotted in darker pink is as the 2023-sample but finishes in 2006Q4 ($T = 236$), the MR-sample, plotted in purple, spans the period from 1950Q1 to 2006Q4 ($T = 228$). The plotted series are standardised by subtracting from each series its first observation in 1980.

1948Q1. One of these samples, hereafter the 2023-sample, finishes in 2023Q3, and the other one, hereafter the 2006-sample, terminates in 2006Q4. Following MR, we use a VAR(4) model with a constant term, a linear and a quadratic trend, and a dummy for 1975Q2 as deterministic terms.

7.2. Verifying identification via heteroskedasticity

We base our structural analysis on the model (2). Hence, we have to sample from the posterior of the structural $\mathbf{B}_0$ matrix which is not identified without further restrictions if the shocks are not heteroskedastic. Even if the shocks are identified, the row ordering and row signs may change in different drawings from the posterior without taking special precautions to prevent that from happening. We therefore follow LE and reorder the rows and adjust their signs such that each draw has the minimum distance to the benchmark $\mathbf{B}_0$ matrix computed from the estimates of the structural parameters from BP to begin with and call this the BP–ordering. More detail on this procedure is provided in Appendix D. Hence, the shocks can be labelled along the lines of BP as an unanticipated tax shock (w_t^{ttr}), a government spending shock (w_t^{gs}), and an additional shock (w_t^{gdp}) capturing unexpected changes in gdp_t not caused by tax or spending shocks. We will label our shocks accordingly

Table 1: Verification of identification through heteroskedasticity of the structural shocks (based on the BP–ordering).

	2023-sample		2006-sample		MR-sample	
w_t^{ttr}	**-21.38**	[4.69]	-1.51	[0.18]	0.32	[0.05]
w_t^{gs}	**-4.62**	[0.79]	-1.32	[0.15]	0.23	[0.05]
w_t^{gdp}	**-63.39**	[6.43]	0.50	[0.03]	0.39	[0.03]

Note: The table reports the log of the Bayes Factors estimated via the log of SDDRs from equation (27) together with numerical standard errors (NSEs) provided in parentheses. Negative values provide evidence against homoskedasticity. Bold font numbers represent cases in which the evidence for heteroskedasticity is positive (values greater than 3 in absolute terms) or strong (greater than 20) on the scale of Kass and Raftery (1995). The NSEs are computed based on 30 subsamples of the original MCMC draws.

although it is, of course, not clear from the outset that the shocks can actually be identified by heteroskedasticity and, thus, we may not get identified shocks with our methodology. If we get identified shocks, they may differ from those in BP or MR in which case our labels may not be meaningful. We will later return to this issue.

The next step in our analysis is to assess whether there are shocks that are identified through heteroskedasticity. Our main tool for that purpose is the SDDR from equation (27). The SDDR values computed for each of the three shocks individually using our three data samples are reported in Table 1. For the 2023-sample, the evidence for heteroskedasticity of all three structural shocks is strong according to the scale proposed by Kass and Raftery (1995). The values of the log Bayes factors shown in the table mean that the posterior mass in favour of heteroskedasticity exceeds 99% for all the shocks. This result provides overwhelming evidence for the identification of all three shocks via heteroskedasticity in the 2023-sample and is robust to many variations in the model prior specification. These variations include perturbations of the hyper-parameters that need to be fixed in our setup. We checked the conclusions for three values of each scale and shape of the prior distribution for ω_n, as well as for three alternative setups for the hyper-parameters for each of the matrices $\mathbf{A}$ and $\mathbf{B}_0$. Each of these alternative setups included cases of stronger and weaker shrinkage than in our

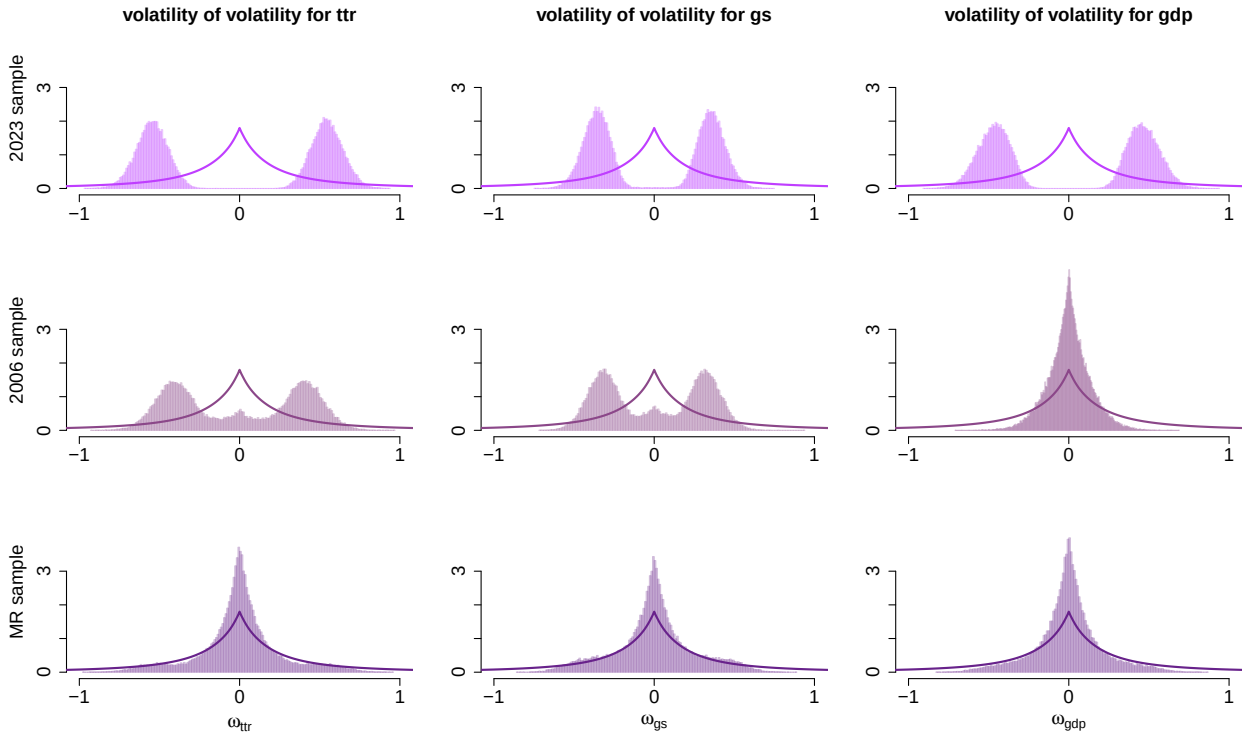
benchmark prior specification.

The evidence for the structural shocks to be identified through heteroskedasticity is much weaker in the 2006-sample. Moreover, the log Bayes factors estimated by the log-SDDRs for the MR-sample are even positive, implying that the posterior mass for homoskedasticity is greater than that for heteroskedasticity. The log-SDDRs are negative for the last two shocks in the 2006-sample that includes eight more observations than the MR-sample from the volatile late 1940s. More specifically, in the 2006-sample, the posterior probability of a heteroskedastic shock w_t^{ttr} is 82%. Obviously, in this case the evidence for identification through heteroskedasticity of the first shock is limited and it is even more limited for the other shocks. These findings are also robust to the perturbations in the values of the prior hyper-parameters.

In Figure 4 we further illustrate how the SDDRs work by plotting the marginal prior versus the marginal posterior densities of the volatility of the log-volatility parameter associated with our three samples. Based on the information from these plots, the SDDRs from equation (27) can be approximated by the ratio of the marginal posterior ordinate at zero to that of the marginal prior density. The figures for the 2023-sample exhibit posterior mass concentrated away from the origin, providing evidence against homoskedasticity. Instead, the posterior mass for the 2006- and MR-samples is concentrated about the hypothesis of homoskedasticity, often more than the prior, thus favouring homoskedasticity. Figure 4 features the marginal prior densities of the volatility of the ω_i parameters and highlights their essential characteristics discussed in Section 5, such as their high concentration about the origin, fat tails, boundedness from above, and flexibility provided by a hierarchical specification, extending the normal prior with fixed variance used by Chan (2018).

Finally, we analyse the sequences of conditional variances of the structural shocks that are required to be clearly distinct for partial identification of the shocks to hold according to Theorem 1. We plot their posterior means together with 90% highest posterior density (HPD) intervals in Figure 5. The conditional variances are visibly time varying for the 2023-sample. The conditional variances of the first shock are significantly different from

Figure 4: Marginal prior (solid line) and posterior (histograms) densities of the volatility of the log-volatility parameter, ω_1, across samples (based on the BP–ordering).

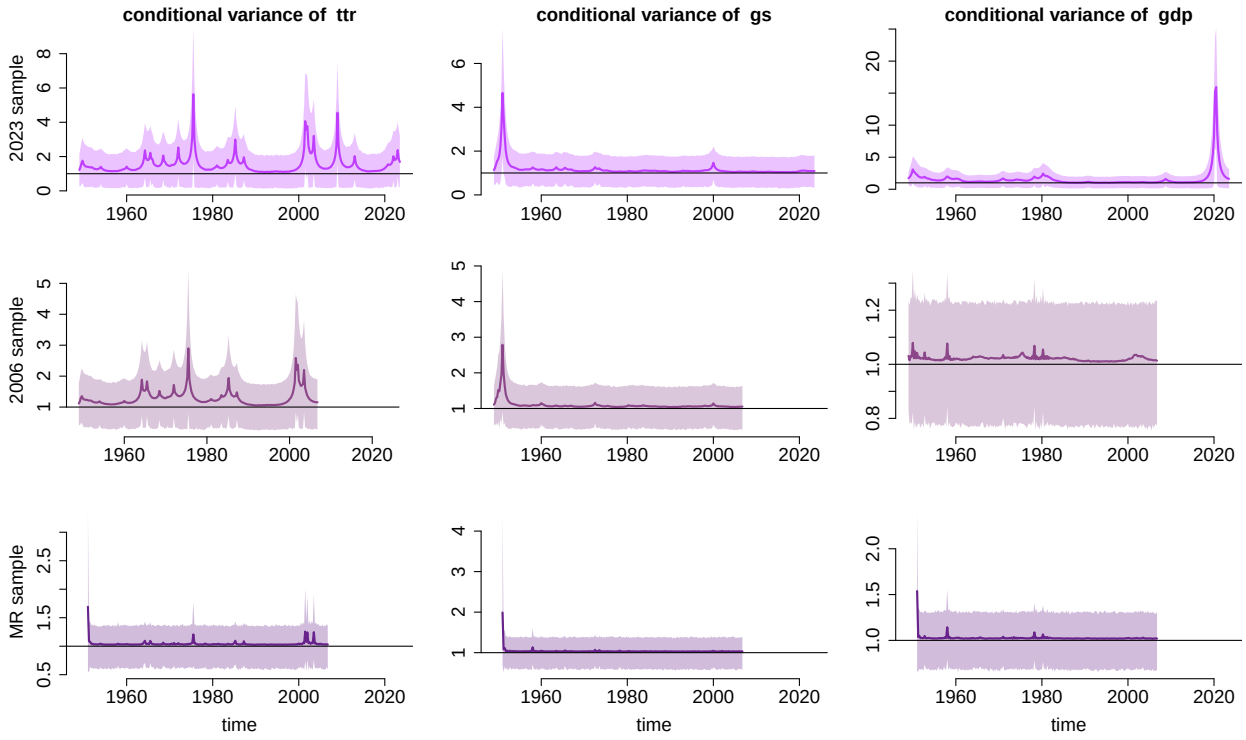


Note: The marginal prior density is estimated by numerical integration as in Gelfand and Smith (1990) using a grid of points from -1.1 to 1.1. They are the same for all samples and shocks. The marginal posterior densities are approximated using histograms. The ratio of these densities at point zero approximates the SDDR in equation (27). Posterior mass less concentrated than the prior mass about zero provides evidence against homoskedasticity.

1 in six periods in that sample including the mid 70s and mid 80s, individual quarters in 2001, 2002, and 2003, as well as in the first quarter of 2009. The variances of the second shock are different from 1 in the first quarter of 1951 only, while those of the third shock have HPD intervals not including 1 in 1950, and in quarters 2 and 3 of 2020. The distinctive occurrence times of high volatility periods for the three shocks provide strong evidence for them to be different in these sequences, further supporting the identification via heteroskedasticity in this sample.

The conditional variances in the 2006-sample are to some extent similar to those from the 2023-sample until 2006. However, at all times the 90% HPD intervals include the

Figure 5: Conditional variance of structural shocks in the three samples (based on the BP–ordering).



Note: The figures plot time-varying conditional variances of the structural shocks. The lines report the posterior mean and the shaded areas 90% HPD intervals. The variances in the first row clearly exhibit non-proportional changes across time. The horizontal black line is set at the value of 1 around which the prior is centred.

value of 1. This is caused by a weaker signal provided from the data in the shorter sample regarding time-varying volatility, which undermines the evidence for identification in the framework of our model. In the MR-sample, the evidence for conditional variances that support identification is even weaker. Thus, the bottom line is that, in the 2023-sample, the shocks are clearly identified through heteroskedasticity, while the evidence for identification through heteroskedasticity is weaker in the 2006-sample and no such evidence is found in the MR-sample.

7.3. Checking alternative ordering rules

One may wonder how much our results depend on the BP–ordering of our draws from the posterior of $\mathbf{B}_0$. Therefore, we have repeated our sampling using the estimates obtained

by MR to order the rows of the $\mathbf{B}_0$ drawings (see Appendix D). The results of the SDDRs based on the MR–ordering are presented in Table 2(a). They paint a similar picture as the results in Table 1. The evidence for shocks identified through heteroskedasticity is overwhelming in the 2023-sample. It is weaker for the 2006-sample and hardly existent in the MR-sample.

Table 2: Verification of identification through heteroskedasticity of the structural shocks (based on alternative orderings).

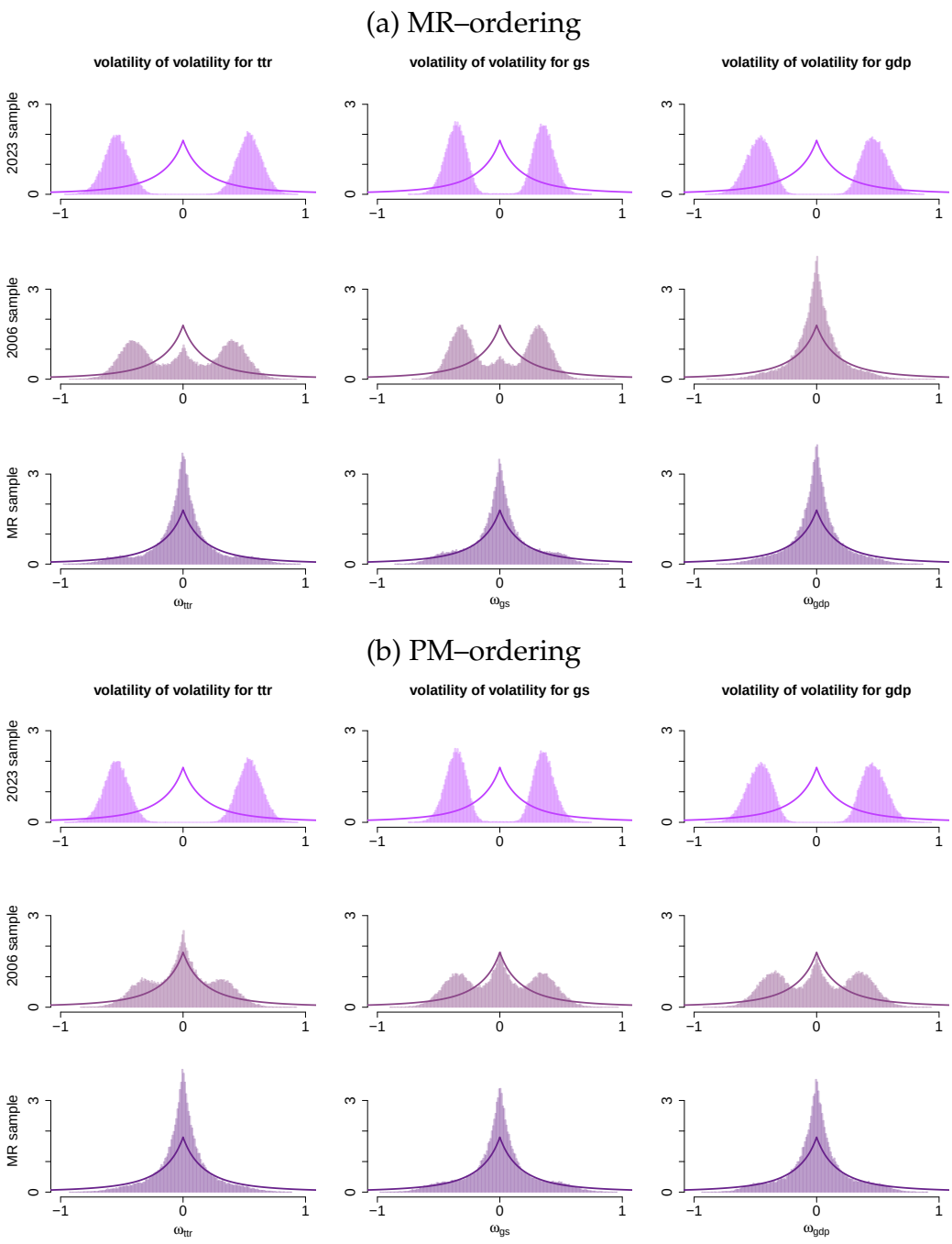
	2023-sample		2006-sample		MR-sample	
(a) MR–ordering						
w_t^{ttr}	**-21.38**	[4.97]	-0.92	[0.14]	0.32	[0.05]
w_t^{gs}	**-4.62**	[0.79]	-1.27	[0.13]	0.24	[0.04]
w_t^{gdp}	**-32.46**	[8.16]	0.38	[0.04]	0.38	[0.03]
(a) PM–ordering						
w_t^{ttr}	**-21.38**	[4.7]	-1.47	[0.2]	0.27	[0.05]
w_t^{gs}	**-4.62**	[0.79]	-1.31	[0.14]	0.52	[0.03]
w_t^{gdp}	**-63.39**	[6.43]	0.35	[0.03]	-0.04	[0.07]

Note: The table reports the log of the Bayes Factors estimated via the log of SDDRs from equation (27) together with numerical standard errors provided in parentheses. The note to Table 1 applies.

In Figure 6(a) we show the marginal prior and posterior densities of the ω_1 parameter. The picture is very similar to that in Figure 4. In other words, the posterior in the 2023-sample has considerable mass away from the origin and, hence, strongly supports identified shocks, while the situation is much less clear for the 2006-sample and, for the MR-sample, where identification is clearly not supported because the prior and posterior densities are both centred at zero and have considerable density mass in the neighbourhood of zero.

Finally, we show the conditional variances based on the MR–ordering in Figure 7(a). Comparing that figure to Figure 5, it can be seen that the conditional variances are again very similar to those in the latter figure. Thus, the choice of $\mathbf{B}_0$ for standardizing

Figure 6: Marginal prior (solid line) and posterior (histograms) densities of the volatility of the log-volatility parameter, ω_1, across samples (based on alternative orderings).



Note: The note to Figure 4 applies.

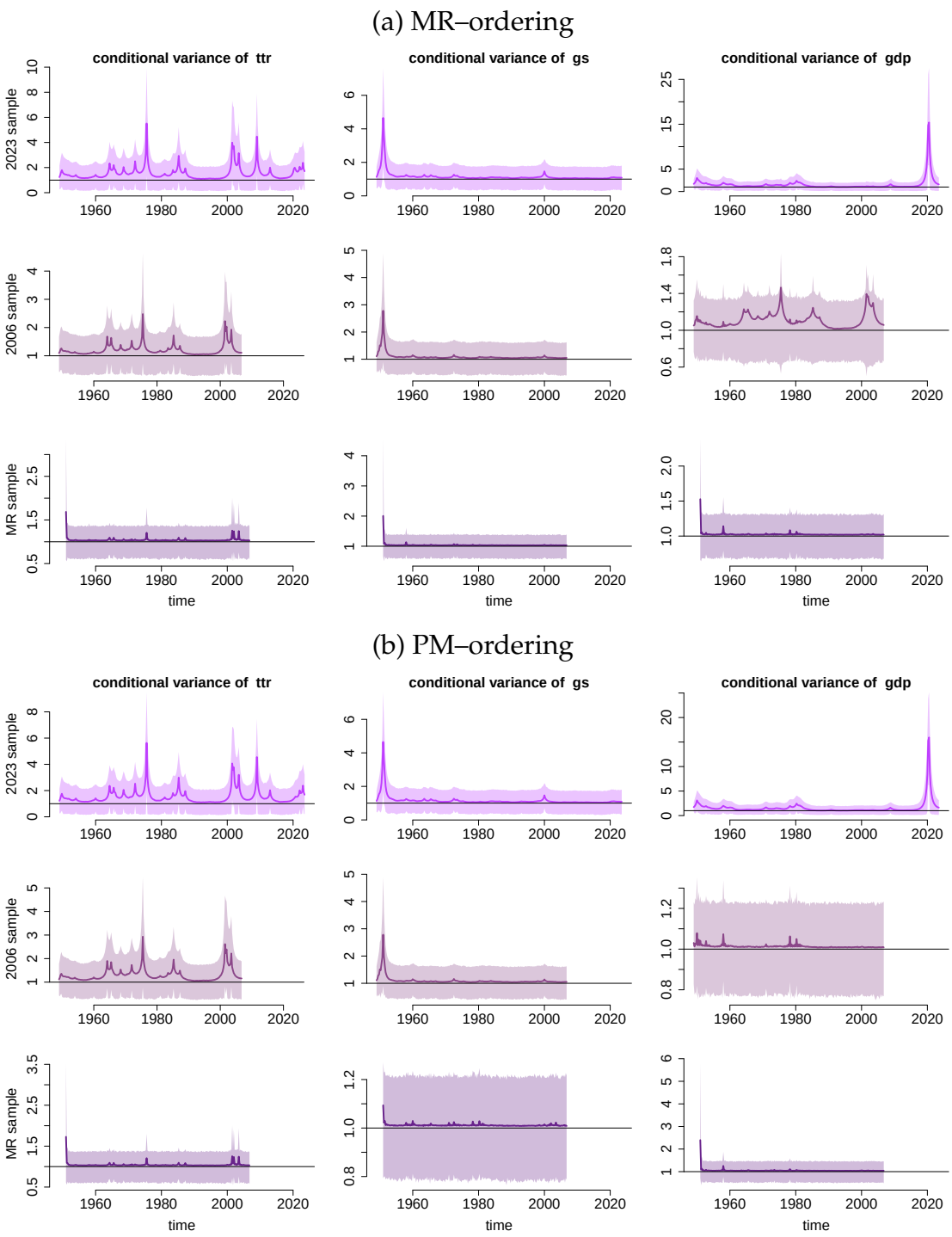
the posterior draws is of limited importance. At least, if there is sufficient conditional heteroskedasticity to ensure identification of the shocks, whether we use the BP– or the MR–ordering for the $\mathbf{B}_0$ drawings is not important.

We emphasize, however, that some kind of standardisation of the $\mathbf{B}_0$ drawings is necessary even if the shocks are all well identified because the structure of the model is invariant to changing the order and sign of the shocks. As long as the standardisation ensures a unique ordering and sign of the shocks, it should have little impact on the samples from the posterior distributions if the shocks are well identified. Therefore, given that for the 2023-sample we can expect to identify all three shocks through heteroskedasticity, we have also used a target matrix $\mathbf{B}_0$ for this sample which is not based on a set of estimates from some alternative identification scheme. Instead, we have used a selected posterior mode as the benchmark $\mathbf{B}_0$ matrix and call it the PM–ordering (see Appendix D for details).

In this case, it is not clear a priori that the ordering of the shocks will be the same as for the BP– and MR–orderings. As the shocks are distinguished by their conditional variances, we consider the conditional variances and order them such that they look similar to those based on the BP– and MR–orderings. In this case, the three distinct variance patterns allow for easily matching them with the shocks from the BP– and MR–orderings such that we can easily label the shocks correspondingly. We present the resulting conditional variances in Figure 7(b).

We have also computed SDDRs and the conditional variances of the three shocks, using the PM–ordering of the $\mathbf{B}_0$ drawings from the posterior. The results are shown in Table 2(b). They strongly support that all three shocks are identified through heteroskedasticity in the 2023-sample. In fact, the SDDR values in Table 2(b) are identical to the corresponding values for the 2023-sample in Table 1. Additionally, the robustness of heteroskedasticity and identification verification to various ordering rules is confirmed by the plots of marginal posterior and prior distributions of the volatility of the log-volatility parameter for the PM–ordering in Figure 6 (b) closely resembling other reported figures of this parameter. Thus, as long as some fixed ordering is used to

Figure 7: Conditional variance of structural shocks in the three samples (based on alternative orderings).



Note: The note to Figure 5 applies.

standardise the drawings from the posterior of $\mathbf{B}_0$, it does not affect the posterior of the conditional variances and, hence, the identification of the shocks.

7.4. The effects of tax shocks

Thus far, we have documented partial identification via heteroskedasticity of the tax shock in two of our samples. As the MR-sample does not support identification through heteroskedasticity of any of the shocks, we do not consider the MR-sample in the following. There is strong evidence for identification in the 2023-sample and much weaker in the 2006-sample. Subsequently we investigate how this reduced level of empirical support for identification affects the impulse responses of the tax shocks on gdp_t. Given that our identification results are robust with respect to different orderings of the posterior drawings, we now focus on the PM-ordering.

Given that heteroskedasticity provides three identified shocks, we begin by investigating which one of them is the tax shock. The properties of the conditional variances of the first shock in the PM–ordering closely resemble those of the tax shock in the BP– and MR–orderings. This fact makes it more likely that the first shock in the PM–ordering is the tax shock as well. We investigate this further and report the correlations between the structural shocks from our estimated models and PM–orderings and the narrative measure of the unanticipated tax shock by Romer and Romer (2010) in Table 3. The results show that the first shock in our models is the most correlated with the narrative measure. This correlation exceeds 0.22 for all the models reported in the table. It is much higher than for the second shock for which the values are -0.022 for the 2023-sample and 0.117 for the 2006-sample. They are still higher than for the third shock for which they are less than -0.15 for both samples. Therefore, irrespectively of the empirical support for identification across both samples, the first shock can be labeled the *tax shock*. The correlations for this shock have similar values as the tax shocks estimated by BP, MR, and LE, according to our reproductions of findings from other papers reported in the last three columns of Table 3.

As our first shock in the PM–ordering can be called the tax shock using its correlation

Table 3: Sample correlations between the narrative tax shock measure used by MR and the estimated structural shocks (based on the PM–ordering).

	2023-sample	2006-sample	BP results	MR results	LE results
w_t^{ttr}	0.224	0.264	0.277	0.298	0.233
w_t^{gs}	-0.022	0.030			
w_t^{gdp}	-0.154	-0.170			

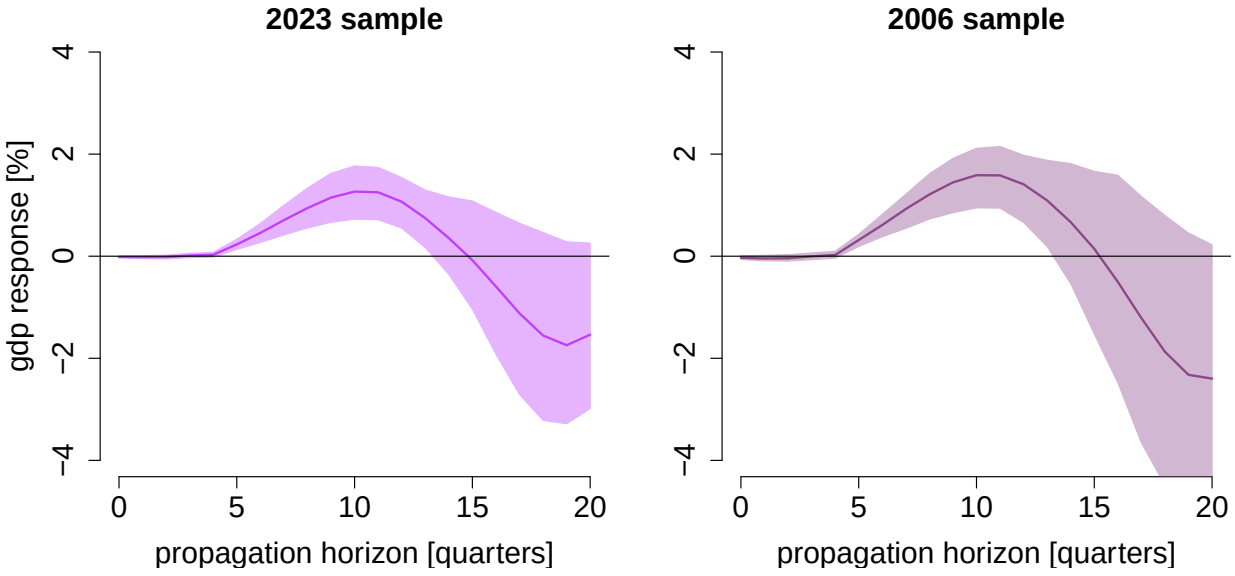
Note: The table reports sample correlations between the narrative measure of tax shocks proposed by Romer and Romer (2010) and used by MR. It indicates that our first shock is the most correlated with the instrument, which validates its labeling as the tax shock. The results in columns 2023-sample and 2006-sample are based on our posterior estimations where we used the posterior mean of the shocks as their estimator. The results in columns BP results, MR results, and LE results are based on our reproduction of the results from MR and LE using the authors' computer code and data.

with a narrative measure, we investigate its dynamic effects on gdp_t, a focal relationship in the other studies. In Figure 8 we report the corresponding impulse responses coming from our estimates for both samples. Following MR and LE, they represent gdp_t responses to a tax shock that reduces ttr_t by one percent of gdp_t in the last quarter of 2006. They share three common features, namely, (i) no effect on impact and for the first four quarters followed by (ii) an increase in gdp_t reaching a peak around 1.27 percent ten quarters after the impact, and (iii) becoming statistically insignificant after around three years. For the two samples they differ by the width of the 68% HDP intervals that are smaller for the 2023-sample and wider for the 2006-sample. This increase in width is driven by both the larger sample size of the 2023-sample and the reduced level of empirical support for identification of the tax shock in the 2006-sample.

Nevertheless, the impulse response estimates from the 2023- and 2006-samples are quite similar to each other. We further compare them to the impulse responses reported in BP[2], MR, and LE for the MR-sample, that is, the original data used by these authors. Figure 9 reports our estimates for the 2023-sample with 95% HDP intervals with the results from the BP, MR, and LE models reporting the maximum likelihood estimates

[2]Our BP results are based on the BP model estimated by MR.

Figure 8: Impulse responses of gross domestic product to a negative tax shock: our estimates (based on the PM–ordering).



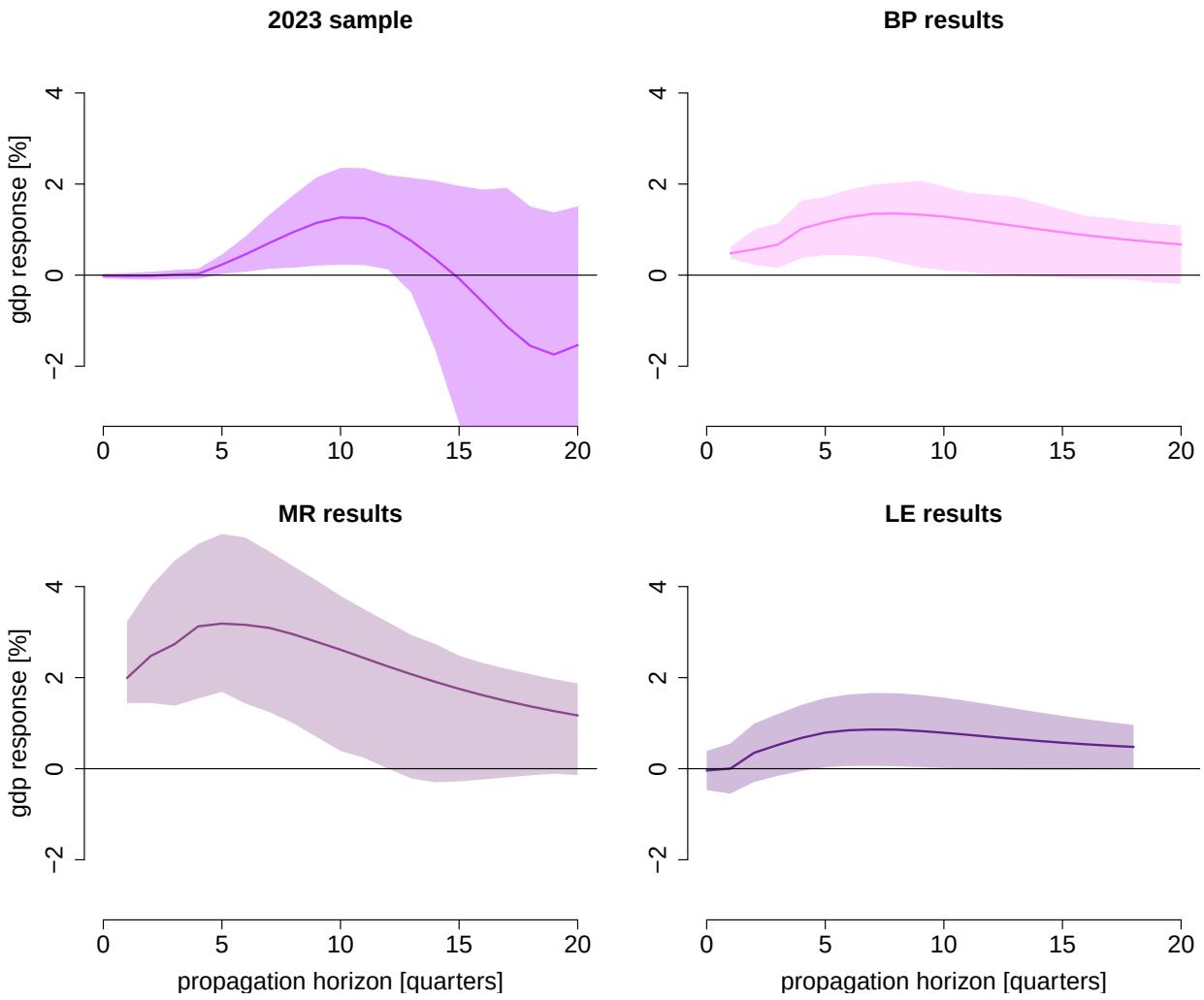
Note: The figure reports impulse responses of gdp_t to a negative tax shock lowering ttr_t by one percent of gdp_t value in the last quarter of 2006. The lines report the posterior medians and the shaded areas the 68% HPD point-wise intervals.

with the 95% confidence intervals. Our results share two features with other estimates, namely, the peak is reached in the mid-horizons, and the statistical significance lost around three years after the impact. Additionally, our peak response is similar to those in BP and LE, whereas MR obtain a larger peak. However, only the impulse responses reported by LE are statistically insignificant on impact and in the following four quarters as in our estimates, while those by BP and MR are positive and significant also on impact. Nevertheless, the conclusions from our estimates do not deviate far from those established in the literature and are obtained by identification via heteroskedasticity and shock labelling using narrative measures only.

8. Conclusions

In this paper, we provided general conditions for the identification of a structural shock via heteroskedasticity in multivariate dynamic structural models. These conditions are

Figure 9: Impulse responses of gross domestic product to a negative tax shock: comparison with other studies.



Note: The figure reports impulse responses of gdp_t to a negative tax shock lowering ttr_t by one percent of gdp_t value in the last quarter of 2006. In the 2023-sample plot, the line reports the posterior median and the shaded area reports the 95% HPD point-wise interval for the PM–ordering. In the remaining plots, the lines report the maximum likelihood estimator and the shaded areas–the 95% point-wise confidence intervals.

applicable to a wide range of heteroskedastic and conditionally heteroskedastic structural vector autoregressions and can also be used if only a subset of the shocks can be identified through heteroskedasticity. We also proposed flexible and easy to compute Savage-Dickey density ratios to verify the identification conditions. This was facilitated by specifying

new priors for time-varying conditional variances of the structural shocks. Such priors are flexible due to a hierarchical specification and ensure the standardisation using a specification centred at homoskedastic structural shocks. As a result, shock identification is indicated by the data and not implied by the prior.

These methods were illustrated by investigating whether the shocks in a system of fiscal variables are identified via heteroskedasticity. Identification is not confirmed for one of our samples but is supported when a longer sample period is used.

Our model is flexible and applicable to a wide range of time series in empirical macroeconomic and financial applications, which is facilitated by the code being available in an **R** package **bsvars** by Woźniak (2024a,b). Additionally, an important extension in which the structural matrix and the volatility of the log-volatility parameters change over time with a Markov process was recently proposed by Camehl and Woźniak (2024). That model facilitates verification through heteroskedasticity within Markov regimes.

References

Barndorff-Nielsen, O.E., 1997. Normal inverse gaussian distributions and stochastic volatility modelling. Scandinavian Journal of statistics 24, 1–13.

Bauwens, L., Lubrano, M., Richard, J., 1999. Bayesian inference in dynamic econometric models. Oxford University Press, USA.

Belmonte, M.A., Koop, G., Korobilis, D., 2014. Hierarchical Shrinkage in Time-Varying Parameter Models: Hierarchical Shrinkage in Time-Varying Parameter Models. Journal of Forecasting 33, 80–94. doi:10.1002/for.2276.

Bertsche, D., Braun, R., 2022. Identification of Structural Vector Autoregressions by Stochastic Volatility. Journal of Business & Economic Statistics 40, 328–341. doi:10.1080/07350015.2020.1813588.

Bitto, A., Frühwirth-Schnatter, S., 2019. Achieving shrinkage in a time-varying parameter model framework. Journal of Econometrics 210, 75–97.

Blanchard, O., 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, 1329–1368. doi:10.1162/003355302320935043.

Cadonna, A., Frühwirth-Schnatter, S., Knaus, P., 2020. Triple the gamma unifying shrinkage prior for variance and variable selection in sparse state space and tvp models. Econometrics 8, 20.

Callealta Barroso, F.J., García-Pérez, C., Prieto-Alaiz, M., 2020. Modelling income distribution using the log Student's t distribution: New evidence for European Union countries. Economic Modelling 89, 512–522. doi:10.1016/j.econmod.2019.11.021.

Camehl, A., Woźniak, T., 2024. Time-varying identification of monetary policy shocks. arXiv preprint arXiv:2311.05883 .

Carriero, A., Clark, T.E., Marcellino, M., 2019. Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors. Journal of Econometrics 212, 137–154. doi:10.1016/j.jeconom.2019.04.024.

Chan, J., Jeliazkov, I., 2009. Efficient Simulation and Integrated Likelihood Estimation in State Space Models. International Journal of Mathematical Modelling and Numerical Optimisation 1, 101–120.

Chan, J.C., Eisenstat, E., 2018. Bayesian model comparison for time-varying parameter vars with stochastic volatility. Journal of Applied Econometrics 33, 509–532.

Chan, J.C.C., 2018. Specification tests for time-varying parameter models with stochastic volatility. Econometric Reviews 37, 807–823. doi:10.1080/07474938.2016.1167948.

Chan, J.C.C., Koop, G., Yu, X., 2024. Large order-invariant bayesian vars with stochastic volatility. Journal of Business & Economic Statistics 42, 825–837. doi:10.1080/07350015.2023.2252039.

Clark, T.E., Ravazzolo, F., 2015. Macroeconomic Forecasting Performance Under Alternative Specification of Time-Varying Volatility. Journal of Applied Econometrics 30, 551–575.

Cogley, T., Sargent, T.J., 2005. Drifts and volatilities: Monetary policies and outcomes in the post WWII US. Review of Economic Dynamics 8, 262–302.

Doan, T., Litterman, R.B., Sims, C.A., 1984. Forecasting and Conditional Projection Using Realistic Prior Distributions. Econometric Reviews 3, 37–41.

Eddelbuettel, D., 2013. Seamless R and C++ Integration with Rcpp. Springer, New York, NY.

Eddelbuettel, D., François, R., Allaire, J., Ushey, K., Kou, Q., Russel, N., Chambers, J., Bates, D., 2011. Rcpp: Seamless r and c++ integration. Journal of statistical software 40, 1–18.

Eddelbuettel, D., Sanderson, C., 2014. RcppArmadillo: Accelerating R with high-performance C++ linear algebra. Computational Statistics & Data Analysis 71, 1054–1063.

Francis, N., Ramey, V.A., 2009. Measures of per capita hours and their implications for the technology-hours debate. Journal of Money, credit and Banking 41, 1071–1097.

Frühwirth-Schnatter, S., Wagner, H., 2010. Stochastic model specification search for Gaussian and partial non-Gaussian state space models. Journal of Econometrics 154, 85–100.

Gelfand, A.E., Smith, A.F.M., 1990. Sampling-Based Approaches to Calculating Marginal Densities. Journal of the American Statistical Association 85, 398–409. doi:10.2307/2289776.

Giannone, D., Lenza, M., Primiceri, G.E., 2015. Prior selection for vector autoregressions. Review of

Economics and Statistics 97, 436–451.

Herwartz, H., Lütkepohl, H., 2014. Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. Journal of Econometrics 183, 104–116.

Hogg, R.V., Klugman, S.A., 1983. On the estimation of long tailed skewed distributions with actuarial applications. Journal of Econometrics 23, 91–102. doi:10.1016/0304-4076(83)90077-5.

Hörmann, W., Leydold, J., 2014. Generating generalized inverse gaussian random variates. Statistics and Computing 24, 547–557.

Hosszejni, D., Kastner, G., 2021. Modeling Univariate and Multivariate Stochastic Volatility in *R* with **stochvol** and **factorstochvol**. Journal of Statistical Software 100.

Jarociński, M., 2024. Estimating the fed's unconventional policy shocks. Journal of Monetary Economics .

Kass, R.E., Raftery, A.E., 1995. Bayes Factors. Journal of the American Statistical Association 90, 773–795. doi:10.1080/01621459.1995.10476572.

Kastner, G., Frühwirth-Schnatter, S., 2014. Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models. Computational Statistics & Data Analysis 76, 408–423.

Kilian, L., Lütkepohl, H., 2017. Structural Vector Autoregressive Analysis. Cambridge University Press, Cambridge.

Lanne, M., Luoto, J., 2021. Gmm estimation of non-gaussian structural vector autoregression. Journal of Business & Economic Statistics 39, 69–81. doi:10.1080/07350015.2019.1629940.

Lanne, M., Lütkepohl, H., 2008. Identifying monetary policy shocks via changes in volatility. Journal of Money, Credit and Banking 40, 1131–1149.

Lanne, M., Lütkepohl, H., Maciejowska, K., 2010. Structural vector autoregressions with Markov switching. Journal of Economic Dynamics and Control 34, 121–131.

Lewis, D.J., 2021. Identifying Shocks via Time-Varying Volatility. The Review of Economic Studies 88, 3086–3124. doi:10.1093/restud/rdab009.

Leydold, J., Hörmann, W., 2017. GIGrvg: Random Variate Generator for the GIG Distribution. URL: https://CRAN.R-project.org/package=GIGrvg. r package version 0.5.

Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer-Verlag, Berlin.

Lütkepohl, H., Meitz, M., Netšunajev, A., Saikkonen, P., 2021. Testing identification via heteroskedasticity in structural vector autoregressive models. The Econometrics Journal 24, 1–22.

Lütkepohl, H., Milunovich, G., 2016. Testing for identification in SVAR-GARCH models. Journal of Economic Dynamics and Control 73, 241–258.

Lütkepohl, H., Netšunajev, A., 2017. Structural vector autoregressions with smooth transition in variances. Journal of Economic Dynamics and Control 84, 43–57.

Lütkepohl, H., Velinov, A., 2016. Structural vector autoregressions: Checking identifying long-run

restrictions via heteroskedasticity. Journal of Economic Surveys 30, 377–392.

Lütkepohl, H., Woźniak, T., 2020. Bayesian inference for structural vector autoregressions identified by Markov-switching heteroskedasticity. Journal of Economic Dynamics and Control 113, 103862. doi:10.1016/j.jedc.2020.103862.

McCausland, W.J., Miller, S., Pelletier, D., 2011. Simulation smoothing for state–space models: A computational efficiency analysis. Computational Statistics & Data Analysis 55, 199–212.

Meade, B., Lafayette, L., Sauter, G., Tosello, D., 2017. Spartan HPC-Cloud Hybrid: Delivering Performance and Flexibility. University of Melbourne doi:https://doi.org/10.4225/49/58ead90dceaaa.

Mertens, K., Ravn, M.O., 2014. A reconciliation of SVAR and narrative estimates of tax multipliers. Journal of Monetary Economics 68, S1–S19. doi:10.1016/j.jmoneco.2013.04.004.

Mountford, A., Uhlig, H., 2009. What are the effects of fiscal policy shocks? Journal of Applied Econometrics 24, 960–992. doi:10.1002/jae.1079.

Netšunajev, A., 2013. Reaction to technology shocks in Markov-switching structural VARs: Identification via heteroskedasticity. Journal of Macroeconomics 36, 51–62.

Olmsted, J., 2017. RcppTN: Rcpp-Based Truncated Normal Distribution RNG and Family. URL: https://CRAN.R-project.org/package=RcppTN. r package version 0.2-2.

Omori, Y., Chib, S., Shephard, N., Nakajima, J., 2007. Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference. Journal of Econometrics 140, 425–449.

Ramey, V.A., 2016. Macroeconomic shocks and their propagation. Handbook of macroeconomics 2, 71–162.

Rigobon, R., 2003. Identification through heteroskedasticity. Review of Economics and Statistics 85, 777–792.

Rigobon, R., Sack, B., 2003. Measuring the reaction of monetary policy to the stock market. Quarterly Journal of Economics 118, 639–669.

Robert, C.P., 1995. Simulation of truncated normal variables. Statistics and computing 5, 121–125.

Romer, C.D., Romer, D.H., 2010. The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks. American Economic Review 100, 763–801. doi:10.1257/aer.100.3.763.

Rubio-Ramírez, J.F., Waggoner, D.F., Zha, T., 2010. Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference. Review of Economic Studies 77, 665–696.

Sentana, E., Fiorentini, G., 2001. Identification, estimation and testing of conditionally heteroskedastic factor models. Journal of Econometrics 102, 143–164.

U.S. Bureau of Economic Analysis, 2024a. Table 1.1.5. Gross Domestic Product. Data. URL: https://apps.bea.gov/iTable/. Retrieved from NIPA Tables [accessed January 2, 2024].

U.S. Bureau of Economic Analysis, 2024b. Table 1.1.9. Implicit Price Deflators for Gross Domestic Product. Data. URL: https://apps.bea.gov/iTable/. Retrieved from NIPA Tables [accessed January 2, 2024].

U.S. Bureau of Economic Analysis, 2024c. Table 3.2. Federal Government Current Receipts and Expenditures.

Data. URL: https://apps.bea.gov/iTable/. Retrieved from NIPA Tables [accessed January 2, 2024].

U.S. Bureau of Economic Analysis, 2024d. Table 3.9.5. Government Consumption Expenditures and Gross Investment. Data. URL: https://apps.bea.gov/iTable/. Retrieved from NIPA Tables [accessed January 2, 2024].

U.S. Bureau of Labor Statistics, 2024. Population Level [CNP16OV]. Data. Federal Reserve Bank of St. Louis. URL: https://fred.stlouisfed.org/series/CNP16OV. Retrieved from FRED database [accessed January 2, 2024].

Verdinelli, I., Wasserman, L., 1995. Computing Bayes Factors Using a Generalization of the Savage-Dickey Density Ratio. Journal of the American Statistical Association 90, 614–618.

Waggoner, D.F., Zha, T., 2003a. A Gibbs sampler for structural vector autoregressions. Journal of Economic Dynamics & Control 28, 349–366.

Waggoner, D.F., Zha, T., 2003b. Likelihood preserving normalization in multiple equation models. Journal of Econometrics 114, 329–347. doi:10.1016/S0304-4076(03)00087-3.

Woźniak, T., 2024a. bsvars: Bayesian Estimation of Structural Vector Autoregressive Models. URL: https://CRAN.R-project.org/package=bsvars. R package version 3.0.0.

Woźniak, T., 2024b. Fast and Efficient Bayesian Analysis of Structural Vector Autoregressions Using the R package bsvars. Unpublished Manuscript. University of Melbourne.

Woźniak, T., Droumaguet, M., 2015. Assessing Monetary Policy Models: Bayesian Inference for Heteroskedastic Structural VARs. University of Melbourne Working Papers Series 2017.

Appendix A. Proofs

Appendix A.1. Proof of Theorem 1

Let $\mathbf{B}_*$ be a matrix that satisfies:

$$\mathbf{\Sigma}_t = \mathbf{B}_* \mathbf{\Lambda}_t \mathbf{B}_*', \qquad t = 0, 1, \ldots.$$

It will be shown that, under the conditions of Theorem 1, the n^{th} column of $\mathbf{B}_*$ must be the same as that of $\mathbf{B}$, except perhaps for a reversal of signs. Without loss of generality, it is assumed in the following that $n = 1$ because this simplifies the notation. In other words, it is shown that the first columns of $\mathbf{B}$ and $\mathbf{B}_*$ are the same except for a reversal of signs if $\sigma_1^2 \neq \sigma_i^2$, $i = 2, \ldots, N$.

There exists a nonsingular $(N \times N)$ matrix $\mathbf{Q}$ such that $\mathbf{B}_* = \mathbf{BQ}$. Using condition (4) for $t = 0$, $\mathbf{Q}$ has to satisfy the relation

$$\mathbf{BB}' = \mathbf{BQQ}'\mathbf{B}'.$$

Multiplying this relation from the left by $\mathbf{B}^{-1}$ and from the right by $\mathbf{B}^{-1'}$ implies that $\mathbf{QQ}' = \mathbf{I}_N$ and, hence, $\mathbf{Q}$ is an orthogonal matrix.

The relations

$$\mathbf{B\Lambda}_t\mathbf{B}' = \mathbf{BQ\Lambda}_t\mathbf{Q}'\mathbf{B}'$$

imply $\mathbf{\Lambda}_t = \mathbf{Q\Lambda}_t\mathbf{Q}'$ and, hence, $\mathbf{Q\Lambda}_t = \mathbf{\Lambda}_t\mathbf{Q}$ for all $t = 0, 1, \ldots$.

Denoting the $(i.j)^{\text{th}}$ element of $\mathbf{Q}$ by q_{ij}, the latter equation implies that

$$q_{n1}\sigma_1^2 = q_{n1}\sigma_n^2, \qquad n = 1, \ldots, N.$$

Hence, since σ_n^2 is different from σ_1^2 for $n = 2, \ldots, N$, we must have $q_{n1} = 0$ for $n = 2, \ldots, N$. Since, $\mathbf{Q}$ is orthogonal, the first column must then be

$$(1, 0, \ldots, 0)' \qquad \text{or} \qquad (-1, 0, \ldots, 0)'$$

which proves the theorem. Q.E.D.

Appendix A.2. Proof of Corollary 1

Consider the setup of Theorem 1 with $\mathbf{B} = \mathbf{B}_0^{-1}$. Then the arguments in the proof of Theorem 1 show that $\mathbf{B}_{0*}^{-1} = \mathbf{B}_0^{-1}\mathbf{Q}$, where $\mathbf{Q}$ is as in the proof of Theorem 1. Hence, $\mathbf{B}_{0*} = \mathbf{Q}'\mathbf{B}_0$, which shows that $\mathbf{B}_{0*}$ and $\mathbf{B}_0$ have the same n^{th} row up to sign. Q.E.D.

Appendix A.3. Proof of Corollary 2

To show that uniqueness of the n^{th} row of $\mathbf{B}_0$ implies a unique n^{th} column of $\mathbf{B}_0^{-1}$ we focus without loss of generality on the first row. If the first row of $\mathbf{B}_0$ is unique, any other admissible $\mathbf{B}_0$ matrix must be of the form $\mathbf{Q}\mathbf{B}_0$, where $\mathbf{Q}$ is an orthogonal matrix of the form:

$$\begin{bmatrix} 1 & \mathbf{0}_{(1\times(N-1))} \\ \mathbf{0}_{((N-1)\times 1)} & \mathbf{Q}_* \end{bmatrix},$$

with $\mathbf{Q}_*$ being an orthogonal $((N-1)\times(N-1))$ matrix. This fact is an easy implication of Theorem 1. Thus, any admissible inverse has the form $\mathbf{B}_0^{-1}Q'$ and, hence, has the same first column as $\mathbf{B}_0^{-1}$. Clearly, the same argument applies for any other row of $\mathbf{B}_0$, meaning that the impact effects of the n^{th} shock are unique if the n^{th} row of $\mathbf{B}_0$ is unique. This fact allows us to do impulse response analysis for a partially identified model. For each identified shock, unique impulse responses are obtained and can be easily computed in the usual way. Q.E.D.

Appendix B. Multivariate prior for stochastic volatility

Our prior assumptions also imply the joint distributions for the sequences of latent variables related to the volatility processes. In what follows, we first define two new multivariate distributions and use them subsequently to state the joint distributions of conditional variances and their logarithms.

Definition 3. (Multivariate normal product distribution) Let x be a scalar zero-mean normally distributed random variable with variance σ^2 that is independent of a $T \times 1$

zero-mean normal vector $\mathbf{Y}$ with covariance $\mathbf{\Sigma}$. Then, a random vector $\mathbf{Z} = x\mathbf{Y}$ follows a T-variate normal product distribution with zero mean and covariance equal to $\sigma^2\mathbf{\Sigma}$, denoted by $\mathbf{Z} \sim \mathcal{NP}_T(\sigma^2\mathbf{\Sigma})$, with density:

$$2^{-\frac{T-1}{2}}\pi^{-\frac{T+1}{2}}\det(\mathbf{\Sigma})^{-\frac{1}{2}}\left(\frac{1}{\sigma^2}\mathbf{Z}'\mathbf{\Sigma}^{-1}\mathbf{Z}\right)^{-\frac{T-1}{4}}K_{-\frac{T-1}{2}}\left(\sqrt{\frac{1}{\sigma^2}\mathbf{Z}'\mathbf{\Sigma}^{-1}\mathbf{Z}}\right). \tag{B.1}$$

□

Definition 4. (Multivariate log normal product distribution) Let a $T \times 1$ random vector $\mathbf{Z}$ follow a multivariate normal product distribution: $\mathbf{Z} \sim \mathcal{NP}_T(\sigma^2\mathbf{\Sigma})$. Then a $T \times 1$ random vector $\mathbf{Q} = \exp(\mathbf{Z})$ obtained by applying the exponent to each of the elements of $\mathbf{Z}$ follows the multivariate log normal product distribution, denoted by $\mathbf{Q} \sim \log\mathcal{NP}_T(\sigma^2\mathbf{\Sigma})$, with density:

$$\begin{aligned}&2^{-\frac{T-1}{2}}\pi^{-\frac{T+1}{2}}\det(\mathbf{\Sigma})^{-\frac{1}{2}}\\&\quad\times\det(\operatorname{diag}(\mathbf{Q}))^{-1}\left(\frac{1}{\sigma^2}\log(\mathbf{Q})'\mathbf{\Sigma}^{-1}\log(\mathbf{Q})\right)^{-\frac{T-1}{4}}K_{-\frac{T-1}{2}}\left(\sqrt{\frac{1}{\sigma^2}\log(\mathbf{Q})'\mathbf{\Sigma}^{-1}\log(\mathbf{Q})}\right).\end{aligned} \tag{B.2}$$

□

Note that the univariate (log-)normal product distributions are special cases of their multivariate versions for $T = 1$. The multivariate distributions are useful to state the following joint distributions for the sequences of volatilities:

Proposition 3. ***(Joint distributions of conditional volatilities)***
Given the prior specification from equations (8)–(12) and (16)–(22), the joint priors for the $T \times 1$ vectors containing the latent process $\mathbf{h}_n$, log-conditional variances $\log\boldsymbol{\sigma}_n^2 = \omega_n\mathbf{h}_n$, and conditional variances $\boldsymbol{\sigma}_n^2 = \exp(\omega_n\mathbf{h}_n)$ are given by the following T-variate normal, normal product, and log normal product distributions:

(a) $\mathbf{h}_n \mid \rho_n \sim \mathcal{N}_T\left(\mathbf{0}_{T\times 1}, \left(\mathbf{H}'_{\rho_n}\mathbf{H}_{\rho_n}\right)^{-1}\right)$,

(b) $\log \sigma_n^2 \mid \rho_n, \sigma_{\omega_n}^2 \sim \mathcal{NP}_T\left(\sigma_{\omega_n}^2 \left(\mathbf{H}_{\rho_n}' \mathbf{H}_{\rho_n}\right)^{-1}\right)$,

(c) $\sigma_n^2 \mid \rho_n, \sigma_{\omega_n}^2 \sim \log \mathcal{NP}_T\left(\sigma_{\omega_n}^2 \left(\mathbf{H}_{\rho_n}' \mathbf{H}_{\rho_n}\right)^{-1}\right)$. □

Appendix C. Gibbs sampler for the estimation of the parameters

This section scrutinises the estimation procedure that belongs to the class of MCMC methods. The assumptions regarding the distribution of residuals and the prior distribution of the parameters of the model result in a convenient and efficient Gibbs sampler that performs excellently even for larger systems of variables.

Appendix C.1. Sampling SVAR parameters

The conjugate prior distribution for matrix $\mathbf{B}_0$ results in a convenient generalised-normal full conditional posterior distribution that is proportional to:

$$p\left(\mathbf{B}_0 \mid \mathbf{y}, \mathbf{A}, \sigma_1^2, \ldots, \sigma_N^2, \gamma_0\right) \propto \det\left(|\mathbf{B}_0|\right)^{\overline{\nu}-N} \exp\left\{-\frac{1}{2}\sum_{n=1}^{N} \mathbf{B}_{0.n} \overline{\mathbf{S}}_n^{-1} \mathbf{B}_{0.n}'\right\} \tag{C.1}$$

$$\overline{\mathbf{S}}_n^{-1} = \mathbf{I}_N/\gamma_{0.n} + \sum_{t=1}^{T} \mathbf{u}_t \mathbf{u}_t'/\sigma_{n.t}^2 \tag{C.2}$$

$$\overline{\nu} = T + \underline{\nu} \tag{C.3}$$

The random number generator from this distribution follows the algorithm by Waggoner and Zha (2003a). Our experience clearly indicates its fast convergence and efficient extraction of the global shape of the posterior distribution, as pointed out by Woźniak and Droumaguet (2015).

In order to sample the autoregressive parameters $\mathbf{A}$, we follow the row-by-row algorithm by Chan et al. (2024) that reduces the number of operations to be performed by the computer by orders of magnitude in comparison to sampling all the parameters at once. Each of the rows, denoted by $\mathbf{A}_n$, is sampled from a conditional multivariate normal distribution given all other rows and parameters, and data. Denote by $\mathbf{A}_0^{(n)}$ an $N \times K$ matrix filled with the elements of matrix $\mathbf{A}$ and zeros in the n^{th} row, and an

$(Np+d)$–vector $\mathbf{x}_t = \begin{bmatrix} \mathbf{y}'_{t-1} & \dots & \mathbf{y}'_{t-p} & d'_t \end{bmatrix}'$. Then the structural-form model from equation (2) can be written as

$$\mathbf{B}_0 \left(\mathbf{y}_t - \mathbf{A}_0^{(n)} \mathbf{x}_t \right) = \left(\mathbf{B}_{0.n} \otimes \mathbf{x}'_t \right) \mathbf{A}'_n + \mathbf{w}_t. \tag{C.4}$$

Define an N–vector $\mathbf{z}_t^{(n)} = \mathbf{B}_0 \left(\mathbf{y}_t - \mathbf{A}_0^{(n)} \mathbf{x}_t \right)$ and an $N \times (NP+d)$ matrix $\mathbf{W}_t^{(n)} = \left(\mathbf{B}_{0.n} \otimes \mathbf{x}'_t \right)$. Then, the full conditional posterior distribution for the vector $\mathbf{A}_n$ is given by:

$$\mathbf{A}'_n \mid \mathbf{y}, \mathbf{A}_0^{(n)}, \mathbf{B}_0, \sigma_1^2, \dots, \sigma_N^2, \gamma_A \sim \mathcal{N}_{Np+d} \left(\overline{\mathbf{V}}_n \overline{\mathbf{A}}_n, \overline{\mathbf{V}}_n \right) \tag{C.5}$$

$$\overline{\mathbf{V}}_n^{-1} = \underline{\mathbf{\Omega}}^{-1} / \gamma_{A.n} + \sum_{t=1}^{T} \mathbf{W}_t^{(n)\prime} \operatorname{diag} \left(\sigma_{1.t}^2, \dots, \sigma_{N.t}^2 \right)^{-1} \mathbf{W}_t^{(n)} \tag{C.6}$$

$$\overline{\mathbf{A}}_n = \underline{\mathbf{\Omega}}^{-1} \underline{\mathbf{A}}'_n / \gamma_{A.n} + \sum_{t=1}^{T} \mathbf{W}_t^{(n)\prime} \operatorname{diag} \left(\sigma_{1.t}^2, \dots, \sigma_{N.t}^2 \right)^{-1} \mathbf{z}_t^{(n)} \tag{C.7}$$

where $\underline{\mathbf{A}}_n$ is the n^{th} row of $\underline{\mathbf{A}}$.

The hierarchy of the structural matrix hyper-parameters $\gamma_{0.n}$, $\underline{s}_{0.n}$, and $\underline{s}_{\gamma_0}$ is sampled from their respective full conditional posterior distributions:

$$\gamma_{0.n} \mid \mathbf{B}_{0.n} \sim \mathcal{IG}2 \left(\underline{s}_{0.n} + \mathbf{B}_{0.n} \mathbf{B}'_{0.n}, \quad \underline{\nu}_0 + N^2 \right) \tag{C.8}$$

$$\underline{s}_{0.n} \mid \gamma_B, \underline{s}_{\gamma_0} \sim \mathcal{G} \left((\underline{s}_{\gamma_0}^{-1} + (2\gamma_{0.n})^{-1})^{-1}, \underline{\nu}_{\gamma_0} + 0.5\underline{\nu}_0 \right) \tag{C.9}$$

$$\underline{s}_{\gamma_0} \mid \underline{s}_0 \sim \mathcal{IG}2 \left(\underline{s}_{s_0} + 2 \sum_{n=1}^{N} \underline{s}_{0.n}, \underline{\nu}_{s_0} + 2N\underline{\nu}_{\gamma_0} \right), \tag{C.10}$$

whereas the hierarchy of the autoregressive hyper-parameters $\gamma_{A.n}$, $\underline{s}_{A.n}$, and $\underline{s}_{\gamma_A}$ is sampled from:

$$\gamma_{A.n} \mid \mathbf{A}_n, \underline{s}_{A.n} \sim \mathcal{IG}2 \left(\underline{s}_{A.n} + \left(\mathbf{A}_n - \underline{\mathbf{A}}_n \right) \underline{\mathbf{\Omega}}^{-1} \left(\mathbf{A}_n - \underline{\mathbf{A}}_n \right)', \quad \underline{\nu}_A + Np + d \right) \tag{C.11}$$

$$\underline{s}_{A.n} \mid \gamma_{A.n}, \underline{s}_{\gamma_A} \sim \mathcal{G} \left((\underline{s}_{\gamma_A}^{-1} + (2\gamma_{A.n})^{-1})^{-1}, \underline{\nu}_{\gamma_A} + 0.5\underline{\nu}_A \right) \tag{C.12}$$

$$\underline{s}_{\gamma_A} \mid \underline{s}_A \sim \mathcal{IG}2 \left(\underline{s}_{s_A} + 2 \sum_{n=1}^{N} \underline{s}_{A.n}, \underline{\nu}_{s_A} + 2N\underline{\nu}_{\gamma_A} \right). \tag{C.13}$$

Appendix C.2. Sampling stochastic volatility parameters

The Gibbs sampler for the parameters of the SV processes results from our prior assumptions described in Section 4 and the assumption regarding the normality of the structural shocks in equation (7). It is facilitated using the auxiliary mixture sampler proposed by Omori et al. (2007). Transform equation (10) by squaring and taking the logarithm of its both sides and plugging in the expression from equation (8) obtaining:

$$\widetilde{w}_{n.t} = \omega_n h_{n.t} + \widetilde{\epsilon}_{n.t}, \tag{C.14}$$

where $\widetilde{w}_{n.t} = \log w_{n.t}^2$ and where $\widetilde{\epsilon}_{n.t} = \log \epsilon_{n.t}^2$. Given the standard normal assumption in equation (9), the distribution of $\widetilde{\epsilon}_{n.t}$ is $\log \chi_1^2$. This non-standard distribution is approximated precisely by a mixture of ten normal distributions defined by Omori et al. (2007). Applying the auxiliary mixture technique makes the linear equation (C.14) conditionally normal given the mixture component indicators, which greatly simplifies the sampling algorithm. This mixture of normals is specified by $s_{n.t} = 1, \dots, 10$ – the mixture component indicator for the n^{th} equation at time t, the normal component probability $\pi_{s_{n.t}}$, mean $\mu_{s_{n.t}}$, and variance $\sigma_{s_{n.t}}^2$. The latter three parameters are fixed and given in Omori et al. (2007), while $s_{n.t}$ augments the parameter space and is estimated. Its prior distribution is multinomial with probabilities $\pi_{s_{n.t}}$. Finally, define $T \times 1$ vectors: $\mathbf{s}_n = \begin{pmatrix} s_{n.1} & \dots & s_{n.T} \end{pmatrix}'$ collecting the realisations of $s_{n.t}$ for all t, $\boldsymbol{\mu}_{\mathbf{s}_n} = \begin{pmatrix} \mu_{s_{n.1}} & \dots & \mu_{s_{n.T}} \end{pmatrix}'$, and $\boldsymbol{\sigma}^2_{\mathbf{s}_n} = \begin{pmatrix} \sigma^2_{s_{n.1}} & \dots & \sigma^2_{s_{n.T}} \end{pmatrix}'$ collecting the nth equation auxiliary mixture means and variances, and $\widetilde{\mathbf{w}}_n = \begin{pmatrix} \widetilde{w}_{n.1} & \dots & \widetilde{w}_{n.T} \end{pmatrix}'$.

Sampling latent volatilities $\mathbf{h}_n$ proceeds independently for each n from the following T-variate normal distribution parameterised following Chan and Jeliazkov (2009) in terms

of its precision matrix $\overline{\mathbf{V}}_{h_n}$ and location vector $\overline{\mathbf{h}}_n$ as:

$$\mathbf{h}_n \mid \mathbf{y}, \mathbf{s}_n, B_0, B_+, \omega_n, \rho_n \sim \mathcal{N}_T\left(\overline{\mathbf{V}}_{h_n}\overline{\mathbf{h}}_n, \overline{\mathbf{V}}_{h_n}\right) \tag{C.15}$$

$$\overline{\mathbf{V}}_{h_n}^{-1} = \omega_n^2 \operatorname{diag}\left(\sigma_{\mathbf{s}_n}^{-2}\right) + \mathbf{H}_{\rho_n}'\mathbf{H}_{\rho_n} \tag{C.16}$$

$$\overline{\mathbf{h}}_n = \omega_n \operatorname{diag}\left(\sigma_{\mathbf{s}_n}^{-2}\right)\left(\widetilde{\mathbf{w}}_n - \boldsymbol{\mu}_{\mathbf{s}_n}\right) \tag{C.17}$$

The distinguishing feature of the precision matrix is that it is tridiagonal, which greatly improves the speed of generating random numbers from this full conditional posterior distribution if only the appropriate simulation smoother proposed by McCausland, Miller and Pelletier (2011) is implemented.

The parameters that are essential for the assessment of identification of the SVAR models, ω_n, are sampled independently from the following normal distribution:

$$\omega_n \mid \mathbf{y}, \mathbf{s}_n, h_n, \sigma_{\omega_n}^2 \sim \mathcal{N}\left(\overline{v}_{\omega_n}\overline{\omega}_n, \overline{v}_{\omega_n}\right) \tag{C.18}$$

$$\overline{v}_{\omega_n}^{-1} = \mathbf{h}_n' \operatorname{diag}\left(\sigma_{\mathbf{s}_n}^{-2}\right)\mathbf{h}_n + \sigma_{\omega_n}^{-2} \tag{C.19}$$

$$\overline{\omega}_n = \mathbf{h}_n' \operatorname{diag}\left(\sigma_{\mathbf{s}_n}^{-2}\right)\left(\widetilde{\mathbf{w}}_n - \boldsymbol{\mu}_{\mathbf{s}_n}\right) \tag{C.20}$$

Next, proceed to the ancillarity-sufficiency interweaving sampler proposed by Kastner and Frühwirth-Schnatter (2014). They show that sampling directly the parameters of the centred SV model leads to an efficient sampler if data is heteroskedastic, but it leads to substantial inefficiencies if data is homoskedastic. On the other hand, sampling directly parameters of the non-centred SV parameterisation leads to efficient sampling for homoskedastic data but not for heteroskedastic series. The solution offering the optimal strategy when the heteroskedasticity is uncertain, and to be verified, is to apply an ancillarity-sufficiency interweaving step in the Gibbs sampler. Our implementation proceeds as follows: Having sampled the random vector $\mathbf{h}_n$ and parameter ω_n, compute the parameters of the centred parameterisation $\tilde{h}_{n.t} = \omega_n h_{n.t}$ and

$\sigma^2_{v_n} = \omega^2_n$. Then, sample $\sigma^2_{v_n}$ from the following full conditional posterior distribution:

$$\sigma^2_{v_n} \mid \mathbf{y}, \tilde{\mathbf{h}}_n, \sigma^2_{\omega_n} \sim \mathcal{GIG}\left(-\frac{T-1}{2}, \tilde{\mathbf{h}}_n' \mathbf{H}_{\rho_n}' \mathbf{H}_{\rho_n} \tilde{\mathbf{h}}_n, \sigma^{-2}_{\omega_n}\right), \tag{C.21}$$

where $\tilde{\mathbf{h}}_n = \begin{pmatrix} \tilde{h}_{n.1} & \dots & \tilde{h}_{n.T} \end{pmatrix}$. Finally, compute $\omega_n = \pm\sqrt{\sigma^2_{v_n}}$ and $h_{n.t} = \frac{1}{\omega_n}\tilde{h}_{n.t}$ and return them as the MCMC draws for these parameters.

The autoregressive parameters of the SV equations are sampled independently from the following truncated normal distribution:

$$\rho_n \mid \mathbf{y}, h_n, \sigma^2_{\omega_n} \sim \mathcal{N}\left(\left(\sum_{t=0}^{T-1} h^2_{n.t}\right)^{-1}\left(\sum_{t=1}^{T} h_{n.t} h_{n.t-1}\right), \left(\sum_{t=0}^{T-1} h^2_{n.t}\right)^{-1}\right)\mathcal{I}\left(|\rho_n| < \sqrt{1-\sigma^2_{\omega_n}}\right). \tag{C.22}$$

This sampler is performed using the algorithm proposed by Robert (1995) and implemented in the **R** package **RcppTN** by Olmsted (2017).

The prior variances of parameter ω_n, $\sigma^2_{\omega_n}$, are *a posteriori* sampled independently from the following generalized inverse Gaussian distribution:

$$\sigma^2_{\omega_n} \mid \mathbf{y}, \omega_n \sim \mathcal{GIG}\left(\underline{A} - \frac{1}{2}, \omega^2_n, \frac{2}{\underline{S}}\right) \tag{C.23}$$

using the algorithm introduced by Hörmann and Leydold (2014) and implemented in the **R** package **GIGrvg** by Leydold and Hörmann (2017).

Finally, the auxiliary mixture indicators $s_{n.t}$ are each sampled independently from a multinomial distribution with the probabilities proportional to the product of the prior probabilities $\pi_{s_{n.t}}$ and the conditional likelihood function.

Appendix C.3. Computational considerations

The computations reproducing our results can be performed using the **R** package **bsvars** by Woźniak (2024a,b) that contains our data set with observations until 2022. It contains compiled code implementing the developed Gibbs sampler as well as the computations for the SDDR and other objects in **C++** using the **R** package **Rcpp** by Eddelbuettel,

François, Allaire, Ushey, Kou, Russel, Chambers and Bates (2011) and Eddelbuettel (2013) for convenient interfacing with **R** and the package **RcppArmadillo** by Eddelbuettel and Sanderson (2014) for algebraic operations and sampling random matrices. The **C++** source code for some low-level utility functions is taken from the open-source package **stochvol** by Hosszejni and Kastner (2021). The computations for this paper were performed at the Spartan HPC-Cloud Hybrid (see Meade, Lafayette, Sauter and Tosello, 2017) at the University of Melbourne.

Appendix D. Row sign and order normalisation

Heteroskedastic SVARs are identified up to the signs and orders of the rows of the structural matrix. Their practical application to the analysis of the sign and order dependent quantities requires transformation of the posterior sample so that it seems drawn from the posterior region corresponding to the selected row signs and order. We follow the normalisation practice by Lewis (2021) and choose the changes of row signs and order of the structural matrix that minimise a distance from the particular posterior draw to the benchmark structural matrix, denoted by $\widehat{\mathbf{B}}_0$. Let an $N \times N$ diagonal scaling matrix $\mathbf{D}$ with 1 or -1 on the main diagonal, and an $N \times N$ permutation matrix $\mathbf{P}$ represent the possible row sign and order transformation of $\mathbf{B}_0^{(s)}$, denoted by $\mathbf{PDB}_0^{(s)}$. We choose those $\mathbf{D}$ and $\mathbf{P}$ that minimise the likelihood-based distance proposed by Jarociński (2024): $\left\{\operatorname{vec}\left[\left(\mathbf{PDB}_0^{(s)} - \widehat{\mathbf{B}}_0\right)'\right]\right\}' \widehat{\mathbf{\Omega}}^{-1} \left\{\operatorname{vec}\left[\left(\mathbf{PDB}_0^{(s)} - \widehat{\mathbf{B}}_0\right)'\right]\right\}$, where $\widehat{\mathbf{\Omega}}$ is the covariance matrix of the asymptotic distribution of the maximum likelihood estimator evaluated at $\widehat{\mathbf{B}}_0$. Having chosen the row signs and order, the appropriately transformed draw of the structural matrix is returned and the equation ordering of the SV parameters and latent variables is adjusted accordingly.

Following Lewis (2021), we construct the benchmark $\widehat{\mathbf{B}}_0$ such that it matches the matrix

product on the left-hand side of the equation

$$\begin{bmatrix} \sigma_{ttr} & 0 & 0 \\ 0 & \sigma_{gs} & 0 \\ 0 & 0 & \sigma_{gdp} \end{bmatrix}^{-1} \begin{bmatrix} 1 & \theta_{gs} & 0 \\ \gamma_{ttr} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -\theta_{gdp} \\ 0 & 1 & -\gamma_{gdp} \\ -\zeta_{ttr} & -\zeta_{gs} & 1 \end{bmatrix} \begin{bmatrix} u_t^{ttr} \\ u_t^{gs} \\ u_t^{gdp} \end{bmatrix} = \begin{bmatrix} w_t^{ttr} \\ w_t^{gs} \\ w_t^{gdp} \end{bmatrix} \tag{D.1}$$

with the parameter values from the appropriate columns of Table 1 in Mertens and Ravn (2014).

The PM–ordering is chosen by drawing first from the posterior of $\mathbf{B}_0$ without paying attention to row ordering and sign. Such a sample from the posterior has modes corresponding to the various possible combinations of row signs and orderings. We pick one of the modes and use it for fixing the row signs and orderings in the posterior sample by choosing the row signs and orderings such that Jarociński's likelihood distance is minimised.