What do Data Say About Time-Variation in Monetary Policy Shock Identification?

Annika Camehl[a], Tomasz Woźniak[b,*]

[a] *Erasmus University Rotterdam*
[b] *University of Melbourne*

February 28, 2023

Abstract

We show that data support time-variation in identification patterns of US monetary policy shocks between January 1960 and May 2022. We look at the monetary policy reaction function in a form of Taylor's rule that is potentially extended by additional indicators. In the regime predominate before 2008, shadow interest rates react contemporaneously also to the term spread while in the regime mainly present after 2008 this role is taken by the money aggregate. To show that, we develop a Bayesian heteroskedastic structural vector autoregressive model with Markov-Switching and data-driven regime-specific identification search. This model enables regime-specific identification of structural shocks, time-varying impulse responses, and a swift way to verify identification through heteroskedasticity within a regime.

Keywords: Stochastic Search Specification Selection, Identification Through Heteroskedasticity, Stochastic Volatility, Markov-Switching
JEL classification: C11, C32, E52

1. Introduction

To disentangle unexpected changes in monetary policy from systematic behavior, researchers wildly apply contemporaneous exclusion restrictions in vector autoregressive models, especially due to their ease of use (e.g., Primiceri, 2005; Sims and Zha, 2006;

Walentin, 2014; Wu and Xia, 2016; Jawadi, Sousa and Traverso, 2017). The restrictions to identify monetary policy shocks commonly apply to the contemporaneous relation of interest rates in the monetary policy reaction function to prices and output, similarly stated in empirical Taylor rules. However, the targeting strategy of the Federal Reserve to reach its dual mandate of price stability and economic output has evolved over the last sixty years driven by changes in the monetary policy instrument, regulatory environment, or importance of the financial sector (see for an overview Boivin, Kiley and Mishkin, 2010). For instance, the monetary policy reaction function might include term spreads, capturing the slope of the yield curve, when short term interest rates are at the zero lower bound to acknowledge the role of stabilizing long term interest rates (Baumeister and Benati, 2013; Liu, Mumtaz, Theodoridis and Zanetti, 2017; Feldkircher and Huber, 2018; Tillmann, 2020). Likewise, before the 1980's change towards interest rate policy but also through large asset purchase programs increasing money supply, monetary policy actions can impact simultaneously interest rates and monetary aggregates (Sims and Zha, 2006; Belongia and Ireland, 2015). Hence, appropriate exclusion restrictions to identify monetary policy shocks might change over time. However, it is standard in the literature to rely on time-invariant identification patterns.

In this paper, we show that data strongly support time-varying identification patterns of US monetary policy shocks. We find strong evidence for including term spreads in the monetary policy reaction function in a regime which dominates before 2008. Here, systematic monetary policy reacts – by adjusting short term interest rates – contemporaneously to changes in output, prices, and term spreads. In contrast, in the regime occurring mainly after 2008, the systematic response of monetary policy adjusts to movements in output, prices, and a monetary aggregate.

To reach these conclusions, we develop a Bayesian heteroskedastic structural vector autoregressive (SVAR) model with Markov-Switching and data-driven regime-specific identification search. Our model has two new features. First, we include Markov-switching time variation in the matrix of contemporaneous coefficients and in a stochastic volatility parameter. The former allows us to estimate the time-varying structural matrix and the latter to check whether this matrix is identified via heteroskedasticity within each regime. Second, we add a data-driven search of regime-specific exclusion restrictions which facilitates data-supported selection of alternative ways to identify monetary policy shocks within each regime. We estimate our model with monthly US data from January 1960 to May 2022, including industrial production, consumer price index, shadow rate, term spread, money, credit spread, and stock prices.

In line with empirical papers arguing for the need of time-variation in coefficients and/or error covariance matrices (among others, Sims and Zha, 2006; Koop, León-González and Strachan, 2009; Primiceri, 2005; Ang, Boivin, Dong and Loo-Kung, 2011; Hubrich and Tetlow, 2015; Liu et al., 2017; Jackson, Owyang and Soques, 2018), our model, first, allows for regime dependent coefficients, implemented via Markov-Switching. Relying on changes across regimes circumvents the issue of overparameterization time-varying parameter models easily suffer from (Sims and Zha, 2006; Sims, Waggoner and Zha, 2008). Moreover, we implement time variation in the shocks' conditional variances modelled via a stochastic volatility (SV) process which has regime-dependent volatility of the volatility parameter. The regime-dependence of the stochastic volatility parameter offers a swift way to verify the identification of the monetary policy reaction function through heteroskedasticity within each regime and, thus, extending the framework proposed by Lütkepohl, Shang, Uzeda and Woźniak (2022a). We rely on a stochastic volatility process since, among others, Cogley and Sargent (2005), Primiceri (2005), Clark (2011), Clark and Ravazzolo (2015), and Chan and Eisenstat (2018) advocate using the SV for modelling volatility in macroeconomic time series benefiting forecasting accuracy. We use the heteroskedastic shocks to identify the structural matrix (as in, e.g., Lütkepohl and Woźniak, 2017; Bertsche and Braun, 2022) as then the identifying exclusion restrictions become overidentifying. In consequence, these restrictions can be verified using our novel data-driven search for regime-specific identifying exclusion restriction patterns.

The data-driven search for regime-specific exclusion restrictions is the second new feature of our model. To facilitate the restriction search, we propose a new stochastic search structural specification selection (S^5) mechanism, as a generalization of the selection priors introduced by George and Mcculloch (1997), George, Sun and Ni (2008) and Koop and Korobilis (2016). Specifically, we set a multinomial prior on a state-dependent indicator variable. This indicator determines the identifying pattern of exclusion restrictions imposed on the instantaneous effect matrix while the remaining parameters are sampled. Thereby, the posterior distributions of a parameter in the structural matrix of contemporaneous effects will show whether an exclusion restriction is supported by the data in a respective regime. With our search mechanism, we challenge the common assumption in the literature of time-invariant exclusion restrictions for the identification of monetary policy shocks. Moreover, we contribute to the discussion on appropriate exclusion restrictions to identify monetary policy shocks by including a wide range of various patterns in our search mechanism. Thereby, we provide evidence on which of the exclusion restrictions are supported by the data within each regime.

To the best of our knowledge, we are the first to introduce a data-driven search to verify alternative ways of identifying the monetary policy shock within each regime. Only few previous papers adopt regime dependent identification patterns, albeit setting restrictions a-priori without further validation. Bacchiocchi and Fanelli (2015) impose two identification patterns with an additive relationship between the two in an SVAR model identified through heteroskedasticity with exogenous switches. Bacchiocchi, Castelnuovo and Fanelli (2017) apply this approach to study the effects of monetary policy and report differences in impulse responses during the Great Moderation compared to previous periods. Kimura and Nakajima (2016) analyze changes in the monetary policy of Japan identifying a monetary policy reaction function in two exogenously given regimes once based on an interest rate rule or on a bank reserves rule.

In our model, we estimate two regimes. The first one predominates in the first part of the sample until 2008. The second regime occurs mainly after 2008 dominating also during the Covid-19 crisis. Within both regimes, we can verify that the monetary policy reaction function is identified via heteroskedasticity.

We find strong support with our S^5 mechanism for including the term spread in the monetary policy reaction function of the first regime. The posterior probability of the S^5

component indicator for the restriction pattern where no exclusion restriction is set on the term spread equals one. In this regime, output, prices, shadow interest rates, and the term spread move simultaneously. Hence, our results support an empirical generalized Taylor rule specification for the monetary policy reaction function in the first regime including term spreads as well. The term spread reveals future expectations of investors and can be used by the central bank as a timely indicator of the state of the economy (Nimark, 2008; Rudebusch and Wu, 2008; Tillmann, 2020).

In the second regime, data select money to appear in the monetary policy reaction function. Again, the evidence is very strong since the posterior probability of the component indicator for those specific exclusion restriction patterns is numerically equal to one. Importantly, we find evidence for a switch in the policy indicator in the second regime. The estimate of a parameter on the shadow interest rates in the monetary policy reaction function in the second regime should be considered not different from zero and this value is included within the 90% posterior density region. This fact combined with a large and positive estimate on the contemporaneous effect of money in the monetary policy reaction function provides evidence that the monetary policy indicator in this regime switches to money.

Allowing for regime-specific identification patterns through the search mechanism has a strong impact on the persistence of the regimes. We find a much less persistent second regime encompassing only a few sample periods in models which allow for Markov-switching in the structural matrix but exclude the stochastic search mechanism. Therefore, the time variation in the monetary policy reaction function is evident only in models with the S^5 and contradicted without it. These findings of our empirical investigation are robust to changing the assumptions regarding the prior distributions and several other features of the model specification.

In what follows, we introduce our empirical approach in Section 2, scrutinise the details of the structural model in Section 3, provide details on the S^5 mechanism in Section 4 as well as on the priors and estimation of the remaining parameters in Section 5, describe the regime-specific verification of identification through heteroskedasticity in Section 6, and report the evidence on our main findings in Section 7.

2. Time-varying identification of US monetary policy shocks

To study time variation in identification of US monetary policy shocks, we estimate our heteroskedastic SVAR model with monthly data from January 1960 to May 2022 for seven endogenous variables collected in vector, $\mathbf{y}_t = (ip_t, cpi_t, R_t, TS_t, m_t, CS_t, sp_t)'$. These variables allow us to specify various identification patterns for monetary policy shocks and to study the impact of unexpected monetary tightening on the economic environment, monetary measures, and financial variables.

Firstly, we include log industrial production (*ip*) and log consumer price index (*cpi*) as measures of real economic activity and price level respectively. We use the shadow rate (*R*) by Wu and Xia (2016) as the monetary policy instrument that captures both conventional and unconventional monetary policy actions. Since the shadow rate series starts in January 1990, we use the federal funds rate for the preceding period.

We include a term spread (*TS*) measured as the 10-year treasury constant maturity rate minus the federal funds rate. This indicator captures the slope of the yield curve. The term spread reacts to the large asset purchase programs after the financial crisis (Baumeister and Benati, 2013; Feldkircher and Huber, 2018; Liu et al., 2017). We use it also as a potential policy indicator to measure unconventional monetary policy and to check if it reacts to monetary policy shifts.

The fifth variable in the model is a monetary aggregate measure. We use log M2 money supply (*m*). Similarly to the term spread, it can be seen as a monetary policy indicator (Christiano, Eichenbaum and Evans, 1999; Leeper and Zha, 2003; Belongia and Ireland, 2015) or as a variable affected by monetary policy shocks (Sims and Zha, 2006; Hubrich and Tetlow, 2015; Liu et al., 2017).

Finally, in order to capture financial conditions, we include a credit spread (*CS*) measured by Moody's seasoned Baa corporate bond yield relative to the yield on the 10-year treasury rate (Gilchrist and Zakrajsek, 2012; Caldara and Herbst, 2019). As the last variable, we take the log of the S&P500 stock price index (*sp*) as a fast-moving price indicator

(Arias, Caldara and Rubio-Ramírez, 2019; Leeper and Zha, 2003). Appendix B gives details on the data sources and transformations.

Based on this set of variables, we can specify different exclusion restrictions on contemporaneous relations in the monetary policy reaction function to identify monetary policy shocks. We allow those restrictions to potentially vary across regimes. Based on our search, we can, first, analyze whether the identification pattern of monetary policy shocks and, hence, which variables affect the systematic response of monetary policy changes over regimes. Second, since the theoretical and empirical literature motivates different specifications of the monetary policy reaction function, we can shed light on which patterns are supported by the data.

We define the following regime-specific identification patterns for monetary policy shocks by setting restrictions on the matrix of contemporaneous effects: where $*$ indicates

Table 1: Regime-specific identification patterns of monetary policy shocks

$$\begin{bmatrix} * & 0 & 0 & 0 & 0 & 0 & 0 \\ * & * & 0 & 0 & 0 & 0 & 0 \\ * & * & * & S^5 & S^5 & S^5 & S^5 \\ * & * & * & * & 0 & 0 & 0 \\ * & * & * & * & * & 0 & 0 \\ * & * & * & * & * & * & 0 \\ * & * & * & * & * & * & * \end{bmatrix} \begin{bmatrix} ip_t \\ cpi_t \\ R_t \\ TS_t \\ m_t \\ CS_t \\ sp_t \end{bmatrix}$$

an unrestricted parameter, 0 a zero restriction, and S^5 that we search for a restriction on this element in each regime. In the search, we allow one parameter denoted by S^5 at a time to enter the interest rate equation. We relax this (allowing multiple parameters to enter at the same time) and extend it to several equations in the sensitivity analysis. Conducting the search within each regime, we can verify whether identification of monetary policy shocks is time-varying. We allow the unrestricted parameters to vary across regimes as well.

The first two rows in the matrix of contemporaneous relations given in Table 1 relate to the production sector. Following standard assumptions, we characterize output and prices as sluggish. These variables react with a lag to changes in all other variables.

Thus, they appear before the monetary policy indicators in the order of variables as is commonly the case in recursive systems (see e.g. Sims, 1986; Bernanke and Blinder, 1992; Sims, 1992; Christiano et al., 1999; Christiano, Eichenbaum and Evans, 2005; Primiceri, 2005; Sims and Zha, 2006; Belongia and Ireland, 2015).

The third row specifies systematic monetary policy behavior with the shadow rate as the policy indicator. We allow for a free search of identifying restrictions on the monetary policy reaction function, fixing the first three elements. The simultaneous systematic response of interest rates to movements in prices and output depicts a version of the standard empirical Taylor rule (Taylor, 1993). When interest rates react contemporaneously to some of the remaining variables, the underlying monetary policy reaction function follows a generalization of the standard Taylor rule accounting for additional variables.

Our restriction pattern in the third row embeds several generalized Taylor rule specifications. If systematic changes in interest rates depend on TS_t, the fourth element in the third row would be non-zero. Term spreads can reveal information on future expectations of participants in the bond market (Nimark, 2008; Tillmann, 2020). Moreover, they can be used as an indicator of expectations on changes to the economic environment as Rudebusch and Wu (2008), Ang et al. (2011), and Vázquez, María-Dolores and Londoño (2013), Hördahl, Remolona and Valente (2020), among others, report a strong link between the yield curve and macroeconomic variables. A shift in the term spreads can give an indication that the economy changes and the central bank should react to that by adjusting interest rates. We also interpret such an instantaneous reaction of interest rates to ts_t as in spirit of the identification strategies for monetary policy shocks used in Baumeister and Benati (2013), Feldkircher and Huber (2018), and Liu et al. (2017). While they assume a negative contemporaneous reaction of the term spread to an exogenous shift in the behavior of the monetary authority, we allow term spreads to potentially enter the monetary policy rule.

No exclusion restriction on the contemporaneous reaction to money indicates that the central bank actions simultaneously affect shadow interest rates and money. Such simultaneity can arise since unconventional monetary policy actions, such as quantitative easing, has lead to an systematic expansion of the money supply (Belongia and Ireland, 2015). Moreover, including money in the monetary policy reaction function indicates that the central bank changes interest rates in response to changes in money since those give an indication on the liquidity of the market. Allowing money to enter the monetary policy reaction function contemporaneously is in line with the identification used by Sims and Zha (2006) and Leeper and Zha (2003). They identify a monetary policy shock by augmenting the standard Taylor rule with an indicator for money supply. Likewise, Belongia and Ireland (2015) demonstrate that data reject excluding money supply from the Taylor rule for their analyzed sample.

We consider that the Fed might set interest rates based on current values of the credit spread, fifth element in the third row. The credit spread can reveal information on financial conditions (Gilchrist and Zakrajsek, 2012; Caldara and Herbst, 2019). The central bank could react to the signals given by movements in the credit spread regarding future economic developments. Gilchrist and Zakrajsek (2012), for example, show that the credit spread has predictive power for economic activity. Caldara and Herbst (2019) argue that restricting a simultaneous effect of credit spreads leads to a mixed monetary policy shock. The shock would be a combination of exogenous shifts in monetary policy and a systematic response of the central bank to changes in financial conditions.

Moreover, interest rates may react contemporaneously to changes in the stock price index. This would follow the assumption that the monetary authority reacts to fast moving variables and thus to the information available within the month (Arias et al., 2019; Leeper and Zha, 2003).

Examples of such generalized Taylor rules can also be found in dynamic general equilibrium or New Keynesian models. In those theoretical models, the monetary authority changes interest rates additionally in response to term spreads (Vázquez et al., 2013), growth in money (Andrés, López-Salido and Vallés, 2006), or changes in credit spread and aggregated credit (Curdia and Woodford, 2010), or stock prices (Fuhrer and Tootell, 2008; Castelnuovo and Nisticò, 2010; Lengnick and Wohltmann, 2016). Likewise, Gertler and Karadi (2011) and Gertler and Karadi (2013) characterize monetary policy by a standard Taylor rule in normal times but when the credit spread rises sharply they allow the central bank to change to credit policy or large-scale asset purchases.

The fourth row represents a reaction function of the term spread. The term spread reacts to output, prices, and interest rate immediately. In the money demand equation, fifth row, money can react contemporaneously to output, prices, interest rate and term spread (see for variours zero restrictions with respect to the money equation Christiano et al., 1999; Leeper and Zha, 2003; Sims and Zha, 2006; Belongia and Ireland, 2015).

As a fast moving variable we order the credit spread second to last (Caldara and Herbst, 2019). We put the commodity price last in order, thus establishing it as an information variable (as in, for example, Christiano et al., 1999; Sims and Zha, 2006).

3. Heteroskedastic SVARs with Markov-Switching and S^5 in the Structural Matrix

In this section, we introduce our heteroskedastic SVAR model with Markov-switching time variation in the structural matrix and search of regime-specific identification patterns. By adding those new features we extend the model by Lütkepohl, Shang, Uzeda and Woźniak (2022b). Before discussing the model, we introduce the following notation. Let $\mathbf{0}_N$ denote an N-vector of zeros, $\boldsymbol{\imath}_N$ a vector of ones, $\mathbf{e}_{n.N}$ the n^{th} column of the identity matrix of order N, denoted by $\mathbf{I}_N$. Let $[\mathbf{X}]_{n\cdot}$ denote the n^{th} row of matrix $\mathbf{X}$ and $[\mathbf{X}]_{\cdot n}$ its n^{th} column. Unless otherwise specified, the time indicator t goes from 1 to T, the variable and equation indicator n – from 1 to N, the Markov-switching regime indicator m – from 1 to M, and the S^5 indicator k_n goes from 1 to K_n.

3.1. Model specification

Consider the following reduced from vector autoregressive (VAR) model:

$$\mathbf{y}_t = \sum_{l=1}^{p} \mathbf{A}_l \mathbf{y}_{t-l} + \mathbf{A}_d \mathbf{d}_t + \boldsymbol{\varepsilon}_t \tag{1}$$

where $\mathbf{y}_t = (ip_t, cpi_t, R_t, TS_t, m_t, CS_t, sp_t)'$ is an $N = 7$-vector of observed variables at time t, $\mathbf{A}_l$ are $N \times N$ autoregressive matrices for lag $l = 1, \ldots, p$, $\mathbf{d}_t$ is a d-vector collecting the values of deterministic terms at time t, $\mathbf{A}_d$ is an $N \times d$ matrix of deterministic terms' slopes, and $\boldsymbol{\varepsilon}_t$ in an N-vector of the reduced-form shocks. We can rewrite the model in equation (1) compactly as:

$$\mathbf{y}_t = \mathbf{A}\mathbf{x}_t + \boldsymbol{\varepsilon}_t \tag{2}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & \dots & \mathbf{A}_p & \mathbf{A}_d \end{bmatrix}$ is of dimension $N \times (Np + d)$ and $\mathbf{x}_t = \begin{bmatrix} \mathbf{y}'_{t-1} & \dots & \mathbf{y}'_{t-p} & \mathbf{d}'_t \end{bmatrix}'$ a $(Np + d)$-vector.

Our model allows for time variation in impulse responses and in identification of structural shocks based on exclusion restrictions. To that end, we specify the following time-varying structural equations relating the reduced- to the structural-form shocks $\mathbf{u}_t$:

$$\mathbf{B}\left(s_t, \boldsymbol{\kappa}(s_t)\right) \boldsymbol{\varepsilon}_t = \mathbf{u}_t \tag{3}$$

where $\mathbf{B}\left(s_t, \boldsymbol{\kappa}(s_t)\right)$ denotes the time-varying $N \times N$ structural matrix. It depends on a regime indicator $s_t = m$ of a discrete Markov process with M states and on a regime-specific collection of S[5] indicators, $\boldsymbol{\kappa}(s_t)$. The latter contains equation-specific indicators, $\kappa_n(s_t)$ for all n, that for each regime select amongst K_n patterns in the exclusion restrictions imposed on the rows of the structural matrix. The Markov process, s_t, is stationary, aperiodic, and irreducible, with a $M \times M$ transition matrix $\mathbf{P}$ and an N-vector of initial values $\boldsymbol{\pi}_0$:

$$s_t \sim \mathcal{M}arkov(\mathbf{P}, \boldsymbol{\pi}_0). \tag{4}$$

For simplicity, denote by $\mathbf{B}_{m.\mathbf{k}}$ the matrix $\mathbf{B}\left(s_t, \boldsymbol{\kappa}(s_t)\right)$ for given realisations of $s_t = m$ and $\boldsymbol{\kappa}(m) = (k_1, \dots, k_N)$.

Following Waggoner and Zha (2003), the exclusion restrictions on the structural matrix are imposed by decomposing each row of the structural matrix, $[\mathbf{B}_{m.\mathbf{k}}]_{n\cdot}$, into a $1 \times r_{n.m.k_n}$ vector $\mathbf{b}_{n.m.k_n}$, collecting the elements to be estimated, and an $r_{n.m.k_n} \times N$ matrix $\mathbf{V}_{n.m.k_n}$, containing zeros and ones placing the elements of $\mathbf{b}_{n.m.k_n}$ at the appropriate spots:

$$[\mathbf{B}_{m.\mathbf{k}}]_{n\cdot} = \mathbf{b}_{n.m.k_n}\mathbf{V}_{n.m.k_n}. \tag{5}$$

We pre-specify the number of parameters to be estimated for every row, regime, and S[5] identification pattern, $r_{n.m.k_n}$, as well as the form of the structural system determined by the matrices $\mathbf{V}_{n.m.k_n}$ as in Table **??**.

The distinguishing feature of our model is its capacity to select structural shocks' identification patterns which are specific to the regimes identified by the Markov process. Such a data-based search of the identification pattern is feasible if the potential exclusion restrictions lead to an over-identified system. Hence, using the S[5] mechanism in row n of $\mathbf{B}$ requires that the n^{th} structural shock is identified within each regime absent the potential additional restrictions we are searching for. We achieve this identification via heteroskedasticity which we explain in Section 4.4.

The structural shocks at time t are contemporaneously and temporarily uncorrelated and jointly conditionally normally distributed given the past observations on vector $\mathbf{y}_t$, denoted by $\mathbf{Y}_{t-1}$ with zero mean and a diagonal covariance matrix:

$$\mathbf{u}_t \mid \mathbf{Y}_{t-1} \sim \mathcal{N}_T\left(\mathbf{0}_N, \text{diag}\left(\boldsymbol{\sigma}_t^2\right)\right) \tag{6}$$

where $\boldsymbol{\sigma}_t^2$ is an N-vector of structural shocks' conditional variances at time t.

Each of the N conditional variances, $\sigma_{n.t}^2$, follows a non-centered Stochastic Volatility process with regime-dependent volatility of the volatility that decomposes the conditional variances into

$$\sigma_{n.t}^2 = \exp\{\omega_n(s_t)h_{n.t}\} \tag{7}$$

where $\omega_n(s_t)$ is the volatility of the volatility parameter defined as plus-minus square-root of the conditional variance of the log-conditional variances $\log \sigma_{n.t}^2$, and $h_{n.t}$ is the log-volatility of the n^{th} structural shock at time t. This specification extends the Stochastic Volatility process by Kastner and Frühwirth-Schnatter (2014), which Lütkepohl et al. (2022b) incorporate into an SVAR model, by the regime-dependence in the parameter $\omega_n(s_t)$. If $\omega_n(m) = 0$, then the n^{th} structural shock is homoskedastic in the m^{th} regime. Additionally, the Markov process s_t drives the time variation in both $\mathbf{B}(s_t, \boldsymbol{\kappa}(s_t))$ and $\omega_n(s_t)$. This facilitates the assessment of identification of the structural matrix via heteroskedasticity in each of the regimes.

The log-volatility follows an autoregressive process:

$$h_{n.t} = \rho_n h_{n.t-1} + v_{n.t}, \qquad \text{and} \qquad v_{n.t} \sim \mathcal{N}(0, 1) \tag{8}$$

with the initial value $h_{n.0} = 0$, where ρ_n is the autoregressive parameter and $v_{n.t}$ is a standard normal innovation.

4. Stochastic Search Structural Specification Selection

This section presents our novel mechanism for the regime-specific data-driven stochastic search of the structural matrix identification. First, we conceptualise it as a hierarchical prior distribution. Next, we explain the estimation procedure and posterior inference on the specification selection.

4.1. S^5 for Markov-switching structural matrix

Our aim is to search for changes in the identification pattern of US monetary policy shocks across states equation-by-equation, thereby, allowing for time-varying identification of structural shocks. To that end, we introduce the S^5 mechanism and implement it via a new prior distribution. In this prior setup, given a fixed S^5 n^{th} equation component indicator, $\kappa_n(m) = k_n$, the unrestricted elements of the structural matrix' rows are normally distributed with zero mean and covariance matrix set to the product of a hyper-parameter γ_B and an identity matrix:

$$\mathbf{b}'_{n.m.k_n} \mid \gamma_B, \kappa_n(m) = k_n \sim \mathcal{N}_{r_{n.m.k_n}}\left(\mathbf{0}_{r_{n.m.k_n}}, \gamma_B \mathbf{I}_{r_{n.m.k_n}}\right). \tag{9}$$

This conditional distribution is equivalent to the normal prior assumed by Waggoner and Zha (2003) for a time-invariant model, and it can be presented as a restricted generalised–normal distribution proposed by Arias, Rubio-Ramírez and Waggoner (2018) with shape parameter set to value N. Furthermore, (9) is combined with a multinomial prior distribution over a single trial for the S^5 component indicator $\kappa_n(m) = k_n \in \{1, \dots, K_n\}$ with flat probabilities equal to $\frac{1}{K_n}$:

$$\kappa_n(m) \sim \mathcal{M}ultinomial\left(K_n^{-1}\boldsymbol{\imath}_{K_n}\right). \tag{10}$$

This combination of the conditional zero-mean normal prior distribution for the elements of the structural matrix from equation (9) with the S^5 component specific identifying restrictions from equation (5), and the marginal multinomial distribution for the S^5 component indicator from equation (10), makes the S^5 mechanism a multi-component generalisation of a spike-and-slab prior by Geweke (1996). Moreover, we consider our approach an extension of the stochastic search variable selection (George and Mcculloch, 1997; George et al., 2008) and, in particular, of the stochastic search specification selection (Koop and Korobilis, 2016). The latter assumes a Gaussian mixture prior distribution on a parameter vector where the mixing weights follow a Binomial prior. The normal components differ in the covariance terms such that one component induces strong shrinkage towards a zero mean vector while the other component allows for an unrestricted estimation due to an uninformative prior. Hence, high posterior mass for or against exclusion restrictions concern all parameters in a vector of coefficients equally, distinguishing the *specification* from the *variable* selection. While our mechanism builds upon the idea of a specification selection, we generalize the set prior distribution with respect to Koop and Korobilis (2016) by allowing for multiple components and by the treatment of the restricted component.

In our S^5 prior specification, in some components, the elements of the structural matrix might be marginally normally distributed with zero mean and variance equal to γ_B with probability K_n^{-1}, and in some other components, they are assigned a Dirac mass at value zero, δ_0, with the same prior probability. Choosing a symmetric normal prior distribution with probability mass around zero makes our prior specification compatible with spike-and-slab priors. In those priors, the distribution of the unrestricted component is centred at the restrictions from the restricted component (see Malsiner-Walli and Wagner, 2018, for a recent review).

The relationship of our specification to the spike-and-slab prior can be best seen by analysing the prior distribution for an $(n, i)^{\text{th}}$ element of the regime-specific matrix $\mathbf{B}(m, \boldsymbol{\kappa}(m))$, denoted by $[\mathbf{B}_m]_{n.i}$, marginalised over $\boldsymbol{\kappa}(m)$. Such an element is restricted to zero in $K_{R.n}$ components, whereas it stays unrestricted and normally distributed in $K_n - K_{R.n}$ of them. Therefore, its prior distribution is a Dirac mass at value zero with probability $\frac{K_{R.n}}{K_n}$, and normal with probability $\frac{K_n - K_{R.n}}{K_n}$:

$$[\mathbf{B}_m]_{n.i} \mid \gamma_B \sim \frac{K_n - K_{R.n}}{K_n} \mathcal{N}(0, \gamma_B) + \frac{K_{R.n}}{K_n} \delta_0, \tag{11}$$

which corresponds exactly to the formulation of a spike-and-slab prior.

Lastly, we set a marginal prior on the structural matrix overall level of shrinkage γ_B that follows an inverted gamma 2 distribution with scale $\underline{s}_B$ and shape $\underline{\nu}_B$:

$$\gamma_B \sim \mathcal{IG}2\left(\underline{s}_B, \underline{\nu}_B\right). \tag{12}$$

The hierarchical prior for the shrinkage parameter ensures flexibility of our prior distribution for the parameters of the structural matrices. Importantly, it avoids arbitrary choices regarding the specification as the level of shrinkage is estimated within the model.

4.2. Bayesian estimation

Our novel model facilitates the estimation of the time-varying structural matrix with regime-specific stochastic specification selection with respect to the identifying pattern of exclusion restrictions. The prior setup combines the flexibility of a hierarchical specification with the interpretability of the spike-and-slab prior. At the same time, it leads to an efficient estimation procedure.

To estimate the parameters in the structural matrix, we use a Gibbs sampler. The full conditional posterior distribution of the unrestricted elements of the structural matrix is generalised-normal, proportional to

$$|\det(\mathbf{B}_{m.\boldsymbol{\kappa}})|^{T_m} \exp\left\{-\frac{1}{2}\sum_{n=1}^{N} \mathbf{b}_{n.m.k_n}\mathbf{V}_{n.m.k_n}\overline{\boldsymbol{\Omega}}_B^{-1}\mathbf{V}'_{n.m.k_n}\mathbf{b}'_{n.m.k_n}\right\} \tag{13}$$

where T_m is the number of observations in regime m, such that $s_t = m$, and with scale matrix $\overline{\boldsymbol{\Omega}}_B$

$$\overline{\boldsymbol{\Omega}}_B^{-1} = \gamma_B^{-1}\mathbf{I}_N + \sum_{t:s_t=m}(\mathbf{y}_t - \mathbf{A}\mathbf{x}_t)(\mathbf{y}_t - \mathbf{A}\mathbf{x}_t)'. \tag{14}$$

Within each of the regimes, m, the sampling algorithm proceeds row-by-row following the algorithm by Waggoner and Zha (2003) and is combined with sampling the S^5 indicators $\boldsymbol{\kappa}(m)$. The rows for which the exclusion restrictions are not subject to stochastic selection, $K_n = 1$, are sampled straight away from the full conditional posterior. For the rows with the stochastic search of exclusion restrictions, we first sample K_n vectors $\mathbf{b}_{n.m.k_n}$ from the full conditional posterior distribution given $\kappa_n(m) = k_n$ for all $k_n \in \{1, \dots, K_n\}$, then we compute the full conditional posterior probabilities of each of these S^5 components, $\overline{p}_{n.m.k_n}$, that are proportional to the product of the likelihood function, $L(\theta \mid \mathbf{Y}_T)$, and the prior distributions:

$$\overline{p}_{n.m.k_n} \propto L(\theta \mid \mathbf{Y}_T, \kappa_n(m) = k_n)p\left(\mathbf{B}_{n.m.k_n} \mid \gamma_B, \kappa_n(m) = k_n\right)p\left(\kappa_n(m) = k_n\right). \tag{15}$$

The realisation of the S^5 indicator is sampled from the multinomial distribution, denoted by k_n, and is returned together with the corresponding vector $\mathbf{b}_{n.m.k_n}$ as the draw from the joint distribution of $\mathbf{b}_{n.m.\kappa_n(m)}$ and $\kappa_n(m)$. Finally, we sample the shrinkage hyper-parameter from its inverted gamma 2 full conditional posterior distribution:

$$\gamma_B \mid \mathbf{B}(s_t, \kappa(s_t)), \kappa(s_t) \sim \mathcal{IG}2\left(\underline{s}_B + \sum_n \sum_m \mathbf{b}_{n.m.k_n} \mathbf{b}'_{n.m.k_n}, \quad \underline{\nu}_B + \sum_n \sum_m r_{n.m.k_n} \right). \tag{16}$$

4.3. Posterior inference for structural identification patterns

Having estimated a model and obtained the sample of S posterior draws, the inference about equation-specific and cross-equation identifying restrictions is straightforward. In order to see which set of exclusion restrictions in the n^{th} equation is best supported by the data estimate the marginal posterior distribution of the indicator $\kappa_n(m)$. This involves the estimation of each S[5] component posterior probability by computing the fraction of posterior draws for which the indicator, $\kappa_n(m)^{(s)}$, takes a particular value k_n, for each of its values from 1 to K_n:

$$\widehat{\Pr}\left[\kappa_n(m) = k_n \mid \mathbf{Y}_T\right] = S^{-1} \sum_{s=1}^{S} \mathcal{I}(\kappa_n(m)^{(s)} = k_n). \tag{17}$$

One could also be interested in a conditional posterior distribution of the n^{th} equation specification given a particular specification of the i^{th} row. That can be assessed by the estimation of the following probabilities for all values k_n can take:

$$\widehat{\Pr}\left[\kappa_n(m) = k_n \mid \mathbf{Y}_T, \kappa_i(m) = k_i^*\right] = S^{-1} \sum_{s=1}^{S} \mathcal{I}\left(\kappa_n(m)^{(s)} = k_n \wedge \kappa_i(m)^{(s)} = k_i^*\right). \tag{18}$$

Finally, if a specific economic interpretation is given to a particular combination of exclusion restrictions for the structural matrix, one can estimate the joint posterior probability of such a specification by computing the fraction of the posterior draws for which this combination of S[5] indicators holds:

$$\widehat{\Pr}\left[\kappa_1(m) = k_n^* \wedge \cdots \wedge \kappa_N(m) = k_i^* \mid \mathbf{Y}_T\right] = S^{-1} \sum_{s=1}^{S} \mathcal{I}\left(\kappa_1(m)^{(s)} = k_n \wedge \cdots \wedge \kappa_N(m)^{(s)} = k_i\right). \tag{19}$$

The decision regarding the model specification selection is left to the investigator. They can decide to choose the specification that is *a posteriori* the most likely or accept it only if its marginal posterior probability exceeds 50%.

4.4. Achieving over-identification of the exclusion restrictions

We do not restrict the exclusion restriction patterns to be lower-triangular structural matrices. Hence, additional parameters to be estimated imply that the system is not identified through zero restrictions. To ensure that our data-driven search of the identification patterns is feasible for the n^{th} row, the additional exclusion restrictions on $[\mathbf{B}_{m.\mathbf{k}}]_{n\cdot}$ need to be over-identifying. To achieve that, we identify the n^{th} structural shock via heteroskedasticity. Hence, we uncover the instantaneous uncorrelated structural shocks based on changes in volatility. To that end, we model heteroskedasticity via our specific stochastic volatility representation as given in equations (7) to (8), (similar to Lütkepohl et al., 2022b). Alternatively, changes in the reduced form covariance can be implemented via standard stochastic volatility processes (with no dependence on the regime as in Bertsche and Braun, 2022), exogenous changes (Rigobon, 2003), (G)ARCH models (Normandin and Phaneuf, 2004; Milunovich and Yang, 2013; Lütkepohl and Milunovich, 2016), Markov switching (Lanne, Lütkepohl and Maciejowska, 2010) or smooth transitions (Lütkepohl and Netšunajev, 2017). The additional exclusion restrictions can provide the lacking economic labeling to the so (statistically) identified structural shocks.

Within each regime, the regime-specific reduced from unconditional covariance matrix, $\Sigma_{\varepsilon.m}$, can be decomposed as

$$\Sigma_{\varepsilon.m} = \mathbf{B}_{m.\mathbf{k}}^{-1} \operatorname{diag}\left(\sigma_t^2\right) \mathbf{B}_{m.\mathbf{k}}^{-1\prime}. \tag{20}$$

Within a regime we can apply the same identification result as derived by Lütkepohl et al. (2022b) using our specific parameterization as given in equation (7). That is, the n^{th} row of $\mathbf{B}_{m.\mathbf{k}}$ is identified if $\omega_n(m) \neq \omega_k(m)\ \forall k \in \{1, \ldots, N\}\backslash\{n\}$. A structural shock is identified through heteroskedasticity if changes in the conditional variances are non-proportional. Note that it is crucial here that our model links the time variation in the structural matrix to the changes in the volatility of the volatility parameter $\omega_n(m)$ via the same Markov process.

The validity of the S^5 mechanism for the n^{th} equation depends on the identification through heteroskedasticity of the n^{th} structural shock within a regime. The n^{th} structural shock is identified through heteroskedasticity in two cases. First, the shock is identified when this shock, and no other, is homoskedastic within a specific regime. Second, in the presence of any other homoskedastic shock in the system, the shock is identified if the n^{th} structural shock is heteroskedastic in that regime. Therefore, heteroskedasticity verification plays a central role in our setup. Subsequently, we first explain the priors and posterior simulation for the relevant parameters and then, we discuss heteroskedasticity testing in Section 6.

5. Prior Distributions and Posterior Sampler

In this section, we discuss the prior specification of the remaining parameters of the model and their estimation. The objective of setting the prior distributions is to combine the analytical feasibility of deriving an efficient Gibbs sampler. The latter is outlined as well.

5.1. Priors for conditional mean parameters

The prior distribution for the autoregressive parameters incorporates the ideas of the Minnesota prior by Doan, Litterman and Sims (1984) combined with the estimation of the level of prior shrinkage as inspired by Giannone, Lenza and Primiceri (2015). The former is implemented via the construction of the prior mean and covariances, the latter by a hierarchical prior. Moreover, it facilitates numerically fast and efficient estimation. Each of the rows of matrix $\mathbf{A}$ follows independently a conditional multivariate normal distribution with the mean set to vector $\underline{\mathbf{m}}_{A.n} = \begin{bmatrix}\mathbf{e}'_{n.N} & \mathbf{0}'_{N(p-1)+d}\end{bmatrix}$ and the covariance $\gamma_A \underline{\mathbf{\Omega}}_A$:

$$[\mathbf{A}]'_{n\cdot} \mid \gamma_A \sim \mathcal{N}_{Np+d}\left(\underline{\mathbf{m}}'_{A.n}, \gamma_A \underline{\mathbf{\Omega}}_A\right) \tag{21}$$

where γ_A is the shrinkage parameter common to all elements of matrix $\mathbf{A}$, and $\underline{\mathbf{\Omega}}_A$ is a diagonal matrix with vector $\left[\mathbf{p}^{-2\prime} \otimes \boldsymbol{\imath}_N' \quad 100\boldsymbol{\imath}_d'\right]'$ on the main diagonal, where $\mathbf{p}$ is a vector containing a sequence of integers from 1 to p. This specification includes the shrinkage level exponentially decaying with the increasing lag order, relatively large prior variances for the deterministic term parameters, and the flexibility of the hierarchical prior that leads to the estimation of the level of shrinkage. We implement the last feature by setting a prior on the overall reduced form parameters shrinkage that follows:

$$\gamma_A \sim \mathcal{IG}2\left(\underline{s}_A, \underline{\nu}_A\right). \tag{22}$$

Specifying the Markov process parameters, we assume that each of the regimes lasts *a priori* on average for 11 months. Hence, each of the rows of the transition probabilities matrix $\mathbf{P}$ follows independently a Dirichlet distribution with unit parameters for all of the row's elements except for the one corresponding to the diagonal element for which the parameter is equal to $1 + d_m$. To assure the prior expected regime duration of 11 months, we set $d_m = 9$ for $M = 2$ regime models, and $d_m = 19$ for those with $M = 3$ regimes. The initial probabilities $\boldsymbol{\pi}_0$ follow a Dirichlet distribution with unit parameters:

$$[\mathbf{P}]_{m\cdot} \sim \mathcal{D}irichlet(\boldsymbol{\imath}_M + d_m \mathbf{e}_{m.M}), \qquad \boldsymbol{\pi}_0 \sim \mathcal{D}irichlet(\boldsymbol{\imath}_M). \tag{23}$$

5.2. *Priors for conditional variance parameters*

The prior specification for the parameters of the Stochastic Volatility process combines the flexibility of the specification via a hierarchical structure with the benefits regarding the normalisation and heteroskedasticity verification following Lütkepohl et al. (2022b).

The prior distribution for the regime-specific volatility of the volatility process, $\omega_n(s_t)$, follows a zero-mean conditional normal distribution given the shrinkage level $\sigma^2_{\omega.n}$:

$$\omega_n(m) \mid \sigma^2_{\omega.n} \sim \mathcal{N}\left(0, \sigma^2_{\omega.n}\right). \tag{24}$$

The level of shrinkage, $\sigma^2_{\omega.n}$, follows a gamma distribution with the scale $\underline{s}$, shape $\underline{a}$, and the expected value $\underline{sa}$:

$$\sigma^2_{\omega.n} \sim \mathcal{G}(\underline{s}, \underline{a}). \tag{25}$$

This prior setup following Lütkepohl et al. (2022b) implies that the marginal prior for $\omega_n(m)$ combines extreme prior probability mass around the restriction for homoskedasticity $\omega_n(m) = 0$ and fat tails. Consequently, it centres the prior probability mass of conditional variances along value one that implies homoskedasticity. Therefore, whenever a shock is identified through heteroskedasticity, the signal comes entirely from data. Finally, this prior specification facilitates Bayesian verification of homoskedasticity which we discuss in detail in Section 6. All these properties hold given some additional assumptions that include $|\rho_n| < 1$, $\sigma^2_{\omega.n}/(1-\rho_n^2) \leq 1$, and $\underline{a} > 0.5$ (see Lütkepohl et al., 2022b). Lastly, the prior distribution for the autoregressive parameter of the log-volatility process is uniform over the stationarity region:

$$\rho_n \sim \mathcal{U}(-1, 1). \tag{26}$$

5.3. Bayesian estimation

Estimation of the parameters considered in this section is implemented by the Gibbs sampler. Our emphasis was put on the selection of the most efficient sampling techniques that would facilitate estimation in larger systems of variables for which the dimension of the parameter space grows rapidly. The Gibbs sampler iteration for the autoregressive parameters, A, follows a numerically fast and efficient equation-by-equation sampler by Chan, Koop and Yu (2022) in which the parameters are drawn from a multivariate normal distribution. The autoregressive overall shrinkage level is sampled from the inverted gamma 2 full conditional posterior distribution.

The rows of the transition probabilities matrix, $\mathbf{P}$, and the initial state probabilities vector, $\boldsymbol{\pi}_0$, are all sampled from independent M-variate Dirichlet full conditional posterior distributions as in Frühwirth-Schnatter (2006). The sampling procedure for the Markov process realisations follows the forward-filtering backward-sampling algorithm proposed by Chib (1996).

Estimation of the Stochastic Volatility process via Gibbs sampler is possible thanks to the auxiliary mixture by Omori, Chib, Shephard and Nakajima (2007a), simulation smoother and precision sampler as presented by McCausland, Miller and Pelletier (2011). Finally, efficient estimation when heteroskedasticity is uncertain requires the implementation of the ancillarity-sufficiency interweaving strategy by Kastner and Frühwirth-Schnatter (2014).

The computer code for the Gibbs sampler is provided for R (R Core Team, 2021) in package **bsvarTVPs** that is available in a public repository.[1] Excellent computational speed of the algorithms is obtained by the application of frontier econometric techniques and compiled code written in C++ using the R packages **Rcpp** by Eddelbuettel, François, Allaire, Ushey, Kou, Russel, Chambers and Bates (2011), **RcppArmadillo** by Eddelbuettel and Sanderson (2014), and **bsvars** by Woźniak (2022).

6. Assessment of heteroskedasticity within regimes

Our model specification explicitly binds the time variation in the structural matrix with the volatility of the volatility parameter $\omega_n(s_t)$. They both depend on the same Markov process s_t and, thus, change the regime simultaneously. This design facilitates the verification of the hypothesis of within-regime homoskedasticity of each of the shocks represented by the restriction:

$$\omega_n(m) = 0. \tag{27}$$

If this restriction holds then the n^{th} structural shock is homoskedastic in the m^{th} regime, which should be verified for all equations and regimes.

As shown by Lütkepohl et al. (2022b) in a time-invariant heteroskedastic model, verifying the restriction of zero volatility of the volatility is a useful tool for making statements

[1] Access the **bsvarTVPs** package at `https://github.com/donotdespair/bsvarTVPs`

regarding the identification of structural shocks. In that setting, detecting a homoskedastic shock, in the presence of another homoskedastic shock, violates the condition of non-proportional changes in conditional variances providing the identification. Consequently, both such shocks are not identified through heteroskedasticity. More generally, at most one homoskedastic shock is allowed for the whole structural matrix to be identified.

Importantly, the reliability of the regime-specific S^5 of the parameters in this equation relies on identification via heteroskedasticity within the regimes. In our setup, with simultaneous regime changes in the structural matrix and the volatility of the volatility parameter, verifying restrictions from equation (27) allows us to validate the S^5 for the n^{th} equation.

To that end, we use the Savage-Dickey Density Ratio by Verdinelli and Wasserman (1995). It is the ratio of the marginal posterior density to the marginal prior density ordinates both evaluated at the restriction:

$$SDDR_{n.m} = \frac{p\left(\omega_n(m) = 0 \mid \mathbf{Y}_T\right)}{p\left(\omega_n(m) = 0\right)}. \tag{28}$$

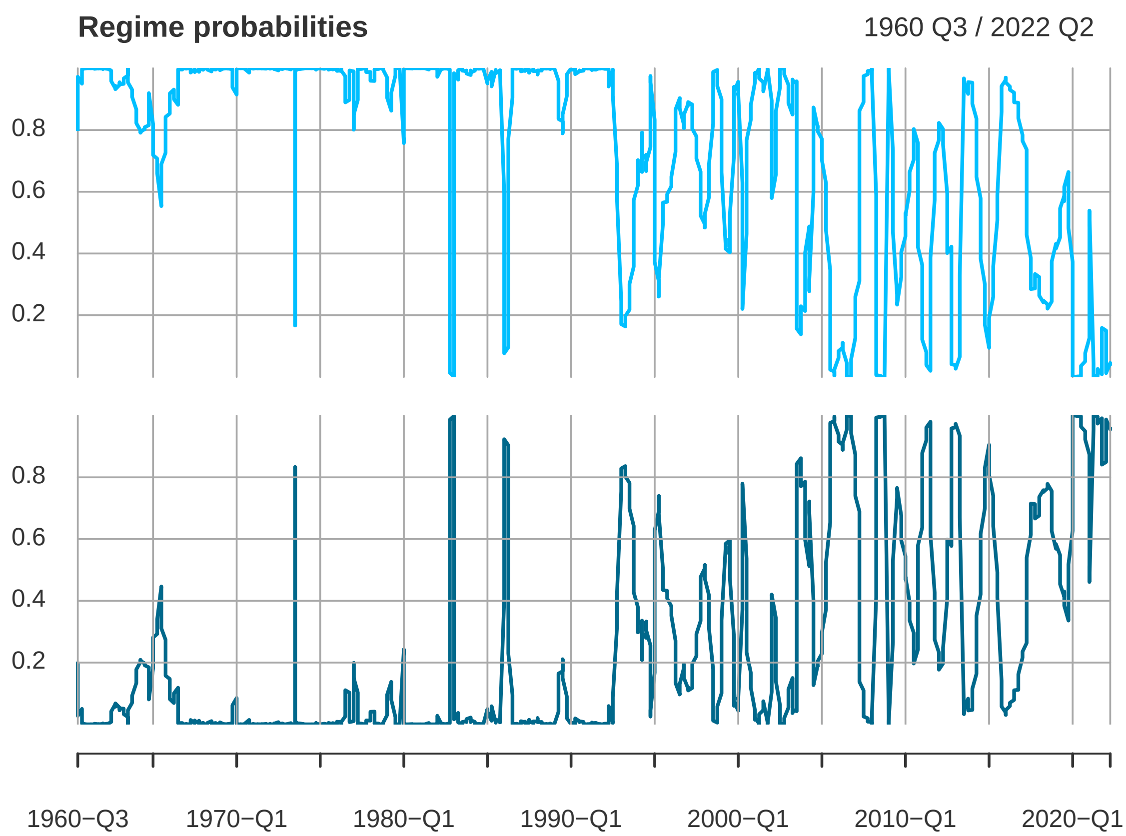
The SDDR is interpreted as the Bayes factor for two models: one with homoskedastic and another heteroskedastic shock in the m^{th} regime. Its value greater than 1 implies a higher concentration of the posterior than of the prior probability mass around the restriction and provides evidence in favour of homoskedasticity. On the contrary, its value of less than 1 and greater than zero indicates heteroskedasticity of the shock in that regime.

Our prior setup for the verified parameter closely follows that by Lütkepohl et al. (2022b) which shrinks the prior probability mass strongly towards the restriction and assures that the prior density function is bounded and heavy-tailed at once. All these features make the verification of the restriction using the SDDR feasible, numerically fast, reliable, and less dependent on the arbitrary prior choices than the heteroskedasticity testing approach proposed by Chan (2018). Note that to date, there does not exist a comprehensive time-varying identification through heteroskedasticity approach, nor does regime-specific heteroskedasticity testing.

The numerator and denominator of the SDDR are computed by numerical integration using the estimator by Gelfand and Smith (1990). The corresponding conditional prior and full conditional posterior distributions of $\omega_{n.m}$ are known up to their densities provided by Lütkepohl et al. (2022b). The computations for the numerator need to be expanded by conditioning on the realisations of the Markov process in the regime-specific case. Those for the denominator remain unchanged.

7. What data say about time-variation in monetary policy shock identification

Figure 1: Regime probabilities of the Markov process



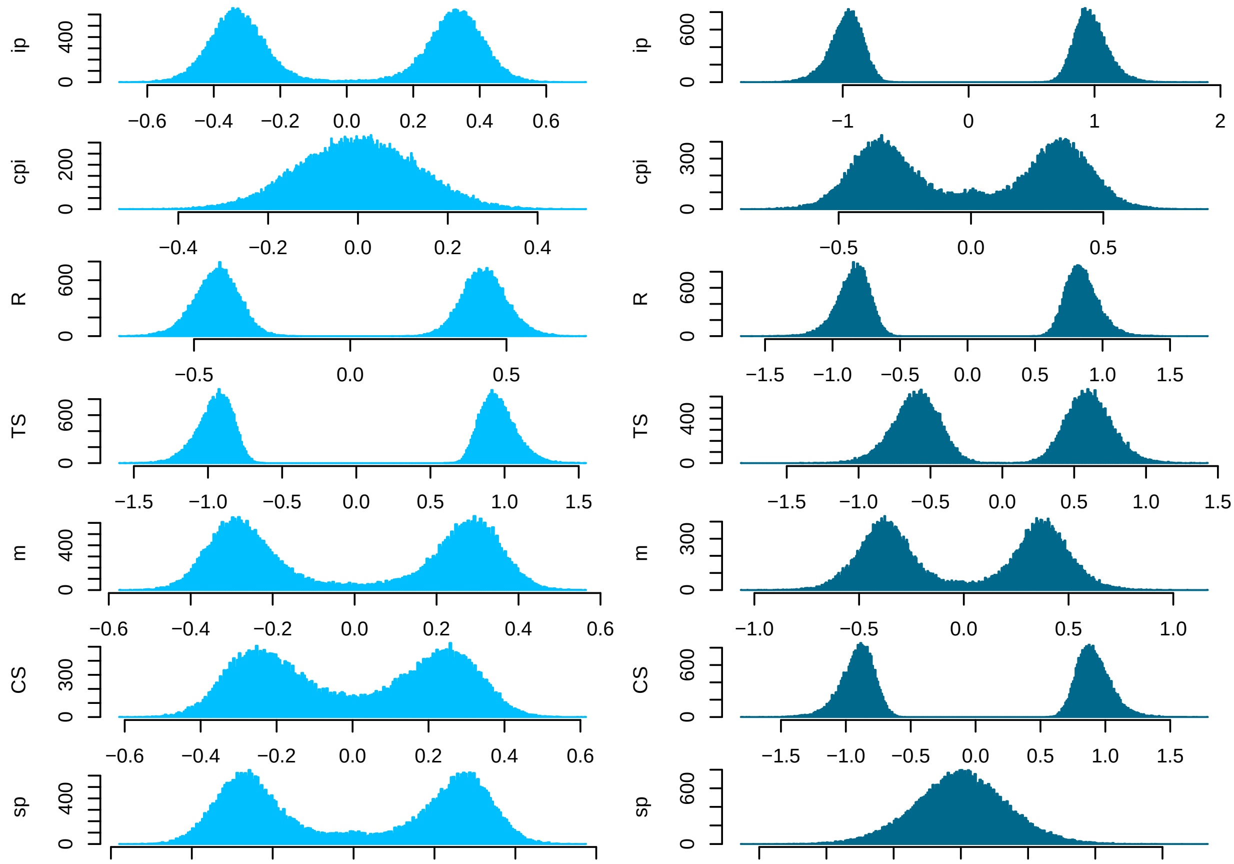
Note: Figure shows the estimated regime probabilities of the first regime (light color) in the top subplot and of the second regime (dark color) in the bottom subplot.

We estimate our model with two regimes.[2] Figure 1 shows the regime probabilities

[2]The computations for this paper were performed at the Spartan HPC-Cloud Hybrid (see Meade,

over time for the two regimes. Regime one is very persistent in the first part of the sample until early 1990s and predominates until 2008. The second regime prevails during the financial and Covid-19 crisis.

Figure 2: Histograms of the regime-specific stochastic volatility parameters ω_n



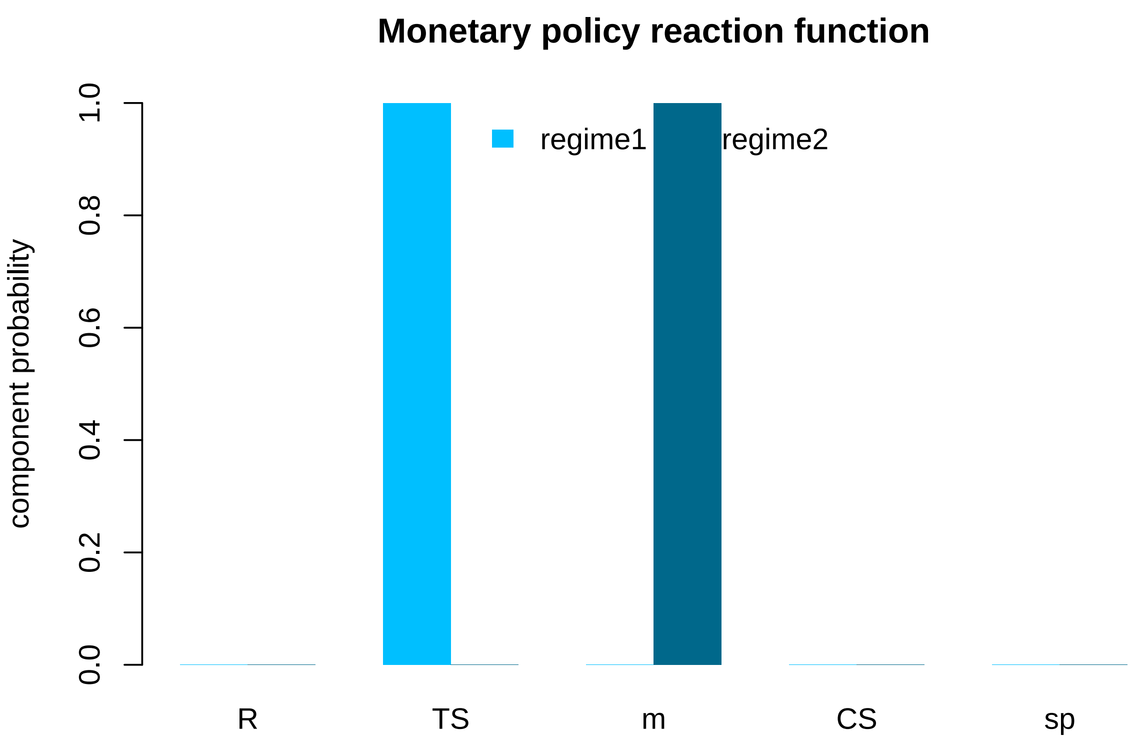
Note: Figure shows the regime-specific volatility of the volatility parameters $\omega_n(m)$ for $m = 1, 2$ (in columns) for all variables (in rows).

Within both regimes, we can verify identification via heteroskedasticity for the monetary policy reaction function. Figure 2 plots regime-specific volatility of the volatility parameters $\omega_n(m)$ for $m = 1, 2$ (in columns) for all variables (in rows). We find clear evidence for heteroskedasticity within regimes for the monetary policy reaction function visualized by the bi-modality of the posterior distribution. Hence, the restriction

Lafayette, Sauter and Tosello, 2017) at the University of Melbourne and at the Dutch national e-infrastructure Lisa Compute Cluster (see Lisa SURFSARA).

search for regime-specific identification pattern in the monetary policy reaction function is possible.

Figure 3: Posterior probabilities of the S^5 mechanism in the monetary policy reaction function



Note: Figure shows the posterior probabilities of $\kappa_n(m)$ for identification pattern in the monetary policy reaction function for regime one (light color) and regime two (dark color).

Next, Figure 3 gives the posterior probabilities of the S^5 mechanism for the five possible options in the shadow rate equation. That is, shadow rates depend contemporaneously on ("R") only output and prices, ("TS") these two and term spread, ("m") these two and money, ("CS") these two and credit spread, and ("sp") these two and stock prices. For the first regime, we find strong data support for the inclusion of term spread. The posterior probability of $\kappa_2(m = 1)$ – S^5 component indicator for the restriction pattern where no restriction is placed on the term spread – is one. Hence, monetary policy reacts simultaneously to changes in the term spread signaling changes in the expectations on future economic developments in regime one.

Several papers argue that term spreads are a target of unconventional monetary policy as long term interest rates are the main transmission channel of the large asset purchase programs (Baumeister and Benati, 2013; Liu et al., 2017; Feldkircher and Huber, 2018; Tillmann, 2020). The monetary authority can target term spreads via the portfolio balance channel; increased asset purchases decrease the supply of long-term securities and lower long-term yields. Gagnon, Raskin, Remache and Sack (2011), Krishnamurthy and Vissing-Jorgensen (2011), and D'Amico and King (2013), among others, provide evidence that the large asset purchase programs during the zero lower bound period affect long-term yields. Our results extend those findings by showing that the monetary policy reaction function of the Fed accounts for the information revealed by term spreads even before the start of unconventional monetary policy since the first regime occurs mainly already before 2008.

Table 2: Posterior means of regime-specific contemporaneous parameters in the monetary policy reaction function

	ip	*cpi*	*R*	*TS*	*m*	*CS*	*sp*
TR regime 1	9.98	-64.78	3.10	3.58	0.00	0	0
sd	7.13	24.07	0.16	0.16	0.00	0	0
lower	2.86	-88.67	2.92	3.41	0.00	0	0
upper	16.98	-40.88	3.25	3.74	0.00	0	0
TR regime 2	-5.87	22.25	1.26	0.00	233.60	0	0
sd	8.82	25.31	0.87	0.00	21.01	0	0
lower	-13.70	-1.27	0.45	0.00	212.86	0	0
upper	3.57	48.52	1.91	0.00	254.01	0	0

Note: Table gives the posterior means of the contemporaneous parameters in the monetary policy reaction function (in columns) for regime one (in row TR regime 1) and two (in row TR regime 2) together with the estimated standard deviation (sd) and the lower and upper bound of the highest density interval.

Contemporaneously, we find evidence for a positive relation between the terms spread and interest rates in regime one. The posterior mean of the contemporaneous coefficient on term spreads in the monetary reaction function is positive with strong evidence that it is different from zero, shown in Table 2.

In the second regime, the monetary authority responds simultaneously to changes in the monetary aggregate. The posterior value of $\kappa_4(m = 2)$ – S^5 component indicator for the restriction pattern where no restriction is placed on money – is one, dark colored bar in Figure 3. The posterior mean of the coefficient of the monetary aggregate in the monetary policy reaction function is positive and different from zero. Notably, our results provide evidence that the shadow rate coefficient is not different from zero. That points into the direction that the monetary policy indicator switches in the second regime to the monetary aggregate.

Our findings are in line with Belongia and Ireland (2015) supporting that after 2008 monetary policy actions lead to changes in the monetary aggregates. While the Fed officially targeted money supply before 1982, Belongia and Ireland (2015) argue that unconventional monetary policy can be seen as attempts to increase money growth. Similar to the strong support of data for money in the monetary policy rule in the second regime, the authors show that excluding a monetary measure from the interest rate rule is rejected by the data in the sample from mid 1960s to 2013. Following a similar argument, Kimura and Nakajima (2016) distinguish between a conventional and unconventional monetary policy shock by setting a zero restriction on bank reserves in the interest rate equation (conventional monetary policy) or having a zero restriction on the interest rate in the bank reserves demand equation (unconventional monetary policy). In both cases, the corresponding opposite element is non-zero.

7.1. Alternative Patterns in the Restriction Search

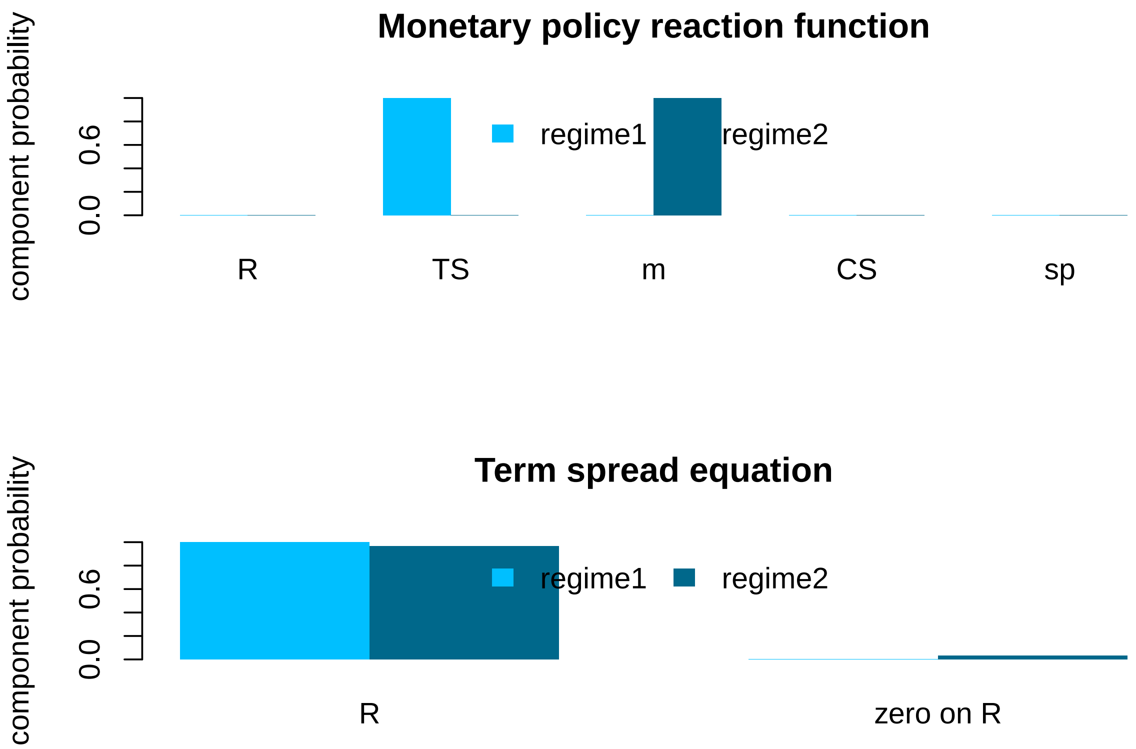
We further investigate the role of the term spread by applying our S^5 mechanism to the monetary policy reaction function and also to the term spread rule. Here, we search for a zero restriction on the simultaneous response of terms spreads to changes in interest rates, S^5 element in the fourth row in the structural matrix given in Table 3.

A potential zero restriction on the contemporaneous reaction of the shadow interest rate in the fourth row is motivated by the zero lower bound period. Baumeister and Benati (2013), Feldkircher and Huber (2018), and Liu et al. (2017) use sign restrictions to identify a term spread shock, as a measure of an unconventional monetary policy shock,

Table 3: Regime-specific identification patterns of monetary policy and term spread shocks

$$
\begin{bmatrix}
* & 0 & 0 & 0 & 0 & 0 & 0 \\
* & * & 0 & 0 & 0 & 0 & 0 \\
* & * & * & S^5 & S^5 & S^5 & S^5 \\
* & * & S^5 & * & 0 & 0 & 0 \\
* & * & * & * & * & 0 & 0 \\
* & * & * & * & * & * & 0 \\
* & * & * & * & * & * & *
\end{bmatrix}
\begin{bmatrix}
ip_t \\ cpi_t \\ R_t \\ TS_t \\ m_t \\ CS_t \\ sp_t
\end{bmatrix}
$$

Figure 4: Posterior probabilities of the S^5 mechanism in the monetary policy reaction function and term spread equation



Note: Figure shows the posterior probabilities of $\kappa_n(m)$ for identification pattern in the monetary policy reaction function (top subplot) and term spread equation (bottom subplot) for regime one (light color) and regime two (dark color). For the term spread equation, "R" gives the posterior indicator for the restriction pattern where no restriction is placed on shadow rates and "zero on R" where an exclusion restriction is placed on shadow rates.

but combine it with a single zero restriction on the short-term interest rate reaction. Thus, when short term rates reach the zero lower bound, the term spread reacts no longer to such zero interest rates.

Allowing for data-driven restriction search in the term spread equation does not change the posterior probabilities for including term spreads in the monetary policy reaction function in regime one and for the monetary aggregate in regime 2, first plot in Figure 4. Data do not support a zero restriction on shadow interest rates in the term spread equation in both regimes. The posterior probability for including shadow interest rates, denoted by "R" in Figure 4, is one in the first regime and very close to one in the second regime and, thus, the probabilities for an exclusion restriction, labeled by "zero on R", are zero. Since we include shadow rates in our model which also move to negative values in the zero lower bond period, support for a zero restriction on the interest rates is less likely. Hence, the term spreads reacts to movements in the shadow rate contemporaneously.

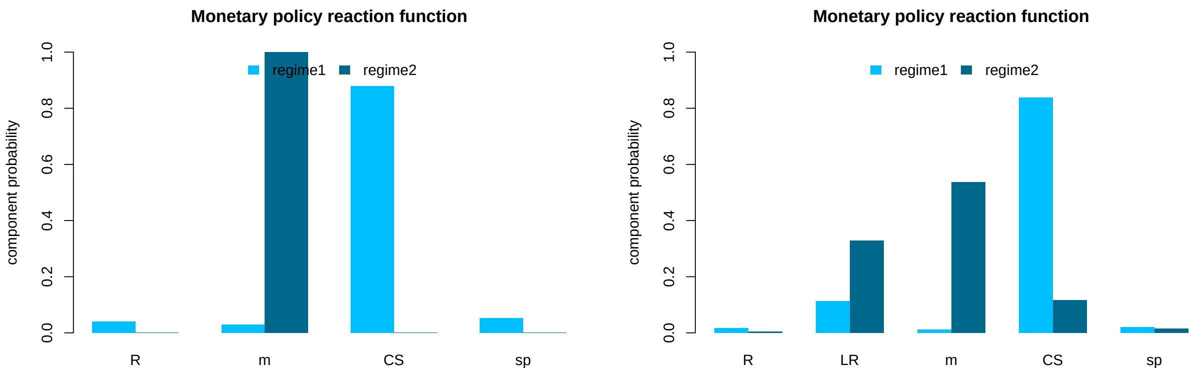
7.2. Robustness analysis

We further investigate the role of the term spread. To that end, we estimate a model without the term spread. In that model the credit spread takes over the role of the term spread. The posterior probabilities of including the credit spread in the monetary policy reaction function is above 80% in the first regime, while money is selected to be included in the second regime, panel (a) in Figure 1. Similarly to the term spread, the credit spread can also signal changes in the economic environments. Hence, it seems that the central bank reacts to early signals on economic developments in the first regime. Next, we estimate a model including long term interest rates (*LR*) instead of the term spread. We find support of around 80% probability to include credit spreads in the monetary policy reaction function in the first regime and above 10% probability to include long term interest rates, panel (b) in Figure 1. The results for the second regime are somehow mixed with highest posterior probability for the monetary aggregate, followed by long term interest rates, and then by credit spreads.

Furthermore, if we exclude the restriction search but allow for Markov-switching

Figure 5: Posterior probabilities of the S^5 mechanism in the monetary policy reaction function for models excluding term spreads

(a) excluding *TS* (b) including *LR* instead of *TS*



Note: Figure shows the posterior probabilities of $\kappa_n(m)$ for identification pattern in the monetary policy reaction function for different models for regime one (light color) and regime two (dark color). The first model does not include the term spread, the second model includes long term interest rates instead of term spreads.

in the matrix of contemporaneous coefficients we find a less persistent second regime. This second regime is based on only a low number of observations. Hence, allowing for time-varying identification of the monetary policy reaction function is crucial to find regime-dependence in the empirical Taylor rule specification.

Moreover, varying the degree of shrinkage in the prior distributions of autoregressive and contemporaneous parameters does not change the finding of including the term spread in the monetary policy reaction function in the first regime.

8. Conclusions

We propose a Bayesian heteroskedastic structural vector autoregressive model with Markov-Switching and data-driven regime-specific identification search. Our model has as new features Markov-switching time-variation in the structural matrix of contemporaneous effects and in a stochastic volatility parameter as well as a data-supported search of regime-specific exclusion restriction patterns. Those new features allow us to study time-variation in monetary policy identification and provide new insights into which exclusion restrictions are supported by the data. We facilitate the search through a new stochastic search structural specification selection mechanism which generalizes selection priors and is feasible by identifying the monetary policy reaction function via heteroskedastic residuals. With our new model we can challenge the common assumption of time-invariant identification of monetary policy shocks.

Our findings show that data support time-variation in US monetary policy shock identification. We find strong evidence for including the term spread in the monetary policy reaction function in the first regime predominating before 2008. In the second regime mainly occurring after 2008, the monetary authority reacts simultaneously to changes in shadow rates and in a monetary aggregate. Hence, our results support regime-specific empirical generalized Taylor rule specifications to model the systematic component of monetary policy.

References

Andrés, J., López-Salido, J.D., Vallés, J., 2006. Money in an Estimated Business Cycle Model of the Euro Area. Economic Journal 116, 457–477.

Ang, A., Boivin, J., Dong, S., Loo-Kung, R., 2011. Monetary Policy Shifts and the Term Structure. The Review of Economic Studies 78, 429–457. doi:10.1093/restud/rdq006.

Arias, J.E., Caldara, D., Rubio-Ramírez, J.F., 2019. The systematic component of monetary policy in SVARs: An agnostic identification procedure. Journal of Monetary Economics 101, 1–13. doi:https://doi.org/10.1016/j.jmoneco.2018.07.011.

Arias, J.E., Rubio-Ramírez, J.F., Waggoner, D.F., 2018. Inference based on structural vector autoregressions identified with sign and zero restrictions: Theory and applications. Econometrica 86, 685–720.

Bacchiocchi, E., Castelnuovo, E., Fanelli, L., 2017. Gimme a Break: Identification and Estimation of the Macroeconomic Effects of Monetary Policy Shocks in the US. Macroeconomic Dynamics .

Bacchiocchi, E., Fanelli, L., 2015. Identification in Structural Vector Autoregressive Models with Structural Changes, with an Application to US Monetary Policy. Oxford Bulletin of Economics and Statistics 77, 761–779.

Baumeister, C., Benati, L., 2013. Unconventional Monetary Policy and the Great Recession: Estimating the Macroeconomic Effects of a Spread Compression at the Zero Lower Bound. International Journal of Central Banking 9, 165–212.

Belongia, M.T., Ireland, P.N., 2015. Interest Rates and Money in the Measurement of Monetary Policy. Journal of Business & Economic Statistics 33, 255–269.

Bernanke, B.S., Blinder, A.S., 1992. The Federal Funds Rate and the Channels of Monetary Transmission. American Economic Review 82, 901–921.

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

Boivin, J., Kiley, M.T., Mishkin, F.S., 2010. Chapter 8 - how has the monetary transmission mechanism evolved over time?, Elsevier. volume 3 of *Handbook of Monetary Economics*, pp. 369 – 422.

Caldara, D., Herbst, E., 2019. Monetary Policy, Real Activity, and Credit Spreads: Evidence from Bayesian Proxy SVARs. American Economic Journal: Macroeconomics 11, 157–92. doi:10.1257/mac.20170294.

Castelnuovo, E., Nisticò, S., 2010. Stock market conditions and monetary policy in a DSGE model for the U.S. Journal of Economic Dynamics and Control 34, 1700–1731. doi:https://doi.org/10.1016/j.jedc.2010.06.028.

Chan, J., Eisenstat, E., 2018. Bayesian model comparison for time-varying parameter VARs with stochastic volatility. Journal of Applied Econometrics .

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

Chan, J.C., Koop, G., Yu, X., 2022. Large Order-Invariant Bayesian VARs with Stochastic Volatility.

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

Chib, S., 1996. Calculating Posterior Distributions and Modal Estimates in Markov Mixture Models. Journal of Econometrics 75, 79–7.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 1999. Monetary Policy Shocks: What Have We Learned and to What End?, in: Taylor, J.B., Woodford, M. (Eds.), Handbook of Macroeconomics. Elsevier. chapter Chapter 1, pp. 65–148.

Christiano, L.J., Eichenbaum, M., Evans, C.L., 2005. Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. Journal of Political Economy 113, 1–45.

Clark, T.E., 2011. Real-time density forecasts from Bayesian vector autoregressions with stochastic volatility. Journal of Business and Economic Statistics 29, 327–341.

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

Cogley, T., Sargent, T.J., 2005. Drifts and volatilities: Monetary policies and outcomes in the post WWII US.

Review of Economic Dynamics 8, 262–302.

Curdia, V., Woodford, M., 2010. Credit Spreads and Monetary Policy. Journal of Money, Credit and Banking 42, 3–35.

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

D'Amico, S., King, T.B., 2013. Flow and stock effects of large-scale treasury purchases: Evidence on the importance of local supply. Journal of Financial Economics 108, 425–448.

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

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

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

Feldkircher, M., Huber, F., 2018. Unconventional U.S. Monetary Policy: New Tools, Same Channels? JRFM 11, 1–31.

Frühwirth-Schnatter, S., 2006. Finite Mixture and Markov Switching Models. Springer Science+Business Media, LLC, New York.

Fuhrer, J., Tootell, G., 2008. Eyes on the prize: How did the fed respond to the stock market? Journal of Monetary Economics 55, 796–805.

Gagnon, J., Raskin, M., Remache, J., Sack, B., 2011. The Financial Market Effects of the Federal Reserve's Large-Scale Asset Purchases. International Journal of Central Banking 7, 3–43.

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

George, E.I., Mcculloch, R.E., 1997. Approaches for bayesian variable selection. Statistica Sinica 7, 339–373.

George, E.I., Sun, D., Ni, S., 2008. Bayesian Stochastic Search for VAR Model Restrictions. Journal of Econometrics 142, 553–580.

Gertler, M., Karadi, P., 2011. A model of unconventional monetary policy. Journal of Monetary Economics 58, 17 – 34.

Gertler, M., Karadi, P., 2013. QE 1 vs. 2 vs. 3. . . : A Framework for Analyzing Large-Scale Asset Purchases as a Monetary Policy Tool. International Journal of Central Banking 9, 5–53.

Geweke, J., 1996. Variable selection and model comparison in regression. In Bayesian Statistics 5 .

Giannone, D., Lenza, M., Primiceri, G.E., 2015. Prior selection for vector autoregressions. Review of Economics and Statistics 97, 436–451.

Gilchrist, S., Zakrajsek, E., 2012. Credit Spreads and Business Cycle Fluctuations. American Economic Review 102, 1692–1720.

Hördahl, P., Remolona, E.M., Valente, G., 2020. Expectations and Risk Premia at 8:30 a.m.: Deciphering the Responses of Bond Yields to Macroeconomic Announcements. Journal of Business & Economic Statistics 38, 27–42. doi:10.1080/07350015.2018.1429278.

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

Hubrich, K., Tetlow, R.J., 2015. Financial stress and economic dynamics: The transmission of crises. Journal of Monetary Economics 70, 100 – 115.

Jackson, L.E., Owyang, M.T., Soques, D., 2018. Nonlinearities, smoothing and countercyclical monetary policy. Journal of Economic Dynamics and Control 95, 136–154.

Jawadi, F., Sousa, R.M., Traverso, R., 2017. On the macroeconomic and wealth effects of unconventional monetary policy. Macroeconomic Dynamics 21, 1189–1204.

Kastner, G., Frühwirth-Schnatter, S., 2014. Ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc estimation of stochastic volatility models. Computational Statistics & Data Analysis 76, 408 – 423. doi:https://doi.org/10.1016/j.csda.2013.01.002. cFEnetwork: The Annals of Computational and Financial Econometrics.

Kimura, T., Nakajima, J., 2016. Identifying conventional and unconventional monetary policy shocks: a latent threshold approach. The B.E. Journal of Macroeconomics 16, 277–300.

Koop, G., Korobilis, D., 2016. Model uncertainty in Panel Vector Autoregressive models. European Economic Review 81, 115–131.

Koop, G., León-González, R., Strachan, R.W., 2009. On the Evolution of the Monetary Policy Transmission Mechanism. Journal of Economic Dynamics and Control 33, 997–1017.

Krishnamurthy, A., Vissing-Jorgensen, A., 2011. The Effects of Quantitative Easing on Interest Rates: Channels and Implications for Policy. Brookings Papers on Economic Activity 42, 215–287.

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

Leeper, E.M., Zha, T., 2003. Modest policy interventions. Journal of Monetary Economics 50, 1673–1700.

Lengnick, M., Wohltmann, H.W., 2016. Optimal monetary policy in a new Keynesian model with animal spirits and financial markets. Journal of Economic Dynamics and Control 64, 148–165.

Liu, P., Mumtaz, H., Theodoridis, K., Zanetti, F., 2017. Changing Macroeconomic Dynamics at the Zero Lower Bound. Journal of Business & Economic Statistics 0, 0–0. arXiv:https://doi.org/10.1080/07350015.2017.1350186.

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

Lütkepohl, H., Netšunajev, A., 2017. Structural vector autoregressions with smooth transition in variances.

Journal of Economic Dynamics and Control 84, 43–57.

Lütkepohl, H., Shang, F., Uzeda, L., Woźniak, T., 2022a. Partial identification of heteroskedastic structural vars: Theory and bayesian inference.

Lütkepohl, H., Shang, F., Uzeda, L., Woźniak, T., 2022b. Partial identification of heteroskedastic structural vars: Theory and bayesian inference.

Lütkepohl, H., Woźniak, T., 2017. Bayesian Inference for Structural Vector Autoregressions with Markov-Switching Heteroskedasticity.

Malsiner-Walli, G., Wagner, H., 2018. Comparing spike and slab priors for bayesian variable selection. arXiv preprint arXiv:1812.07259 .

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

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

Milunovich, G., Yang, M., 2013. On Identifying Structural VAR Models via ARCH Effects. Journal of Time Series Econometrics 5, 117–131.

Nimark, K., 2008. Monetary policy with signal extraction from the bond market. Journal of Monetary Economics 55, 1389–1400. doi:https://doi.org/10.1016/j.jmoneco.2008.09.004.

Normandin, M., Phaneuf, L., 2004. Monetary policy shocks : Testing identification conditions under time-varying conditional volatility. Journal of Monetary Economics 51, 1217–1243.

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

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

Primiceri, G.E., 2005. Time Varying Structural Vector Autoregressions and Monetary Policy. The Review of Economic Studies 72, 821–852.

R Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria. URL: https://www.R-project.org/.

Rigobon, R., 2003. Identification Through Heteroskedasticity. The Review of Economics and Statistics 85, 777–792.

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

Rudebusch, G.D., Wu, T., 2008. A Macro-Finance Model of the Term Structure, Monetary Policy and the Economy. The Economic Journal 118, 906–926. doi:10.1111/j.1468-0297.2008.02155.x.

Sims, C.A., 1986. Are forecasting models usable for policy analysis? Quarterly Review 10, 2–16.

Sims, C.A., 1992. Interpreting the macroeconomic time series facts : The effects of monetary policy. European

Economic Review 36, 975–1000.

Sims, C.A., Waggoner, D.F., Zha, T., 2008. Methods for inference in large multiple-equation Markov-switching models. Journal of Econometrics 146, 255–274.

Sims, C.A., Zha, T., 2006. Were There Regime Switches in U.S. Monetary Policy? American Economic Review 96, 54–81.

Taylor, J.B., 1993. Discretion versus policy rules in practice. Carnegie-Rochester Conference Series on Public Policy 39, 195–214.

Tillmann, P., 2020. Monetary Policy Uncertainty and the Response of the Yield Curve to Policy Shocks. Journal of Money, Credit and Banking 52, 803–833. doi:https://doi.org/10.1111/jmcb.12657.

Verdinelli, I., Wasserman, L., 1995. Computing bayes factors using a generalization of the savage-dickey density ratio. Journal of the American Statistical Association 90, 614–618.

Vázquez, J., María-Dolores, R., Londoño, J.M., 2013. On the informational role of term structure in the US monetary policy rule. Journal of Economic Dynamics and Control 37, 1852–1871. URL: https://www.sciencedirect.com/science/article/pii/S0165188913000705, doi:https://doi.org/10.1016/j.jedc.2013.04.002.

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

Walentin, K., 2014. Business cycle implications of mortgage spreads. Journal of Monetary Economics 67, 62–77.

Woźniak, T., 2022. bsvars: Bayesian Estimation of Structural Vector Autoregressive Models. URL: https://cran.r-project.org/package=bsvars. r package version 1.0.1.9000.

Wu, J.C., Xia, F.D., 2016. Measuring the macroeconomic impact of monetary policy at the zero lower bound. Journal of Money, Credit and Banking 48, 253–291. arXiv:https://onlinelibrary.wiley.com/doi/pdf/10.1111/jmcb.12300.

Appendix A. Gibbs sampler for Autoregressive and Stochastic Volatility Parameters

Estimation of the structural matrix, the associated S^5 component, and the corresponding prior shrinkage parameter was presented in Section 4.2. In this appendix, we scrutinise the estimation procedure for the remaining parameters including the autoregressive matrix and its prior shrinkage parameter, and the components of the Stochastic Volatility model. 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 in terms of efficiency, mixing, and computational speed even for larger systems of variables.

Sampling autoregressive parameters

We follow Chan et al. (2022) and sample the autoregressive parameters of the $\mathbf{A}$ matrix row-by-row, where the n^{th} row is denoted by $[\mathbf{A}]_{n\cdot}$. Denote by $\mathbf{A}_{n=0}$ the $\mathbf{A}$ matrix with its n^{th} row set to zeros. To derive the sampler for the row of the autoregressive matrix, rewrite the structural form equation (3) in an equivalent form as:

$$\mathbf{B}\left(s_t, \kappa(s_t)\right)\left(\mathbf{y}_t - \mathbf{A}_{n=0}\mathbf{x}_t\right) = \left([\mathbf{B}\left(s_t, \kappa(s_t)\right)]_{\cdot n} \otimes \mathbf{x}_t'\right)[\mathbf{A}]_{n\cdot}' + \boldsymbol{\varepsilon}_t \tag{A.1}$$

Define an $n \times 1$ vector $\mathbf{z}_t^n = \mathbf{B}\left(s_t, \kappa(s_t)\right)\left(\mathbf{y}_t - \mathbf{A}_{n=0}\mathbf{x}_t\right)$ and an $N \times (Np + d)$ matrix $\mathbf{Z}_t^n = \left([\mathbf{B}\left(s_t, \kappa(s_t)\right)]_{\cdot n} \otimes \mathbf{x}_t'\right)$ and rewrite equation (A.1) as:

$$\mathbf{z}_t^n = \mathbf{Z}_t^n[\mathbf{A}]_{n\cdot}' + \boldsymbol{\varepsilon}_t, \tag{A.2}$$

where the structural shocks follow the normal distribution defined in equation (6). This form of a model combined with the conditionally normal prior distribution results in the multivariate normal full conditional posterior distribution for the rows of $\mathbf{A}$ given by:

$$[\mathbf{A}]'_{n\cdot} \mid \mathbf{Y}_T, \mathbf{A}_{n=0}, \mathbf{B}\left(s_t, \kappa(s_t)\right), \sigma_1^2, \dots \sigma_T^2, \gamma_A \sim \mathcal{N}_{Np+d}\left(\overline{\mathbf{\Omega}}_{A.n}\overline{\mathbf{m}}'_{A.n}, \overline{\mathbf{\Omega}}_{A.n}\right) \tag{A.3}$$

$$\overline{\mathbf{\Omega}}_{A.n}^{-1} = \gamma_A^{-1}\underline{\mathbf{\Omega}}_A^{-1} + \sum_t \mathbf{Z}_t^{n\prime} \operatorname{diag}\left(\sigma_t^2\right)^{-1} \mathbf{Z}_t^n \tag{A.4}$$

$$\overline{\mathbf{m}}'_{A.n} = \gamma_A^{-1}\underline{\mathbf{\Omega}}_A^{-1}\underline{\mathbf{m}}'_{A.n} + \sum_t \mathbf{Z}_t^{n\prime} \operatorname{diag}\left(\sigma_t^2\right)^{-1} \mathbf{z}_t^n \tag{A.5}$$

The sampler maintains all the computational advantages that Chan et al. (2022) enumerates. The overall autoregressive shrinkage hyper-parameter is sampled from its inverted gamma 2 full conditional posterior distribution:

$$\gamma_A \mid \mathbf{A} \sim \mathcal{IG}2\left(\underline{s}_A + \sum_n \left([\mathbf{A}]_{n\cdot} - \underline{\mathbf{m}}_{A.n}\right)\underline{\mathbf{\Omega}}_A^{-1}\left([\mathbf{A}]_{n\cdot} - \underline{\mathbf{m}}_{A.n}\right)', \quad \underline{\nu}_A + N(Np+d)\right) \tag{A.6}$$

Sampling stochastic volatility parameters

Gibbs sampler for the parameters of the SV processes results from our prior assumptions described in Section 5.2 and the assumption regarding the normality of the structural shocks in equation (6). It is facilitated using the auxiliary mixture sampler proposed by Omori, Chib, Shephard and Nakajima (2007b).

Specify the n^{th} structural shock as:

$$u_{n.t} = \exp\left\{\frac{1}{2}\omega_n(s_t)h_{n.t}\right\} z_{n.t}, \tag{A.7}$$

where $z_{n.t}$ is a standard normal innovation. Transform this equation by squaring and taking the logarithm of both sides obtaining:

$$\tilde{u}_{n.t} = \omega_n(s_t)h_{n.t} + \tilde{z}_{n.t}, \tag{A.8}$$

where $\tilde{u}_{n.t} = \log u_{n.t}^2$ and $\tilde{z}_{n.t} = \log z_{n.t}^2$. The distribution of $\tilde{z}_{n.t}$ is $\log \chi_1^2$. This non-standard distribution is approximated precisely by a mixture of ten normal distributions defined by Omori et al. (2007b). Applying the auxiliary mixture technique makes the linear equation (A.8) conditionally normal given the mixture component indicators, which greatly simplifies the sampling algorithm. This mixture of normals is specified by $q_{n.t} = 1, \dots, 10$ – the mixture component indicator for the n^{th} equation at time t, the normal component probability $\pi_{q_{n.t}}$, mean $\mu_{q_{n.t}}$, and variance $\sigma^2_{q_{n.t}}$. The latter three parameters are fixed and given in Omori et al. (2007b), while $q_{n.t}$ augments the parameter space and is estimated. Its prior distribution is multinomial with probabilities $\pi_{q_{n.t}}$.

Define $T\times 1$ vectors: $\mathbf{h}_n = \begin{bmatrix} h_{n.1} & \dots & h_{n.T} \end{bmatrix}'$ collecting the log-volatilities, $\mathbf{q}_n = \begin{bmatrix} q_{n.1} & \dots & q_{n.T} \end{bmatrix}'$ collecting the realizations of $q_{n.t}$ for all t, $\boldsymbol{\mu}_{\mathbf{q}_n} = \begin{bmatrix} \mu_{q_{n.1}} & \dots & \mu_{q_{n.T}} \end{bmatrix}'$, and $\boldsymbol{\sigma}^2_{\mathbf{q}_n} = \begin{bmatrix} \sigma^2_{q_{n.1}} & \dots & \sigma^2_{q_{n.T}} \end{bmatrix}'$ collecting the n^{th} equation auxiliary mixture means and variances, $\widetilde{\mathbf{U}}_n = \begin{bmatrix} \widetilde{u}_{n.1} & \dots & \widetilde{u}_{n.T} \end{bmatrix}'$, and $\boldsymbol{\omega}_n = \begin{bmatrix} \omega_n(s_1) & \dots & \omega_n(s_T) \end{bmatrix}'$ collecting the volatility of the volatility parameters according to their current time assignment based on the sampled realizations of the Markov process. Whenever a subscript on these vectors is extended by the Markov process' regime indicator m it means that the vector contains only the T_m observations for t such that $s_t = m$, e.g., $\widetilde{\mathbf{U}}_{n.m}$. Finally, define a $T \times T$ matrix $\mathbf{H}_{\rho_n}$ with ones on the main diagonal, value $-\rho_n$ on the first subdiagonal, and zeros elsewhere, and a $T_m \times T_m$ matrix $\mathbf{H}_{\rho_n.m}$ with rows and columns selected according to the Markov state allocations such that $s_t = m$.

Sampling latent volatilities $\mathbf{h}_n$ proceeds independently for each n from the following T-variate normal distribution parameterized following Chan and Jeliazkov (2009) in terms of its precision matrix $\overline{\mathbf{V}}^{-1}_{\mathbf{h}_n}$ and location vector $\overline{\mathbf{h}}_n$ as:

$$\mathbf{h}_n \mid \mathbf{Y}_T, \mathbf{s}_n, \mathbf{q}_n, \mathbf{B}, \mathbf{A}, \boldsymbol{\omega}_n, \rho_n \sim \mathcal{N}_T\left(\overline{\mathbf{V}}_{\mathbf{h}_n}\overline{\mathbf{h}}_n, \overline{\mathbf{V}}_{\mathbf{h}_n}\right) \tag{A.9}$$

$$\overline{\mathbf{V}}^{-1}_{\mathbf{h}_n} = \operatorname{diag}\left(\boldsymbol{\omega}^2_n\right)\operatorname{diag}\left(\boldsymbol{\sigma}^{-2}_{\mathbf{q}_n}\right) + \mathbf{H}'_{\rho_n}\mathbf{H}_{\rho_n} \tag{A.10}$$

$$\overline{\mathbf{h}}_n = \operatorname{diag}\left(\boldsymbol{\omega}_n\right)\operatorname{diag}\left(\boldsymbol{\sigma}^{-2}_{\mathbf{q}_n}\right)\left(\widetilde{\mathbf{u}}_n - \boldsymbol{\mu}_{\mathbf{q}_n}\right) \tag{A.11}$$

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 et al. (2011) is implemented.

The regime-dependent volatility of the volatility parameters that are essential for the assessment of identification of the SVAR models, $\omega_n(m)$, are sampled independently from the following normal distribution:

$$\omega_n(m) \mid \mathbf{Y}_m, \mathbf{s}_{n.m}, \mathbf{s}_{n.m}, \mathbf{h}_{n.m}, \sigma^2_{\omega_n} \sim \mathcal{N}\left(\overline{v}_{\omega_{n.m}}\overline{\omega}_{n.m}, \overline{v}_{\omega_{n.m}}\right) \tag{A.12}$$

$$\overline{v}^{-1}_{\omega_{n.m}} = \mathbf{h}'_{n.m}\operatorname{diag}\left(\sigma^{-2}_{\mathbf{q}_n.m}\right)\mathbf{h}_{n.m} + \sigma^{-2}_{\omega_n} \tag{A.13}$$

$$\overline{\omega}_n = \mathbf{h}'_{n.m}\operatorname{diag}\left(\sigma^{-2}_{\mathbf{q}_n.m}\right)\left(\widetilde{\mathbf{u}}_{n.m} - \boldsymbol{\mu}_{\mathbf{q}_n.m}\right) \tag{A.14}$$

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

$$\rho_n \mid \mathbf{Y}_T, \mathbf{h}_n \sim \mathcal{N}\left(\left(\sum_{t=0}^{T-1} h^2_{n.t}\right)^{-1}\left(\sum_{t=1}^{T} h_{n.t}h_{n.t-1}\right), \left(\sum_{t=0}^{T-1} h^2_{n.t}\right)^{-1}\right)\mathcal{I}\left(|\rho_n| < 1\right). \tag{A.15}$$

using the algorithm proposed by Robert (1995).

The prior variances of parameters $\omega_n(m)$, $\sigma^2_{\omega_n}$, are *a posteriori* independent and sampled from the following generalized inverse Gaussian full conditional posterior distribution:

$$\sigma^2_{\omega_n} \mid \mathbf{Y}_T, \omega_n(1), \dots, \omega_n(M) \sim \mathcal{GIG}\left(M\underline{A} - \frac{1}{2}, \sum_m \omega^2_n(m), \frac{2}{\underline{S}}\right) \tag{A.16}$$

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

Finally, proceed to the ancillarity-sufficiency interweaving sampler proposed by Kastner and Frühwirth-Schnatter (2014). They show that sampling directly the parameters of the centered 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-centered SV parameterization 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 ancillarity-sufficiency interweaving step in the Gibbs sampler. Our implementation proceeds as follows: Having sampled random vector $\mathbf{h}_n$ and parameters $\omega_n(m)$, compute the parameters of the centered parameterization $\tilde{h}_{n.t} = \omega_n(m)h_{n.t}$ and $\sigma^2_{\upsilon_n} = \omega^2_n(m)$. Then, sample $\sigma^2_{\upsilon_n.m}$ from the generalised inverse Gaussian full conditional posterior distribution:

$$\sigma^2_{\upsilon_n.m} \mid \mathbf{Y}_m, \tilde{\mathbf{h}}_{n.m}, \sigma^2_{\omega_n} \sim \mathcal{GIG}\left(-\frac{T_m - 1}{2}, \tilde{\mathbf{h}}'_{n.m}\mathbf{H}'_{\rho_n.m}\mathbf{H}_{\rho_n.m}\tilde{\mathbf{h}}_{n.m}, \sigma^{-2}_{\omega_n}\right) \tag{A.17}$$

using the algorithm introduced by Hörmann and Leydold (2014). Resample ρ_n using the full conditional posterior distribution from equation (A.15) where the vector $\mathbf{h}_n$ is replaced by $\tilde{\mathbf{h}}_n$. Finally, compute $\omega_n(m) = \pm\sqrt{\sigma^2_{\upsilon_n.m}}$ and $h_{n.t} = \frac{1}{\omega_n(m)}\tilde{h}_{n.t}$ and return them as the MCMC draws for these parameters.

Appendix B. Data

ip	log industrial production (INDPRO), Total Index, Index 2017=100, Seasonally Adjusted, FRED database
cpi	log consumer price index (CPALTT01USM661S), Total All Items, Index 2015=100, Seasonally Adjusted, FRED database
R	shadow rates of Wu and Xia (2016) and before January 1990 federal funds rate (DFF), FRED database
TS	$ts = r_L - r$, 10-year Treasury constant maturity rate minus federal funds rate (T10YFFM), FRED database
m	log M2 (M2SL), Billions of Dollars, Seasonally Adjusted, FRED database
CS	Moody's seasoned Baa corporate bond yield relative to yield on 10-year treasury constant maturity (BAA10YM), FRED database
sp	log S&P500 stock price index, monthly average of closing price, yahoo finance
LR	long-term 10-year government bond yields (IRLTLT01USM156N), FRED database