DEPARTMENT OF ECONOMICS WORKING PAPER SERIES

The multivariate simultaneous unobserved components model and identification via heteroskedasticity

Mengheng Li
Ivan Mendieta-Muñoz

Working Paper No: 2019-06

May 2019

University of Utah
Department of Economics
260 Central Campus Drive
Gardner Commons, RM 4100
Tel: (801) 581-7481
http://www.econ.utah.edu

The multivariate simultaneous unobserved components model and identification via heteroskedasticity

Mengheng Li
Economics Discipline Group, University of Technology Sydney.
Email: mengheng.li@uts.edu.au

Ivan Mendieta-Muñoz
Department of Economics, University of Utah.
Email: ivan.mendietamunoz@utah.edu

Abstract

We propose a multivariate simultaneous unobserved components framework to determine the two-sided interactions between structural trend and cycle innovations. We relax the standard assumption in unobserved components models that trends are only driven by permanent shocks and cycles are only driven by transitory shocks by considering the possible spillover effects between structural innovations. The direction of spillover has a structural interpretation, whose identification is achieved via heteroskedasticity. We provide identifiability conditions and develop an efficient Bayesian MCMC procedure for estimation. Empirical implementations for both Okun's law and the Phillips curve show evidence of significant spillovers between trend and cycle components.

Keywords: Unobserved components, Identification via heteroskedasticity, Trends and cycles, Permanent and transitory shocks, State space models, Spillover structural effects.

JEL Classification: C11, C32, E31, E32, E52.

The multivariate simultaneous unobserved components model and identification via heteroskedasticity

Mengheng Li* Ivan Mendieta-Muñoz†

May 2019

Abstract

We propose a multivariate simultaneous unobserved components framework to determine the two-sided interactions between structural trend and cycle innovations. We relax the standard assumption in unobserved components models that trends are only driven by permanent shocks and cycles are only driven by transitory shocks by considering the possible spillover effects between structural innovations. The direction of spillover has a structural interpretation, whose identification is achieved via heteroskedasticity. We provide identifiability conditions and develop an efficient Bayesian MCMC procedure for estimation. Empirical implementations for both Okun's law and the Phillips curve show evidence of significant spillovers between trend and cycle components.

Keywords: *Unobserved components, Identification via heteroskedasticity, Trends and cycles, Permanent and transitory shocks, State space models, Spillover structural effects.*
JEL Classification: C11; C32; E31; E32; E52.

*Business School, University of Technology Sydney. Mailing address: Office CB08.09.016, Building 8, Dr. Chau Chak Wing Building, 14–28 Ultimo Road, Ultimo, NSW 2007, Sydney, Australia. Email: mengheng.li@uts.edu.au

†Department of Economics, University of Utah. Mailing address: Office 4230, Suite 4100, 260 Central Campus Drive, Gardner Commons, Salt Lake City, Utah 84112, USA. Email: ivan.mendietamunoz@utah.edu

1 Introduction

In structural time series analysis, unobserved components (UC) models decompose a (vector) time series into a trend and cycle component driven by permanent and transitory shocks, respectively (see Harvey and Shephard, 1993; Durbin and Koopman, 2012). Basic UC models assume no correlation between the permanent trend and transitory cycle innovations, thus neglecting the possible interactions among shocks *a priori*. Different studies have shown that this assumption does not hold empirically. Morley et al. (2003) introduces the correlated UC model which allows for correlation between permanent and transitory shocks, showing (i) that the zero-correlation restriction between permanent and transitory shocks can be rejected and (ii) that trend innovations are strongly negatively correlated with cycle innovations in their application to U.S. output. Among others, univariate applications of correlated UC models include Dungey et al. (2015) who document correlation between trend output and output gap and discuss its implications for determining the nature of recessions; and Hwu and Kim (2019) who show that the incorporation of correlation between trend inflation and inflation gap leads to improved forecasting performance. Regarding multivariate applications, Basistha (2007) estimates inflation and output trends that are correlated with cyclical fluctuations for Canada; Morley (2007) finds correlation between permanent and transitory movements in aggregate income and consumption for the U.S.; and Sinclair (2009) finds evidence suggesting both within-series and cross-series trend-cycle correlations. Further, Mitra and Sinclair (2012) model correlation between all the contemporaneous trend and cycle shocks using output for the G-7 countries.

However, as discussed by Morley (2007) and Weber (2011), when shocks are correlated, the UC model is no longer structural because, similar to a reduced-form VAR model with non-diagonal covariance matrix, further restrictions are needed in order to recover the structural shocks. Theoretically, it is possible to achieve identification by assuming that cycle shocks (commonly associated with aggregate demand fluctuations, monetary policies, etc.) are neutral in the long-run, an assumption consistent with the seminal work of Blanchard and Quah (1989) and in line with different interpretations of causality running from trend to cycle.[1] Empirically, however, a decision between the two potential

[1]Nevertheless, it is worth mentioning that various economic theories also permit different types of long-run non-neutrality at the theoretical level, see Keating (2013a,b).

directions of causality crucially depends on the ability to identify the possible simultaneous effects from the data. As Fisher et al. (2016) point out regarding the interpretation and identification in structural VARs, whether shocks are permanent or transitory is at large a subjective choice made by econometricians.[2]

The present paper introduces the so-called multivariate simultaneous unobserved components (MSUC) model in order to study the interactions between permanent and transitory shocks. The proposed framework extends the extant literature in three ways. First, we show that the model maintains the identifiability of the structural form without imposing any kind of short-run or long-run restrictions on the contemporaneous structural matrix.[3] Thus, besides within-series and cross-series trend-cycle correlations, we identify the *direction of correlation* by tracing back the correlation of reduced-form residuals to structural spillovers of *both* trend shocks to the cycle component *and* of cycle shocks to the trend component. Our contribution is related to that of Weber (2011), who derived the spillovers by over-identifying a correlated UC model in a univariate analysis for US industrial production. With respect to the latter, our research: (i) identifies the possible spillovers considering a multivariate framework; and (ii) allows for a more general identification strategy because it does not require the strong assumption that all reduced-form shocks have unit variance, which may lead to biased results if there exists heteroskedasticity, thereby providing a clearer interpretation of the identified structural shocks.

Second, the identification of the proposed MSUC model is achieved by explicitly considering heteroskedasticity, an important feature of macroeconomic time series, as documented by the literature on the "Great Moderation" and by the recent literature on structural VARs with heteroskedasticity.[4] Lanne et al. (2010) and Herwartz and Lütkepohl (2014) show that it is possible to identify the structural matrix if there exists enough variation in volatility by assuming a Markov regime switching for the variances

[2]For example, Keating (2013a,b) also shows that the structural interpretation of the Blanchard and Quah (1989) decomposition (in which transitory and permanent shocks are identified by assuming the long-run restriction that only permanent shocks affect the level of output) changes if transitory shocks are assumed to affect the price level in a different manner.

[3]In structural VARs (see Sims, 1980, Amisano and Giannini, 2012, among many others), it is usually believed that the correlations of reduced-form shocks are derived from the contemporaneous effects between variables, and this interpretation is also used in UC models.

[4]In brief, it is accepted that macroeconomic volatility has experienced a major reduction since the mid-1980s. The "Great Moderation" continued until the Great Financial Crisis in 2007, when volatility increased again, although it seems to have decreased again in recent years.

of structural shocks (SVAR-MRS). Netsunajev (2013), Velinov and Chen (2015) and Lütkepohl and Netšunajev (2017) (among others) document successful implementations of this approach. Following this growing strand of literature, we assume that the change in the covariance matrix of reduced-form shocks in the context of UC models results from heteroskedasticity, which allows us to identify the complete spillover effects without imposing further restrictions on the structural matrix.

Third, we develop a Bayesian sampling scheme that directly generates posterior draws for the structural parameters, similar to Wozniak and Droumaguetb (2015) and Lütkepohl and Wozniak (2017) in SVAR-MRS models. Because the number of parameters and model flexibility grows exponentially, Bayesian methods for estimating SVAR-MRS have been attempted. For example, Kulikov and Netšunajev (2013) and Lanne et al. (2016) develop methods for sampling the reduced-form parameters from their posterior distribution and transform draws into structural parameters; therefore, these methods generate posterior distributions of exactly-identified structural parameters, which limits the model's applicability if over-identifying restrictions are of interest. Although different from this literature, we are able to show the identifiability of MSUC models in two steps: (i) the reduced-form model, *i.e.* the correlated UC model, is identifiable based on Trenker and Weber (2016); and (ii) as in Lanne et al. (2010) and Lütkepohl and Wozniak (2017), similar conditions are given to ensure the identifiability of the structural matrix.

Lütkepohl and Wozniak (2017) argue that the identification of structural VARs via heteroskedasticity can be understood as unrestricted since samples are generated directly for structural parameters, thus providing a basis for Bayesian tests for over-identification restrictions. Although this argument also applies to our setting, it should be noted that the Markov chain Monte Carlo (MCMC) procedure developed in the literature of SVAR-MRS models falls short because of the following reasons: (i) the "dependent variables" in MSUC models (*i.e.* the trend and cycle components) are not observed, so they need to be inferred from the available data; (ii) trends (which are by definition non-stationary) appear in the reduced-form model together with stationary cycle components; (iii) unlike Bayesian VARs, where one can simplify the sampling procedure by conditioning on initial observations, initialisations of UCs still need to be inferred from data; and (iv) the volatility of structural shocks corresponds to different

components in MSUC models, so it can be considered as permanent and transitory volatility (Stock and Watson, 2007; Shephard, 2015; Li and Koopman, 2018), while a Markov regime switching imposes abrupt regime switches with identical timing, which is arguably a questionable assumption.

The efficient MCMC algorithm developed in this paper deals jointly with the first and second points above by sampling the trends and cycles component-by-component and the contemporaneous structural matrix row-by-row recursively. This algorithm uses the fact that drawing from the joint posterior of UCs and the structural matrix can be facilitated by an efficient Metropolis-Hastings (MH) within Gibbs sampler. One novelty of the proposed method is that effectively we only apply univariate MH algorithms for one element in each row of the structural matrix that does not have a standard posterior distribution, thereby reducing the dimensionality problem in SVAR-MRS models into a simple univariate problem and maintaining a very high acceptance rate. The third concern is solved by using state space methods that impute initial values from their conditional posterior distributions to the reduced-form model implied by the MSUC model. Finally, because we consider both Markov regime switching volatility and stochastic volatility (where the volatility of structural shocks changes at each point in time), our method allows for higher flexibility when modelling permanent and transitory volatilities.

The remainder of the paper is organised as follows. Section 2 introduces the MSUC model and relates the latter to some relevant examples in the literature in order to motivate the model specification. Section 3 provides the identifiability conditions for the MSUC model and discusses the new MCMC algorithm for Bayesian inference. Section 4 conducts an empirical study on the interactions between permanent and transitory shocks by considering MSUC versions of Okun's law and the Phillips curve for the US economy, finding evidence of significant within-components and cross-components spillovers. Section 5 discusses some relevant policy implications derived from the empirical findings. Finally, we conclude in Section 6.

2 Multivariate simultaneous unobserved components

As mentioned in the previous section, the correlation between innovations can be attributed to the contemporaneous effects between unobserved components, as in SVAR

models. We propose the following state space representation of a MSUC model:

$$y_t = Z\tau_t + c_t, \tag{1}$$

$$A \begin{pmatrix} \tau_{t+1} \\ c_{t+1} \end{pmatrix} = \Phi(L_p) \begin{pmatrix} \tau_t \\ c_t \end{pmatrix} + e_t, \quad e_t \sim N(0, \Sigma), \tag{2}$$

for $t = 1, ..., T$, where $y_t \in \mathbb{R}^N$ is the vector of observations; $\tau_t \in \mathbb{R}^k$ contains non-stationary components; $c_t \in \mathbb{R}^N$ collects stationary cycles; and Z is an $N \times k$ selection matrix such that $Z\tau_t$ is the trend of y_t.[5] Notice that if y_t only has a vector random walk trend, then $k = N$ and Z becomes an identity matrix. Importantly, A is a matrix with ones on its diagonal and $\Phi(L_p) = \Phi_1 L + ... + \Phi_p L^p$ is a p-th order lag polynomial with $(N + k)$-dimensional coefficient matrices Φ_i, $i = 1, ..., p$ that captures the transition dynamics between lagged components. Moreover, $e_t \in \mathbb{R}^{N+k}$ is a vector of uncorrelated structural errors, *i.e.* $\Sigma = \text{diag}(\sigma_1^2, ..., \sigma_{N+k}^2)$, consisting of permanent shocks that drive the long-run or trend component τ_t and transitory shocks that drive the cycles c_t.

The state transition (2) can be written in a more familiar form

$$\begin{pmatrix} \tau_{t+1} \\ c_{t+1} \end{pmatrix} = B \begin{pmatrix} \tau_t \\ c_t \end{pmatrix} + \epsilon_t, \quad \epsilon_t = A^{-1} e_t \sim N(0, \Omega), \tag{3}$$

where $\Omega = A^{-1}\Sigma A^{-1'}$ and

$$B = \begin{bmatrix} B_\tau & 0 \\ 0 & B_c(L_p) \end{bmatrix} = A^{-1}\Phi(L_p). \tag{4}$$

It is important to clarify two points. First, the model represented by equations (1) and (2) and the one shown in (1) and (3) are observationally equivalent. Although the latter is similar to the structural time series models that use unobserved components (see Harvey and Shephard (1993) and Durbin and Koopman (2012, Chapter 2)), it cannot be considered as "structural" because the elements contained in ϵ_t are correlated and, thus, they represent reduced-form shocks. Most of the studies reviewed in Section 1 estimate

[5] There is ample evidence that permanent shocks are ubiquitous as many economic time series exhibit non-stationary behaviour. For example, a close examination of over 130 aggregate measures of economic variables as in Stock and Watson (2005) and McCracken and Ng (2016) reveals that most of the time series are integrated, and many of them are integrated of order larger than one. This means that a significant part of the driving force behind economic variables is of permanent nature.

the model given by equations (1) and (3) using Kalman filter and maximum likelihood.

Second, if y_t has a vector random walk trend, B_τ is an identity matrix; but B_τ can also be specified to model other non-stationary dynamics or to incorporate structural relationships between trends. For example, the Phillips curve unobserved components model proposed by Harvey (2011), where the vector trend component $\tau_t = (\tau_{1,t}, \tau_{g,t}, \tau_{2,t})'$ consists of the output trend $\tau_{1,t}$, which is an integrated random walk of order 2, indicating that the growth rate of trend output $\tau_{g,t}$ itself is a random walk (see also Stock and Watson, 1998), while the trend inflation $\tau_{2,t}$ is a random walk. Another example is the modelling of the IS curve in Laubach and Williams (2003) and Holston et al. (2017), who estimate the natural rate of interest as the trend component of real rate driven by the trend output growth rate (see also Galí, 2015b, Chapter 3). This specification states that the first difference of trend output growth and the natural rate of interest are cointegrated.

Apparently, $\Phi(L_p)$ is determined by the structural matrix A and the autoregression coefficient matrix $B_c(L_p)$.[6] Since our main interest consists in testing for spillovers between permanent and transitory shocks, we do not restrict A for identification as in numerous structural VAR models. By contrast, we apply the principles of identification of VAR models via heteroskedasticity discussed in Lanne et al. (2010) and Herwartz and Lütkepohl (2014) to our proposed simultaneous unobserved components model. The idea is straightforward. Suppose that the estimated covariance matrix of the reduced-form innovations $\Omega = A^{-1}\Sigma A^{-1'}$ can adopt two values, say Ω_1 and Ω_2. This gives us $K^2 + K$ distinct parameters in total, where $K = N + k$. If we assume that the relevant changes come only from the volatility of structural shocks, we will have $K^2 - K$ elements in A and $2K$ variances in Σ_1 and Σ_2 to estimate. Therefore, it is possible to provide exact identification of the system. The next section introduces the estimation procedure developed for the proposed MSUC models and for an extension that considers stochastic volatility, which allows to consider that Σ_t changes at each point in time.

[6]Note that it is also possible to estimate the $\Phi(L_p)$ that corresponds to a reduced-form transition matrix B with non-zero upper-right block, *i.e.* assuming that the right-hand side of equation (5) is non-zero. This may indicate that previous cycle components influence the trend. We do not consider this possibility because such effects (if any) are expected to be small as the trend component should be highly persistent with low-frequency movements. In our empirical applications, we did not find statistical evidence suggesting non-zero upper-right block in B, so we only focus on the contemporaneous effect.

3 Identification and estimation via heteroskedasticity

3.1 Identifiability

As discussed by Trenker and Weber (2016), a prerequisite for identifying A is being able to identify Ω as in the reduced-form system (3). From $B = A^{-1}\Phi(L_p)$ we have

$$\begin{aligned}
(A^{-1})_{\tau\tau}\Phi(L_p)_{\tau\tau} + (A^{-1})_{\tau 2}\Phi(L_p)_{c\tau} &= B_\tau, \\
(A^{-1})_{c\tau}\Phi(L_p)_{\tau\tau} + (A^{-1})_{cc}\Phi(L_p)_{2\tau} &= 0, \\
(A^{-1})_{c\tau}\Phi(L_p)_{\tau c} + (A^{-1})_{cc}\Phi(L_p)_{cc} &= B_c(L_p), \\
(A^{-1})_{\tau\tau}\Phi(L_p)_{\tau c} + (A^{-1})_{\tau c}\Phi(L_p)_{cc} &= 0.
\end{aligned} \tag{5}$$

X_{ij}, $i, j = c, \tau$, indicates the ij-th block of matrix X, where the subscripts τ and c stand for the rows or columns corresponding to the trend τ and cycle c, respectively. Using matrix block inversion formula, it can be shown that

$$\begin{aligned}
\Phi(L_p)_{\tau\tau} &= A_{\tau\tau}B_\tau, \quad \Phi(L_p)_{\tau c} = A_{\tau c}B_c(L_p), \\
\Phi(L_p)_{c\tau} &= A_{c\tau}B_\tau, \quad \Phi(L_p)_{cc} = A_{cc}B_c(L_p).
\end{aligned} \tag{6}$$

Since B_τ is constant, to determine $\Phi(L_p)$ is to determine $B_c(L_p)$. It turns out that the latter is key to the identifiability of Ω in the reduced-form system. For simplicity, we assume that the trend component is modelled as a vector random walk, *i.e.* $B_\tau = I_N$; and let the lag polynomial $B_c(L_p)$ be $B_{c,1}L + ... + B_{c,p}L^p$ with the $N \times N$ coefficient matrix $B_{c,i}$, $i = 1, ..., p$. System (1) and (3) leads to a VARIMA$(p, 1, p)$ process for y_t:

$$[I_N - B_c(L_p)]\Delta y_t = [I_N - B_c(L_p)]\epsilon_{\tau,t} + \Delta\epsilon_{c,t} = \Theta(L_p)u_t, \tag{7}$$

where $\Theta(L_p) = I_N + \Theta_1 L + ... + \Theta_p L^p$ is a N-dimensional lag polynomial of order p for the i.i.d. shocks $u_t \sim N(0, \Sigma_u)$.[7] As long as the VAR and VMA polynomials are not mutually cancelling (that is, there exists no common roots to $|I_N - B_c(L_p)| = 0$ and to $|\Theta(L_p)| = 0$) and provided that all roots to $|I_N - B_c(L_p)| = 0$ are outside the unit circle, the VAR

[7] Representing the original process for y_t as a VARIMA process is simply a result of the Granger's lemma (see Lütkepohl, 1984, Athanasopoulos et al., 2016).

components can be uniquely identified from the reduced VARIMA system (Dufour and Pelletier, 2005). If B_τ is not an identity matrix (for example, when an integrated random walk trend is present as in Harvey (2011) and Holston et al. (2017)) it is necessary to take higher order differences to have an identifiable reduced-form VARIMA(p, d, p) with $d \geq 2$.

3.1.1 Identification of trend and cycle covariance matrix

From the reduced-form VARIMA model (7), the identification of $B_c(L_p)$ is straightforward. The identification of Ω is achieved by linking the autocovariance matrices of $[I_N - B_c(L_p)]\epsilon_{\tau,t} + \Delta\epsilon_{c,t}$ to those of $\Theta(L_p)u_t$ (that is, a multivariate Yule-Walker equations procedure). Denote the VMA autocovariance by $\Gamma_s = \sum_{i=0}^{p-s} \Theta_{i+s}\Sigma_u\Theta_i'$, $s = 1, 2, ...$, such that $\Gamma_s = 0$ for $s > p$. The order condition for identification under $p \geq 2$ is satisfied because Γ_0, Γ_1, ..., Γ_p provide $N(N+1)/2 + N^2p$ parameters from which Ω is recovered. Clearly, a relevant rank condition in relation to the system of equations linking the reduced-form and structural matrices needs to be met (see also, Morley et al., 2003). To proceed, we assume the following.

Assumption 1. *In the MSUC model (1) and (3), cycles are stable, i.e. the characteristic polynomial $|I_N - B_c(L_p)|$ has all roots outside the unit circle. Furthermore, besides that $B_{c,1}$ is of full rank, there exists a vector of coefficients $c_i \in \mathbb{R}$, $i = 2, ..., p$, such that $\bar{B} = \sum_{i=1}^{p} c_i B_{c,i}$ is of full rank.*

As pointed out by Trenker and Weber (2016), this assumption is not primitive, which means that estimating the model while restricting identifiability via a rank condition is not feasible. However, one can easily adopt a rule of thumb by simply estimating the model shown in (1) and (3) with varying $p \geq 2$ and checking if a full rank $\bar{B}$ matrix can be found, say by choosing $c_i = 1$ for $i = 1, ..., p$. We have not encountered any problems regarding the rank condition in all the empirical applications.

Let D_N denote the $N^2 \times \frac{1}{2}N(N+1)$ duplication matrix[8] and define the "projector" $D_N^+ = (D_N'D_N)^{-1}D_N'$. Let Ω be partitioned with diagonal blocks Ω_τ and Ω_c and off-diagonal blocks $\Omega_{\tau,c}$ and $\Omega_{\tau,c}'$. The system of equations linking reduced-form

[8] The duplication matrix is such that for any N-dimensional square matrix A, $\text{vec}(A) = D_N\text{vech}(A)$.

autocovariances to structural parameters in Ω can be written as

$$\mathcal{Y} = \mathcal{X}\beta,$$

where $\mathcal{Y} = [\text{vec}(\Gamma_0)'D_N^{+\prime}, \text{vec}(\Gamma_1)', ..., \text{vec}(\Gamma_p)']'$, $\beta = [\text{vech}(\Omega_\tau)', \text{vec}(\Omega_c)', \text{vec}(\Omega_{\tau,c})']'$, and the $(N^2p+\frac{1}{2}N^2+\frac{1}{2}N)\times\frac{1}{2}(5N^2+N)$ matrix $\mathcal{X}$ only involves $B_{c,i}$, $i = 1, ..., p$. Identifiability is stated in the following proposition.

Proposition 1. *Under Assumption 1, rank($\mathcal{X}$)=$\frac{1}{2}(5N^2 + N)$; β and Ω are uniquely determined.*

The proof is given in Appendix A.

3.1.2 Identification of contemporaneous matrix

Identification of the reduced-form innovation covariance matrix Ω does not grant identification of the contemporaneous matrix A because the structural transition matrix $\Phi(L_p)$ in (2) does not necessarily have a block structure.

To recover the structural matrix A from the identified reduced-form system (1) and (3), we assume heteroskedasiticity of the structural shocks e_t.[9] Two forms of heteroskedasticity are considered: Markov regime switching and stochastic volatility. The former assumes two-volatility states Σ_1 and Σ_2 with a transition matrix governing the probability of switching between the two states;[10] whereas the latter assumes volatility changes smoothly at each point in time, namely Σ_t for $i = 1, ..., T$.

As mentioned previously, the minimum requirement to identify A is to have two distinct volatility states. Let $s_t = \{1, 2\}$ for $t = 1, ..., T$ indicate the volatility state, and define $\Sigma_{s_t} = \text{diag}(\sigma^2_{1,s_t}, ..., \sigma^2_{K,s_t})$ and the ratio of variances

$$\omega_{s_t} = \left(\frac{\sigma^2_{1,s_t}}{\sigma^2_{1,1}}, ..., \frac{\sigma^2_{K,s_t}}{\sigma^2_{K,1}}\right)'.$$

So ω_1 is a K-dimensional vector of ones. The definition of variance ratio facilitates the discussion on identification of A. Let $\mathbb{1}.$ denote the indicator function. We make the

[9] See Lanne et al. (2010), Herwartz and Lütkepohl (2014), and Lütkepohl and Wozniak (2017) for examples that use heteroskedasticity to identify contemporaneous effects in VAR analysis.

[10] For notational simplicity, a two-state volatility regime model is considered, but extensions to multiple states is straightforward.

following assumption regarding the volatility states of the reduced-form system.

Assumption 2. *In the MSUC model (1) and (3), there exists* $T_1 = \sum_{t=1}^{T} \mathbb{1}_{s_t=1}$ *such that* $T_1 \geq K+1$ *and* $T_2 = \sum_{t=1}^{T} \mathbb{1}_{s_t=2} = T - T_1 \geq K+1$*; furthermore,* $E(\epsilon_t \epsilon_t') = \Omega_1$ *for all* t *such that* $s_t = 1$ *and* $E(\epsilon_t \epsilon_t') = \Omega_2$ *for all* t *such that* $s_t = 2$*.*

Assumption 2 guarantees that Ω_i, $i = 1, 2$ is positive definite such that there exists a Gramian decomposition of the form

$$\Omega_i = A^{-1} \Sigma_i A^{-1\prime}, \qquad i = 1, 2. \tag{8}$$

This assumption is line with the vast literature on the "Great Moderation" (see also Justiniano and Primiceri, 2008). The following proposition provides the conditions for the determination of A.

Proposition 2. *Under Assumption 1 and Assumption 2,* A *is uniquely identified if and only if* ω_2 *has distinct elements, or*

$$\frac{\sigma_{i,2}^2}{\sigma_{i,1}^2} \neq \frac{\sigma_{j,2}^2}{\sigma_{j,1}^2}, \qquad \text{for } i, j \in \{1, ..., K\}, \text{ and } i \neq j.$$

The proof is given in Appendix B.

Assumption 2 suggests global identification of A under specified conditions. In VAR analysis, Lütkepohl and Wozniak (2017) argue that this setup offers an advantage over other literature on identification via heteroskedasticity that uses the so-called B-models, resulting in locally identified shocks that are unique up to sign and ordering (Lütkepohl, 2005; Lanne et al., 2010). Such a model assumes that the link between structural and reduced-form shocks is $e_t = B^{-1}\epsilon_t$. Local uniqueness of B relies on normalisation of volatility in one state, e.g. $E(e_1 e_1') = I_K$. One potential issue regarding this specification in a Markov regime switching framework is the scaling problem associated with the label switching of states that may lead to an irregular shape of posterior distribution (Droumaguet et al., 2015). Furthermore, working with the contemporaneous matrix A in the current setup also facilitates Bayesian estimation and inference, as shown in the next section. Although proposition 2 relies on the heterogeneity in volatility states, it does not preclude the case where one structural shock is homoskedastic ($\sigma_{1,1}^2 = \sigma_{1,2}^2$). Proposition 2 guarantees identification as long as there is no proportional change in the

variances.

Unlike VAR models, identification via heteroskedasticity in MSUC models is not merely a statistical identification procedure since it also delivers structural errors with economic interpretation because of the trend-cycle decomposition achieved by UC models. In the spirit of Beveridge and Nelson (1981), the structural permanent shocks only drive trend components, and transitory shocks only drive cycle components. Component-to-component spillovers are modelled via the structural matrix A.

3.2 Bayesian estimation and inference

The sampling of the MSUC model has two main blocks: (i) unobserved components and (ii) model parameters. In what follows, we explain the sampling procedure and discuss how it differs from the Bayesian estimation and inference for VAR models identified via heteroskedasticity. The main complications come from the fact that components are unobserved and non-stationary. Therefore, it is not possible to assume that the first p observations are deterministic, which is common in Bayesian VARs. In our case, we have to consider them as "extra parameters" which have to be imputed from their conditional posterior. To see this, define the Kp-dimensional vector x_t as

$$x_t = (\tau_t', c_t', \tau_{t-1}, c_{t-1}', ..., \tau_{t-p+1}, c_{t-p+1}')', \qquad t = 1, ..., T.$$

Apparently, x_1 involves "historical" trends $\tau_{2-p}, ..., \tau_0$ and cycles $c_{2-p}, ..., c_0$ if $p \geq 2$. Even if we restrict ourselves to data $\{y_t\}_{t=p+1}^{T}$, we cannot treat x_{p+1} as deterministic because by definition it is a vector of unobserved random variables. We choose to impute the initial state vector such that our estimates are based on the correct posterior; and we define the transition matrix Ξ as

$$\Xi = \begin{bmatrix} \Xi_1 & \Xi_2 & \dots & \Xi_{p-1} & \Xi_p \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

where Ξ_i, $i = 1, ..., p$ is $K \times K$ such that

$$\Xi_1 = \begin{bmatrix} B_\tau & \mathbf{0}_{k\times N} \\ \mathbf{0}_{N\times k} & B_{c,1} \end{bmatrix}, \quad \Xi_i = \begin{bmatrix} \mathbf{0}_{k\times k} & \mathbf{0}_{k\times N} \\ \mathbf{0}_{N\times k} & B_{c,i} \end{bmatrix} \text{ for } i = \{2, ..., p\},$$

with $\mathbf{0}$ denoting a matrix of zeros with subscripted dimension. $B_{c,i}$ is the i-th coefficient matrix of the VAR(p) cycles in reduced-form and defined via (6). Therefore, the reduced-form model can be written as the following state space model,

$$y_t = \Lambda x_t, \qquad x_{t+1} = \Xi x_t + R\epsilon_t, \qquad \epsilon_t \sim N(0, \Omega_{s_t}), \tag{9}$$

with the loading matrix Λ and selection matrix R defined by $\Lambda = [Z : I_N : \mathbf{0}_{N\times K(p-1)}]$ and $R = [I_K : \mathbf{0}_{K\times K(p-1)}]'$, where $[M_1 : M_2]$ denotes horizontal concatenation of matrix M_1 and M_2.[11] It is clear that when drawing x_1 from its conditional posterior distribution we automatically impute $\tau_{2-p}, ..., \tau_0$ and $c_{2-p}, ..., c_0$ such that other parameters are drawn based on their correct conditional posterior, *i.e.* without risking ill-conditioned initialisations.

It is noteworthy that the above model is linear and Gaussian, conditional on the volatility states $S_T = \{s_1, ..., s_T\}$. We assume a two-state Markov regime switching framework for volatility (*i.e.* high and low volatility), equipped with a 2×2 transition matrix `P` whose i, j-entry denotes the probability of switching from state i to j, $i, j = 1, 2$. Model (9) is completed by assuming the following initialisation for the unobserved components:

$$\begin{aligned} x_1 &\sim N(\vec{x}_1, V_1), \\ \vec{x}_1 &= (\vec{\tau}_1', \vec{c}_1', ..., \vec{\tau}_{2-p}', \vec{c}_{2-p}')' = \mathbf{0}_{Kp\times 1}, \\ V_1 &= \begin{bmatrix} V_{\tau,1,1} & 0 & \dots & 0 & 0 \\ 0 & V_{c,1,1} & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & V_{\tau,1,2-p} & 0 \\ 0 & 0 & \dots & 0 & V_{c,1,2-p} \end{bmatrix}. \end{aligned}$$

[11] Note that one can model measurement errors in the observation equation in (9); but since such errors are also considered as transitory one can easily incorporate them into the cycle components.

The initial covariance matrix V_1 is such that $V_{\tau,1}$, which has diagonal blocks $V_{\tau,1,i} = M \cdot I_k$, $i = 1, 0, .., 2-p$ with M being a sufficiently large number such that gives the trends an uninformative initialisation (Durbin and Koopman, 1997); and $V_{c,1}$ denotes the covariance matrix from the ergodic distribution of the cycles consisting of blocks $V_{c,1,i}$, $i = 1, 0, ...2-p$. It follows that

$$\text{vec}(V_{c,1}) = (I_{Np} - B_c^* \otimes B_c^*)^{-1} \text{vec}\left(\sum_{i=1}^{2} \pi_i R_c \Omega_i R_c'\right). \tag{10}$$

In the above formulation, the $Np \times Np$ matrix B_c^* is the transition matrix of the VAR(p) cycles written in companion form. That is,

$$B_c^* = \begin{bmatrix} B_{c,1} & B_{c,2} & \dots & B_{c,p-1} & B_{c,p} \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 \end{bmatrix},$$

where $B_{c,i} = A_{cc}^{-1}\Phi_{i,cc}$ as in (6); $\pi = (\pi_1, \pi_2)'$ is the stationary distribution of the Markov process s_t such that $\pi = \pi \mathtt{P}$; and R_c contains rows of R for cycle components in x_t.

The conditional likelihood can be computed using the Kalman filter via prediction error decomposition (Durbin and Koopman, 2012, chapter 4). To facilitate the discussion on an efficient sequential procedure for sampling the contemporaneous matrix A and the unobserved components x_t, we repeat the seminal Kalman recursion.

Define $a_t = E(x_t|Y_{t-1})$ and $P_t = \text{Var}(x_t|Y_{t-1})$ where $Y_t = \{y_1, ..., y_t\}$. Let $a_1 = \vec{x}_t$ and $P_1 = V_1$ initialise the filter. We have the following recursion:

$$\begin{aligned} v_t &= y_t - \Lambda a_t, & F_t &= \Lambda P_t \Lambda', \\ a_{t+1} &= \Xi a_t + G_t v_t, & P_{t+1} &= \Xi P_t (\Xi - G_t \Lambda)' + R \Omega_t R', \end{aligned} \tag{11}$$

for $t = 1, ..., T$, where $G_t = T P_t \Lambda' F_t^{-1}$ is the Kalman gain. Notice that Ω_t depends on s_t, and as a result so do all other variables in the above recursion. Let $\boldsymbol{\theta}$ collect all model

parameters. The conditional likelihood is given by

$$p(Y_T|S_T,\boldsymbol{\theta}) = \exp\left(-\frac{TN}{2}\log 2\pi - \frac{1}{2}\sum_{t=1}^{T}\log|F_t| - \frac{1}{2}\sum_{t=1}^{T} v_t' F_t^{-1} v_t\right).$$

3.2.1 Prior distribution

For the prior distribution of the first volatility state, each $\sigma_{i,1}^2$ is assumed to be $IG(\alpha_v, \beta_v)$-distributed with the inverse gamma (IG) shape parameter $\alpha_v = 0.5$ and scale parameter $\beta_v = 0.5$ such that the prior is noninformative (no moment exists). The prior of each variance ratio $\omega_{i,2}$ is $IG(\alpha_\omega, \beta_\omega)$-distributed with $\alpha_\omega = 0.5$ and $\beta_\omega = 1.5$ such that the prior is noninformative and its mode equals 1. This is to say that at the mode we have homoskedasticity and A is not identified. By this design we limit the influence of priors on identification of structural matrix A to the minimum.

To simplify the notation, we assume that τ_t is a vector of random walk trend components, *i.e.* $k = N$ and $Z = I_N$.[12] The novelty of our MCMC algorithm relies on a recursive sampling scheme for A. To facilitate this, the prior of A is specified column-by-column for the following transformation of the unrestricted elements in each column. Notice that (2) implies

$$\begin{aligned}\tau_{i,t} =& - A_{i,1}\tau_{1,t} - ... - A_{i,i-1}\tau_{i-1,t} - A_{i,i+1}\tau_{i+1,t} - ... - A_{i,N}\tau_{N,t} \\ & - A_{i,N+1}c_{1,t} - ... - A_{i,N+i}c_{i,t} - ... - A_{i,K}c_{N,t} + \sum_{j=1}^{N}\Phi_{1,i,j}\tau_{j,t-1} \\ & + \sum_{j=N+1}^{K}\Phi_{1,i,j}c_{j-N,t-1} + ... + \sum_{j=1}^{N}\Phi_{p,i,j}\tau_{j,t-p} + \sum_{j=N+1}^{K}\Phi_{p,i,j}c_{j-N,t-p} + e_{i,t-1},\end{aligned}$$

for $i = 1, ..., N$, where $X_{i,j}$ denotes the ij-th element of matrix X. Similarly, we can have such an expression for the cycle components. Moving lagged variables to the left-hand side and noting $c_{i,t} = y_{i,t} - \tau_{i,t}$, the above equation becomes

$$\begin{aligned}\tau_{i,t}^* =& \delta_{i,1}\tau_{1,t} + ... + \delta_{i,i-1}\tau_{i-1,t} + \delta_{i,i+1}\tau_{i+1,t} + ... + \delta_{i,N}\tau_{N,t} \\ & + \delta_{i,N+1}c_{1,t} + ... + \delta_{i,N+i}y_{i,t}^* + ... + \delta_{i,K}c_{N,t} + (\delta_{i,N+i} - 1)e_{i,t-1},\end{aligned} \tag{12}$$

where $\tau_{i,t}^* = \tau_{i,t} + x_{i,t-1}^*$, $y_{i,t}^* = y_{i,t} - x_{i,t-1}^*$ with $x_{i,t-1}^* = \sum_{j=1}^{N}\Phi_{1,i,j}\tau_{j,t-1} +$

[12] It is straightforward to relax this assumption and to allow for more general specifications of trend components as discussed in Section 2.

$\sum_{j=N+1}^{K} \Phi_{1,i,j} c_{j-N,t-1} + ... + \sum_{j=1}^{N} \Phi_{p,i,j} \tau_{j,t-p+1} + \sum_{j=N+1}^{K} \Phi_{p,i,j} c_{j-N,t-p+1}$ with $\Phi_{n,i,j}$ denoting the ij-th element of Φ_n for $n = 1, ..., p$, and

$$\delta_{i,j} = -\frac{A_{i,j}}{1 - A_{i,N+i}} = -A_{i,j}(1 + \delta_{i,N+i}), \quad j \in \{1, ..., i-1, i+1, ..., K\}. \tag{13}$$

In (12), only $\delta_{i,N+i}$ appears multiplicatively with the error term, so we can easily implement a Gibbs sampler for all $\delta_{i,j}$'s, $j \neq N + i$ conditional on $\delta_{i,N+i}$, and an efficient accept-reject Metropolis-Hastings (MH) algorithm for $\delta_{i,N+i}$ because $A_{i,j}$ can be recovered with draws of $\delta_{i,j}$'s.[13] We assume a hierarchical normal prior $N(0, \gamma_j)$ for element $\delta_{i,j}$, $i = 1, ..., K$, where γ_j is $IG(\alpha_\gamma, \beta_\gamma)$-distributed with $\alpha_\gamma = \beta_\gamma = 0.5$. The shrinkage parameter γ_j controls the tightness of the normal prior which is a random variable itself to be determined by the data. This design assumes *a priori* that shocks from an unobserved component spill over to other components with similar magnitude, while allowing for heterogeneity of magnitude across components. *A posterior*, however, the hierarchical prior leaves ample room for data to speak.

The prior of the autoregressive transition matrix $\Phi = [\Phi_1 : ... : \Phi_p]$ is specified conditional on A, amounting to specifying a prior of the VAR coefficient matrix $B_c = [B_{c,1} : ... : B_{c,p}]$ for the VAR(p) cycles. Define the $K \times Kp$ matrix $W = [D : \mathbf{0}_{K \times K(p-1)}]$, where D is a diagonal matrix with k ones and N numbers within $[0, 1)$ on its diagonal suggesting that the first k coefficients of each row of Φ correspond to persistent trend components and the following N coefficients are for the cycle components (less persistent than integrated trend process). Define the $Kp \times Kp$ diagonal matrix L with diagonal $(\mathbf{1}_{1\times K}, \frac{1}{4}\mathbf{1}_{1\times K}, \frac{1}{9}\mathbf{1}_{1\times K}, ..., \frac{1}{p^2}\mathbf{1}_{1\times K})'$ where $\mathbf{1}$ denotes a matrix of ones with subscripted dimension. The prior of Φ'_{i-} is $N(A_{i-}W, \gamma_\Phi L)$, where X_{i-} denotes the i-th row of matrix X, for $i = 1, ..., K$; and the shrinkage parameter γ_Φ is $IG(\alpha_\gamma, \beta_\gamma)$-distributed.[14]

The prior of the i-th row, $i = 1, 2$ of the Markov transition probability matrix `P` for two-volatility-state is given a 2-dimensional Dirichlet distribution Dir_2 with off-diagonal

[13] This is an appealing feature embedded in MSUC models that Bayesian structural VARs do not have, as shown in the next section.

[14] This is similar to the Minnesota prior of Doan et al. (1984), which imposes a tightening pattern on the variances of coefficients for lagged variables.

parameter $e_j = 1$ and diagonal parameter $e_i = 10$, *i.e.*

$$\mathtt{P}'_{1-} \sim Dir_2(e_1, e_2), \quad \mathtt{P}'_{2-} \sim Dir_2(e_2, e_1).$$

This design assumes *a priori* that $p(s_t = i|s_{t-1} = j) = p(s_t = j|s_{t-1} = i) = 1/11$ and $p(s_t = i|s_{t-1} = i) = p(s_t = j|s_{t-1} = j) = 10/11$ for $i, j = 1, 2$ and $i \neq j$, meaning that the volatility states are persistent over time.

In summary, we have the following prior distribution of θ

$$\begin{aligned} p(\boldsymbol{\theta}) =& p(\gamma_\Phi) \times \prod_{i=1}^{K} p(\gamma_i) \times \prod_{j=1}^{K} \prod_{i=1;i\neq j}^{K} p(\delta_{i,j}|\gamma_j) \times \prod_{i=1}^{2} p(\mathtt{P}'_{i-}) \\ & \times \prod_{i}^{K} p(\Phi'_{i-}|A_{i-}, \gamma_\phi) p(\sigma^2_{i,1}) p(\omega_{i,2}), \end{aligned}$$

where component-specific distributions are given by

$$\begin{aligned} \gamma_\Phi &\sim IG(\alpha_\gamma, \beta_\gamma), \\ \gamma_i &\sim IG(\alpha_\gamma, \beta_\gamma), \quad i = 1, ..., K, \\ \delta_{i,j}|\gamma_j &\sim N(0, \gamma_j), \quad j = 1, ..., K, \quad i = 1, ..., j-1, j+1, ...K, \\ \Phi'_{i-}|A_{i-}, \gamma_\Phi &\sim N(A_{i-}W, \gamma_\Phi L), \quad i = 1, ..., K, \\ \sigma^2_{i,1} &\sim IG(\alpha_v, \beta_v), \quad i = 1, ..., K, \\ \omega_{i,2} &\sim IG(\alpha_\omega, \beta_\omega), \quad i = 1, ..., K, \\ \mathtt{P}'_{1-} &\sim Dir_2(e_1, e_2), \quad \mathtt{P}'_{2-} \sim Dir_2(e_2, e_1). \end{aligned}$$

3.2.2 Sampling procedure and inference

The Bayesian sampling procedure iterates over the following five sampling blocks to generate posterior draws:

1. Sample A and x_t from $p(\{\delta_{i,j}\}_{j=1,j\neq i}^{K}, \{\tau_{i,t}\}_{t=1}^{T}|Y_{i,T}, \{\gamma_j\}_{j=1,j\neq i}^{K}, \sigma^2_{i,1}, \omega_{i,2}, S_T)$ for $i = 1, ..., N$, where $Y_{i,T} = \{y_{i,1}, ..., y_{i,T}\}$;

2. Sample Φ from $p(\Phi'_{i-}|A_{i-}, \gamma_\Phi, X_T, \sigma^2_{i,1}, \omega_{i,2}, S_T)$ for $i = 1, ..., K$, where $X_T = \{x_1, ..., x_T\}$;

3. Sample Σ_1 from $p(\sigma^2_{i,1}|X_T, A, \Phi, \omega_{i,2}, S_T)$ and sample ω_2 from

$p(\omega_{i,2}|X_T, A, \Phi, \sigma^2_{i,1}, S_T)$ for $i = 1, ..., K$;

4. Sample S_T from $p(S_T|X_T, A, \Phi, \Sigma_1, \Sigma_2, \mathtt{P})$ and sample $\mathtt{P}$ from $p(\mathtt{P}'_{i,-}|X_T, A, \Phi, \Sigma_1, \Sigma_2, S_T)$ for $i = 1, 2$;

5. Sample γ_Φ from $p(\gamma_\Phi|\Phi)$ and sample $\{\gamma_i\}_{i=1}^K$ from $p(\gamma_i|A_{-i})$ for $i = 1, ..., K$, where A_{-i} denotes the i-th column of A.

One novelty of our proposed sampling procedure is to sample rows of A and elements of x_t sequentially to complete the sampling in Block 1. In the literature of structural VAR models identified using heteroskedasticity (Canova and Pérez Forero, 2015 and Lütkepohl and Wozniak, 2017), a random walk sampler is usually employed to sample $\vec{A}$ as a whole in an accept-reject MH algorithm. To see why a random walk sampler is inefficient in the MSUC models, let the structural state equation (2) be rewritten as

$$A\dot{x}_{t+1} = \Phi x_t + e_t, \quad \dot{x}_{t+1} = (\tau'_{t+1}, c'_{t+1})', \tag{14}$$

and define $\tilde{x}_{t+1} = (\dot{x}'_{t+1} \otimes I_K)h - \Phi x_t$ and $\hat{x}_t = -(\dot{x}'_{t+1} \otimes I_K)H$. Let $a \in \mathbb{R}^{K(K-1)}$ collect unrestricted elements in A via $\text{vec}(A) = Ha + h$ so that H and h contain zeros and ones. Without considering initialisation, the latter approach would rely on the following conditional posterior distribution

$$p(A|X_T, \Phi, \Sigma_1, \Sigma_2, S_T) \propto |\det(A)|^{T-1} \exp\left\{-\frac{1}{2}\sum_{t=1}^{T-1}(\tilde{x}_{t+1} - \hat{x}_t a)'\Sigma^{-1}_{s_t}(\tilde{x}_{t+1} - \hat{x}_t a)\right\} p(a).$$

The above posterior prompts one to use a MH algorithm with multivariate normal or Student's t-proposal around $(\sum_{t=1}^{T-1} \hat{x}'_t\Sigma^{-1}_{s_t}\hat{x}_t)^{-1} \sum_{t=1}^{T-1} \hat{x}'_t\Sigma^{-1}_{s_t}\tilde{x}_{t+1}$ with scale matrix $(\sum_{t=1}^{T-1} \hat{x}'_t\Sigma^{-1}_{s_t}\hat{x}_t)^{-1}$, depending on a chosen prior for a. Due to the factor $|\det(A)|^{T-1}$ outside the Gaussian kernel, a small change generated by the random walk sampler would potentially lead to big differences in the posterior ordinate, meaning the Markov chain can get stuck very easily, especially when the dimension of A is large. The inefficiency worsens in MSUC models because $\tilde{x}_t$ and $\hat{x}_t$ are unobserved and imputed from data. Apparently, the product form in (14) is the major source of inefficiency. Our approach, however, circumvents the dimensionality issue based on the transformation (12).

Suppressing dependence on other parameters, we notice that draws from the conditional posterior distribution $p(A_{1-}, ..., A_{K-}, \tau_{1,T}, c_{1,T}, ..., \tau_{N,T}, c_{N,T}|Y_T)$ for A and x_t in Block 1, where $\tau_{i,T} = \{\tau_{i,1}, ..., \tau_{i,T}\}$ and $c_{i,T} = \{c_{i,1}, ..., c_{i,T}\}$, can be facilitated by a Gibbs sampler that iterates over

$$\tau_{i,T}, c_{i,T}|\bullet \sim p(\tau_{i,T}, c_{i,T}|A_{i-}, ..., A_{K-}, \tau_{1,T}, c_{1,T}, ..., \tau_{i-1,T}, c_{i-1,T}, \tau_{i+1,T}, c_{i+1,T}, ..., \tau_{N,T}, c_{N,T}, Y_T), \tag{15}$$

$$A_{i-}|\bullet \sim p(A_{i-}|A_{1-}, ..., A_{i-1-}, A_{i+1-}, ..., A_{K-}, \tau_{1,T}, c_{1,T}, ..., \tau_{N,T}, c_{N,T}, Y_T), \tag{16}$$

for $i = 1, ..., N$. It is interesting to note that the conditional posterior in (15) is implied by a univariate MSUC model only involving $Y_{i,T}$. The latter can be represented by the following state space model:

$$\begin{aligned} y_{i,t} &= \tau_{i,t} + c_{i,t}, \\ A_i^* \begin{pmatrix} \tau_{i,t+1} \\ c_{i,t+1} \end{pmatrix} &= d_t^i + \Phi_{i,1}^* \begin{pmatrix} \tau_{i,t} \\ c_{i,t} \end{pmatrix} + ... + \Phi_{i,p}^* \begin{pmatrix} \tau_{i,t-p+1} \\ c_{i,t-p+1} \end{pmatrix} + \begin{pmatrix} e_{i,t} \\ e_{N+i,t} \end{pmatrix}, \end{aligned}$$

where

$$A_i^* = \begin{bmatrix} 1 & A_{i,N+i} \\ A_{N+i,i} & 1 \end{bmatrix}, \quad \Phi_{i,n}^* = \begin{bmatrix} \Phi_{n,i,i} & \Phi_{n,i,N+i} \\ \Phi_{n,N+i,i} & \Phi_{n,N+i,N+i} \end{bmatrix}, \quad n = 1, ..., p,$$

and $d_t^i = (d_{\tau,t}^i, d_{c,t}^i)'$ is a known but time-variant vector conditional on draws of other unobserved components, defined by

$$\begin{aligned} d_{\tau,t}^i &= f_{\tau,t}^i + \sum_{n=1}^{p} \left(\sum_{j=1, j\neq i}^{N} [\Phi_{n,i,j}\tau_{j,t+1-n} + \Phi_{n,i,N+j}c_{j,t+1-n}] \right), \\ d_{c,t}^i &= f_{c,t}^i + \sum_{n=1}^{p} \left(\sum_{j=1, j\neq i}^{N} [\Phi_{n,N+i,j}\tau_{j,t+1-n} + \Phi_{n,N+i,N+j}c_{j,t+1-n}] \right), \\ f_{\tau,t}^i &= -\sum_{j=1, j\neq i}^{N} [A_{i,j}\tau_{j,t+1} + A_{i,N+j}c_{j,t+1}] \\ f_{c,t}^i &= -\sum_{j=1, j\neq i}^{N} [A_{N+i,j}\tau_{j,t+1} + A_{N+i,N+j}c_{j,t+1}]. \end{aligned}$$

The above is a linear Gaussian state space model with time-varying predetermined

intercept $(A_i^*)^{-1}d_t^i$ and state covariance matrix $(A_i^*)^{-1}\text{diag}(\omega_{i,s_t}\sigma_i^2, \omega_{N+i,s_t}\sigma_{N+i}^2)(A_i^*)^{-1'}$ which can be cast into a reduced-form similar to (9). We adopt the simulation smoother developed by Durbin and Koopman (2002) for the above system to generate draws of $\tau_{i,T}$ and $c_{i,T}$.[15]

It is worth noticing that, once conditional on other rows of A, the conditioning set in the conditional posterior distribution of A in (16) is the same as when $Y_{j,T}$, $j \neq i$ dropped. This unique feature of MSUC models allows us to sample A row-by-row with transformation (13) and using (12). For each $i = 1, ..., K$, the sampling of $\delta_{i,j}$ for $j = 1, ...i-1, i+1, ..., K$ consists of two steps. In the first step, (12) boils down to a standard Bayesian least square problem given $\delta_{i,N+i}$. In the second step, we can use a MH algorithm based on importance sampling for the single parameter $\delta_{i,N+i}$ conditional on δ_{i1}, also using (12).[16] The proposal density is constructed from the first and second order derivative of its conditional posterior distribution. Further details on these sampling procedures are given in Appendix D.

Instead of using a MH algorithm for sampling A all together, our proposed procedure effectively separates the sampling of A into pieces while maintaining a high acceptance rate, which depends on the effectively costless task of finding an univariate proposal density for $\delta_{i,N+i}$. Table 1 shows the result of a small simulation study where we compare our algorithm with the MH algorithm of sampling A as a whole using a random walk sampler as in Bayesian VARs with heteroskedasticity (Canova and Pérez Forero, 2015). As the dimension increases, the number of unrestricted elements in A grows rapidly, rendering the random walk sampler inefficient. The fact that when $N = 5$ only 12% of the draws are accepted highlights the drawback of using a MH algorithm for sampling A as a whole because of its multiplication by x_t, which itself needs to be drawn. Our proposed algorithm is much less sensitive to dimensionality by iteratively sampling A_{i-}. Although one may argue that not sampling A in one step leads to poor mixing, we see from Table 1 that this is not the case since the algorithm maintains high acceptance rate and functionality when N increases.[17]

[15] The choice of sampler does not affect final results. The precision sampler of Chan and Jeliazkov (2009) designed for UC models does not gain much efficiency in our case due to correlation between trends and cycles. We thus adopt the sampler by Durbin and Koopman (2002) because of its easy implementation and speediness. Alternatively, one can use the one proposed by De Jong and Shephard (1995) or the one by Carter and Kohn (1994).

[16] Alternatively, one can apply a Griddy-Gibbs sampler; but we find that the basic Griddy-Gibbs sampler generates draws of $\delta_{i,N+i}$ with very high autocorrelation.

[17] The sampling of other system parameters is relatively standard, taking into account an extra accept-

Table 1: Comparison of two samplers for structural matrix A in MSUC models

$\dim(y_t) : N$	1	2	3	4	5
$\dim(a) : K(K-1)$	2	12	30	56	90
Sample A row-by-row with parameter transformation					
IF	6.22	8.41	7.63	11.92	10.81
acceptance rate	0.94	0.92	0.82	0.82	0.74
Sample A as a whole with random walk sampler					
IF	14.93	22.67	28.93	89.27	> 300
acceptance rate	0.88	0.64	0.60	0.46	0.12

[1] We simulate y_t with different dimensions from a MSUC model with heteroskedasticity (Markov switching between two volatility regimes). y_t consist of a (vector) random walk trend and (V)AR(2) cycles with diagonal coefficient matrices in its reduced-form. Both samplers generate a Markov chain with length 30,000, with the initial 2,000 iterations discarded, keeping parameters other than A at their DGP values. [2] IF is the inefficiency factor computed using Parzen window with bandwidth 1000. IF measures efficiency based on the speed of decaying for the autocorrelation function.

3.2.3 Stochastic volatility

Two-state volatility regime switching satisfies the identification condition, and extensions with more than two regimes can be estimated similarly. The assumption of abrupt volatility regime switches and switches occurring at the same time for all structural shocks, however, can be restrictive. Depending on the nature of shocks, transitory or permanent, volatility changes may take place in different magnitudes and at different times (Fisher et al., 2016). Thus, we consider stochastic volatility an agnostic way for modelling potentially different patterns of volatility changes where the adjustment occurs at each point in time.

In brief, we assume the following:

Assumption 3. *The logarithm of each structural volatility series is a random walk. That is, for* $i = 1, ..., K$ *and* $t = 1, ..., T-1$,

$$\log \sigma_{i,t+1} = \log \sigma_{i,t} + \rho_i \xi_{i,t}, \quad \xi_{i,t} \sim N(0,1),$$

with diffuse initialisation $\log \sigma_{i,1} \sim N(0, \bar{M})$, *where* $\bar{M}$ *is a big number.*

The volatility of volatility parameter ρ_i is unknown and needs to be estimated. Similar

reject MH step that corrects for initialisation; details are reported in Appendix D.

to the Markov regime switching case, the identification requires changes of volatility over time, which is warranted (*a.s.*) via the random walk specification; therefore we have

Proposition 3. *Under Assumption 1 and Assumption 3, A is uniquely identified.*

The proof is given in Appendix C.

In this case, we have[18]

$$\Omega_t = A^{-1}\Sigma_t A^{-1\prime}, \quad \Sigma_t = \text{diag}(\sigma_{1,t}^2, ..., \sigma_{K,t}^2).$$

In our specification we have $A_{i-}\dot{x}_{t+1} = \Phi_{i-}x_t + e_{i,t}$ with $e_{i,t} \sim N(0, \sigma_{i,t}^2)$ for $i = 1, ..., K$. For the sampling procedure, we adopt the 7-component Gaussian mixture method of Kim et al. (1998) to draw from $p(\log \sigma_{i,t} | \{e_{i,t}\}_{t=1}^{T-1}, \rho_i, S_T)$ and $p(s_t | \{\log \sigma_{i,t}\}_{t=1}^{T-1}, \{e_{i,t}\}_{t=1}^{T-1})$ iteratively, where $s_t = \{1, ..., 7\}$ here is a tabulated indicator process selecting the mixing component.[19]

3.2.4 Bayesian test for no spillover

In order to test for spillover effects between the structural shocks, we consider the Bayes factor computed using the Savage-Dicky density ratio (SDDR). Suppose that we are interested in testing the null hypothesis $H_0 : A_{i,j} = 0$ (or, equivalently, $H_0 : \delta_{i,j} = 0$ based on (13)), that is to say, the structural shock driving the j-th component does not spillover to the i-th component. The SDDR is given by

$$SDDR_{i,j} = \frac{p(\delta_{i,j} = 0 | Y_T)}{p(\delta_{i,j} = 0)},$$

[18]Note that here we require that all volatility series are time-varying. Bertsche et al. (2018) show that in structural autoregressive models the same setup allows one structural shock that has constant volatility. If more than one structural shocks have constant volatility, then the corresponding block A^{-1} has to be upper diagonal. We, however, do not discuss such cases in this paper.

[19]The modelling of volatility dynamics is deemed as of secondary importance if the focus is to study the component-to-component spillover effect via estimating the contemporaneous structural matrix A. Furthermore, any "intermediate" cases where the number of volatility changes is within $[3, T-1]$ can be accommodated by, *e.g.*, the Dirichelet process mixture model of Bauwens et al. (2017). It is reasonable to assume that, if the structural matrix A estimated under two-state and stochastic volatility are numerically close (which is what find in our empirical study), any other "intermediate" cases will deliver similar results.

where the numerator and denominator indicate the posterior and prior density of $\delta_{i,j}$ evaluated at zero, respectively. The denominator is given by

$$p(\delta_{i,j}=0)=\int_{\gamma_j} p(\delta_{i,j}=0|\gamma_j)p(\gamma_j)d\gamma_j=\frac{1}{\sqrt{2\pi}}E(\gamma_j^{-1/2})=\frac{\Gamma(\alpha_\gamma+0.5)}{\Gamma(\alpha_\gamma)\sqrt{2\pi\beta_\gamma}},$$

where $\Gamma(.)$ denotes the gamma function. The numerator does not have an analytical expression, but an unbiased estimate is easy to compute using the importance-weighted marginal posterior density estimator (IWMDE) $\hat{p}(\delta_{i,j}=0|\gamma_j)$ of Chen (1994). Let $\varphi^{(m)}$ denote the m-th draw from the MCMC sample of all parameters and unobserved components; let $\delta_{i,j}^{(m)}$ denote the m-th draw from the sample of $\delta_{i,j}$; and let $\varphi_{-(i,j)}^{(m)}$ denote $\varphi^{(m)}$ but without $\delta_{i,j}^{(m)}$. The IWMDE is given by

$$\hat{p}(\delta_{i,j}=0|Y_T)=\frac{1}{M}\sum_{m=1}^{M} g(\delta_{i,j}^{(m)}|\varphi_{-(i,j)}^{(m)})\frac{p(Y_T|\delta_{i,j}=0,\varphi_{-(i,j)}^{(m)})p(\delta_{i,j}=0,\varphi_{-(i,j)}^{(m)})}{p(Y_T|\varphi^{(m)})p(\varphi^{(m)})}, \tag{17}$$

where M the total number of draws and where $g(\delta_{i,j}^{(m)}|\varphi_{-(i,j)}^{(m)})$ is the proposal distribution used for drawing $\delta_{i,j}^{(m)}$ given in Section 3.2.2. Importantly, we omit $\tau_i^{(m)}$ and $c_i^{(m)}$ for $i=1,...,N$ when computing $p(Y_T|\delta_{i,j}=0,\varphi_{-(i,j)}^{(m)})$ and $p(Y_T|\varphi^{(m)})$, but instead apply the Kalman recursion (11) to integrate out all unobserved components. This Rao-Blackwellisation step is essential to reducing Monte Carlo noise. Since $E(\hat{p}(\delta_{i,j}=0|y_T)=p(\delta_{i,j}|Y_T)$, we calculate $\widehat{SDDR}_{i,j}=\hat{p}(\delta_{i,j}=0|Y_T)/p(\delta_{i,j}=0)$ which is an unbiased estimator of $SDDR_{i,j}$.

The $\widehat{SDDR}_{i,j}$s for all i and j are computed once the MCMC is stopped. This procedure is easy to implement for testing no spillover effects between unobserved components or structural shocks. In our empirical study, the $\widehat{SDDR}_{i,j}$s are compared with the scale proposed by Kass and Raftery (1995) for determining Bayesian statistical significance.

4 Empirical results

In this section, we empirically test for spillovers between permanent and transitory shocks in the US economy considering MSUC versions of Okun's law and the Phillips curve. Okun's law studies the relationship between output and unemployment; while the Phillips curve links output and inflation.

4.1 Data and the MSUC model

Let $y_t = (y_{1,t}, y_{2,t})'$, where $y_{1,t}$ is (the natural logarithm of) real GDP and $y_{2,t}$ is either the unemployment rate or the CPI headline inflation rate. Figure 1 shows the time series plots for the period 1970Q1-2018Q3 used in our analysis.

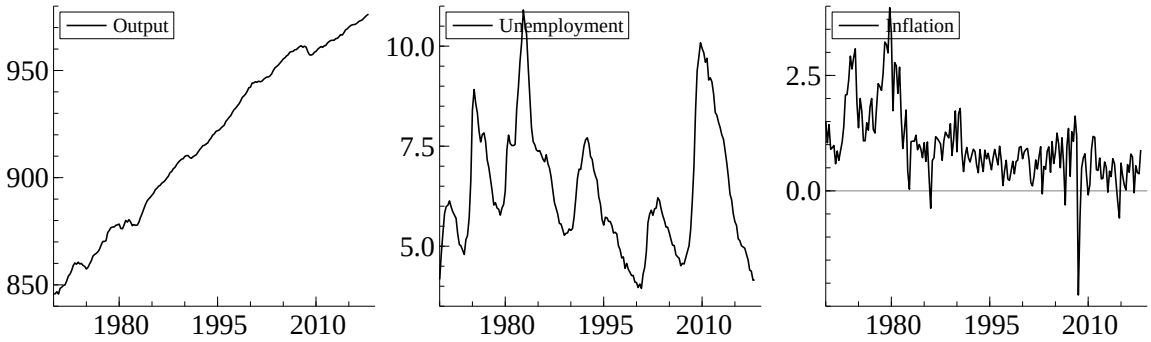

Figure 1: USA, 1970Q1-2018Q3. Time series plots of output, unemployment rate and inflation. Output is the natural logarithm of real GDP; unemployment rate is the civilian unemployment rate; and inflation is the CPI headline inflation rate.

We consider that each quarterly time series is composed of a non-stationary trend and a stationary cycle, *i.e.* $y_{i,t} = \tau_{i,t} + c_{i,t}$, $i = 1, 2$. Augmented Dickey-Fuller tests suggests that the output series is an $I(2)$ process, in line with the findings of Stock and Watson (1998) and Harvey (2011). Furthermore, the KPSS test (Charemza and Syczewska, 1998) rejects the null of trend stationarity for $\Delta y_{1,t}$, where Δ denotes the difference operator, thus indicating that the growth rate of trend output is a non-stationary process. Hence, we consider an integrated random walk for the output trend as in Stock and Watson (1998), Harvey (2011) and Holston et al. (2017), which means that the drift term measuring the growth rate of trend output is a random walk itself:

$$\tau_{1,t+1} = \tau_{g,t} + \tau_{1,t} + \epsilon_{\tau_1,t},$$
$$\tau_{g,t+1} = \tau_{g,t} + \epsilon_{\tau_g,t},$$

for $t = 1, ..., T-1$. This specification for output characterises two types of permanent shocks that drive output trend in the reduced-form models: $\epsilon_{\tau_1,t}$ and $\epsilon_{\tau g,t}$, which affect the level of the trend and the growth rate of the trend, respectively. Perron and Wada (2009), Basistha and Startz (2008) and Sinclair (2009) (among others) assume that the trend growth rate exhibits a one-time structural break. We do not follow this approach

because the break is assumed to be exogenous; by contrast, we consider the possibility that trend growth can be affected by both structural permanent and transitory shocks.[20]

We assume a VAR(2) process for the cycle components in the reduced-form specification.[21] The structural equation describing the dynamics of unobserved components is given by

$$A\begin{pmatrix}\tau_{1,t+1}\\ g_{t+1}\\ \tau_{2,t+1}\\ c_{1,t+1}\\ c_{2,t+1}\end{pmatrix} = \Phi_1\begin{pmatrix}\tau_{1,t}\\ g_{t}\\ \tau_{2,t}\\ c_{1,t}\\ c_{2,t}\end{pmatrix} + \Phi_2\begin{pmatrix}\tau_{1,t-p+1}\\ g_{t-p+1}\\ \tau_{2,t-p+1}\\ c_{1,t-p+1}\\ c_{2,t-p+1}\end{pmatrix} + \begin{pmatrix}e_{\tau_1,t}\\ e_{g,t}\\ e_{\tau_2,t}\\ e_{c_1,t}\\ e_{c_2,t}\end{pmatrix}, \quad e_t \sim N(0, \Sigma_{s_t}),$$

where $\Sigma_{s_t} = \text{diag}(\sigma^2_{\tau_1,t}, ..., \sigma^2_{c_2,t})$. Regarding the volatility dynamics, we consider both a two-regime Markov switching model, *i.e.* $s_t = 1, 2$, and a stochastic volatility model with $s_t = t$ and $\log(\sigma_{i,t})$ being a random walk process for all i.

4.2 Estimation results

Proposition 2 states that the identifiability of the structural matrix A relies on the variation of the volatility of structural shocks. As mentioned above, we considered two different forms of heteroskedasticity for the MSUC version of Okun's law and the Phillips curve: a two-state Markov regime switching model and a stochastic volatility model.[22]

Figure 2 shows the posterior estimate of the volatility state indicator s_t obtained from the MSUC versions of Okun's law and the Phillips curve. The models clearly identify two volatility states. The high volatility state takes place before the mid-1980's, followed by a fall in volatility — which is consistent with the "Great Moderation" literature, showing up again during the 2008 Great Financial Crisis (GFC), and declining again since 2010.[23]

[20] We also conduct a Bayesian test for non-zero variance $\sigma^2_{\tau_g}$ of $\epsilon_{\tau g,t}$ by applying the non-centred parametrisation of Frühwirth-Schnatter and Wagner (2010) to a univariate UC model for output. The posterior distribution of σ_{τ_g} (with a non-centred transformation) unambiguously shows two clusters of probability mass, suggesting non-zero innovations that drive trend growth.

[21] We also increased the VAR order to 4 and 8; but our main conclusions remained unaltered. These results are not reported, but are available on request.

[22] We also investigated the reduced-form system by considering a full innovation covariance matrix; and we report those results in Appendix E in order to focus on the description of the spillover effects in this section.

[23] Appendix E also shows the posterior estimates of the stochastic volatility series. Note that the Phillips

Thus, both the stochastic volatility model and the two-regime Markov regime switching model capture the high volatility period before 1985 and during 2008.

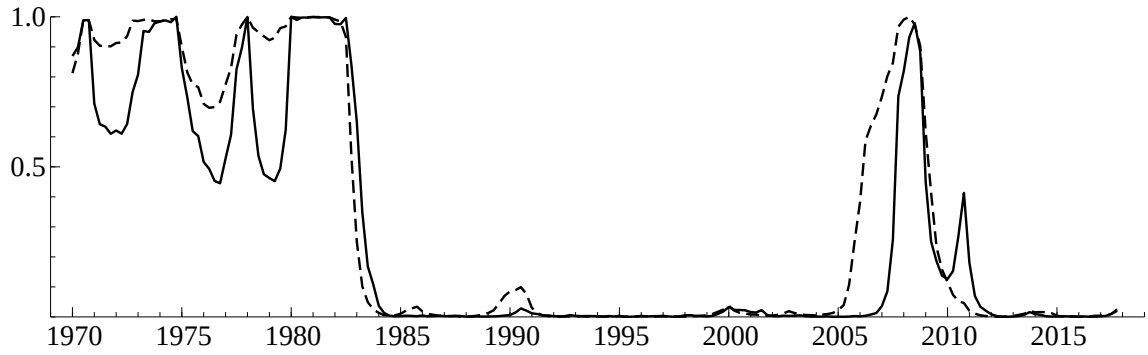

Figure 2: Estimated volatility regimes of MSUC models. The solid and dash lines indicate the posterior probability of being in the high volatility state based on Okun's law and the Phillips curve, respectively.

The posterior estimate of the structural matrix A is reported in Table 2 together with the SDDR testing for $H_0 : A_{i,j} = 0$, introduced in Section 3.2.4. As discussed previously, the correlation between component innovations in correlated UC models results from multiplying A^{-1} on both sides of the structural state transition (2). This means that A^{-1} as in $\epsilon_t = A^{-1}e_t$ captures the interplay between structural shocks or how they spillover to unobserved components, which is reported in Table 3. To summarise, Table 2 reports the statistical significance of the within-components and cross-components spillovers; while 3 reports the direction and degree of the respective spillovers.

Tables 2 and 3 show several interesting results. First, all the diagonal elements of A^{-1} are positive, which means that we can pair identified structural shocks with their corresponding components. Although we do not restrict all elements on the diagonal of A^{-1} to be positive *a priori*, the MCMC algorithm never delivers negative values for the diagonal elements of A^{-1}. This enables the structural interpretation of identified permanent and transitory shocks; but we also notice this may not be true in other contexts.[24]

curve suggests a more persistent state dynamics, because inflation is more responsive to volatility changes. Regarding Okun's law, only trend unemployment experiences a clear volatility spike during the GFC, while evidence from the other components seems to be more ambiguous. With respect to the Phillips curve, trend inflation is not exposed to high volatility during the GFC, while the other components become quite volatile during that period.

[24]For example, if $(A^{-1})_{1,1}$ were negative, the contemporaneous impulse response of $\tau_{1,t}$ to positive of $e_{\tau_1,t}$ would be negative, implying an unconventional interpretation of the structural shock $e_{\tau_1,t}$ driving $\tau_{1,t}$. In structural VARs, A^{-1} is usually restricted to have ones on its diagonal; such specification however

Table 2: Estimated contemporaneous structural matrix obtained from the MSUC model

	Okun's law					The Phillips curve				
Two volatility states Markov regime switching										
	$\tau_{1,t}$	$\tau_{g,t}$	$\tau_{2,t}$	$c_{1,t}$	$c_{2,t}$	$\tau_{1,t}$	$\tau_{g,t}$	$\tau_{2,t}$	$c_{1,t}$	$c_{2,t}$
$\tau_{1,t}$	1	0.05 −5.1	−0.07 0.3	0.84 **−24.3**	0.23 **−17.2**	1	0.23 −7.2	−0.08 **−16.2**	0.42 **−31.5**	0.04 1.2
$\tau_{g,t}$	0.02 −0.8	1	−0.12 **−14.5**	0.00 0.7	0.18 **−16.9**	0.12 −8.2	1	−0.48 **−18.6**	0.35 **−13.0**	−0.23 **−34.3**
$\tau_{2,t}$	0.03 −0.4	0.04 0.7	1	0.29 **−11.8**	0.08 −3.7	0.12 −9.3	0.21 **−23.0**	1	−0.32 **−25.6**	0.30 **−37.1**
$c_{1,t}$	−0.18 **−14.8**	−0.27 **−18.7**	0.31 **−32.5**	1	0.22 **−16.5**	−0.34 **−14.4**	−0.32 **−26.1**	0.21 **−12.0**	1	−0.54 **−16.8**
$c_{2,t}$	−0.00 0.4	−0.01 1.1	0.74 **−19.3**	0.65 **−11.5**	1	0.23 **−10.4**	−0.00 −0.1	0.14 −8.2	−0.27 **−16.6**	1
Stochastic volatility										
$\tau_{1,t}$	1	0.06 1.4	−0.11 −8.1	0.52 **−14.2**	0.27 **−24.7**	1	0.01 0.2	−0.03 0.3	0.30 **−17.8**	−0.00 −0.0
$\tau_{g,t}$	0.02 −1.1	1	−0.12 **−15.5**	0.01 0.4	0.26 **−19.4**	0.01 −0.4	1	−0.62 **−32.6**	0.48 **−24.1**	−0.45 **−16.3**
$\tau_{2,t}$	0.11 −8.6	0.03 1.4	1	0.22 −8.1	0.06 −3.2	0.17 −4.0	0.28 **−16.2**	1	−0.46 **−18.6**	0.22 **−22.4**
$c_{1,t}$	−0.20 **10.4**	−0.23 **−17.0**	0.21 −9.2	1	0.21 **−35.2**	−0.32 **−21.8**	−0.18 **−17.2**	0.21 **−11.6**	1	−0.44 **−31.2**
$c_{2,t}$	−0.00 2.3	0.00 −0.1	0.65 **14.4**	0.56 **−27.8**	1	0.18 −8.2	−0.02 0.0	0.18 −8.4	−0.26 **−21.7**	1

Reported is the posterior mean estimate of the structural matrix A obtained from the MSUC model for Okun's law and the Phillips curve. Below the posterior mean is the SDDR for testing $H_0 : A_{i,j} = 0$, with boldface number indicating strong evidence against H_0. The upper panel shows estimates identified considering the two-volatility-state Markov regime switching, while the bottom panel shows estimates identified considering stochastic volatility.

Table 3: Estimated spillover effects of shocks obtained from the MSUC model

	Okun's law					The Phillips curve				
Two volatility states Markov regime switching										
	$e_{\tau_1,t}$	$e_{\tau_g,t}$	$e_{\tau_2,t}$	$e_{c_1,t}$	$e_{c_2,t}$	$e_{\tau_1,t}$	$e_{\tau_g,t}$	$e_{\tau_2,t}$	$e_{c_1,t}$	$e_{c_2,t}$
$\epsilon_{\tau_1,t}$	0.86	-0.26	0.35	-0.75	-0.11	0.99	-0.36	0.03	-0.37	-0.33
$\epsilon_{\tau_g,t}$	-0.04	0.98	-0.01	-0.10	0.21	-0.21	0.93	0.44	-0.08	0.05
$\epsilon_{\tau_2,t}$	-0.07	-0.10	1.08	-0.23	-0.04	0.03	-0.11	0.91	0.28	-0.14
$\epsilon_{c_1,t}$	0.18	0.27	-0.12	1.01	-0.20	0.16	0.30	-0.13	0.96	0.62
$\epsilon_{c_2,t}$	-0.06	-0.10	-0.72	-0.49	1.16	-0.19	0.18	-0.17	0.31	1.27
Stochastic volatility										
$\epsilon_{\tau_1,t}$	0.89	-0.15	0.30	-0.41	-0.21	0.94	-0.07	0.07	-0.25	-0.16
$\epsilon_{\tau_g,t}$	-0.05	0.97	-0.05	-0.14	0.30	-0.16	0.85	0.50	-0.07	0.24
$\epsilon_{\tau_2,t}$	-0.14	-0.06	1.01	-0.15	-0.01	0.01	-0.15	0.82	0.43	-0.06
$\epsilon_{c_1,t}$	0.20	0.22	-0.03	1.01	-0.21	0.22	0.21	-0.14	0.90	0.53
$\epsilon_{c_2,t}$	-0.02	-0.09	-0.65	-0.47	1.12	-0.11	0.11	-0.19	0.20	1.18

Reported is the posterior mean of the inverse of structural matrix, *i.e.* A^{-1}. The upper panel shows estimates identified considering the two-volatility-state Markov regime switching, while the bottom panel shows estimates identified considering stochastic volatility.

Second, the estimates obtained from the models that consider stochastic volatility have the same sign and similar magnitudes compared with those that consider the regime switching volatility, thus suggesting that our results are robust to different specifications of volatility dynamics.[25]

Third, the majority of both the within-components spillovers and the cross-components spillovers are significant. This suggests relevant interactions between trend and cycles for GDP, unemployment and inflation.

Regarding the output within-components spillovers, we find significant positive spillovers from trend-to-cycle and negative cycle-to-trend spillovers when both Okun's law and the Phillips curve are considered. However, the element $A_{2,4}$ is significant only in the Phillips curve estimation, indicating that only output cycle shocks measured via the Phillips curve specification (and not via Okun's law) generate negative spillover

cannot facilitate the proposed efficient sampling scheme introduced in Section 3.2.2.

[25]It is worth mentioning that, since the identification of A relies on the assumption that the time-variation of the covariance matrix of reduced-form shocks only comes from heteroskedasticity of structural shocks, it can be argued that a constant contemporaneous relationship between components may not hold. To examine this, we estimated the MSUC models using data prior to the GFC and report the results in Appendix E. It can be observed that the GFC does not change the overall magnitude of the elements in A and, thus, our discussion is also robust to the inclusion of the GFC period.

effects to the trend output growth rate. Likewise, the negative spillovers from cycle-to-trend are larger than the respective positive ones from trend-to-cycle: the elements $(A^{-1})_{1,4}$ are larger than $(A^{-1})_{4,1}$ in Table 3. On the other hand, with respect to the within-components spillovers for the unemployment rate (inflation rate) we find negative statistically significant spillovers from trend innovations (cycle innovations) to cycle (trend).

Regarding the cross-components spillovers for Okun's law, we find that: (i) GDP and unemployment exhibit statistically significant spillovers if the cyclical components are considered; (ii) the cycle-to-cycle spillover from GDP to unemployment is larger than the one of unemployment to GDP: the element $(A^{-1})_{5,4}$ is larger than $(A^{-1})_{4,5}$; (iii) the spillovers between GDP and unemployment are much smaller when the trend innovations are considered: there are no significant spillovers from output trend or output growth rate trend to unemployment trend, and only a small spillover from unemployment trend to GDP trend growth. As for the results for the Phillips curve, we find statistically significant cycle-to-cycle spillovers from GDP to inflation and vice versa —with the spillover from inflation cycle to GDP cycle being larger since $(A^{-1})_{4,5} > (A^{-1})_{5,4}$— and also statistically significant trend-to-trend spillovers —with the ones running from inflation trend to the permanent components of GDP being larger.[26]

Finally, the implied correlation between reduced-form shocks considering both the high and low volatility regimes is reported in Table 4. Compared with results estimated directly from the reduced-form model, which is reported in Appendix E, we observe that Υ_i, $i = 1, 2$ is smaller here. This is because we assume that *a priori* each structural shock affects other components in a similar way via the hierarchical prior. This is not a restriction but rather a more realistic assumption in order to facilitate the structural identification rather than the unstructured reduced-form model; and it is reassuring to corroborate that both correlation matrices have the same sign.

The results in Table 4 allow us to show more clearly the direction of correlation associated with the cross-components spillovers by linking these results to Tables 2 and 3. Consider the long-run Okun's law, illustrated via the negative correlation between

[26]Notice that other differences between GDP-unemployment and GDP-inflation interactions can also be found. For example, the element $A_{2,5}$ is significant both in Okun's law and the Phillips curve; however, $A_{1,5}$ is significant only in the Okun's law estimation. This suggests that unemployment and inflation cycles influence different long-run components of output. Note also that if we had modelled the change in trend growth rate as an exogenous structural break, we would not have been able to identify these effects.

Table 4: Correlation matrix of UC innovations obtained from the MSUC models[1]

	Okun's law					The Phillips curve				
	$\epsilon_{\tau_1,t}$	$\epsilon_{\tau_g,t}$	$\epsilon_{\tau_2,t}$	$\epsilon_{c_1,t}$	$\epsilon_{c_2,t}$	$\epsilon_{\tau_1,t}$	$\epsilon_{\tau_g,t}$	$\epsilon_{\tau_2,t}$	$\epsilon_{c_1,t}$	$\epsilon_{c_2,t}$
$\epsilon_{\tau_1,t}$	1	-0.12	0.16	-0.26	0.09	1	-0.46	0.02	-0.18	-0.47
$\epsilon_{\tau_g,t}$	-0.14	1	-0.04	-0.03	0.16	-0.45	1	0.24	-0.06	0.14
$\epsilon_{\tau_2,t}$	0.15	-0.06	1	-0.43	-0.24	0.03	0.26	1	0.16	-0.18
$\epsilon_{c_1,t}$	-0.25	0.02	-0.43	1	-0.59	-0.22	-0.04	0.12	1	0.66
$\epsilon_{c_2,t}$	0.10	0.10	-0.26	-0.61	1	-0.48	0.13	-0.20	0.69	1
Partial R-square[2]										
High vol.	0.09	0.05	0.57	0.71	0.66	0.39	0.34	0.31	0.58	0.65
Low vol.	0.08	0.04	0.62	0.75	0.71	0.38	0.33	0.30	0.59	0.66

[1] Reported is the correlation matrix Υ_i estimated from the MSUC model for Okun's law and the Phillips curve. The matrix is computed as $\Upsilon_i = (\text{diag}\Sigma_i)^{-1/2}\Sigma_i(\text{diag}\Sigma_i)^{-1/2}$ for volatility state $i = 1, 2$. The upper and lower triangular parts of the matrix indicate the correlation matrix under the high and low volatility regime, respectively. [2] Reported is the partial R-squared under both volatility regimes, computed from the inverse of correlation matrix. It measures the proportion of variation of one shock that is explained by other shocks.

shocks to trend growth rate $\epsilon_{\tau_g,t}$ and to trend unemployment rate $\epsilon_{\tau_2,t}$. By comparing $(A^{-1})_{3,2}$ and $(A^{-1})_{2,3}$, it is clear that the negative correlation in Table 4 is caused by a unit structural shock of trend unemployment rate leading to a 0.1 unit decrease in trend growth rate as shown in Table 3 (considering the estimated effect obtained from the two-volatility-state Markov regime switching), not the other way around. In the same vein, the long-run Phillips curve is present considering trend inflation and both trend growth rate and trend output, but the main channel is the positive spillover from structural shocks of trend inflation to trend growth rate. On the other hand, the short-run relationships are captured via $\Upsilon_{4,5} < 0$ (or $\Upsilon_{5,4} < 0$) for Okun's law and $\Upsilon_{4,5} > 0$ (or $\Upsilon_{5,4} > 0$) for the Phillips curve. However, comparing $(A^{-1})_{4,5}$ with $(A^{-1})_{5,4}$ for Okun's law we observe that the most important negative spillover comes from output cycle shocks to unemployment cycle; whereas for the Phillips curve the most important spillover is the positive effect of inflation cycle shocks to output cycle.

The bottom panel of Table 4 also reports the partial R-squared of reduced-form shocks in high and low volatility states. The latter is computed using the correlation matrix only (Anufriev and Panchenko, 2015), which measures how much variation of one shock is explained by the variation of other shocks. Comparing Okun's law with the Phillips

curve, shocks to inflation components tend to explain more variation in shocks to trend output and trend growth rate. Additionally, approximately 50% and 30% variation in shocks to trend unemployment and trend inflation, respectively, are explained by other shocks (including shocks to output cycle).

5 Implications for macro modelling and policy making

This section briefly discusses some of the most important implications derived from our findings for policy making and macroeconomic modelling. First, our results suggest the existence of statistically significant spillovers between permanent and transitory components for GDP in both directions, negative spillovers from permanent to transitory components for the unemployment rate, and negative spillovers from cycle to trend components for the inflation rate. These findings indicate that each variable under consideration exhibits its own relevant spillovers that are necessary to consider when evaluating the trade-offs of different economic policies. For example, the positive trend-to-cycle spillover for GDP indicates that a productivity shock that might permanently increase the level or growth rate of GDP would also affect positively the transitory movement of the GDP series, contradicting the strong negative correlation between permanent and transitory shocks for output found by Morley et al. (2003) and Sinclair (2009). Nevertheless, our analysis also shows that the observed negative correlation between trend innovations and cycles for output is derived from the strong negative spillover from cycle shocks to trend GDP. This highlights that the correlations obtained from the correlated UC model are insufficient to provide a detailed analysis of the structural interactions between permanent and transitory innovations.

Second, the finding that transitory structural shocks can influence negatively permanent components is related to the current discussion of hysteresis effects (Galí, 2015a; Blanchard et al., 2015). Broadly speaking, hysteresis refers to the possibility that transitory shocks (derived from different aggregate demand fluctuations) can affect permanent components (derived from different supply side fluctuations). Our results show the existence of negative spillovers from cycle-to-trend for GDP and the inflation rate. The presence of hysteresis effects in these variables has important implications for

the design of optimal monetary policy because, as Galí (2015a) and Galí (2016) discuss, hysteresis implies that real variables may experience permanent deviations from their efficient levels, even in response to shocks that are transitory. Moreover, as Ascari and Ropele (2009) discuss, much of the literature on monetary policy rules considers theoretical New Keynesian models that assume zero-inflation steady rate. They also show that, if a positive trend inflation is considered, then the structural equations and the determinacy region of these models experience substantial changes. In other words, a positive trend inflation drastically alters the Taylor principle. Our computations of the trend inflation rate show that the latter has been positive during the whole period, which indicates that the standard New Keynesian framework could be improved by explicitly incorporating the spillovers between transitory and permanent shocks.

Finally, our empirical results corroborate the slow recovery in the labour market and some of the discussions of the literautre on jobless recoveries (see, for example, Galí (2015a)). Figure 3 shows the posterior estimates of the trend and cycle components for Okun's law. The trend unemployment rate experiences a hump-shaped behaviour during the 1980's, while the trend growth rate fluctuates around a gradual decline. It is clear that both trend unemployment and its cycle increased during the GFC. A similar phenomenon can be observed by looking at Figure 4, which shows the estimates obtained from the MSUC model considering the Phillips curve. Both trend inflation and its cycle were affected by the GFC, and the latter seems to have affected the permanent component of output more than its cycle, compared to the results obtained from Okun's law. In order to provide a comprehensive explanation of these differences, it may be necessary to consider larger information sets with more macroeconomic variables instead of bivariate systems. This is beyond the scope of this paper and is left for future research.

6 Concluding remarks

We proposed a multivariate simultaneous unobserved components (MSUC) model that attributes the correlation in a correlated unobserved components model to the contemporaneous spillover effects between components, each driven by a combination of permanent and transitory structural shocks. Therefore, the direction of spillover has a structural interpretation, whose identification is achieved via heteroskedasticity. In order

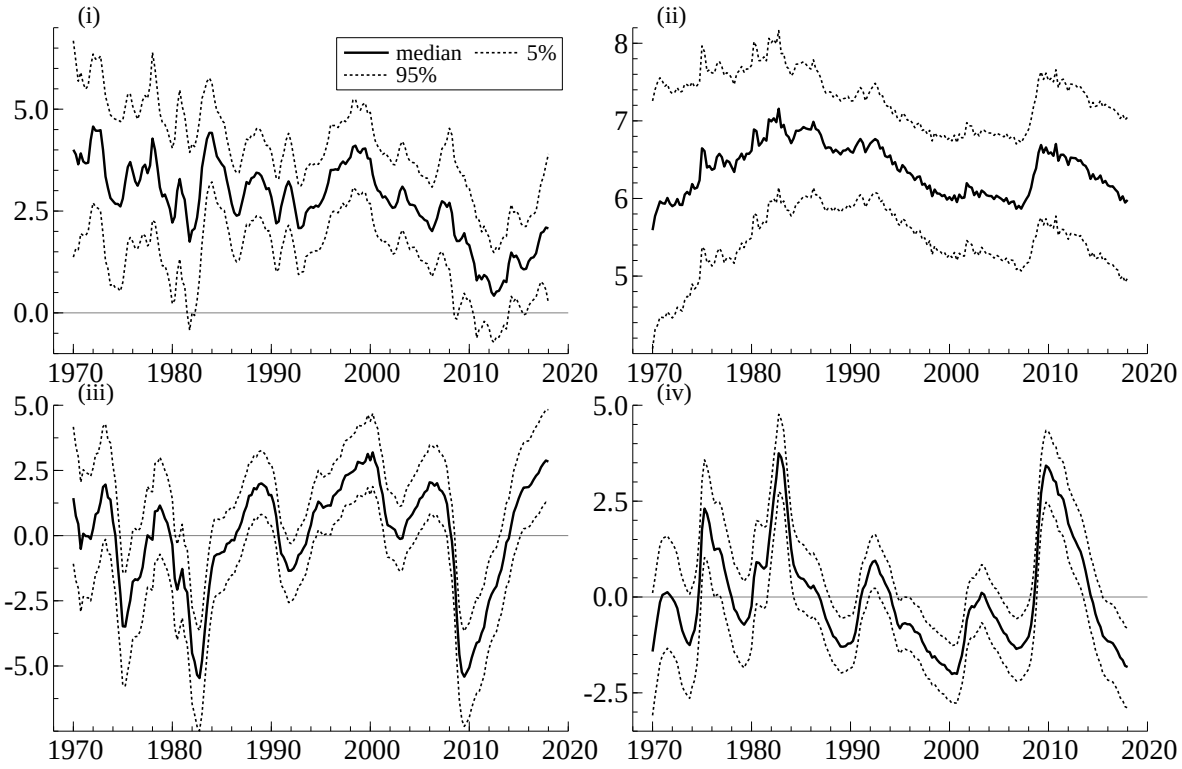

Figure 3: Estimated components of the MSUC model of Okun's law. The solid and dashed lines indicate the posterior mean and 5/95-th percentile, respectively of (i): Annual trend growth rate of output. (ii): Trend unemployment rate. (iii): Output cycle. (iv): Unemployment cycle.

to estimate the MSUC model, we developed an efficient Bayesian estimation procedure that maintains functionality despite the fact that the number of free elements in the contemporaneous structural matrix grows exponentially with dimensionality. In our empirical study, we estimate the MSUC versions of Okun's law and the Phillips curve, considering both Markov regime switching volatility states and stochastic volatility. The results show robust evidence of significant within-components and cross-components spillovers, thus indicating that it is inappropriate to treat permanent and transitory components as independent. Our findings shed new light regarding the policy debates on the output-inflation and output-unemployment interactions, and also regarding the trade-offs between short- and long-run. At the empirical level, future work may try to incorporate time-variation in the structural matrix under stochastic volatility; while at the theoretical level the development of general equilibrium models that allow for spillover effects between trends and cycles seems to be a promising avenue to capture relevant empirical dynamics.

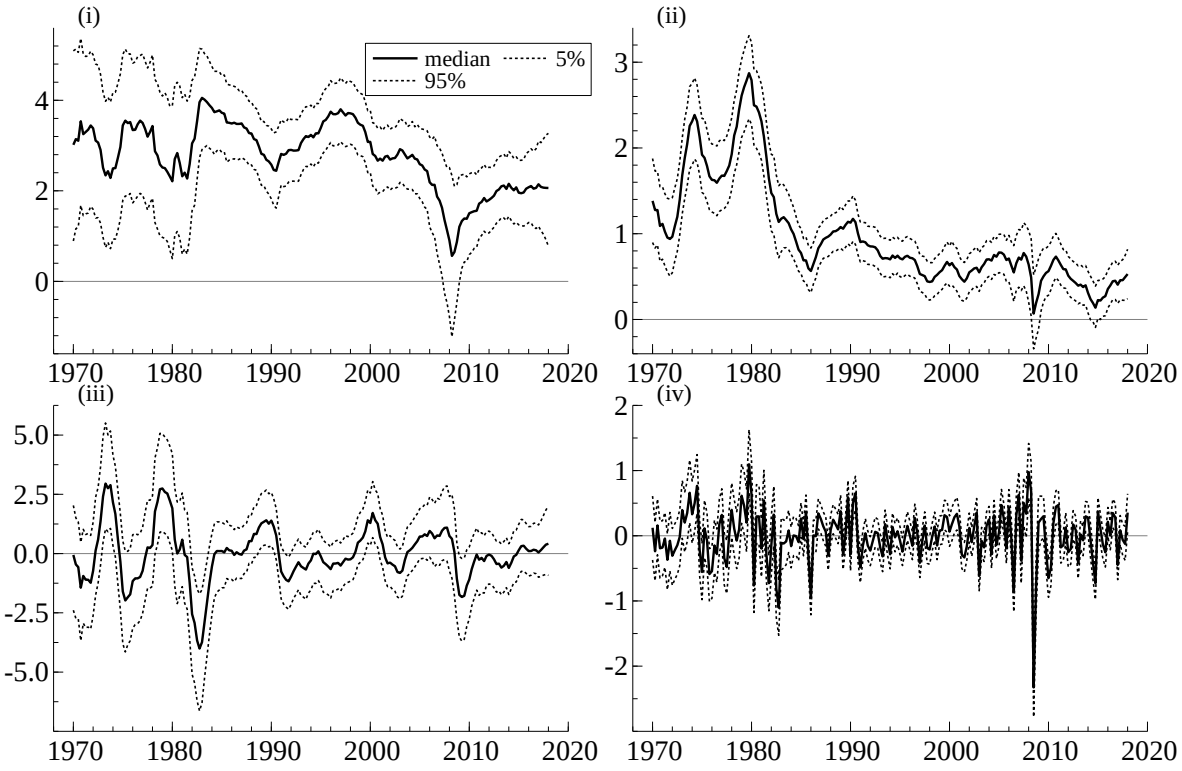

Figure 4: Estimated components of the MSUC model of the Phillips curve. The solid and dashed lines indicate the posterior mean and 5/95-th percentile, respectively of (i): Annual trend growth rate of output. (ii): Trend inflation. (iii): Output cycle. (iv): Inflation cycle.

References

Amisano, G. and C. Giannini (2012). *Topics in structural VAR econometrics*. Springer Science & Business Media.

Anufriev, M. and V. Panchenko (2015). Connecting the dots: Econometric methods for uncovering networks with an application to the Australian financial institutions. *Journal of Banking & Finance 61*, S241–S255.

Ascari, G. and T. Ropele (2009). Trend inflation, Taylor principle, and indeterminacy. *Journal of Money, Credit and Banking 41*(8), 1557–1584.

Athanasopoulos, G., D. S. Poskitt, F. Vahid, and W. Yao (2016). Determination of long-run and short-run dynamics in EC-VARMA models via canonical correlations. *Journal of Applied Econometrics 31*(6), 1100–1119.

Basistha, A. (2007). Trend-cycle correlation, drift break and the estimation of trend and cycle in Canadian gdp. *Canadian Journal of Economics/Revue canadienne d'économique 40*(2), 584–606.

Basistha, A. and R. Startz (2008). Measuring the nairu with reduced uncertainty: a multiple-indicator common-cycle approach. *The Review of Economics and Statistics 90*(4), 805–811.

Basturk, N., S. Grassi, L. F. Hoogerheide, A. Opschoor, and H. K. Van Dijk (2017). The R package MitISEM: efficient and robust simulation procedures for Bayesian inference. *Tinbergen Institute Discussion Paper*.

Bauwens, L., J.-F. Carpantier, and A. Dufays (2017). Autoregressive moving average infinite hidden Markov-switching models. *Journal of Business & Economic Statistics 35*(2), 162–182.

Bertsche, D., R. Braun, et al. (2018). Identification of structural vector autoregressions by stochastic volatility. *Department of Economics Working Paper, University of Konstanz*.

Beveridge, S. and C. R. Nelson (1981). A new approach to decomposition of economic time series into permanent and transitory components with particular attention to measurement of the 'business cycle'. *Journal of Monetary economics 7*(2), 151–174.

Blanchard, O., E. Cerutti, and L. Summers (2015). Inflation and activity–two explorations and their monetary policy implications. *National Bureau of Economic Research*, No. w21726.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*(4), 655–673.

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

Carter, C. K. and R. Kohn (1994). On gibbs sampling for state space models. *Biometrika 81*(3), 541–553.

Chan, J. C. and I. Jeliazkov (2009). Efficient simulation and integrated likelihood estimation in state space models. *International Journal of Mathematical Modelling and Numerical Optimisation 1*(1-2), 101–120.

Charemza, W. W. and E. M. Syczewska (1998). Joint application of the Dickey-Fuller and KPSS tests. *Economics Letters 61*(1), 17–21.

Chen, M.-H. (1994). Importance-weighted marginal Bayesian posterior density estimation. *Journal of the American Statistical Association 89*(427), 818–824.

Chib, S. (1996). Calculating posterior distributions and modal estimates in markov mixture models. *Journal of Econometrics 75*(1), 79–97.

De Jong, P. and N. Shephard (1995). The simulation smoother for time series models. *Biometrika 82*(2), 339–350.

Doan, T., R. Litterman, and C. Sims (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews 3*(1), 1–100.

Droumaguet, M., A. Warne, and T. Wozniak (2015). Granger causality and regime inference in Bayesian Markov-switching VARs. *ECB Working Paper* (1794).

Dufour, J.-M. and D. Pelletier (2005). Practical methods for modelling weak VARMA processes: identification, estimation and specification with a macroeconomic application. *Preprint*.

Dungey, M., J. P. Jacobs, J. Tian, and S. van Norden (2015). Trend in cycle or cycle in trend? New structural identifications for unobserved-components models of US real GDP. *Macroeconomic Dynamics 19*(4), 776–790.

Durbin, J. and S. J. Koopman (1997). Monte Carlo maximum likelihood estimation for non-Gaussian state space models. *Biometrika 84*(3), 669–684.

Durbin, J. and S. J. Koopman (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika 89*(3), 603–616.

Durbin, J. and S. J. Koopman (2012). *Time series analysis by state space methods*, Volume 38. Oxford University Press.

Fisher, L. A., H.-S. Huh, and A. R. Pagan (2016). Econometric methods for modelling systems with a mixture of I(1) and I(0) variables. *Journal of Applied Econometrics 31*(5), 892–911.

Frühwirth-Schnatter, S. and H. Wagner (2010). Stochastic model specification search for Gaussian and partial non-Gaussian state space models. *Journal of Econometrics 154*(1), 85–100.

Galí, J. (2015a). Hysteresis and the European unemployment problem revisited. *National Bureau of Economic Research Working Paper*.

Galí, J. (2015b). *Monetary policy, inflation, and the business cycle: an introduction to the new Keynesian framework and its applications*. Princeton University Press.

Galí, J. (2016). Insider-outsider labor markets, hysteresis and monetary policy. *Universitat Pompeu Fabra, Department of Economics Working Papers* (1506).

Harvey, A. (2011). Modelling the Phillips curve with unobserved components. *Applied Financial Economics 21*(1-2), 7–17.

Harvey, A. C. and N. Shephard (1993). Structural time series models. *Handbook of Statistics,(edited by GS Maddala, CR Rao and HD Vinod) 11*, 261–302.

Herwartz, H. and H. Lütkepohl (2014). Structural vector autoregressions with Markov switching: Combining conventional with statistical identification of shocks. *Journal of Econometrics 183*(1), 104–116.

Holston, K., T. Laubach, and J. C. Williams (2017). Measuring the natural rate of interest: International trends and determinants. *Journal of International Economics 108*, S59–S75.

Hwu, S.-T. and C.-J. Kim (2019). Estimating trend inflation based on unobserved components model: Is it correlated with the inflation gap? *Journal of Money, Credit and Banking*.

Justiniano, A. and G. E. Primiceri (2008). The time-varying volatility of macroeconomic fluctuations. *American Economic Review 98*(3), 604–41.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the american statistical association 90*(430), 773–795.

Keating, J. W. (2013a). Interpreting permanent shocks to output when aggregate demand may not be neutral in the long run. *Journal of Money, Credit and Banking 45*(4), 747–756.

Keating, J. W. (2013b). What do we learn from Blanchard and Quah decompositions of output if aggregate demand may not be long-run neutral? *Journal of Macroeconomics 38*, 203–217.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic volatility: likelihood inference and comparison with ARCH models. *The review of economic studies 65*(3), 361–393.

Kulikov, D. and A. Netšunajev (2013). *Identifying monetary policy shocks via heteroskedasticity: A Bayesian approach*. Eesti Pank.

Lanne, M., J. Luoto, et al. (2016). Data-driven inference on sign restrictions in Bayesian structural vector autoregression. Technical report, Department of Economics and Business Economics, Aarhus University.

Lanne, M., H. Lütkepohl, and K. Maciejowska (2010). Structural vector autoregressions with Markov switching. *Journal of Economic Dynamics and Control 34*(2), 121–131.

Laubach, T. and J. C. Williams (2003). Measuring the natural rate of interest. *Review of Economics and Statistics 85*(4), 1063–1070.

Li, M. and S. J. Koopman (2018). Unobserved components with stochastic volatility in us inflation: Estimation and signal extraction. *Tinbergen Institute Discussion Paper*.

Lütkepohl, H. (1984). Linear transformations of vector ARMA processes. *Journal of Econometrics 26*(3), 283–293.

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Springer Science & Business Media.

Lütkepohl, H. and A. Netšunajev (2017). Structural vector autoregressions with heteroskedasticity: A review of different volatility models. *Econometrics and statistics 1*, 2–18.

Lütkepohl, H. and T. Wozniak (2017). Bayesian inference for structural vector autoregressions identified by Markov-switching heteroskedasticity. *DIW Berlin German Institute for Economic Research Discussion Paper 1707*.

McCracken, M. W. and S. Ng (2016). FRED-MD: A monthly database for macroeconomic research. *Journal of Business & Economic Statistics 34*(4), 574–589.

Mitra, S. and T. M. Sinclair (2012). Output fluctuations in the G-7: An unobserved components approach. *Macroeconomic dynamics 16*(3), 396–422.

Morley, J. C. (2007). The slow adjustment of aggregate consumption to permanent income. *Journal of Money, Credit and Banking 39*(2-3), 615–638.

Morley, J. C., C. R. Nelson, and E. Zivot (2003). Why are the Beveridge-Nelson and unobserved-components decompositions of GDP so different? *Review of Economics and Statistics 85*(2), 235–243.

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

Oh, K. H., E. Zivot, and D. Creal (2008). The relationship between the Beveridge-Nelson decomposition and unobserved components models with correlated shocks. *Journal of Econometrics 146*, 207–219.

Perron, P. and T. Wada (2009). Let's take a break: Trends and cycles in US real GDP. *Journal of monetary Economics 56*(6), 749–765.

Shephard, N. (2015). Martingale unobserved component models. In S. J. Koopman and N. Shephard (Eds.), *Unobserved Components and Time Series Econometrics*, pp. Chapter 10. Oxford University Press.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: Journal of the Econometric Society*, 1–48.

Sinclair, T. M. (2009). The relationships between permanent and transitory movements in US output and the unemployment rate. *Journal of Money, Credit and Banking 41*(2-3), 529–542.

Stock, J. H. and M. W. Watson (1998). Median unbiased estimation of coefficient variance in a time-varying parameter model. *Journal of the American Statistical Association 93*(441), 349–358.

Stock, J. H. and M. W. Watson (2005). Implications of dynamic factor models for VAR analysis. *National Bureau of Economic Research Working Paper*.

Stock, J. H. and M. W. Watson (2007). Why has us inflation become harder to forecast? *Journal of Money, Credit and banking 39*, 3–33.

Trenker, C. and E. Weber (2016). On the identification of multivariate correlated unobserved components models. *Economics Letters 138*(1), 15–18.

Velinov, A. and W. Chen (2015). Do stock prices reflect their fundamentals? New evidence in the aftermath of the financial crisis. *Journal of Economics and Business 80*, 1–20.

Weber, E. (2011). Analyzing U.S. output and the great moderation by simultaneous unobserved components. *Journal of Money, Credit and Banking 43*(8), 1579–1597.

Wozniak, T. and M. Droumaguetb (2015). Assessing monetary policy models: Bayesian inference for heteroskedastic structural VARs. Technical report.

Appendix A Proof of Proposition 1

From the autocovariance function of the VARIMA$(p,1,d)$ representation, $\mathcal{X}$ only involves the coefficient matrices of the VAR(p) cycles. Let C_{N^2} denote the commutation matrix such that for any $N \times N$ matrix M the relation $C_{N^2}\text{vec}(M) = \text{vec}(M')$ holds; then $\mathcal{X}$ is given by

$$\mathcal{X} = \begin{bmatrix} D_N^+ \left(I_{N^2} + \sum_{i=1}^{p} B_{c,i} \otimes B_{c,i}\right) D_N & 2D_N^+ & 2D_N^+(I_{N^2} + I_N \otimes B_1) \\ \left(-I_N \otimes B_{c,1} + \sum_{i=1}^{p-1} B_{c,i} \otimes B_{c,i+1}\right) D_N & -I_{N^2} & -C_{N^2} - I_N \otimes B_{c,1} + I_N \otimes B_{c,2} \\ \left(-I_N \otimes B_{c,2} + \sum_{i=1}^{p-2} B_{c,i} \otimes B_{c,i+2}\right) D_N & \mathbf{0}_{N^2 \times N^2} & -I_N \otimes B_{c,2} + I_N \otimes B_{c,3} \\ \vdots & \vdots & \vdots \\ (-I_N \otimes B_{c,p-1} + B_{c,1} \otimes B_{c,p}) D_N & \mathbf{0}_{N^2 \times N^2} & -I_N \otimes B_{c,p-1} + I_N \otimes B_{c,p} \\ (-I_N \otimes B_{c,p}) D_N & \mathbf{0}_{N^2 \times N^2} & -I_N \otimes B_{c,p} \end{bmatrix}.$$

To show that the $(N^2p + \frac{1}{2}N^2 + \frac{1}{2}N) \times \frac{1}{2}(5N^2 + N)$ matrix $\mathcal{B}$ is of full rank, it suffices to show that via elementary row operations the transformed matrix has $\frac{1}{2}(5N^2+N)$ non-zero rows so that β is uniquely determined, and so is Ω. Notice that $\text{vec}(\Omega_c)'$ has $N(N-1)/2$ elements shown up twice; but since $\mathcal{Y}$ and $\mathcal{X}$ are implied by the correlated UC model, it automatically guarantees such a structure.

Let $\mathcal{X}_{ij}^+$ denote the ij-th block of matrix $\mathcal{X}$ detailed above, $i = 1, ..., p+1$ and $j = 1, 2, 3$; let $\mathcal{X}^*$ denote its transformation via elementary row operations with block $\mathcal{X}_{ij}^*$. If $\mathcal{X}^+$ has rank $\frac{1}{2}(5N^2 + N)$, we should be able to construct $\mathcal{X}^*$ such that $\mathcal{X}_{33}^*$ is of full rank N^2. According to Assumption 1, we can construct a coefficient sequence ρ_i for $i = 3, ..., p+1$ with

$$\rho_i = \begin{cases} -c_{i-1}, & \text{for } i = 3, \\ -\sum_{j=2}^{i-1} c_j, & \text{for } 4 \le i \le p+1. \end{cases}$$

This allows for $\mathcal{X}_{33}^*$ to be constructed by

$$\mathcal{X}_{33}^* = \sum_{i=3}^{p+1} \rho_i \mathcal{X}_{i3}^+ = I_N \otimes \bar{B},$$

which is of full rank N^2 under Assumption 1. So we align $\mathcal{X}_{3j}^*$ with $\sum_{i=3}^{p+1} \rho_i \mathcal{X}_{ij}^+$ for $j = 1, 2$. Then we construct

$$\mathcal{X}_{2j}^* = \mathcal{X}_{2j}^+ + \sum_{i=3}^{p+1} (\rho_i + 1)\mathcal{X}_{ij}^+, \quad j = 1, 2, 3.$$

And the first row block of $\mathcal{X}^*$ is the same as that of $\mathcal{X}^+$.

According to Morley et al. (2003) and Trenker and Weber (2016), $\mathcal{X}^*$ has a rank deficit only if there exists a $\frac{1}{2}(N+1)N \times 1$ vector f_1 and a $N^2 \times 1$ vector f_2 such that

$$(f_1'(\mathcal{X}^*_{13} + 2D_N^+\mathcal{X}^*_{23}) - f_2'\mathcal{X}^*_{33})D_N = \mathbf{0}_{1\times\frac{1}{2}N(N+1)}, \tag{18}$$

$$f_1'(\mathcal{X}^*_{11} + 2D_N^+\mathcal{X}^*_{21}) - f_2\mathcal{X}^*_{31} = \mathbf{0}_{1\times\frac{1}{2}N(N+1)}. \tag{19}$$

It can be easily verified that $\mathcal{X}^*_{13} + 2D_N^+\mathcal{X}^*_{23} = 2D_N^+\mathcal{X}^*_{33}$; so (18) gives $f_2' = 2f_1'D_N^+$. With the constructed full rank $\mathcal{X}^*_{33} = (I_N \otimes \bar{B})$, it follows from (19) that we need

$$f_1'\big(D_N^+[(I_N - \sum_{i=1}^{p} B_{c,i}) \otimes (I_N - \sum_{i=1}^{p} B_{c,i})]D_N\big) = \mathbf{0}_{1\times\frac{1}{2}N(N+1)}$$

to ensure rank deficit. But this is not possible under the condition that the VAR(p) cycles are stable. This means the rank of $\mathcal{X}^*$ must be full, and so is $\mathcal{X}^+$.

Appendix B Proof of Proposition 2

It follows from Assumption 1 that the following decomposition holds

$$\Omega_1^{-1} = C'C, \quad \Omega_2 = C'\text{diag}(\omega_{1,2}, ..., \omega_{K,2})C.$$

Lanne et al. (2010) shows that C, with diagonal elements $c_{11}, ..., c_{KK}$, is unique up to row order and sign swap if and only if for all $i, j \in \{1, ..., K\}$, $i \neq j$ we have $\omega_{i,2} \neq \omega_{j,2}$.

This result carries over to our setting by defining $A = \text{diag}(c_{11}, ..., c_{KK})^{-1}C$ so that it has unit diagonal. This means that A is a normalised version of C and its row ordering and sign are determined since the normalisation is done via its diagonal elements. By replacing C with $\text{diag}(c_{11}, ..., c_{KK})A$ we have $\Omega_i^{-1} = A'\Sigma_i^{-1}A$ for $i = 1, 2$, or $\Omega_i = A^{-1}\Sigma_i A^{-1'}$.

Appendix C Proof of Proposition 3

Similar to above, writing $\Omega_1^{-1} = C'C$ yields

$$\Omega_t = C'\text{diag}(\sigma^2_{1,t}/\sigma^2_{1,1}, ..., \sigma^2_{K,t}/\sigma^2_{K,1})C, \quad t = 2, ..., T.$$

Based on Bertsche et al. (2018), C with diagonal elements $c_{11}, ..., c_{KK}$ is identified up to row order and sign swap *a.s.* under the random walk specification. Defining $A = \text{diag}(c_{11}, ..., c_{KK})^{-1}C$, so that it has unit diagonal, and replacing C by $\text{diag}(c_{11}, ..., c_{KK})A$, we have $\Omega_t^{-1} = A'\Sigma_t^{-1}A$. Uniqueness is achieved by respecting that A is a normalised version of C and its row ordering and sign are determined. The result also carries over to the case where there exists one structural shock having constant volatility, as proven by Bertsche et al. (2018).

Appendix D Details of the sampling procedure

Sampling $\delta_{i,j}$.

Firstly, notice that if $\delta_{i,j}$ is $N(0, \gamma_j)$-distributed, a new draw $\delta_{i1}^* = (\delta_{i,1}, ..., \delta_{i,i-1}, \delta_{i,i+1}, ..., \delta_{i,N+i-1}, \delta_{N+i+1}, ..., \delta_{i,k})'$ can be generated from $N(\mu_{i1}^\delta, \Psi_{i1}^\delta)$, where

$$\Psi_{i1}^\delta = \left(\sum_{t=2}^{T} \frac{\tilde{x}_t \tilde{x}_t'}{(\delta_{i,N+i} - 1)^2 \sigma_{i,s_{t-1}}^2 \omega_{i,s_{t-1}}} + \text{diag}(\gamma_1, ...\gamma_{i-1}, \gamma_{i+1}, ...\gamma_{N+i-1}, \gamma_{N+i+1}, ..., \gamma_K) \right)^{-1},$$

$$\mu_{i1}^\delta = \Psi_{i1}^\delta \sum_{t=2}^{T} \frac{\tilde{x}_t(\tau_{i,t}^* - \delta_{i,N+i} y_{i,t}^*)}{(\delta_{i,N+i} - 1)^2 \sigma_{i,s_{t-1}}^2 \omega_{i,s_{t-1}}}.$$

If the initialisation of unobserved components is omitted, the new draw is accepted with probability of one, but not from the correct conditional posterior. Notice that in (10), B_c^* and Ω_i, $i = 1, 2$ are functions of A, and thus of δ_{i1}^* conditional on other parameters, so with initialisation considered, the draw is accepted with probability

$$\min \left\{ \frac{|V_{c,1}(\delta_{i1}^*)|^{-1/2} \exp(-\frac{1}{2} x_{c,1}' V_{c,1}(\delta_{i1}^*)^{-1} x_{c,1})}{|V_{c,1}(\delta_{i1})|^{-1/2} \exp(-\frac{1}{2} x_{c,1}' V_{c,1}(\delta_{i1})^{-1} x_{c,1})}, 1 \right\},$$

where δ_{i1} (without asterisk) is the previous draw in the Markov chain and $x_{c,1}$ denotes the cycle components in x_1, *i.e.* $x_{c,1} = (c_1, ..., c_{2-p})'$.

Secondly, defining $\tilde{y}_{i,t} = \tau_{i,t}^* - \delta_{i,1}\tau_{1,t} - ... - \delta_{i,i-1}\tau_{i-1,t} - \delta_{i,i+1}\tau_{i+1,t} - ... - \delta_{i,N+1}c_{1,t} - ... - \delta_{i,K}c_{N,t}$, (12) becomes

$$\tilde{y}_{i,t} = \delta_{i,N+1} y_{i,t}^* + (\delta_{i,N+i} - 1) e_{i,t-1}.$$

Ignoring initialisation, the conditional posterior follows

$$p(\delta_{i,N+1}|\boldsymbol{\cdot}) \propto (\delta_{i,N+1}-1)^{-T+1} \exp\left(-\frac{1}{2}\sum_{t=2}^{T}\frac{(\tilde{y}_{i,t}-\delta_{i,N+i}y^*_{i,t})^2}{(\delta_{i,N+i}-1)^2\sigma^2_{i,s_{t-1}}\omega_{i,s_{t-1}}} - \frac{1}{2}\frac{\delta^2_{i,N+i}}{\gamma_{N+i}}\right). \quad (20)$$

This non-standard univariate distribution can be well-approximated by a Student's t-proposal with mean μ^{δ}_{i2} and scale parameter Ψ^{δ}_{i2} equal to the mode and curvature around the mode of the above density function.[27] That is, we apply the Newton's method

$$\delta^{(n+1)}_{i,N+i} = \delta^{(n)}_{i,N+i} - \frac{p'(\delta^{(n)}_{i,N+i}|\boldsymbol{\cdot})}{p''(\delta^{(n)}_{i,N+i}|\boldsymbol{\cdot})}$$

to find the mode iteratively, with $\delta^{(1)}_{i,N+i} = \mathrm{Cov}(\tilde{y}_{i,t}, y^*_{i,t})/\mathrm{Var}(y^*_{i,t})$ and

$$\begin{aligned}
p'(\delta_{i,N+i}|\boldsymbol{\cdot}) &= -(T-1)\frac{1}{\delta_{i,N+i}-1} + (\delta_{i,N+i}-1)^{-3}\sum_{t=2}^{T}\frac{(\tilde{y}_{i,t}-\delta_{i,N+1}y^*_{i,t})^2}{\sigma^2_{i,s_{t-1}}\omega_{i,s_{t-1}}} \\
&\quad + (\delta_{i,N+i}-1)^{-2}\sum_{t=2}^{T}\frac{(\tilde{y}_{i,t}-\delta_{i,N+1}y^*_{i,t})y^*_{i,t}}{\sigma^2_{i,s_{t-1}}\omega_{i,s_{t-1}}} - \frac{1}{\gamma_{N+i}}\delta_{i,N+i}, \\
p''(\delta_{i,N+i}|\boldsymbol{\cdot}) &= (T-1)(\delta_{i,N+i}-1)^{-2} - 3(\delta_{i,N+i}-1)^{-4}\sum_{t=2}^{T}\frac{(\tilde{y}_{i,t}-\delta_{i,N+1}y^*_{i,t})^2}{\sigma^2_{i,s_{t-1}}\omega_{i,s_{t-1}}} \\
&\quad - 4(\delta_{i,N+i}-1)^{-3}\sum_{t=2}^{T}\frac{(\tilde{y}_{i,t}-\delta_{i,N+1}y^*_{i,t})y^*_{i,t}}{\sigma^2_{i,s_{t-1}}\omega_{i,s_{t-1}}} - \frac{1}{\gamma_{N+i}},
\end{aligned}$$

until some convergence criterion is met. Let μ^{δ}_{i2} denote the mode and define $\Psi^{\delta}_{i2} = -1/p''(\mu^{\delta}_{i2}|\boldsymbol{\cdot})$. A new draw $\delta^*_{i,N+i}$ is generated from a Student' t-distribution $T(\mu^{\delta}_{i2}, \Psi^{\delta}_{i2}, \nu)$, with the degrees of freedom ν arbitrarily chosen. The draw is accepted with probability

$$\min\left\{\frac{|V_{c,1}(\delta^*_{i,N+i})|^{-1/2}\exp\left(-\frac{1}{2}x'_{c,1}V_{c,1}(\delta^*_{i,N+i})^{-1}x_{c,1}\right)p(\delta^*_{i,N+i}|\boldsymbol{\cdot})T(\delta_{i,N+i};\mu^{\delta}_{i2},\Psi^{\delta}_{i2},\nu)}{|V_{c,1}(\delta_{i,N+i})|^{-1/2}\exp\left(-\frac{1}{2}x'_{c,1}V_{c,1}(\delta_{i,N+i})^{-1}x_{c,1}\right)p(\delta_{i,N+i}|\boldsymbol{\cdot})T(\delta^*_{i,N+i};\mu^{\delta}_{i2},\Psi^{\delta}_{i2},\nu)}, 1\right\},$$

where $p(\delta_{i,N+i}|\boldsymbol{\cdot})$ is the density kernel in (20) and $T(\delta_{i,N+i};\mu^{\delta}_{i2},\Psi^{\delta}_{i2},\nu)$ is the density of constructed Student's t-proposal evaluated at $\delta_{i,N+i}$. Once $\delta_{i,j}$ and $\delta_{i,N+i}$ are generated, A_{i-} is computed using (13).

[27] More efficient proposals such as a mixture of Student's t-distributions can be easily constructed (Basturk et al., 2017) due to the fact that we deal with a univariate distribution. We find that a simple Student's t-proposal suffices.

Sampling Φ, $\sigma^2_{i,1}$, $\omega_{i,2}$, $\mathtt{P}$, γ_Φ *and* γ_i.

Φ is done independently equation-by-equation. Based on (4), we notice that only the sampling of Φ_{cc} is needed (*i.e.* the autoregressive coefficient matrix for the VAR(p) cycles). Using (14), standard Bayesian calculation shows that posterior draws of Φ'_{i-} can be generated from $N(A_{i-}\mu_i^\Phi, \Psi_i^\Phi)\mathbf{1}_{\{||B_c^*||<1\}}$, where

$$\Psi_i^\Phi = \left(\sum_{t=1}^{T-1} x_t'\Sigma_{s_t}^{-1}x_t + \frac{1}{\gamma_\Phi}L^{-1}\right)^{-1},$$
$$\mu_i^\Phi = \Psi_i^\Phi\left(\sum_{t=1}^{T-1} x_t'\Sigma_{s_t}^{-1}\dot{x}_{t+1} + \frac{1}{\gamma_\Phi}L^{-1}W\right),$$

and $||B_c^*||$ denotes the largest eigenvalue in absolute value of B_c^* so that the indicator function $\mathbf{1}_{\{||B_c^*||<1\}}$ guarantees that the VAR(p) cycles are stationary. Taking initialisation into account, the draw is accepted with probability

$$\min\left\{\frac{|V_{c,1}(\Phi_{i-}^*)|^{-1/2}\exp(-\frac{1}{2}x_{c,1}'V_{c,1}(\Phi_{i-}^*)^{-1}x_{c,1})}{|V_{c,1}(\Phi_{i-})|^{-1/2}\exp(-\frac{1}{2}x_{c,1}'V_{c,1}(\Phi_{i-})^{-1}x_{c,1})}, 1\right\}.$$

For $i = 1, ..., K$, the volatility parameter $\sigma^2_{i,1}$ can be sampled from

$$IG\left(\alpha_v + \frac{T}{2}, \beta_v + \sum_{t=1}^{T-1}\frac{(A_{i-}\dot{x}_{t+1} - \Phi_{i-}x_t)^2}{2\omega_{i,s_t}}\right)$$

and the variance ratio $\omega_{i,2}$ can be sampled from

$$IG\left(\alpha_\omega + \frac{T_2}{2}, \beta_\omega + \sum_{t\in T_2}\frac{(A_{i-}\dot{x}_{t+1} - \Phi_{i-}x_t)^2}{2\sigma^2_{i,s_t}}\right).$$

These draws are accepted with probability

$$\min\left\{\frac{|V_{c,1}(\sigma^{2*}_{i,1}\omega^*_{i,j})|^{-1/2}\exp(-\frac{1}{2}x_{c,1}'V_{c,1}(\sigma^{2*}_{i,1}\omega^*_{i,j})^{-1}x_{c,1})}{|V_{c,1}(\sigma^2_{i,1}\omega_{i,j})|^{-1/2}\exp(-\frac{1}{2}x_{c,1}'V_{c,1}(\sigma^2_{i,1}\omega_{i,j})^{-1}x_{c,1})}, 1\right\}, \quad j = 1, 2.$$

The posterior draws of transition probability are sampled via a 2-dimensional Dirichlet distribution. Specifically, we draw $\mathtt{P}'_{1-}$ and $\mathtt{P}'_{2-}$ from

$$Dir_2(e_1+\sum_{t=2}^{T}\mathbf{1}_{s_{t-1}=1,s_t=1}, e_2+\sum_{t=2}^{T}\mathbf{1}_{s_{t-1}=1,s_t=2}), \quad Dir_2(e_2+\sum_{t=2}^{T}\mathbf{1}_{s_{t-1}=2,s_t=1}, e_1+\sum_{t=2}^{T}\mathbf{1}_{s_{t-1}=2,s_t=2}),$$

respectively, where $\sum_{t=2}^{T} \mathbf{1}_{s_{t-1}=j, s_t=i}$ counts the number of transitions from volatility regime j to i. The new draw is accepted with probability

$$\min \left\{ \frac{|V_{c,1}(\mathtt{P}^*)|^{-1/2} \exp(-\frac{1}{2} x'_{c,1} V_{c,1}(\mathtt{P}^*)^{-1} x_{c,1})}{|V_{c,1}(\mathtt{P})|^{-1/2} \exp(-\frac{1}{2} x'_{c,1} V_{c,1}(\mathtt{P})^{-1} x_{c,1})}, 1 \right\}.$$

Based on the updated transition probability, the index process of two Markov regimes S_T is sampled using the forward filter and backward simulation smoother of Chib (1996) with initialisation taken into account. Through $e_t = A\dot{x}_{t+1} - \Phi x_t$ and $e_t \sim N(0, \Sigma_{s_t})$, the algorithm utilises $p(S_T|e_1, ..., e_T) = \prod_{t=1}^{T} p(s_t|Y_T, s_{t+1}, ..., s_T, \mathtt{P})$ and $p(s_t|Y_T, s_{t+1}, ..., s_T, \mathtt{P}) \propto p(s_t|e_t)p(s_{t+1}|s_t, \mathtt{P})$ by a forward recursion that determines $p(s_t|e_t)$ and a backward recursion that draws posterior samples of S_T.

Finally, the posterior draws of shrinkage parameters are directly generated from an inverse gamma distribution due to conjugacy. We draw γ_Φ from $IG(\alpha_\gamma + \frac{K^2 p}{2}, \beta_\gamma + \frac{1}{2}\sum_{i=1}^{K}(\Phi_{i-} - A_{i-}W)(\Phi_{i-} - A_{i-}W)')$ and γ_i from $IG(\alpha_\gamma + \frac{K-1}{2}, \beta_\gamma + \frac{1}{2}A'_{-i}A_{-i})$ for $i = 1, ..., K$.

Appendix E Complementary empirical results

E.1 Reduced-form correlations

In this section we investigate the reduced-form models for Okun's law and the Phillips curve. Assuming constant volatility, *i.e.* $\Sigma_{s_t} = \Sigma$ for all t, the reduced-form dynamics for the unobserved components is given by

$$\begin{pmatrix} \tau_{t+1} \\ c_{t+1} \end{pmatrix} = \begin{bmatrix} B_\tau & 0 \\ 0 & B_1 \end{bmatrix} \begin{pmatrix} \tau_t \\ c_t \end{pmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & B_2 \end{bmatrix} \begin{pmatrix} \tau_{t-1} \\ c_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{\tau,t} \\ \epsilon_{c,t} \end{pmatrix}, \quad \epsilon_t \sim N(0, \Omega),$$

with 5×5 full covariance matrix Ω and B_τ as in Laubach and Williams (2003) and Holston et al. (2017). For simplicity, we assume that B_1 and B_2 are 2×2 diagonal matrices. The identification condition 1 is satisfied as long as $B_2 \neq -B_1$. Morley (2007), Oh et al. (2008), Weber (2011) and Trenker and Weber (2016) call this a correlated UC model, which is a linear and Gaussian state space model. We estimate this specification using Kalman filter and maximum likelihood.

Table 5 shows the estimated correlation matrix Υ implied by the covariance matrix

Table 5: Correlation matrix of UC innovations from reduced-form models

	Okun's law					Phillips curve				
	$\epsilon_{\tau_1,t}$	$\epsilon_{\tau_g,t}$	$\epsilon_{\tau_2,t}$	$\epsilon_{c_1,t}$	$\epsilon_{c_2,t}$	$\epsilon_{\tau_1,t}$	$\epsilon_{\tau_g,t}$	$\epsilon_{\tau_2,t}$	$\epsilon_{c_1,t}$	$\epsilon_{c_2,t}$
$\epsilon_{\tau_1,t}$	1	-0.19	-0.65**	0.45**	0.87**	1	-0.15	0.19*	-0.72**	-0.42**
$\epsilon_{\tau_g,t}$		1	-0.43**	0.60**	0.67**		1	0.12	-0.26*	0.24**
$\epsilon_{\tau_2,t}$			1	-0.19*	-0.63**			1	0.76**	-0.16*
$\epsilon_{c_1,t}$				1	-0.88**				1	0.36**
$\epsilon_{c_2,t}$					1					1

[1] We report the correlation matrix Υ estimated from a correlated unobserved components model for Okun's law and the Phillips curve. The covariance matrix Σ of the innovations is unconstrained, and we compute $\Upsilon = (\text{diag}\Sigma)^{-1/2}\Sigma(\text{diag}\Sigma)^{-1/2}$ where diagΣ is a diagonal matrix with the diagonal of Σ.
[2] Estimation is done via Kalman filter and maximum likelihood. ** and * indicate statistical significance at the 5% and 10% levels, respectively.

Σ for both specifications. With respect to Okun's law, Sinclair (2009) also estimates a correlated UC model, but without considering the permanent shocks $\epsilon_{g,t}$. She finds that both within-series and cross-series trend-cycle correlations are high in absolute values, all exceeding 0.85. Specifically, the cross-series correlation between permanent components of output and unemployment rate is -0.96, and the one between transitory components is -0.98; while the within-series correlations between permanent and transitory components are -0.85 for the GDP series and -0.97 for the unemployment rate series, respectively.

Different from Sinclair (2009), our specification shows that the long-run Okun's law (that is, the cross-series correlation between permanent components of output and unemployment) is attributed to how innovations of trend unemployment $\epsilon_{\tau_2,t}$ are correlated to both the trend growth rate $\epsilon_{\tau_g,t}$ and trend output $\epsilon_{\tau_1,t}$; while the short-run Okun's law is captured by the cross-series correlation between transitory components of output and unemployment. The left panel of Table 5 shows that both long-run and short-run Okun's laws are smaller in magnitude than the results found by Sinclair (2009). Two further results stand out: (i) in contrast to Sinclair (2009) and Morley et al. (2003), we find positive within-series trend-cycle correlations for GDP; (ii) in line with Sinclair (2009), we find a negative within-series trend-cycle correlation for the unemployment rate, although the coefficient is also smaller in our results.

On the other hand, Harvey (2011) estimates a simplified version of our UC Phillips curve by imposing zeros on all off-diagonal elements of Υ, except for $\Upsilon_{4,5} = \Upsilon_{5,4}$.[28]

[28] His model can be considered as structural because innovations to trends and cycles are orthogonal.

From the right panel of Table 5, we observe evidence suggesting a long-run Phillips curve between trend inflation and trend growth rate of output (trend output) of approximately $\Upsilon_{2,3} = 0.12$ ($\Upsilon_{1,3} = 0.19$).[29] We also find that there exists a significant short-run Phillips curve coefficient between inflation cycle and output cycle with $\Upsilon_{4,5} = 0.36$. Finally, the within-series trend-cycle correlations both for GDP and the inflation rate are negative, although only the former is statistically significant.

Figures 5 and 6 show the extracted components using the Kalman smoother under the assumptions of unrestricted covariance matrix and restricted covariance matrix —which only allows for a short-run Okun's law or Phillips curve. The restricted models deliver estimated components that are smoother than the unrestricted models, because the latter omits the potential interactions between permanent and transitory shocks. As Oh et al. (2008) and Sinclair (2009) show, it is possible that the estimated components have larger volatility than the time series itself if the components are negatively correlated. This is clear if we compare the estimated trend growth rate from both models: unrestricted models yield more volatile trends that are expected to be affected by other shocks than the restricted one.[30] Likewise, the unrestricted model for Okun's law yields a very volatile trend unemployment rate (see also Sinclair 2009), following the time series of unemployment rate closely. This could be counter-intuitive, because it implies that unemployment is largely composed of a permanent component, leaving little room for cyclical fluctuations.

E.2 Robustness checks

Figures 7 and 8 show the posterior estimates of the time evolution of the volatility of structural shocks in our MSUC versions of Okun's law and the Phillips curve, identified via stochastic volatility. The stochastic volatility specification serves as a robustness check

Although the cycles are correlated via $\Upsilon_{4,5}$, the regression lemma between two Gaussian variables indicates that the effect of output cycle on inflation cycle is determined by $\Upsilon_{4,5}\sigma_{c_2}/\sigma_{c_1}$.

[29] Notice, however, that only $\Upsilon_{1,3}$ is statistically significant.

[30] It is worth mentioning that the $\tau_{g,t}$ obtained from the restricted model suggests very smooth transitions around 2008. First, this may suggest that the trend growth rate of output has been pushed down by the Great Financial Crisis. Second, the time path smoothness is likely due to the pathology associated with underestimating the signal-to-noise ratio of the trend component in structural state space time series models using maximum likelihood, as pointed out by Stock and Watson (1998). This, however, is not a concern in correlated UC models with unrestricted covariance matrix, because other system's shocks spill over to the trend and affect its variation, which means that we can detect enough variation in long-run components.

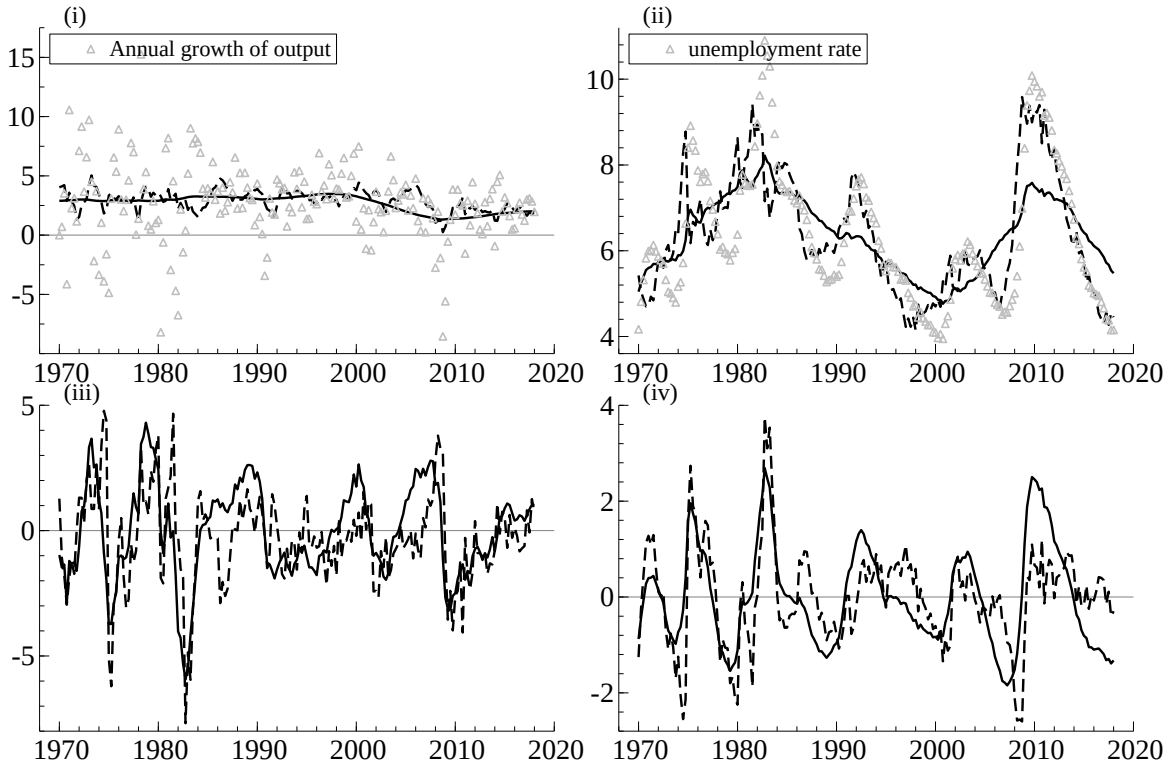

Figure 5: Reduced-form UC model for Okun's law with unrestricted and restricted covariance matrix. (i): Annual trend growth rate of output. Triangles represent the annual growth rate of output, *i.e.* $4\Delta y_t$. (ii): Trend unemployment rate. Triangles represent the actual unemployment rate. (iii): Output cycle. (iv): Unemployment cycle. Straight lines show the estimates from the UC model with restricted covariance matrix; dashed lines indicate the estimates from the correlated UC model (*i.e.*, UC model with unrestricted covariance matrix).

for the MSUC model identified via the two-volatility-state Markov switching regime.

Table 6 reports the posterior mean estimate of the structural matrix A in the MSUC model using the data up until the GFC. This serves as a robustness check in order to corroborate if the inclusion of GFC period introduces any changes in A. The results do not suggest significant differences from Table 2, which comes from the fact that even if the high volatility brought about by the GFC is ignored, the data still shows two distinct volatility states, namely the high volatility state in the 1980s and the low volatility state after the "Great Moderation".

Table 6: Estimated contemporaneous structural matrix prior to the Great Financial Crisis

	Okun's law					The Phillips curve				
Two volatility states Markov regime switching										
	$\tau_{1,t}$	$\tau_{g,t}$	$\tau_{2,t}$	$c_{1,t}$	$c_{2,t}$	$\tau_{1,t}$	$\tau_{g,t}$	$\tau_{2,t}$	$c_{1,t}$	$c_{2,t}$
$\tau_{1,t}$	1	0.02 −1.4	0.02 0.4	0.67 **−17.8**	0.27 **−21.4**	1	0.17 −3.6	−0.04 −8.4	0.46 **−24.8**	0.10 −4.2
$\tau_{g,t}$	0.03 −1.1	1	−0.09 **−16.6**	0.02 1.3	0.21 **−22.3**	0.14 **−11.2**	1	−0.55 **−30.4**	0.26 **−14.2**	−0.22 **−17.3**
$\tau_{2,t}$	0.02 1.2	0.08 −2.7	1	0.35 **−32.7**	0.13 −8.6	0.09 −3.0	0.27 **−18.3**	1	−0.38 **−12.7**	0.22 **−15.8**
$c_{1,t}$	−0.25 **−36.9**	−0.20 **−14.4**	0.30 **−27.3**	1	0.24 **−20.8**	−0.37 **−23.6**	−0.22 **−19.5**	0.25 **−31.4**	1	−0.61 **−28.1**
$c_{2,t}$	0.03 0.2	−0.04 −3.2	0.66 **−17.7**	0.58 **−18.4**	1	0.22 **−25.6**	−0.01 −0.1	0.09 −4.4	−0.23 **−21.0**	1
Stochastic volatility										
$\tau_{1,t}$	1	0.02 −0.0	−0.05 0.3	0.61 **−25.6**	0.31 **−12.8**	1	0.10 −6.4	−0.06 3.4	0.37 **−33.2**	0.04 −2.3
$\tau_{g,t}$	0.08 −4.9	1	−0.14 **−16.1**	0.05 2.3	0.30 **−33.4**	0.08 −9.2	1	−0.52 **−18.2**	0.34 **−26.9**	−0.30 **−11.5**
$\tau_{2,t}$	0.04 −2.0	0.05 0.8	1	0.28 **−14.3**	0.11 −4.5	0.11 −6.8	0.22 **−32.7**	1	−0.42 **−24.1**	0.31 **−30.3**
$c_{1,t}$	−0.17 **−17.5**	−0.14 **−21.6**	0.33 **−11.5**	1	0.18 **−12.4**	−0.23 **−18.2**	−0.21 **−24.0**	0.18 **−22.9**	1	−0.54 **−33.8**
$c_{2,t}$	−0.00 0.2	−0.02 0.3	0.47 **−11.8**	0.55 **−19.8**	1	0.20 −2.3	−0.04 1.1	0.14 −9.2	−0.18 **−17.2**	1

Reported is the posterior mean estimate of the structural matrix A obtained from the MSUC model for Okun's law and the Phillips curve, using data prior to the GFC. Below the posterior mean is the SDDR for testing $H_0 : A_{i,j} = 0$, with boldface number indicating strong evidence against H_0. The upper panel shows estimates identified considering the two-volatility-state Markov regime switching, while the bottom panel shows estimates identified considering stochastic volatility.

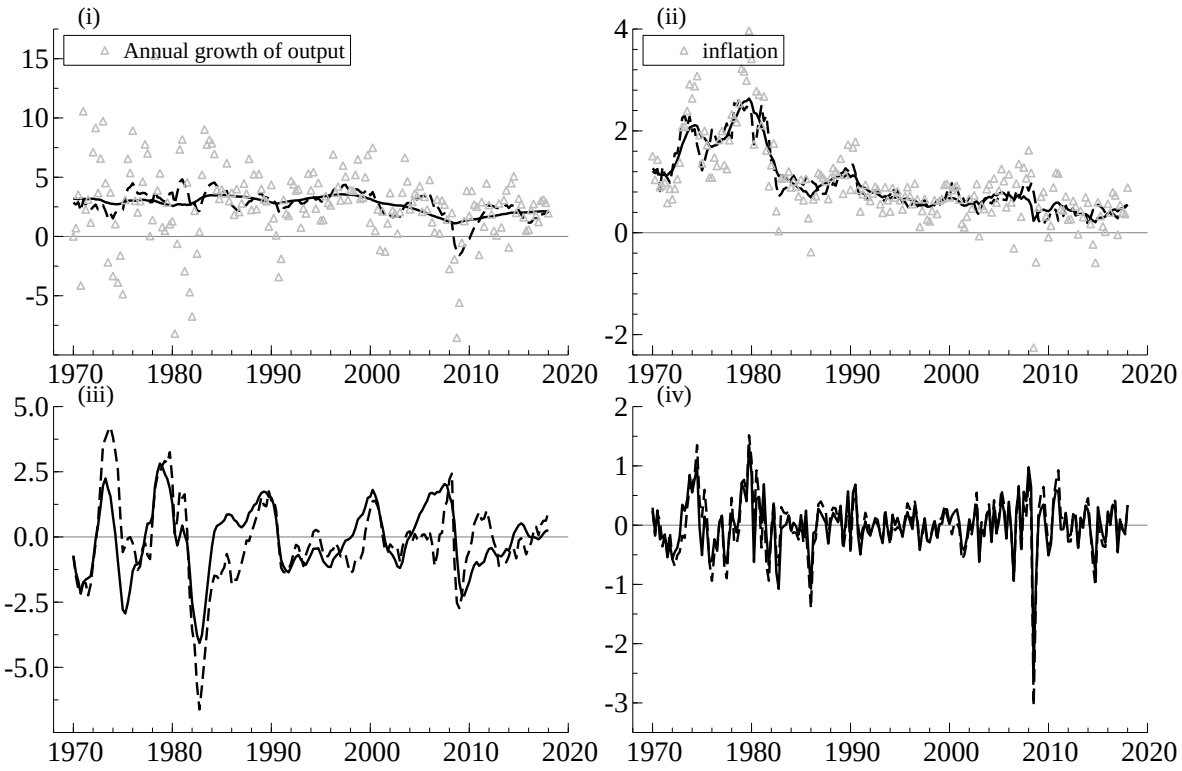

Figure 6: Reduced-form UC model for the Phillips curve with unrestricted and restricted covariance matrix. (i): Annual trend growth rate of output. Triangles represent the annual growth of output, *i.e.* $4\Delta y_t$. (ii): Trend inflation. Triangles represent the CPI headline inflation rate. (iii): Output cycle. (iv): Inflation cycle. Straight lines show the estimates from the UC model with restricted covariance matrix; dashed lines indicate the estimates from the correlated UC model (*i.e.*, UC model with unrestricted covariance matrix).

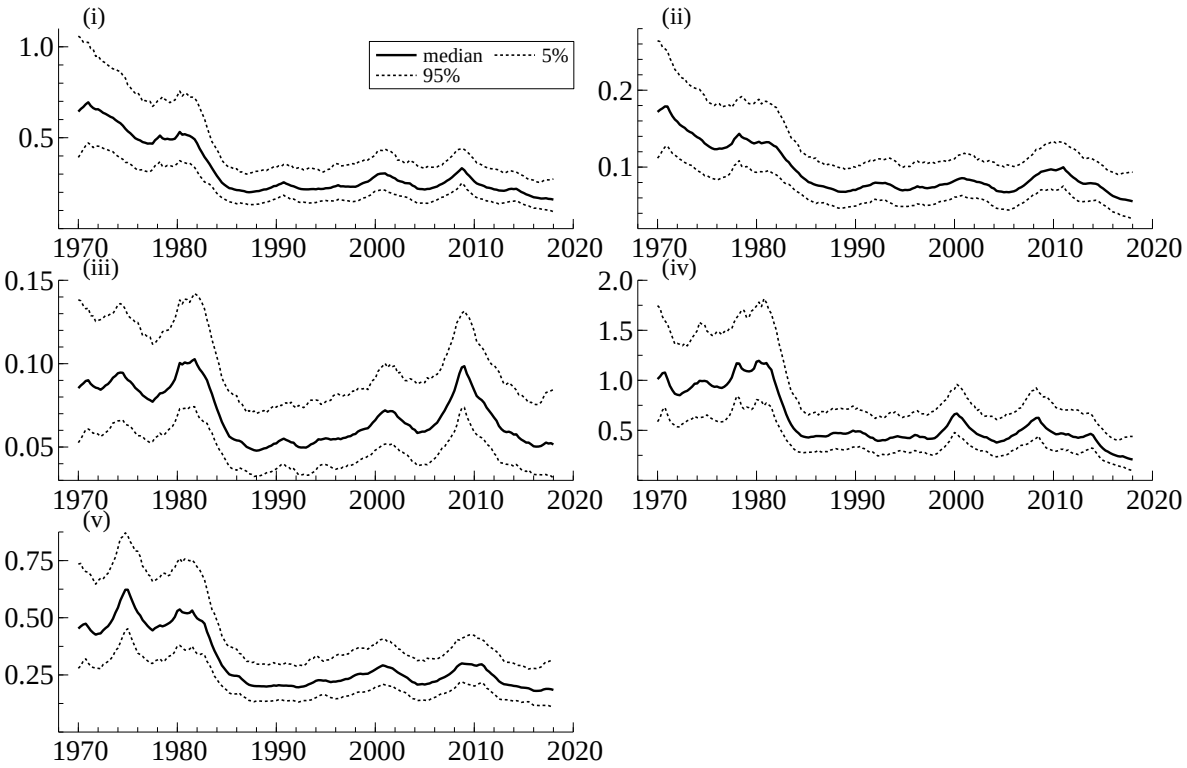

Figure 7: Estimated stochastic volatility of the MSUC model. The solid and dashed lines indicate the posterior median and the 5-th/95-th percentile, respectively, of stochastic volatility $\sigma_{i,t}$ with i being: (i) Trend output; (ii) Annual trend growth rate of output; (iii) Trend unemployment rate; (iv) Output cycle; and (v) Unemployment cycle.

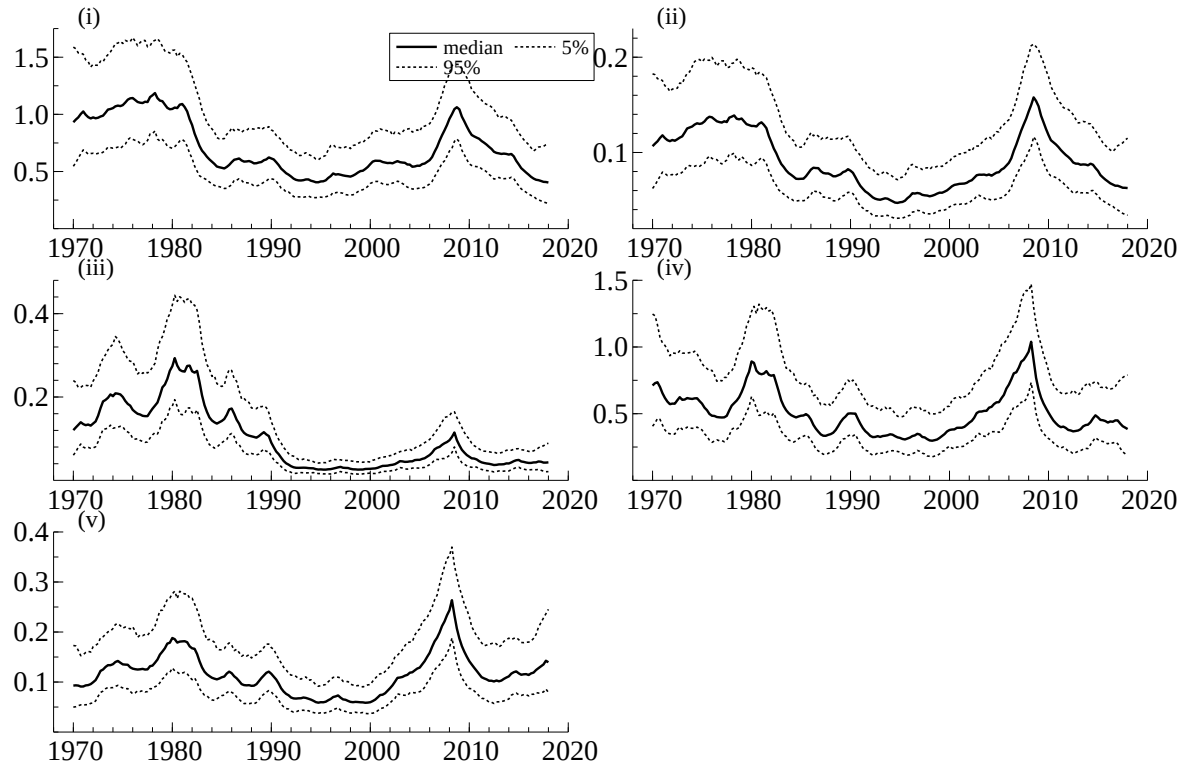

Figure 8: Estimated stochastic volatility of the MSUC model. The solid and dashed lines indicate the posterior median and the 5-th/95-th percentile, respectively of stochastic volatility $\sigma_{i,t}$ with i being: (i) Trend output; (ii) Annual trend growth rate of output; (iii) Trend inflation; (iv) Output cycle; and (v) Inflation cycle.