

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

Benny Hartwig\*

## Abstract

**Abstract:** The ordering of variables is often considered to be negligible for the estimates of the reduced-form covariance matrix of the Cholesky multivariate stochastic volatility model. This paper shows that this procedure imposes systematically different dynamic restriction across alternative orderings on the covariance matrix when the ratio of reduced-form volatilities is time-varying. Consequently, conclusion drawn from this model also hinge on the deemed unimportant selected ordering of variables. This paper illustrates these effects for a small-scale macroeconometric model and proposes the dynamic correlation Cholesky multivariate stochastic volatility model 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](mailto: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

An important but often neglected property of the Cholesky multivariate stochastic volatility (CMSV) model of Primiceri (2005) and Tsay (2005) is that the estimated reduced-form covariance matrix may be sensitive to the ordering of variables.<sup>1</sup> Exceptions are Primiceri (2005); Koop, León-González, and Strachan (2009); Nakajima and Watanabe (2011); Lopes, McCulloch, and Tsay (2012). However, Figure 1 illustrates that the full spectrum of this property is not sufficiently explored. It shows that the estimated covariances are sensitive and substantially different across alternative orderings in Primiceri’s (2005) application.<sup>2,3</sup> Specifically, estimates diverge during the stagflation period when individual residuals exhibit some mild non-common volatility pattern. In contrast, estimates of volatility are hardly affected by this property.

This paper argues that the channels through which alternative ordering affect the estimated covariance matrix are not well understood. Moreover, it stresses that the differences between alternative estimates depicted in Figure 1 are not arbitrary but obey a systematic pattern that is likely to be present in many empirical applications. In addition, it emphasizes that this type of sensitivity for the estimates must not be ignored for applications which use the estimated reduced-form covariance matrix as an input. Such applications include structural VARs with CMSV that identify structural shocks with short-run (but no triangular restriction), long-run, or sign restrictions as well as portfolio optimization and risk management problems. This property is especially problematic for inference as alternative estimates may give rise to ambiguous conclusions which impede robust inference.

This paper makes several novel contributions. It identifies a time-varying ratio of reduced-form volatilities as the main channel through which alternative ordering impose different dynamic restrictions on the covariance matrix. Under the true ordering, the parameter of contemporaneous relation evolves linearly in the CMSV model. In any alternative ordering, however, the implied dynamics of this parameter

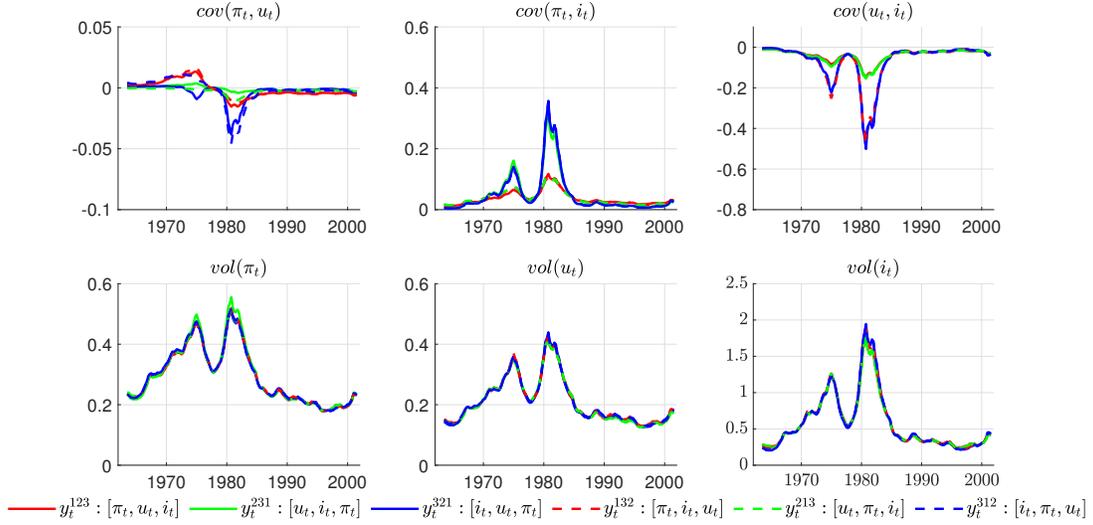
---

<sup>1</sup>See Primiceri (2005); Cogley and Sargent (2005); Asai, McAleer, and Yu (2006)

<sup>2</sup>Estimates are based on Algorithm 2 of the corrigendum by Del Negro and Primiceri (2015).

<sup>3</sup>Primiceri (2005) reported that his results are not sensitive to alternative orderings, however, this conclusion might be a burden from the incorrect Gibbs sampler used at the time.

Figure 1: Estimated contemporaneous reduced-form covariance matrix



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

are nonlinear. It is driven by the correlation process and the ratio of volatilities, which is log-normally distributed. Consequently, the nonlinear dynamics of the implied process cannot be well captured by a linear process of an analogously setup CMSV model. Moreover, when the data is generated by a separate evolving marginal volatility and correlation process then the estimated covariances are systematically different across alternative ordering. Specifically, the ratio of volatilities driving the parameter of contemporaneous relation is inverted in a reordering. This kind of transformation, therefore, induces a different sort of dynamics in the parameter.

Besides just explicating the problem, this paper proposes the dynamic correlation Cholesky multivariate stochastic volatility (DC-Cholesky MSV, or DC-CMSV) model in spirit of Engle (2002) as a robust alternative. The evolution of the covariance matrix is specified by separate volatility and correlation dynamics. The correlation dynamics are modelled through a pseudo CMSV model on the standardized data,

which features a constant ratio of reduced-form volatilities. Moreover, 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. Thus, estimates of the DC-CMSV model are almost insensitive to alternative orderings.

A notable feature of the DC-CMSV model is that the implied parameter of contemporaneous relation is not restricted to be linear but may also capture nonlinear structural changing conditions. In addition, the model can easily be implemented into existing routines. Moreover, the estimation of the model remains simple as traditional Kalman filter methods or the fast band-precision matrix routines of Chan and Jeliazkov (2009) can be used for inference purposes.

Next, simulation evidence shows that the estimates of the CMSV model are more distinct across alternative ordering, the stronger the idiosyncratic volatility patterns in the data. By construction, the estimates of the DC-CMSV model are hardly affected by this type of patterns in the data.

Last, this paper demonstrates that restrictions imposed by a particular variable ordering on the estimated reduced-form covariance may be so decisive that it may drive conclusions and yields results that are inconsistent across alternatives. In particular, the robustness of Primiceri's (2005) results are reconsidered under the assumption that the ordering of variables can be ignored, which is common practice in the literature. Alternative estimates of this exercise provide strong evidence for an alternative scenario: the U.S. systematic interest rate response to inflation and unemployment was substantially more aggressive during the stagflation period. In contrast, the original results indicate that the response was largely muted. Besides, estimates of DC-CMSV version of the model draw an unambiguous conclusion under all possible ordering. The results suggest that the reaction function was modestly more aggressive. This evidence is consistent with the finding of Sims and Zha (2006).

The findings of this paper relate to several strands in the literature. First, it formalizes the conditions when and why the lack of rotational invariance may matter for a data set at hand. It advances the argument of Christopher Sims in Cogley and Sargent (2005) and Asai, McAleer, and Yu (2006) by the fact that the dependence between volatilities and correlations particularly matters when the ratio of

volatilities varies over time. Moreover, the results of the paper show that rotational non-invariance of the prior plays 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 provide evidence that ordering of variables may play a role in structural inference. Bognanni (2018) shows for Baumeister and Peersman's (2013) application that the estimated effects of an oil supply shock on U.S. real activity are sensitive to the chosen ordering as well. The author interprets the choice of selecting a particular ordering as arbitrary and regards it as an additional source of model and parameter uncertainty. The findings of this paper, however, qualify this view. Particularly, alternative estimates must be interpreted as a byproduct of the model as it generally imposes alternative dynamic restrictions on the reduced-form covariance matrix. Thus, the finding of this paper question the validity of using estimates of CMSV model as an input for two-step identified SVARs.

Third, the estimates of 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 alternative approaches proposed by Primiceri (2005) or Nakajima and Watanabe (2011) suffer from immense or even intractable computational burdens. In fact, these methods need to explore all  $n!$  ( $n$  factorial) possible models.

Fourth, the DC-CMSV model is an attractive alternative to Wishart or inverted Wishart stochastic volatility, which yield rotationally invariant estimates of the covariance matrix. See for instance Uhlig (1997), Bognanni (2018), Chan, Doucet, León-González, and Strachan (2018) in the context of (TVP)-VARs with MSV or Philipov and Glickman (2006); Asai and McAleer (2009) in the context of MSV models. As pointed out by Primiceri (2005), this class of models, however, allow for less flexible dynamics of the covariance matrix. Particularly, they either allow for integrated dynamics or simple autoregressive dynamics of order one.

The rest of this paper proceeds as follows. Section 2 studies properties of the CMSV model under alternative orderings and under an alternative data generating process. Section 3 introduces the DC-CMSV model. Section 4 studies model properties using a Monte Carlo simulation. Section 5 reconsiders Primiceri's (2005) application in more detail. Section 6 concludes this paper.

## 2 On Cholesky Multivariate Stochastic Volatility

Consider the bivariate vector  $y_t \sim N(0, \Sigma_t)$  with time-varying covariance matrix. Table 1 presents two common multivariate stochastic volatility models for  $\Sigma_t$ , which mainly differ by their chosen factorisation of  $\Sigma_t$ .

Table 1: Multivariate stochastic volatility model of  $\Sigma_t$

Panel (a) CMSV model

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

where

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

Reparameterized vector of observations

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

State dynamics

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

Innovations

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

Prior distribution

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

Panel (b) DC-MSV model

$$\Sigma_t = D_t R_t D_t$$

where

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

Reparameterized vector of observations

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

State dynamics

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

Innovations

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

Prior distribution

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

Panel (a) presents the Cholesky multivariate stochastic volatility (CMSV) model which is based on Primiceri's (2005) model.<sup>4</sup> The name of the model derives from the fact that it specifies the dynamics of the triangular factorisation of  $\Sigma_t$  rather than specifying the dynamics of  $\Sigma_t$  directly. Panel (b) explicates the dynamic correlation multivariate stochastic volatility (DC-MSV) model of Yu and Meyer (2006).<sup>5</sup> This model specifies individual volatility and correlation dynamics to span the evolution of  $\Sigma_t$ . This factorisation is denoted as the volatility-correlation factorisation of  $\Sigma_t$ . The DC-MSV model is chosen as an alternative data generating process because the ordering of variables has no effect on the estimated reduced-form covariance matrix.<sup>6</sup>

The following analysis is restricted to the bivariate case because of tractability reasons.<sup>7</sup> Nevertheless, these properties are considered to be representative for the  $n$ -dimensional case. Specifically because the relationship between individual parameters of  $\Sigma_t$  and the parameters under these alternative factorization of  $\Sigma_t$  does not fundamentally change in higher dimensions.

## 2.1 Some properties of the Cholesky MSV model

Let  $y_t$  be generated by the CMSV model with covariance matrix  $\Sigma_t$ . Then, it follows from the triangular factorisation of  $\Sigma_t$  that the mapping from model parameters  $\{g_{1,t}, g_{2,t}, a_t\}$  to  $\{\sigma_{11,t}^2, \sigma_{22,t}^2, \sigma_{12,t}, \rho_t\}$  the elements and functions of the reduced-form covariance matrix  $\Sigma_t$  is given by

$$\begin{aligned}\sigma_{11,t}^2 &= \exp(2g_{1,t}), & \sigma_{22,t}^2 &= \exp(2g_{2,t}) + a_t^2 \exp(2g_{1,t}), \\ \sigma_{12,t} &= a_t \exp(2g_{1,t}) & \rho_t &= a_t \frac{\sigma_{11,t}}{\sigma_{22,t}}\end{aligned}$$

where  $\sigma_{ii,t}^2$  is the variance of the  $i$ th element of  $\Sigma_t$  for  $i = 1, 2$ ,  $\sigma_{12,t}$  is the covariance,  $\rho_t$  is the correlation and  $a_t$  is the contemporaneous relation.

---

<sup>4</sup>As in Koop, León-González, and Strachan (2009), a diagonal covariance matrix for the innovations of stochastic volatility is assumed to enhance comparability between both models.

<sup>5</sup>Alternative state dynamics are assumed to mimic the dynamics of the CMSV model.

<sup>6</sup>For a proof, see Appendix A.2

<sup>7</sup>The DC-MSV model cannot be easily generalized to higher dimensions for  $n \geq 3$ .

Then, the transition equation for the implied correlation process,  $\rho_t$ , is given by

$$\rho_t = \rho_{t-1} \frac{\exp(\epsilon_{1,t}^g)}{\exp(\epsilon_{2,t}^{g^{**}})} + \epsilon_t^a \frac{\sigma_{11,t}}{\sigma_{22,t}}$$

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

**Property** ( $\Sigma_t$  under CMSV model). Let  $y_t$  be generated by a CMSV model with  $\Sigma_t$ , then some important properties for the elements and functions of  $\Sigma_t$  are

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

For proof, see Appendix A.1

These properties give rise to two important consideration for the CMSV model as a data generating process. First, the model rules out common reduced-form volatility dynamics. This may be restrictive for some applications such as term-structure modelling. Second, the assumption of a smoothly evolving contemporaneous relation implies that there are rapid changes in the correlation pattern when volatility clusters idiosyncratically. Stated differently, the model interprets abrupt changes in relative volatilities as the dominant driver of changing correlation. Therefore, this assumption may put substantial dynamic restriction on the covariance matrix.

Next, let  $\tilde{y}_t = Py_t$  be the reordered vector of variables where  $P$  is a permutation matrix and let  $\tilde{\Sigma}_t = P\Sigma_tP'$  be the covariance matrix with permuted elements. Analogously, the triangular factorisation of  $\tilde{\Sigma}_t = \tilde{A}_t^{-1}\tilde{D}_t\tilde{A}_t'^{-1}$  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

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

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

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

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

**Property** (Reordering in CMSV model). Let  $\Sigma_t^*$  be the 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

- $\Sigma_t^*$  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 analogous constructed parameters  $\{g_{1,t}^*, g_{2,t}^*, a_t^*\}$  increases in the variability of the ratio of reduced-form variances  $\frac{\sigma_{11,t}^2}{\sigma_{22,t}^2}$

For proof, see Appendix A.1

To put it differently, the ordering of variables induces a dynamic structure in  $\Sigma_t$  that cannot be replicated by an analogously set up CMSV model for any alternative ordering of variables. The CMSV model imposes alternative dynamic restrictions on the reduced-form covariance matrix under alternative ordering. Hence, the choice of the ordering of variables is non trivial in the CMSV model.

While  $\Sigma_t^*$  and  $\tilde{\Sigma}_t$  cannot have the same dynamic structure, the dynamics may be quite similar or 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 more close to be roughly constant. However, when there are idiosyncratic volatility patterns, then the distance grows larger.

Above statements allow for some remarks about the CMSV model in the literature. Primiceri (2005) pointed out that the ordering of variables matters for  $\Sigma_t$  because the prior distribution of  $\Sigma_t$  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 dynamic dependence of model parameters in conjunction with the factorisation of the covariance matrix leads to a non observational equivalent prior density for  $\Sigma_t$ .<sup>8</sup> This discussion clarifies when the ordering of variables is important for inference. Precisely, it matters when there are idiosyncratic volatility patterns.

These results also shed light on the discussion of Asai, McAleer, and Yu (2006) and a comment raised by Christopher Sims in Cogley and Sargent (2005). They conjecture that not separating volatility and correlation dynamics may impose some dynamic restrictions on the covariance matrix. Particularly, the CMSV model rules out common volatility pattern, which induces some nonlinear correlation patterns. Also, because volatility pattern are not common, alternative orderings impose different dynamic restriction on the covariance matrix.

## 2.2 The Cholesky MSV model and the DC-MSV model

Let  $y_t$  be generated by the DC-MSV model with covariance matrix  $\Sigma_t$ . Then, it follows from the volatility-correlation decomposition of  $\Sigma_t$  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 the reduced-form covariance matrix  $\Sigma_t$  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  $i$ th element of  $\Sigma_t$  for  $i = 1, 2$ ,  $\sigma_{12,t}$  is the covariance,  $\rho_t$  is the correlation,  $a_t$  is the contemporaneous relation implied under  $\Sigma_t$  and  $\tilde{a}_t$  is the contemporaneous relation implied under  $\tilde{\Sigma}_t = P\Sigma_t P'$ .

Then, the transition equations for the implied contemporaneous relations under

---

<sup>8</sup>Primiceri (2005) elaborates on the triangular factorisation in the contemporaneous case.

the two alternative ordering  $a_t$  and  $\tilde{a}_t$  are given by

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

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

Under the DC-MSV model, the contemporaneous relation is driven by the correlation and the ratio of volatilities which is defined by second divided by first ordered variable. Thus, the ordering of variables plays a particular for the implied dynamic evolution for this parameter.

Notice the volatility-correlation decomposition of  $\Sigma_t$  allows for the specification of very general volatility dynamics. For instance, the DC-MSV model may be setup to feature a purely idiosyncratic (as specified in Table 1) or exhibit some commonalities or a completely common volatility pattern.

**Property** ( $\Sigma_t$  under DC-MSV model). Let  $y_t$  be generated by a DC-MSV model with  $\Sigma_t$ , then some important properties for the elements and functions of  $\Sigma_t$  are

1. the correlation  $\rho_t$  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  $\rho_t$  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 but have different dynamics

For proof, see Appendix A.2

Comparing the dynamic properties of  $\Sigma_t$ , the CMSV model assumes a linear Gaussian process for the contemporaneous relation but implies a nonlinearly evolving correlation process; whereas the DC-MSV model implies nonlinear dynamics for the contemporaneous relation and approximate linear Gaussian dynamics for the correlation in some specified range. In other words, the correlation acts somewhat as a degree of freedom in the CMSV model, while the contemporaneous relation gets this role in the DC-MSV model.

**Property** (DC-MSV, CMSV and implied covariances). Let  $\Sigma_t$  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

In other words, when  $y_t$  is generated by the DC-MSV model, then the implied covariance by a CMSV model are systematically different across alternative orderings. Particularly, they represent an upper and lower bound of the true covariance parameter under the DC-MSV model.

**Property** (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 posterior mean and variance of  $a_t$  and  $\tilde{a}_t$  depend on the distance between the sequence of  $y_{1,t}^2$  and  $y_{2,t}^2$ .

For proof, see Appendix A.2

The posterior distribution of the parameter of contemporaneous relation is not same across alternative ordering in the CMSV model because the information of the data is interpreted differently across alternative orderings. For this reason, the model produces different estimates of the covariance matrix under alternative orderings. This observation raises the natural question by how much this channel drives estimates of the posterior distribution apart.

To quantify this distance, a Monte Carlo simulation is conducted with 250 replications. A dynamic correlation model with unitary variance is used to simulate 1,000 observations or four years of daily data

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

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:  $\rho_t = \text{mod}(t/200)$ .

These correlation processes 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.

The CMSV model is estimated with fixed hyperparameters and common prior distribution. The MCMC estimation produce 35000 samples of which 15000 are reserved for the burnin period.

To gauge the model fit, the average mean absolute error (MAE) of the estimated quantity under both orderings is reported. It is defined as

$$MAE(X^{ORD}, X^0) = \frac{1}{I} \sum_i \left( \frac{1}{T} \sum_t |X_t^{ORD(i)} - X_t^0| \right)$$

where  $X^{ORD(i)}$  denotes the parameter estimate of the model with order  $i = 1, 2$ . ORD(1) and ORD(2) denote the ordering  $(y_{1,t}, y_{2,t})$  and  $(y_{2,t}, y_{1,t})$ , 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, which is defined as

$$MAD(X^{ORD(1)}, X^{ORD(2)}) = \frac{1}{T} \sum_t |X_t^{ORD(1)} - X_t^{ORD(2)}|$$

Table 2 presents the results from this Monte Carlo simulation. Turning to the precision of the estimates, the MAE figures indicate that both estimated  $a_t$  and  $\tilde{a}_t$  fit the true correlation equally well. However, the implied correlation figure of the CMSV

model is always more precise than the estimate of contemporaneous relation. This is not surprising as the model is designed to produce valid draws of a covariance matrix and not of a correlation matrix. In other words, estimates of the contemporaneous relation are less precise than an implied correlation estimate because the estimation routine does not accounts for the bounds of correlation.

Table 2: Precision and discrepancy of posterior median estimates

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

The table shows the forecast accuracy (MAE) and distance (MAD) for estimated correlation  $\rho_t$  and contemporaneous relation  $a_t$  and  $\tilde{a}_t$ . A bold figure highlights the best model in each panel and row.

Next, the MAD figures indicate that the distance between alternative estimates of  $a_t$  and  $\tilde{a}_t^*$  obtained under different orderings of variable are of the same magnitude. The average distance among all simulated processes is 0.026 which is rather small. However, the distance between 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. Further, the distance between estimated correlation is the smallest with 0.016 among all estimates. This indicates that even though the likelihood drives pseudo estimates of correlation  $a_t$  and  $\tilde{a}_t$  apart, it does, however, not martially affect the correlation estimate.

Therefore, this Monte Carlo simulation provides evidence that the estimated correlation of the CMSV model fits the data well and is almost invariant to a rotation of variables for homoskedastic data.

### 3 The DC-Cholesky MSV Model

The property that the estimated correlation of a CMSV model is almost rotationally invariant for homoskedastic data is an attractive feature of the model. This property can be exploited to construct a new multivariate stochastic volatility with separate volatility and correlation dynamics in the spirit of Engle (2002). This section presents the details of the dynamic correlation Cholesky multivariate stochastic volatility (DC-Cholesky MSV or DC-CMSV) model.

Let  $y_t$  be a mean zero vector process with a time-varying covariance matrix  $\Sigma_t$

$$y_t \sim N(0, \Sigma_t) \quad (1)$$

Then,  $\Sigma_t$  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

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

Then, it follows

$$y_t = D_t \epsilon_t, \quad \epsilon_t \sim N(0, R_t) \quad (2)$$

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

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

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)

$$R_t = Q_t^{*-1/2} Q_t Q_t^{*-1/2} \quad (3)$$

$$Q_t = A_t^{*-1} D_t^* D_t^{*'} A_t^{*-1} \quad (4)$$

$$Q_t^* = \text{diag}[\text{vecd}(Q_t)] \quad (5)$$

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

$$h_t = h_{t-1} + \epsilon_t^h, \quad (6)$$

$$a_t^* = a_{t-1}^* + \epsilon_t^{a^*}, \quad (7)$$

$$h_t^* = h_{t-1}^* + \epsilon_t^{h^*} \quad (8)$$

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

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

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

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

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

Next, the Gibbs sampling algorithm for the DC-CMSV model is presented, which builds on the notation and results from Chan (2017).

**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  $s^{(r)} \sim (s|y, h^{(r-1)})$
  - Draw  $h^{(r)} \sim (h|y, s^{(r)}, W^{(r-1)}, h_0^{(r-1)})$
  - Draw  $W^{(r)} \sim (W|h^{(r)}, h_0^{(r-1)})$
  - Draw  $h_0^{(r)} \sim (h_0|y, h^{(r)}, W^{(r)})$
  - Draw  $\epsilon^{(r)} \sim (\epsilon|y, h^{(r)})$
  
2. posterior draws from  $p(a^*, a_0^*, S^*|\epsilon^{(r)}, h^{*(r-1)})$ 
  - Draw  $a^{*(r)} \sim (a^*|\epsilon^{(r)}, a_0^{*(r-1)}, S^{*(r-1)}, h^{*(r-1)})$
  - Draw  $S^{*(r)} \sim (S^*|\epsilon^{(r)}, a^{*(r)}, a_0^{*(r-1)})$
  - Draw  $a_0^{*(r)} \sim (a_0^*|\epsilon^{(r)}, a^{*(r)}, S^{*(r)})$
  
3. posterior draws from  $p(s^*, h^*, W^*, h_0^*|\epsilon^{(r)}, a^{*(r)})$ 
  - Draw  $s^{*(r)} \sim (s^*|\epsilon^{(r)}, h^{*(r-1)})$
  - Draw  $h^{*(r)} \sim (h^*|\epsilon^{(r)}, s^{*(r)}, W^{*(r-1)}, h_0^{*(r-1)})$
  - Draw  $W^{*(r)} \sim (W^*|h^{*(r)}, h_0^{*(r-1)})$
  - Draw  $h_0^{*(r)} \sim (h_0^*|\epsilon^*, h^{*(r)}, W^{*(r)})$

Relative to the traditional 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 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 correlation matrix  $R_t$  by equation (3)-(5). The draw of  $R_t$  in conjunction with  $D_t$  is used to span the covariance matrix  $\Sigma_t$ .

Next, merits and drawbacks of the DC-CMSV model are highlighted. First, the posterior distributions of the marginal volatilities is rotationally invariant. Next, if the process for marginal volatilities is correctly specified, then the estimated correlation matrix is almost rotationally invariant. Thus, the estimated covariance matrix

under the DC-CMSV model is almost rotationally invariant as well. Moreover, parameter of contemporaneous relations may accommodate nonlinear dynamics.

Nevertheless, the appealing properties come at the cost of increased computational complexity. Particularly, the computational costs increase as the volatility series needs to be sampled twice instead of once, i.e. the independent volatilities and the auxiliary volatilities for the estimation of the correlation matrix. This increased computational complexity, however, is not substantial threat to the applicability of the model. In principle, the marginal volatilities can be sampled by parallel routines.

Notice when there is strong prior belief that the data is actually generated by the CMSV model with a particular ordering, then the estimates of the DC-CMSV model may be substantially off the true values. This property arises as the model imposes a roughly linear restriction on the correlation process, which, however, cannot be used to obtain the true nonlinear movements of the correlation. In this case, the CMSV model is the preferred model.

## 4 Simulation Evidence

This section conducts a Monte Carlo simulation to evaluate forecast accuracy and the sensitivity to alternative ordering of variables for the CMSV model and the DC-CMSV model. For comparison, the integrated dynamic conditional correlation (IDCC) model of Engle (2002) is added as a benchmark.

As data generating process (DGP), the stationary bivariate stochastic volatility model with known correlation structure of Asai and McAleer (2009) is used to simulate 1000 observation, or four years of daily financial data. 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.

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

then use them for each correlation process,

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

which are the same as in Section 2.2.

The hyperparameters  $(k_S, k_W)$  and  $(k_W, k_S^*, k_W^*)$  are all set to 0.1. The MCMC estimation produce 35000 samples of which 15000 are reserved for the burnin period. The IDDC model is estimated by the quasi-maximum likelihood method using the MFE Toolbox of Sheppard (2013).

All statistics are based on 250 Monte Carlo replications. The forecast accuracy of the estimated parameter is evaluated by the average mean absolute error (MAE) over all possible permutations. The average MAE is computed for the estimated correlation, covariance and value-at-risk of an equally weighted portfolio.

$$VaR_t = 1.65 \sqrt{\omega^2 \cdot V_{1,t} + (1 - \omega)^2 \cdot V_{2,t} + 2 \cdot \omega \cdot (1 - \omega) \cdot \rho_t \cdot V_{1,t}^{\frac{1}{2}} \cdot V_{2,t}^{\frac{1}{2}}}$$

where  $V_{i,t} = \exp(h_{i,t})$  is the variance for  $i = 1, 2$ ,  $\rho_t$  is the correlation and  $\omega = 0.5$  is the weight of the portfolio.

Next, the sensitivity of estimated parameters is assessed by distance and correlation metrics. As distance metrics, the mean absolute difference (MAD) and the root mean square difference (RMSD) are reported. When the MAD and the RMSD diverge substantially, then this indicates that there are periods when the distance between estimated parameters is unusually large.

The similarity of the parameter estimates 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.

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

where  $\Delta$  denotes the first difference operator.

## 4.1 Results

To illustrate the sensitivity, Figure 2 shows estimated posterior median of the ratio of volatilities and parameter of contemporaneous relation for a selected ordering and both models in the upper panel, and estimated correlation and covariance for a selected model and both orderings in the lower panel for one replication of the sine correlation process.

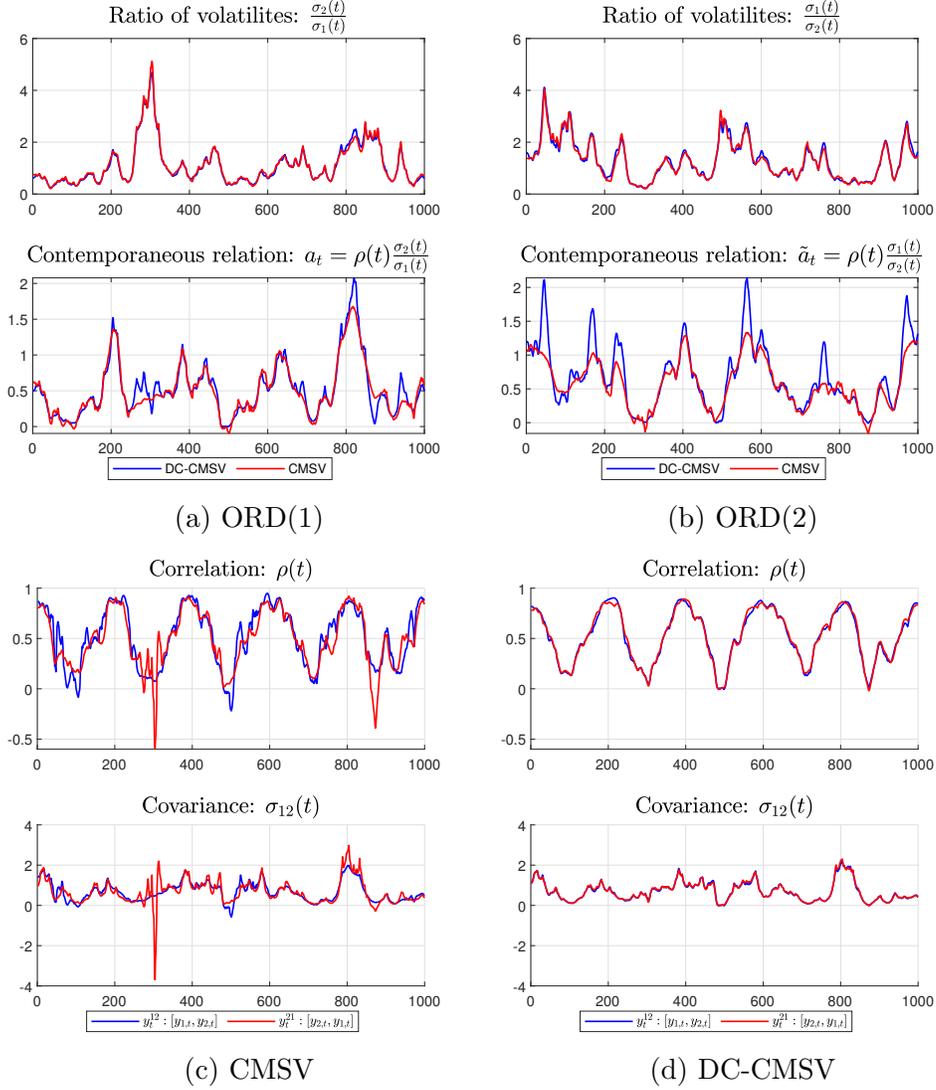
The figure shows that the estimated ratio of volatilities is similar across models and that it exhibits nonlinear patterns over time. For the parameter of contemporaneous relation, however, there are marked differences across models. Especially, estimates of the DC-CMSV model exhibit nonlinear dynamics, which is linked to the ratio of volatilities. However, when these nonlinearities are not properly captured, as in the CMSV model, then the estimated correlation and covariance may exhibit systematic differences across orderings. Particularly, the estimated correlation and covariance appear more smooth in ORD(1), while it exhibits some wobbliness and spikes in ORD(2). This pattern is the flipside of a rather spiky path of the contemporaneous relation in ORD(1) but smooth path in ORD(2).

Table 3-5 present the performance evaluation of the estimated correlation, covariance, and value-at-risk, respectively. Notice the IDCC model is independent of the ordering of variables, hence, no distance and similarity metrics are reported.

Table 3 shows that the DC-CMSV produces the most precise estimates of the correlation for all except the fast sine correlation process. Here the CMSV performs best. For the benchmark DGP, the gains in accuracy are moderate. However, the forecast accuracy of the CMSV model deteriorates substantially for the high volatility DGP, while the figures of the CMSV and the DC-CMSV are similar for the low volatility DGP. Moreover, the forecast accuracy of the DC-CMSV remains similar under alternative scales for the innovations to stochastic volatility.

Next, the MAD and the RMSD for the CMSV indicate that there may be a considerable distance between the estimated correlation across alternative orderings. Particularly, the RMSD for the benchmark DGP documents that a distance of 0.04 to 0.11 for the estimated correlation path is not unusual. These figures substantially in-

Figure 2: Conditional covariance matrix (sine, benchmark DGP)



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

Table 3: Estimated correlation

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.040	<b>0.030</b>	0.045	0.031	<b>0.025</b>	0.041	<b>0.031</b>	<b>0.307</b>	-0.284
sine	0.096	<b>0.086</b>	0.150	0.063	<b>0.018</b>	0.088	<b>0.023</b>	0.559	<b>0.943</b>
fastsine	<b>0.249</b>	0.256	0.256	0.082	<b>0.011</b>	0.108	<b>0.014</b>	0.382	<b>0.911</b>
step	0.078	<b>0.061</b>	0.083	0.044	<b>0.017</b>	0.065	<b>0.022</b>	0.499	<b>0.768</b>
ramp	0.118	<b>0.110</b>	0.168	0.067	<b>0.020</b>	0.096	<b>0.027</b>	0.567	<b>0.942</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.060	<b>0.040</b>	0.072	0.048	<b>0.025</b>	0.076	<b>0.031</b>	<b>0.304</b>	-0.269
sine	0.118	<b>0.090</b>	0.155	0.108	<b>0.015</b>	0.146	<b>0.020</b>	0.278	<b>0.947</b>
fastsine	<b>0.248</b>	0.256	0.258	0.144	<b>0.008</b>	0.185	<b>0.011</b>	0.127	<b>0.924</b>
step	0.103	<b>0.067</b>	0.095	0.080	<b>0.014</b>	0.120	<b>0.018</b>	0.291	<b>0.788</b>
ramp	0.138	<b>0.114</b>	0.179	0.112	<b>0.017</b>	0.154	<b>0.022</b>	0.297	<b>0.945</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	<b>0.024</b>	0.029	0.029	<b>0.023</b>	0.037	<b>0.029</b>	<b>0.139</b>	-0.247
sine	0.086	<b>0.083</b>	0.147	0.043	<b>0.021</b>	0.060	<b>0.027</b>	0.772	<b>0.942</b>
fastsine	<b>0.253</b>	0.256	0.256	0.039	<b>0.014</b>	0.052	<b>0.017</b>	0.667	<b>0.901</b>
step	0.064	<b>0.057</b>	0.076	0.029	<b>0.017</b>	0.040	<b>0.022</b>	0.669	<b>0.792</b>
ramp	0.110	<b>0.108</b>	0.166	0.045	<b>0.023</b>	0.066	<b>0.031</b>	0.763	<b>0.940</b>

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

flate and deflate for the high volatility DGP and the low volatility DGP, respectively. In contrast, analogous figures for the DC-CMSV are hardly affected by a different scale of idiosyncratic volatility patterns in the simulated data.

Turing to the similarity metric, the FD metric is fairly below one for the estimated correlation of the CMSV model. This indicates that changes of the estimated correlation paths feature some idiosyncratic components. For the DC-CMSV model,

the FD is substantially higher, but also somewhat lower than one. For the case of constant correlation, the metric is negative which suggest that the estimated correlation path exhibit some mirror type behavior. However, the FD figure should not be interpreted in isolation. In fact, the small MAD and RMSD figures indicate that estimated correlations are indeed very similar for the DC-CMSV.

For the estimated covariance, Table 4 shows that the DC-MSV produces for

Table 4: Estimated covariance

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.313	<b>0.285</b>	0.374	0.169	<b>0.031</b>	0.240	<b>0.044</b>	0.378	<b>0.970</b>
sine	0.212	<b>0.199</b>	0.304	0.102	<b>0.021</b>	0.153	<b>0.030</b>	0.668	<b>0.981</b>
fastsine	<b>0.331</b>	<b>0.331</b>	0.350	0.106	<b>0.013</b>	0.159	<b>0.018</b>	0.622	<b>0.990</b>
step	0.248	<b>0.224</b>	0.303	0.109	<b>0.021</b>	0.167	<b>0.030</b>	0.566	<b>0.981</b>
ramp	0.232	<b>0.221</b>	0.315	0.104	<b>0.024</b>	0.158	<b>0.034</b>	0.642	<b>0.978</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.436	<b>0.389</b>	0.548	0.207	<b>0.037</b>	0.349	<b>0.057</b>	0.469	<b>0.988</b>
sine	0.288	<b>0.253</b>	0.406	0.179	<b>0.021</b>	0.307	<b>0.033</b>	0.572	<b>0.993</b>
fastsine	0.403	<b>0.396</b>	0.440	0.210	<b>0.012</b>	0.351	<b>0.019</b>	0.467	<b>0.997</b>
step	0.348	<b>0.301</b>	0.433	0.169	<b>0.021</b>	0.299	<b>0.033</b>	0.562	<b>0.994</b>
ramp	0.312	<b>0.281</b>	0.421	0.183	<b>0.023</b>	0.313	<b>0.037</b>	0.553	<b>0.992</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.238	<b>0.217</b>	0.259	0.153	<b>0.026</b>	0.201	<b>0.034</b>	0.192	<b>0.931</b>
sine	0.170	<b>0.163</b>	0.240	0.081	<b>0.023</b>	0.114	<b>0.030</b>	0.665	<b>0.963</b>
fastsine	<b>0.296</b>	0.296	0.301	0.049	<b>0.015</b>	0.068	<b>0.020</b>	0.768	<b>0.969</b>
step	0.190	<b>0.177</b>	0.224	0.089	<b>0.019</b>	0.130	<b>0.026</b>	0.473	<b>0.955</b>
ramp	0.190	<b>0.185</b>	0.251	0.085	<b>0.025</b>	0.121	<b>0.034</b>	0.637	<b>0.957</b>

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

Table 5: Estimated value-at-risk

## (a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.252	<b>0.236</b>	0.306	0.115	<b>0.011</b>	0.149	<b>0.014</b>	0.496	<b>0.992</b>
sine	0.212	<b>0.207</b>	0.281	0.072	<b>0.009</b>	0.097	<b>0.012</b>	0.820	<b>0.994</b>
fastsine	0.235	<b>0.231</b>	0.284	0.069	<b>0.006</b>	0.091	<b>0.008</b>	0.840	<b>0.996</b>
step	0.227	<b>0.217</b>	0.285	0.075	<b>0.009</b>	0.104	<b>0.011</b>	0.723	<b>0.994</b>
ramp	0.218	<b>0.213</b>	0.286	0.075	<b>0.010</b>	0.105	<b>0.014</b>	0.762	<b>0.993</b>

## (b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.329	<b>0.307</b>	0.427	0.133	<b>0.012</b>	0.184	<b>0.015</b>	0.663	<b>0.997</b>
sine	0.282	<b>0.271</b>	0.390	0.113	<b>0.009</b>	0.158	<b>0.012</b>	0.856	<b>0.998</b>
fastsine	0.300	<b>0.290</b>	0.390	0.121	<b>0.006</b>	0.162	<b>0.008</b>	0.847	<b>0.998</b>
step	0.301	<b>0.285</b>	0.400	0.108	<b>0.008</b>	0.156	<b>0.011</b>	0.805	<b>0.998</b>
ramp	0.286	<b>0.276</b>	0.394	0.114	<b>0.009</b>	0.161	<b>0.013</b>	0.828	<b>0.998</b>

## (c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.198	<b>0.185</b>	0.222	0.114	<b>0.010</b>	0.144	<b>0.013</b>	0.263	<b>0.980</b>
sine	0.169	<b>0.165</b>	0.211	0.069	<b>0.011</b>	0.091	<b>0.014</b>	0.671	<b>0.983</b>
fastsine	0.196	<b>0.195</b>	0.218	0.037	<b>0.007</b>	0.049	<b>0.009</b>	0.849	<b>0.989</b>
step	0.179	<b>0.172</b>	0.209	0.068	<b>0.009</b>	0.094	<b>0.011</b>	0.571	<b>0.985</b>
ramp	0.175	<b>0.172</b>	0.215	0.076	<b>0.012</b>	0.101	<b>0.016</b>	0.578	<b>0.978</b>

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

almost all cases the most precise estimates. The MAD and RMSD indicate that the distance between estimated covariances is small, while estimates of the CMSV may exhibit substantial differences. Strikingly, the FD statistics is very close to one for the estimates of the DC-MSV. This indicates that the estimated covariances are almost identical across ordering. Since the estimated volatilities of the DC-MSV are by construction independent of the ordering, this table provides strong evidence that the estimated covariance matrix of the DC-MSV is almost rotationally invariant.

Moreover, the similarity and distance metric indicate that the estimates of the CMSV are more different the stronger the idiosyncratic volatility patterns.

For value-at-risk, Table 5 draws an unambiguous picture. The DC-MSV produces the most precise estimates, which are virtually indistinguishable across alternative orderings. The estimates of the CMSV, however, are less precise and it is not unusual that estimates diverge substantially across orderings. From the perspective of risk management, this is clearly an undesirable property as alternative estimates may indicate very different losses for the portfolio over time.

## 4.2 Robustness

The influence of the hyperparameters for the innovation variance is different across models and across orderings. Therefore, it is not clear whether this choice resembles a fair model comparison. For this reason, the model and their hyperparameter are re-estimated using the algorithm of Amir-Ahmadi, Matthes, and Wang (2018).

Table C.1 – C.3 in Appendix C.1 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, in general, slightly smaller than the chosen hyperparameters. Broadly, the performance metrics improve slightly in all dimensions and for both models.

Another concern is misspecification. The true DGP assumes stationary volatility dynamics. This form of misspecification may affect the ability of the DC-CMSV model to take out the nonlinearities of the data, which is important to obtain almost rotationally invariant estimates. Therefore, both models are re-estimated assuming stationary law of motion for volatility and parameter of contemporaneous relation. For this exercise, the stationary DCC model of Engle (2002) is added as a benchmark.

Table C.4 – C.6 in Appendix C.2 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 under 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.

## 5 Empirical Application

This section reviews the robustness of the results for Primiceri (2005)'s application in the light of alternative orderings. The introductory chart (Figure 1) shows that the estimated covariance are sensitive to alternative orderings and exhibit marked differences during the stagflation period. These differences may have an effect on the structural analysis for this particular application.

To investigate this issue, the structural analysis is replicated for all possible permutations of the variables along the lines of a two-step identified structural VAR as described in Rubio-Ramirez, Waggoner, and Zha (2010). The reduced-form parameters of the model are estimated under all possible orderings, then all parameters are reordered back into the original position of variables. Then, the structural model is identified via the Cholesky decomposition of the covariance matrix in each period. All estimates are based on 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. Also, 70,000 draws of the Gibbs sampler are produced, while the first 20,000 are discarded in the burn-in period.

Figure B.1 and B.2 in Appendix B depict the posterior median of the estimated time-varying VAR parameters for each variable in the respective column. The estimates are rather similar across alternative ordering for the original model, while those of the DC-CMSV version of the model are virtually indistinguishable.

Next, Figure B.3 and B.4 in Appendix B show the posterior median of the estimated correlation, covariance, and volatility of the reduced-form residuals. For the original model, the estimated correlation and covariance exhibit pronounced differences under the original model. The difference across volatility estimates is however rather small. In contrast, all estimates of the DC-CMSV version of the model are virtually indistinguishable across alternative orderings. Taking stock, these properties of the estimated reduced-form parameters clearly indicate that results of the structural analysis may only change when an estimate of the covariance is involved.

Also, notice the estimates of the DC-CMSV version of the model are similar to a model average over the estimates of all possible orderings of the original model. Thus, this figure demonstrates that alternative estimates of the original model are not arbitrary different, but exhibit some systematic differences, which averages out over possible orderings. Moreover, already small variations in the ratio of volatilities may lead to systematically different estimates. This provides evidence that this property is not only relevant in simulations but also for empirical applications.

Having documented general differences across reduced-form parameter estimates, the consequences are analyzed in more detail for two particular orderings. Specifically, the estimates of the original ordering and the reverse ordering are contrasted as differences between estimated covariances are most pronounced. Turning to the results, the estimates of U.S. systematic interest rate response for inflation and unemployment exhibit significant differences across these two particular ordering. This is hardly surprising as these estimates use the estimated covariance as an input. There are no marked difference for the other exercises in Primiceri's (2005) application.<sup>9</sup>

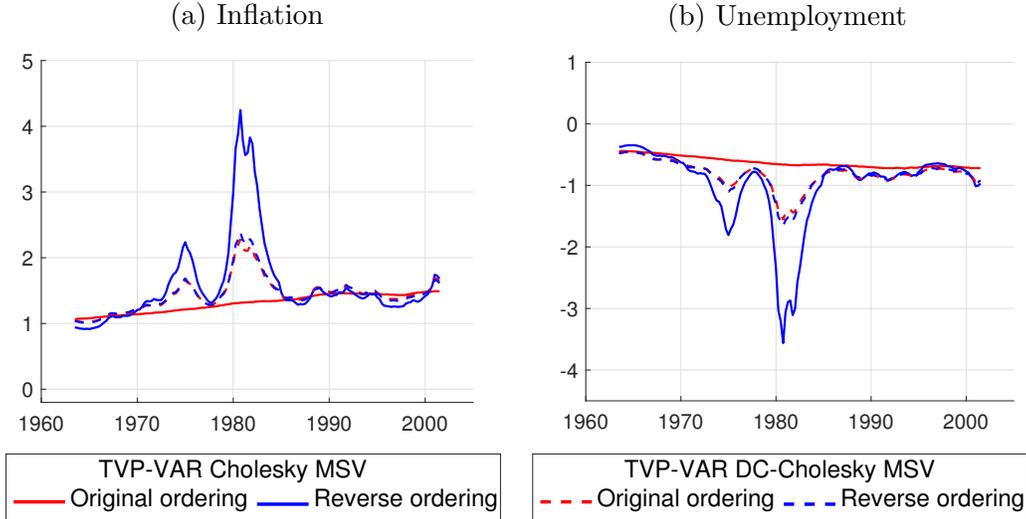
Figure 3 shows the estimated long-run U.S. systematic interest rate response to inflation and unemployment for both the original model and the DC-CMSV version of the model. In particular, the estimates of the original model provide evidence for two equally plausible but mutually exclusive conclusions on how U.S. systematic monetary policy reacted during the stagflation period. The original estimates provide evidence for a muted response, while those under the reverse ordering point to a drastically changing and aggressive response. As a result, this chart provides evidence that the ordering of variables may be so decisive that it may alter conclusions drawn from this model. Therefore, the choice of the ordering is not negligible in the CMSV model. It must be considered as an additional identifying assumption on the dynamic evolution of the reduced-form covariance matrix which may be or may be not attractive feature for this application.

Besides, estimates of DC-CMSV version of the model draw an unambiguous conclusion under all possible ordering. The results suggest that the reaction function was modestly more aggressive. This evidence is consistent with the finding in Sims

---

<sup>9</sup>A comparison for all exercises is documented in Appendix B (Figure B.5 – B.12).

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



The figure depicts interest rate response to a 1% permanent increase of inflation and unemployment rate under alternative orderings using the two-step procedure.

and Zha (2006). Particularly, this paper provides strong evidence that there were regime switches for the conduct of monetary policy during the period of stagflation.

## 6 Conclusion

This paper studied the dynamic properties of the estimated reduced-form time-varying covariance matrix of the CMSV model under alternative orderings of variables. It found that this model imposes alternative dynamic restrictions across orderings when the ratio of volatilities varies over time. The DC-CMSV model was proposed as a robust alternative. The reduced-form estimates of this model are almost rotationally invariant. The model also allows for nonlinear dynamics between contemporaneous relations. For the investigated example, it was illustrated that rotational non-invariance is not negligible and that estimates under alternative orderings may lead to ambiguous conclusions. This property is likely to be important for many empirical applications as it is not uncommon that financial and economic time series to exhibit individual volatility dynamics. Thus, these findings suggest that the

estimates based on the CMSV model should be interpreted with some caution. In addition, the relatively small costs of computing a more robust estimate using the DC-CMSV model seems sensible for most empirical applications.

The main finding that specific state dynamics impose restrictions on the reduced-form estimates may not only be limited to this state space model. Also, the linear dynamic factor model with time-varying factor loadings and stochastic volatility may suffer from similar restrictions. Future research should investigate whether rotational non-invariance in this class of models is also driven by a too 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.
- 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 (2013): “Time-varying effects of oil supply shocks on the US economy,” *American Economic Journal: Macroeconomics*, 5(4), 1–28.
- 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*.
- CHAN, J. (2017): “Notes on Bayesian Macroeconometrics,” .
- CHAN, J. C., A. DOUCET, R. LEÓN-GONZÁLEZ, AND R. W. STRACHAN (2018): “Multivariate stochastic volatility with co-heteroscedasticity,” .
- 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., 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, 10.1093/restud/rdv024.
- 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.

- 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,” *Discussion papers*.
- 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,” .
- PHILIPPOV, A., AND M. E. GLICKMAN (2006): “Multivariate Stochastic Volatility via Wishart Processes,” *Journal of Business & Economic Statistics*, 24(3), 313–328.
- 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.
- SHEPPARD, K. (2013): *MFE MATLAB Function Reference Financial Econometrics*.
- 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*. John Wiley & Sons, 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 “ $\Sigma_t$  under CMSV model” .*

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

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

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

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

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

$$\sigma_{11,t}^2 = \sigma_{11,t-1}^2 \exp(2\epsilon_{1,t}^g). \tag{A.3}$$

Combining (A.2) with (A.3) gives

$$\begin{aligned}
 \frac{1 + b_t a_{t-1}^2}{1 + b_t (a_{t-1} + \epsilon_t^a)^2} \sigma_{11,t-1}^2 \exp(2\epsilon_{2,t}^g) &= \sigma_{11,t-1}^2 \exp(2\epsilon_{1,t}^g) \\
 \frac{1 + b_t a_{t-1}^2}{1 + b_t (a_{t-1} + \epsilon_t^a)^2} \exp(2\epsilon_{2,t}^g) &= \exp(2\epsilon_{1,t}^g). \tag{A.4}
 \end{aligned}$$

Since  $\epsilon_t^a$  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  $\rho_t$  depends on reduced-form parameters by

$$\begin{aligned}
\rho_t &= a_t \frac{\sigma_{11,t}}{\sigma_{22,t}} \\
&= a_{t-1} \frac{\sigma_{11,t-1}}{\sigma_{22,t-1}} \frac{\exp(\epsilon_{1,t}^g)}{\exp(\epsilon_{2,t}^{g^{**}})} + \epsilon_t^a \frac{\sigma_{11,t}}{\sigma_{22,t}} \\
&= \rho_{t-1} \frac{\exp(\epsilon_{1,t}^g)}{\exp(\epsilon_{2,t}^{g^{**}})} + \epsilon_t^a \frac{\sigma_{11,t}}{\sigma_{22,t}}
\end{aligned} \tag{A.5}$$

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 is log-normally distributed, it follows from (A.5) that the correlation evolves nonlinearly. ■

*Subproof of Claim (3).* Trivial. 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

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

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 structure induced into  $\Sigma_t^*$  and  $\tilde{\Sigma}_t$  are different.

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

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

When the ratio of variances is constant, then there is a mapping from the parameters of the analogous set up CMSV model to the implied transformed parameters. However, since the ratio of variances cannot be constant for this DGP, this mapping does not exist. The analogous set up state equations then depart stronger from the true parameters

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

the higher variability of the ratio of variances.

□

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

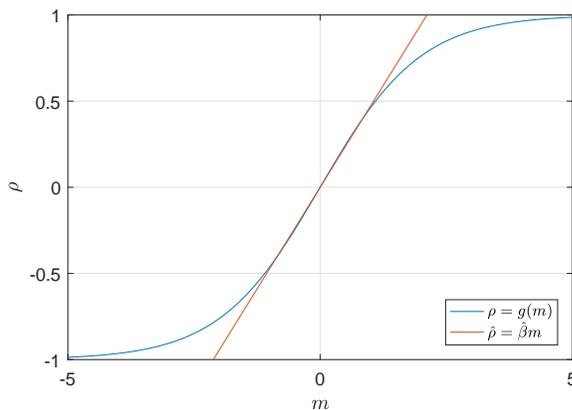
**Property** (Rotational invariance of  $\Sigma_t$  under DC-MSV model). Let  $y_t$  be generated by the DC-MSV model with covariance matrix  $\Sigma_t$ . Define the vector of variables with exchanged rows  $\tilde{y}_t$  and the permuted covariance matrix  $\tilde{\Sigma}_t = P\Sigma_t P'$ . Analogously,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

*Subproof of Claim (1).* For  $g : m \rightarrow \rho$  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  $\rho$  on  $m$  is  $5.5e-5$ . Figure A.1 compares the linear prediction for  $\rho$  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 A.1: Fisher transformation: mapping between  $\rho$  and  $m$



■

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

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

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

$$a_t = a_{t-1} + \eta_t^\rho c, \quad \tilde{a}_t = \tilde{a}_{t-1} + \eta_t^\rho \frac{1}{c}, \quad (\text{A.6})$$

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

$$a_t = a_{t-1} \frac{\exp(\eta_{2,t}^h)}{\exp(\eta_{1,t}^h)} + \eta_t^\rho \frac{\exp(h_{2,t})}{\exp(h_{1,t})}. \quad (\text{A.7})$$

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

$$a_t = a_{t-1} + (a_{t-1} + \eta_t^\rho \frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})})(\eta_{2,t}^h - \eta_{1,t}^h) + \eta_t^\rho \frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})}. \quad (\text{A.8})$$

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

$$error_t = a_t - \hat{a}_t \quad (\text{A.9})$$

where  $a_t$  and  $\hat{a}_t$  denote the resulting parameter under (A.7) and (A.8), respectively. The bias associated with this linear transition function is given by

$$bias_t = 1_{\{a_t > a_{t-1}\}}(\hat{a}_t - a_t) - (1 - 1_{\{a_t > a_{t-1}\}})(\hat{a}_t - a_t) \quad (\text{A.10})$$

where  $1_{\{a_t > a_{t-1}\}}$  is an indicator function, which ensures the correct sign of the bias.<sup>10</sup>

Figure A.2 illustrates the quantitative effects of innovations to stochastic volatility and innovation 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  $\frac{\exp(h_{2,t-1})}{\exp(h_{1,t-1})} = 1$ .<sup>11</sup> 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 (A.8) is further restricted to exhibit homoskedastically and normally distributed innovations, then these dynamic restrictions become tighter such that the approximation error and the bias becomes larger.

Next, when the order of variables is changed, then  $\tilde{a}_t$  has a similar functional form as  $a_t$  in (A.7), but the ratio of volatilities is inverted

$$\tilde{a}_t = \tilde{a}_{t-1} \frac{\exp(\eta_{1,t}^h)}{\exp(\eta_{2,t}^h)} + \eta_t^\rho \frac{\exp(h_{1,t})}{\exp(h_{2,t})}.$$

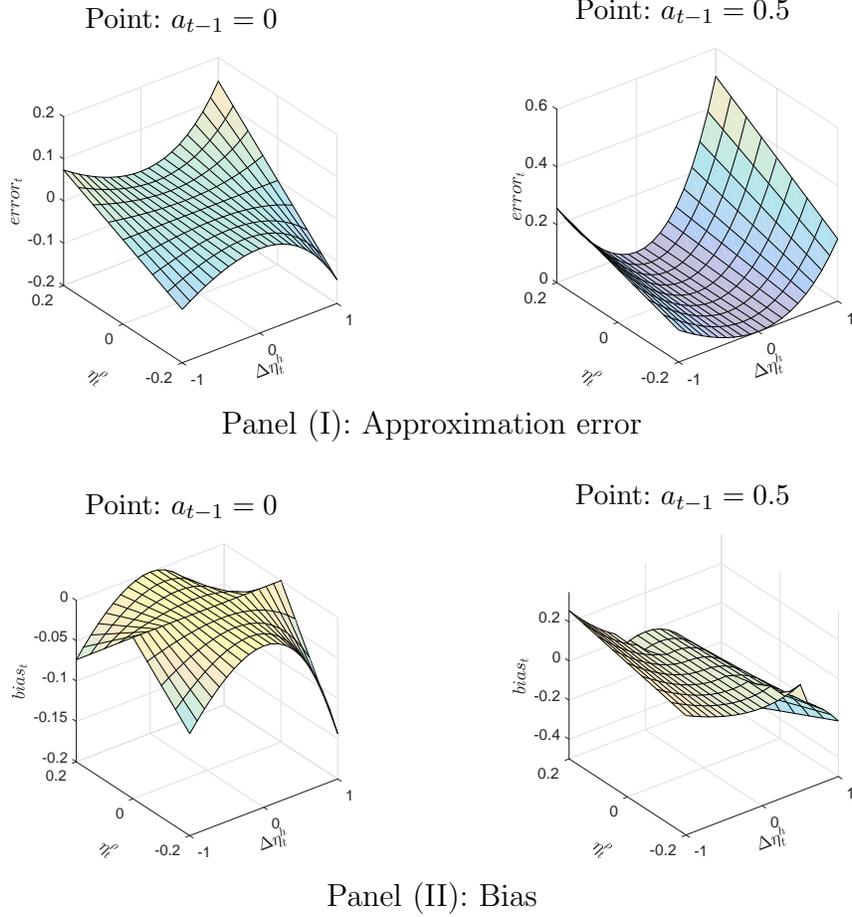
Consequently, when the ratio of volatilities increases then the equation of  $a_t^*$  underestimates the true transition of  $a_t$  in the original ordering, while the dynamic

---

<sup>10</sup>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.

<sup>11</sup>Notice the surface plots for a positive value of  $a_{t-1}$  are a reflection for negative value of  $a_{t-1}$ .

Figure A.2: Approximation error and bias of linearization



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 ratio of volatility.

equation of  $\tilde{a}_t^*$  mechanically overestimates the true transition of  $\tilde{a}_t$  in the alternative ordering.

Since the covariance term  $\sigma_{12,t}^*$  and  $\tilde{\sigma}_{12,t}^*$  is proportional to  $\{a_t^*, g_{1,t}^*\}$  and  $\{\tilde{a}_t^*, \tilde{g}_{2,t}^*\}$ , and  $h_{1,t} = g_{1,t}^*$  and  $h_{2,t} = \tilde{g}_{2,t}^*$  is left unrestricted, it follows that the bias in the contemporaneous parameter carries over to the covariance terms.  $\square$

*Proof of Property “Posterior distribution of  $a_t$  and  $\tilde{a}_t$  under homoscedasticity”.*

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 permutation matrix exchanging rows, the respective covariance matrices  $\Sigma_t = A_t^{-1}D_tD_t'A_t^{-1'}$  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  $\{1, \tilde{g}_t, \tilde{a}_t\}$ . Because of 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.<sup>12</sup>

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

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

---

<sup>12</sup>Notice that this does not imply that the elements in  $\Sigma_t$  and  $\tilde{\Sigma}_t$  have the same distribution.

where

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

Then, since  $\bar{a} = \tilde{a}$  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 solution for the individual elements in  $a$  and  $\tilde{a}$  differ, while they add up to the same sum in the time-invariant case.<sup>13</sup> Consequently, the likelihood information leads to a rotationally non-invariant posterior distribution for the time-varying parameter.  $\square$

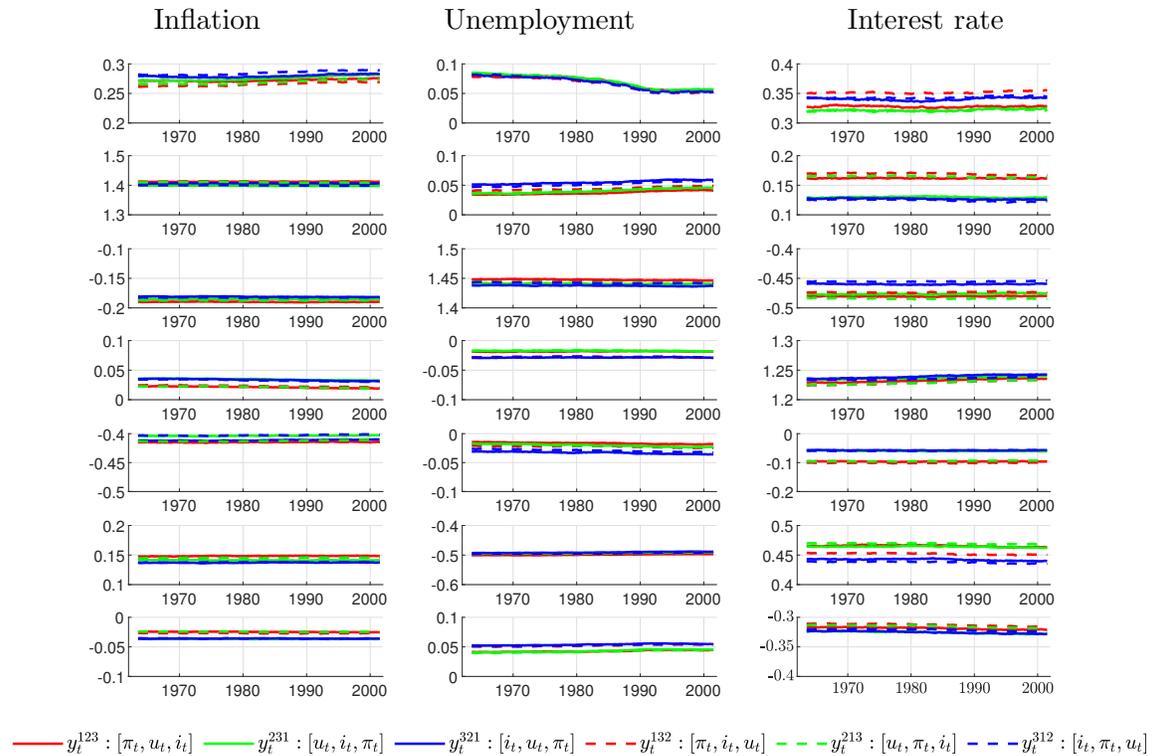
---

<sup>13</sup>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  $K_a = \tilde{K}_a$ .  $\bar{a} = \tilde{a}$  as  $\bar{a} = \frac{1}{\exp(2g_2)} \sum_{t=1}^T y_{1,t}y_{2,t}$  and  $\tilde{a} = \frac{1}{\exp(2g_1)} \sum_{t=1}^T y_{1,t}y_{2,t}$ . Hence, the posterior distribution of  $a$  and  $\tilde{a}$  is the same under alternative orderings of the variables.

## B Figures

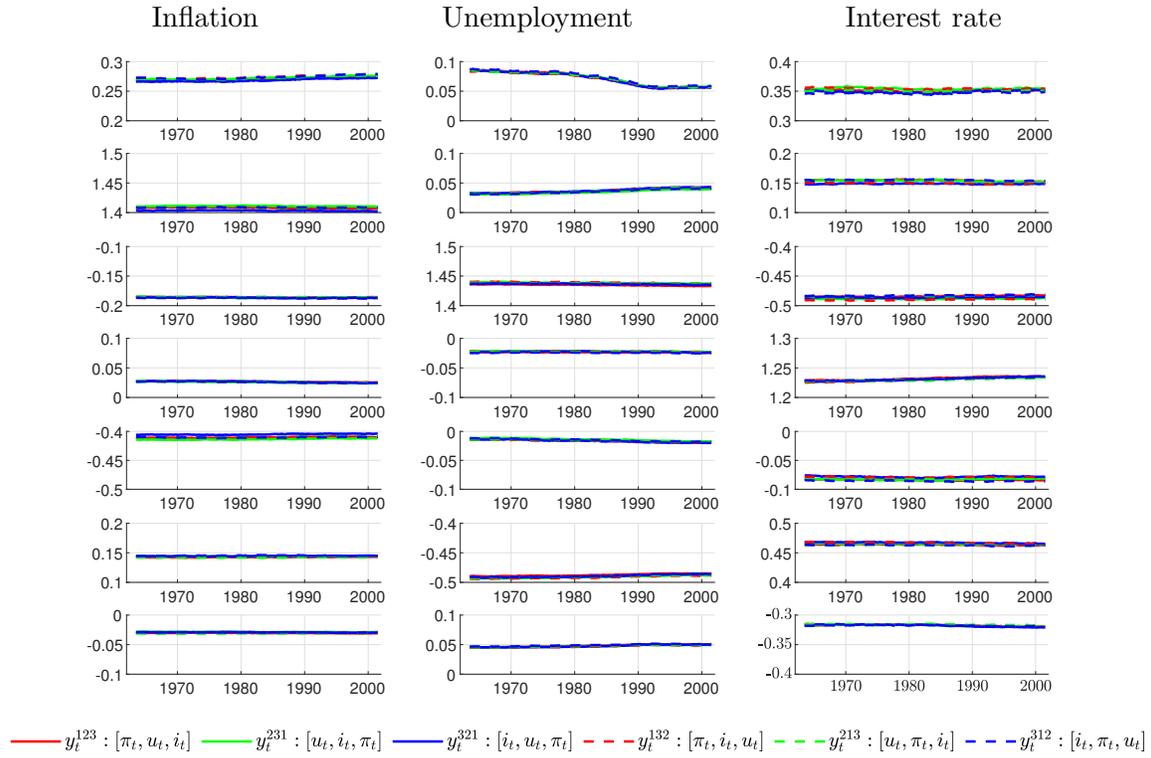
### B.1 Sensitivity of reduced-form parameters

Figure B.1: TVP-VAR CMSV



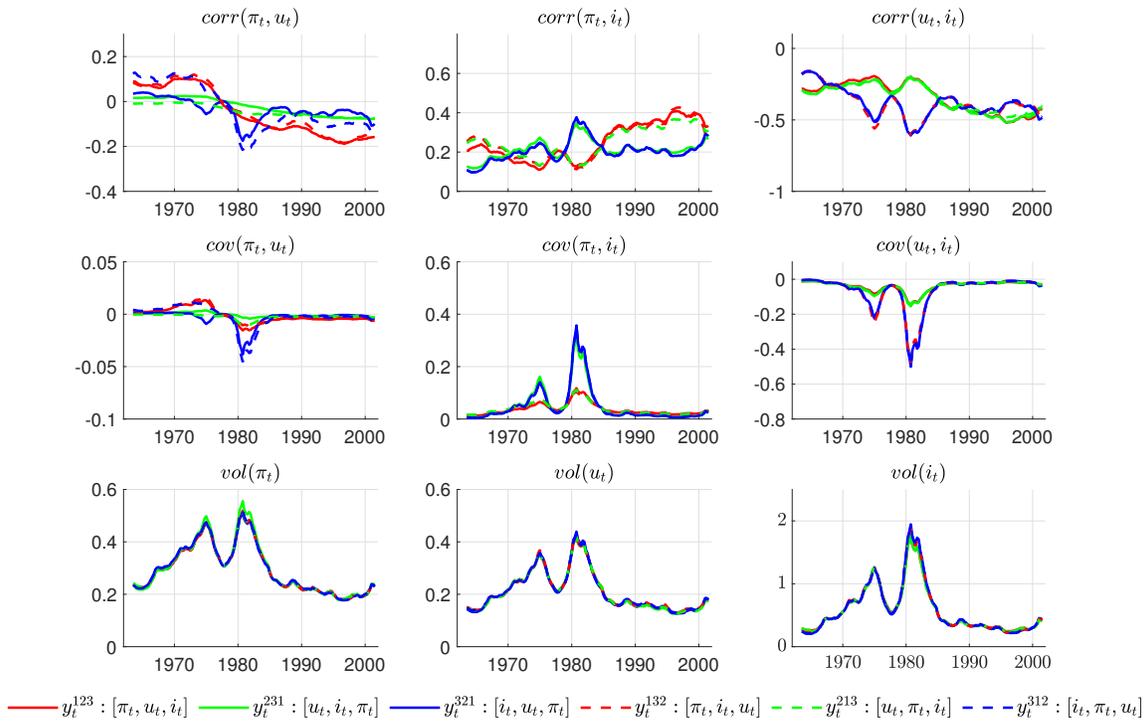
The figure depicts the posterior median of the time-varying VAR parameters for each equation in the respective column for alternative orderings.

Figure B.2: TVP-VAR DC-CMSV



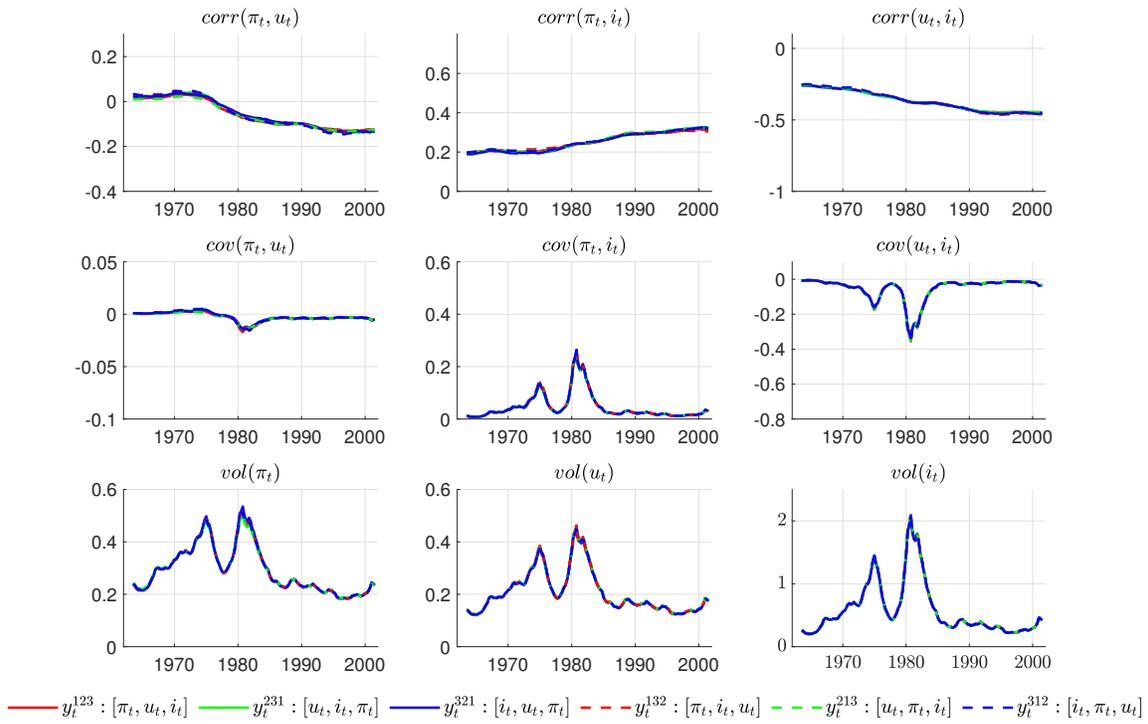
The figure depicts the posterior median of the time-varying VAR parameters for each equation in the respective column for alternative orderings.

Figure B.3: TVP VAR CMSV



The figure depicts posterior median of the correlation, covariance, and standard deviation of the reduced-form residual for alternative orderings. INF denotes the inflation rate, UNEMP denotes the unemployment rate and IR denotes the interest rate.

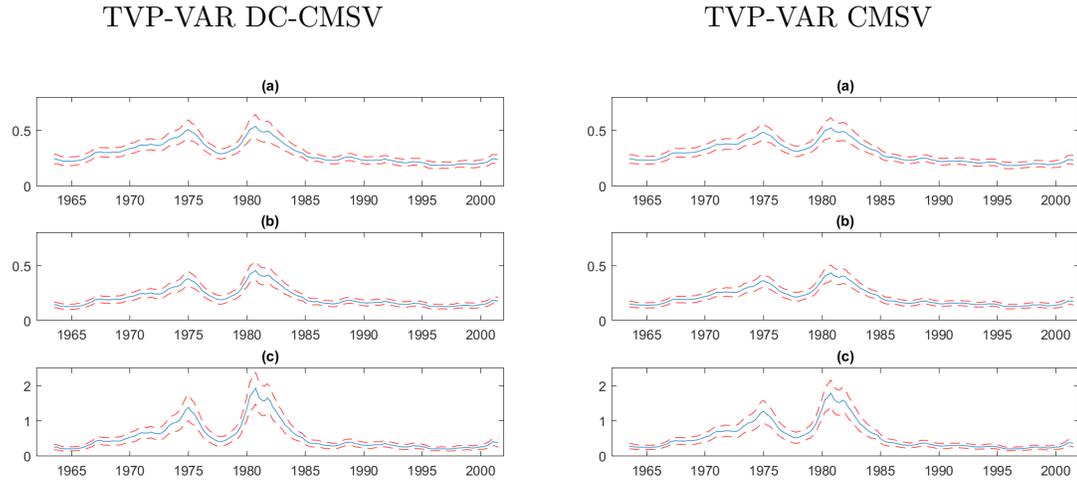
Figure B.4: TVP-VAR DC-CMSV



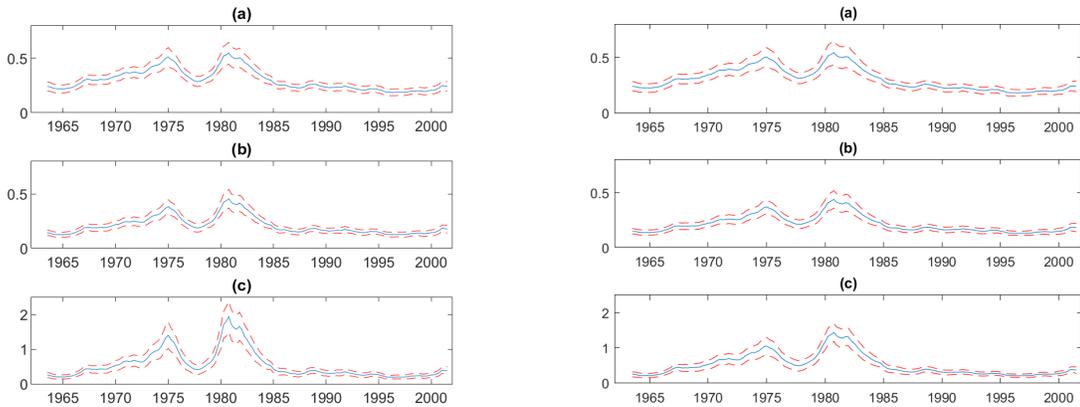
The figure depicts posterior median of the correlation, covariance, and standard deviation of the reduced-form residual for alternative orderings. INF denotes the inflation rate, UNEMP denotes the unemployment rate and IR denotes the interest rate.

## B.2 Sensitivity of structural analysis

Figure B.5: Replication Figure 1



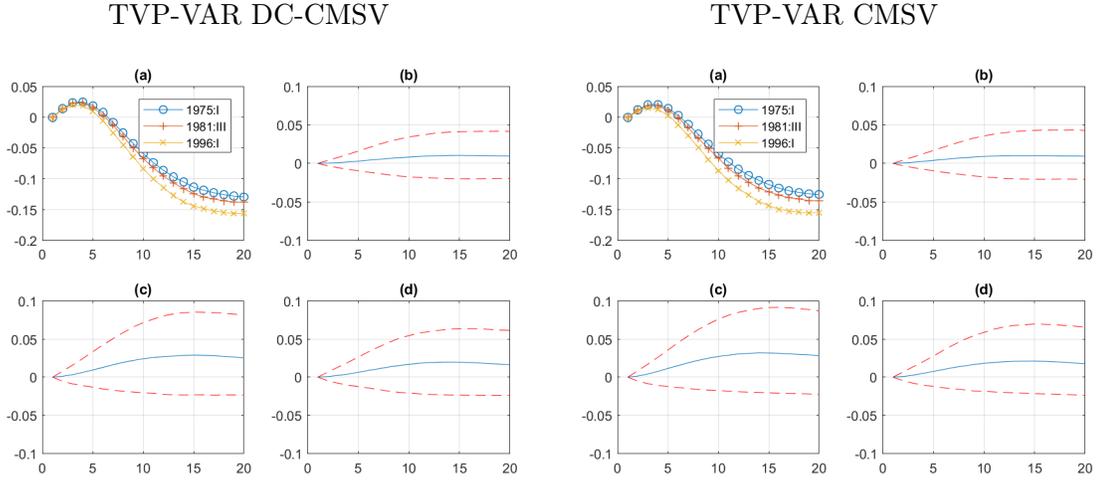
Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



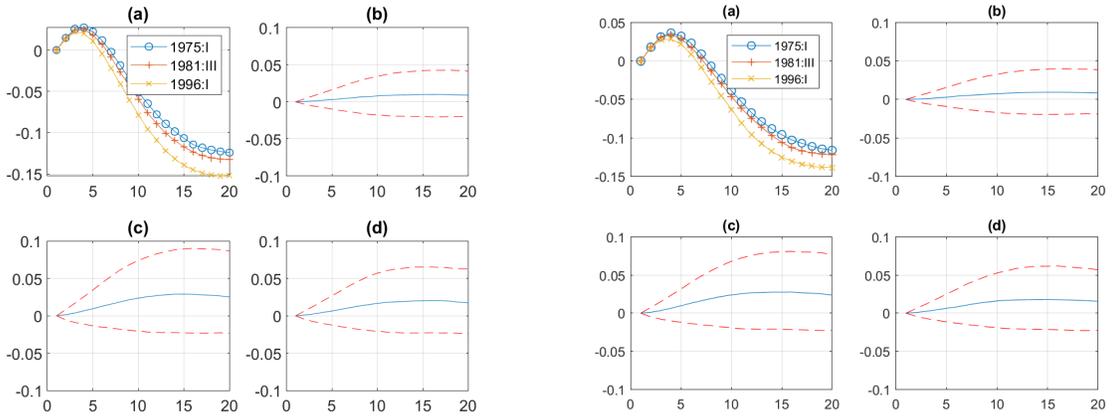
Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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 B.6: Replication Figure 2



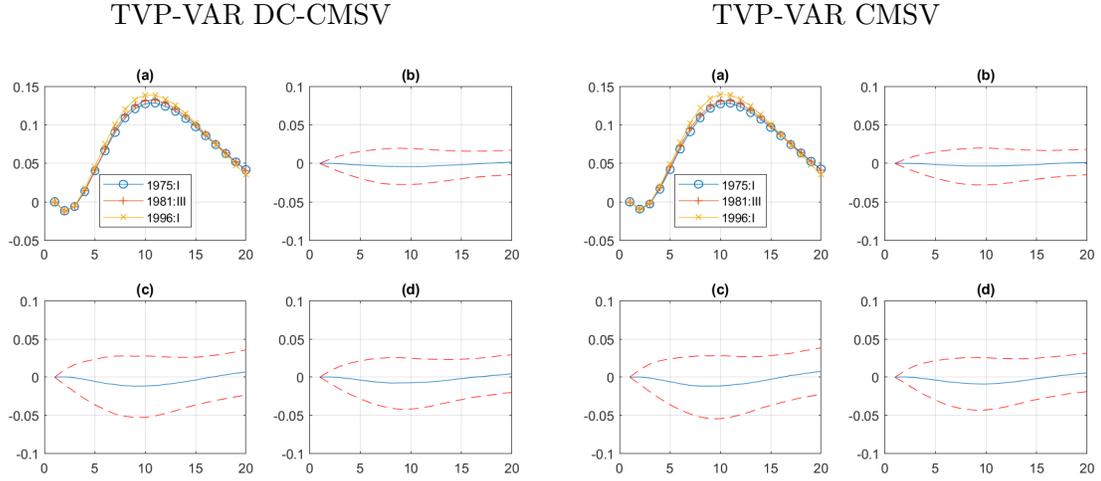
Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



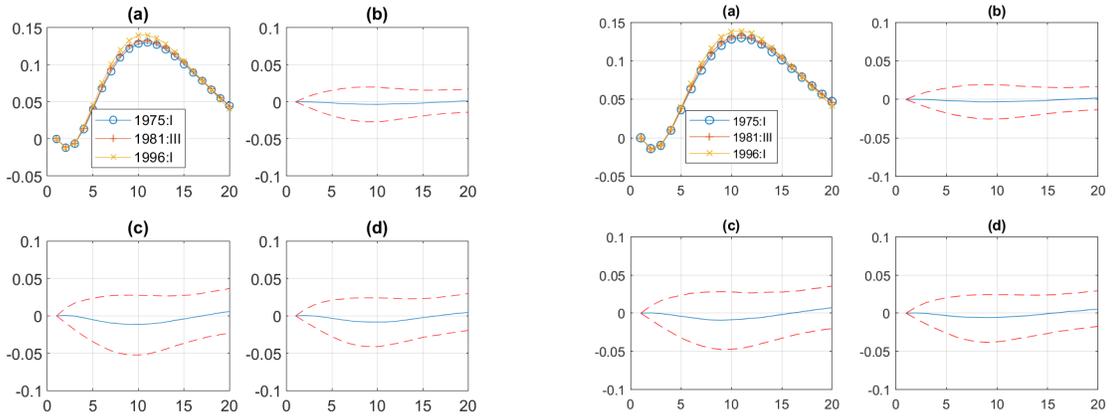
Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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 B.7: Replication Figure 3



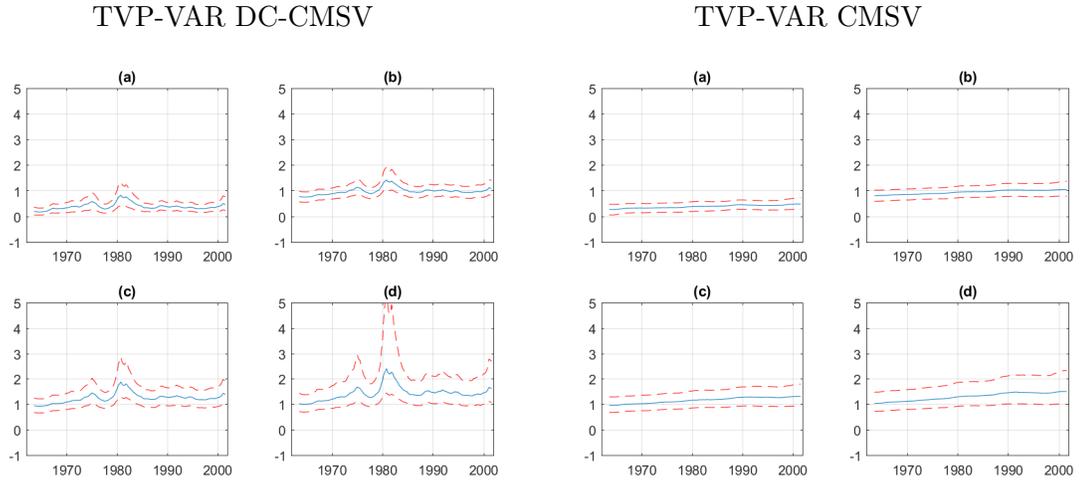
Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



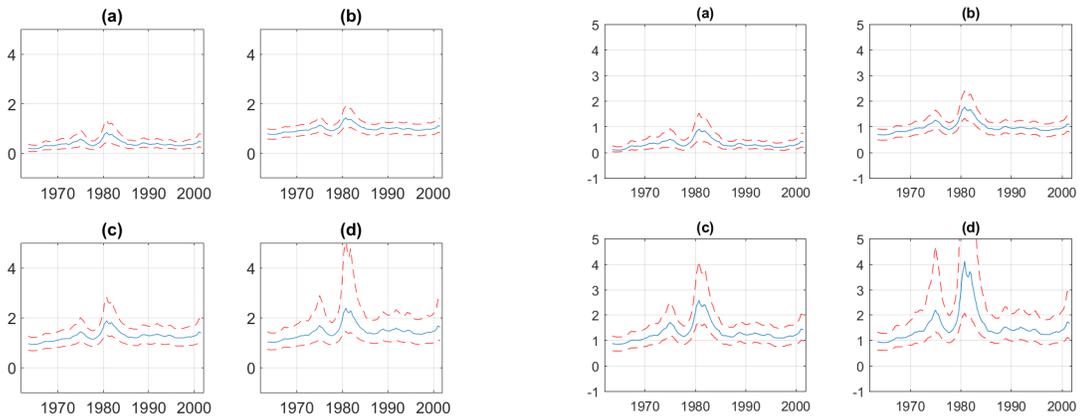
Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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 B.8: Replication Figure 4



Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$

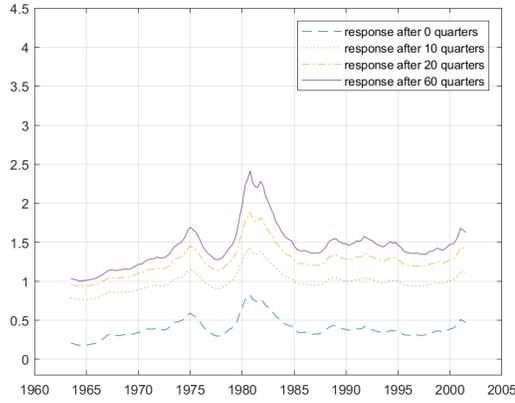


Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

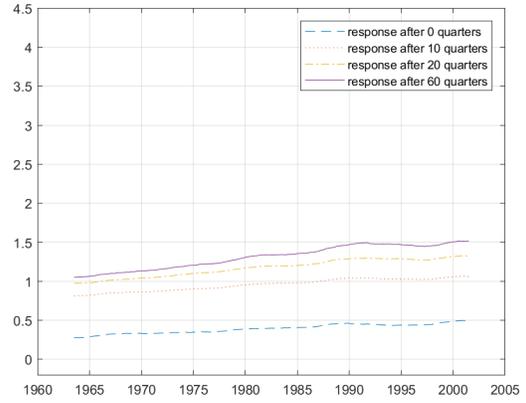
The figure depicts interest rate response to a 1% permanent increase of inflation with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

Figure B.9: Replication Figure 5

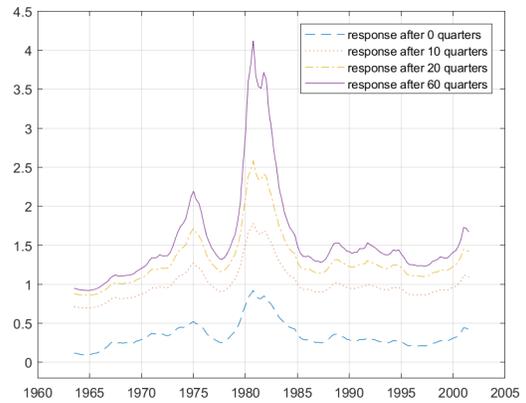
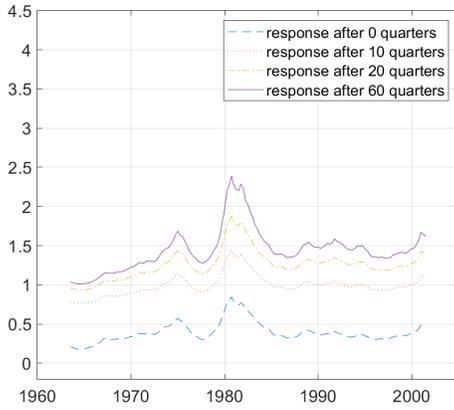
TVP-VAR DC-CMSV



TVP-VAR CMSV



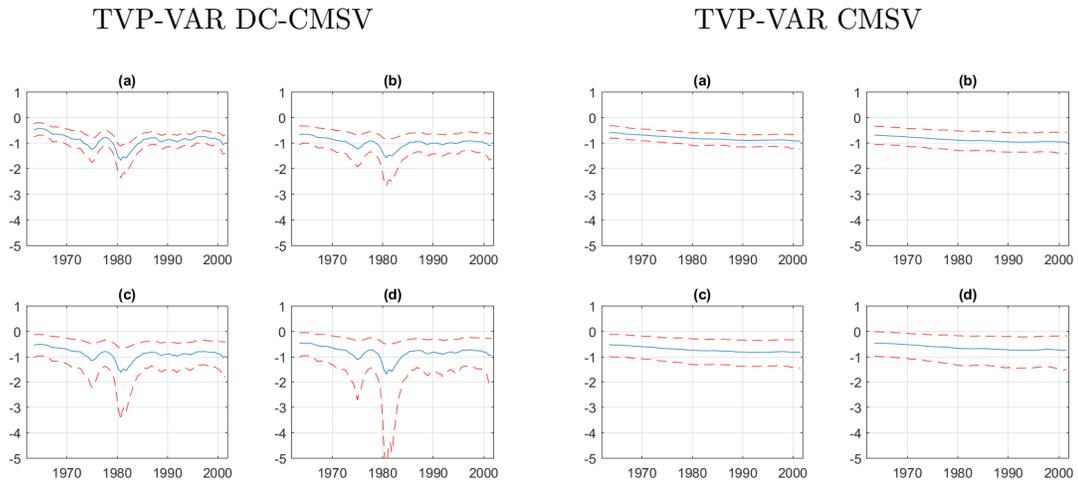
Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



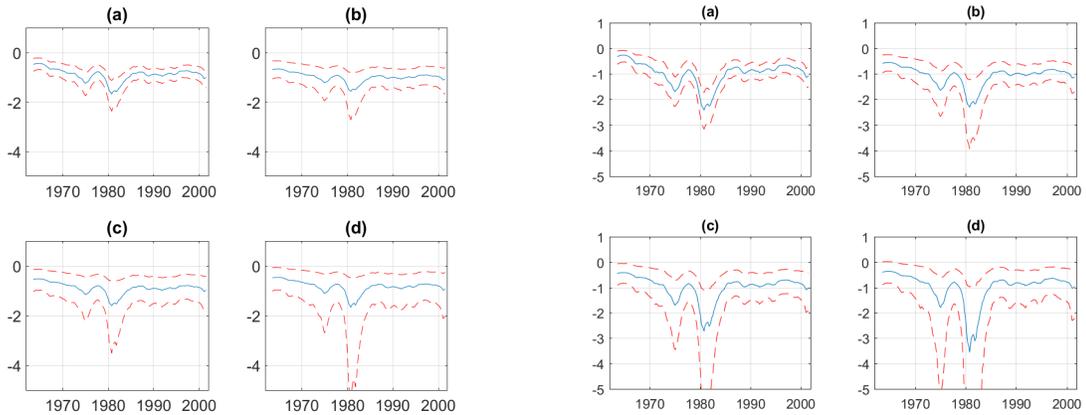
Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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

Figure B.10: Replication Figure 6



Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$

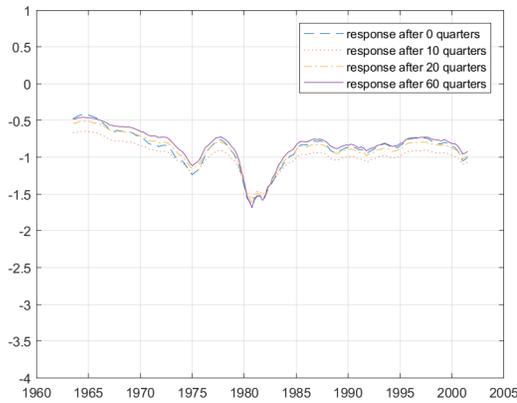


Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

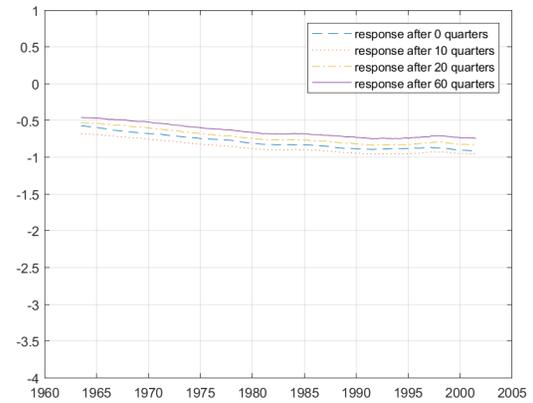
The figure depicts interest rate response to a 1% permanent increase of unemployment rate with 16th and 84th percentiles. (a) Simultaneous response, (b) response after 10 quarters, (c) response after 20 quarters, (d) response after 60 quarters.

Figure B.11: Replication Figure 7

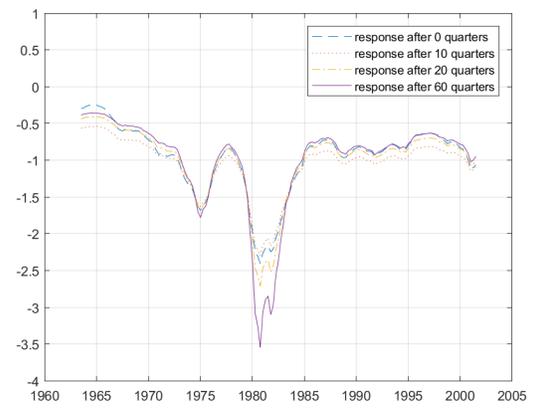
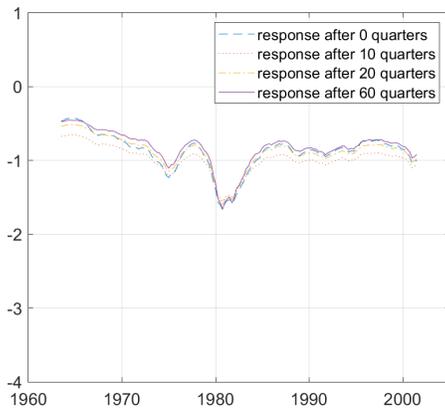
TVP-VAR DC-CMSV



TVP-VAR CMSV



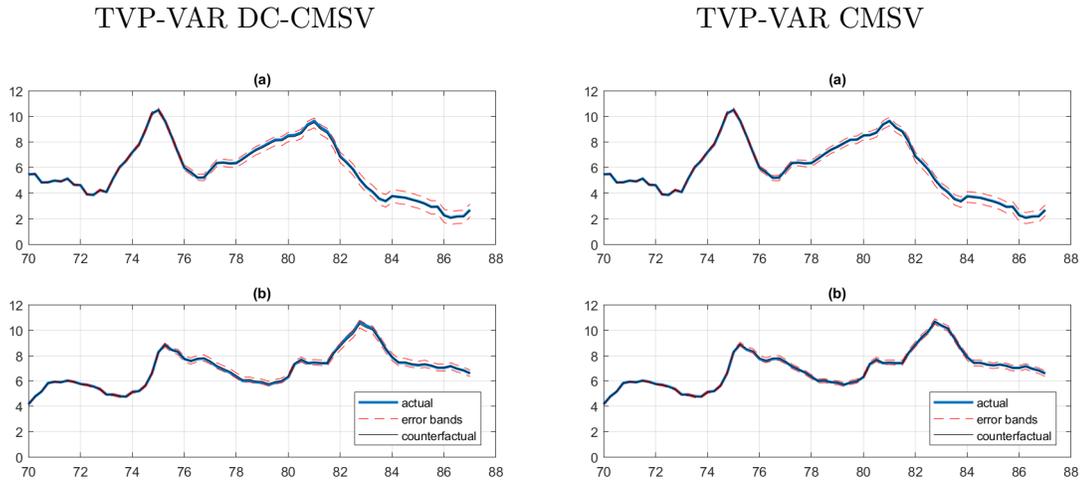
Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



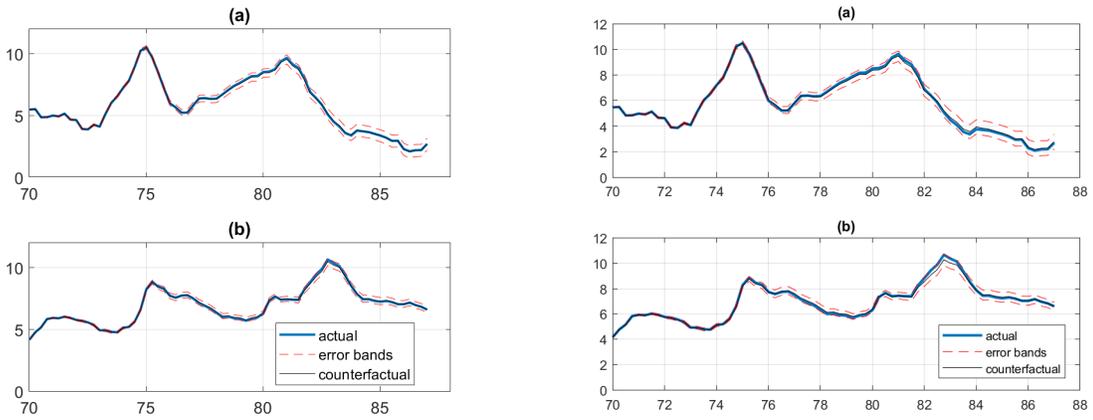
Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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

Figure B.12: Replication Figure 8



Panel (I): Original ordering,  $y_t = [\pi_t, u_t, i_t]$



Panel (II): Reverse ordering,  $y_t = [i_t, u_t, \pi_t]$

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

# C Tables

## C.1 Robustness: estimated hyperparameters

Table C.1: Estimated correlation

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.041	<b>0.030</b>	0.045	0.029	<b>0.023</b>	0.038	<b>0.029</b>	<b>0.371</b>	-0.219
sine	0.098	<b>0.086</b>	0.150	0.064	<b>0.018</b>	0.090	<b>0.023</b>	0.534	<b>0.942</b>
fastsine	<b>0.245</b>	0.256	0.256	0.087	<b>0.008</b>	0.115	<b>0.010</b>	0.390	<b>0.931</b>
step	0.081	<b>0.062</b>	0.083	0.045	<b>0.016</b>	0.068	<b>0.021</b>	0.493	<b>0.795</b>
ramp	0.119	<b>0.110</b>	0.168	0.068	<b>0.020</b>	0.098	<b>0.027</b>	0.546	<b>0.942</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.061	<b>0.039</b>	0.072	0.048	<b>0.025</b>	0.076	<b>0.031</b>	<b>0.308</b>	-0.216
sine	0.119	<b>0.089</b>	0.155	0.106	<b>0.015</b>	0.146	<b>0.020</b>	0.275	<b>0.945</b>
fastsine	<b>0.244</b>	0.257	0.258	0.142	<b>0.006</b>	0.185	<b>0.008</b>	0.176	<b>0.936</b>
step	0.105	<b>0.068</b>	0.095	0.080	<b>0.015</b>	0.121	<b>0.019</b>	0.287	<b>0.801</b>
ramp	0.139	<b>0.113</b>	0.179	0.111	<b>0.018</b>	0.154	<b>0.023</b>	0.296	<b>0.944</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	<b>0.024</b>	0.029	0.024	<b>0.021</b>	0.031	<b>0.026</b>	<b>0.296</b>	-0.181
sine	0.087	<b>0.083</b>	0.147	0.043	<b>0.020</b>	0.061	<b>0.026</b>	0.748	<b>0.942</b>
fastsine	<b>0.251</b>	0.256	0.256	0.051	<b>0.010</b>	0.067	<b>0.012</b>	0.587	<b>0.926</b>
step	0.066	<b>0.057</b>	0.076	0.029	<b>0.015</b>	0.042	<b>0.019</b>	0.671	<b>0.826</b>
ramp	0.111	<b>0.108</b>	0.166	0.046	<b>0.023</b>	0.068	<b>0.030</b>	0.742	<b>0.942</b>

The table shows the statistics for the estimated correlation of the MSV model with estimated hyperparameters. Accuracy metric is MAE, discrepancy metrics are MAD and RMSD, and similarity metric is FD. A bold figure highlights the best model in each panel and row.

Table C.2: Estimated covariance

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.314	<b>0.288</b>	0.374	0.155	<b>0.030</b>	0.224	<b>0.042</b>	0.445	<b>0.976</b>
sine	0.215	<b>0.200</b>	0.304	0.102	<b>0.021</b>	0.154	<b>0.030</b>	0.666	<b>0.983</b>
fastsine	<b>0.330</b>	0.331	0.350	0.112	<b>0.010</b>	0.170	<b>0.014</b>	0.607	<b>0.993</b>
step	0.251	<b>0.225</b>	0.303	0.107	<b>0.020</b>	0.165	<b>0.028</b>	0.591	<b>0.985</b>
ramp	0.235	<b>0.222</b>	0.315	0.105	<b>0.024</b>	0.161	<b>0.035</b>	0.636	<b>0.979</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.438	<b>0.393</b>	0.548	0.202	<b>0.038</b>	0.344	<b>0.059</b>	0.486	<b>0.987</b>
sine	0.292	<b>0.255</b>	0.406	0.177	<b>0.021</b>	0.308	<b>0.033</b>	0.566	<b>0.993</b>
fastsine	0.401	<b>0.395</b>	0.440	0.207	<b>0.009</b>	0.353	<b>0.014</b>	0.472	<b>0.998</b>
step	0.353	<b>0.303</b>	0.433	0.169	<b>0.021</b>	0.303	<b>0.034</b>	0.557	<b>0.994</b>
ramp	0.316	<b>0.283</b>	0.421	0.182	<b>0.024</b>	0.315	<b>0.039</b>	0.544	<b>0.991</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.238	<b>0.219</b>	0.259	0.134	<b>0.024</b>	0.177	<b>0.032</b>	0.349	<b>0.956</b>
sine	0.172	<b>0.165</b>	0.240	0.077	<b>0.022</b>	0.109	<b>0.029</b>	0.689	<b>0.967</b>
fastsine	<b>0.295</b>	0.296	0.301	0.062	<b>0.011</b>	0.085	<b>0.014</b>	0.738	<b>0.983</b>
step	0.192	<b>0.178</b>	0.224	0.082	<b>0.017</b>	0.119	<b>0.023</b>	0.574	<b>0.970</b>
ramp	0.192	<b>0.185</b>	0.251	0.082	<b>0.024</b>	0.119	<b>0.034</b>	0.649	<b>0.961</b>

The table shows the statistics for the estimated covariance of the MSV model with estimated hyperparameters. A bold figure highlights the best model in each panel and row.

Table C.3: Estimated value-at-risk

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.252	<b>0.237</b>	0.306	0.105	<b>0.011</b>	0.137	<b>0.014</b>	0.570	<b>0.994</b>
sine	0.213	<b>0.207</b>	0.281	0.071	<b>0.009</b>	0.096	<b>0.012</b>	0.826	<b>0.995</b>
fastsine	0.236	<b>0.229</b>	0.284	0.073	<b>0.005</b>	0.098	<b>0.007</b>	0.813	<b>0.997</b>
step	0.227	<b>0.217</b>	0.285	0.074	<b>0.008</b>	0.102	<b>0.011</b>	0.748	<b>0.995</b>
ramp	0.218	<b>0.213</b>	0.286	0.075	<b>0.011</b>	0.105	<b>0.014</b>	0.766	<b>0.994</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.330	<b>0.308</b>	0.427	0.129	<b>0.012</b>	0.180	<b>0.016</b>	0.682	<b>0.997</b>
sine	0.282	<b>0.271</b>	0.390	0.111	<b>0.009</b>	0.156	<b>0.012</b>	0.849	<b>0.998</b>
fastsine	0.300	<b>0.289</b>	0.390	0.121	<b>0.006</b>	0.164	<b>0.008</b>	0.827	<b>0.998</b>
step	0.301	<b>0.285</b>	0.400	0.108	<b>0.009</b>	0.156	<b>0.011</b>	0.799	<b>0.998</b>
ramp	0.286	<b>0.276</b>	0.394	0.113	<b>0.010</b>	0.161	<b>0.013</b>	0.817	<b>0.998</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	IDCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.198	<b>0.186</b>	0.222	0.100	<b>0.009</b>	0.126	<b>0.012</b>	0.428	<b>0.987</b>
sine	0.169	<b>0.165</b>	0.211	0.062	<b>0.010</b>	0.083	<b>0.013</b>	0.731	<b>0.987</b>
fastsine	0.196	<b>0.193</b>	0.218	0.044	<b>0.006</b>	0.058	<b>0.007</b>	0.834	<b>0.993</b>
step	0.178	<b>0.172</b>	0.209	0.062	<b>0.008</b>	0.086	<b>0.010</b>	0.672	<b>0.990</b>
ramp	0.175	<b>0.171</b>	0.215	0.072	<b>0.012</b>	0.096	<b>0.016</b>	0.627	<b>0.982</b>

The table shows the statistics for the estimated value-at-risk of the MSV model with estimated hyperparameters. A bold figure highlights the best model in each panel and row.

## C.2 Robustness: stationary state dynamics

Table C.4: Estimated correlation

(a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.041	<b>0.028</b>	0.049	<b>0.028</b>	0.029	0.038	<b>0.036</b>	<b>0.437</b>	-0.370
sine	0.101	<b>0.092</b>	0.141	0.065	<b>0.021</b>	0.092	<b>0.027</b>	0.508	<b>0.901</b>
fastsine	<b>0.224</b>	0.229	0.231	0.118	<b>0.046</b>	0.157	<b>0.058</b>	0.470	<b>0.694</b>
step	0.081	<b>0.063</b>	0.088	0.047	<b>0.022</b>	0.069	<b>0.028</b>	0.486	<b>0.700</b>
ramp	0.121	<b>0.114</b>	0.159	0.070	<b>0.027</b>	0.100	<b>0.035</b>	0.526	<b>0.875</b>

(b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.063	<b>0.035</b>	0.079	0.048	<b>0.033</b>	0.078	<b>0.042</b>	<b>0.348</b>	-0.345
sine	0.122	<b>0.095</b>	0.147	0.111	<b>0.019</b>	0.151	<b>0.024</b>	0.249	<b>0.903</b>
fastsine	0.235	<b>0.232</b>	0.234	0.165	<b>0.041</b>	0.215	<b>0.052</b>	0.226	<b>0.702</b>
step	0.106	<b>0.069</b>	0.106	0.084	<b>0.022</b>	0.125	<b>0.028</b>	0.281	<b>0.694</b>
ramp	0.141	<b>0.118</b>	0.168	0.116	<b>0.025</b>	0.160	<b>0.032</b>	0.275	<b>0.877</b>

(c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.030	<b>0.023</b>	0.031	0.025	<b>0.024</b>	0.032	<b>0.031</b>	<b>0.364</b>	-0.368
sine	0.090	<b>0.088</b>	0.137	0.043	<b>0.023</b>	0.060	<b>0.029</b>	0.730	<b>0.905</b>
fastsine	<b>0.215</b>	0.226	0.230	0.092	<b>0.052</b>	0.120	<b>0.065</b>	0.650	<b>0.688</b>
step	0.065	<b>0.058</b>	0.077	0.031	<b>0.020</b>	0.043	<b>0.026</b>	0.656	<b>0.740</b>
ramp	0.113	<b>0.111</b>	0.155	0.046	<b>0.029</b>	0.068	<b>0.038</b>	0.725	<b>0.878</b>

The table shows the statistics for the estimated correlation of the MSV model with stationary state dynamics for volatility and the time-varying parameter. A bold figure highlights the best model in each panel and row.

Table C.5: Estimated covariance

## (a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.314	<b>0.286</b>	0.380	0.152	<b>0.038</b>	0.217	<b>0.053</b>	0.519	<b>0.934</b>
sine	0.215	<b>0.201</b>	0.294	0.101	<b>0.024</b>	0.152	<b>0.034</b>	0.685	<b>0.977</b>
fastsine	0.311	<b>0.303</b>	0.345	0.158	<b>0.052</b>	0.243	<b>0.076</b>	0.517	<b>0.804</b>
step	0.249	<b>0.224</b>	0.303	0.107	<b>0.027</b>	0.162	<b>0.037</b>	0.639	<b>0.977</b>
ramp	0.235	<b>0.222</b>	0.307	0.105	<b>0.032</b>	0.160	<b>0.045</b>	0.650	<b>0.957</b>

## (b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.440	<b>0.394</b>	0.562	0.195	<b>0.051</b>	0.331	<b>0.082</b>	0.547	<b>0.951</b>
sine	0.293	<b>0.257</b>	0.394	0.181	<b>0.025</b>	0.312	<b>0.040</b>	0.573	<b>0.990</b>
fastsine	0.394	<b>0.367</b>	0.446	0.236	<b>0.053</b>	0.408	<b>0.088</b>	0.437	<b>0.874</b>
step	0.353	<b>0.303</b>	0.433	0.170	<b>0.031</b>	0.301	<b>0.049</b>	0.590	<b>0.988</b>
ramp	0.317	<b>0.284</b>	0.412	0.185	<b>0.033</b>	0.319	<b>0.053</b>	0.551	<b>0.980</b>

## (c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.236	<b>0.216</b>	0.263	0.139	<b>0.029</b>	0.183	<b>0.038</b>	0.414	<b>0.900</b>
sine	0.171	<b>0.164</b>	0.230	0.076	<b>0.024</b>	0.105	<b>0.032</b>	0.724	<b>0.956</b>
fastsine	<b>0.264</b>	0.266	0.288	0.125	<b>0.053</b>	0.174	<b>0.072</b>	0.555	<b>0.735</b>
step	0.190	<b>0.176</b>	0.222	0.088	<b>0.022</b>	0.124	<b>0.029</b>	0.626	<b>0.961</b>
ramp	0.191	<b>0.185</b>	0.242	0.083	<b>0.031</b>	0.117	<b>0.042</b>	0.671	<b>0.922</b>

The table shows the statistics for the estimated covariance of the MSV model with stationary state dynamics for volatility and the time-varying parameter. A bold figure highlights the best model in each panel and row.

Table C.6: Estimated value-at-risk

## (a) Benchmark DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.252	<b>0.236</b>	0.308	0.107	<b>0.013</b>	0.138	<b>0.017</b>	0.628	<b>0.986</b>
sine	0.212	<b>0.206</b>	0.280	0.072	<b>0.011</b>	0.098	<b>0.014</b>	0.841	<b>0.994</b>
fastsine	0.235	<b>0.224</b>	0.288	0.110	<b>0.023</b>	0.146	<b>0.030</b>	0.595	<b>0.921</b>
step	0.227	<b>0.217</b>	0.286	0.077	<b>0.011</b>	0.104	<b>0.014</b>	0.780	<b>0.994</b>
ramp	0.218	<b>0.212</b>	0.285	0.077	<b>0.014</b>	0.106	<b>0.018</b>	0.782	<b>0.987</b>

## (b) High Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.332	<b>0.309</b>	0.431	0.131	<b>0.015</b>	0.181	<b>0.020</b>	0.725	<b>0.992</b>
sine	0.283	<b>0.271</b>	0.390	0.115	<b>0.010</b>	0.160	<b>0.013</b>	0.856	<b>0.998</b>
fastsine	0.302	<b>0.285</b>	0.395	0.140	<b>0.021</b>	0.190	<b>0.028</b>	0.774	<b>0.975</b>
step	0.302	<b>0.286</b>	0.400	0.112	<b>0.011</b>	0.159	<b>0.015</b>	0.817	<b>0.997</b>
ramp	0.288	<b>0.276</b>	0.393	0.116	<b>0.013</b>	0.164	<b>0.017</b>	0.827	<b>0.997</b>

## (c) Low Volatility DGP

	MAE			MAD		RMSD		FD	
	CMSV	DC-CMSV	DCC	CMSV	DC-CMSV	CMSV	DC-CMSV	CMSV	DC-CMSV
const	0.197	<b>0.183</b>	0.223	0.105	<b>0.011</b>	0.133	<b>0.014</b>	0.480	<b>0.971</b>
sine	0.168	<b>0.163</b>	0.208	0.064	<b>0.011</b>	0.084	<b>0.015</b>	0.777	<b>0.985</b>
fastsine	0.191	<b>0.184</b>	0.219	0.100	<b>0.026</b>	0.128	<b>0.034</b>	0.406	<b>0.826</b>
step	0.178	<b>0.170</b>	0.209	0.070	<b>0.010</b>	0.094	<b>0.012</b>	0.709	<b>0.988</b>
ramp	0.175	<b>0.169</b>	0.213	0.075	<b>0.015</b>	0.099	<b>0.020</b>	0.657	<b>0.964</b>

The table shows the statistics for the estimated value-at-risk of the MSV model with stationary state dynamics for volatility and the time-varying parameter. A bold figure highlights the best model in each panel and row.