

Partial Identification of Heteroskedastic Structural Vector Autoregressions: Theory and Bayesian Inference

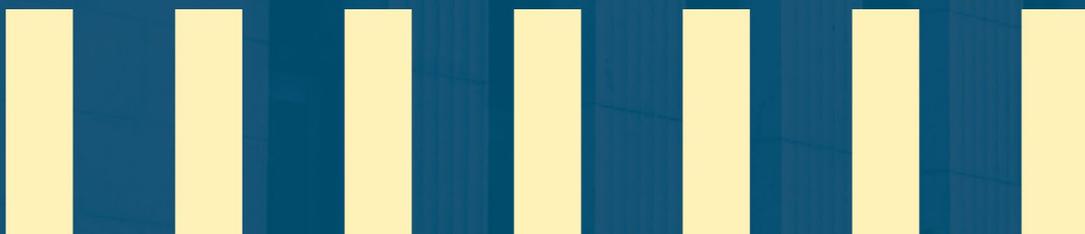
Helmut Lütkepohl
Freie Universität Berlin and DIW Berlin

Fei Shang
South China University of Technology and
Yuexiu Capital Holdings Group

Luis Uzeda
Canadian Economic Analysis Department
Bank of Canada
luzedagarcia@bankofcanada.ca

Tomasz Woźniak
University of Melbourne
tomasz.wozniak@unimelb.edu.au

Bank of Canada staff working papers provide a forum for staff to publish work-in-progress research independently from the Bank's Governing Council. This research may support or challenge prevailing policy orthodoxy. Therefore, the views expressed in this paper are solely those of the authors and may differ from official Bank of Canada views. No responsibility for them should be attributed to the Bank.



Acknowledgements

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

Abstract

We consider structural vector autoregressions that are identified through stochastic volatility. Our analysis focuses on whether a particular structural shock can be identified through heteroskedasticity without imposing any sign or exclusion restrictions. Three contributions emerge from our exercise: (i) a set of conditions that ensures the matrix containing structural parameters is either partially or globally unique; (ii) a shrinkage prior distribution for the conditional variance of structural shocks, centred on the hypothesis of homoskedasticity; and (iii) a statistical procedure for assessing the validity of the conditions outlined in (i). Our shrinkage prior ensures that the evidence for identifying a structural shock relies predominantly on the data and is less influenced by the prior distribution. We demonstrate the usefulness of our framework through a fiscal structural model and a series of simulation exercises.

Topics: Econometric and statistical methods; Fiscal policy

JEL codes: C11, C12, C32, E62

Résumé

Nous étudions des modèles vectoriels autorégressifs structurels qui comportent une volatilité stochastique. Nous cherchons à déterminer s'il est possible de repérer un choc structurel particulier à l'aide de l'hétéroscédasticité sans imposer des contraintes de signe et d'exclusion. Trois apports ressortent de notre analyse : i) un ensemble de conditions qui garantissent que la matrice comportant des paramètres structurels est partiellement ou globalement unique, ii) une mesure de rétrécissement liée à la distribution a priori pour la variance conditionnelle des chocs structurels, axée sur l'hypothèse de l'homoscédasticité, et iii) une méthode statistique permettant d'évaluer la validité des conditions énoncées au point i). Notre mesure de rétrécissement a priori garantit que les éléments probants utilisés pour repérer un choc structurel reposent principalement sur les données et sont moins influencés par la distribution a priori. Nous démontrons l'utilité de notre cadre au moyen d'un modèle budgétaire structurel et d'une série d'exercices de simulation.

Sujets : Méthodes économétriques et statistiques; Politique budgétaire

Codes JEL : C11, C12, C32, E62

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-normalized 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, one is often interested in identifying a specific shock, such as a monetary or 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 hand, 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 can 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, Koop and Yu, 2024](#)).

The first 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, which we provide more details on in Section 2. We further show that a shock having such a unique sequence of (conditional) variances leads to globally identified impulse response functions.

A second contribution of this paper is a formal characterization of the marginal prior for the conditional variance of structural shocks, derived from both conventional and alternative parameterizations of stochastic volatility processes. The latter involves (i) a specific hierarchical-prior setup, which we propose and (ii) the non-centred specification of state-space models, as proposed by [Kastner and Frühwirth-Schnatter \(2014\)](#), which we adapt to the context of SVAR models. In particular, 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 through heteroskedasticity for two reasons. First, they normalize the SVAR model facilitating the identification and estimation of conditional variances and the structural matrix. Second, this setup requires the evidence in favour of heteroskedasticity to rely more heavily on data. A Monte Carlo study shows that our approach is effective in achieving both objectives.

Importantly, the conditions for partial identification that we derive can be verified. In this regard, the third contribution of this paper is the development of a Savage-Dickey density ratio (SDDR) for the hypothesis of homoskedasticity. More precisely, we provide the conditions under which the analysis using the SDDR is feasible. Our verification procedure generalizes [Lütkepohl and Woźniak's \(2020\)](#) procedure with a more flexible process for conditional variance and extends that of [Chan \(2018\)](#) to SVARs.

In addition to the above-mentioned Monte Carlo study, we illustrate our methods by applying and comparing our approach to popular fiscal SVARs in the literature, namely those of [Blanchard and Perotti \(2002\)](#), [Mertens and Ravn \(2014\)](#), and [Lewis \(2021\)](#). Using our proposed framework, we find evidence of identification through heteroskedasticity for all shocks we consider in an extended sample covering the period up to 2023. Our empirical assessment relies on estimation via a Gibbs sampler that employs state-of-the-

art techniques, including the structural matrix sampler by [Waggoner and Zha \(2003a\)](#), autoregressive slope row-by-row sampling by [Chan et al. \(2024\)](#), the auxiliary mixture technique by [Omori, Chib, Shephard and Nakajima \(2007\)](#), and the ancillarity-sufficiency interweaving strategy by [Kastner and Frühwirth-Schnatter \(2014\)](#), which enables efficient simulation smoothing when heteroskedasticity is uncertain. All estimation procedures are accessible via the **R** package **bsvars** by [Woźniak \(2024a,b\)](#) implementing our algorithms using **C++**, which speeds up the computations by orders of magnitude.

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 generalized 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 non-parametric approach to identify shocks through heteroskedasticity and a test based on the autocorrelation of the reduced-form residuals that assumes non-proportional changes in the 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 the SDDR to verify identification using Markov-switching heteroskedasticity in a model with an arbitrary number of regimes.

The remainder of this paper is structured as follows. Section 2 discusses heteroskedastic SVARs and the general conditions we propose to identify them through heteroskedasticity. Section 3 presents the two approaches we consider for stochastic-volatility modeling. Section 4 characterizes the marginal prior for the conditional variance of the structural shocks that emerges from these approaches. Section 5 presents the conditions to test identification using the SDDR approach within our proposed framework. A Monte Carlo study and an empirical application to a fiscal SVAR are discussed in Sections 6 and 7, respectively. Section 8 concludes. An appendix

contains details on proofs, estimation procedures and additional results.

Notation

The following notation applies to the main text and technical appendix: \mathbf{y} denotes the available data, \mathbf{I}_N is the identity matrix of order N , $\mathbf{0}_{N \times N}$ and $\mathbf{1}_N$ are a matrix of zeros and a vector of ones of the indicated dimensions, respectively, the operator $\text{diag}(\cdot)$ puts the vector provided as its argument on the main diagonal of a diagonal matrix, the 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 that 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 stands 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 $\mathcal{IG2}$, 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 based on the variances of the structural shocks. Our results are applicable within a broad class of heteroskedastic SVAR models. Similarly general identification results can be found in [Lewis \(2021\)](#). These, however, are formulated as rank conditions to be verified by a frequentist test. In contrast, we state our conditions so as to facilitate their Bayesian verification.

Consider the following reduced-form VAR model of order p :

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

where \mathbf{y}_t is an N -dimensional vector of observable time series variables, \mathbf{A}_i , $i = 1, \dots, p$, 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 $\mathbf{u}_t = (u_{1,t}, \dots, u_{N,t})'$ 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 , using the $N \times N$ contemporaneous effects matrix \mathbf{B}_0 ,

$$\mathbf{B}_0 \mathbf{u}_t = \mathbf{w}_t, \quad (2)$$

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

$$\mathbf{\Lambda}_t = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2), \quad (3)$$

where the $\sigma_{n,t}^2$ are the unconditional or conditional variances. The assumptions for \mathbf{w}_t imply that the unconditional or conditional covariance matrices of the residuals \mathbf{u}_t , $\mathbb{E}[\mathbf{u}_t \mathbf{u}_t']$ and $\mathbb{E}[\mathbf{u}_t \mathbf{u}_t' | \mathbf{u}_{t-1}, \mathbf{u}_{t-2}, \dots]$, respectively, may be time-varying and are denoted by $\mathbf{\Sigma}_t$.

It is well-known that the structural matrix \mathbf{B}_0 is not identified without additional restrictions. Below, we state general conditions for partial identification of some of the parameters of \mathbf{B}_0 .

Theorem 1. *Let $\mathbf{\Sigma}_t$, $t = 0, 1, \dots$, be a sequence of positive definite $N \times N$ matrices and $\mathbf{\Lambda}_t = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2)$ 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}_0 such that*

$$\mathbf{\Sigma}_t = \mathbf{B}_0^{-1} \mathbf{\Lambda}_t \mathbf{B}_0^{-1'}, \quad t = 0, 1, \dots \quad (4)$$

Let $\sigma_n^2 = (1, \sigma_{n,1}^2, \sigma_{n,2}^2, \dots)$ be a possibly infinite dimensional vector. Then the n^{th} row of \mathbf{B}_0 is unique up to sign if $\sigma_n^2 \neq \sigma_i^2 \quad \forall i \in \{1, \dots, N\} \setminus \{n\}$.

Proof. The proof is given in [Appendix A.1](#). □

Note that the vectors σ_j^2 contain value one for the variance in period 0. This specific parameterization gives the elements of the vector σ_j^2 the interpretation of variances relative to the variances for $t = 0$. We are using relative variances in our theorem because it makes it easier to state the result, is in line with our normalization of the structural model, and also leads directly to the verification procedure discussed in [Section 5](#).

The theorem generalizes [Theorem 1](#) of [Lütkepohl and Woźniak \(2020\)](#), which presents an analogous result for structural errors with volatility changes generated by a homogeneous Markov switching process with finitely many volatility states. Our [Theorem 1](#) provides a general result on identification of a single equation through heteroskedasticity and also applies, for instance, if the volatility changes are generated by a different Markov process for each shock. 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. Our [Theorem 1](#) 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 \(2016\)](#), and [Lütkepohl and Netšunajev \(2017\)](#) are special cases of [Theorem 1](#) (see also [Lemma 1](#) in [Appendix A.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 are given by $\Sigma_t = \mathbf{B}_0^{-1} \Lambda_t \mathbf{B}_0^{-1'}$, where $\Lambda_t = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2)$ 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 generalizes results for full identification in [Sentana and Fiorentini \(2001\)](#), [Lewis \(2021\)](#), and [Bertsche and Braun \(2022\)](#) to the case of partial identification.

Structural impulse responses are computed from the reduced-form impulse responses $\Phi_i, i = 1, \dots$ obtained recursively as $\Phi_i = \sum_{j=1}^i \mathbf{A}_j \Phi_{i-j}$ for $i = 1, 2, \dots$, with $\Phi_0 = \mathbf{I}_N$ from the reduced-form VAR slope coefficients $\mathbf{A}_1, \dots, \mathbf{A}_p$ using $\mathbf{A}_j = 0$ for $j > p$ ([Lütkepohl, 2005](#), Section 2.1.2). The structural impulse responses are the elements of the matrices $\Theta_i = \Phi_i \mathbf{B}_0^{-1}, i = 0, 1, \dots$. Thus, for computing them, the structural matrix \mathbf{B}_0 is needed. In particular, if just one shock is identified through heteroskedasticity, the following result formalizes the implications of Theorem 1 for impulse response analysis:

Corollary 1. *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 Φ_i by the n^{th} column of \mathbf{B}_0^{-1} .*

Proof. See [Appendix A.2](#). □

3. Two approaches to parameterize stochastic volatility

The previous section discussed the conditions required to identify SVARs under a broad class of time-varying volatility models. We now shift our focus to heteroskedastic SVARs, where shocks are modeled and identified through stochastic volatility. Specifically, we examine two parameterizations of this process: the centred and non-centred approaches.

The centred parameterization

To fix ideas, we first present the more conventional (centred) parameterization for modeling stochastic volatility. In this setup, each diagonal element of Λ_t in (3) is

parameterized as follows:

$$\sigma_{n,t}^2 = \exp(\tilde{h}_{n,t}), \quad (5)$$

$$\tilde{h}_{n,t} = \rho_n \tilde{h}_{n,t-1} + \tilde{v}_{n,t} \quad \text{s.t.} \quad \rho_n \in (-1, 1), \quad (6)$$

$$\tilde{v}_{n,t} \sim \mathcal{N}(0, \omega_n^2) \quad \text{for } n = 1, \dots, N \text{ and } t = 1, \dots, T. \quad (7)$$

Theorem (1) implies that identification depends on the sequence of conditional variances, $\{\sigma_{n,t}^2\}_{t=1}^T$, associated with a shock being unique and not proportional to any sequence of volatilities from another shock in the system. These conditions are satisfied under (5)–(7), as this representation assumes that the log-volatility state variable $\tilde{h}_{n,t}$ (and thus $\sigma_{n,t}^2$) evolves stochastically as a stationary autoregressive process.

In this setup, assessing identification through stochastic volatility reduces to testing whether $\omega_n^2 = 0$. If that is the case, the corresponding shock is homoskedastic. Conversely, if $\omega_n^2 \neq 0$, then by construction $\{\sigma_{n,t}^2\}_{t=1}^T$ is unique, and the condition for identification in Theorem 1 is satisfied.

Nevertheless, implementing statistical tests for $\omega_n^2 = 0$ is challenging because zero lies at the boundary of the parameter space for ω_n^2 . Moreover, Bayesian methods that estimate stochastic volatility models under the centred parameterization typically use an inverse-gamma prior for ω_n^2 , whose domain is undefined at zero (see [Appendix B](#)). To address these issues, we adopt the non-centred parameterization for $\sigma_{n,t}^2$ discussed next.

The non-centred parameterization

To obtain the non-centred parameterization of $\sigma_{n,t}^2$, akin to [Kastner and Frühwirth-Schnatter \(2014\)](#) and [Chan \(2018\)](#), we first define $\tilde{h}_{n,t} = \omega_n h_{n,t}$. This transforms (5)–(7) into

$$\sigma_{n,t}^2 = \exp(\omega_n h_{n,t}), \quad (8)$$

$$h_{n,t} = \rho_n h_{n,t-1} + v_{n,t} \quad \text{s.t.} \quad \rho_n \in (-1, 1), \quad (9)$$

$$v_{n,t} \sim \mathcal{N}(0, 1) \quad \text{for } n = 1, \dots, N \text{ and } t = 1, \dots, T. \quad (10)$$

We further assume that $h_{n,0} = 0$, ensuring $\sigma_{n,0}^2 = 1$ to satisfy the normalization condition in Theorem 1.

Importantly, the conditions in Theorem 1 hold under more general specifications for SV innovations, that is, when the innovations driving $h_{n,t}$ and $\tilde{h}_{n,t}$ are correlated, given that correlation does not imply proportionality of $\{\sigma_{n,t}^2\}_{t=1}^T$ across equations. Even perfect correlation of volatility shocks $v_{n,t}$ and $\tilde{v}_{n,t}$ does not imply proportional changes of the conditional variances $\sigma_{n,t}^2$ due to the non-linear transformation in Equations (8) and (5), respectively. This proportionality arises only when the shocks are perfectly correlated, parameters ρ_n and ω_n^2 are equal to their counterparts across equations, and $\omega_n^2 \neq 0$, an extreme scenario excluded in our framework.

While there is a one-to-one mapping between the representations in (5)–(7) and (8)–(10), they have markedly different implications for (i) the marginal prior distribution of $\sigma_{n,t}^2$ and (ii) the feasibility of proposing a statistical method to assess shock identification. These implications are elaborated in Sections 4 and 5.

Two more comments are in order. First, unlike the centred parameterization, the non-centred representation for $\sigma_{n,t}^2$ is cast in terms of the standard deviation parameter ω_n instead of ω_n^2 . Given that ω_n can take both positive and negative values on a real line, the non-centred approach allows us to elicit priors for which ω_n is defined at zero. As shown later in Section 4.1, we propose a conditionally normal prior for ω_n centred at zero, which implies a gamma prior for ω_n^2 . The gamma prior allocates more mass near $\omega_n^2 = 0$ compared to the inverse-gamma prior typically used in the centred approach (see Chan, 2018). Consequently, the non-centred approach enables stronger shrinkage toward homoskedasticity.

Second, it is easy to see from (8) that the likelihood function is invariant to sign at the $(\omega_n, h_{n,t})$ ordinate. This follows from the fact that both $(\omega_n, h_{n,t})$ and $(-\omega_n, -h_{n,t})$ yield the same value for $\sigma_{n,t}^2$. Consequently, the posterior for ω_n may be bimodal or unimodal around zero. Bimodality will only occur if ω_n (and, consequently ω_n^2) is far from zero. Therefore, bimodality of the posterior for ω_n provides evidence that $\sigma_{n,t}^2 \neq 0$, supporting the identification of structural shocks through stochastic volatility. For the purpose of

identification, a bimodal (as opposed to unimodal) posterior for ω is desirable. We return to this point in the context of our empirical application in Section 7.

Having distinguished two approaches to model $\sigma_{n,t}^2$, for the remainder of this paper, we adopt the non-centred parameterization unless explicitly stated otherwise.

4. The marginal prior for $\sigma_{n,t}^2$

Recall that Theorem 1 is stated in terms of $\sigma_{n,t}^2$. In the context of Bayesian estimation, $\sigma_{n,t}^2$ can be characterized through its marginal prior distribution. However, as shown in the previous section, $\sigma_{n,t}^2$ is a non-linear function of $h_{n,t}$, ω_n , and ρ . This non-linearity complicates the assessment of how the choice of priors for these variables affects the marginal prior for $\sigma_{n,t}^2$.

To shed light on this matter, this section provides a detailed examination of the marginal prior for $\sigma_{n,t}^2$. This is achieved in two steps. First, the priors for the parameters that underlie $\sigma_{n,t}^2$, namely ω_n and ρ , are specified. Second, the proposed marginal prior for $\sigma_{n,t}^2$ is characterized, illustrating how it ensures centring and shrinkage towards a homoskedastic SVAR. In what follows, we focus on the characterization of a univariate prior for $\sigma_{n,t}^2$ and provide more general results for a multivariate distribution of $\{\sigma_{n,t}^2\}_{t=1}^T$ in [Appendix B](#).

4.1. Priors for the parameters underlying $\sigma_{n,t}^2$

Once again, the parameters associated with the non-centred representation of $\sigma_{n,t}^2$ are ω_n , the essential parameter in our setup that determines whether $\sigma_{n,t}^2$ changes over time, and ρ_n , the autoregressive parameter of the latent process $h_{n,t}$. We assume the following hierarchical prior structure for these parameters:

$$\omega_n \mid \sigma_{\omega_n}^2 \sim \mathcal{N}(0, \sigma_{\omega_n}^2), \quad (11)$$

$$\sigma_{\omega_n}^2 \mid \rho_n \sim \mathcal{G}(\underline{S}, \underline{A}) \mathcal{I}(0 < \sigma_{\omega_n}^2 < 1 - \rho_n^2), \quad (12)$$

$$\rho_n \mid \sigma_{\omega_n}^2 \sim \mathcal{U}\left(-\sqrt{1 - \sigma_{\omega_n}^2}, \sqrt{1 - \sigma_{\omega_n}^2}\right), \quad (13)$$

where $\sigma_{\omega_n}^2$ denotes the prior variance of ω_n .

The prior specification for ω_n in (11) extends the one proposed by Chan (2018). Specifically, instead of fixing the prior variance as in Chan (2018), we adopt a hierarchical prior in which $\sigma_{\omega_n}^2$ follows the gamma distribution stated in (12). Consequently, our specification allows for the estimation of $\sigma_{\omega_n}^2$, making the prior for ω_n less dependent on arbitrary choices. Moreover, based on the results of Bitto and Frühwirth-Schnatter (2019) and Cadonna, Frühwirth-Schnatter and Knaus (2020), marginalizing the prior for ω_n over $\sigma_{\omega_n}^2$ yields a prior that combines extreme shrinkage towards homoskedasticity with heavy tails. The latter accommodates heteroskedasticity when it arises from strong data signals.

We complement the three priors above with the following three restrictions:

$$\frac{\sigma_{\omega_n}^2}{1 - \rho_n^2} \leq 1, \quad (14)$$

$$\underline{A} > 0.5, \quad (15)$$

$$|\rho_n| < 1. \quad (16)$$

Restriction (14) ensures the desired level of centring and shrinkage in our proposed marginal prior for $\sigma_{n,t}^2$, which we show formally in Section 4.2. It restricts the prior variances from Proposition 1 presented below in the limit $\lim_{t \rightarrow \infty} \sigma_{\omega_n}^2 \frac{1 - \rho_n^{2t}}{1 - \rho_n^2} = \sigma_{\omega_n}^2 / (1 - \rho_n^2)$ to ensure that the condition holds for variances at all periods t . The restriction in (15) determines the marginal prior for ω_n , making it particularly suitable for our setup. Restriction (16) is standard and ensures that $h_{n,t}$ in (9) is stationary. Additionally, Restrictions (16) and (14) determine the bounds for ρ_n as expressed in the uniform prior for ρ_n in (13). Similarly, the truncation of the gamma prior for $\sigma_{\omega_n}^2$ stated in (12) arises from Restriction (14). We provide further elaboration on these restrictions later in this section.

4.2. Characterizing the marginal prior for $\sigma_{n,t}^2$

This section provides a detailed description of the marginal prior for $\sigma_{n,t}^2$. As discussed in Section 1, a key feature of our prior setup for $\sigma_{n,t}^2$ is to ensure that this prior is not only centred on the hypothesis of a homoskedastic SVAR but also provides shrinkage toward it. In this regard, Definitions 1 and 2, along with Proposition 1, presented below, will be instrumental in structuring the approach to achieving these objectives.

Definition 1. Normal product distribution

Let x and y denote two independent zero-mean normally distributed random variables with variances σ_x^2 and σ_y^2 , respectively. Then, random variable $z = xy$ follows the normal product distribution with zero mean and variance $\sigma_z^2 = \sigma_x^2\sigma_y^2$, denoted by $z \sim \mathcal{NP}(\sigma_z^2)$, and density function given by $\frac{1}{\pi\sqrt{\sigma_z^2}}K_0\left(\frac{|z|}{\sqrt{\sigma_z^2}}\right)$. \square

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 a change of variables.

Definition 2. Log-normal product distribution

Let random variable z follow the normal product distribution with variance σ_z^2 . Then, random variable $q = \exp(z)$ follows the log-normal product distribution, denoted $q \sim \log \mathcal{NP}(\sigma_z^2)$, with density given by: $\frac{1}{\pi\sqrt{\sigma_z^2}q}K_0\left(\frac{|\log q|}{\sqrt{\sigma_z^2}}\right)$. \square

Based on the results from Definitions 1 and 2, we can state Proposition 1:

Proposition 1. Auxiliary results on conditional distributions for $h_{n,t}$ and $\sigma_{n,t}^2$

Given the prior specification from Equations (8)–(7) and (11)–(15), the marginal priors for the latent process $h_{n,t}$, log-conditional variances $\log \sigma_{n,t}^2 = \omega_n h_{n,t}$, and conditional variances $\sigma_{n,t}^2 = \exp(\omega_n h_{n,t})$ are given by the following normal, normal product, and log normal product distributions:

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

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

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

Proof. **(a)** The result is based on the properties of a normal compound distribution that facilitates the integration of $\int p(h_{n,t}, h_{n,t-1}, \dots, h_{n,1}) d(h_{n,t-1}, \dots, h_{n,1})$, where the joint distribution under the integral is constructed from the conditional distributions $h_{n,t} \mid h_{n,t-1}, \dots, h_{n,1} \sim \mathcal{N}(\rho_n h_{n,t-1}, 1)$ and using $h_{n,0} = 0$. **(b)** The result is obtained directly by applying Definition 1, the result **(a)** and the prior in Expression (11). Point **(c)** is obtained as a straightforward consequence of the first two results and Definition 2. \square

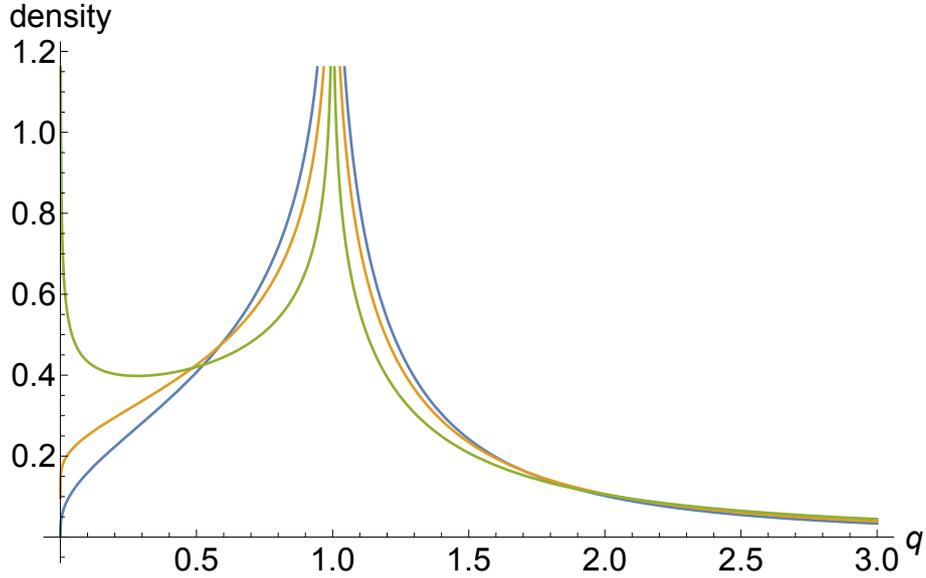
The introduced results facilitate centring and shrinking our prior for $\sigma_{n,t}^2$ toward a homoskedastic SVAR, which is useful because it ensures that evidence supporting heteroskedasticity – and, consequently, the identification of a shock – must come from the data. Moreover, it provides an alternative strategy for normalizing SVARs that does not rely on common approaches, such as setting the diagonal elements of \mathbf{B}_0 to one or imposing that the expected value of $\sigma_{n,t}^2$ equals one. Both of these approaches complicate the derivation of an efficient Bayesian estimation algorithm. In what follows, we discuss how we achieve centring and shrinking of our prior for $\sigma_{n,t}^2$.

To centre the prior for $\sigma_{n,t}^2$ around the hypothesis of homoskedasticity, we must ensure that the log-normal product distribution characterizing $\sigma_{n,t}^2$ has a single pole at the value 1. This follows directly from two points: (i) the normalization condition in Theorem 1, which sets $\sigma_{n,0}^2 = 1$, and (ii) the fact that homoskedasticity in our setup corresponds to setting $\omega_n = 0$, which implies $\sigma_{n,t}^2 = \exp(\omega_n h_{n,t}) = 1$, as discussed in Section 3. Property 1, presented below, establishes when the log-normal product for $\sigma_{n,t}^2$ is proper and has a single pole at 1.

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 when its variance satisfies $\sigma_z^2 \leq 1$. In this case, the value of the density function approaches infinity when its argument, q , goes to 1 and approaches 0 when q goes to 0 from the right.

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



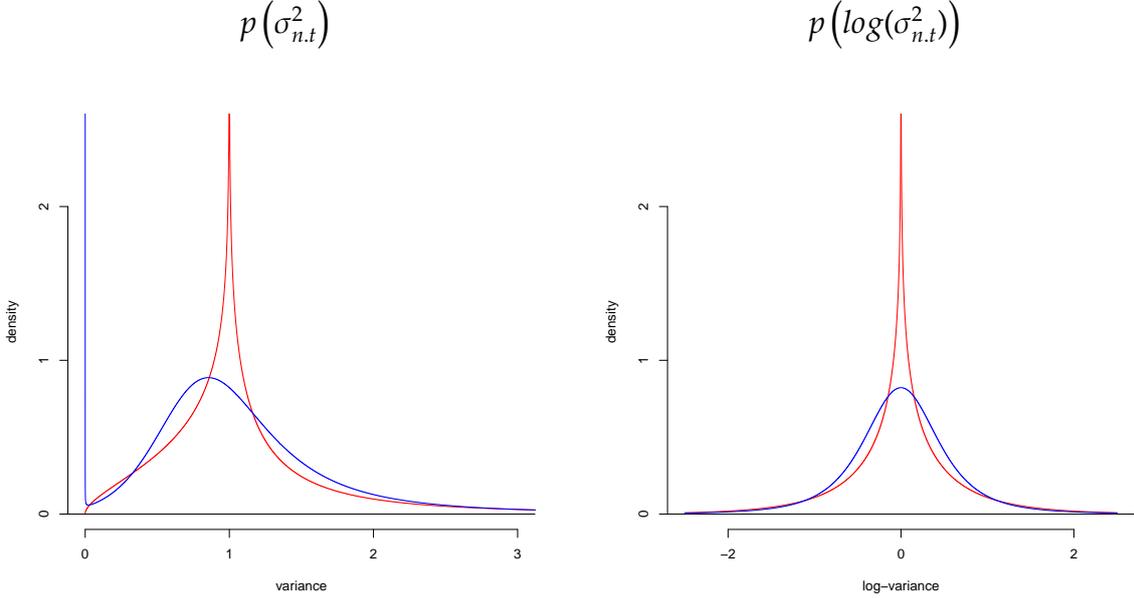
Note: The blue, orange, and green lines correspond to the densities for the values of the scale parameter σ_z^2 equal to 0.8, 1, and 1.5, respectively.

If $\sigma_z^2 > 1$, this distribution has an additional pole at 0, hence approaching infinity as q goes to either 0 or 1.

Figure 1 illustrates Property 1. A few points are worth highlighting. First, note that the condition for a single pole at one was stated in (14), which denotes a restriction on the variance of the prior limiting distribution for $\sigma_{n,t}^2$, as characterized in Proposition 1. This restriction ensures that the inequality in (14) holds for all t . Second, the single-pole-at-one condition implies a strong concentration of the prior probability mass for $\sigma_{n,t}^2$ at the value corresponding to the homoskedasticity of the structural shocks. This prior is equation invariant and, thus, it supports our claim that, at the prior mode, the SVAR model is not identified through heteroskedasticity.

The prior shrinkage for $\sigma_{n,t}^2$ is also achieved through Property 1 via the inequality restriction in (14). Specifically, this restriction prevents the prior probability mass for $\sigma_{n,t}^2$ from being distributed more evenly over the interval from 0 to 1, as would occur in the presence of an additional pole at zero as for the density plotted in green in Figure 1.

Figure 2: The marginal priors for $\sigma_{n,t}^2$ and $\log(\sigma_{n,t}^2)$ in their centred and non-centred parameterizations



Note: The densities in red and blue correspond to the non-centred and centred approaches, respectively, to model $\sigma_{n,t}^2$, as discussed in Section 3.

The centering and shrinkage effects resulting from our prior setup are more evident in Figure 2, which compares the marginal prior distributions for $\sigma_{n,t}^2$ and $\log(\sigma_{n,t}^2)$ based on their centred and non-centred parameterizations.¹

Notably, the marginal prior for $\sigma_{n,t}^2$ in the non-centred case, denoted by the red line in Figure 2, inherits the properties from the log-normal product conditional prior for $\sigma_{n,t}^2$ given ρ_n and $\sigma_{\omega_n}^2$, stated in Proposition 1(c) with a less-than-one restriction on its variance from Expression (14). These properties are convergence to value zero when the conditional variance goes to zero from the right, a pole at 1, strong shrinkage toward the prior mode, and heavy tails.

¹The marginal priors in Figure 2 are computed using the numerical integration of Gelfand and Smith (1990) in two steps. In the first one for the non-centred parameterization, a sample of S draws is obtained from the prior distributions, denoted by $\{\rho_n^{(s)}, \sigma_{\omega_n}^{2(s)}\}_{s=1}^S$. In the second step, the marginal prior ordinates at pre-specified points, denoted by ζ_g for $g = 1, \dots, G$, are each computed by $\widehat{p}(\sigma_{n,t}^2 = \zeta_g) = S^{-1} \sum_{s=1}^S p(\sigma_{n,t}^2 = \zeta_g | \rho_n^{(s)}, \sigma_{\omega_n}^{2(s)})$. Appropriate modifications reflecting the prior assumptions are made for the centred parameterization computations.

In contrast, in the centred case, the marginal prior for $\sigma_{n,t}^2$ has the properties of the log-Student-t distribution revisited in [Appendix B.3](#), that is, less concentration around the hypothesis of homoskedasticity, a mode at one, and a pole at zero. This log-Student-t distribution arises from the inverse gamma prior for ω_n^2 featured by the centred parameterization and the exponential transformation of the log-volatility to conditional variance. In summary, the prior distribution for $\sigma_{n,t}^2$ in the centred parameterization favours heteroskedasticity, implies shock identification even in the absence of time-varying volatility, and does not support the normalization of the conditional variances at one.

5. Bayesian Identification Verification

Recall from [Section 3](#) that in our non-centred setup, identifying a given shock n through heteroskedasticity involves assessing the restriction $\omega_n = 0$. If this restriction holds, then $h_{n,t} = 0$ and $\sigma_{n,t}^2 = 1$ for all t , which corresponds to a homoskedastic shock. Conversely, if $\omega_n \neq 0$, then the conditions in [Theorem 1](#) ensure shock identification through heteroskedasticity.

To assess $\omega_n = 0$, we adopt a Bayesian approach by comparing the fit of homoskedastic and a partially heteroskedastic SVAR using the Bayes factor. To compute the Bayes factor, we use the SDDR approach, which defines the Bayes factor as

$$BF_{homosk} = \frac{p(\omega_n = 0|\mathbf{y})}{p(\omega_n \neq 0)}. \quad (17)$$

The numerator in [\(17\)](#) is computed using numerical integration methods based on the estimator proposed by [Gelfand and Smith \(1990\)](#). This approach only requires the full conditional posterior distribution of ω_n to be known up to its probability density function, which is normal, and the posterior draws from the unrestricted model ($\omega_n \neq 0$). Consequently, computation of [\(17\)](#) requires estimation only under the unrestricted, that is, heteroskedastic SVAR. [Appendix D](#) provides a detailed description of the evaluation of the marginal posterior density $p(\omega_n = 0|\mathbf{y})$.

The denominator $p(\omega_n = 0)$ involves the marginal prior, which is obtained by integrating out $\sigma_{\omega_n}^2$ from the hierarchical-prior structure of ω_n , discussed in Section 4.1, where $\omega_n | \sigma_{\omega_n}^2 \sim \mathcal{N}(0, \sigma_{\omega_n}^2)$ and $\sigma_{\omega_n}^2 \sim \mathcal{G}(\underline{S}, \underline{A})$. Proposition 2 formalizes this marginal prior as follows:

Proposition 2. Density of the marginal prior $p(\omega_n)$

The marginal prior density function for parameter ω_n obtained by marginalizing the joint prior distribution $p(\omega_n, \sigma_{\omega_n}^2)$ over $\sigma_{\omega_n}^2$, $p(\omega_n) = \int_0^\infty p(\omega_n | \sigma_{\omega_n}^2) p(\sigma_{\omega_n}^2) d\sigma_{\omega_n}^2$, where the priors $p(\omega_n | \sigma_{\omega_n}^2)$ and $p(\sigma_{\omega_n}^2)$ are given by Expressions (11) and (12), respectively. This yields

$$p(\omega_n) = \frac{|\omega_n|^{\underline{A}-\frac{1}{2}} K_{\underline{A}-\frac{1}{2}}\left(\sqrt{\frac{2}{\underline{S}}}|\omega_n|\right)}{\sqrt{\pi} (\sqrt{2})^{\underline{A}-\frac{3}{2}} \Gamma(\underline{A}) (\sqrt{\underline{S}})^{\underline{A}+\frac{1}{2}}}. \quad (18)$$

□

Proof. The integration proceeds by recognizing the constant and kernel and applies to the latter, which is facilitated using the normalizing constant of the generalized inverse Gaussian distribution provided by [Barndorff-Nielsen \(1997\)](#). □

To compute the Bayes factor using the SDDR approach, it is crucial that the marginal prior $p(\omega_n)$ is bounded at $\omega_n = 0$. Property (2) establishes that the existence of this bound depends on the hyperparameter \underline{A} :

Property 2. Upper bound of the marginal prior density for ω_n (see [Cadonna et al., 2020](#), Theorem 2).

$$\lim_{\omega_n \rightarrow 0} p(\omega_n) = \begin{cases} \infty & \text{for } 0 < \underline{A} \leq 0.5 \\ \frac{1}{\sqrt{2\pi\underline{S}}(\underline{A}^2 - \frac{1}{4})} \frac{\Gamma(\underline{A} + \frac{3}{2})}{\Gamma(\underline{A})} & \text{for } \underline{A} > 0.5 \end{cases} \quad (19)$$

Thus, the marginal prior density $p(\omega_n)$ is bounded from above if $\underline{A} > 0.5$, as required by Restriction (15). Accordingly, we set $\underline{A} = 1$, reducing the gamma prior to an exponential

distribution, consistent with the Bayesian Lasso prior considered by [Belmonte, Koop and Korobilis \(2014\)](#). Other choices are possible and are reviewed by [Cadonna et al. \(2020\)](#). Additionally, we set the hyper-parameter $\underline{S} = 0.05$, ensuring that nearly all prior probability mass for $\sigma_{\omega_n}^2$ lies within the interval $(0, 1)$.

6. A Monte Carlo study

An important question for practitioners is the performance of the verification procedure for identification of the structural shocks in finite samples under a misspecified variance process. Additionally, we look into the capacity of our procedure to normalize the structural parameters. We conduct a comparative Monte Carlo study to shed some light on these questions. More specifically, we compare the performance of our model using the hierarchical prior assumptions presented in Section 4.1 with one based on the prior distribution used by [Chan \(2018\)](#) featuring the zero-mean normal prior for ω_n as in (11) with the prior variance fixed to $\sigma_{\omega_n}^2 = 10$. The latter specification is complemented by a uniform prior for the autoregressive parameter $\rho_n \sim \mathcal{U}(-1, 1)$ and violates the scaling Restriction (14), potentially undermining the normalization.²

Our analysis estimates the two models for many artificially generated data sets. Estimation details on how to sample from the posterior densities for the states and parameters in the Monte Carlo setup presented here are discussed in [Appendix C](#). All our data-generating processes (DGPs) are bivariate and share the structural equation $\mathbf{B}_0 \mathbf{y}_t = \mathbf{w}_t$, uncorrelated structural shocks, $\mathbf{w}_t \sim \mathcal{N}_2(\mathbf{0}_2, \text{diag}(\sigma_{1,t}^2, \sigma_{2,t}^2))$, and the structural matrix set to $\mathbf{B}_0 = \begin{bmatrix} 100 & 80 \\ -20 & 200 \end{bmatrix}$ inspired by the parameter estimates from our empirical example in Section 7. They differ in the specification of the volatility process for which three alternatives are considered:

SV: where $\sigma_{n,t}^2 = \exp(\tilde{h}_{n,t})$ with $\tilde{h}_{n,t} = 0.92\tilde{h}_{n,t-1} + \tilde{v}_{n,t}$ and $\tilde{v}_{n,t} \sim \mathcal{N}(0, 0.25)$, which is equivalent to its non-centred version given by $\sigma_{n,t}^2 = \exp(\sqrt{0.25}h_{n,t})$ and the same AR(1) equation written for $h_{n,t}$ but with a standard normal shock.

²We acknowledge that Chan's prior was proposed for reduced-form models not requiring normalization.

GARCH: where the conditional variances follow a GARCH(1,1) equation, $\sigma_{n,t}^2 = 0.02 + 0.28w_{n,t-1}^2 + 0.7\sigma_{n,t-1}^2$ and $\sigma_{n,0}^2 = 1$.

Markov switching heteroskedasticity (MSH): where $\sigma_{n,t}^2 = \sigma_{n,s_t}^2$, s_t is a two-state Markov process with transition probabilities $\mathbf{P} = \begin{bmatrix} 0.98 & 0.02 \\ 0.02 & 0.98 \end{bmatrix}$, for $s_t = 1$, $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}_2, I_2)$ and for $s_t = 2$, $\mathbf{w}_t \sim \mathcal{N}(\mathbf{0}_2, \text{diag}(20, 10))$.

In other words, only the first volatility model corresponds to the SV model assumed for our Bayesian algorithms while the assumptions underlying our methods are violated if the volatility changes are generated by the two other models. This setup allows us to explore the robustness of our methods against misspecification.

For each volatility process, we generate data using four different scenarios: (1) both shocks are homoskedastic, (2) the first shock is homoskedastic while the second shock is heteroskedastic, (3) the first shock is heteroskedastic while the second shock is homoskedastic, (4) both shocks are heteroskedastic. For homoskedastic shocks we set $\sigma_{n,t}^2 = 1 \forall t$, while the heteroskedastic shocks are generated by the three different volatility models. We use two sample sizes, $T \in \{780, 260\}$, corresponding to 65 years of monthly and quarterly data, respectively. Finally, the study is based on one hundred simulated data sets for each of the four scenarios.

Table 1 reports the rejection rates for two approaches to decide on homoskedasticity of the first shock. These rejection rates are computed based on two strategies to construct critical values, which we refer to as l -value and q -value hereafter following [Benjamini and Hochberg \(1995\)](#) and [Storey \(2002\)](#). In the l -value approach reported in Panel A, we use the decision-theory consistent approach and reject homoskedasticity if $BF_{homosk} < 1$, that is, when more than 50 % of the posterior probability is assigned to heteroskedasticity. In the q -value approach, the critical value is set to the fifth percentile of the posterior odds ratio, BF_{homosk} from Equation (17), computed under the true null hypothesis $\omega_1 = 0$. Consequently, the rejection rates in the first row of Panel B of Table 1 are fixed at 5 %.

The rejection rates reported in Panel A of Table 1 show that our approach in most

Table 1: Simulation results: Rejection rates for homoskedasticity using our prior vs. [Chan \(2018\)](#)

		Our prior			Chan (2018) prior		
T	homoskedastic shocks in each DGP	DGPs			DGPs		
		SV	GARCH	MSH	SV	GARCH	MSH
Panel A: l -value approach							
780	shocks 1 & 2	0.00	0.00	0.00	0.00	0.00	0.00
	shock 1	0.00	0.00	0.02	0.01	0.00	0.01
	shock 2	0.98	0.80	0.22	0.95	0.71	0.18
	none	1.00	0.83	0.56	0.98	0.73	0.49
260	shocks 1 & 2	0.00	0.00	0.00	0.00	0.00	0.00
	shock 1	0.01	0.01	0.00	0.02	0.00	0.01
	shock 2	0.57	0.31	0.19	0.55	0.27	0.30
	none	0.78	0.37	0.41	0.74	0.28	0.61
Panel B: q -value approach							
780	shocks 1 & 2	0.05	0.05	0.05	0.05	0.05	0.05
	shock 1	0.08	0.13	0.09	0.07	0.11	0.10
	shock 2	0.99	0.94	0.53	0.99	0.94	0.54
	none	1.00	0.97	0.91	1.00	0.98	0.90
260	shocks 1 & 2	0.05	0.05	0.05	0.05	0.05	0.05
	shock 1	0.07	0.12	0.10	0.07	0.10	0.07
	shock 2	0.85	0.55	0.37	0.82	0.59	0.39
	none	0.95	0.68	0.78	0.96	0.69	0.78

Note: The table reports rejection rates for the hypothesis of homoskedasticity in the first shock, i.e., $\mathcal{H}_0 : \omega_1 = 0$ using simulated data. The rates are calculated based on 100 realizations of DGPs each with the following characteristics: sample sizes: $T \in \{260, 780\}$; volatility processes: SV, GARCH, MSH; homoskedastic shock arrangements: shocks 1 & 2, shock 1, shock 2, none. For a homoskedastic shock, the variance is set to $\sigma_{\eta}^2 = 1$.

cases outperforms the one based on Chan's prior. However, the latter features a better performance for the Markov-switching DGP for the smaller sample size. Overall, both methods detect homoskedasticity of the first shock very well and exhibit high performance in rejecting it when the first shock is heteroskedastic with SV or GARCH

volatility. Rejecting the null hypothesis becomes more cumbersome when the shock follows a Markov-switching volatility process. These findings are confirmed by the rejection rates reported in Panel B, where both priors result in very similar rejection rates.

Next, we compare the capacity of the two models with different priors for ω_n to normalize the system. We pointed out that the prior by Chan (2018) violates the scaling restriction (14). This implies a pole at zero in the prior for the conditional variance that is similar to the one featured by the density depicted in green in Figure 1. Due to this violation, the conditional variance is not normalized about the value 1 and implies a corresponding distortion in the values of the elements of the structural matrix.

In Table 2, we report averaged relative root-mean-squared errors (\overline{RMSE}) between the posterior mean estimates of the structural matrix for the two models, denoted by $\bar{\mathbf{B}}_0^{(our)}$ and $\bar{\mathbf{B}}_0^{(Chan)}$, relative to its corresponding true values, $\bar{\mathbf{B}}_0^{(true)}$. The values of RMSE computed for individual generated time series for a model with Chan's prior to that with our prior are given by:

$$RMSE = \sqrt{\frac{\sum_{n=1}^N \sum_{i=1}^N (\bar{\mathbf{B}}_{0,ni}^{(Chan)} - \mathbf{B}_{0,ni}^{(true)})^2}{\sum_{n=1}^N \sum_{i=1}^N (\bar{\mathbf{B}}_{0,ni}^{(our)} - \mathbf{B}_{0,ni}^{(true)})^2}}. \quad (20)$$

Therefore, a value greater than 1 means that the model with Chan's prior delivers less precise estimates than a model with our prior, while values within the interval $(0, 1)$ indicate that Chan's prior results in superior estimates. The averaged value of \overline{RMSE} reported in Table 2 is a sample mean of $RMSE$ s computed for the 100 simulated data sets. The results in Table 2 clearly reflect the distortions in the parameter estimates due to using Chan's prior for ω_n and, hence, not normalizing the estimated conditional variances. If at least one shock is heteroskedastic, all \overline{RMSE} s are clearly greater than one.

Our Monte Carlo study leads to two major conclusions. The first one is that our proposed method is very well suited for investigating the identification through heteroskedasticity. It performs well in many situations even if the volatility model does

Table 2: Average relative root-mean-squared errors for the structural matrix for models with our prior distribution for ω_n and that by [Chan \(2018\)](#), for SV data generating process

T	homoskedastic shocks in each DGP	\overline{RMSE}
780	shocks 1 & 2	0.90
	shock 1	103.58
	shock 2	19.32
	none	74.67
260	shocks 1 & 2	0.98
	shock 1	8.53
	shock 2	2.84
	none	9.24

Note: The table reports relative root-mean-squared errors between the posterior mean and the true values of the structural matrix parameters for models with alternative prior distributions for ω_n computed as in Expression (20). The reported values greater than 1 indicate a higher root-mean-squared error in the model with the prior by [Chan \(2018\)](#).

not match the assumed SV volatility model. In many cases our approach outperforms the procedure based on the prior proposed by [Chan \(2018\)](#). In terms of \overline{RMSE} , our method is considerably more precise than Chan’s procedure in estimating the structural parameters. Thus, we can endorse our method for applied work and present an empirical example in Section 7.

We acknowledge that there are other Bayesian and frequentist methods that have been proposed for checking identification through heteroskedasticity in structural VAR analysis such as those by [Lanne and Saikkonen \(2007\)](#), [Lütkepohl and Milunovich \(2016\)](#), [Lewis \(2021\)](#), and [Lütkepohl and Woźniak \(2020\)](#). We have also performed simulations for all of them and report the results in [Appendix G](#). As those methods feature null hypotheses that are different from our approach, they are not directly comparable to our approach. Some of the methods are designed for specific volatility models and partly perform very well in some situations. However, none of them outperforms our approach in all situations and they have clear deficiencies for some of the scenarios considered in our simulation design. Therefore, none of them is generally preferable to our approach.

7. Empirical application: 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.

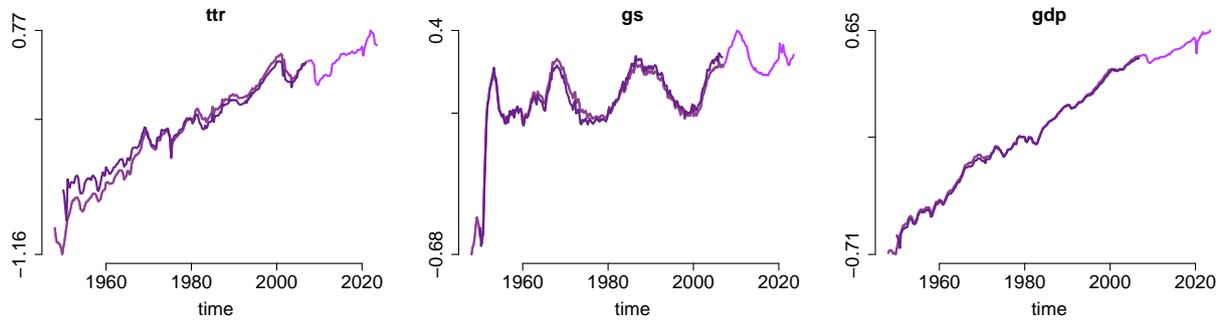
To illustrate the methods developed in the previous sections, we will consider a fiscal SVAR model in which the unanticipated tax shock has been identified in different ways. These alternative identification strategies include, for example, [Blanchard and Perotti \(2002\)](#) (henceforth BP), who use 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. We use the MR model as our benchmark to illustrate the use of our methodology for identifying the tax shock through 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 lengths 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

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 standardized by subtracting from each series its first observation in 1980.

MR-sample, uses the data from MR and LE that is downloaded directly from Karel Mertens' website.³ Following the data construction described by MR, total tax revenue, government spending, and gross domestic product, as well as the GDP deflator are taken from NIPA Tables numbers 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 \(2009\)](#) 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 1948Q1. One of these samples, hereafter the 2023-sample, ends in 2023Q3, and the other one, hereafter the 2006-sample, ends 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.

³The spreadsheet is available at https://karelmertenscom.files.wordpress.com/2017/09/jme2014_data.xls

Table 3: Verification of identification through heteroskedasticity of the structural shocks (based on the BP-ordering)

	2023-sample		2006-sample		MR-sample	
w_t^{tr}	-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 (17) together with numerical standard errors (NSEs) provided in brackets. 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.

7.2. Verifying identification through heteroskedasticity

We base our structural analysis on 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 homoskedastic. Even if the shocks are identified, the row ordering and row signs may change in different drawings from the posterior if one does not take 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 details on this procedure are provided in [Appendix E](#). Hence, the shocks can be labeled along the lines of BP as an unanticipated tax shock (w_t^{tr}), 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 although it is, of course, not clear from the outset if the shocks can be identified through heteroskedasticity with our methodology. If they can, they may still differ from those in BP and MR, in which case our labels may not be meaningful. We will return to this issue later.

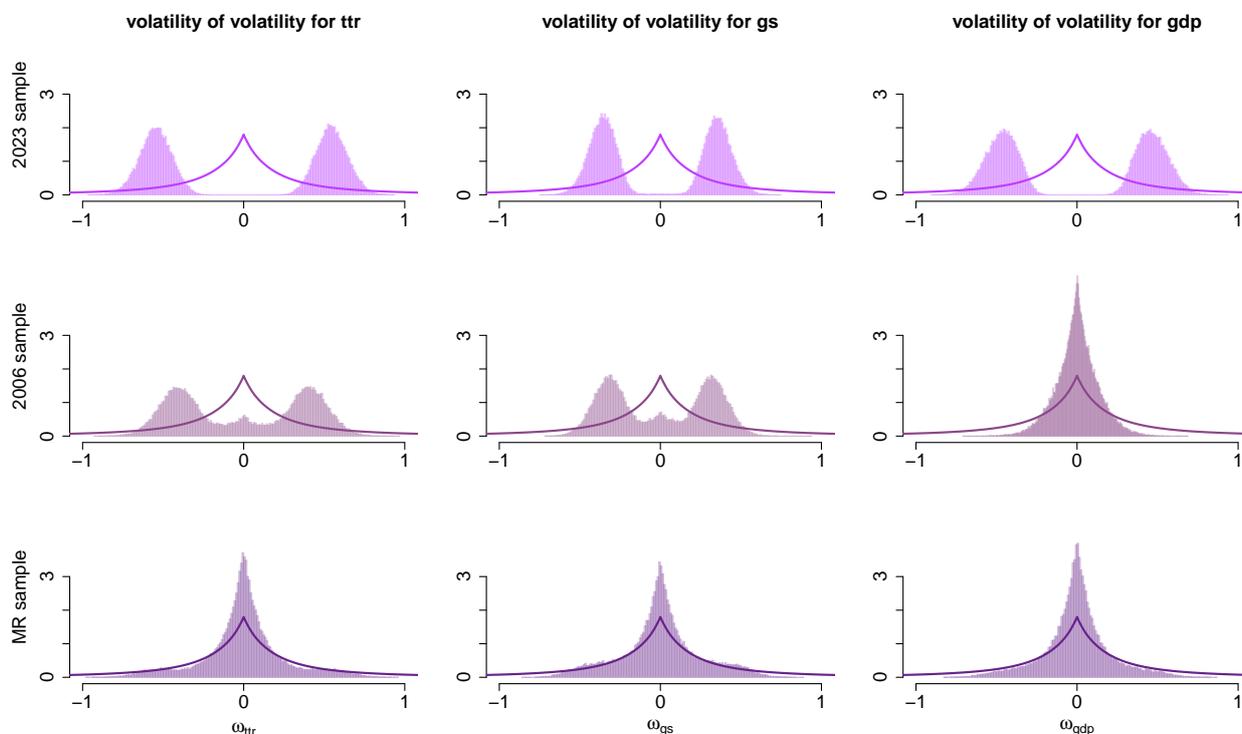
The next step in our analysis is to assess whether there are shocks that are identified through heteroskedasticity. Our main tool for that purpose is the SDDR from

Equation (17). The SDDR values computed for each of the three shocks individually using our three data samples are reported in Table 3. 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 Table 3 indicate that the posterior mass in favour of heteroskedasticity exceeds 99% for all the shocks. This result provides strong evidence for the identification of all three shocks through 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 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, which includes eight more observations than the MR-sample from the volatile late 1940s. More specifically, in the 2006-sample, the posterior probability of the heteroskedastic shock w_t^{tr} 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 ω_n associated with our three samples. Based on the information from these plots, the SDDRs from Equation (17) 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 and the bi-modality discussed in Section 3, providing evidence against

Figure 4: Marginal prior (solid line) and posterior (histograms) densities of ω_1 across samples (based on the BP-ordering)



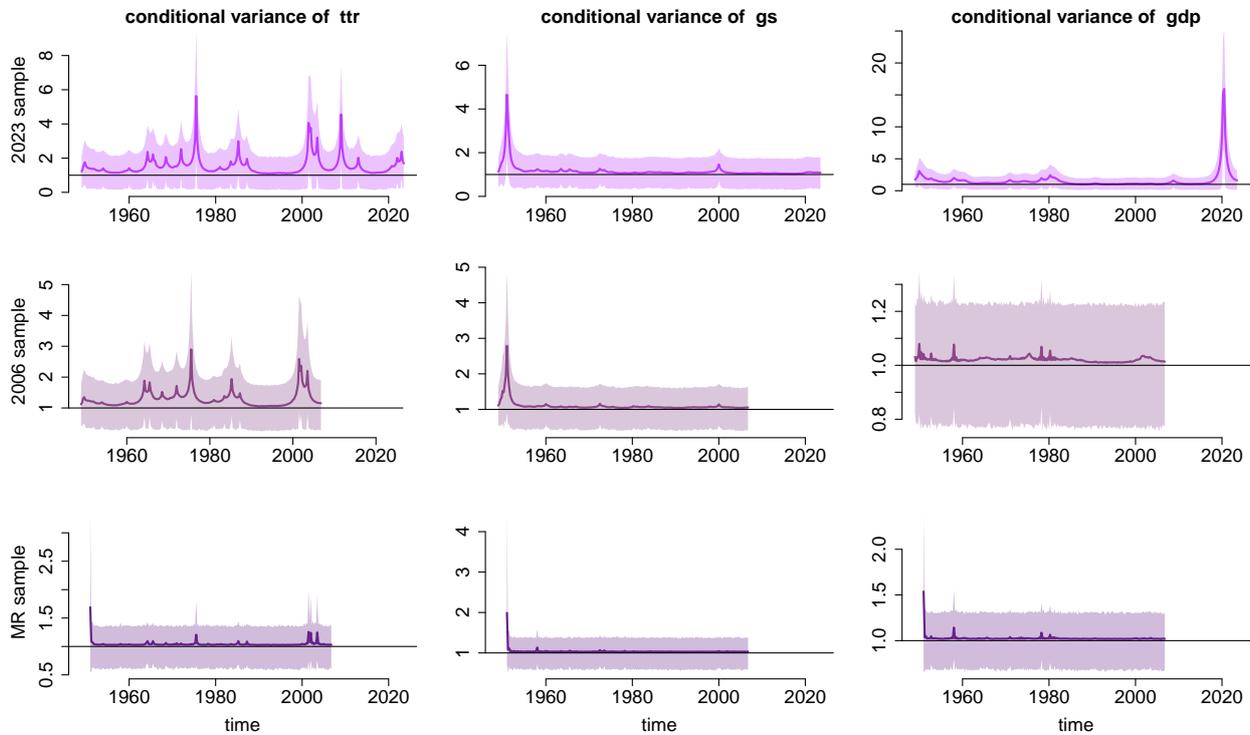
Note: The marginal prior density is estimated by numerical integration as in [Gelfand and Smith \(1990\)](#) using a grid of points from -1.1 to 1.1. They are the same for all samples and shocks. The marginal posterior densities are approximated using histograms. The ratio of these densities at point zero approximates the SDDR in Equation (17). Posterior mass less concentrated than the prior mass about zero provides evidence against homoskedasticity.

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.

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 5.⁴ The conditional variances are visibly time varying for the 2023-sample. The conditional variances of the first shock are significantly different from

⁴We compare these variance trajectories to those for a model with the centred SV process in [Appendix F](#).

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.

1 in six periods in that sample, including the mid-70s and mid-80s, individual quarters in 2001, 2002, and 2003, and 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 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 through heteroskedasticity in this sample. In particular, this evidence supports our claim that the first shock is identified as its conditional variances evolve non-proportionally to those of other shocks.

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 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 E](#)). The results of the SDDRs based on the MR-ordering are presented in [Table 4\(a\)](#). They paint a similar picture as the results in [Table 3](#). 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.

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 bi-modal 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 normalizing 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

Table 4: 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]

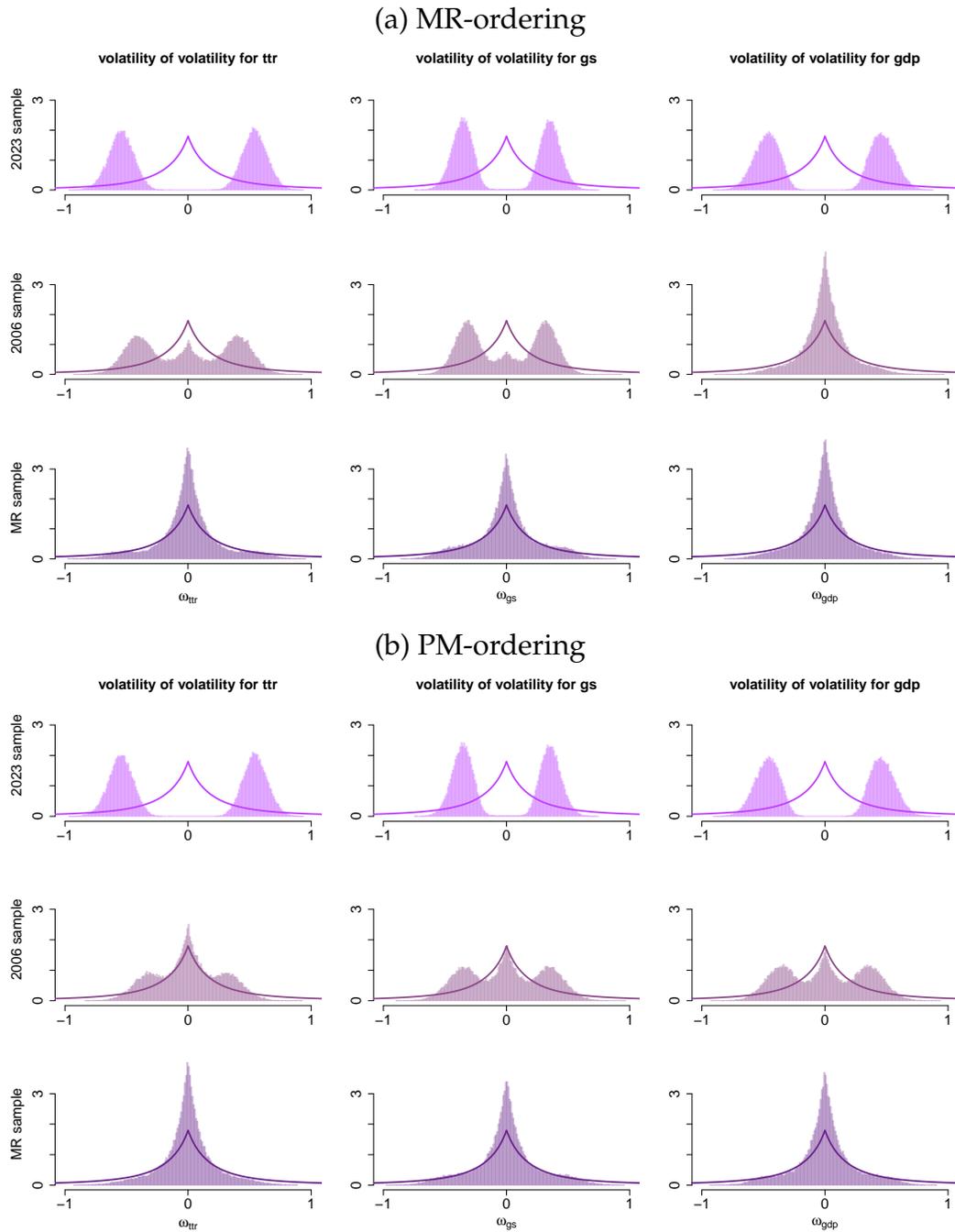
Note: The table reports the log of the Bayes factors estimated via the log of SDDRs from Equation (17) together with numerical standard errors provided in brackets. The note to Table 3 applies.

MR-ordering for the \mathbf{B}_0 drawings is not important.

We emphasize, however, that some kind of normalization 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 normalization 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 that 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 E](#) for details).

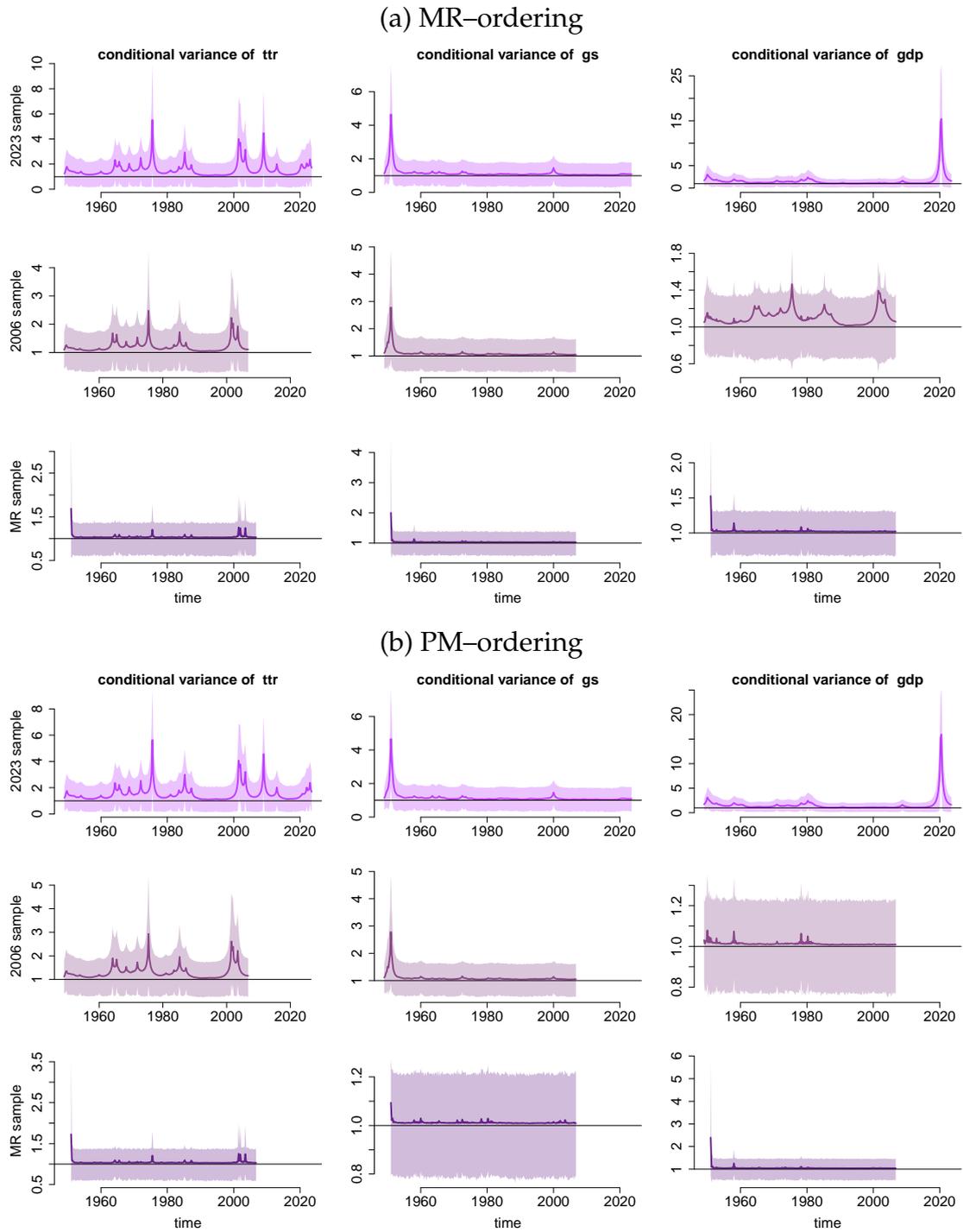
In this case, it is not clear a priori that the ordering of the shocks will be the same as for the BP- and MR-orderings. As the shocks are distinguished by their conditional variances, we consider the conditional variances and order them such that they look similar to those based on the BP- and MR-orderings. In this case, the three distinct variance patterns

Figure 6: Marginal prior (solid line) and posterior (histograms) densities of ω_1 across samples (based on alternative orderings). The note to Figure 4 applies.



allow for easily matching them with the shocks from the BP- and MR-orderings so we can

Figure 7: Conditional variance of structural shocks in the three samples (based on alternative orderings). The note to Figure 5 applies.



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 4(b). They strongly support that all three shocks are identified through heteroskedasticity in the 2023-sample. In fact, the SDDR values in Table 4(b) are identical to the corresponding values for the 2023-sample in Table 3. Additionally, the robustness of heteroskedasticity and identification verification to various ordering rules is confirmed by the plots of marginal posterior and prior distributions of ω_n 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 normalize 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 through 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 evidence 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 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 5. The

Table 5: Correlations between the narrative tax shock measure by [Romer and Romer \(2010\)](#) and other selected measures (based on the PM-ordering)

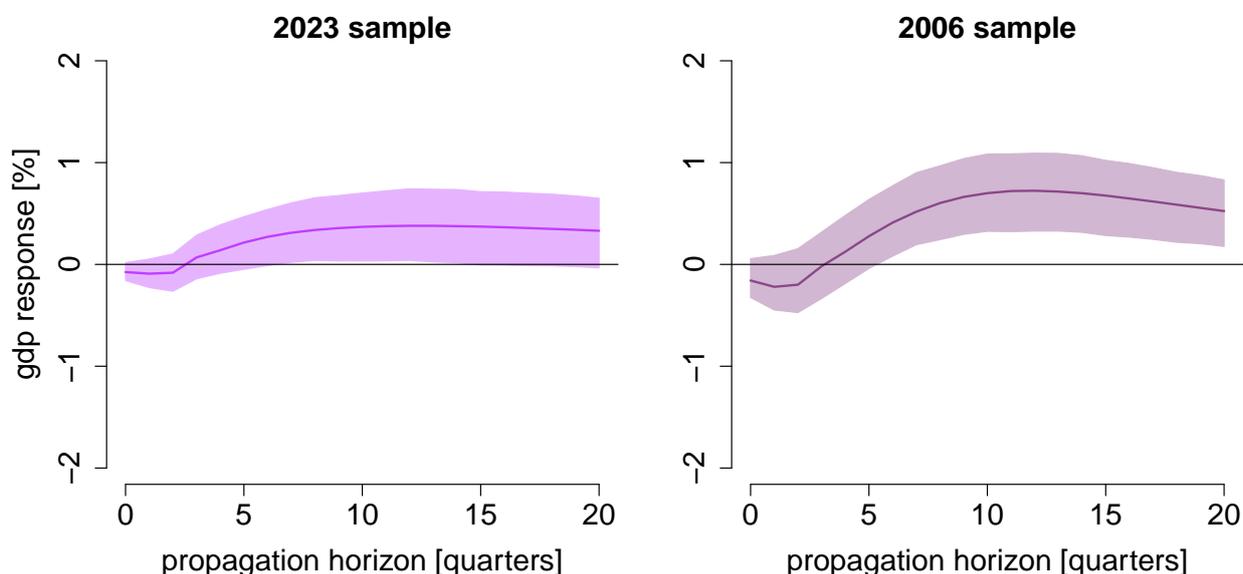
	2023-sample	2006-sample	BP results	MR results	LE results
w_t^{ttr}	0.224	0.264	0.277	0.298	0.233
w_t^{gs}	-0.022	0.030			
w_t^{gdp}	-0.154	-0.170			

Note: The table reports sample correlations between the narrative measure of tax shocks proposed by [Romer and Romer \(2010\)](#) and used by MR. The results in the 2023-sample and 2006-sample columns are based on our posterior estimations, where we used the posterior mean of the shocks as their estimator. The results in the BP results, MR results, and LE results columns are based on our reproduction of the results from MR and LE using the authors' computer codes and data.

results show that the first shock in our models is, albeit modestly, the most correlated one with the narrative measure of [Romer and Romer \(2010\)](#). Such a low correlation is, however, in line with the *weak instrument* observation by [Ramey \(2016\)](#) and exceeds 0.22 for all the models reported in Table 5. It is also higher than for the second shock, for which the values are -0.022 for the 2023-sample and 0.03 for the 2006-sample, and for the third shock for which the reported correlations are less than -0.15 for both samples. Notably, the correlations for the first shock have similar values as the tax shocks estimated by BP, MR, and LE, reported in the last three columns of Table 5. Therefore, our correlation analysis provides some support that the first shock could be interpreted as the *tax shock*.

Next, we investigate the dynamic effects of the tax shock identified with our approach on gdp_t . Figure 8 reports the corresponding impulse responses for both the 2006- and 2023- samples. Following MR and LE, they represent gdp_t responses to a tax shock that reduces ttr_t by 1% of gdp_t . Our impulse response results in Figure 8 share two common features: (i) no effect on impact and over the first six quarters and (ii) an increase in gdp_t reaching a peak thirteen quarters after the impact at a value around 0.38% for the 2023-sample and 0.73 for the 2006 one. The shorter sample shock is more persistent as its effect stays significant even five years after impact, whereas that for the longer sample

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



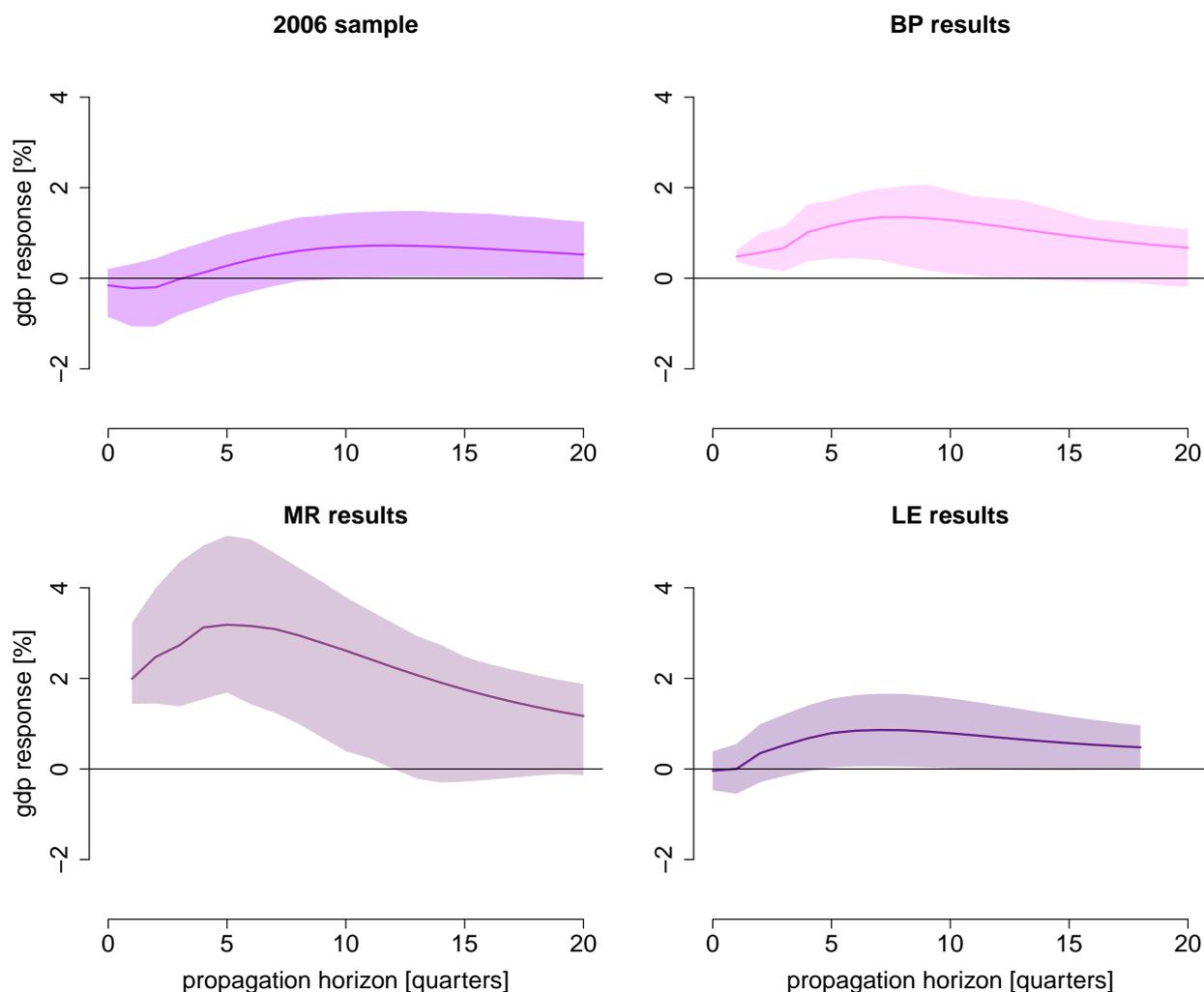
Note: The figure reports impulse responses of gdp_t to a negative tax shock lowering ttr_t by 1% of the gdp_t value in the last quarter of 2006. The lines report the posterior medians and the shaded areas the 68% HPD point-wise intervals.

dies out after 3.5 years.

Nevertheless, the shapes of the impulse responses from the 2023- and 2006-samples are quite similar to each other. We further compare them to the impulse responses reported in BP,⁵ MR, and LE for the MR-sample, that is, the original data used by these authors. Figure 9 reports our estimates for the 2006-sample with 90% HDP intervals with the results from the BP, MR, and LE models reporting the maximum likelihood estimates 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 is lost around three or four 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.

⁵Our BP results are based on the BP model estimated by MR.

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 1% of the 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 90% 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.

Nevertheless, the conclusions from our estimates do not deviate far from those established in the literature and are obtained by identification through heteroskedasticity and shock labelling using narrative measures only.

8. Conclusions

In this paper, we provided general conditions for identifying a structural shock through heteroskedasticity in multivariate dynamic structural models. These conditions apply 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 a flexible and easy-to-compute Bayes factor to verify the identification conditions. This was enabled by our analysis of marginal priors for conditional variances of the structural shocks. Such priors are flexible due to a hierarchical specification and ensure normalization using a specification centred at homoskedastic structural shocks. As a result, shock identification through time-varying volatility relies more heavily on the data and is less susceptible to prior.

These methods were applied in a Monte Carlo simulation confirming that our prior distributions lead to a strong performance of the identification verification procedure and precise estimation. In an applied example, we show that the unanticipated tax shock in the U.S.A. is identified through heteroskedasticity reliably.

However, 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 the **R** package **bsvars** by [Woźniak \(2024a,b\)](#). Additionally, an important extension in which the structural matrix and the conditional standard deviation of the SV equation change over time with a Markov process was recently proposed by [Camehl and Woźniak \(2024\)](#). This model facilitates verification through heteroskedasticity within Markov regimes.

References

- Barndorff-Nielsen, Ole E**, “Normal Inverse Gaussian Distributions and Stochastic Volatility Modelling,” *Scandinavian Journal of statistics*, 1997, 24 (1), 1–13.
- Barroso, Francisco Javier Callealta, Carmelo García-Pérez, and Mercedes Prieto-Alaiz**, “Modelling Income Distribution Using the Log Student’s t Distribution: New Evidence for European Union Countries,” *Economic Modelling*, July 2020, 89, 512–522.
- Bauwens, Luc, Michel Lubrano, and J.F. Richard**, *Bayesian Inference in Dynamic Econometric Models*, Oxford University Press, USA, 1999.
- Belmonte, Miguel A.G., Gary Koop, and Dimitris Korobilis**, “Hierarchical Shrinkage in Time-Varying Parameter Models: Hierarchical Shrinkage in Time-Varying Parameter Models,” *Journal of Forecasting*, 2014, 33 (1), 80–94.
- Benjamini, Yoav and Yosef Hochberg**, “Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing,” *Journal of the Royal Statistical Society: Series B (Methodological)*, 1995, 57 (1), 289–300.
- Bertsche, Dominik and Robin Braun**, “Identification of Structural Vector Autoregressions by Stochastic Volatility,” *Journal of Business & Economic Statistics*, 2022, 40 (1), 328–341.
- Bitto, Angela and Sylvia Frühwirth-Schnatter**, “Achieving Shrinkage in a Time-Varying Parameter Model Framework,” *Journal of Econometrics*, 2019, 210 (1), 75–97.
- Blanchard, O. and R. Perotti**, “An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output,” *The Quarterly Journal of Economics*, November 2002, 117 (4), 1329–1368.
- Cadonna, Annalisa, Sylvia Frühwirth-Schnatter, and Peter Knaus**, “Triple the Gamma Unifying Shrinkage Prior for Variance and Variable Selection in Sparse State Space and TVP Models,” *Econometrics*, 2020, 8 (2), 20.
- Camehl, Annika and Tomasz Woźniak**, “Time-Varying Identification of Monetary Policy Shocks,” *arXiv preprint arXiv:2311.05883*, 2024.
- Carriero, Andrea, Todd E. Clark, and Massimiliano Marcellino**, “Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate priors,” *Journal of Econometrics*, 2019, 212 (1), 137–154.
- Chan, Joshua and Ivan Jeliazkov**, “Efficient Simulation and Integrated Likelihood Estimation in State Space Models,” *International Journal of Mathematical Modelling and Numerical Optimisation*, 2009, 1, 101–120.
- Chan, Joshua C. C.**, “Specification Tests for Time-Varying Parameter Models with Stochastic Volatility,” *Econometric Reviews*, 2018, 37 (8), 807–823.
- , **Gary Koop, and Xuewen Yu**, “Large Order-Invariant Bayesian VARs with Stochastic Volatility,” *Journal of Business & Economic Statistics*, 2024, 42 (2), 825–837.

- Clark, Todd E. and Francesco Ravazzolo**, “Macroeconomic Forecasting Performance Under Alternative Specification of Time-Varying Volatility,” *Journal of Applied Econometrics*, 2015, 30, 551–575.
- Cogley, Timothy and Thomas J. Sargent**, “Drifts and Volatilities: Monetary Policies and Outcomes in the Post WWII US,” *Review of Economic Dynamics*, 2005, 8, 262–302.
- Doan, Thomas, Robert B. Litterman, and Christopher A. Sims**, “Forecasting and Conditional Projection Using Realistic Prior Distributions,” *Econometric Reviews*, 1984, 3 (June 2013), 37–41.
- Eddelbuettel, Dirk**, *Seamless R and C++ Integration with Rcpp*, New York, NY: Springer, 2013.
- **and Conrad Sanderson**, “RcppArmadillo: Accelerating R with high-performance C++ linear algebra,” *Computational Statistics & Data Analysis*, March 2014, 71, 1054–1063.
- **, Romain François, J Allaire, Kevin Ushey, Qiang Kou, N Russel, John Chambers, and D Bates**, “Rcpp: Seamless R and C++ integration,” *Journal of statistical software*, 2011, 40 (8), 1–18.
- Francis, Neville and Valerie A Ramey**, “Measures of Per Capita Hours and Their Implications for the Technology-Hours Debate,” *Journal of Money, credit and Banking*, 2009, 41 (6), 1071–1097.
- Gelfand, Alan E. and Adrian F. M. Smith**, “Sampling-Based Approaches to Calculating Marginal Densities,” *Journal of the American Statistical Association*, jun 1990, 85 (410), 398–409.
- Giannone, Domenico, Michele Lenza, and Giorgio E. Primiceri**, “Prior Selection for Vector Autoregressions,” *Review of Economics and Statistics*, 2015, 97 (2), 436–451.
- Herwartz, H. and H. Lütkepohl**, “Structural Vector Autoregressions with Markov Switching: Combining Conventional with Statistical Identification of Shocks,” *Journal of Econometrics*, 2014, 183, 104–116.
- Hogg, Robert V. and Stuart A. Klugman**, “On the Estimation of Long Tailed Skewed Distributions with Actuarial Applications,” *Journal of Econometrics*, 1983, 23 (1), 91–102.
- Hörmann, Wolfgang and Josef Leydold**, “Generating Generalized Inverse Gaussian Random Variates,” *Statistics and Computing*, 2014, 24 (4), 547–557.
- Hosszejni, Darjus and Gregor Kastner**, “Modeling Univariate and Multivariate Stochastic Volatility in R with `stochvol` and `factorstochvol`,” *Journal of Statistical Software*, 2021, 100 (12).
- Izzeldin, Marwan, Mike G. Tsionas, and Panayotis G. Michaelides**, “Multivariate Stochastic Volatility with Large and Moderate Shocks,” *Journal of the Royal Statistical Society Series A: Statistics in Society*, June 2019, 182 (3), 887–917.
- Jarociński, Marek**, “Estimating the Fed’s unconventional policy shocks,” *Journal of Monetary Economics*, 2024.
- Kass, Robert E. and Adrian E. Raftery**, “Bayes Factors,” *Journal of the American Statistical Association*, 1995, 90 (430), 773–795.
- Kastner, Gregor and Sylvia Frühwirth-Schnatter**, “Ancillarity-Sufficiency Interweaving Strategy (ASIS) for Boosting MCMC Estimation of Stochastic Volatility Models,” *Computational Statistics & Data Analysis*,

2014, 76, 408–423.

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

Lanne, M. and H. Lütkepohl, “Identifying Monetary Policy Shocks via Changes in Volatility,” *Journal of Money, Credit and Banking*, 2008, 40, 1131–1149.

—, —, and **K. Maciejowska**, “Structural Vector Autoregressions with Markov Switching,” *Journal of Economic Dynamics and Control*, 2010, 34, 121–131.

Lanne, Markku and Jani Luoto, “GMM Estimation of Non-Gaussian Structural Vector Autoregression,” *Journal of Business & Economic Statistics*, January 2021, 39 (1), 69–81.

— and **Pentti Saikkonen**, “A Multivariate Generalized Orthogonal Factor GARCH Model,” *Journal of Business & Economic Statistics*, 2007, 25 (1), 61–75.

Lewis, Daniel J, “Identifying Shocks via Time-Varying Volatility,” *The Review of Economic Studies*, 2021, 88 (6), 3086–3124.

Leydold, Josef and Wolfgang Hörmann, *GIGrvg: Random Variate Generator for the GIG Distribution* 2017. R package version 0.5.

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

Lütkepohl, Helmut and Aleksei Netsunajev, “Structural Vector Autoregressions with Smooth Transition in Variances,” *Journal of Economic Dynamics and Control*, 2017, 84, 43–57.

— and **Anton Velinov**, “Structural Vector Autoregressions: Checking Identifying Long-run Restrictions via Heteroskedasticity,” *Journal of Economic Surveys*, 2016, 30, 377–392.

— and **Tomasz Woźniak**, “Bayesian Inference for Structural Vector Autoregressions Identified by Markov-Switching Heteroskedasticity,” *Journal of Economic Dynamics and Control*, 2020, 113, 103862.

Lütkepohl, Helmut, Mika Meitz, Aleksei Netsunajev, and Pentti Saikkonen, “Testing identification via heteroskedasticity in structural vector autoregressive models,” *The Econometrics Journal*, 2021, 24 (1), 1–22.

Lütkepohl, Helmut and George Milunovich, “Testing for identification in SVAR-GARCH models,” *Journal of Economic Dynamics and Control*, 2016, 73, 241–258.

McCausland, William J, Shirley Miller, and Denis Pelletier, “Simulation Smoothing for State-Space models: A Computational Efficiency Analysis,” *Computational Statistics & Data Analysis*, 2011, 55 (1), 199–212.

Meade, B., L. Lafayette, G. Sauter, and D. Tosello, “Spartan HPC-Cloud Hybrid: Delivering Performance and Flexibility,” *University of Melbourne*, 2017.

Mertens, Karel and Morten O. Ravn, “A Reconciliation of SVAR and Narrative Estimates of Tax Multipliers,” *Journal of Monetary Economics*, 2014, 68, S1–S19.

- Mountford, Andrew and Harald Uhlig**, “What Are the Effects of Fiscal Policy Shocks?,” *Journal of Applied Econometrics*, 2009, 24 (6), 960–992.
- Netšunajev, A.**, “Reaction to Technology Shocks in Markov-switching Structural VARs: Identification via Heteroskedasticity,” *Journal of Macroeconomics*, 2013, 36, 51–62.
- Olmsted, Jonathan**, *RcppTN: Rcpp-Based Truncated Normal Distribution RNG and Family* 2017. R package version 0.2-2.
- Omori, Yasuhiro, Siddhartha Chib, Neil Shephard, and Jouchi Nakajima**, “Stochastic Volatility with Leverage: Fast and Efficient Likelihood Inference,” *Journal of Econometrics*, 2007, 140 (2), 425–449.
- Ramey, Valerie A.**, “Macroeconomic Shocks and Their Propagation,” *Handbook of macroeconomics*, 2016, 2, 71–162.
- Rigobon, R.**, “Identification Through Heteroskedasticity,” *Review of Economics and Statistics*, 2003, 85, 777–792.
- and **B. Sack**, “Measuring the Reaction of Monetary Policy to the Stock Market,” *Quarterly Journal of Economics*, 2003, 118, 639–669.
- Robert, Christian P.**, “Simulation of Truncated Normal Variables,” *Statistics and computing*, 1995, 5 (2), 121–125.
- Romer, Christina D and David H Romer**, “The Macroeconomic Effects of Tax Changes: Estimates Based on a New Measure of Fiscal Shocks,” *American Economic Review*, 2010, 100 (3), 763–801.
- Rubio-Ramírez, Juan F, Daniel F. Waggoner, and Tao Zha**, “Structural Vector Autoregressions: Theory of Identification and Algorithms for Inference,” *Review of Economic Studies*, 2010, 77 (2), 665–696.
- Sentana, E. and G. Fiorentini**, “Identification, Estimation and Testing of Conditionally Heteroskedastic Factor Models,” *Journal of Econometrics*, 2001, 102, 143–164.
- Storey, John D.**, “A direct approach to false discovery rates,” *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2002, 64 (3), 479–498.
- Tsionas, Mike G.**, “A Non-Iterative (Trivial) Method for Posterior Inference in Stochastic Volatility Models,” *Statistics & Probability Letters*, July 2017, 126, 83–87.
- U.S. Bureau of Economic Analysis**, “Table 1.1.5. Gross Domestic Product,” Data 2024. Retrieved from NIPA Tables [accessed January 2, 2024].
- , “Table 1.1.9. Implicit Price Deflators for Gross Domestic Product,” Data 2024. Retrieved from NIPA Tables [accessed January 2, 2024].
- , “Table 3.2. Federal Government Current Receipts and Expenditures,” Data 2024. Retrieved from NIPA Tables [accessed January 2, 2024].
- , “Table 3.9.5. Government Consumption Expenditures and Gross Investment,” Data 2024. Retrieved from NIPA Tables [accessed January 2, 2024].

- U.S. Bureau of Labor Statistics**, "Population Level [CNP16OV]," Data, Federal Reserve Bank of St. Louis 2024. Retrieved from FRED database [accessed January 2, 2024].
- Waggoner, Daniel F. and Tao Zha**, "A Gibbs Sampler for Structural Vector Autoregressions," *Journal of Economic Dynamics & Control*, 2003, 28 (2), 349–366.
- and —, "Likelihood Preserving Normalization in Multiple Equation Models," *Journal of Econometrics*, June 2003, 114 (2), 329–347.
- Woźniak, Tomasz**, *bsvars: Bayesian Estimation of Structural Vector Autoregressive Models* 2024. R package version 3.2.
- , "Fast and Efficient Bayesian Analysis of Structural Vector Autoregressions Using the R package bsvars," Working Paper, University of Melbourne 2024.
- Woźniak, Tomasz and Matthieu Droumaguet**, "Assessing Monetary Policy Models: Bayesian Inference for Heteroskedastic Structural VARs," *University of Melbourne Working Papers Series*, 2015, 2017.

Appendix A. Proofs

Appendix A.1. Proof of Theorem 1

We first prove the following lemma.

Lemma 1. Let Σ_t , $t = 0, 1, \dots$, be a sequence of positive definite $N \times N$ matrices and $\Lambda_t = \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2)$ a sequence of $N \times N$ diagonal matrices with $\Lambda_0 = \mathbf{I}_N$. Suppose there exists a nonsingular $N \times N$ matrix \mathbf{B} such that

$$\Sigma_t = \mathbf{B}\Lambda_t\mathbf{B}', \quad t = 0, 1, \dots \quad (\text{A.1})$$

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

Proof. Let \mathbf{B}_* be a matrix that satisfies: $\Sigma_t = \mathbf{B}_*\Lambda_t\mathbf{B}_*'$, $t = 0, 1, \dots$. It will be shown that, under the conditions of Lemma 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 $\sigma_1^2 \neq \sigma_i^2$, $i = 2, \dots, N$.

There exists a nonsingular $N \times N$ matrix \mathbf{Q} such that $\mathbf{B}_* = \mathbf{B}\mathbf{Q}$. Using $\Sigma_0 = \mathbf{B}\mathbf{B}'$, \mathbf{Q} has to satisfy the relation

$$\mathbf{B}\mathbf{B}' = \mathbf{B}\mathbf{Q}\mathbf{Q}'\mathbf{B}'.$$

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

The relations

$$\mathbf{B}\Lambda_t\mathbf{B}' = \mathbf{B}\mathbf{Q}\Lambda_t\mathbf{Q}'\mathbf{B}'$$

imply $\Lambda_t = \mathbf{Q}\Lambda_t\mathbf{Q}'$ and, hence, $\mathbf{Q}\Lambda_t = \Lambda_t\mathbf{Q}$ for all $t = 0, 1, \dots$

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

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

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

$$(1, 0, \dots, 0)' \quad \text{or} \quad (-1, 0, \dots, 0)'$$

which proves the lemma. □

Now consider the setup of Theorem 1 with $\mathbf{B} = \mathbf{B}_0^{-1}$. Then the arguments in the proof of Lemma 1 show that $\mathbf{B}_{0*}^{-1} = \mathbf{B}_0^{-1}\mathbf{Q}$, where \mathbf{Q} is as in the proof of Lemma 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.2. Proof of Corollary 1

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 $\mathbf{B}_0^{-1}\mathbf{Q}'$ 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. Priors

Appendix B.1. 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 (see [Tsonas, 2017](#); [Izzeldin, Tsonas and Michaelides, 2019](#), on the role of joint specification in posterior inference).

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:

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

□

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:

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

□

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)–(10) and (11)–(15), the joint priors for the $T \times 1$ vectors containing the latent process \mathbf{h}_n , log-conditional variances $\log \sigma_n^2 = \omega_n \mathbf{h}_n$, and conditional variances $\sigma_n^2 = \exp(\omega_n \mathbf{h}_n)$ are given by the following T -variate normal, normal product, and log normal product distributions:

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

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

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

where $\mathbf{h}_n = (h_{n,1} \ \dots \ h_{n,T})'$ is a $T \times 1$ vector and \mathbf{H}_{ρ_n} is a $T \times T$ matrix with ones on the main diagonal, with $-\rho_n$ on the first subdiagonal, and with zeros elsewhere.

Appendix B.2. 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 generalized-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 n th row of the matrices \mathbf{B}_0 and \mathbf{A} , respectively. Then the prior distribution for matrix \mathbf{B}_0 is proportional to

$$p(\mathbf{B}_0 \mid \gamma_0) \propto \det(|\mathbf{B}_0|)^{\nu-N} \exp \left\{ -\frac{1}{2} \sum_{n=1}^N \frac{1}{\gamma_{0,n}} \mathbf{B}_{0,n} \mathbf{B}'_{0,n} \right\}. \quad (B.3)$$

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 (B.3) 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 $\underline{\mathbf{A}} = \left[\mathbf{D} \quad \mathbf{0}_{N \times (N(p-1)+d)} \right]$, 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 $\underline{\mathbf{\Omega}} = \text{diag}\left(\mathbf{p}^{-1'} \otimes \mathbf{I}'_N \quad 100\mathbf{I}'_d\right)$, 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 $\gamma_{0.n}$:

$$\gamma_{0.n} \mid \underline{s}_{0.n} \sim \text{IG2}\left(\underline{s}_{0.n}, \underline{\nu}_0\right), \quad \underline{s}_{0.n} \mid \underline{s}_{\gamma_{0.n}} \sim \mathcal{G}\left(\underline{s}_{\gamma_{0.n}}, \underline{\nu}_{\gamma_{0.n}}\right), \quad \underline{s}_{\gamma_{0.n}} \sim \text{IG2}\left(\underline{s}_{s_0}, \underline{\nu}_{s_0}\right), \quad (\text{B.4})$$

$$\gamma_{A.n} \mid \underline{s}_{A.n} \sim \text{IG2}\left(\underline{s}_{A.n}, \underline{\nu}_A\right), \quad \underline{s}_{A.n} \mid \underline{s}_{\gamma_{A.n}} \sim \mathcal{G}\left(\underline{s}_{\gamma_{A.n}}, \underline{\nu}_{\gamma_{A.n}}\right), \quad \underline{s}_{\gamma_{A.n}} \sim \text{IG2}\left(\underline{s}_{s_A}, \underline{\nu}_{s_A}\right). \quad (\text{B.5})$$

We set $\underline{\nu}_0, \underline{\nu}_{\gamma_{0.n}}, \underline{s}_{s_0}$, 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 $\underline{\nu}_A, \underline{\nu}_{\gamma_{A.n}}, \underline{s}_{s_A}$, 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.

Appendix B.3. Prior for conditional variances in a centred SV model

Consider a centred SV model from equations (5)–(7) with the inverted-gamma 2 prior for the SV conditional variance parameter:

$$\omega_n^2 \mid \sigma_{\omega_n}^2 \sim \mathcal{IG}2(\sigma_{\omega_n}^2, \underline{\nu}). \quad (\text{B.6})$$

Then,

- (a) the log-volatilities, $\tilde{h}_{n,t}$, follow a Student-t marginal prior distribution (see [Bauwens, Lubrano and Richard, 1999](#)),
- (b) the conditional variances, $\sigma_{n,t}^2$, follow a log-Student-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 $\exp\left\{-\frac{1}{2}\left[\underline{\nu} + 1 - \sqrt{(\underline{\nu} + 1)^2 - 4\sigma_{\omega_n}^2}\right]\right\}$ iff $\sigma_{\omega_n}^2 < \frac{(\underline{\nu}+1)^2}{4}$,
- (e) the prior distribution stated in (b) has a median at point 1 only if $\underline{\nu}$ goes to infinity. \square

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

Appendix C. Gibbs sampler for the estimation of the parameters

This section scrutinizes 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 generalized-normal full conditional posterior distribution that is proportional to:

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

$$\bar{\mathbf{S}}_n^{-1} = \mathbf{I}_N / \gamma_{0,n} + \sum_{t=1}^T \mathbf{u}_t \mathbf{u}'_t / \sigma_{n,t}^2 \quad (\text{C.2})$$

$$\bar{v} = T + \underline{v} \quad (\text{C.3})$$

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 $\mathbf{x}_t = \left[\mathbf{y}'_{t-1} \quad \dots \quad \mathbf{y}'_{t-p} \quad d'_t \right]'$. Then the structural-form model from equation

(2) can be written as

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

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

$$\mathbf{A}'_n | \mathbf{y}, \mathbf{A}_0^{(n)}, \mathbf{B}_0, \sigma_1^2, \dots, \sigma_{N'}^2, \gamma_A \sim \mathcal{N}_{Np+d}(\bar{\mathbf{V}}_n \bar{\mathbf{A}}_n, \bar{\mathbf{V}}_n) \quad (\text{C.5})$$

$$\bar{\mathbf{V}}_n^{-1} = \underline{\mathbf{\Omega}}^{-1} / \gamma_{A.n} + \sum_{t=1}^T \mathbf{W}_t^{(n)'} \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2)^{-1} \mathbf{W}_t^{(n)} \quad (\text{C.6})$$

$$\bar{\mathbf{A}}_n = \underline{\mathbf{\Omega}}^{-1} \underline{\mathbf{A}}'_n / \gamma_{A.n} + \sum_{t=1}^T \mathbf{W}_t^{(n)'} \text{diag}(\sigma_{1,t}^2, \dots, \sigma_{N,t}^2)^{-1} \mathbf{z}_t^{(n)} \quad (\text{C.7})$$

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}_{γ_0} is sampled from their respective full conditional posterior distributions:

$$\gamma_{0.n} | \mathbf{B}_{0.n} \sim \text{IG2}(\underline{s}_{0.n} + \mathbf{B}_{0.n} \mathbf{B}'_{0.n}, \underline{\nu}_0 + N^2) \quad (\text{C.8})$$

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

$$\underline{s}_{\gamma_0} | \underline{s}_0 \sim \text{IG2}\left(\underline{s}_{s_0} + 2 \sum_{n=1}^N \underline{s}_{0.n} \underline{\nu}_{s_0} + 2N\underline{\nu}_{\gamma_0}\right), \quad (\text{C.10})$$

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

$$\gamma_{A.n} | \mathbf{A}_n, \underline{s}_{A.n} \sim \text{IG2}(\underline{s}_{A.n} + (\mathbf{A}_n - \underline{\mathbf{A}}_n) \underline{\mathbf{\Omega}}^{-1} (\mathbf{A}_n - \underline{\mathbf{A}}_n)', \underline{\nu}_A + Np + d) \quad (\text{C.11})$$

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

$$\underline{s}_{\gamma_A} | \underline{s}_A \sim \text{IG2}\left(\underline{s}_{s_A} + 2 \sum_{n=1}^N \underline{s}_{A.n} \underline{\nu}_{s_A} + 2N\underline{\nu}_{\gamma_A}\right). \quad (\text{C.13})$$

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 normality assumption for the structural shocks ($w_{n,t}$). It is facilitated by using the auxiliary mixture sampler proposed by [Omori et al. \(2007\)](#). To this end, note that each structural shock can be written as:

$$w_{n,t} = \sqrt{\sigma_{n,t}^2} \epsilon_{n,t}, \quad (\text{C.14})$$

$$\epsilon_{n,t} \sim \mathcal{N}(0, 1). \quad (\text{C.15})$$

By squaring and taking the logarithm of both sides of Equation (C.14) and remembering that we define $\sigma_{n,t}^2 = \exp(\omega_n h_{n,t})$, we have:

$$\tilde{w}_{n,t} = \omega_n h_{n,t} + \tilde{\epsilon}_{n,t}, \quad (\text{C.16})$$

where $\tilde{w}_{n,t} = \log w_{n,t}^2$ and $\tilde{\epsilon}_{n,t} = \log \epsilon_{n,t}^2$. Given the standard normal assumption for $\epsilon_{n,t}$ in (C.15), the distribution of $\tilde{\epsilon}_{n,t}$ 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.16) 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: $\mathbf{s}_n = (s_{n,1} \ \dots \ s_{n,T})'$ collecting the realisations of $s_{n,t}$ for all t , $\boldsymbol{\mu}_{\mathbf{s}_n} = (\mu_{s_{n,1}} \ \dots \ \mu_{s_{n,T}})'$, and $\boldsymbol{\sigma}_{\mathbf{s}_n}^2 = (\sigma_{s_{n,1}}^2 \ \dots \ \sigma_{s_{n,T}}^2)'$ collecting the n^{th} equation auxiliary mixture means and variances, and $\tilde{\mathbf{w}}_n = (\tilde{w}_{n,1} \ \dots \ \tilde{w}_{n,T})'$.

Sampling latent volatilities \mathbf{h}_n proceeds independently for each n from the following T -variate normal distribution parameterized following [Chan and Jeliaskov \(2009\)](#) in terms

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

$$\mathbf{h}_n \mid \mathbf{y}, \mathbf{s}_n, B_0, B_+, \omega_n, \rho_n \sim \mathcal{N}_T(\bar{\mathbf{V}}_{h_n} \bar{\mathbf{h}}_n, \bar{\mathbf{V}}_{h_n}) \quad (\text{C.17})$$

$$\bar{\mathbf{V}}_{h_n}^{-1} = \omega_n^2 \text{diag}(\sigma_{s_n}^{-2}) + \mathbf{H}'_{\rho_n} \mathbf{H}_{\rho_n} \quad (\text{C.18})$$

$$\bar{\mathbf{h}}_n = \omega_n \text{diag}(\sigma_{s_n}^{-2})(\tilde{\mathbf{w}}_n - \boldsymbol{\mu}_{s_n}) \quad (\text{C.19})$$

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:

$$\omega_n \mid \mathbf{y}, \mathbf{s}_n, h_n, \sigma_{\omega_n}^2 \sim \mathcal{N}(\bar{v}_{\omega_n} \bar{\omega}_n, \bar{v}_{\omega_n}) \quad (\text{C.20})$$

$$\bar{v}_{\omega_n}^{-1} = \mathbf{h}'_n \text{diag}(\sigma_{s_n}^{-2}) \mathbf{h}_n + \sigma_{\omega_n}^{-2} \quad (\text{C.21})$$

$$\bar{\omega}_n = \mathbf{h}'_n \text{diag}(\sigma_{s_n}^{-2})(\tilde{\mathbf{w}}_n - \boldsymbol{\mu}_{s_n}) \quad (\text{C.22})$$

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

$\sigma_{v_n}^2 = \omega_n^2$. Then, sample $\sigma_{v_n}^2$ from the following full conditional posterior distribution:

$$\sigma_{v_n}^2 \mid \mathbf{y}, \tilde{\mathbf{h}}_n, \sigma_{\omega_n}^2 \sim \mathcal{GIG}\left(-\frac{T-1}{2}, \tilde{\mathbf{h}}_n' \mathbf{H}'_{\rho_n} \mathbf{H}_{\rho_n} \tilde{\mathbf{h}}_n, \sigma_{\omega_n}^{-2}\right), \quad (\text{C.23})$$

where $\tilde{\mathbf{h}}_n = (\tilde{h}_{n,1} \dots \tilde{h}_{n,T})$. Finally, compute $\omega_n = \pm \sqrt{\sigma_{v_n}^2}$ 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:

$$\rho_n \mid \mathbf{y}, h_n, \sigma_{\omega_n}^2 \sim \mathcal{N}\left(\left(\sum_{t=0}^{T-1} h_{n,t}^2\right)^{-1} \left(\sum_{t=1}^T h_{n,t} h_{n,t-1}\right), \left(\sum_{t=0}^{T-1} h_{n,t}^2\right)^{-1}\right) \mathcal{I}\left(|\rho_n| < \sqrt{1 - \sigma_{\omega_n}^2}\right). \quad (\text{C.24})$$

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 $\omega_n, \sigma_{\omega_n}^2$, are *a posteriori* sampled independently from the following generalized inverse Gaussian distribution:

$$\sigma_{\omega_n}^2 \mid \mathbf{y}, \omega_n \sim \mathcal{GIG}\left(\underline{A} - \frac{1}{2}, \omega_n^2, \frac{2}{\underline{S}}\right) \quad (\text{C.25})$$

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. Computing the Savage-Dickey density ratio

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 $\bar{\omega}_n$ and variance \bar{v}_{ω_n} specified in equations (C.20)–(C.22). 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):

$$\widehat{p}(\omega_n = 0 \mid \mathbf{y}) = \frac{1}{S} \sum_{s=1}^S f_{\mathcal{N}}\left(0; \bar{\omega}_n^{(s)}, \bar{v}_{\omega_n}^{(s)}\right), \quad (\text{D.1})$$

where $f_{\mathcal{N}}$ denotes the density function of a normal distribution, whereas $\bar{\omega}_n^{(s)}$ and $\bar{v}_{\omega_n}^{(s)}$ 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.

Appendix E. Row sign and order normalization

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 normalization practice by [Lewis \(2021\)](#) and choose the changes of row signs and order of the structural matrix that minimize 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 minimize the likelihood-based distance proposed by [Jarociński \(2024\)](#): $\{\text{vec}[(\mathbf{PDB}_0^{(s)} - \widehat{\mathbf{B}}_0)']\}' \widehat{\mathbf{\Omega}}^{-1} \{\text{vec}[(\mathbf{PDB}_0^{(s)} - \widehat{\mathbf{B}}_0)']\}$, 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

$$\begin{bmatrix} \sigma_{ttr} & 0 & 0 \\ 0 & \sigma_{gs} & 0 \\ 0 & 0 & \sigma_{gdp} \end{bmatrix}^{-1} \begin{bmatrix} 1 & \theta_{gs} & 0 \\ \gamma_{ttr} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 0 & -\theta_{gdp} \\ 0 & 1 & -\gamma_{gdp} \\ -\zeta_{ttr} & -\zeta_{gs} & 1 \end{bmatrix} \begin{bmatrix} u_t^{ttr} \\ u_t^{gs} \\ u_t^{gdp} \end{bmatrix} = \begin{bmatrix} w_t^{ttr} \\ w_t^{gs} \\ w_t^{gdp} \end{bmatrix} \quad (\text{E.1})$$

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 minimized.

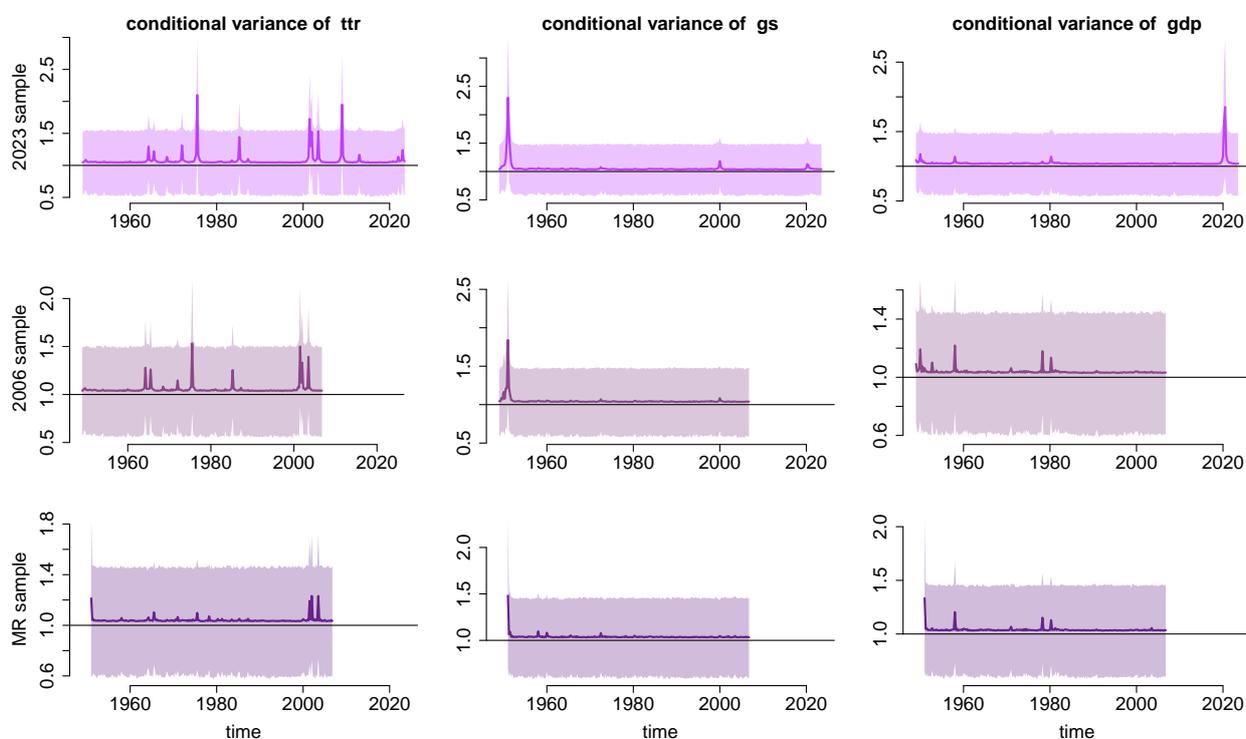
Appendix F. Volatility plots for alternative SV model specification

For the sake of comparison, we report the conditional variance plots for the SVAR model with an alternative specification for the SV process. We estimated models for all considered data with our priors for matrices \mathbf{B}_0 and \mathbf{A} , but with the centred SV model by [Chan et al. \(2024\)](#) that is specified similarly to that used by [Cogley and Sargent \(2005\)](#). The model features SV equations (7)–(5) together with the prior distributions set following [Chan et al. \(2024\)](#) as:

$$\omega_n^2 \sim IG2(1, 3), \quad \text{and} \quad \rho_n \sim \mathcal{U}(-1, 1). \quad (\text{F.1})$$

Figure F.10 reports the conditional variances. They have similar shapes as those reported in Figure 5 for our non-centred parameterisation of the SV model. In particular,

Figure F.10: Conditional variance of structural shocks in the three samples for a model with the centred SV specification. The note to Figure 5 applies.



the periods in which the conditional variances are significantly greater than 1 overlap to a large extent with those for our model for the 2023 sample. Nevertheless, the periods of high volatility are estimated to be more modest than for our model as the variance values in these periods are much lower here. We attribute this phenomenon to the thinner right tail of the prior entertained by the centred parameterisation. Note that a more flexible hierarchical prior structure, for instance, estimating the scale of the inverse gamma prior in (F.1) instead of setting it to one, could give this specification greater flexibility.

Appendix G. Rejection rates for other identification through heteroskedasticity tests

In this section, we report the rejection rates of other procedures that verify identification through heteroskedasticity and consider different hypotheses. Consequently, they are not directly comparable to the results for our procedure reported in Section 6.

Table G.6: Simulation Results: Rejection Rates for Proportional Variance Changes in a Markov-Switching Heteroskedasticity Model Using the Procedure by [Lütkepohl and Woźniak \(2020\)](#)

T	homoskedastic shocks in each DGP	DGPs		
		SV	GARCH	MSH
<i>l</i> -value approach				
780	shocks 1 & 2	1.00	1.00	1.00
	shock 1	0.99	1.00	0.97
	shock 2	0.92	0.94	0.99
	none	0.46	0.82	0.51
260	shocks 1 & 2	1.00	1.00	1.00
	shock 1	1.00	1.00	0.99
	shock 2	0.98	0.99	1.00
	none	0.80	0.94	0.51

Note: The table reports rejection rates for the hypothesis of proportional changes in conditional variances $\mathcal{H}_0 : \sigma_{1,s_t=2}^2 / \sigma_{2,s_t=2}^2 = 1$ investigated using $\ln SDDR$ by [Lütkepohl and Woźniak \(2020\)](#). The rates are calculated based on 100 realisations of DGPs each with the following characteristics: Sample sizes: $T \in \{260, 780\}$; Volatility processes: SV, GARCH, MSH; Homoskedastic shock arrangements: shocks 1 & 2, shock 1, shock 2, none. For a homoskedastic shock the variance is set to $\sigma_n^2 = 1$.

The first method is also a Bayesian method and was proposed by [Lütkepohl and Woźniak \(2020\)](#). It investigates identification in the context of a MSH model by checking whether the shock variances in different volatility regimes are sufficiently different for identification through heteroskedasticity. It does so by considering ratios of variances. Table [G.6](#) reports the rejection rates for the l -value approach for the hypothesis that the variances of the structural shocks are not proportional in a homogeneous two-regime Markov-switching heteroskedasticity model represented by the restriction involving the ratio of conditional variances from the second regime $\sigma_{1,s_t=2}^2 / \sigma_{2,s_t=2}^2 = 1$ and investigated using the Bayes factor (see [Lütkepohl and Woźniak \(2020\)](#) for details). This restriction is sufficient because the conditional variances in the first regime are equal to one.

The procedure performs very well when one shock is hetero- and the other homoskedastic. Its performance significantly weakens when both shocks are heteroskedastic, especially for a heterogeneous Markov process. The Bayes factor performs badly when all shocks are homoskedastic. This is attributed to the fact that [Lütkepohl and Woźniak \(2020\)](#) assume a stationary Markov process requiring non-zero occurrences of each regime. Given this assumption, in homoskedastic data, the second regime picks up a few outlying observations with the regime-specific variances much higher than in the first regime. This leads to the rejection of the hypothesis. Note that [Lütkepohl and Woźniak \(2020\)](#) recommend verifying heteroskedasticity first, for which they provide another Bayesian procedure. Finally, due to this behaviour of the procedure for homoskedastic data, we were not able to calibrate the critical value for the q -value approach.

Table [G.7](#) reports the rejection rates for the frequentist tests by [Lanne and Saikkonen \(2007\)](#) and [Lütkepohl and Milunovich \(2016\)](#) used by [Bertsche and Braun \(2022\)](#) and that proposed by [Lewis \(2021\)](#). Not all of these values are comparable with those reported in [Table 1](#) due to different hypotheses verified and the lack of critical values for Bayesian procedures. However, due to our design of the simulation consisting of checking the procedures performance using the same generated data sets the rejection rates in the Bayesian q -value approach and the size-adjusted power simulations are directly

comparable across the Panel B of tables 1 and G.7.

Lanne and Saikkonen (2007) proposed two alternative versions of a LM-type test based on Portmanteau test statistics to determine the heteroskedasticity rank defined as the number of independent univariate GARCH processes in the system. Lütkepohl and Milunovich (2016) proposed another LM test for the same set of null hypotheses for a differently specified structural model. We test the null hypothesis that the heteroskedasticity rank is equal to one which implies identification of the structural shocks for our DGPs. The test proposed by Lewis (2021) verifies the rank order of a specifically constructed matrix involving conditional variances. In order to investigate the hypothesis representing an identified system in our bivariate DGPs, we test the null hypothesis of the rank order being equal to one which again implies identification of the shocks.

The results indicate that the test by Lewis (2021) and the Q_1 test by Lanne and Saikkonen (2007) exhibit excellent size properties as their empirical sizes nearly perfectly match the nominal size of the tests based on critical values for a 5% level test. The Q_1 test by Lanne and Saikkonen (2007) is consistently undersized and the test by Lanne and Saikkonen (2007) is consistently oversized in all our samples (see the first row in Panel A of Table G.7). As the test is oversized, its empirical power is somewhat inflated and it is not surprising that its empirical power exceeds that of the Lewis (2021) test. However, even if a small sample correction of the tests is performed as in the size-adjusted power simulation approach in Panel B of Table G.7, the power of the Q_2 test by Lanne and Saikkonen (2007) is superior in many cases. Exceptions are some DGPs based on MSH processes.

The overall conclusion of our simulations is that none of the procedures works perfectly and is superior for all the scenarios considered in our simulations.

Table G.7: Simulation Results: Rejection Rates for Tests by [Lanne and Saikkonen \(2007\)](#) (LS2007), [Lütkepohl and Milunovich \(2016\)](#) (LM2016), and [Lewis \(2021\)](#). The note to [Table 1](#) applies.

T	homoskedastic shocks in each DGP	Q_1 by LS2007			Q_2 by LS2007			by LM2016			by Lewis (2021)		
		SV	GARCH	MSH	SV	GARCH	MSH	SV	GARCH	MSH	SV	GARCH	MSH
Panel A: based on asymptotic critical values for a 5% level test													
780	shocks 1 & 2	0.05	0.05	0.05	0.02	0.02	0.02	0.10	0.10	0.10	0.05	0.05	0.05
	shock 1	0.87	0.43	0.88	0.92	1.00	0.84	0.93	1.00	0.81	0.74	0.73	0.65
	shock 2	0.99	0.96	0.97	0.91	1.00	0.93	0.91	1.00	0.92	0.94	0.95	0.81
	none	0.98	1.00	0.99	0.99	1.00	0.99	0.99	1.00	0.99	0.77	0.87	0.02
260	shocks 1 & 2	0.07	0.07	0.07	0.04	0.04	0.04	0.12	0.12	0.12	0.04	0.04	0.04
	shock 1	0.50	0.24	0.50	0.69	0.80	0.49	0.66	0.81	0.48	0.30	0.34	0.33
	shock 2	0.69	0.64	0.77	0.60	0.83	0.58	0.63	0.86	0.49	0.63	0.55	0.50
	none	0.77	0.91	0.84	0.87	0.99	0.81	0.83	0.99	0.78	0.43	0.72	0.05
Panel B: based on the the 5th percentile of the simulated test statistics													
780	shocks 1 & 2	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	shock 1	0.87	0.43	0.88	0.93	1.00	0.88	0.92	1.00	0.74	0.77	0.74	0.69
	shock 2	0.99	0.96	0.97	0.94	1.00	0.93	0.89	1.00	0.86	0.95	0.95	0.82
	none	0.98	1.00	0.99	0.99	1.00	0.99	0.99	1.00	0.99	0.78	0.87	0.02
260	shocks 1 & 2	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05	0.05
	shock 1	0.48	0.24	0.48	0.70	0.81	0.51	0.50	0.68	0.31	0.33	0.37	0.33
	shock 2	0.69	0.64	0.76	0.60	0.83	0.59	0.51	0.77	0.30	0.64	0.61	0.53
	none	0.77	0.91	0.84	0.87	0.99	0.81	0.66	0.98	0.71	0.44	0.74	0.06