Hartwig, Benny

Conference Paper

Robust Inference in Time-Varying Structural VAR Models: The DC-Cholesky Multivariate Stochastic Volatility Model

Beiträge zur Jahrestagung des Vereins für Socialpolitik 2020: Gender Economics

Provided in Cooperation with:
Verein für Socialpolitik / German Economic Association

Robust Inference in Time-Varying Structural VAR Models: The DC-Cholesky Multivariate Stochastic Volatility Model

Benny Hartwig*

Abstract

This paper investigates how the ordering of variables affects properties of the time-varying covariance matrix in the Cholesky multivariate stochastic volatility model. It establishes that systematically different dynamic restrictions are imposed when the ratio of volatilities is time-varying. Simulations demonstrate that estimated covariance matrices become more divergent when volatility clusters idiosyncratically. It is illustrated that this property is important for empirical applications. Specifically, alternative estimates on the evolution of U.S. systematic monetary policy and inflation-gap persistence indicate that conclusions may critically hinge on a selected ordering of variables. The dynamic correlation Cholesky multivariate stochastic volatility model is proposed as a robust alternative.

Keywords: Model uncertainty, Multivariate stochastic volatility, Dynamic correlations, Monetary policy, Structural VAR.

JEL: C11, C32, E32, E52.

* Goethe University Frankfurt and Deutsche Bundesbank, Wilhelm-Epstein-Strasse 14, 60431 Frankfurt am Main, Germany. E-mail: benny.hartwig@bundesbank.de

I gratefully acknowledge comments and conversations with Michael Binder, Esteban Prieto, Christian Schumacher, Christiane Baumeister, Christoph Meinerding, Harald Uhlig, Elmar Mertens, Joshua Chan, Mu-Chun Wang and Frank Schorfheide. The views expressed in this paper are those of the authors and do not necessarily coincide with the views of the Deutsche Bundesbank or the Eurosystem.

1 Introduction

Technological innovations, secular trends and policy changes are a few of many factors that shape economic interactions and business cycles over time. To detect and analyze their dynamic relations, time-varying parameter VARs with Cholesky multivariate stochastic volatility (TVP-VAR with CMSV) developed by Primiceri (2005) and Cogley and Sargent (2005) have become a widely established tool in the literature. In fact, a vast body of empirical research has been generated using this class of models.[1]

Even though it is well documented that the time-varying covariance matrix may be sensitive to the ordering of variables in the CMSV model,[2] this property is often ignored and left unchecked in many empirical studies.[3] Such an approach, however, runs the risk of empirical results hinging on a selected ordering of variables and of alternative estimates leading to different conclusions. This paper argues that this property has not been sufficiently explored and must not be ignored in empirical applications.

To illustrate the extent of this type of model uncertainty, Figure 1 shows alternative posterior median estimates of volatilities and covariances of the reduced-form residuals in Primiceri's (2005) application.[4] The figure shows that estimated covariances are sensitive and substantially different across alternative orderings of variables. Specifically, estimates strongly diverge during the stagflation period when the reduced-form residuals exhibit some mildly non-common volatility pattern. This is an undesirable feature as the impulse response functions analyzed in Primiceri's (2005) application depend on both volatilities and covariances. To understand the nature behind the difference across estimates, this paper investigates how the ordering of variables affects properties of the time-varying covariance matrix in the CMSV model.

This paper makes several novel contributions. First, it identifies a time-varying ratio of reduced-form volatilities as to how alternative orderings impose different dynamic restrictions on the time-varying covariance matrix. In the CMSV model, the parameter

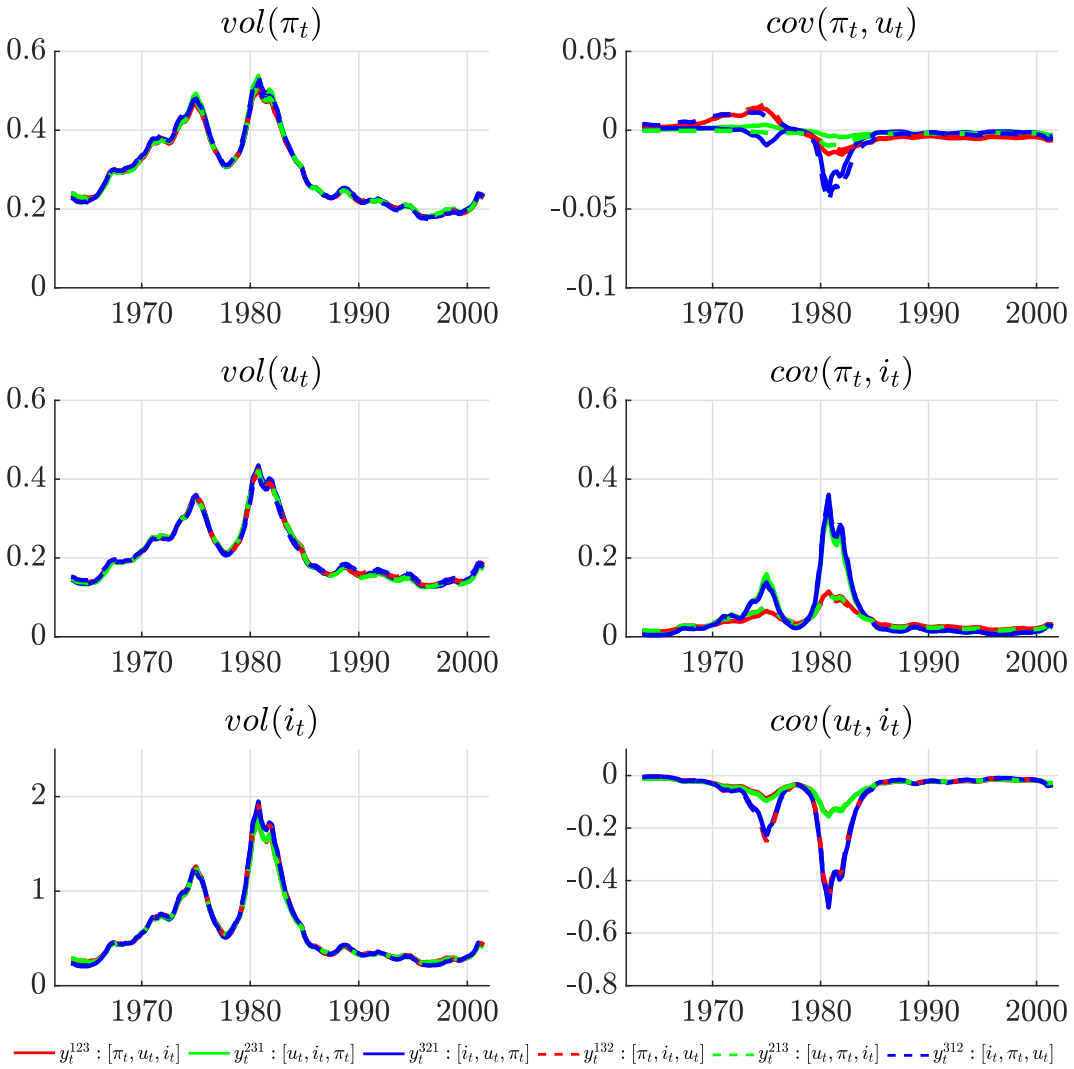
[1]For instance, see the follow-up work of Benati and Surico (2008); Galí and Gambetti (2015) on U.S. monetary policy; Gali and Gambetti (2009); Benati (2008) on great moderation; Baumeister and Peersman (2013a); Baumeister and Peersman (2013b) on oil markets; Gambetti and Musso (2017); Prieto, Eickmeier, and Marcellino (2016) on macro-finance relations; and Mumtaz and Zanetti (2013); Cogley, Primiceri, and Sargent (2010) on TVP-VARs with more complex CMSV versions.

[2]See Primiceri (2005); Cogley and Sargent (2005); Asai, McAleer, and Yu (2006).

[3] Koop, León-González, and Strachan (2009); Nakajima and Watanabe (2011); Lopes, McCulloch, and Tsay (2012); Chan, Doucet, León-González, and Strachan (2018); Bognanni (2018) are exceptions.

[4]Alternative estimates are based on Algorithm 2 in Del Negro and Primiceri (2015) and are obtained from TVP-VARs with CMSV, in which the order of variables has been exchanged.

Figure 1: Estimated contemporaneous reduced-form covariance matrices



The figure shows the posterior median of the covariance (cov) and volatility (vol) of the reduced-form residual of inflation (π_t), unemployment (u_t) and the interest rate (i_t) for all possible orderings in the TVP-SVAR with CMSV.

of contemporaneous relation evolves linearly for a specific variable ordering. When the order of variables is exchanged, the implied dynamics of this parameter are nonlinear. They are determined by the correlation process and a time-varying ratio of volatilities, which is log-normally distributed. These alternative properties of the state process cannot be well captured by an analogously set-up CMSV model and, thus, the model imposes alternative dynamic restrictions on the time-varying covariance matrix.

Moreover, assuming an alternative data generating process that separates volatility and correlation dynamics, the covariance estimates of a CMSV model are systematically different across alternative orderings. This occurs due to the ratio of volatilities driving the parameter of contemporaneous relation being inverted in a reordering, which is associated with different dynamic properties. Monte Carlo simulations show that estimates of the time-varying covariance matrix in the CMSV model become more distinct, when volatility is less common and exhibits pronounced idiosyncratic volatility patterns.

The second contribution of this paper is the introduction of the dynamic correlation Cholesky multivariate stochastic volatility (DC-Cholesky MSV or DC-CMSV) model in the spirit of Engle (2002) as an ordering robust alternative. The DC-CMSV model specifies individual processes for volatilities and correlations to model the time-varying covariance matrix. The correlation dynamics are modelled via a CMSV model on the standardized data, which features a constant ratio of reduced-form volatilities. Simulations and empirical evidence presented in this paper show that the lack of rotational invariance becomes an empirically negligible property for the DC-CMSV model.

A notable feature of the DC-CMSV model is that the parameter of contemporaneous relation is not restricted to evolving linearly but may capture relations between variables that change nonlinearly over time. Moreover, the DC-CMSV approach can be easily implemented into existing CMSV routines. Estimation of the model remains simple, as traditional Kalman filter methods or fast band-precision matrix routines of Chan and Jeliazkov (2009) can be used for inference purposes.

The third contribution of the paper is to demonstrate that restrictions imposed by a particular variable ordering on the time-varying covariance matrix may be so decisive that one may arrive at alternative conclusions. This property is illustrated for Primiceri's (2005) and Cogley, Primiceri, and Sargent's (2010) application.

Regarding Primiceri's (2005) application, alternative estimates on how U.S. monetary policy evolved suggest that the interest rate response to inflation and unemployment was substantially more aggressive during the stagflation period. These estimates are in stark contrast to those presented in the corrigendum of Del Negro and Primiceri (2015), which indicate a largely muted response by the Fed. Apart from that, alternative estimates from a TVP-VAR with DC-CMSV are virtually indistinguishable under all possible orderings. The results from this newly proposed model suggest that the reaction of U.S. systematic monetary policy was modestly more aggressive during the stagflation period, which is consistent with the findings of Sims and Zha (2006).

Estimates from Cogley, Primiceri, and Sargent's (2010) model are very sensitive to the ordering of variables, as the introduction of CMSV heteroskedasticity to the

time-varying VAR parameters on top of the residuals introduces substantial parameter uncertainty. Alternative estimates for inflation-gap persistence suggest that it gradually declined over the sample and did not significantly increase during the stagflation period. Overall, estimates from different models considered in this paper suggest that the majority of empirical evidence provided by Cogley, Primiceri, and Sargent (2010) can be qualitatively confirmed but that there may be a broad range of possible values.

The findings of this paper relate to several strands in the literature. First, the paper formalizes when and why the ordering of variables may matter for a data set at hand. The paper advances the argument of Asai, McAleer, and Yu (2006) and the suggestion made by Christopher Sims to Cogley and Sargent (2005) on p. 11 by the fact that the dependence between volatilities and correlations particularly matters when the ratio of volatilities varies over time. Moreover, evidence provided in this paper suggests that differences in prior distributions play a subordinate role in explaining the sensitivity of the estimates. This possibility was discussed by Primiceri (2005) and Bognanni (2018).

Second, this paper is not the first to demonstrate that variable ordering may play a role in inference. Bognanni (2018) shows for Baumeister and Peersman's (2013b) application that the estimated effects of an oil supply shock on U.S. real activity are sensitive to the chosen ordering as well. Bognanni's (2018) argues that the selection of a variable ordering is an arbitrary choice and should be considered as an additional source of model and parameter uncertainty. What the findings of this paper add is that the CMSV model generally imposes alternative dynamic restrictions on the time-varying covariance matrix. Thus, estimates from this model should not be used as an input for two-step identified time-varying structural VARs, as detailed in Primiceri (2005).

Third, estimates of the time-varying covariance matrix from the DC-CMSV model can be considered as an effective model average over all alternative estimates of the CMSV model. This is an attractive feature, as approaches proposed by Primiceri (2005) or Nakajima and Watanabe (2011) suffer from immense or even intractable computational burdens. These methods need to explore all $n!$ (n factorial) possible models.

Fourth, the DC-CMSV model is an attractive alternative to (inverted) Wishart stochastic volatility models, which are insensitive to the ordering of variables.[5] The caveat of this model class is however that they allow for less flexible covariance matrix dynamics. Particularly, they allow for either integrated or simple autoregressive dynamics, but not for a combination of it or more general processes, see Primiceri

[5]For instance, see Uhlig (1997), Bognanni (2018), and Chan, Doucet, León-González, and Strachan (2018) in the context of (TVP)-VARs with MSV or Philipov and Glickman (2006) and Asai and McAleer (2009) in the context of MSV models.

(2005).

The rest of this paper proceeds as follows. Section 2 derives some properties of the CMSV model under alternative orderings and under an alternative data generating process. Section 3 introduces the DC-CMSV model as a robust alternative. Section 4 conducts a Monte Carlo study to assess properties of the CMSV and DC-CMSV model. Section 5 reconsiders Primiceri's (2005) and Cogley, Primiceri, and Sargent's (2010) application in more detail. Section 6 concludes the paper.

2 On Cholesky Multivariate Stochastic Volatility

This section investigates how the ordering of variables affects properties of the time-varying covariance matrix, Σ_t, in the Cholesky multivariate stochastic volatility (CMSV) model. The name of the model is derived from the fact that it specifies the dynamics of the parameters from the triangular factorization of Σ_t rather than specifying the dynamics of Σ_t directly. Moreover, this section compares the data generating process (DGP) of key state parameters of the CMSV model with the dynamic correlation multivariate stochastic volatility (DC-MSV) model of Yu and Meyer (2006). This alternative MSV model specifies individual volatility and correlation dynamics to span the evolution of Σ_t, which is denoted as the volatility-correlation factorization of Σ_t. The DC-MSV model is chosen as an alternative data generating process because the time-varying covariance matrix is invariant to the ordering of variables[6] and laws of motion for state parameters are comparable across both models.

The following analysis is restricted to the bivariate case for tractability reasons. Nevertheless, the properties discussed below are considered to be representative for the n-dimensional case. This is specifically because the relationship between individual parameters of Σ_t and the parameters under these alternative factorizations of Σ_t does not fundamentally change in higher dimensions.

2.1 Some properties of the Cholesky MSV model

Now, the CMSV model is presented, which builds upon Primiceri's (2005) model.[7] Let y_t be a vector process of dimension 2×1 that is mean zero and has time-varying covariance matrix Σ_t of dimension 2×2.

$$y_t \sim N(0, \Sigma_t) \tag{1}$$

[6]For proof, see Appendix A.2

[7]Around the same time, Tsay (2005) independently introduced the CMSV approach in the financial econometrics literature, see the survey of Asai, McAleer, and Yu (2006).

Then, without loss of generality, the triangular factorization of Σ_t is given by

$$\Sigma_t = A_t^{-1} D_t D_t' A_t'^{-1} \tag{2}$$

where A_t is a lower triangular matrix and D_t is a diagonal matrix

$$A_t = \begin{bmatrix} 1 & 0 \\ a_t & 1 \end{bmatrix}, \qquad D_t = \begin{bmatrix} \exp(g_{1,t}) & 0 \\ 0 & \exp(g_{2,t}) \end{bmatrix}.$$

It follows that the vector of observables can be rewritten as

$$y_t = A_t^{-1} D_t u_t^C, \quad u_t^C \sim N(0, I_2) \tag{3}$$

A decomposition of the covariance matrix resulting in (3) is convenient because it allows for efficient estimation of covariance matrices, see e.g., Pourahmadi (1999). Primiceri (2005) exploits this fact and proposes modelling the coefficients in (3) rather than of (1). This is a valid strategy because there is a one-to-one mapping between Σ_t and its triangular factorization given by A_t and D_t. Notice that the ordering of variables in y_t, which determines the structure of A_t and D_t, is not necessarily related to an identification scheme but is simply a convenient way to decompose Σ_t for estimation.

The model's state parameters are assumed to be Gaussian random walks

$$\begin{aligned} g_t &= g_{t-1} + \epsilon_t^g, \quad \epsilon_t^g \sim N(0, G) \\ a_t &= a_{t-1} + \epsilon_t^a, \quad \epsilon_t^a \sim N(0, \sigma_a^2) \end{aligned}$$

where G is a diagonal covariance matrix as proposed by Koop, León-González, and Strachan (2009). This assumption is commonly used in the applied literature and it is made here to enhance tractability. All innovations are assumed to be jointly normally distributed with the following variance covariance matrix

$$Var\left(\begin{bmatrix} u_t^C \\ \epsilon_t^a \\ \epsilon_t^g \end{bmatrix}\right) = \begin{bmatrix} I_2 & 0 & 0 \\ 0 & \sigma_a^2 & 0 \\ 0 & 0 & G \end{bmatrix}.$$

To complete the specification of the model, common prior distributions are assumed for the initial state parameters and variances

$$\begin{aligned} a_0 &\sim N(\mu_a, V_a), & \sigma_a^2 &\sim IG(\nu_S, k_S^2), \\ g_0 &\sim N(\mu_g, V_g), & \sigma_{g,i}^2 &\sim IG(\nu_g, k_G^2), \forall i = 1, 2 \end{aligned}$$

Let y_t be generated by the CMSV model with Σ_t. From the triangular factorisation of Σ_t, it follows that the mapping from model parameters $\{g_{1,t}, g_{2,t}, a_t\}$ to $\{\sigma^2_{11,t}, \sigma^2_{22,t}, \sigma_{12,t}, \rho_t\}$ the elements and functions of Σ_t is given by

$$\sigma^2_{11,t} = \exp(2g_{1,t}), \qquad \sigma^2_{22,t} = \exp(2g_{2,t}) + a_t^2 \exp(2g_{1,t}),$$
$$\sigma_{12,t} = a_t \exp(2g_{1,t}) \qquad \rho_t = a_t \frac{\sigma_{11,t}}{\sigma_{22,t}}$$

where $\sigma^2_{ii,t}$ is the variance of the ith element of Σ_t for $i = 1, 2$, $\sigma_{12,t}$ is the covariance, ρ_t is the correlation and a_t is the contemporaneous relation.

Then, the model implied state equation for the correlation process, ρ_t, is given by

$$\rho_t = \rho_{t-1} \frac{\exp(\epsilon^g_{1,t})}{\exp(\epsilon^{g**}_{2,t})} + \epsilon^a_t \frac{\sigma_{11,t}}{\sigma_{22,t}}$$

where $\epsilon^{g**}_{2,t} \equiv \log(\sigma_{22,t}) - \log(\sigma_{22,t-1})$.

Property 1 (Σ_t under CMSV model)**.** Let y_t be generated by the CMSV model with Σ_t, then

1. the ratio of reduced-form volatilities $\frac{\sigma_{22,t}}{\sigma_{11,t}}$ is time-varying,
2. the correlation ρ_t evolves nonlinearly,
3. the contemporaneous relation a_t evolves linearly.

For proof, see Appendix A.1.

These properties give rise to two important considerations when using the CMSV model as a data generating process. First, the model rules out common reduced-form volatility dynamics. Second, the assumption of a smoothly evolving contemporaneous relation implies that correlation patterns may rapidly change when volatility clusters idiosyncratically. Stated differently, the model interprets abrupt changes in relative volatilities as the dominant driver of changing correlation.

Let $\tilde{y}_t = Py_t$ be the vector of variables with exchanged rows where P is a permutation matrix satisfying $P \neq I_2$ and $I_2 = P'P$. I_2 is the identity matrix. Let $\tilde{\Sigma}_t = P\Sigma_t P'$ be the covariance matrix with permuted elements. Analogously, the triangular factorsation of $\tilde{\Sigma}_t = \tilde{A}_t^{-1}\tilde{D}_t\tilde{D}_t\tilde{A}_t'^{-1}$ implies that the mapping from model parameters $\{g_{1,t}, g_{2,t}, a_t\}$ to $\{\tilde{g}_{1,t}, \tilde{g}_{2,t}, \tilde{a}_t\}$, the transformed model parameters for $\tilde{y}_t$, is given by

$$\exp(2\tilde{g}_{1,t}) = \sigma^2_{22,t}, \qquad \exp(2\tilde{g}_{2,t}) = \sigma^2_{11,t} - \tilde{a}_t^2\sigma^2_{22,t}, \qquad \tilde{a}_t = a_t \frac{\sigma^2_{11,t}}{\sigma^2_{22,t}}.$$

Then, the model-implied state equation for the analogously defined parameter of contemporaneous relation, $\tilde{a}_t$, is given by

$$\tilde{a}_t = \tilde{a}_{t-1} \frac{\exp\left(2\epsilon^g_{1,t}\right)}{\exp\left(2\epsilon^{g**}_{2,t}\right)} + \epsilon^a_t \frac{\sigma^2_{11,t}}{\sigma^2_{22,t}},$$

where $\epsilon^{g**}_{2,t} \equiv \log(\sigma_{22,t}) - \log(\sigma_{22,t-1})$.

Property 2 (Reordering in CMSV model)**.** Let Σ^*_t be the time-varying covariance matrix of an analogously set-up CMSV model on $\tilde{y}_t$ with model parameters $\{g^*_{1,t}, g^*_{2,t}, a^*_t\}$, then

- Σ^*_t and $\tilde{\Sigma}_t$ cannot have the same dynamic structure, and
- the average distance between transformed implied parameters $\{\tilde{g}_{1,t}, \tilde{g}_{2,t}, \tilde{a}_t\}$ and analogously constructed parameters $\{g^*_{1,t}, g^*_{2,t}, a^*_t\}$ increases with the variability of the ratio of reduced-form variances $\frac{\sigma^2_{11,t}}{\sigma^2_{22,t}}$.

For proof, see Appendix A.1.

To put it differently, the ordering of variables induces a dynamic structure in Σ_t that cannot be replicated by an analogously set-up CMSV model for any alternative ordering of variables. The CMSV model therefore imposes different dynamic restrictions on the time-varying covariance matrix under alternative orderings. As a consequence, the ordering of variables is a nontrivial choice in the CMSV model.

While Σ^*_t and $\tilde{\Sigma}_t$ cannot have the same dynamic structure, the dynamics may be similar or may diverge substantially. This distance depends on the volatility pattern of the data. Specifically, the distance is smaller when the volatility pattern of the individual series exhibits strong commonalities. In this incidence, the ratio of reduced-form variances becomes closer to being roughly constant. However, when there are idiosyncratic volatility patterns, then this distance grows larger. Notice, an analytical quantification of this distance is not readily available. Section 4 provides some quantification of this distance by means of a Monte Carlo simulation with a variety of alternative data generating processes.

Above statements allow for some clarification on remarks that appear in the literature about properties of the time-varying covariance matrix in the CMSV model. Primiceri (2005) points out that the ordering of variables matters for Σ_t because the prior distribution of Σ_t is not rotationally invariant. Particularly, he shows that the individual elements of the covariance matrix have alternative distributions under different orderings of the variables (see footnote 5). Nevertheless, he suggests that it is not

a priori clear how inference is affected and that the effect might vary from case to case. Relatedly, Bognanni (2018) argues that the introduction of the dynamic dependence of model parameters in conjunction with the factorization of the covariance matrix leads to a non-observational equivalent prior density for Σ_t.[8] The discussion above clarifies when the ordering of variables is important for inference. Specifically, it matters when there are idiosyncratic volatility patterns.

These results also shed light on the discussion of Asai, McAleer, and Yu (2006) and a suggestion made by Christopher Sims to Cogley and Sargent (2005) on p. 11. They conjecture that not separating volatility and correlation dynamics may impose some dynamic restrictions on the time-varying covariance matrix. Particularly, the CMSV model rules out common volatility patterns, which induce some nonlinear correlation dynamics. Also, because volatility patterns are not common, alternative orderings impose different dynamic restrictions on the time-varying covariance matrix.

2.2 The Cholesky MSV model and the DC-MSV model

Alternatively, the dynamics of Σ_t in (1) may be modelled by a DC-MSV model in the spirit of Yu and Meyer (2006). Particularly, this model differs from the CMSV model by the choice of the factorization for Σ_t. In particular, it models the dynamics of the covariance matrix by decomposing Σ_t into individual processes for volatility and correlation, which is given by

$$\Sigma_t = D_t R_t D_t \tag{4}$$

where R_t is a correlation matrix and D_t is a diagonal matrix

$$R_t = \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix}, \qquad D_t = \begin{bmatrix} \exp(h_{1,t}) & 0 \\ 0 & \exp(h_{2,t}) \end{bmatrix}.$$

It follows that the vector of observables can be written as

$$y_t = D_t u_t^{DC}, \quad u_t^{DC} \sim N(0, R_t). \tag{5}$$

A decomposition of the covariance matrix resulting in (5) has been introduced by Bollerslev (1990) and Engle (2002) as a parsimonious modelling alternative to fully parameterized multivariate GARCH models for estimating large dynamic covariance matrices. One main distinguishing characteristic between these factorisations is that in (4) volatilities and correlations are modelled as independent processes, whereas in (2) they are modelled jointly.

[8] Primiceri (2005) elaborates on the triangular factorization in the contemporaneous case.

The model's state parameters are assumed to evolve as Gaussian random walks

$$\begin{aligned} h_t &= h_{t-1} + \eta_t^h, \quad \eta_t^h \sim N(0, W) \\ m_t &= m_{t-1} + \eta_t^m, \quad \eta_t^m \sim N(0, \sigma_m^2) \\ \rho_t &= \frac{\exp(m_t) - 1}{\exp(m_t) + 1}, \quad \eta_t^\rho \equiv \rho_t - \rho_{t-1} \end{aligned}$$

where W is a diagonal covariance matrix and m_t is an auxiliary process that is mapped into a correlation process using the Fisher transformation. All innovations are assumed to be jointly normally distributed with the following variance covariance matrix.

$$Var\left(\begin{bmatrix} u_t^{DC} \\ \eta_t^m \\ \eta_t^h \end{bmatrix}\right) = \begin{bmatrix} R_t & 0 & 0 \\ 0 & \sigma_m^2 & 0 \\ 0 & 0 & W \end{bmatrix}.$$

To complete model specification, common prior distributions are assumed.

$$\begin{aligned} m_0 &\sim N(\mu_m, V_m), & \sigma_m^2 &\sim IG(\nu_m, k_m^2), \\ h_0 &\sim N(\mu_h, V_h), & \sigma_{h,i}^2 &\sim IG(\nu_h, k_W^2), \forall i = 1, 2. \end{aligned}$$

Let y_t be generated by the DC-MSV model with Σ_t. From the volatility-correlation decomposition of Σ_t, it follows that the mapping from model parameters $\{h_{1,t}, h_{2,t}, m_t\}$ to $\{\sigma_{11,t}^2, \sigma_{22,t}^2, \sigma_{12,t}, \rho_t, a_t, \tilde{a}_t\}$, the elements and functions of Σ_t, is given by

$$\begin{aligned} \sigma_{11,t}^2 &= \exp(2h_{1,t}), & \sigma_{22,t}^2 &= \exp(2h_{2,t}) \\ \sigma_{12,t} &= \rho_t \exp(h_{1,t}) \exp(h_{2,t}), & \rho_t &= \frac{\exp(m_t) - 1}{\exp(m_t) + 1}, \\ a_t &= \rho_t \frac{\sigma_{22,t}}{\sigma_{11,t}}, & \tilde{a}_t &= \rho_t \frac{\sigma_{11,t}}{\sigma_{22,t}} \end{aligned}$$

where $\sigma_{ii,t}^2$ is the variance of the ith element of Σ_t for $i = 1, 2$, $\sigma_{12,t}$ is the covariance, ρ_t is the correlation, a_t and $\tilde{a}_t$ are the respective parameters of contemporaneous relation implied under Σ_t and $\tilde{\Sigma}_t = P\Sigma_t P'$, respectively.

Then, model-implied state equations for a_t and $\tilde{a}_t$ are given by

$$a_t = a_{t-1} \frac{\exp(\epsilon_{2,t-1}^h)}{\exp(\epsilon_{1,t-1}^h)} + \eta_t^\rho \frac{\sigma_{22,t}}{\sigma_{11,t}}, \qquad \tilde{a}_t = \tilde{a}_{t-1} \frac{\exp(\epsilon_{1,t-1}^h)}{\exp(\epsilon_{2,t-1}^h)} + \eta_t^\rho \frac{\sigma_{11,t}}{\sigma_{22,t}}$$

where $\eta_t^\rho \equiv \rho_t - \rho_{t-1}$.

Under the DC-MSV model, the parameters of contemporaneous relations a_t and $\tilde{a}_t$ are driven by the correlation as well as the respective ratio of volatilities $\frac{\sigma_{22,t}}{\sigma_{11,t}}$ and $\frac{\sigma_{11,t}}{\sigma_{22,t}}$.

These ratios are determined by the ordering of the variables in y_t and $\tilde{y}_t$. Specifically, they are defined by the volatility of the variable ordered in the second position in the vector of variables divided by the one ordered in the first position. Thus, variable orderings play a particular role for the implied evolution of a_t and $\tilde{a}_t$.

Notice, the volatility-correlation decomposition of Σ_t allows for specifying very general volatility dynamics. For instance, the DC-MSV model may be set up to feature a purely idiosyncratic volatility pattern (as specified above) or to exhibit some commonalities or a completely common volatility pattern.

Property 3 (Σ_t under DC-MSV model)**.** Let y_t be generated by the DC-MSV model with Σ_t, then

1. the correlation ρ_t evolves approximately linearly Gaussian for $\rho_t \in (-0.5, 0.5)$,
2. when the ratio of reduced-form volatilities $\frac{\sigma_{22,t}}{\sigma_{11,t}}$ is constant, then a_t and $\tilde{a}_t$ are solely driven by ρ_t and have the same dynamics up to a scalar,
3. when the ratio of reduced-form volatilities $\frac{\sigma_{22,t}}{\sigma_{11,t}}$ is time-varying, then a_t and $\tilde{a}_t$ evolve nonlinearly and have different dynamic properties.

For proof, see Appendix A.2.

Comparing the properties of Σ_t implied by both models, the CMSV model assumes a linear Gaussian process for the parameter of contemporaneous relation but implies a nonlinearly evolving correlation process. The DC-MSV model, in contrast, implies nonlinear dynamics for the parameter of contemporaneous relation and approximately linear Gaussian dynamics for the correlation in some specified range. In other words, the correlation process acts somewhat as a degree of freedom in the CMSV model, while the parameter of contemporaneous relation acquires this role in the DC-MSV model.

Property 4 (DC-MSV, CMSV and implied covariances)**.** Let Σ_t be generated by the DC-MSV model. Then, the implied dynamics of the covariance $\sigma_{12,t}$, approximated by the state equations of the CMSV model for y_t, is underestimated when the ratio of volatilities increases; while it is mechanically overestimated, when it is approximated by the state equations of the CMSV model for $\tilde{y}_t$, as the ratio of volatilities is inverted.

For proof, see Appendix A.2.

Therefore, when y_t is generated by the DC-MSV model, the covariances implied by a CMSV model are systematically different across alternative orderings. In particular, estimates of the CMSV model represent an upper and lower bound of the true covariance parameter under the DC-MSV model.

Property 5 (Posterior distribution of a_t and $\tilde{a}_t$ under homoskedasticity)**.** Let y_t be generated by a bivariate dynamic correlation model with constant unitary variances on the main diagonal. Then, the difference of posterior mean and variance of a_t and $\tilde{a}_t$ implied under a respective CMSV model is induced by the likelihood and not the prior. The difference between the posterior mean and variance of a_t and $\tilde{a}_t$ depends on the distance between the sequence of $y_{1,t}^2$ and $y_{2,t}^2$.

For proof, see Appendix A.2.

Under this alternative data generating process, the implied posterior distribution of a_t and $\tilde{a}_t$ is not the same as the data is interpreted differently across alternative orderings in the CMSV model. For this reason, the model produces different estimates of the time-varying covariance matrix under alternative orderings. However, the prior distribution of a_t and $\tilde{a}_t$ is the same across alternative orderings. This suggests that the posterior distribution of the time-varying covariance matrix should be largely insensitive to the ordering of variables. Section 4 provides evidence that differences across posterior estimates are visually negligible.

The property of the CMSV model that prior and posterior distribution of the time-varying covariance matrix are largely insensitive to the ordering of variables when data is homoskedastic suggests that one may use the CMSV approach as an ordering robust data generating process for the correlation process. To construct a fully specified time-varying covariance matrix, this process for the correlation can be combined with a separate volatility model in the spirit of Engle (2002). This idea is exploited in the next section to construct a new multivariate stochastic volatility model that is largely insensitive to the ordering of variables and may be used for higher dimensional systems of variables.[9]

3 The DC-Cholesky MSV Model

This section presents the details of the dynamic correlation Cholesky multivariate stochastic volatility (DC-Cholesky MSV or DC-CMSV) model. The DC-CMSV model uses a separate volatility model for the data and models the correlation dynamics of the standardized data with the CMSV approach.

Let y_t be a $n \times 1$ dimensional vector process that is mean zero and has a time-varying

[9] The DC-MSV model of Yu and Meyer (2006) cannot be easily generalized to higher dimensions for $n \geq 3$. Asai and McAleer (2009) present an alternative DC-MSV model that is applicable to higher dimensions but uses an inverted Wishart process to model correlation dynamics.

covariance matrix Σ_t of dimension $n \times n$

$$y_t \sim N(0, \Sigma_t). \tag{6}$$

Then, Σ_t may be decomposed into marginal volatilities and correlations by

$$\Sigma_t \;=\; D_t R_t D_t$$

where D_t is a diagonal matrix with volatilities and R_t is a correlation matrix

$$D_t \;=\; \begin{bmatrix} \exp(h_{1,t}) & 0 & \dots & 0 \\ 0 & \exp(h_{2,t}) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \exp(h_{n,t}) \end{bmatrix}, R_t = \begin{bmatrix} 1 & \rho_{2,1,t} & \dots & \rho_{n,1,t} \\ \rho_{2,1,t} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho_{n,n-1,t} \\ \rho_{n,1,t} & \dots & \rho_{n,n-1,t} & 1 \end{bmatrix}.$$

It follows that the vector observables can be rewritten as

$$y_t = D_t \epsilon_t, \quad \epsilon_t \sim N(0, R_t). \tag{7}$$

Then, an auxiliary positive definite matrix is estimated on the standardized data

$$\epsilon_t = A_t^{*-1} D_t^* e_t, e_t \sim N(0, I_n)$$

where

$$A_t^* \;=\; \begin{bmatrix} 1 & 0 & \dots & 0 \\ a_{2,1,t}^* & 1 & \dots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ a_{n,1,t}^* & \dots & a_{n,n-1,t}^* & 1 \end{bmatrix}, D_t^* = \begin{bmatrix} \exp(h_{1,t}^*) & 0 & \dots & 0 \\ 0 & \exp(h_{2,t}^*) & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \exp(h_{n,t}^*) \end{bmatrix}$$

which is transformed to a correlation matrix using the formulas of Engle (2002)

$$R_t = Q_t^{*-\frac{1}{2}} Q_t Q_t^{*-\frac{1}{2}} \tag{8}$$

$$Q_t = A_t^{*-1} D_t^* D_t^{*'} {A_t}^{*'-1} \tag{9}$$

$$Q_t^* = \text{diag}[\text{vecd}(Q_t)] \tag{10}$$

where $\text{vecd}(Q_t)$ selects the diagonal of Q_t.

Let a_t^* be the lower off-diagonal elements of A_t^* (stacked by rows) and h_t and h_t^* be the vector of log volatilities on the diagonal of the matrix D_t and D_t^*, respectively.

Assume that the state dynamics evolve as a random walk

$$h_t = h_{t-1} + \epsilon_t^h, \tag{11}$$
$$a_t^* = a_{t-1}^* + \epsilon_t^{a^*}, \tag{12}$$
$$h_t^* = h_{t-1}^* + \epsilon_t^{h^*}. \tag{13}$$

All innovations of the model are assumed to be joint normal.

$$V = Var\left(\begin{bmatrix} e_t \\ \epsilon_t^h \\ \epsilon_t^{a^*} \\ \epsilon_t^{h^*} \end{bmatrix}\right) = \begin{bmatrix} I_n & 0 & 0 & 0 \\ 0 & W & 0 & 0 \\ 0 & 0 & S^* & 0 \\ 0 & 0 & 0 & W^* \end{bmatrix}$$

where I_n is an identity matrix, S^* is a block diagonal matrix, $W = diag([\sigma_{h,1}^2, ..., \sigma_{h,n}^2])$ and $W^* = diag([\sigma_{h,1}^{*2}, ..., \sigma_{h,n}^{*2}])$ are positive definite matrices.

Assume independent prior distribution for h_0, a_0^*, h_0^*, W, S_i^*, W^*.

$$\begin{aligned} h_0 &\sim N(\mu_h, V_h), & \sigma_{h,i}^2 &\sim IG(\nu_h, k_W^2), \forall i = 1, ..., n, \\ a_0^* &\sim N(\mu_a^*, V_a^*), & S_i &\sim IW(\nu_{S,i}^*, k_S^{*2} \cdot I_i), \forall i = 1, ..., n-1, \\ h_0^* &\sim N(\mu_h^*, V_h^*), & \sigma_{h,i}^{*2} &\sim IG(\nu_h^*, k_W^{*2}), \forall i = 1, ..., n. \end{aligned}$$

Next, the Gibbs sampling algorithm for the DC-CMSV model is presented, which builds on the notation and results from Chan (2017). Stochastic volatility is sampled using the auxiliary mixture sampler of Kim, Shephard, and Chib (1998).

Algorithm: Gibbs sampling algorithm for the DC-CMSV model
Pick some initial values for $h^{(0)}$, $W^{(0)}$, $h_0^{(0)}$, $\epsilon^{*(0)}$, $a^{(0)}$, $S^{(0)}$, $a_0^{(0)}$, $h^{(0)}$, $W^{(0)}$ and $h_0^{(0)}$. Then, repeat the steps from $r = 1$ to R:

1. Posterior draws from $p(s, h, W, h_0, \epsilon|y)$
   - Draw $s^{(r)} \sim (s|y, h^{(r-1)})$ (seven point distribution)
   - Draw $h^{(r)} \sim (h|y, s^{(r)}, W^{(r-1)}, h_0^{(r-1)})$ (multivariate normal)
   - Draw $W^{(r)} \sim (W|h^{(r)}, h_0^{(r-1)})$ (independent inverse Gamma)
   - Draw $h_0^{(r)} \sim (h_0|y, h^{(r)}, W^{(r)})$ (independent normal)
   - Draw $\epsilon^{(r)} \sim (\epsilon|y, h^{(r)})$ (transform data to standardized normal)
2. Posterior draws from $p(a^*, a_0^*, S^*|\epsilon^{(r)}, h^{*(r-1)})$

- Draw $a^{*(r)} \sim (a^*|\epsilon^{(r)}, a_0^{*(r-1)}, S^{*(r-1)}, h^{*(r-1)})$ (multivariate normal)
- Draw $S^{*(r)} \sim (S^*|\epsilon^{(r)}, a^{*(r)}, a_0^{*(r-1)})$ (inverse Wishart)
- Draw $a_0^{*(r)} \sim (a_0^*|\epsilon^{(r)}, a^{*(r)}, S^{*(r)})$ (multivariate normal)

3. Posterior draws from $p(s^*, h^*, W^*, h_0^*|\epsilon^{(r)}, a^{*(r)})$

- Draw $s^{*(r)} \sim (s^*|\epsilon^{(r)}, h^{*(r-1)})$ (seven point distribution)
- Draw $h^{*(r)} \sim (h^*|\epsilon^{(r)}, s^{*(r)}, W^{*(r-1)}, h_0^{*(r-1)})$ (multivariate normal)
- Draw $W^{*(r)} \sim (W^*|h^{*(r)}, h_0^{*(r-1)})$ (independent inverse Gamma)
- Draw $h_0^{*(r)} \sim (h_0^*|\epsilon^*, h^{*(r)}, W^{*(r)})$ (independent normal)

In contrast to the traditional Gibbs sampler of the CMSV model, the DC-CMSV sampler first estimates the marginal volatility components of D_t and standardizes the observed data. Then, a pseudo time-varying covariance matrix is estimated on the standardized data with the CMSV model. The parameters A_t^* and D_t^* are then transformed into an estimate of the time-varying correlation matrix R_t by equations (8)–(10). The draw of R_t in conjunction with D_t is used to span the evolution of the time-varying covariance matrix Σ_t.

Next, some of the merits and drawbacks of the DC-CMSV model are discussed. First, the posterior distributions of the marginal volatilities are independent of the ordering of variables. Second, when the process for marginal volatilities is correctly specified, the posterior distribution of the time-varying correlation matrix is largely insensitive to the ordering of variables. Consequently, posterior estimates of the time-varying covariance matrix of the DC-CMSV model are almost insensitive to the ordering of variables. Third and in contrast to the CMSV model, DC-CMSV model implied parameters of contemporaneous relations may capture nonlinear instead of linear dynamics between variables.

Nevertheless, these appealing properties come at the cost of increased computational complexity. In particular, the computational costs increase on account of the need for the volatility series to be sampled twice instead of once, i.e. the independent volatilities and the auxiliary volatilities for the estimation of the correlation matrix. However, the increased computational complexity of the model remains manageable on modern multi-core computers as volatility sampling can be parallelized, see Lopes, McCulloch, and Tsay (2012).

4 Monte Carlo Study

This section conducts a Monte Carlo study to quantify how the ordering of variables affects the posterior estimates of the time-varying covariance matrix in the CMSV model and DC-CMSV model when data is homoskedastic and heteroskedastic. Properties of posterior estimates are characterized by in-sample fit, distance and similarity metrics.

4.1 Properties of Σ_t under homoskedasticity

According to Property 5, the posterior mean and variance of a_t and $\tilde{a}_t$ are driven apart by the influence of the likelihood and not the prior. To study the importance of this property, a bivariate dynamic correlation model with a known correlation structure and unitary variances is assumed as the data generating process to simulate 250 samples of 1,000 observations each.[10] This data generating process has been chosen as it implies that the prior distribution of a_t and $\tilde{a}_t$, as defined in Section 2.2, are the same and that the evolution of these parameters is solely driven by the evolution of the correlation ρ_t, see Property 3.2.

The data generating process is given by

$$\begin{array}{rcl} r_{1,t} & = & \nu_{1,t} \\ r_{2,t} & = & \nu_{2,t} \end{array}, \quad \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \right)$$

where the process of conditional correlation uses the specification of Engle (2002)

1. Constant: $\rho_t = 0.9$,
2. Sine: $\rho_t = 0.5 + 0.4\cos(2\pi t/200)$,
3. Fast Sine: $\rho_t = 0.5 + 0.4\cos(2\pi t/20)$,
4. Step: $\rho_t = 0.9 - 0.5I(t > 500)$,
5. Ramp: $\rho_t = \mod(t/200)$.

These processes for the conditional correlation were chosen by Engle (2002) as they exhibit various types of rapid changes, gradual changes, and periods of constancy. Some of the processes appear to be mean reverting, while others have abrupt changes.

[10] A sample size of 1,000 is not yet realistic for macro application, however, it is chosen to mitigate the effect of a small sample on the estimates.

The CMSV model is estimated with fixed hyperparameters (k_S, k_W), which are set to 0.1 and common prior distribution. The MCMC estimation produces 35,000 samples of which 15,000 are reserved for the burn-in period.

The in-sample fit for parameters of interest estimated under different orderings is measured by the average mean absolute error (MAE). MAE statistics are computed for the implied correlation ρ_t from estimates of Σ_t and for the parameter of contemporaneous relation a_t and $\tilde{a}_t$ across alternative orderings. The MAE is defined as

$$MAE(X^{O\bar{R}D}, X^0) = \frac{1}{M}\sum_{i}^{M}\left(\frac{1}{T}\sum_{t}^{T}|X_t^{ORD(i)} - X_t^0|\right) \tag{14}$$

where $X^{ORD(i)}$ denotes the parameter estimate of the model with variable ordering $i = 1, 2$ for $M = 2$. ORD(1) and ORD(2) denote the ordering $y_t = (y_{1,t}, y_{2,t})'$ and $\tilde{y}_t = (y_{2,t}, y_{1,t})'$, respectively. X^0 denotes the true value of the parameter.

The distance between alternative estimates is measured by the mean absolute difference (MAD) of the parameter of interest. MAD statistics are computed for $\rho_t, a_t, \tilde{a}_t$ and $a_t - \tilde{a}_t$. The latter denotes the difference between posterior median estimates of a_t and $\tilde{a}_t$. The MAD is defined as

$$MAD(X^{ORD(1)}, X^{ORD(2)}) = \frac{1}{T}\sum_{t}^{T}|X_t^{ORD(1)} - X_t^{ORD(2)}| \tag{15}$$

Table 1 presents the results from this Monte Carlo simulation. Turning to the in-sample fit for the estimated parameters, MAE statistics indicate that posterior median estimates of a_t and $\tilde{a}_t$ fit the true correlation equally well. Estimates of the implied correlation ρ_t, however, are always more precise than estimates for the parameters of contemporaneous relation. This is not surprising as the CMSV model is designed to produce valid draws of a covariance matrix and not of a correlation matrix.

MAD statistics indicate that the distance between alternative estimates of a_t and $\tilde{a}_t$ obtained under different orderings are of the same magnitude. The average distance among all simulated processes is 0.026, which is rather small. However, the distance between the difference of estimated posterior medians of a_t and $\tilde{a}_t$, $a_t - \tilde{a}_t$, is not small, with an average of 0.08. Thus, estimates of a_t and $\tilde{a}_t$ exhibit some alternative patterns. Furthermore, the distance across implied correlation estimates ρ_t is the smallest among all considered parameters, at 0.016. This indicates that even though the likelihood drives pseudo estimates of correlation a_t and $\tilde{a}_t$ apart, it does not, however, substantially affect estimates of the implied correlation ρ_t.

Table 1: Precision and discrepancy of posterior median estimates

	MAE			MAD			
	ρ_t	a_t	$\tilde{a}_t$	ρ_t	$a_t - \tilde{a}_t$	a_t	$\tilde{a}_t$
const	**0.016**	0.043	0.043	**0.008**	0.084	0.018	0.019
sine	**0.080**	0.092	0.091	**0.022**	0.086	0.035	0.034
fastsine	**0.256**	0.257	0.257	**0.016**	0.070	0.020	0.020
step	**0.049**	0.065	0.066	**0.010**	0.076	0.018	0.018
ramp	**0.106**	0.117	0.116	**0.023**	0.087	0.037	0.037

The table shows the forecast accuracy (MAE) and distance (MAD) for estimated correlation ρ_t, the contemporaneous relation a_t and $\tilde{a}_t$ across different orderings and the difference between latter parameters $a_t - \tilde{a}_t$. A bold figure highlights the lowest statistics for all parameters considered in each panel and for each DGP in each row.

Therefore, this Monte Carlo simulation provides evidence that the estimated correlation of the CMSV model fits the data well and is almost insensitive to the ordering of variables for homoskedastic data. Here, "almost" means that the distance between alternative posterior median estimates is small, around or less than 0.025.[11]

4.2 Properties of Σ_t under heteroskedasticity

To investigates the sensitivity of the time-varying covariance matrix under heteroskedasticity, stochastic volatility is introduced to the dynamic correlation process. According to Property 2, the more pronounced the movements in the ratio of reduced-form volatilities of the data, the more significant dynamic restriction are imposed on the time-varying covariance matrix in the CMSV model. By construction, the DC-CMSV model integrates out stochastic volatility before estimating the correlation process. Therefore, the structure of the model ensures that different degrees of idiosyncratic volatility patterns do not drive estimates of the time-varying covariance matrix apart.

The process for stochastic volatility follows Asai and McAleer (2009), who assume very persistent stochastic volatility dynamics for their dynamic correlation model, similar to the original GARCH specification in Engle (2002). In addition, a scale parameter c^i is introduced to simulate different degrees of idiosyncratic volatility patterns. Specifically, three different scales are considered $c^i = \{1, 2, 0.5\}$ with $i = \{BM, H, L\}$, which are denoted as benchmark, high volatility, low volatility DGP, respectively. By in-

[11] The chosen level for the threshold is arbitrary, however, an MAD of 0.025 indicates that differences between correlation estimates are hardly visible on the possible range for $\rho_t \in [-1, 1]$.

creasing the scale of the innovations, the process of stochastic volatilities becomes more nonlinear and, hence, the ratio of volatilities exhibits more time-variation.

The data generating process for stochastic volatility is defined as

$$\begin{array}{rcl} h_{1,t+1} & = & 0.98h_{1,t} + \eta_{1,t+1} \\ h_{2,t+1} & = & 0.95h_{2,t} + \eta_{2,t+1} \end{array}, \quad \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, c^i \begin{bmatrix} 0.166^2 & 0 \\ 0 & 0.26^2 \end{bmatrix} \right)$$

then use them for each correlation process,

$$\begin{array}{rcl} r_{1,t} & = & \nu_{1,t} \exp(0.5h_{1,t}) \\ r_{2,t} & = & \nu_{2,t} \exp(0.5h_{2,t}) \end{array}, \quad \begin{pmatrix} \nu_{1,t} \\ \nu_{2,t} \end{pmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \right)$$

which are the same as in Section 4.1.

The CMSV model and DC-CMSV model are estimated with fixed hyperparameters (k_S, k_W) and (k_W, k_S^*, k_W^*) that are all set to 0.1. The MCMC estimation produces 35,000 samples of which 15,000 are reserved for the burn-in period. Appendix B complements the presented results with two robustness exercises.[12]

The in-sample fit of the estimated parameter is evaluated by the MAE. MAE statistics are computed for estimated correlations and covariances. The sensitivity of estimated parameters due to alternative orderings is assessed by distance and correlation metrics. Reported as distance metrics are the MAD and the root mean square difference (RMSD), which is similarly defined as the MAD. When the RMSD statistic grows substantially larger than the MAD statistic, then this indicates that there are periods when the distance between estimated parameters is unusually large.

The similarity of estimated parameters is measured by the correlation of the first difference of estimated parameters (FD). The first difference rather than the level of the estimates is used because the latter induces spurious correlation due to common trends in the level series.

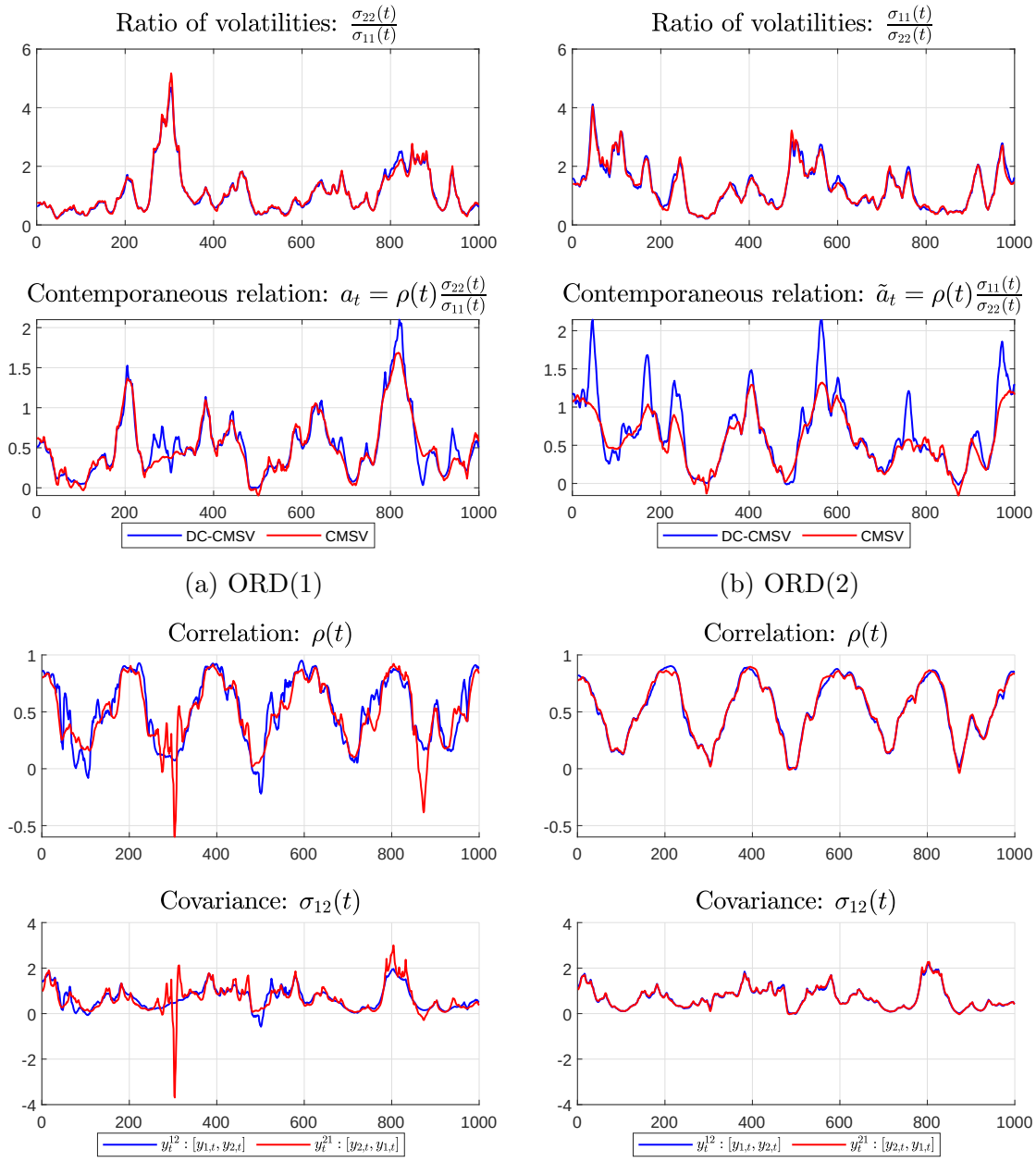
The FD statistics is defined as

$$FD(X^{ORD(1)}, X^{ORD(2)}) = \text{corr} \left(\Delta X_t^{ORD(1)}, \Delta X_t^{ORD(2)} \right)$$

where Δ denotes the first difference operator.

[12] In the first exercise, the hyperparameters (k_S, k_W) and (k_W, k_S^*, k_W^*) are estimated with the recent algorithm of Amir-Ahmadi, Matthes, and Wang (2018) to allow for a fair comparison across models and orderings as well as to control for different degrees of time-variation associated with alternative DGPs. The second exercise addresses the concern that the random walk transition equation is misspecified as the true DGP assumes that volatilities and some correlation processes evolve stationary. To control for this, the models are re-estimated assuming autoregressive laws of motion for volatilities and time-varying parameters. Overall, the results presented are robust to these sensitivity checks.

Figure 2: Time-varying covariance matrix (sine, benchmark DGP)



The figure shows the posterior median of the ratio of volatilities and contemporaneous relation for a selected ordering for both models in the upper panel. The lower panel shows the correlation and the covariance for a selected model for both orderings.

To illustrate the sensitivity of Σ_t under both models, Figure 2 shows the posterior median of the ratio of volatilities and parameter of contemporaneous relation for a

selected ordering of variables and both models in the upper panel, and the correlation and the covariance for a selected model and both orderings in the lower panel for one replication of the sine correlation process.

The upper panel of the figure shows that the ratio of volatilities is similar across models and exhibits nonlinear patterns over time. For the parameter of contemporaneous relation, however, there are marked differences across models. Especially, estimates of the DC-CMSV model exhibit nonlinear dynamics, which are linked to the movement of the ratio of volatilities. Moving to the lower panel, estimated correlations and covariances of the CMSV model exhibit systematic differences across alternative orderings. The estimates especially diverge when the ratio of volatilities suddenly moves substantially. This happens as the CMSV model cannot properly capture the nonlinear dynamics of the parameter of contemporaneous relation. Moreover, estimated correlations and covariances of the DC-CMSV model lie somewhere between alternative estimates of the CMSV model. According to Property 4, this is to be expected as estimates of the CMSV model can be regarded as an upper and lower bound of the true comovement parameters when data is generated from a dynamic correlation model.

Table 2 presents statistics for estimated correlations. The table shows that the DC-CMSV model produces the most precise correlation estimates for all except the fastsine correlation process. Here, estimates of the CMSV model are more precise. However, the in-sample fit of the CMSV model deteriorates substantially for the high volatility DGP, while the statistics of both models are similar for the low volatility DGP. Moreover, the absolute value of the in-sample fit statistics for the DC-CMSV model remains similar under different degrees of idiosyncratic volatility patterns.

Differences between MAD and RMSD statistics for the CMSV model indicate that there may be periods when there is a considerable distance across alternative correlation estimates. Particularly, RMSD statistics for the benchmark DGP show that a distance of 0.04 to 0.11 for the estimated correlation path is not unusual. These statistics substantially inflate and deflate for the high volatility DGP and the low volatility DGP, respectively. In contrast, analogous statistics for the DC-CMSV model are hardly affected by different scales of idiosyncratic volatility patterns in the simulated data.

FD statistics are fairly far below one for estimated correlations of the CMSV model. This indicates that changes in estimated correlation paths feature some idiosyncratic components. For the DC-CMSV model, FD statistics are substantially higher, but are still not close to one. For the case of constant correlation, the metric is negative, which suggests that estimated correlation paths move in opposite directions. However, FD statistics should not be interpreted in isolation. In fact, the low level of MAD and

RMSD statistics indicates that estimated correlations are indeed very similar for the DC-CMSV model.

Table 2: Estimated correlation

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.040	**0.030**	0.031	**0.025**	0.041	**0.031**	**0.307**	-0.284
sine	0.096	**0.086**	0.063	**0.018**	0.088	**0.023**	0.559	**0.943**
fastsine	**0.249**	0.256	0.082	**0.011**	0.108	**0.014**	0.382	**0.911**
step	0.078	**0.061**	0.044	**0.017**	0.065	**0.022**	0.499	**0.768**
ramp	0.118	**0.110**	0.067	**0.020**	0.096	**0.027**	0.567	**0.942**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.060	**0.040**	0.048	**0.025**	0.076	**0.031**	**0.304**	-0.269
sine	0.118	**0.090**	0.108	**0.015**	0.146	**0.020**	0.278	**0.947**
fastsine	**0.248**	0.256	0.144	**0.008**	0.185	**0.011**	0.127	**0.924**
step	0.103	**0.067**	0.080	**0.014**	0.120	**0.018**	0.291	**0.788**
ramp	0.138	**0.114**	0.112	**0.017**	0.154	**0.022**	0.297	**0.945**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	**0.024**	0.029	**0.023**	0.037	**0.029**	**0.139**	-0.247
sine	0.086	**0.083**	0.043	**0.021**	0.060	**0.027**	0.772	**0.942**
fastsine	**0.253**	0.256	0.039	**0.014**	0.052	**0.017**	0.667	**0.901**
step	0.064	**0.057**	0.029	**0.017**	0.040	**0.022**	0.669	**0.792**
ramp	0.110	**0.108**	0.045	**0.023**	0.066	**0.031**	0.763	**0.940**

(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to stochastic volatility. A bold figure highlights the best model in each panel and row.

Turning to estimated covariances, Table 3 shows that the DC-CMSV model produces the most precise estimates for almost all considered correlation DGPs considered. MAD and RMSD statistics indicate that the distance between estimated covariances is small, while estimates from the CMSV model may exhibit substantial differences. Strik-

ingly, FD statistics of the DC-CMSV model are very close to one across all different correlation DGPs. Since the marginal volatilities in the DC-CMSV model are by construction independent of the ordering, these statistics show that estimated covariances are largely insensitive to the ordering of variables. Moreover, similarity and distance statistics indicate that the estimated covariances of the CMSV model substantially diverge as the idiosyncratic volatility patterns in the data grow stronger.

Table 3: Estimated covariance

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.313	**0.285**	0.169	**0.031**	0.240	**0.044**	0.378	**0.970**
sine	0.212	**0.199**	0.102	**0.021**	0.153	**0.030**	0.668	**0.981**
fastsine	**0.331**	**0.331**	0.106	**0.013**	0.159	**0.018**	0.622	**0.990**
step	0.248	**0.224**	0.109	**0.021**	0.167	**0.030**	0.566	**0.981**
ramp	0.232	**0.221**	0.104	**0.024**	0.158	**0.034**	0.642	**0.978**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.436	**0.389**	0.207	**0.037**	0.349	**0.057**	0.469	**0.988**
sine	0.288	**0.253**	0.179	**0.021**	0.307	**0.033**	0.572	**0.993**
fastsine	0.403	**0.396**	0.210	**0.012**	0.351	**0.019**	0.467	**0.997**
step	0.348	**0.301**	0.169	**0.021**	0.299	**0.033**	0.562	**0.994**
ramp	0.312	**0.281**	0.183	**0.023**	0.313	**0.037**	0.553	**0.992**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.238	**0.217**	0.153	**0.026**	0.201	**0.034**	0.192	**0.931**
sine	0.170	**0.163**	0.081	**0.023**	0.114	**0.030**	0.665	**0.963**
fastsine	**0.296**	0.296	0.049	**0.015**	0.068	**0.020**	0.768	**0.969**
step	0.190	**0.177**	0.089	**0.019**	0.130	**0.026**	0.473	**0.955**
ramp	0.190	**0.185**	0.085	**0.025**	0.121	**0.034**	0.637	**0.957**

(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to of stochastic volatility. A bold figure highlights the best model in each panel and row.

Thus, this Monte Carlo simulation provides evidence that CMSV model produces less precise and more divergent estimates of the time-varying covariance matrix as the idiosyncratic volatility patterns in the data grow stronger. It also demonstrates for a variety of data generating processes that estimates of the DC-CMSV model are almost insensitive to alternative orderings of variables.

5 Empirical Application

This section reviews two empirical applications to illustrate how a particular ordering of variables may drive conclusions. The first study to be reviewed is on time variation in U.S. monetary policy by Primiceri (2005) and the second is on inflation-gap persistence in the U.S. by Cogley, Primiceri, and Sargent (2010). In these studies, different versions of the CMSV model are used in conjunction with a time-varying parameter VAR model to study changes in the dynamic relationships between inflation, the unemployment rate and the interest rate. Moreover, this section compares these estimates to those coming from a DC-CMSV model developed in this paper.

5.1 Time variation in U.S. monetary policy

This section reviews the results for Primiceri's (2005) application in the light of alternative orderings. Figure 1 demonstrates that estimated covariances are sensitive to alternative orderings and exhibit marked differences during the stagflation period. Since estimated quantities in the structural analysis depend on the time-varying covariance matrix, some results may change under alternative orderings.

Before investigating this issue, the estimation of reduced-form time-varying parameter VAR models and the identification of structural parameters are briefly discussed, building upon the arguments in Primiceri (2005). Then, details for replicating the structural analysis using estimated parameters from these alternative VARs are provided. Assuming that ordering of variables is negligible, note that estimates from these alternative VARs should imply nearly the same reduced-form dynamics for the data and, thus, give rise to the same conclusions. The presented procedure is tailored to utilize the replication files of Del Negro and Primiceri (2015).

Estimation of reduced-form TVP-VAR: Let y_t be the vector of endogenous variables of dimension $n \times 1$ and let $\tilde{y}_t = Py_t$ be the vector of endogenous variables with exchanged rows where P is a permutation matrix that satisfies $P'P = I_n$. Assume that $\tilde{y}_t$ evolves according to a time-varying parameter VAR model

$$\tilde{y}_t = \tilde{c}_t + \tilde{B}_{1,t}\tilde{y}_{t-1} + ... + \tilde{B}_{k,t}\tilde{y}_{t-k} + \tilde{u}_t, \quad \tilde{u}_t \sim N(0, \tilde{\Sigma}_t) \tag{16}$$

where $\tilde{c}_t$ is of dimension $n \times 1$, $\tilde{B}_{i,t}$, $i = 1, ..., k$ is of dimension $n \times n$ and $\tilde{u}_t$ is of dimension $n \times 1$ or more compactly written in vectorized form

$$\tilde{y}_t = \tilde{X}_t' \tilde{B}_t + \tilde{u}_t, \quad \tilde{u}_t \sim N(0, \tilde{\Sigma}_t) \tag{17}$$

where $\tilde{X}_t' = I_n \otimes [1, \tilde{y}_{t-1}', ..., \tilde{y}_{t-k}']$ and $\tilde{B}_t = vec([\tilde{c}_t, \tilde{B}_{1,t}, ..., \tilde{B}_{k,t}]')$.

To model the evolution of $\tilde{\Sigma}_t$, the CMSV approach is used. Specifically, $\tilde{\Sigma}_t$ is decomposed by a triangular factorization, which is defined as

$$\tilde{\Sigma}_t = \tilde{A}_t^{-1} \tilde{D}_t \tilde{D}_t' \tilde{A}_t^{-1'}. \tag{18}$$

It follows that (17) can be rewritten as

$$\tilde{y}_t = \tilde{X}_t' \tilde{B}_t + \tilde{A}_t^{-1} \tilde{D}_t \tilde{\epsilon}_t, \quad \tilde{\epsilon}_t \sim N(0, I_n). \tag{19}$$

where state parameters and prior distributions are as in Primiceri (2005).

Identification: Next, consider the following structural VAR

$$\tilde{y}_t = \tilde{X}_t' \tilde{B}_t + \tilde{\Xi}_t \tilde{\epsilon}_t, \quad \tilde{\epsilon}_t \sim N(0, I_n),$$

which may differ from (19) because $\tilde{\Xi}_t$ is not necessarily lower triangular. Assuming that there are sufficient identification restrictions that meet some regularity conditions, the parameters in $\tilde{\Xi}_t$ may be exactly, partially or set identified, see Rubio-Ramirez, Waggoner, and Zha (2010) and Arias, Rubio-Ramírez, and Waggoner (2018).

The first step is to estimate (19) using the algorithms described in Del Negro and Primiceri (2015), then obtain posterior draws of the reduced-form time-varying VAR coefficients $\tilde{B}_t$'s and time-varying covariance matrices $\tilde{\Sigma}_t$'s in (17). The second step is to estimate $\tilde{\Xi}_t$ that satisfies

$$\tilde{\Xi}_t \tilde{\Xi}_t' = \tilde{\Sigma}_t$$

for each draw of $\tilde{\Sigma}_t$.[13]

Replication of structural analysis: Primiceri (2005) identifies a monetary policy shock by imposing zero restriction on the contemporaneous reaction of inflation and the unemployment rate. This is a triangular identification scheme for the variable ordering $y_t = [\pi_t, u_t, i_t]'$, which he also uses to estimate the parameters in (19) and hence, $\tilde{\Xi}_t = \tilde{A}_t^{-1} \tilde{D}_t$ for $P = I_n$. Therefore (19) is the structural VAR of interest. However,

[13] $\tilde{\Xi}_t$ may be estimated by using the algorithms of Arias, Rubio-Ramírez, and Waggoner (2018).

alternative estimates of $\tilde{\Xi}_t$ may be obtained by estimating reduced-form parameters first and then estimating the structural parameters in $\tilde{\Xi}_t$, which is not lower triangular for any other admissible permutation matrix except the identity matrix.

To replicate the structural analysis for alternative variable orderings, first $\tilde{B}_t$'s and $\tilde{\Sigma}_t$'s in (17) are estimated by making use of (18) and (19). Then, it is useful to reorder all parameters of the VAR model for $\tilde{y}_t$ in (17) such that they satisfy the variable ordering of y_t. The reordered reduced-form time-varying parameter VAR model can be obtained by pre-multiplying (16) by P'

$$P'\tilde{y}_t = P'\tilde{c}_t + P'\tilde{B}_{1,t}PP'\tilde{y}_{t-1} + ... + P'\tilde{B}_{k,t}PP'\tilde{y}_{t-k} + P'\tilde{u}_t, \quad P'\tilde{u}_t \sim N(0, P'\tilde{\Sigma}_t P)$$

which can be rewritten in compact vectorized form as

$$P'\tilde{y}_t = \bar{P}'\tilde{X}_t'\bar{P}'\bar{P}\tilde{B}_t + P'\tilde{u}_t, \quad P'\tilde{u}_t \sim N(0, P'\tilde{\Sigma}_t P)$$

where $\bar{P} = P \otimes \hat{P}$ and $\hat{P} = \begin{bmatrix} 1 & 0_{1\times(n\cdot k)} \\ 0_{(n\cdot k)\times 1} & [I_k \otimes P] \end{bmatrix}$. Then, the time-varying parameter VAR with original coefficient ordering is given by

$$y_t = X_t'B_t + u_t, \quad u_t \sim N(0, \Sigma_t)$$

where $X_t = \bar{P}'\tilde{X}_t'\bar{P}'$, $B_t = \bar{P}\tilde{B}_t$, $\Sigma_t = P'\tilde{\Sigma}_t P$ and $u_t = P'\tilde{u}_t$. Then, Ξ_t is estimated using a Cholesky decomposition of Σ_t. A_t^{-1} and D_t are obtained by solving $\Xi_t = A_t^{-1}D_t$.[14]

The time-varying parameter VAR model with CMSV (denoted as CMSV-TVP-VAR) is estimated using Algorithm 2 of Del Negro and Primiceri (2015), which is the approximate mixture sampler for stochastic volatility. In addition, a DC-CMSV version of the model is estimated for comparison (denoted as DC-CMSV-TVP-VAR). The Monte Carlo estimation produces 70,000 draws from the Gibbs sampler, while the first 20,000 are discarded in the burn-in period.

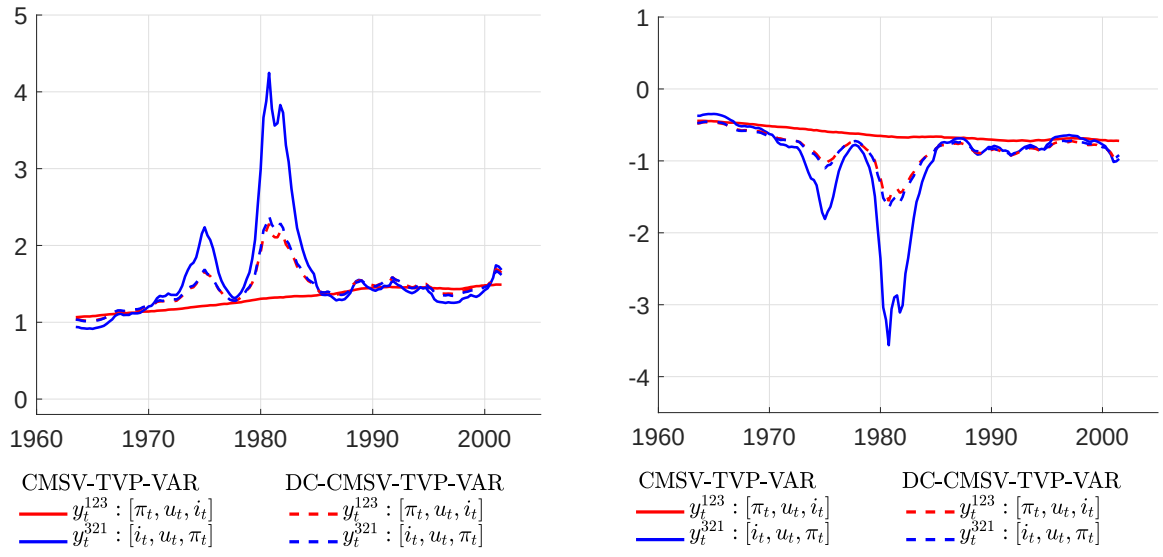
Estimated reduced-form TVP-VAR models: Before turning to the results of the structural analysis, it is instructive to assess the sensitivity of the $\tilde{B}_t$'s and the $\tilde{\Sigma}_t$'s. Figures 8 and 9 in Appendix C depict the posterior median of the estimated time-varying VAR parameters for each variable in the respective column. Alternative estimates for the $\tilde{B}_t$'s of the CMSV-TVP-VAR exhibit some minor differences across alternative ordering but overall they are rather similar. The estimated VAR coefficients of the DC-CMSV-TVP-VAR, however, are virtually indistinguishable.

[14]Note $D_t = \text{diag}[\text{vecd}(\Xi_t)]$, hence, $A_t^{-1} = \Xi_t \text{diag}[\text{vecd}(\Xi_t)]^{-1}$.

Next, Figures 10 and 11 in Appendix C show the posterior median of the estimated correlation, covariance, and volatility of the reduced-form residuals. For the CMSV-TVP-VAR model, the estimated correlation and covariance exhibit pronounced differences during the stagflation period. Differences across volatility estimates are, however, rather small. In contrast, estimates of the DC-CMSV-TVP-VAR model are virtually indistinguishable across alternative orderings. Moreover, estimates of the CMSV-TVP-VAR model are systematically different across orderings in the sense that they are below or above estimates of the DC-CMSV-TVP-VAR model, see Property 4. Taking stock, these properties indicate that potential differences in estimated quantities of the structural analysis are primarily driven by differences in estimated $\tilde{\Sigma}_t$'s but not $\tilde{B}_t$'s.

Revisiting the structural analysis: Having documented general differences across reduced-form parameter estimates, consequences are analyzed in more detail for two particular orderings. Specifically, estimates from the original variable ordering $y_t^{123} = [\pi_t, u_t, i_t]'$ and the reverse variable ordering $y_t^{321} = [i_t, u_t, \pi_t]'$ are contrasted because differences between estimated covariances are the most pronounced.

Figure 3: Long-run U.S. systematic interest rate response



(a) Inflation (b) Unemployment

The figure depicts interest rate response to a 1% permanent increase in inflation at the left panel (a) and the unemployment rate at the right panel (b) under alternative orderings.

Figure 3 shows the estimated long-run U.S. systematic interest rate response to inflation and unemployment for both the CMSV-TVP-VAR model and the DC-CMSV-

TVP-VAR model. In particular, the estimates of the former model provide evidence for two equally plausible but mutually exclusive conclusions as to how U.S. systematic monetary policy reacted during the stagflation period. Estimates obtained from the original variable ordering, y_t^{123}, provide evidence for a muted response, while those obtained under the reverse ordering of variables, y_t^{321}, point to a drastically changing and aggressive response. Thus, the choice of variable ordering may have a substantial effect on the estimates and may lead to alternative conclusions. This is clearly an undesirable property as estimates are associated with substantial model uncertainty that cannot be easily controlled for. Regarding the remaining structural analysis in Primiceri (2005), there are no marked differences for the other exercises as these do not strongly depend on the estimated covariances.[15]

Furthermore, estimates of the DC-CMSV-TVP-VAR model point to an unambiguous conclusion under all possible orderings. The estimates suggest that the reaction function was modestly more aggressive during the stagflation period. This evidence is consistent with the finding in Sims and Zha (2006), who provide strong evidence for regime switches in terms of how monetary policy was conducted during the period of stagflation.

5.2 Inflation-gap persistence in the U.S.

Next, the sensitivity of Cogley, Primiceri, and Sargent's (2010) empirical results are investigated using the replication files in the AEJ: Macroeconomics. The focus of this study differs from most applications of CMSV-TVP-VARs in that it uses this class of models to summarize reduced-form dynamic properties of variables. Specifically, the study is concerned with inflation-gap persistence but also report measures of trend-inflation, volatility, the conditional expectation of inflation based on unemployment news as well as Phillips correlations.

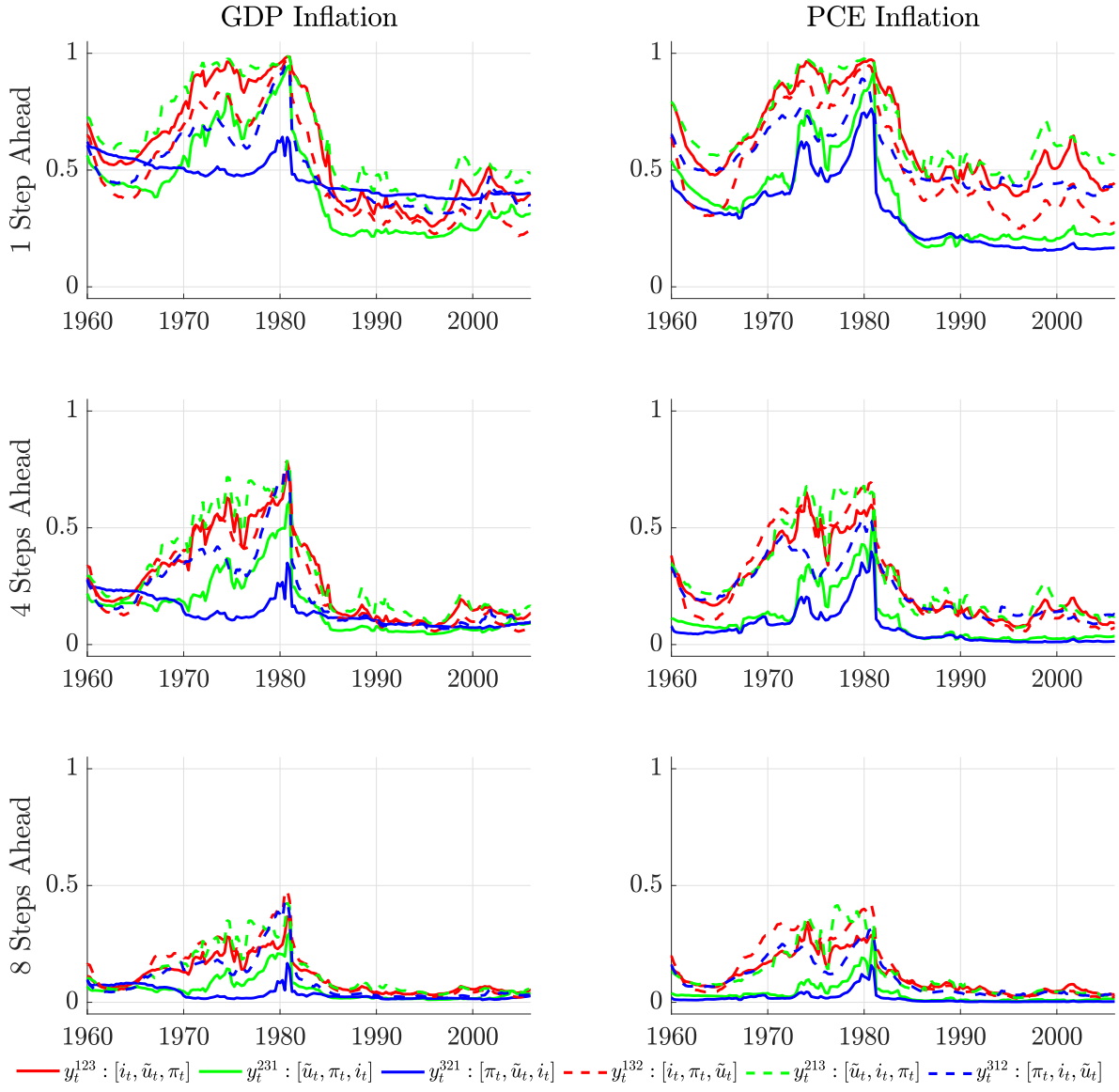
Cogley, Primiceri, and Sargent (2010) extend the TVP-VAR model of Cogley and Sargent (2005) by introducing heteroskedasticity to the time-varying VAR parameters, the $\tilde{B}_t$'s. These models are denoted as the CPS-TVPSV-VAR and the CS-TVP-VAR, respectively. Both models differ from the CMSV-TVP-VAR of Primiceri (2005) by the fact that they assume a constant lower triangular matrix of contemporaneous relations $\tilde{A}^{-1}$ instead of time-varying $\tilde{A}_t^{-1}$ in (18).[16] Note, the authors estimate their model with

[15] A comparison of all exercises is documented in Appendix C (Figures 12 – 19).

[16] By restricting the parameters of contemporaneous relations to being constant, the model-implied process for the correlation is solely determined by the volatility pattern and the selected ordering of variables. Recall that a reordering of variables leads to inversion for the ratio of volatilities.

the variable ordering $y_t^{123} = [i_t, \tilde{u}_t, \pi_t]'$, where $\tilde{u}_t$ is the logit of the unemployment rate, i.e. which is the reverse ordering as compared to Primiceri's (2005) application.[17]

Figure 4: $R^2_{j,t}$ statistics



The figure depicts the posterior median for $R^2_{j,t}$ statistics based on the CPS-TVPSV-VAR, with estimates for GDP inflation shown in the left panel and those for PCE inflation shown in the right panel.

Figure 4 shows posterior median estimates of the $R^2_{j,t}$ statistics, which are used as

[17] Cogley, Primiceri, and Sargent (2010) select this ordering as they follow Cogley and Sargent (2005), who document that this ordering minimizes the drift in the $\tilde{B}_t$'s.

a measure of inflation-gap persistence, for GDP inflation and PCE inflation obtained under all possible orderings of the variables. The graph shows that alternative estimates may lead to a different conclusion as to how inflation-gap persistence in the U.S. evolved during the stagflation period. Particularly, estimates from the alternative ordering $y_t^{321} = [\pi_t, \tilde{u}_t, i_t]$ for GDP inflation suggest that persistence declined gradually over time. This is in stark contrast to the estimates coming from y_t^{123} which indicate that persistence significantly surged during the Great Inflation period and declined after 1980. In addition to qualitative patterns differing, quantitative differences across estimates are substantial. For instance, estimated persistence ranges from 0.6 up to around one for the one-step ahead horizon of both GDP inflation and PCE inflation in the mid-1970s. This marked difference in estimates arises due to estimates of trend-inflation and volatility being substantially different, see Figures **??** and **??** in Appendix D.

Based on this evidence, the question of how much the introduction of CMSV heteroskedasticity induces sensitivity to the estimates arises. To answer this question, estimates from three alternative models are considered. For the first two models, heteroskedasticity for innovations of the $\tilde{B}_t$'s is shut-off. The first model is the CS-TVP-VAR of Cogley and Sargent (2005) and the second is a DC-CMSV version of the CS-TVP-VAR, denoted as the DC-CMSV-TVP-VAR. The third model assumes independent volatility dynamics for the innovations of the $\tilde{B}_t$'s and assumes DC-CMSV for the VAR residuals.[18] This model is denoted as the DC-CMSV-TVPSV-VAR.

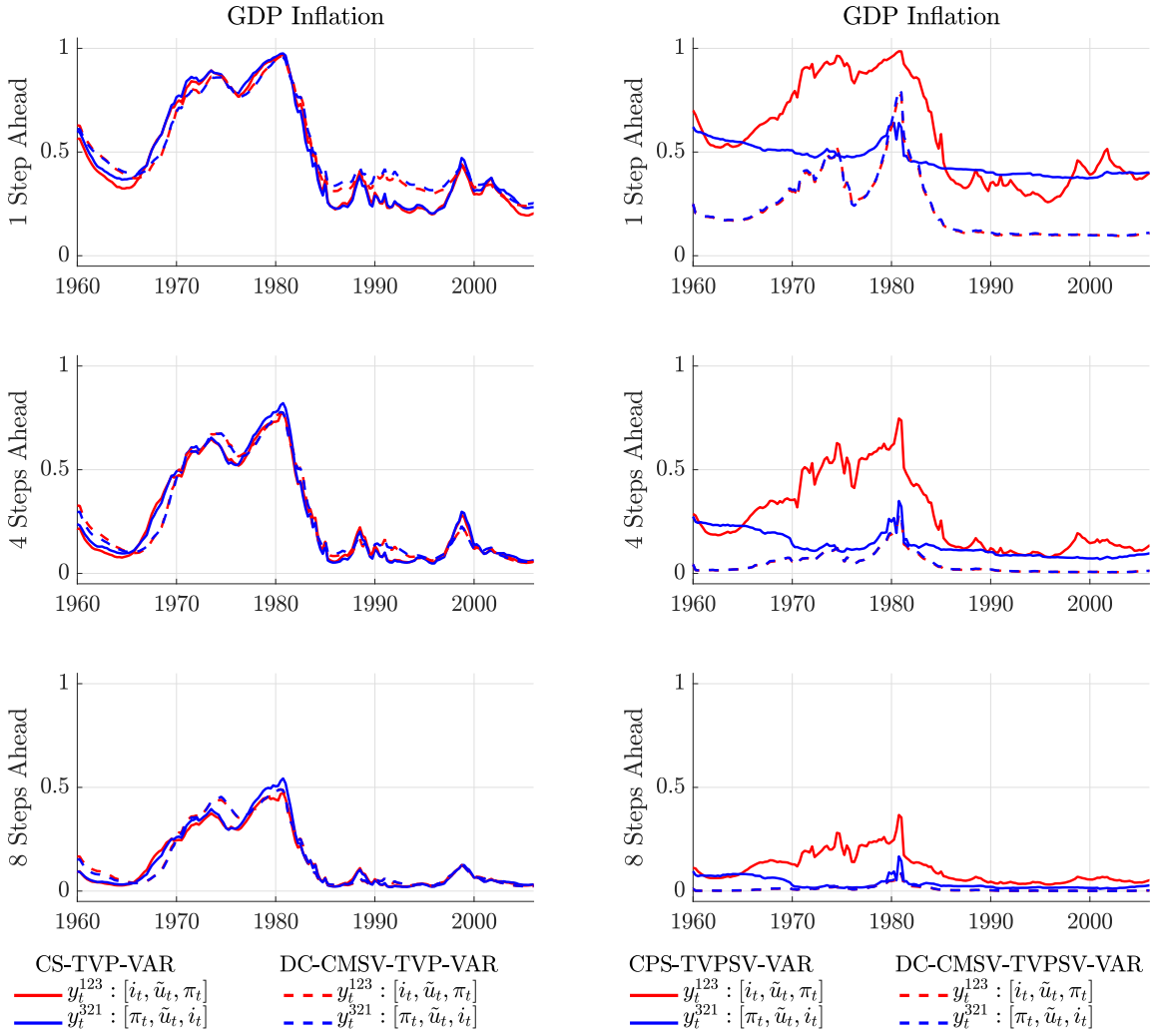
Figure 5 shows posterior medians of $R^2_{j,t}$ statistics for GDP inflation for the variable ordering y_t^{123} and y_t^{321}, with homoskedastic and heteroskedastic innovations for $\tilde{B}_t$'s in the left panel (a) and the right panel (b), respectively. Estimates of $R^2_{j,t}$ statistics from the CS-TVP-VAR and the DC-CMSV-TVP-VAR in panel (a) indicate that the ordering of variables has a limited effect on the estimated path and magnitude of inflation-gap persistence.[19] Thus, shutting off CMSV heteroskedasticity for innovations of the $\tilde{B}_t$'s eliminates the bulk of model and parameter uncertainty for this quantity of interest. Apart from that, estimates from both models suggest that inflation-gap persistence in the U.S. surged during the stagflation period and dropped significantly after the 1980s, confirming the result in Cogley, Primiceri, and Sargent (2010).

Next, estimates from the DC-CMSV-TVPSV-VAR in panel (b) show that estimated inflation-gap persistence is substantially lower and exhibits a moderately different pattern around the stagflation period. Particularly, it features less stickiness in terms

[18]Independent volatility processes instead of a full time-varying covariance matrix for the parameter innovations are assumed to limit model and computational complexity.

[19]Estimates from alternative orderings are similar.

Figure 5: $R^2_{j,t}$ statistics



(a) Homoskedastic innovations for $\tilde{B}_t$'s

(b) Heteroskedastic innovations for $\tilde{B}_t$'s

The figure depicts posterior median of $R^2_{j,t}$ statistics for GDP inflation from TVP-VARs with homoskedastic and heteroskedastic parameter innovations for $\tilde{B}_t$'s at the left panel (a) and the right panel (b), respectively.

of the level shifts depicted in panel (a). Also, notice that for the CPS-TVPSV-VAR reduced-form estimates are affected not only by variable ordering but also by parameter ordering. For instance, when innovations of the parameter of constants, $\tilde{c}_t$'s, have different volatility patterns, then exchanging the order of parameters affects estimates as well. Therefore, this comparison illustrates that modelling heteroskedasticity via the CMSV approach may lead to very different estimates of $\tilde{B}_t$'s and $\tilde{\Sigma}_t$'s, which may

eventually alter inference on reduced-form properties for variables of interest.

Appendix D presents further sensitivity analysis across models for trend-inflation, volatility, the adjustment of inflation expectation to unemployment news and Phillips correlations. Overall, estimated quantities of the CPS-TVPSV-VAR model exhibit sometimes subtle and sometimes substantial differences across alternative variable orderings. Comparing these results with those coming from the CS-TVP-VAR of Cogley and Sargent (2005), most ambiguity between estimates vanishes. It only remains relevant for quantities that involve estimates of the covariances such as conditional inflation expectations and Phillips correlations. Furthermore, estimates from these alternative models indicate that most of the empirical evidence provided by Cogley, Primiceri, and Sargent (2010) can be qualitatively confirmed but that there may be a broad range of possible values.

6 Conclusion

This paper studied the consequences of exchanging the order of variables on the dynamic properties of the time-varying covariance matrix in the CMSV model. The paper found that alternative dynamic restrictions are imposed on the time-varying covariance matrix when the ratio of reduced-form volatilities is time-varying. Simulations demonstrated that the stronger the idiosyncratic volatility pattern in the data, the more divergent the estimates of the time-varying covariance matrix. The DC-CMSV model was proposed as a robust alternative, which produces almost rotationally invariant estimates of the time-varying covariance matrix. An important feature of the DC-CMSV model is that the parameters of contemporaneous relation may evolve nonlinearly. For the two empirical applications considered, it was illustrated that variable orderings may substantially affect estimates and may give rise to alternative conclusions.

The results of this paper suggest that the ordering of variables is a nontrivial choice when estimating a time-varying covariance matrix via the CMSV approach. Moreover, the sensitivity of the estimates documented in this paper is likely to be present in other empirical applications as well since it is not uncommon for financial and economic time series to exhibit individual volatility dynamics. For this reason, estimates based on the CMSV model should be interpreted with some caution and may not as robust as they seem. In addition, the relatively small costs of computing a more robust estimate using the DC-CMSV model seem worthwhile for the purposes of most empirical applications.

The finding that specific assumptions about the state dynamics coupled with the chosen factorization of the time-varying covariance matrix may impose different restrictions across alternative variable orderings may not only be limited to this state

space model. The linear dynamic factor model with time-varying factor loadings and stochastic volatility may also suffer from similar restrictions. Future research should investigate whether rotational non-invariance in this class of models may be also driven by an overly restrictive evolution of the state or factor dynamics.

References

Amir-Ahmadi, P., C. Matthes, and M.-C. Wang (2018): "Choosing Prior Hyperparameters: With Applications to Time-Varying Parameter Models," *Journal of Business & Economic Statistics*, pp. 1–13.

Arias, J. E., J. F. Rubio-Ramírez, and D. F. Waggoner (2018): "Inference Based on Structural Vector Autoregressions Identified With Sign and Zero Restrictions: Theory and Applications," *Econometrica*, 86(2), 685–720.

Asai, M., and M. McAleer (2009): "The structure of dynamic correlations in multivariate stochastic volatility models," *Journal of Econometrics*, 150(2), 182–192.

Asai, M., M. McAleer, and J. Yu (2006): "Multivariate stochastic volatility: a review," *Econometric Reviews*, 25(2-3), 145–175.

Baumeister, C., and G. Peersman (2013a): "The role of time-varying price elasticities in accounting for volatility changes in the crude oil market," *Journal of Applied Econometrics*, 28(7), 1087–1109.

——— (2013b): "Time-varying effects of oil supply shocks on the US economy," *American Economic Journal: Macroeconomics*, 5(4), 1–28.

Benati, L. (2008): "The "Great Moderation" in the United Kingdom," *Journal of Money, Credit and Banking*, 40(1), 121–147.

Benati, L., and P. Surico (2008): "Evolving U.S. Monetary Policy and the Decline of Inflation Predictability," *Journal of the European Economic Association*, 6(2-3), 634–646.

Bognanni, M. (2018): "A Class of Time-Varying Parameter Structural VARs for Inference under Exact or Set Identification," *FRB Cleveland Working Paper no. 18-11*.

Bollerslev, T. (1990): "Modelling the coherence in short-run nominal exchange rates: a multivariate generalized ARCH model," *The Review of Economics and Statistics*, pp. 498–505.

Chan, J. (2017): "Notes on Bayesian Macroeconometrics," unpublished.

Chan, J. C., A. Doucet, R. León-González, and R. W. Strachan (2018): "Multivariate stochastic volatility with co-heteroscedasticity," unpublished.

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.

Cogley, T., G. E. Primiceri, and T. J. Sargent (2010): "Inflation-Gap Persistence in the US," *American Economic Journal: Macroeconomics*, 2(1), 43–69.

Cogley, T., and T. J. Sargent (2005): "Drifts and volatilities: monetary policies and outcomes in the post WWII US," *Review of Economic Dynamics*, 8(2), 262–302.

Del Negro, M., and G. E. Primiceri (2015): "Time Varying Structural Vector Autoregressions and Monetary Policy: A Corrigendum," *The Review of Economic Studies*, 82(4), 1342–1345.

Engle, R. (2002): "Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models," *Journal of Business & Economic Statistics*, 20(3), 339–350.

Galí, J., and L. Gambetti (2015): "The effects of monetary policy on stock market bubbles: Some evidence," *American Economic Journal: Macroeconomics*, 7(1), 233–257.

Gali, J., and L. Gambetti (2009): "On the Sources of the Great Moderation," *American Economic Journal: Macroeconomics*, 1(1), 26–57.

Gambetti, L., and A. Musso (2017): "Loan supply shocks and the business cycle," *Journal of Applied Econometrics*.

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.

Koop, G., R. León-González, and R. W. Strachan (2009): "On the evolution of the monetary policy transmission mechanism," *Journal of Economic Dynamics and Control*, 33(4), 997–1017.

Lopes, H. F., R. E. McCulloch, and R. Tsay (2012): "Cholesky stochastic volatility models for high-dimensional time series," unpublished.

Mumtaz, H., and F. Zanetti (2013): "The Impact of the Volatility of Monetary Policy Shocks," *Journal of Money, Credit and Banking*, 45(4), 535–558.

Nakajima, J., and T. Watanabe (2011): "Bayesian analysis of time-varying parameter vector autoregressive model with the ordering of variables for the Japanese economy and monetary policy," unpublished.

Philipov, A., and M. E. Glickman (2006): "Multivariate Stochastic Volatility via Wishart Processes," *Journal of Business & Economic Statistics*, 24(3), 313–328.

Pourahmadi, M. (1999): "Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation," *Biometrika*, 86(3), 677–690.

Prieto, E., S. Eickmeier, and M. Marcellino (2016): "Time Variation in Macro-Financial Linkages," *Journal of Applied Econometrics*, 31(7), 1215–1233.

Primiceri, G. E. (2005): "Time varying structural vector autoregressions and monetary policy," *The Review of Economic Studies*, 72(3), 821–852.

Rubio-Ramirez, J. F., D. F. Waggoner, and T. Zha (2010): "Structural vector autoregressions: Theory of identification and algorithms for inference," *Review of Economic Studies*, 77(2), 665–696.

Sims, C. A., and T. Zha (2006): "Were There Regime Switches in U.S. Monetary Policy?," *American Economic Review*, 96(1), 54–81.

Tsay, R. S. (2005): *Analysis of financial time series*, Wiley Series in Probability and Statistics. John Wiley & Sons, New Jersey, 2 edn.

Uhlig, H. (1997): "Bayesian Vector Autoregressions with Stochastic Volatility," *Econometrica*, 65(1), 59–73.

Yu, J., and R. Meyer (2006): "Multivariate Stochastic Volatility Models: Bayesian Estimation and Model Comparison," *Econometric Reviews*, 25(2-3), 361–384.

A Proofs

A.1 Some properties of the Cholesky MSV model

Proof of Property "Σ_t under CMSV model" .

Subproof of Claim (1). Define $b_t = \frac{\sigma_{11,t}^2}{\sigma_{22,t}^2}$. Then,

$$\begin{aligned}
\sigma_{22,t}^2 &= \exp(2g_{2,t}) + a_t^2 b_t \sigma_{22,t}^2 \\
&= \frac{1}{1 + b_t a_t^2} \exp(2g_{2,t}) \\
&= \frac{1}{1 + b_t(a_{t-1} + \epsilon_t^a)^2} \exp(2g_{2,t-1}) \exp(2\epsilon_{2,t}^g) \\
&= \frac{1 + b_t a_{t-1}^2}{1 + b_t(a_{t-1} + \epsilon_t^a)^2} \sigma_{22,t-1}^2 \exp(2\epsilon_{2,t}^g). \qquad (20)
\end{aligned}$$

Using $\sigma_{11,t}^2 = b_t \sigma_{22,t}^2$, it follows that

$$\sigma_{11,t}^2 = \frac{1 + b_t a_{t-1}^2}{1 + b(a_{t-1} + \epsilon_t^a)^2} \sigma_{11,t-1}^2 \exp(2\epsilon_{2,t}^g). \qquad (21)$$

However, the state equation for $\sigma_{11,t}^2$ is given by

$$\sigma_{11,t}^2 = \sigma_{11,t-1}^2 \exp(2\epsilon_{1,t}^g). \qquad (22)$$

Combining (21) with (22) gives

$$\begin{aligned}
\frac{1 + b_t a_{t-1}^2}{1 + b_t(a_{t-1} + \epsilon_t^a)^2} \sigma_{11,t-1}^2 \exp(2\epsilon_{2,t}^g) &= \sigma_{11,t-1} \exp(2\epsilon_{1,t}^g) \\
\frac{1 + b_t a_{t-1}^2}{1 + b(a_{t-1} + \epsilon_t^a)^2} \exp(2\epsilon_{2,t}^g) &= \exp(2\epsilon_{1,t}^g). \qquad (23)
\end{aligned}$$

Since ϵ_t^a is independent of $\{\epsilon_{1,t}^g, \epsilon_{2,t}^g\}$, b_t must be time-varying to ensure that this equation holds in every period. Thus, the ratio of volatilities is not constant. ■

Subproof of Claim (2). The correlation ρ_t depends on reduced-form parameters by

$$\begin{aligned}
\rho_t &= a_t \frac{\sigma_{11,t}}{\sigma_{22,t}} \\
&= a_{t-1} \frac{\sigma_{11,t-1}}{\sigma_{22,t-1}} \frac{\exp(\epsilon_{1,t}^g)}{\exp(\epsilon_{2,t}^{g**})} + \epsilon_t^a \frac{\sigma_{11,t}}{\sigma_{22,t}} \\
&= \rho_{t-1} \frac{\exp(\epsilon_{1,t}^g)}{\exp(\epsilon_{2,t}^{g**})} + \epsilon_t^a \frac{\sigma_{11,t}}{\sigma_{22,t}} \qquad (24)
\end{aligned}$$

where $\epsilon_{2,t}^{g**} \equiv \log(\sigma_{22,t}) - \log(\sigma_{22,t-1})$. Then, because the ratio of volatilities is time-varying and log-normally distributed, it follows from (24) that the correlation evolves nonlinearly. ■

Subproof of Claim (3). Follows directly from the definition. ■

□

Proof of Property "Reordering in CMSV model".

Under $\tilde{\Sigma}_t$, the true parameter $\tilde{a}_t$ is given by

$$\tilde{a}_t = a_t \frac{\sigma_{11,t}^2}{\sigma_{22,t}^2}.$$

Then, the transition equation for the implied contemporaneous relation parameter, $\tilde{a}_t$ is given by

$$\tilde{a}_t = \tilde{a}_{t-1} \frac{\exp\left(2\epsilon_{1,t}^{g}\right)}{\exp\left(2\epsilon_{2,t}^{g**}\right)} + \epsilon_t^a \frac{\sigma_{11,t}^2}{\sigma_{22,t}^2},$$

where $\epsilon_{2,t}^{g**} \equiv \log(\sigma_{22,t}) - \log(\sigma_{22,t-1})$.

The time-varying ratio of reduced-form variances implies that $\tilde{a}_t$ evolves nonlinearly. Specifically, the transition from $\tilde{a}_{t-1}$ to $\tilde{a}_t$ is leveraged or dampened by the innovations to stochastic volatility as well as the ratio of variances itself. In contrast, the state equation of $\tilde{a}_t^*$ is a Gaussian random walk. Consequently, the true dynamics of $\tilde{a}_t$ cannot be obtained by the state equation of $\tilde{a}_t^*$. Hence, the dynamic structures induced into Σ_t^* and $\tilde{\Sigma}_t$ are different.

The parameters under the CMSV model set up analogously for $\tilde{\Sigma}_t^*$ are given by

$$\begin{aligned} \tilde{\sigma}_{22,t}^{*2} &= \exp(2\tilde{g}_{2,t}^*), & \tilde{\sigma}_{11,t}^{*2} &= \exp(2\tilde{g}_{1,t}^*) + (\tilde{a}_t^*)^2 \exp(2\tilde{g}_{2,t}^*) \\ \tilde{\sigma}_{12,t}^* &= \tilde{a}_t^* \tilde{\sigma}_{22,t}^{*2}, & \tilde{\rho}_t^* &= \tilde{a}_t^* \frac{\tilde{\sigma}_{22,t}^*}{\tilde{\sigma}_{11,t}^*} \end{aligned}$$

Since the ratio of variances cannot be constant for this DGP, the dynamic path of these parameters departs from the dynamic path of the true parameters

$$\tilde{\sigma}_{12,t} = \tilde{a}_t \tilde{\sigma}_{11,t}^2 = \sigma_{12,t}, \quad \tilde{\rho}_t = a_t \frac{\sigma_{11,t}}{\sigma_{22,t}} = \rho_t,$$

as the variability of the ratio of variances increases.

□

A.2 The Cholesky MSV model and the DC-MSV model

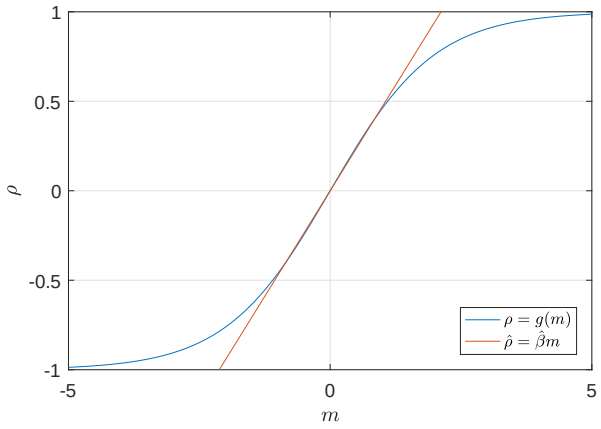
Property 6 (Rotational invariance of Σ_t under DC-MSV model)**.** Let y_t be generated by the DC-MSV model with covariance matrix Σ_t. Define the vector of variables with exchanged rows $\tilde{y}_t$ and the permuted covariance matrix $\tilde{\Sigma}_t = P\Sigma_t P'$. Analogously, define $\tilde{\Sigma}_t^* = D_t^* R_t^* D_t^*$, the covariance matrix of $\tilde{y}_t = Py_t$ where P is a permutation matrix. Then, $\tilde{\Sigma}_t = \tilde{\Sigma}_t^*$, i.e. the reduced-form parameters of Σ_t are independent of the ordering of the variables.

Proof. If $\tilde{\Sigma}_t = \tilde{\Sigma}_t^*$, then $P'\tilde{\Sigma}_t^* P = P'D_t^* PP'R_t^* PP'D_t^* P$ since $D_t = P'D_t^* P$ (D_t^* is diagonal) and $R_t = P'R_t^* P = R_t^*$ (R_t^* is symmetric). It follows that $\Sigma_t = P'\tilde{\Sigma}_t^* P$. □

Proof of Property "Σ_t under DC-MSV model".

Subproof of Claim (1). For $g : m \to \rho$ and $m \in [-1.1, 1.1]$ we have $\rho = g(m) \in [-0.5, 0.5]$. On this interval, the MSE of a linear regression of ρ on m is 5.5e-5. Figure 6 compares the linear prediction for ρ on the interval for $m \in [-5, 5]$. The figure indicates that when $|m| > 1.1$, the approximation error increases substantially as the function g becomes more nonlinear. Thus, for $\rho \in (-0.5, 0.5)$, the mapping $g(m)$ is approximately linearly and the innovations are approximately Gaussian.

Figure 6: Fisher transformation: mapping between ρ and m



■

Subproof of Claim (2). The transition equations for the implied contemporaneous relations under the two alternative orderings a_t and $\tilde{a}_t$ are given by

$$a_t = a_{t-1}\frac{\exp\left(\epsilon_{2,t-1}^h\right)}{\exp\left(\epsilon_{1,t-1}^h\right)} + \eta_t^\rho \frac{\sigma_{22,t}}{\sigma_{11,t}}, \qquad \tilde{a}_t = \tilde{a}_{t-1}\frac{\exp\left(\epsilon_{1,t-1}^h\right)}{\exp\left(\epsilon_{2,t-1}^h\right)} + \eta_t^\rho \frac{\sigma_{11,t}}{\sigma_{22,t}}$$

where $\eta_t^\rho \equiv \rho_t - \rho_{t-1}$. Then, when the ratio of reduced-form volatilities is constant, that is, $\frac{\sigma_{22,t}}{\sigma_{11,t}} = c \; \forall t, c > 0$, it follows that

$$a_t = a_{t-1} + \eta_t^\rho c, \qquad \tilde{a}_t = \tilde{a}_{t-1} + \eta_t^\rho \frac{1}{c}, \tag{25}$$

Thus, the dynamic evolution of a_t and of $\tilde{a}_t$ are driven by the correlation process. Since $\tilde{a}_t = a_t \frac{1}{c^2}$, the dynamic evolution of $\tilde{a}_t$ and of a_t are the same up to a positive scalar. ■

Subproof of Claim (3). When the ratio of reduced-form volatilities is time-varying, then the transition from a_{t-1} to a_t is nonlinear as it is scaled by the log-normally distributed ratio of volatilities. Then, after a reordering of variables, the influence of the ratio of volatilities on the contemporaenous relation is inverted. This means that the distance from a_{t-1} to a_t and from $\tilde{a}_{t-1}$ to $\tilde{a}_t$ is not symmetric. Therefore, a_t and $\tilde{a}_t$ obey different nonlinear dynamics. ■

□

Proof of Property "DC-MSV, CMSV and implied covariances".

The true dynamic structure for the contemporaneous relation is given by

$$a_t = a_{t-1} \frac{\exp\left(\eta_{2,t}^h\right)}{\exp\left(\eta_{1,t}^h\right)} + \eta_t^\rho \frac{\exp\left(h_{2,t}\right)}{\exp\left(h_{1,t}\right)}. \tag{26}$$

This equation substantially differs from the linear Gaussian process of a_t^*. Specifically, it features state dependent time-varying parameters, non-normal and heteroskedastic innovations that may leverage or dampen the transition from a_{t-1} to a_t.

To quantify the impact of these nonlinearities, the equation is linearized using a first order Taylor series expansion with information up to $t-1$, i.e. $a^0 = a_{t-1}$, $\eta^{\rho,0} = 0$, $\eta_i^{h,0} = E(\eta_{i,t}^h) = 0$ and $h_i^0 = h_{i,t-1}$ for $i = 1, 2$. The linearization is given by

$$a_t = a_{t-1} + (a_{t-1} + \eta_t^\rho \frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})})(\eta_{2,t}^h - \eta_{1,t}^h) + \eta_t^\rho \frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})}. \tag{27}$$

This linearization features an approximation error, except when the innovations to stochastic volatility offset each other, i.e. the ratio of volatilities is constant.

The approximation error is defined as

$$error_t = a_t - \hat{a}_t \tag{28}$$

where a_t and $\hat{a}_t$ denote the resulting parameter under (26) and (27), respectively. The bias associated with this linear transition function is given by

$$bias_t = 1_{\{a_t > a_{t-1}\}}(\hat{a}_t - a_t) - (1 - 1_{\{a_t > a_{t-1}\}})(\hat{a}_t - a_t) \tag{29}$$

where $1_{\{a_t > a_{t-1}\}}$ is an indicator function, which ensures the correct sign of the bias.[20]

Figure 7 illustrates the quantitative effects of innovations to stochastic volatility and of innovations to correlation on the approximation error and the bias for two initial points $(a_{t-1} = 0, \rho_{t-1} = 0)$ and $(a_{t-1} = 0.5, \rho_{t-1} = 0.5)$ with $\frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})} = 1$.[21] The true parameter a_t moves on an exponential hyperplane while the first order approximation $\hat{a}_t$ moves on a linear hyperplane, which touches the true hyperplane from below (above, indefinite) for positive (negative, zero) values of a_{t-1}.

For the first point, when there are non-offsetting innovations to stochastic volatility, then the approximation error is non-negative (non-positive) as the true value a_t is above (below) the initial value a_{t-1}. Consequently, the bias is negative in either direction. In other words, the first order approximation underestimates any transition from this point. For the second point, the approximation error is, in general, non-negative since the first order approximation touches the exponential hyperplane from below. Thus, when the ratio of volatilities increases, the true value increases and the first order approximation underestimates this transition. In contrast, it generally overestimates the transition when the ratio of volatilities decreases.

Note that when the linear approximation in (27) is further restricted to exhibit homoskedastically and normally distributed innovations, then these dynamic restrictions become tighter such that the approximation error and the bias become larger.

Next, when the order of variables is changed, then $\tilde{a}_t$ has a similar functional form as a_t in (26), but the ratio of volatilities is inverted

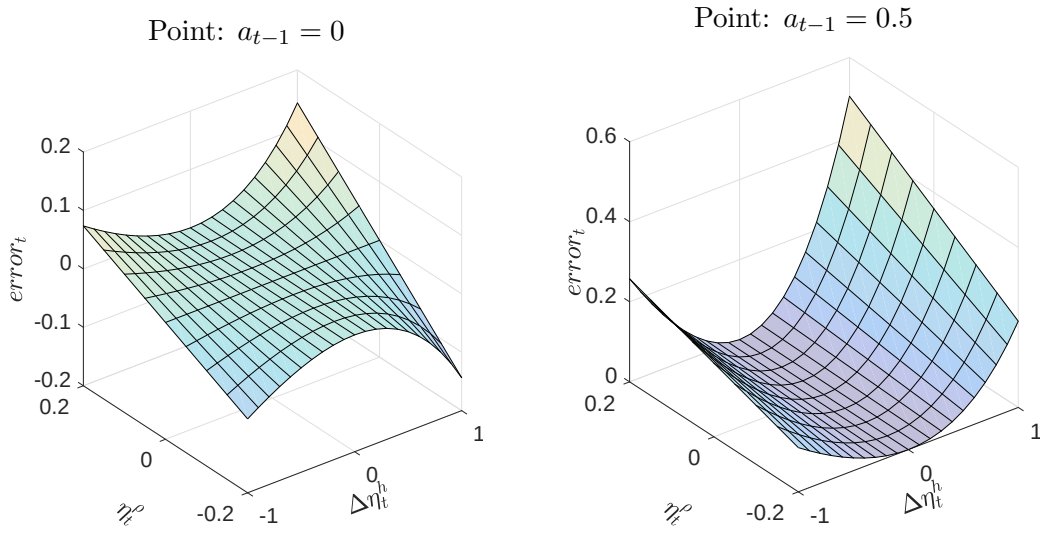
$$\tilde{a}_t = \tilde{a}_{t-1}\frac{\exp(\eta^h_{1,t})}{\exp(\eta^h_{2,t})} + \eta^\rho_t \frac{\exp(h_{1,t})}{\exp(h_{2,t})}.$$

Consequently, when the ratio of volatilities increases then the equation of a^*_t underestimates the true transition of a_t in the original ordering, while the dynamic equation of $\tilde{a}^*_t$ mechanically overestimates the true transition of $\tilde{a}_t$ in the alternative ordering.

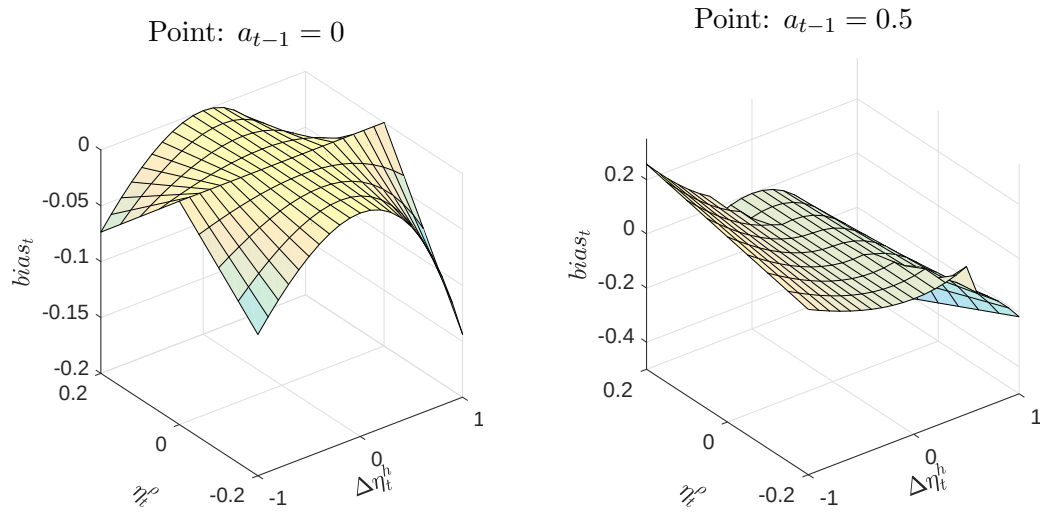
[20] For instance, when the true value falls and the approximate value falls by even more but both remain positive, then the transition is overstated. However, a bias function without sign correction assigns a negative value, indicating underestimation.

[21] Notice the surface plots for a positive value of a_{t-1} are a reflection for negative value of a_{t-1}.

Figure 7: Approximation error and bias of linearization



Panel (I): Approximation error



Panel (II): Bias

The figure shows the approximation error and the bias for two initial values of $(a_{t-1} = 0, \rho_{t-1} = 0)$ and $(a_{t-1} = 0.5, \rho_{t-1} = 0.5)$ with $\frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})} = 1$. The x-axis and the y-axis show the range of the innovations to correlation and of the ratio of volatility.

The covariance terms $\sigma^*_{12,t}$ and $\tilde{\sigma}^*_{12,t}$ are proportional to $\{a^*_t, g^*_{1,t}\}$ and $\{\tilde{a}^*_t, \tilde{g}^*_{2,t}\}$, respectively. $h_{1,t} = g^*_{1,t}$ and $h_{2,t} = \tilde{g}^*_{2,t}$ are left unrestricted. It follows that the bias in the contemporaneous parameter carries over to the covariance terms. □

Proof of Property "Posterior distribution of a_t and $\tilde{a}_t$ under homoskedasticity".

Assume that y_t is generated by

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} \sim N\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & \rho_t \\ \rho_t & 1 \end{bmatrix} \right).$$

Define for y_t and $\tilde{y}_t = Py_t$, where P is a permutation matrix exchanging rows, the respective covariance matrices $\Sigma_t = A_t^{-1} D_t D_t' A_t^{-1'}$ and $\tilde{\Sigma}_t = \tilde{A}_t^{-1} \tilde{D}_t \tilde{D}_t' \tilde{A}_t^{-1'}$. In addition, assume the variance of the first element on the diagonal of D_t and $\tilde{D}_t$ is equal to one. The parameters associated with y_t are $\{1, g_t, a_t\}$ and those with $\tilde{y}_t$ are $\{1, \tilde{g}_t, \tilde{a}_t\}$. Due to the special structure of the DGP, it follows that the prior distribution of $a_t = \tilde{a}_t$ and $g_t = \tilde{g}_t$ is invariant to rotation of variables.[22]

Turning to inference, suppose the posterior draw for the initial value $a_0 = \tilde{a}_0$, the variance of the time-varying parameter $S = \tilde{S}$ and the variance of the transformed second variable $g_t = \tilde{g}_t \ \forall t$.

Using the results in Chan (2017), the posterior distribution of a is given by

$$(a|y, D, a_0, S) \sim N(K_a^{-1}\bar{a}, K_a^{-1})$$

where

$$K_a = \begin{bmatrix} \frac{2}{S} + \frac{{y_{1,1}}^2}{\exp(g_1)} & -\frac{1}{S} & 0 & 0 & \dots & 0 \\ -\frac{1}{S} & \frac{2}{S} + \frac{{y_{1,2}}^2}{\exp(g_2)} & -\frac{1}{S} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -\frac{1}{S} & \frac{2}{S} + \frac{{y_{1,T-1}}^2}{\exp(g_{T-1})} & -\frac{1}{S} \\ 0 & 0 & 0 & 0 & -\frac{1}{S} & \frac{2}{S} + \frac{{y_{1,T}}^2}{\exp(g_T)} \end{bmatrix}, \bar{a} = \begin{bmatrix} \frac{a_0}{S} - \frac{y_{1,1}y_{2,1}}{\exp(g_1)} \\ -\frac{y_{1,2}y_{2,2}}{\exp(g_2)} \\ \vdots \\ -\frac{y_{1,T-1}y_{2,T-1}}{\exp(g_{T-1})} \\ -\frac{y_{1,T}y_{2,T}}{\exp(g_T)} \end{bmatrix}.$$

while the posterior distribution of $\tilde{a}$ is given by

$$(a|\tilde{y}, D, a_0, S) \sim N(\tilde{K}_a^{-1}\tilde{\bar{a}}, \tilde{K}_a^{-1})$$

where

$$\tilde{K}_a = \begin{bmatrix} \frac{2}{S} + \frac{{y_{2,1}}^2}{\exp(g_1)} & -\frac{1}{S} & 0 & 0 & \dots & 0 \\ -\frac{1}{S} & \frac{2}{S} + \frac{{y_{2,2}}^2}{\exp(g_2)} & -\frac{1}{S} & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & -\frac{1}{S} & \frac{2}{S} + \frac{{y_{2,T-1}}^2}{\exp(g_{T-1})} & -\frac{1}{S} \\ 0 & 0 & 0 & 0 & -\frac{1}{S} & \frac{2}{S} + \frac{{y_{2,T}}^2}{\exp(g_T)} \end{bmatrix}, \tilde{\bar{a}} = \begin{bmatrix} \frac{a_0}{S} - \frac{y_{2,1}y_{1,1}}{\exp(g_1)} \\ -\frac{y_{2,2}y_{1,2}}{\exp(g_2)} \\ \vdots \\ -\frac{y_{2,T-1}y_{1,T-1}}{\exp(g_{T-1})} \\ -\frac{y_{2,T}y_{1,T}}{\exp(g_T)} \end{bmatrix}$$

[22] Notice that this does not imply that the elements in Σ_t and $\tilde{\Sigma}_t$ have the same distribution.

Then, since $\bar{a} = \tilde{\bar{a}}$ but $K_a \neq \tilde{K}_a$ unless $y_{1,t}^2 = y_{2,t}^2 \ \forall t$, it follows that the posterior distribution of a and $\tilde{a}$ is different. Note that the backward solutions for the individual elements in a and $\tilde{a}$ differ, whereas they add up to the same sum in the time-invariant case.[23] Consequently, the likelihood information leads to a rotationally non-invariant posterior distribution for the time-varying parameter. □

[23]If a is time-invariant, then $K_a = \frac{1}{\exp(2g_2)} \sum_{t=1}^{T} y_{1,t}^2$ and $\tilde{K}_a = \frac{1}{\exp(2g_1)} \sum_{t=1}^{T} y_{2,t}^2$ with $\exp(g_1) = \exp(g_2)$ and $\sum_{t=1}^{T} y_{1,t}^2 = \sum_{t=1}^{T} y_{2,t}^2$ implies that $K_a = \tilde{K}_a$. $\bar{a} = \tilde{\bar{a}}$ as $\bar{a} = \frac{1}{\exp(2g_2)} \sum_{t=1}^{T} y_{1,t} y_{2,t}$ $\tilde{\bar{a}} = \frac{1}{\exp(2g_1)} \sum_{t=1}^{T} y_{1,t} y_{2,t}$. Hence, the posterior distribution of a and $\tilde{a}$ is the same under alternative orderings of the variables.

B Additional Monte Carlo simulation

B.1 Robustness: estimated hyperparameters

The influence of the hyperparameters on the innovation variance is different across models, orderings and alternative DGPs. Thus, it is not clear whether this prior specification for the hyperparameters resembles a fair model comparison. For this reason, the CMSV and DC-CMSV models and their hyperparamaters are re-estimated using the algorithm of Amir-Ahmadi, Matthes, and Wang (2018).

Tables 4 – 5 show the results for all three DGPs. Overall, the choice of the hyperparameters has a limited effect on the results as the posterior median of the estimated hyperparameters is close to, generally, slightly smaller than the chosen hyperparameters. Broadly speaking, the performance metrics improve slightly in all dimensions and for both models.

Table 4: Estimated correlation

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.041	**0.030**	0.029	**0.023**	0.038	**0.029**	**0.371**	-0.219
sine	0.098	**0.086**	0.064	**0.018**	0.090	**0.023**	0.534	**0.942**
fastsine	**0.245**	0.256	0.087	**0.008**	0.115	**0.010**	0.390	**0.931**
step	0.081	**0.062**	0.045	**0.016**	0.068	**0.021**	0.493	**0.795**
ramp	0.119	**0.110**	0.068	**0.020**	0.098	**0.027**	0.546	**0.942**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.061	**0.039**	0.048	**0.025**	0.076	**0.031**	**0.308**	-0.216
sine	0.119	**0.089**	0.106	**0.015**	0.146	**0.020**	0.275	**0.945**
fastsine	**0.244**	0.257	0.142	**0.006**	0.185	**0.008**	0.176	**0.936**
step	0.105	**0.068**	0.080	**0.015**	0.121	**0.019**	0.287	**0.801**
ramp	0.139	**0.113**	0.111	**0.018**	0.154	**0.023**	0.296	**0.944**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	**0.024**	0.024	**0.021**	0.031	**0.026**	**0.296**	-0.181
sine	0.087	**0.083**	0.043	**0.020**	0.061	**0.026**	0.748	**0.942**
fastsine	**0.251**	0.256	0.051	**0.010**	0.067	**0.012**	0.587	**0.926**
step	0.066	**0.057**	0.029	**0.015**	0.042	**0.019**	0.671	**0.826**
ramp	0.111	**0.108**	0.046	**0.023**	0.068	**0.030**	0.742	**0.942**

(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to stochastic volatility. A bold figure highlights the best model in each panel and row.

Table 5: Estimated covariance

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.314	**0.288**	0.155	**0.030**	0.224	**0.042**	0.445	**0.976**
sine	0.215	**0.200**	0.102	**0.021**	0.154	**0.030**	0.666	**0.983**
fastsine	**0.330**	0.331	0.112	**0.010**	0.170	**0.014**	0.607	**0.993**
step	0.251	**0.225**	0.107	**0.020**	0.165	**0.028**	0.591	**0.985**
ramp	0.235	**0.222**	0.105	**0.024**	0.161	**0.035**	0.636	**0.979**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.438	**0.393**	0.202	**0.038**	0.344	**0.059**	0.486	**0.987**
sine	0.292	**0.255**	0.177	**0.021**	0.308	**0.033**	0.566	**0.993**
fastsine	0.401	**0.395**	0.207	**0.009**	0.353	**0.014**	0.472	**0.998**
step	0.353	**0.303**	0.169	**0.021**	0.303	**0.034**	0.557	**0.994**
ramp	0.316	**0.283**	0.182	**0.024**	0.315	**0.039**	0.544	**0.991**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.238	**0.219**	0.134	**0.024**	0.177	**0.032**	0.349	**0.956**
sine	0.172	**0.165**	0.077	**0.022**	0.109	**0.029**	0.689	**0.967**
fastsine	**0.295**	0.296	0.062	**0.011**	0.085	**0.014**	0.738	**0.983**
step	0.192	**0.178**	0.082	**0.017**	0.119	**0.023**	0.574	**0.970**
ramp	0.192	**0.185**	0.082	**0.024**	0.119	**0.034**	0.649	**0.961**

(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to stochastic volatility. A bold figure highlights the best model in each panel and row.

B.2 Robustness: stationary state dynamics

Another concern might be misspecification of the volatilities. The true DGP assumes stationary volatility dynamics. This form of misspecification may affect the ability of the DC-CMSV model to control for heteroskedasticity of the data, which is important to obtain almost rotationally invariant estimates. Therefore, both models are re-estimated assuming a stationary law of motion for volatility and parameter of contemporaneous relation.

Tables 6 – 7 show the results for all three DGPs. The statistics indicate that the main result is not affected, however, some features stand out. The estimated covariance and value-at-risk are slightly more accurate than those in the main results. However, distance and similarity metrics indicate more distinct estimates. This result is the consequence of a more distinct posterior distribution of the parameter of contemporaneous relation under an autoregressive process. Nevertheless, all estimates are broadly similar under the DC-CMSV model.

Table 6: Estimated correlation

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.041	**0.028**	**0.028**	0.029	0.038	**0.036**	**0.437**	-0.370
sine	0.101	**0.092**	0.065	**0.021**	0.092	**0.027**	0.508	**0.901**
fastsine	**0.224**	0.229	0.118	**0.046**	0.157	**0.058**	0.470	**0.694**
step	0.081	**0.063**	0.047	**0.022**	0.069	**0.028**	0.486	**0.700**
ramp	0.121	**0.114**	0.070	**0.027**	0.100	**0.035**	0.526	**0.875**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.063	**0.035**	0.048	**0.033**	0.078	**0.042**	**0.348**	-0.345
sine	0.122	**0.095**	0.111	**0.019**	0.151	**0.024**	0.249	**0.903**
fastsine	0.235	**0.232**	0.165	**0.041**	0.215	**0.052**	0.226	**0.702**
step	0.106	**0.069**	0.084	**0.022**	0.125	**0.028**	0.281	**0.694**
ramp	0.141	**0.118**	0.116	**0.025**	0.160	**0.032**	0.275	**0.877**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	**0.023**	0.025	**0.024**	0.032	**0.031**	**0.364**	-0.368
sine	0.090	**0.088**	0.043	**0.023**	0.060	**0.029**	0.730	**0.905**
fastsine	**0.215**	0.226	0.092	**0.052**	0.120	**0.065**	0.650	**0.688**
step	0.065	**0.058**	0.031	**0.020**	0.043	**0.026**	0.656	**0.740**
ramp	0.113	**0.111**	0.046	**0.029**	0.068	**0.038**	0.725	**0.878**

(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to stochastic volatility. A bold figure highlights the best model in each panel and row.

Table 7: Estimated covariance

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.314	**0.286**	0.152	**0.038**	0.217	**0.053**	0.519	**0.934**
sine	0.215	**0.201**	0.101	**0.024**	0.152	**0.034**	0.685	**0.977**
fastsine	0.311	**0.303**	0.158	**0.052**	0.243	**0.076**	0.517	**0.804**
step	0.249	**0.224**	0.107	**0.027**	0.162	**0.037**	0.639	**0.977**
ramp	0.235	**0.222**	0.105	**0.032**	0.160	**0.045**	0.650	**0.957**

(a) Benchmark DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.440	**0.394**	0.195	**0.051**	0.331	**0.082**	0.547	**0.951**
sine	0.293	**0.257**	0.181	**0.025**	0.312	**0.040**	0.573	**0.990**
fastsine	0.394	**0.367**	0.236	**0.053**	0.408	**0.088**	0.437	**0.874**
step	0.353	**0.303**	0.170	**0.031**	0.301	**0.049**	0.590	**0.988**
ramp	0.317	**0.284**	0.185	**0.033**	0.319	**0.053**	0.551	**0.980**

(b) High Volatility DGP

	MAE		MAD		RMSD		FD	
	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.236	**0.216**	0.139	**0.029**	0.183	**0.038**	0.414	**0.900**
sine	0.171	**0.164**	0.076	**0.024**	0.105	**0.032**	0.724	**0.956**
fastsine	**0.264**	0.266	0.125	**0.053**	0.174	**0.072**	0.555	**0.735**
step	0.190	**0.176**	0.088	**0.022**	0.124	**0.029**	0.626	**0.961**
ramp	0.191	**0.185**	0.083	**0.031**	0.117	**0.042**	0.671	**0.922**

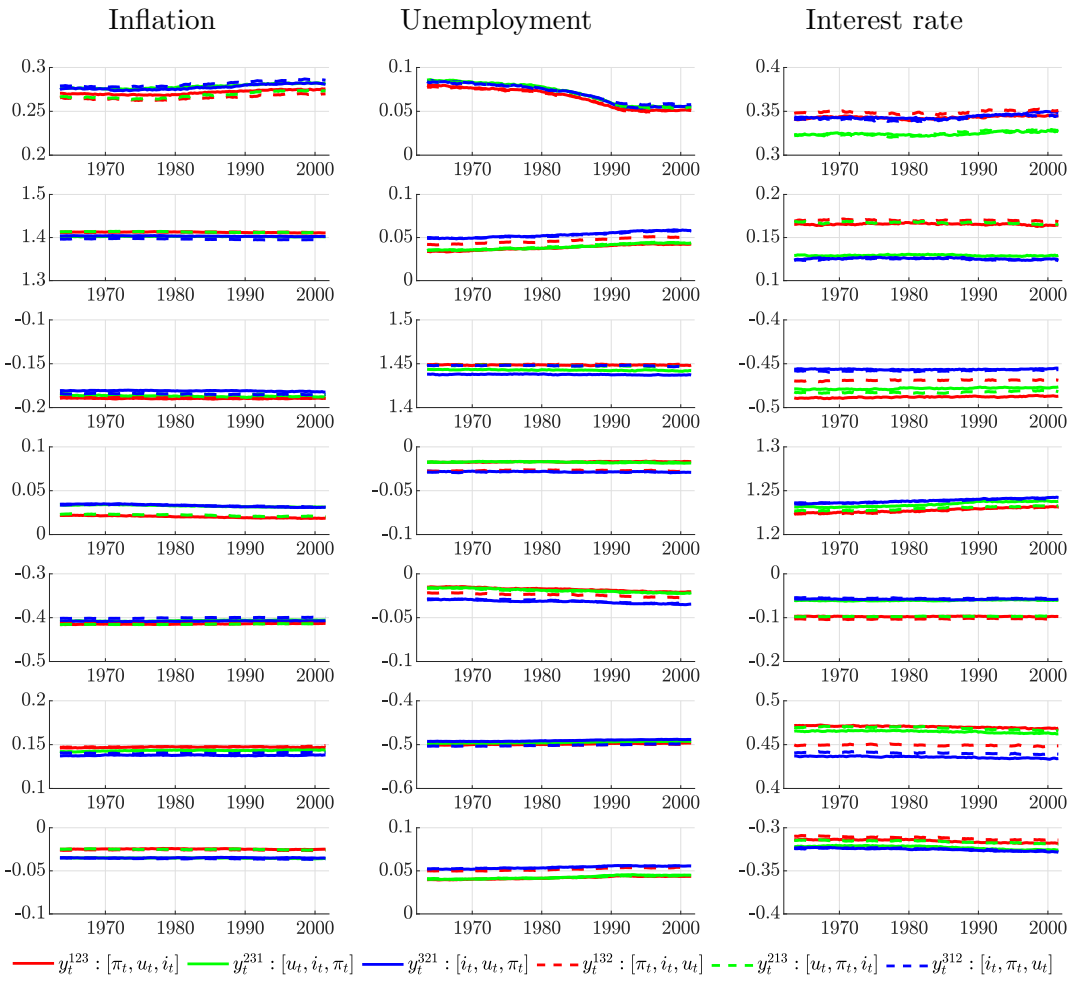
(c) Low Volatility DGP

The table shows the performance metrics for different scales for innovations to stochastic volatility. A bold figure highlights the best model in each panel and row.

C Revisiting Primiceri's (2005) application

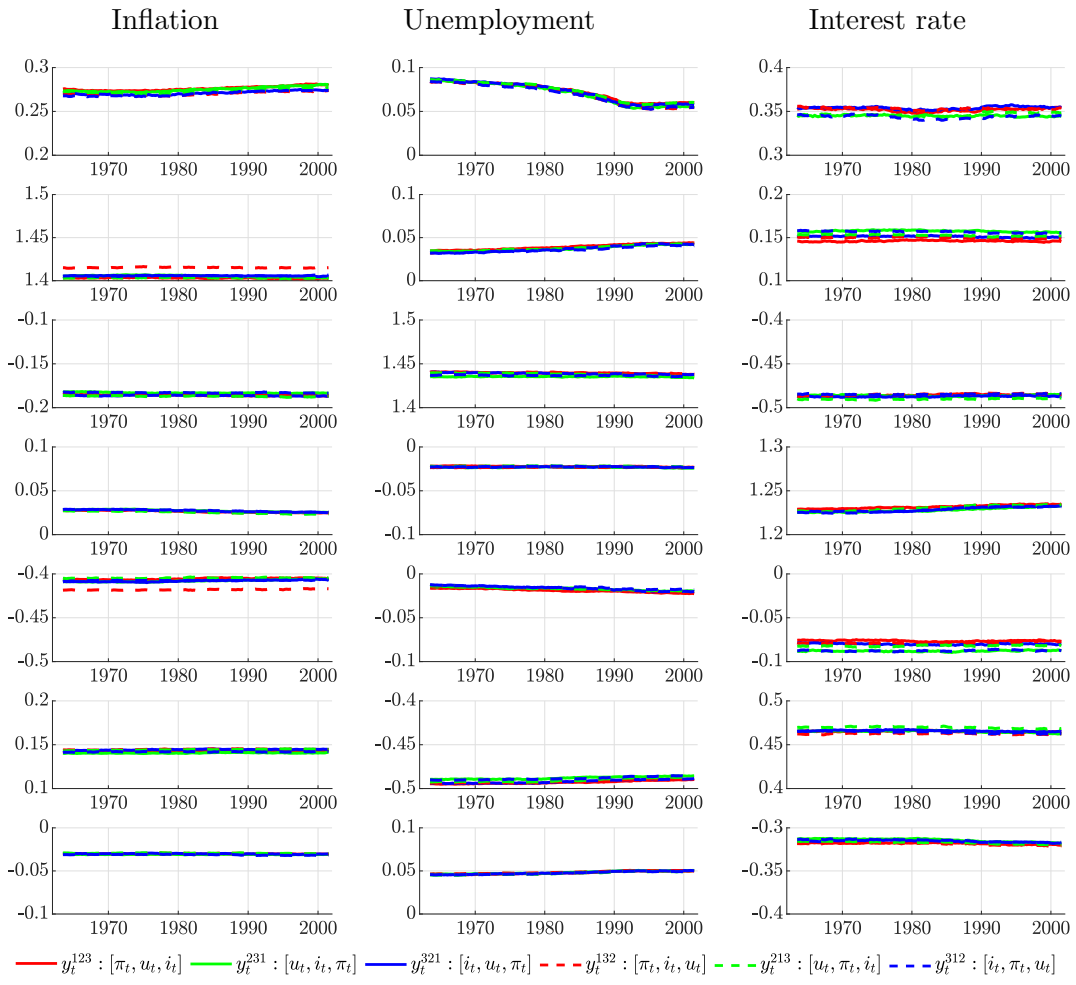
C.1 Sensitivity of reduced-form parameters

Figure 8: Estimated $\tilde{B}_t$'s (CMSV-TVP-VAR)



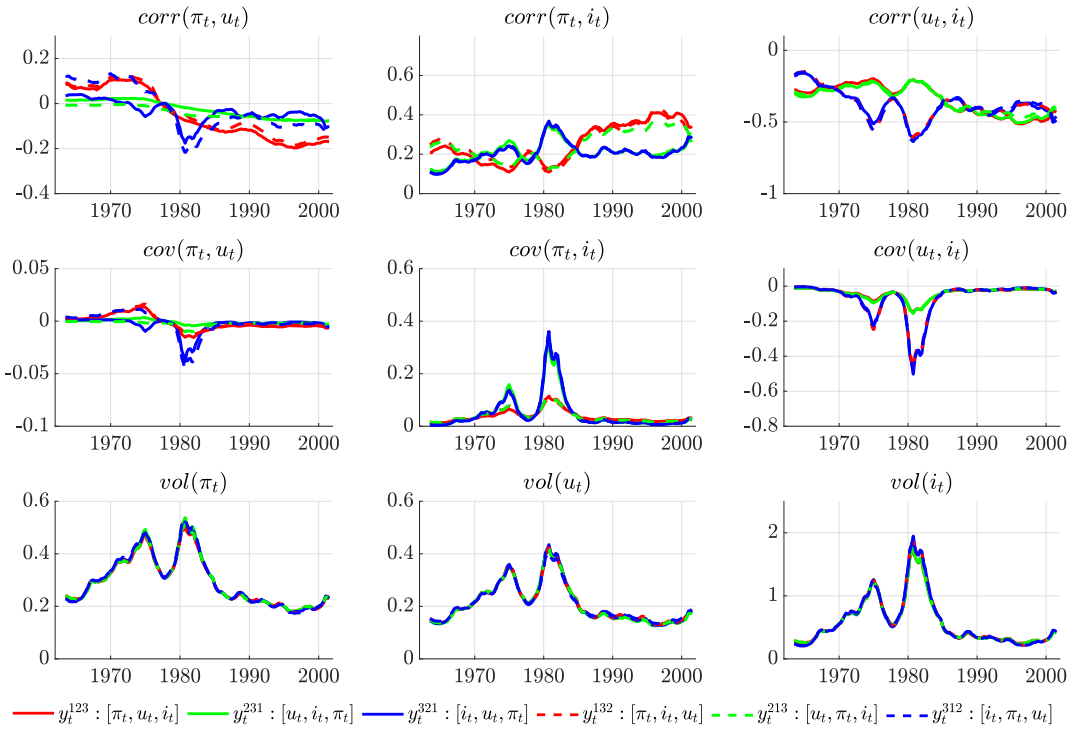
The figure depicts the posterior median of the time-varying VAR parameters for each equation in the respective column for all possible variable orderings obtained from the CMSV-TVP-VAR.

Figure 9: Estimated $\tilde{B}_t$'s (DC-CMSV-TVP-VAR)



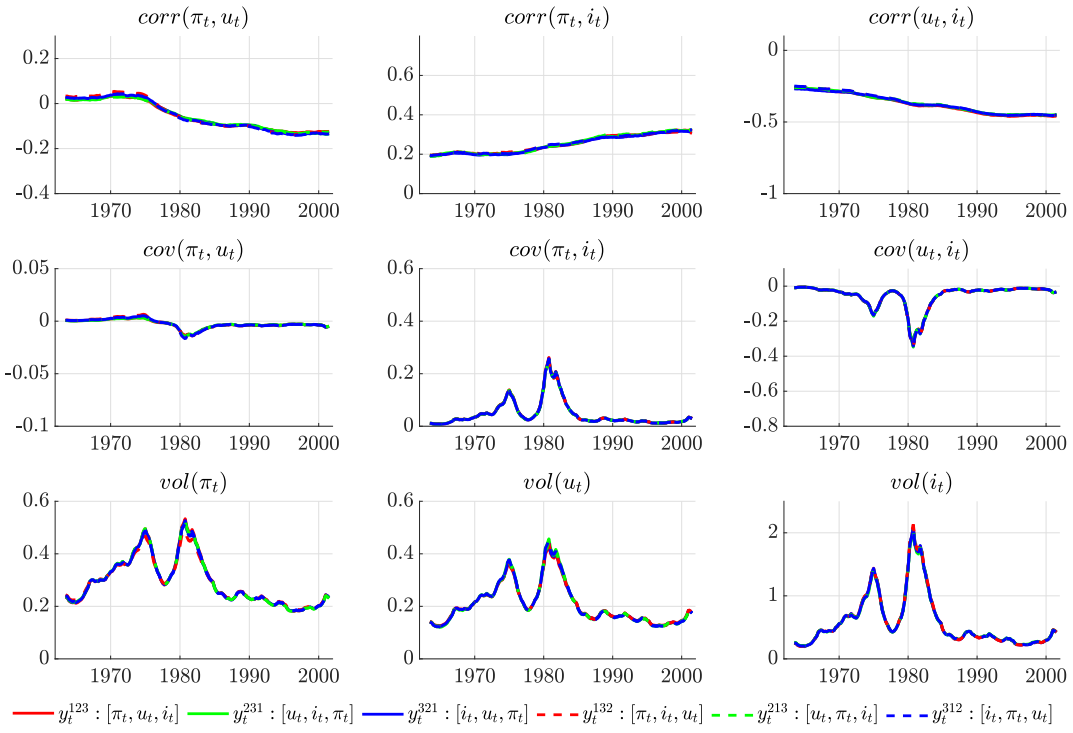
The figure depicts the posterior median of the time-varying VAR parameters for each equation in the respective column for all possible variable orderings obtained from the DC-CMSV-TVP-VAR.

Figure 10: Estimated $\tilde{\Sigma}_t$'s (CMSV-TVP-VAR)



The figure shows the posterior median of the correlation (corr), the covariance (cov) and the volatility (vol) of the reduced-form residual of inflation (π_t), unemployment (u_t) and the interest rate (i_t) for all possible orderings obtained from the CMSV-TVP-VAR.

Figure 11: Estimated $\tilde{\Sigma}_t$'s (DC-CMSV-TVP-VAR)

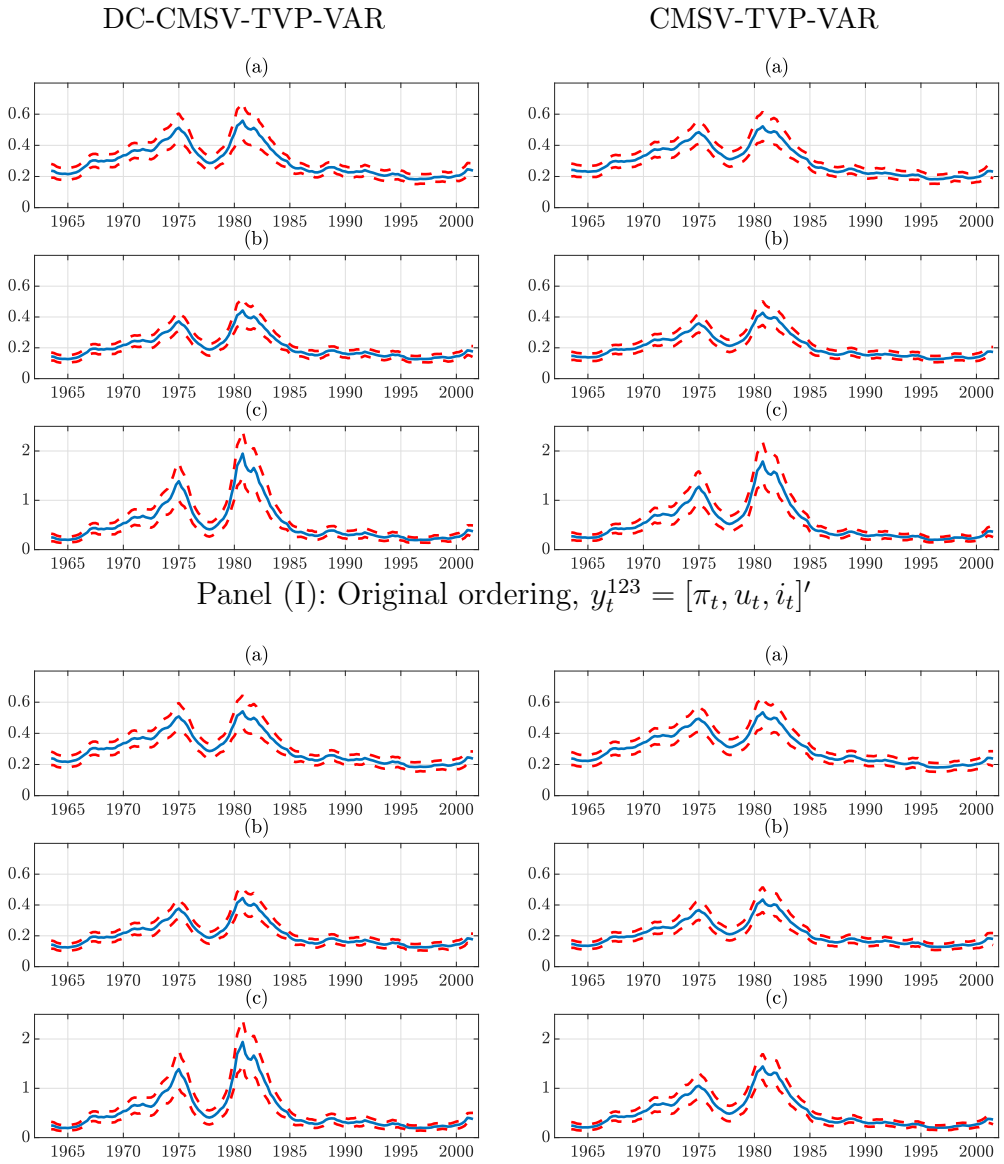


The figure shows the posterior median of the correlation (corr), the covariance (cov) and the volatility (vol) of the reduced-form residual of inflation (π_t), unemployment (u_t) and the interest rate (i_t) for all possible orderings obtained from the DC-CMSV-TVP-VAR.

C.2 Sensitivity of structural analysis

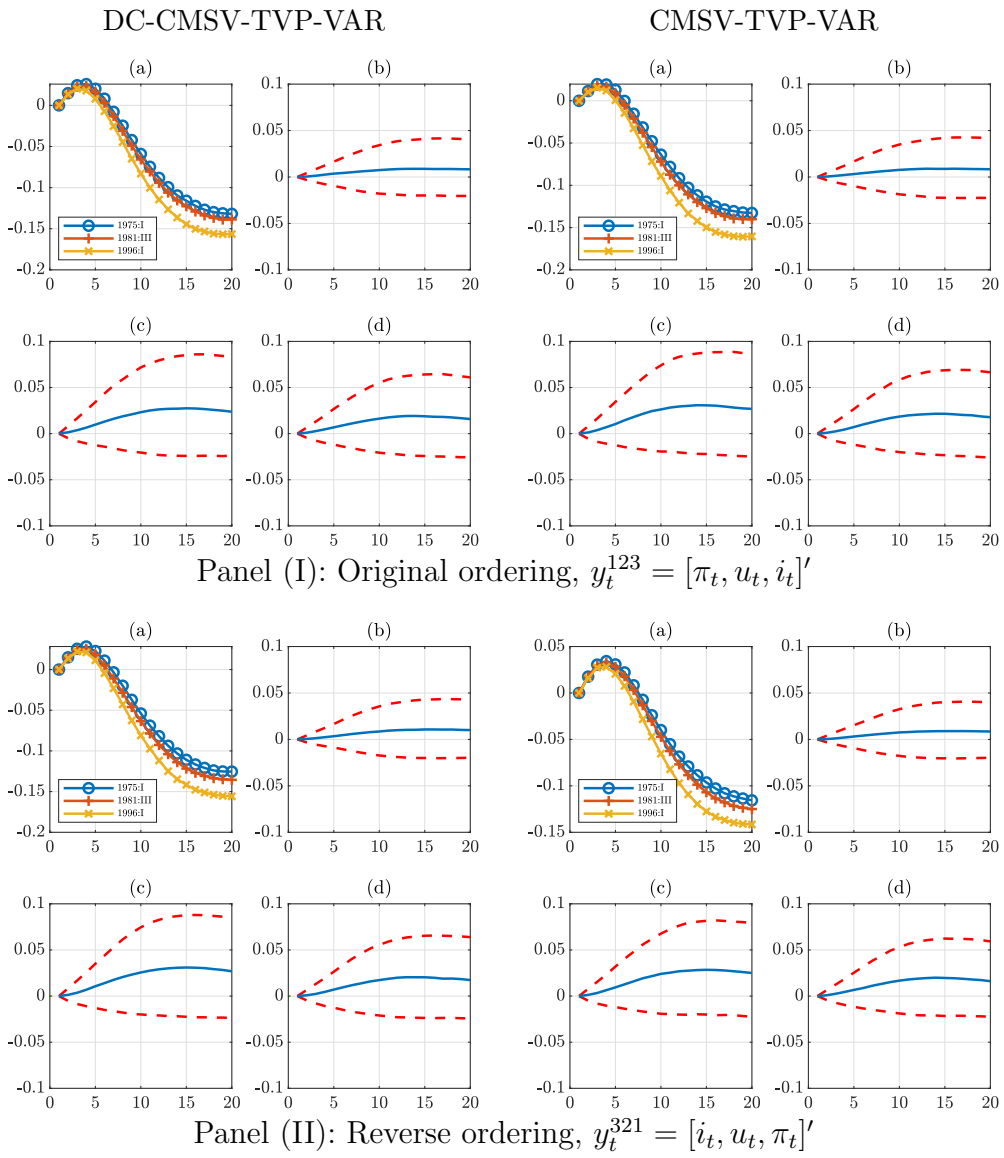
Sensitivity structural analysis

Figure 12: Replication of Figure 1



The figure depicts the posterior mean, 16th and 84th percentiles of the standard deviation of (a) the residuals of the inflation equation, (b) the residuals of the unemployment equation and (c) the residuals of the interest rate equation or monetary policy shocks.

Figure 13: Replication Figure 2

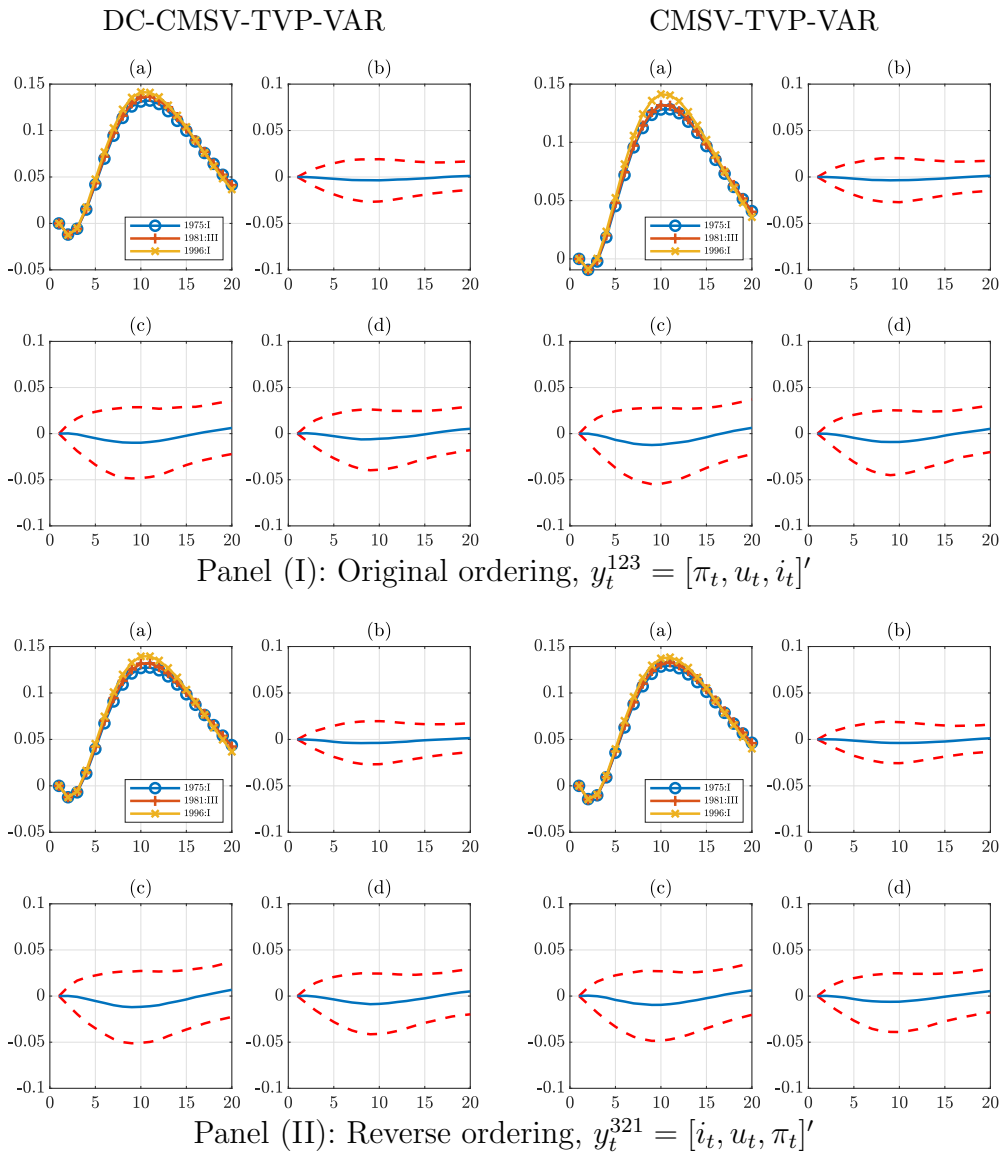


Panel (I): Original ordering, $y_t^{123} = [\pi_t, u_t, i_t]'$

Panel (II): Reverse ordering, $y_t^{321} = [i_t, u_t, \pi_t]'$

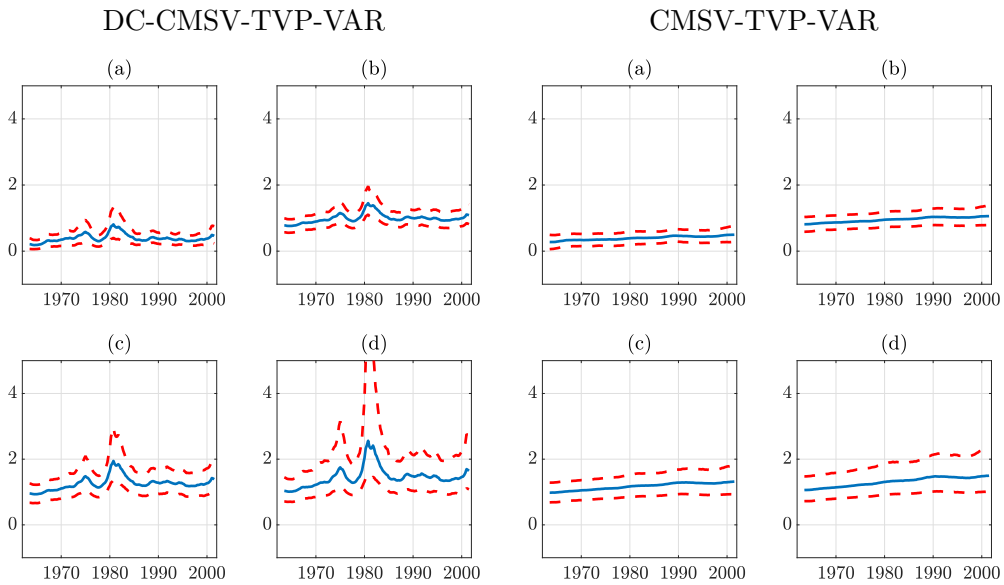
The figure depicts (a) impulse response of inflation to monetary policy shocks in 1975:I, 1981:III, and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

Figure 14: Replication Figure 3

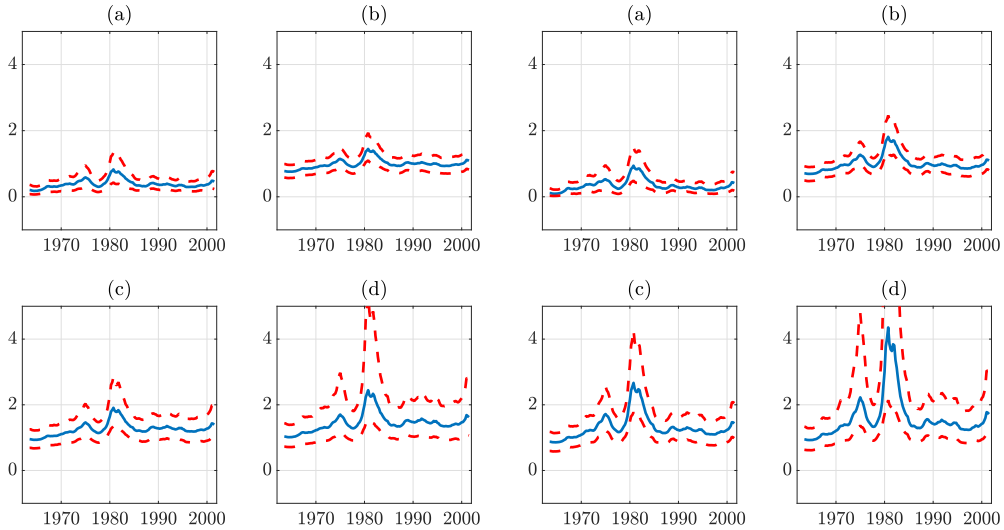


The figure depicts (a) impulse response of unemployment to monetary policy shocks in 1975:I, 1981:III, and 1996:I, (b) difference between the responses in 1975:I and 1981:III with 16th and 84th percentiles, (c) difference between the responses in 1975:I and 1996:I with 16th and 84th percentiles, (d) difference between the responses in 1981:III and 1996:I with 16th and 84th percentiles.

Figure 15: Replication Figure 4



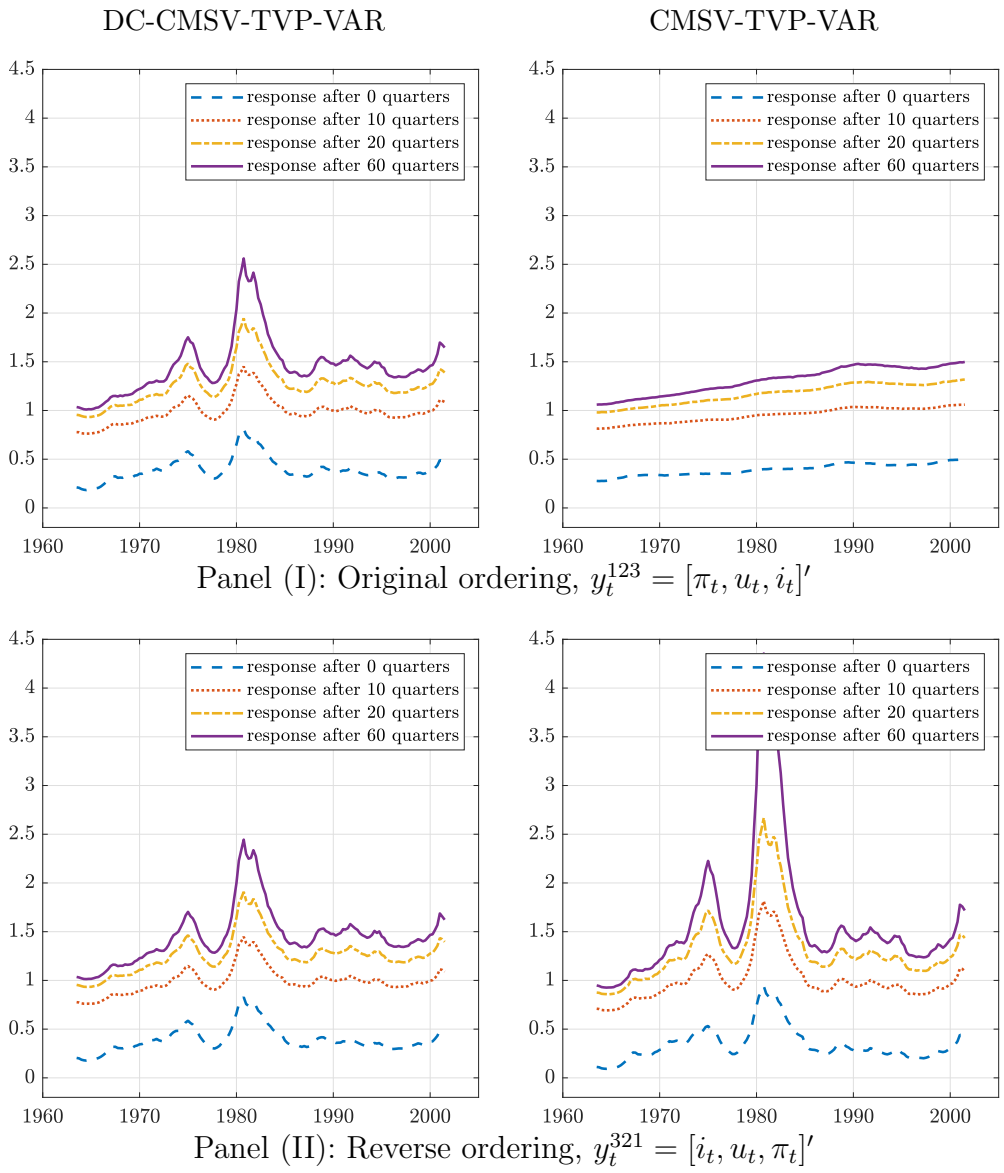
Panel (I): Original ordering, $y_t^{123} = [\pi_t, u_t, i_t]'$



Panel (II): Reverse ordering, $y_t^{321} = [i_t, u_t, \pi_t]'$

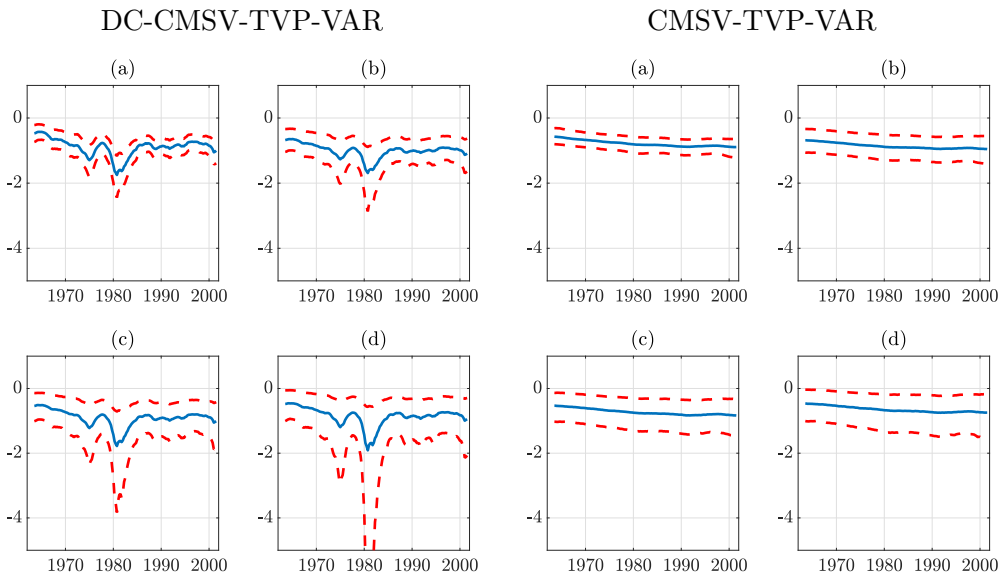
The figure depicts the interest rate response to a 1% permanent increase in inflation with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

Figure 16: Replication Figure 5

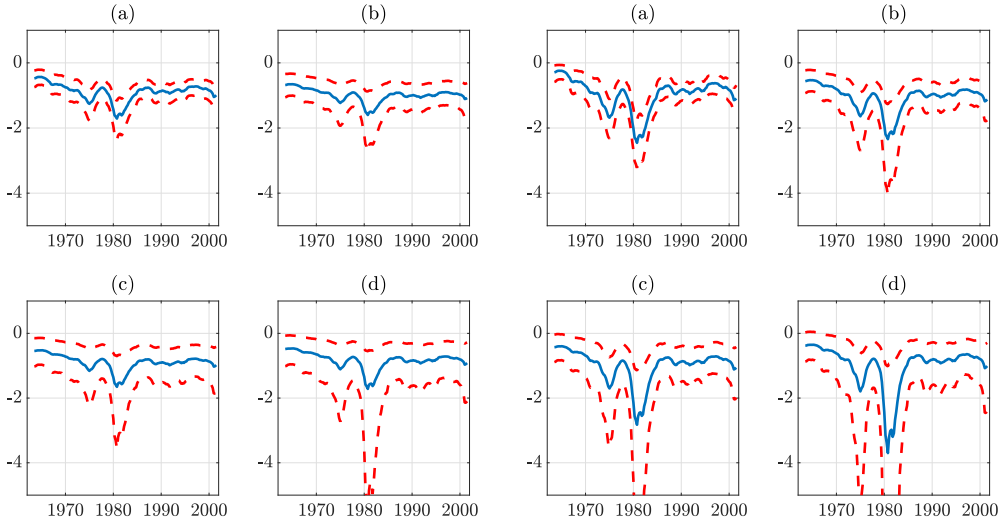


The figure depicts the interest rate response to a 1% permanent increase in inflation.

Figure 17: Replication Figure 6



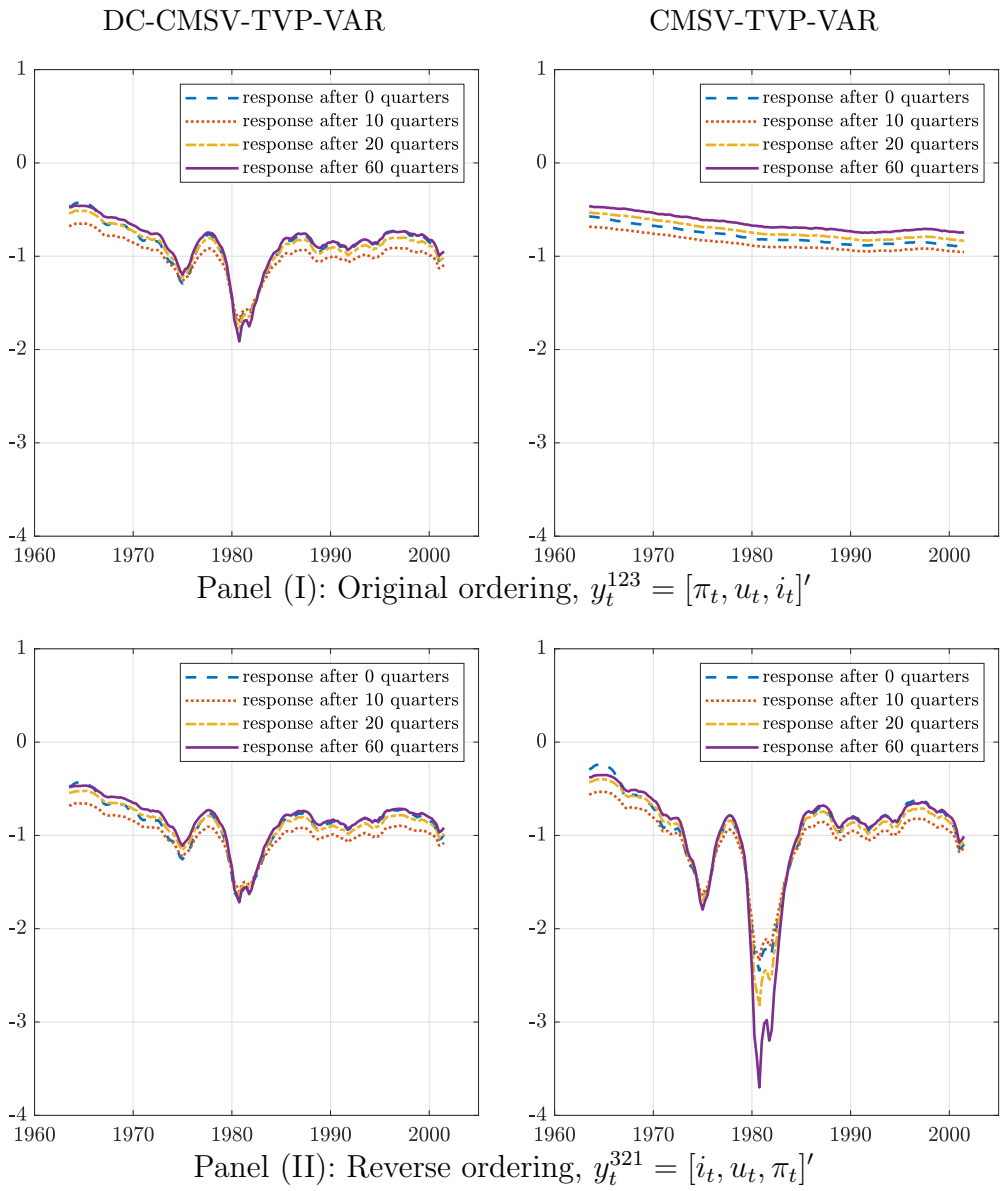
Panel (I): Original ordering, $y_t^{123} = [\pi_t, u_t, i_t]'$



Panel (II): Reverse ordering, $y_t^{321} = [i_t, u_t, \pi_t]'$

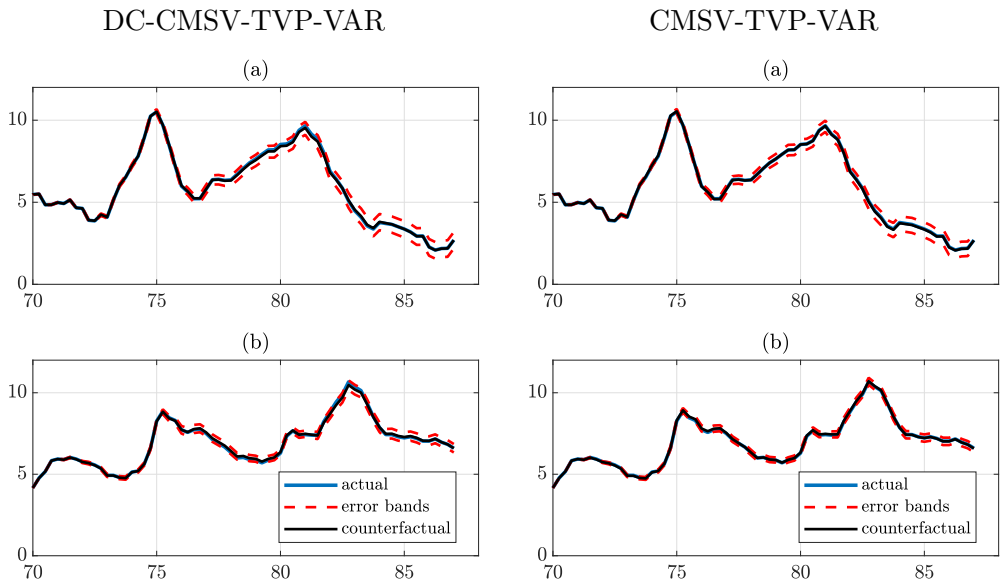
The figure depicts the interest rate response to a 1% permanent increase in the unemployment rate with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

Figure 18: Replication Figure 7

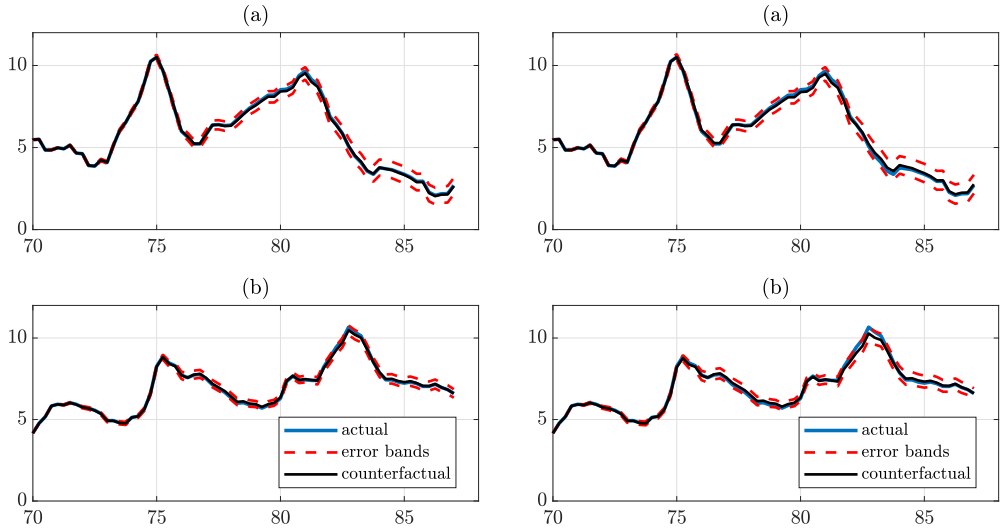


The figure depicts the interest rate response to a 1% permanent increase in the unemployment rate.

Figure 19: Replication Figure 8



Panel (I): Original ordering, $y_t^{123} = [\pi_t, u_t, i_t]'$

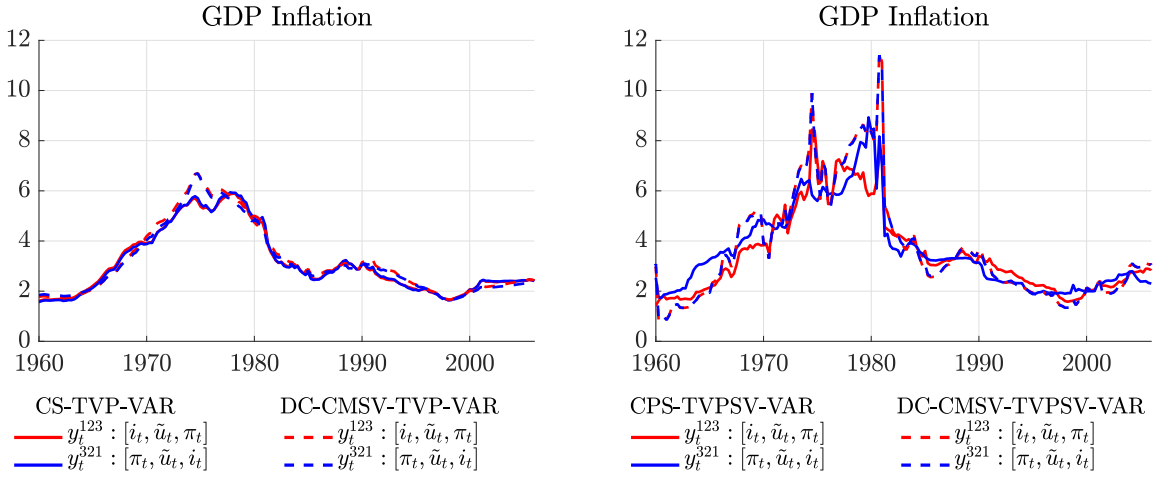


Panel (II): Reverse ordering, $y_t^{321} = [i_t, u_t, \pi_t]'$

The figure depicts a counterfactual historical simulation drawing the parameters of the monetary policy rule from their 1991-1992 posterior. (a) Inflation, (b) unemployment.

D Revisiting Cogley, Primiceri, and Sargent's (2010) application

Figure 20: Trend inflation

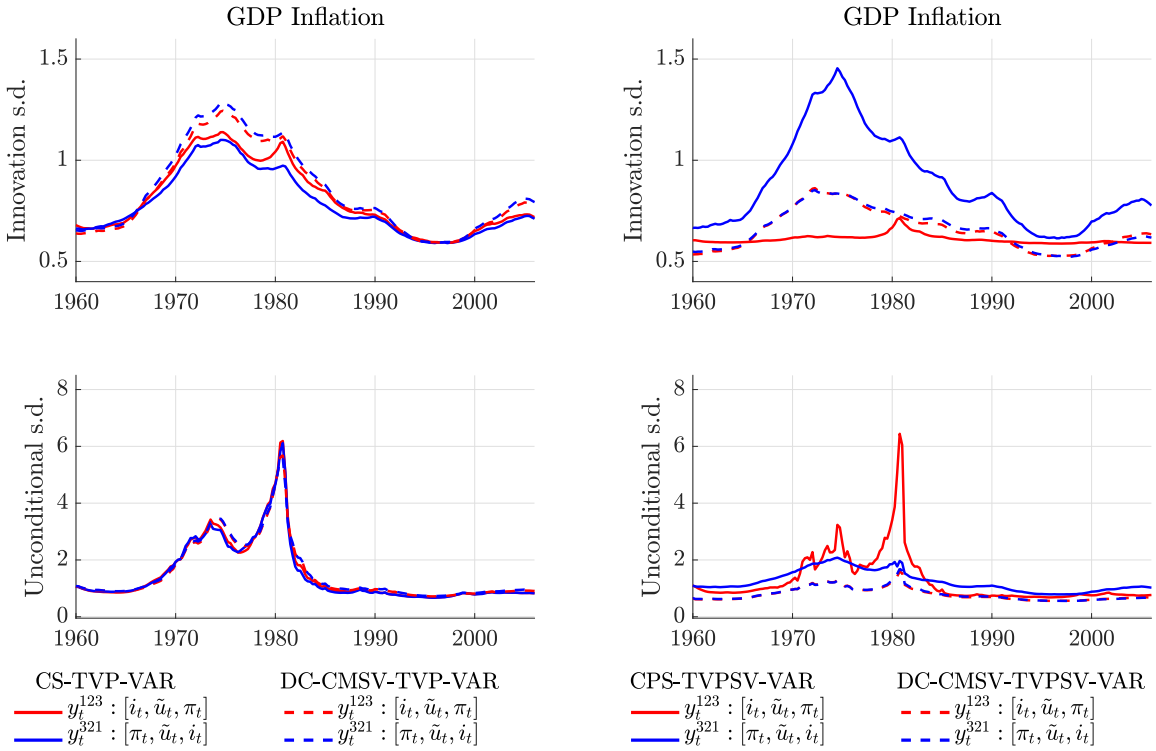


(a) Homoskedastic innovations for $\tilde{B}_t$'s

(b) Heteroskedastic innovations for $\tilde{B}_t$'s

The figure depicts the posterior median for trend inflation of the GDP deflator from TVP-VARs with homoskedastic and heteroskedastic parameter innovations for $\tilde{B}_t$'s in the left panel (a) and the right panel (b), respectively.

Figure 21: Inflation volatility

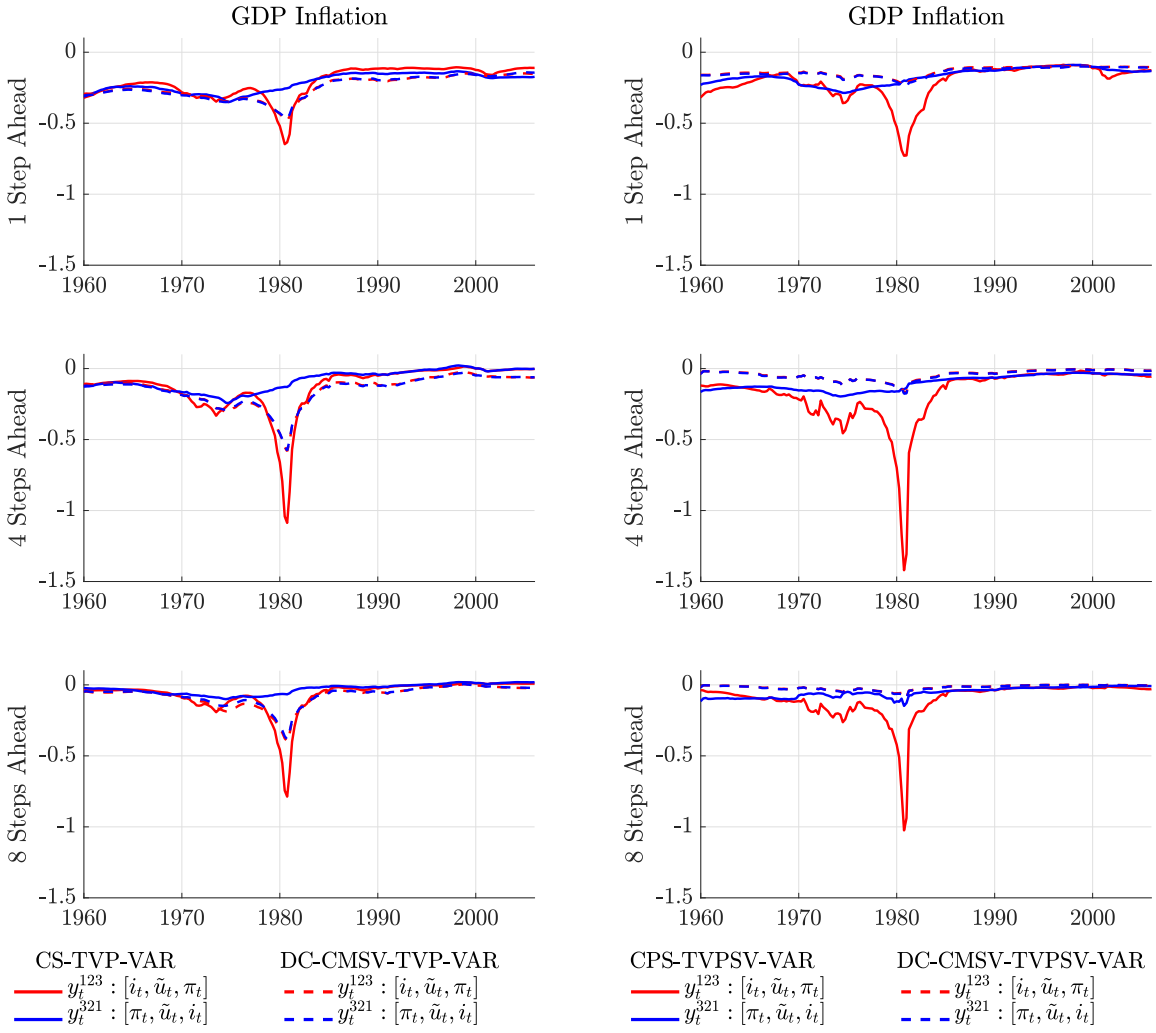


(a) Homoskedastic innovations for $\tilde{B}_t$'s

(b) Heteroskedastic innovations for $\tilde{B}_t$'s

The figure depicts the posterior median for inflation volatility of the GDP deflator from TVP-VARs with homoskedastic and heteroskedastic parameter innovations for $\tilde{B}_t$'s in the left panel (a) and the right panel (b), respectively.

Figure 22: The effect of unemployment news on expected inflation

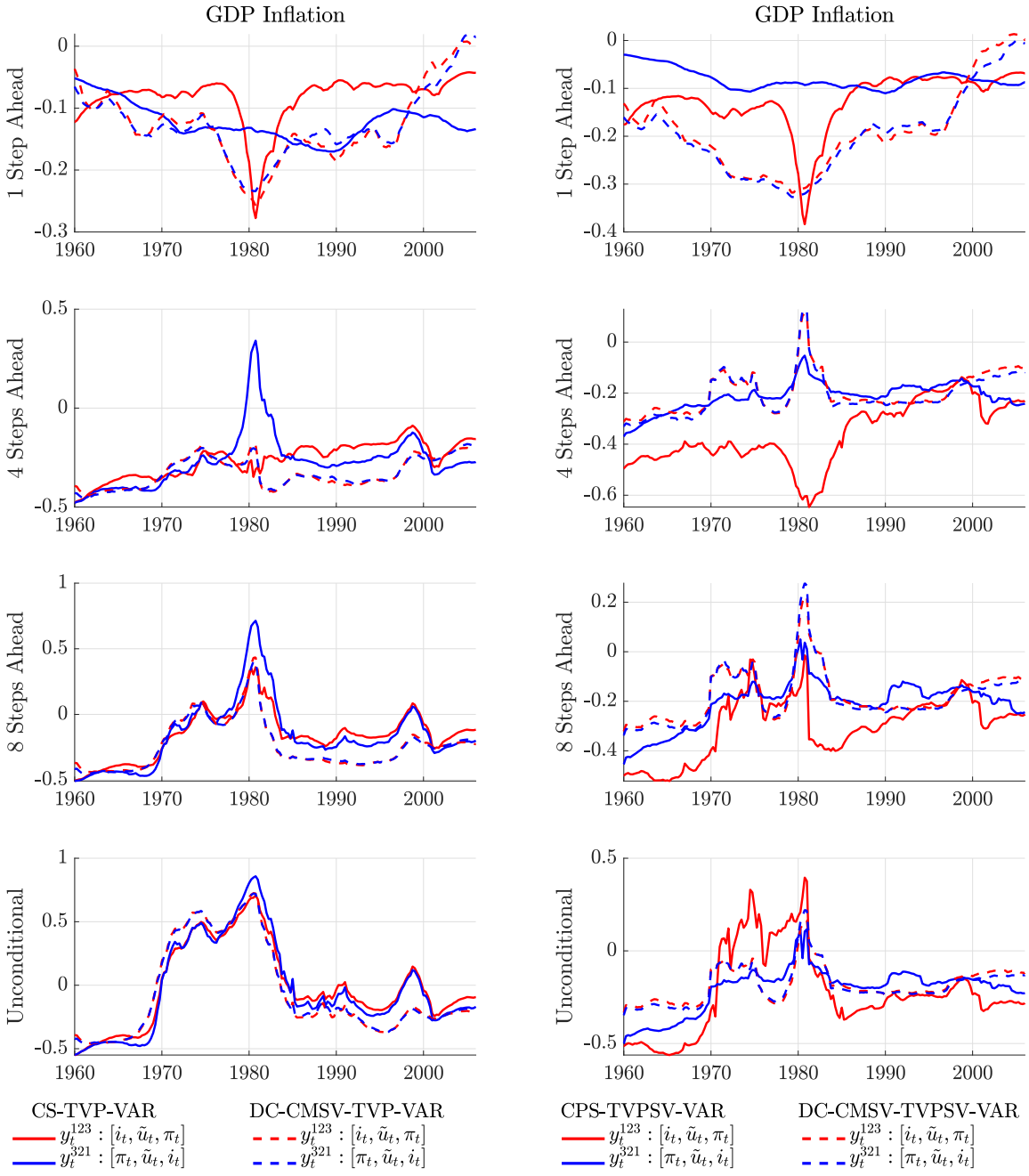


(a) Homoskedastic innovations for $\tilde{B}_t$'s

(b) Heteroskedastic innovations for $\tilde{B}_t$'s

The figure depicts the posterior median for expected GDP price deflator inflation based on a one standard deviation increase of the logit of the unemployment rate from TVP-VARs with homoskedastic and heteroskedastic parameter innovations for $\tilde{B}_t$'s in the left panel (a) and the right panel (b), respectively.

Figure 23: Conditional and unconditional Phillips correlations



(a) Homoskedastic innovations for $\tilde{B}_t$'s

(b) Heteroskedastic innovations for $\tilde{B}_t$'s

The figure depicts the posterior median for conditional and unconditional Phillips correlations from TVP-VARs with homoskedastic and heteroskedastic parameter innovations for $\tilde{B}_t$'s in the left panel (a) and the right panel (b), respectively.