Appendix A Proofs

A.1 Proof of Lemma 1

Consider M positive definite $(d \times d)$ covariance matrices Ω_m, $m = 1, ..., M$ and suppose B is any invertible $(d \times d)$ matrix such that $B^{-1}\Omega_m B'^{-1}$ are diagonal matrices with strictly positive diagonal entries. It follows that $B^{-1}\Omega_1 B'^{-1} = \Lambda_m^{-1} B^{-1}\Omega_m B'^{-1}$ for some $(d \times d)$ diagonal matrices Λ_m^{-1}, $m = 2, ..., M$, that have strictly positive diagonal entries. Elementary matrix algebra then shows that these identities are equivalent to $B\Lambda_m = \Omega_m\Omega_1^{-1}B$, $m = 2, ..., M$. Thus, the matrices B and Λ_m solve the eigenvalue problem of $\Omega_m\Omega_1^{-1}$ with the diagonal of $\Lambda_m = \text{diag}(\lambda_{m1}, ..., \lambda_{md})$ containing the strictly positive eigenvalues and the columns of B being the related eigenvectors. Since this holds for any invertible $(d \times d)$ matrix B that simultaneously diagonalizes the covariance matrices, it is also a necessary property of a time-varying B-matrix B_t.

Suppose also BA solves the eigenvalue problems of $\Omega_m\Omega_1^{-1}$, $m = 2, ..., M$, for some invertible $(d \times d)$ matrix A. That is, $BA\Lambda_m = \Omega_m\Omega_1^{-1}BA$ which is equivalent to $A\Lambda_m A^{-1} = B^{-1}\Omega_m\Omega_1^{-1}B$. But since $B^{-1}\Omega_m\Omega_1^{-1}B = \Lambda_m$, this implies that $A\Lambda_m A^{-1} = \Lambda_m$, which is equivalent to $A\Lambda_m = \Lambda_m A$. Thus, $\lambda_{mi}a_{ij} = \lambda_{mj}a_{ij}$ where a_{ij} is the ijth element of A. It follows that $a_{ij} = 0$ if $\lambda_{mi} \neq \lambda_{mj}$ for some m, implying that A is diagonal matrix under Assumption 1, and BA multiplies each of the columns of B by a scalar. It is well known that eigenvalues of a matrix are unique (up to order), but since the diagonal elements of Λ_m can be in any order, so can the related eigenvectors that are the columns of B. That is, B is unique up to scalar multiples and ordering of its columns. Since the above holds for any appropriate B-matrices B and BA, it holds also for a time-varying B-matrix B_t at each t.■

A.2 Proof of Proposition 1

Lemma 1 shows that the B-matrix B_t is unique up to scalar multiples and reordering of its columns. Suppose the conditional covariance matrix of the structural error is normalized to a constant diagonal matrix with strictly positive diagonal entries, say C. That is, $\sum_{m=1}^{M} \alpha_{m,t} B_t^{-1} \Omega_m B_t^{\prime -1} = C$, which is equivalent to $\sum_{m=1}^{M} \alpha_{m,t} \Omega_m = B_t C B_t'$. Suppose that this identity also holds with another B-matrix, $B_t E_t$, where E_t is a possibly time-varying, invertible $(d \times d)$ matrix. We have $\sum_{m=1}^{M} \alpha_{m,t} (B_t E_t)^{-1} \Omega_m (B_t E_t)^{\prime -1} = C$, which is equivalent to $\sum_{m=1}^{M} \alpha_{m,t} \Omega_m = (B_t E_t) C (B_t E_t)'$. Thus, $B_t C B_t' = (B_t E_t) C (B_t E_t)'$. By Lemma 1, the B-matrix is unique up to scalar multiples and reordering of its columns, so with a given ordering of the columns, E_t is a diagonal matrix. It then follows from $B_t C B_t' = (B_t E_t) C (B_t E_t)'$ that $C = E_t C E_t$, which in turn implies $c_i = e_{t,i}^2 c_i$, where c_i and $e_{t,i}$ are the ith diagonal elements of C and E_t, respectively. Therefore, $e_{t,i} = \pm 1$, implying that with a given ordering of the columns, (for each t) B_t is unique up to changing all signs in a column. Therefore, B_t is unique up ordering of its columns and changing all signs in a column.■

A.3 Proof of Proposition 2

Let $\Omega_1, ..., \Omega_M$ be positive definite covariance matrices. We consider the decomposition $\Omega_1 = WW'$ and $\Omega_m = W \Lambda_m W', \ \ m = 2, ..., M$, where $\Lambda_m = \text{diag}(\lambda_{m1}, ..., \lambda_{md})$, $\lambda_{mi} > 0$ $(i = 1, .., d)$, contains the eigenvalues of $\Omega_m \Omega_1^{-1}$ in the diagonal and the columns of the nonsingular W are the related eigenvectors. The decomposition always exists when $M = 2$ (see, e.g., Muirhead, 1982, Theorem A9.9) but not necessarily when $M \geq 3$. In the following, we assume the covariance matrices satisfy the decomposition.

Repeating some of the proof in Lanne, Lütkepohl, and Maciejowska (2010, p. 130; see also the proof of Theorem A9.9 in Muirhead, 1982) for convenience, suppose that we also have $\Omega_1 = DD'$ and $\Omega_m = D \Lambda_m D'$, $m = 2, ..., M$, for some nonsingular $(d \times d)$ matrix D. Because $D^{-1} W W' D'^{-1} = D^{-1} \Omega_1 D'^{-1} = I_d$, the matrix $Q' \equiv D^{-1} W$ is orthogonal, and hence, $D = WQ$

and $\Lambda_m Q = Q\Lambda_m$. It follows that $\lambda_{mi} q_{ij} = \lambda_{mj} q_{ij}$ where q_{ij} is the ijth element of Q. Thus, $q_{ij} = 0$ if $\lambda_{mi} \neq \lambda_{mj}$ for some m. Assuming that this (Condition (1)) is satisfied by the last $d_1 \in \{1, ..., d\}$ eigenvalues, it follows that Q is a block-diagonal matrix with two blocks in the diagonal. Denoting $d_0 \equiv d - d_1$, the first block is a $(d_0 \times d_0)$ matrix and the second one is a $(d_1 \times d_1)$ diagonal matrix with $q_{d_0+1,d_0+1}, ..., q_{d,d}$ in the diagonal (if $d_1 = d$, Q simply reduces to a diagonal matrix).

As the blocks in the diagonal of an orthogonal block-diagonal matrix are orthogonal and the real eigenvalues of a diagonal orthogonal matrix are ± 1, it follows that the real eigenvalues of the second block in the diagonal of Q are ± 1. Then, because the eigenvalues of a block-diagonal matrix are the eigenvalues of the blocks in the diagonal, and eigenvalues of a diagonal matrix are its diagonal elements (and Q is real), it must be that $q_{d_0+1,d_0+1}, ..., q_{d,d}$ are ± 1.

Thus, because $D = WQ$, the last d_1 columns of W are unique up to changing all signs in a column for given Λ_m, $m = 2, ..., M$. Since Λ_m are unique up to ordering of the diagonal elements and Condition (2) fixes a unique ordering for the last d_1 columns of W and hence also for the related eigenvalues $\lambda_{mi}, i > d_0$, the last d_1 columns of the B-matrix (3.5) (in the main paper) are uniquely identified up to changing all signs in a column. Finally, Condition (3) fixes the signs in the last d_1 columns of W and consequently of B_t, implying that the last d_1 columns of the B-matrix are (globally) unique for given mixing weights $\alpha_{1,t}, ..., \alpha_{M,t}$. Moreover, if $d_1 = d$, the decomposition (3.4) (in the main paper) of $\Omega_1, ..., \Omega_M$ is (globally) unique.■

A.4 Proof of Proposition 3

Consider the matrix decomposition of Ω_m, $m = 1, ..., M$, of Proposition 1. It is shown in the proof of Proposition 1 that any $(d \times d)$ matrix D that also satisfies $\Omega_1 = DD'$ and $\Omega_m = D\Lambda_m D'$, $m = 2, ..., M$, can be presented as $D = WQ$ where Q is orthogonal and $q_{ij} = 0$ when $\lambda_{mi} \neq \lambda_{mj}$ for some m. Then, observe that the jth column of WQ is a linear combination of the columns of W, with the multiplier of the ith column given by q_{ij}. Denoting $d_0 \equiv d - d_1$, it follows that if

$\lambda_{mi} = \lambda_{mj}$ for $i \neq j > d_0$ and all m, but for all $l \notin \{i, j\}$, $\lambda_{ml} \neq \lambda_{mj}$ for some m, the jth column of WQ is a linear combination of the ith and jth columns of W. But if the jth column (of W and WQ) obeys a zero restriction where the ith column obeys a strict sign restriction (Condition (4)), the multiplier q_{ij} must be zero. That is, under the conditions of Proposition 3, with $j = d_0 + 1$ and $i < j$, we have $q_{l,d_0+1} = 0$ for all $l \neq d_0 + 1$ and $q_{lk} = 0$ for all $l, k = d_0 + 2, .., d$ such that $l \neq k$.

By the above discussion, when $d_1 > 1$, Q is a block-diagonal matrix with two blocks in the diagonal: the first one being a $(d_0 + 1 \times d_0 + 1)$ matrix

$$\tilde{Q} \equiv \begin{bmatrix} q_{1,1} & \cdots & q_{1,d_0} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ q_{d_0,1} & \cdots & q_{d_0,d_0} & 0 \\ q_{d_0+1,1} & \cdots & q_{d_0+1,d_0} & q_{d_0+1,d_0+1} \end{bmatrix} \tag{A.1}$$

and the second one a $(d_1 - 1 \times d_1 - 1)$ diagonal matrix with $q_{d_0+2,d_0+2}, ..., q_{d,d}$ in the diagonal. When $d_1 = 1$, we simply have $Q = \tilde{Q}$ where $\tilde{Q}$ is as in (A.1). Consequently, for $k > d_0$ the kth column of WQ equals to the kth column of W multiplied by $q_{k,k}$. It then remains to show that $q_{k,k} = \pm 1$ for all $k = d_0 + 1, ..., d$, after which global uniqueness of the last d_1 columns of the B-matrix can be concluded with arguments similar to the proof of Proposition 2.

Because only the last element of the last column of $\tilde{Q}$ is nonzero, the minors of the elements $q_{d_0+1,1}, ..., q_{d_0+1,d_0}$ are singular. Therefore, it follows from the cofactor presentation of the inverse of $\tilde{Q}$ (e.g., Muirhead, 1982, Appendices A4 and A5) that only the last element in the last column of the inverse of $\tilde{Q}$ is nonzero. Since $\tilde{Q}$ is orthogonal, as it is the upper-left block of the block-diagonal orthogonal matrix Q, its transpose is also its inverse. Hence, only the last element in the last column of the transpose of $\tilde{Q}$ is nonzero. Also, by the definition of $\tilde{Q}$, only the last element in last row of the transpose of $\tilde{Q}$ is nonzero. That is, the transpose is of the form

$$\tilde{Q}' = \begin{bmatrix} q_{1,1} & \cdots & q_{d_0,1} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ q_{1,d_0} & \cdots & q_{d_0,d_0} & 0 \\ 0 & \cdots & 0 & q_{d_0+1,d_0+1} \end{bmatrix}, \tag{A.2}$$

implying that

$$\tilde{Q} = \begin{bmatrix} q_{1,1} & \cdots & q_{1,d_0} & 0 \\ \vdots & \ddots & \vdots & \vdots \\ q_{d_0,1} & \cdots & q_{d_0,d_0} & 0 \\ 0 & \cdots & 0 & q_{d_0+1,d_0+1} \end{bmatrix}. \tag{A.3}$$

The matrix Q is therefore an orthogonal block-diagonal matrix with two blocks in the diagonal. The first block is the upper-left $(d_0 \times d_0)$ submatrix of $\tilde{Q}$ and the second block is the $(d_1 \times d_1)$ diagonal matrix with $q_{d_0+1,d_0+1}, ..., q_{d,d}$ in the diagonal. Now reasoning similar to the proof of Proposition 2 shows that $q_{k,k} = \pm 1$ for $k = d_0 + 1, ..., d$.■

Appendix B A Monte Carlo study

This section presents the results from the small scale Monte Carlo study studying the estimator's finite sample performance on estimating the structural parameters. Because estimation of our SVAR model can be tricky and computationally demanding, we assume the following simplistic setup. We consider two-dimensional SVAR models with two volatility regimes and autoregressive order $p = 1$. That is, the following model is employed:

$$y_t = \phi_0 + A_1 y_{t-1} + W(\alpha_{1,t} I_d + \alpha_{2,t} \Lambda_2)^{1/2} e_t, \tag{B.1}$$

where e_t (2×1) is the structural error (which has a mixture normal distribution), $\Lambda_2 = \text{diag}(\lambda_2)$ and $\alpha_{m,t}$, $m = 1, 2$ are the mixing weights defined in Equation (2.3) in the main paper.

We use the parameter values $\phi_0 = (0, 0)$, $\text{vec}(A_1) = (0.3, 0.1, 0.1, 0.3)$, and $\alpha_1 = 0.6$ throughout and consider three different sets of parameter values for $\text{vec}(W)$ and λ_2. In the first model (Model 1 in Table 1 below), we specify $\text{vec}(W) = (1.0, -0.5, 0.5, 0.80)$ and $\lambda_2 = (0.8, 1.8)$ as example of parameter values with nonzero elements of the impact matrix and eigenvalues $\lambda_{21}, \lambda_{22}$ that are not very close to each other. In the second model (Model 2 in Table 1), the upper-right element of W is set to zero. Other than that, parameter values identical to Model 1 are used in order to

study how the estimation accuracy changes when the impact matrix satisfies a recursive structure but it is not imposed in the estimation. Finally, in the third model (Model 3 in Table 1), we set $\lambda_2 = (1.5, 1.8)$ but use the same values in W as in Model 2 to study how the estimation accuracy changes if the eigenvalues are close to each other. Model 3 is interesting because the parameters are not identified if the eigenvalues are exactly equal to each other, unless (for instance) the recursive structure is imposed in the estimation.

For each specification of vec(W) and Λ_2, we simulated 500 samples of lengths 250, 500, 1000, and 2000 from the SVAR process. For each sample, we then estimated the reduced form VAR, $p = 1, M = 2$, model, decomposed the covariance matrices as in Equation (3.4) of the main paper, and compared the estimates to the true parameter values. Because the model identifies only up to the ordering of the regimes and ordering and signs of the columns of W, we compared the estimates obtained from each ordering and signs to the true parameter values. Then, we chose the ordering and signs that minimized the sum of the absolute values of the differences between the true parameter values and the estimates (which is a standard procedure in the literature, e.g., Gouriéroux, Monfort, and Renne, 2017).

The results from the Monte Carlo study are summarized in Table 1. The first column shows the true parameter values and rest of the four columns show the estimation results for the samples of length 250, 500, 1000, and 2000, respectively. In each of the four columns, the biases, i.e., the means of the estimates over the Monte Carlo repetitions minus the true parameter values are presented first, and next to the biases are the standard deviations of the estimates in parentheses. The results show that estimation accuracy is reasonable, although there is some finite sample bias. The estimates, nevertheless, become more accurate as the sample size increases. There is no noticeable difference in the estimation accuracy between Model 1 and Model 2, which is expected, as the zero impact response was not imposed in the estimation. Comparison to the results for Model 3, in turn, shows that having the true λ_{21} and λ_{22} close to each other seems to decrease the estimation accuracy. Particularly λ_{21} and λ_{22} are imprecisely estimated at small samples. The estimation accuracy of

$(\text{vec}(W), \lambda_2)$	$T = 250$	$T = 500$	$T = 1000$	$T = 2000$
Model 1				
1.00	-0.04 (0.28)	-0.03 (0.23)	-0.02 (0.20)	-0.01 (0.13)
-0.50	0.11 (0.42)	0.08 (0.36)	0.06 (0.30)	0.02 (0.20)
0.50	-0.10 (0.38)	-0.06 (0.34)	-0.06 (0.28)	-0.03 (0.19)
0.80	0.00 (0.25)	0.00 (0.22)	-0.01 (0.18)	-0.02 (0.13)
0.80	-0.14 (0.63)	-0.07 (0.57)	-0.07 (0.37)	-0.03 (0.32)
1.80	0.02 (1.25)	0.04 (1.08)	0.13 (0.83)	0.13 (0.66)
Model 2				
1.00	-0.10 (0.27)	-0.06 (0.19)	-0.03 (0.13)	-0.02 (0.12)
-0.50	0.01 (0.44)	0.01 (0.34)	0.02 (0.27)	0.02 (0.19)
0.00	-0.04 (0.38)	-0.01 (0.33)	-0.02 (0.25)	-0.01 (0.17)
0.80	-0.05 (0.26)	-0.03 (0.23)	-0.02 (0.17)	-0.02 (0.12)
0.80	-0.21 (0.59)	-0.15 (0.48)	-0.08 (0.35)	-0.04 (0.26)
1.80	0.05 (1.15)	0.10 (1.07)	0.16 (0.84)	0.15 (0.67)
Model 3				
1.00	-0.02 (0.19)	-0.04 (0.18)	-0.06 (0.16)	-0.08 (0.12)
-0.50	0.07 (0.45)	0.08 (0.41)	0.06 (0.39)	0.06 (0.33)
0.00	-0.06 (0.44)	-0.06 (0.43)	-0.04 (0.41)	-0.02 (0.37)
0.80	0.01 (0.26)	-0.02 (0.24)	-0.05 (0.25)	-0.06 (0.22)
1.50	-0.43 (0.97)	-0.31 (0.83)	-0.15 (0.67)	-0.03 (0.45)
1.80	-0.18 (1.24)	-0.02 (1.07)	0.11 (1.02)	0.14 (0.55)

Table 1: Results from the Monte Carlo study on the estimator's finite sample performance of estimating the structural parameters. The first column shows the true parameter values and rest of the four columns show the estimation results for samples of length 250, 500, 1000, and 2000, respectively. In each column, the biases, i.e., the means of the estimates over the 500 Monte Carlo repetitions minus the true parameter values are presented first, and next to the biases are the standard deviations of the estimates in parentheses.

the parameters in W decreases less, but the standard deviations of the estimates are relatively large also with the samples of length 2000.

Appendix C Details on the empirical application

C.1 Model selection and adequacy

The maximum likelihood (ML) estimation of the models, quantile residual diagnostics, estimation of impulse response functions are carried out with the CRAN distributed R package gmvarkit (Virolainen, 2018) that accompanies this paper. The R package gmvarkit also contains the dataset

studied in the empirical application to facilitate reproduction of our results. The estimation is based on the exact log-likelihood function.

Since the autoregressive dynamics are linear, we use residuals for studying remaining autocorrelation by employing the standard (adjusted) Portmanteau test. However, because the standardized residuals are not the empirical counterparts of the standardized reduced form shocks, we also employ the multivariate quantile residuals proposed by Kalliovirta and Saikkonen (2010) to further study the adequacy of our model (see also the related paper by Kalliovirta, 2012, for discussion on quantile residual based model diagnostics in a univariate setting). For a correctly specified model, the empirical counterparts of the quantile residuals are asymptotically independent with multivariate standard normal distributions and can hence be used for graphical analysis in a similar manner to the conventional standardized Pearson residuals (Kalliovirta and Saikkonen, 2010, Lemma 3).

We start by estimating linear Gaussian VARs with the autoregressive orders $p = 1, ..., 12$. BIC is minimized by the order $p = 1$, HQIC by the order $p = 2$, and AIC by the order $p = 3$, suggesting that the appropriate autoregressive order is likely relatively small. The (adjusted) Portmanteau test for remaining autocorrelation in the residuals taking into account 20 lags obtains the p-value 0.432, suggesting that the autocorrelation structure of the data is adequately captured by the linear VAR $p = 3$ model. In order to assess whether the residuals are heteroskedastic, we apply the Portmanteau test to the squared (standardized) residuals and obtain a very small p-value (0.00002). To investigate further, Figure 1 depicts the auto- and crosscorrelation functions of the squared (standardized) residuals, which clearly show that there is heteroskedasticity remaining in the residuals. The conclusion of clearly heteroskedastic residuals is not sensitive to the order p nor to using unstandardized residuals (the Portmanteau test applied for squared residuals obtains a p-value that is very small for orders $p = 1, 2, ..., 12$; graphical diagnostic figures are not shown for brevity for other orders than $p = 3$).

Hence, we estimate SVAR(p) models with two volatility regimes ($M = 2$) and $p = 1, ..., 4$. BIC, HQIC and AIC are by the order $p = 2$. The adequacy of model with autoregressive order $p = 2$ is,

Model	Log-lik	BIC	HQIC	AIC
$p = 1, M = 1$	-4.597	9.816	9.576	9.416
$p = 2, M = 1$	-4.435	9.824	9.457	9.211
$p = 3, M = 1$	-4.363	10.012	9.517	9.185
$p = 4, M = 1$	-4.312	10.242	9.620	9.202
$p = 1, M = 2$	-3.743	8.336	8.009	7.789
$p = 2, M = 2$	-3.540	8.262	7.808	7.503
$p = 3, M = 2$	-3.495	8.504	7.922	7.531
$p = 4, M = 2$	-3.495	8.836	8.127	7.650
$p = 4, M = 3$	-3.216	8.381	7.632	7.129

Table 2: The log-likelihoods and values of the information criteria divided by the number of observations for the discussed models.

however, rejected by the Portmanteau test (p-value 0.007). The Portmanteau test also rejects the adequacy of the model with autoregressive order $p = 3$ with the 5% level of significance (p-value 0.038). The model with autoregressive order $p = 4$, however, passes the Portmanteau test with the 5% level of significance (p-value 0.065), so it is preferred over the lower order model despite of the higher values of the information criteria (presented in Table 2 together with the log-likelihoods). In terms of information criteria, the SVAR(4) model with two volatility regimes is, nonetheless, clearly superior to the linear one-regime models.

In order to further study the adequacy of our two-regime SVAR(4) model, we examine the quantile residual time series, sample auto- and crosscorrelation functions of the quantile residuals and squared quantile residuals, and normal quantile-to-quantile plots. The sample auto- and crosscorrelation functions (presented in Figure 2) show that there is not much auto- or crosscorrelation in the quantile residuals. A moderate sized autocorrelation coefficients (ACC) at the lag 11 sticks out for the producer price index. There are also some other moderate sized coefficients, but given that in total of 316 correlation coefficients are presented, some of them are expected to be moderate sized for an IID process as well.

Since there is no remaining auto- or crosscorrelation in the quantile residuals, the auto- and crosscorrelation functions of the squared quantile residuals can be used to evaluate the model's adequacy to the capture the conditional heteroskedasticity in the series. The sample auto- and crosscorrelation functions of the squared quantile residuals are presented in Figure 2. The GDP, GDP deflator,

and PPI each have at least one large ACC in their autocorrelation functions, but ACCs of the interest rate variable are reasonable. There is a moderate large coefficient at lag 12 in the crosscorrelation function of the interest rate variable and GDP deflator. Notably, the very large ACC in the first lag of the ACF of the squared quantile residuals of GDP appears to be due to the extreme volatility caused by the COVID-19 crisis, as it disappears when the COVID-19 period is excluded from the quantile residuals (not shown). Our two-regime SVAR $p = 4$ model is therefore somewhat inadequate to capture the conditional heteroskedasticity in the series.

The quantile residual time series (the top panels of Figure 4) also show some heteroskedasticity and several outliers in the quantile residuals. There is a particularly large (marginal) quantile residual of the GDP in the beginning of the COVID-19 crisis, when the COVID-19 lockdown caused a fast and vast drop in the GDP growth. We do not view this large negative quantile residual of the GDP as an inadequacy, however, as the COVID-19 drop is known to be caused by an exceptionally large exogenous shock, and therefore a large (quantile) residual is expected for a correctly specified model. The normal quantile-quantile-plots (the bottom panels of Figure 4) show that the marginal quantile residual distributions have excess kurtosis but are quite symmetric. The quantile residuals of GDP deflator seem slightly skewed to the right and GDP slightly to the left, however.

In our view, the overall adequacy of the model is decent enough for further analysis, although some of the conditional heteroskedasticity of the series is not captured (largely due to the extreme volatility in the COVID-19 period). The model's capability to capture the conditional heteroskedasticity and marginal distribution of the series can be improved by adding a third regime, so we estimate a three-regime SVAR $p = 4$ model. Introducing the third volatility regime to the model turns out to improve the fit and decrease the values of the information criteria (presented in Table 2). Since the matrix decomposition (Equation (3.4) in the main paper) assumed for the reduced form error covariance matrices of the structural model does not necessarily exists for three-regime models, we test its existence using the likelihood ratio test. The restrictions required for the decomposition obtain the p-value 0.116 from the test and are not thereby rejected. The auto- and crosscorrela-

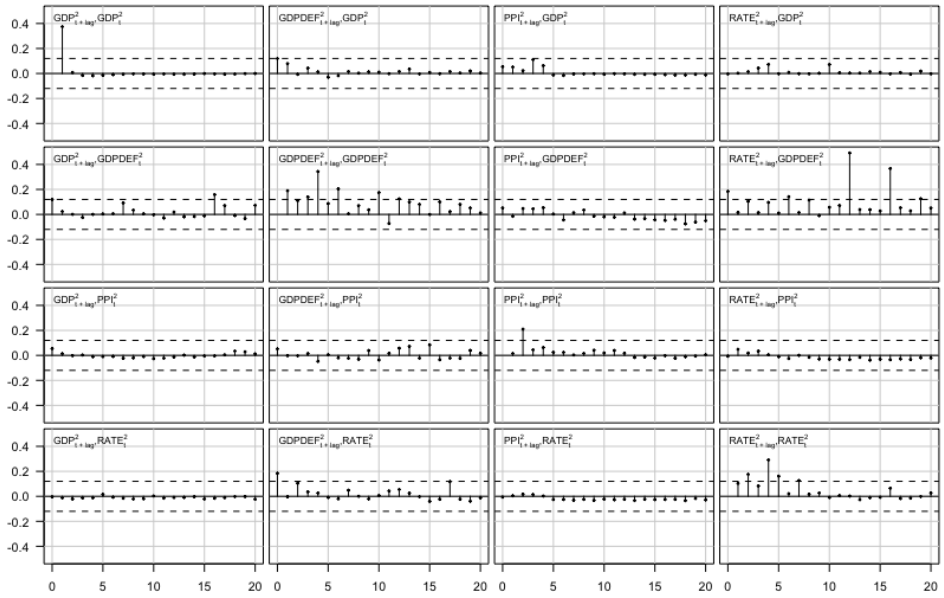

Figure 1: Auto- and crosscorrelation functions of the squared standardized residuals of the fitted linear Gaussian VAR $p = 3$ model for the lags $0, 1, ..., 20$. The lag zero autocorrelation coefficients are omitted, as they are one by convention. The blue dashed lines are the 95% bounds $\pm 1.96/\sqrt{T}$ ($T = 267$ as the first $p = 3$ observations were used as the initial values) for autocorrelations of IID observations.

tion functions of the squared quantile residuals, depicted in Figure 5, however, show that some of the conditional heteroskedasticity is still not captured by the three-regime model. The adequacy of the three-regime model is also rejected by the Portmanteau test for remaining autocorrelation (p-value 0.0005). Moreover, since the three-regime model also induces a long-run price puzzle to the impulse response functions (presented in Figure 6, dot-dashed line with plus signs), the more parsimonious two-regime model of order $p = 4$ is preferred as the main specification.

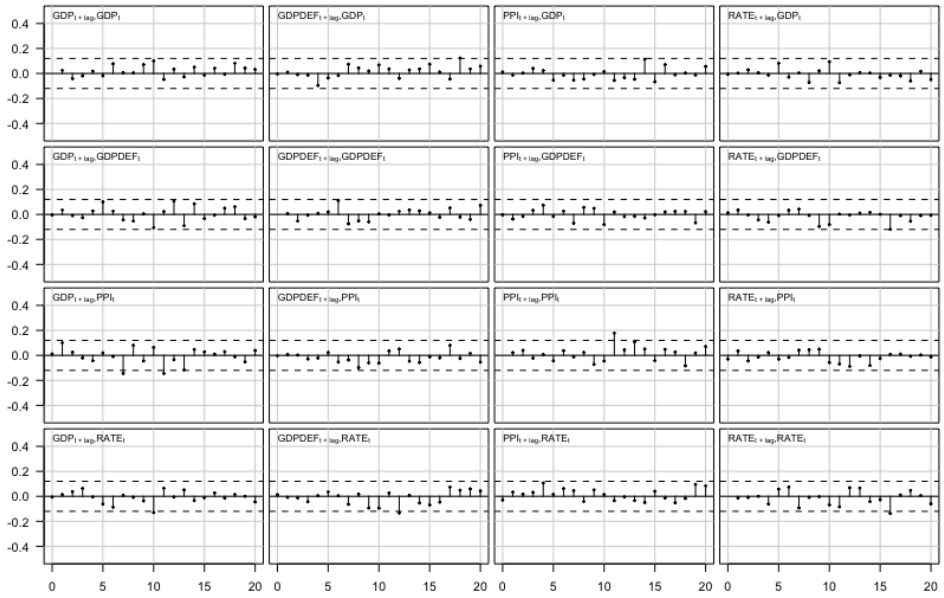

Figure 2: Auto- and crosscorrelation functions of the quantile residuals of the fitted two-regime SVAR $p = 4$ model for the lags $0, 1, ..., 20$. The lag zero autocorrelation coefficients are omitted, as they are one by convention. The blue dashed lines are the 95% bounds $\pm 1.96/\sqrt{T}$ ($T = 266$ as the first $p = 4$ observations were used as the initial values) for autocorrelations of IID observations.

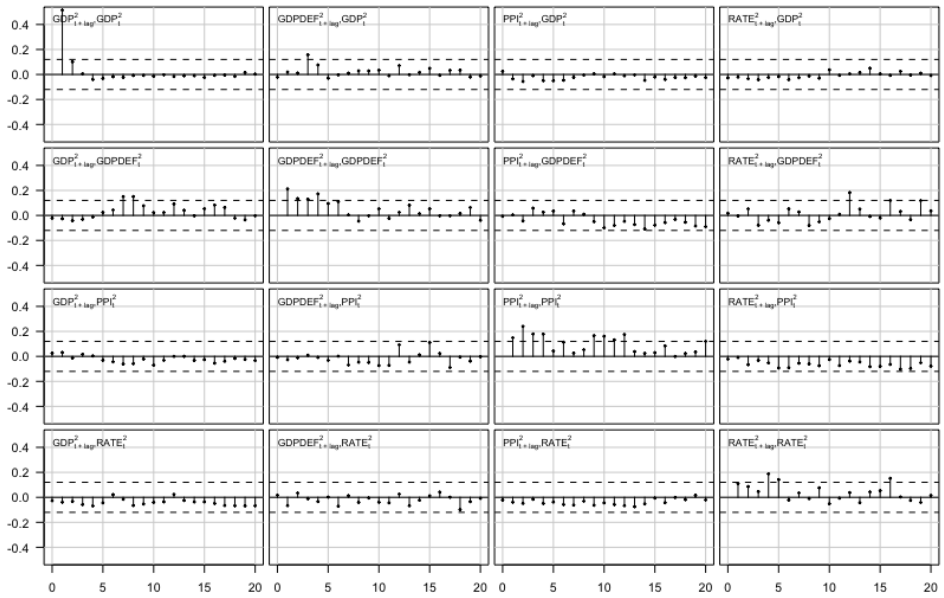

Figure 3: Auto- and crosscorrelation functions of the squared quantile residuals of the fitted two-regime SVAR $p = 4$ model for the lags $0, 1, ..., 20$. The lag zero autocorrelation coefficients are omitted, as they are one by convention. The blue dashed lines are the 95% bounds $\pm 1.96/\sqrt{T}$ ($T = 266$ as the first $p = 4$ observations were used as the initial values) for autocorrelations of IID observations.

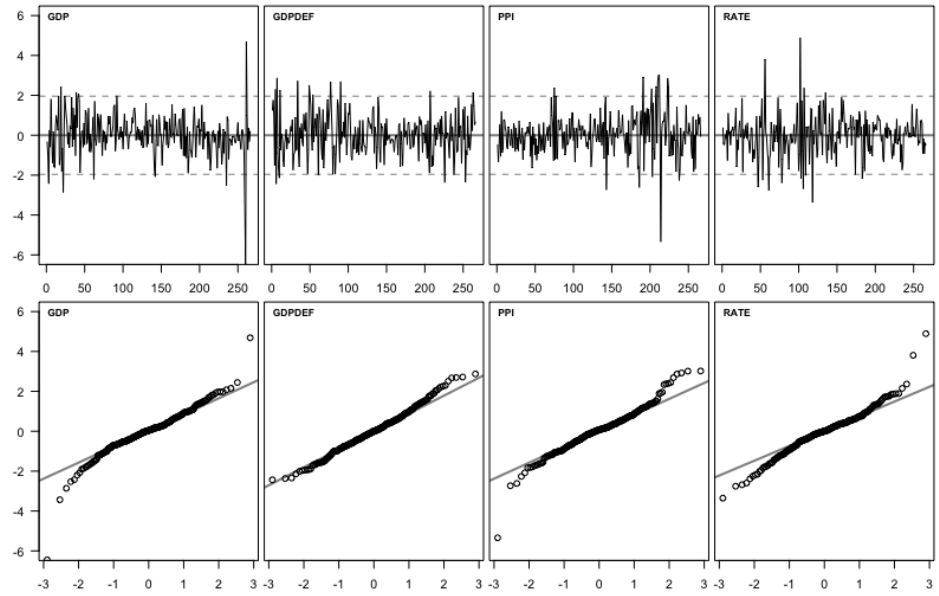

Figure 4: Quantile residual time series and normal quantile-quantile-plots of the fitted two-regime SVAR $p = 4$ model.

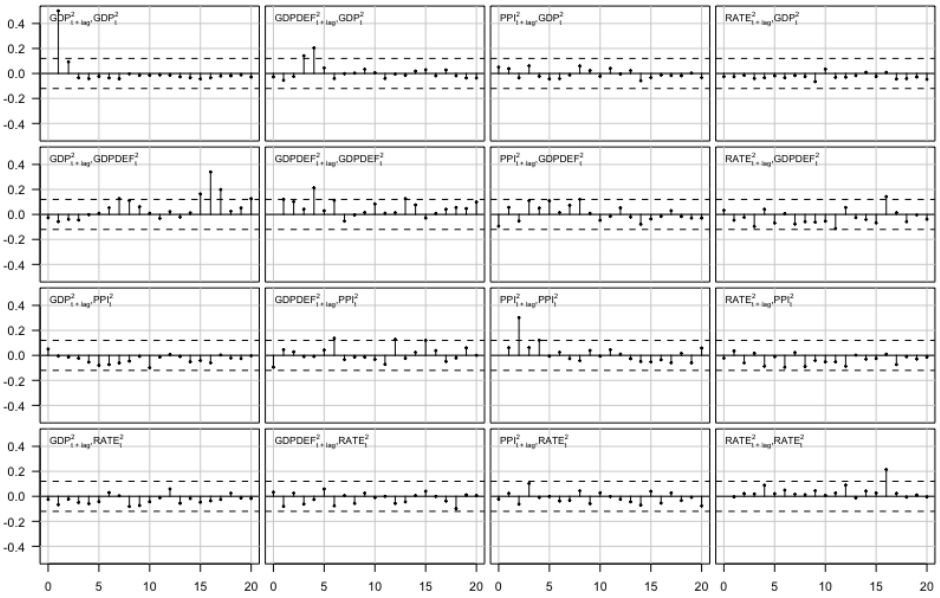

Figure 5: Auto- and crosscorrelation functions of the squared quantile residuals of the fitted three-regime SVAR $p = 4$ model for the lags $0, 1, ..., 20$. The lag zero autocorrelation coefficients are omitted, as they are one by convention. The blue dashed lines are the 95% bounds $\pm 1.96/\sqrt{T}$ ($T = 266$ as the first $p = 4$ observations were used as the initial values) for autocorrelations of IID observations.

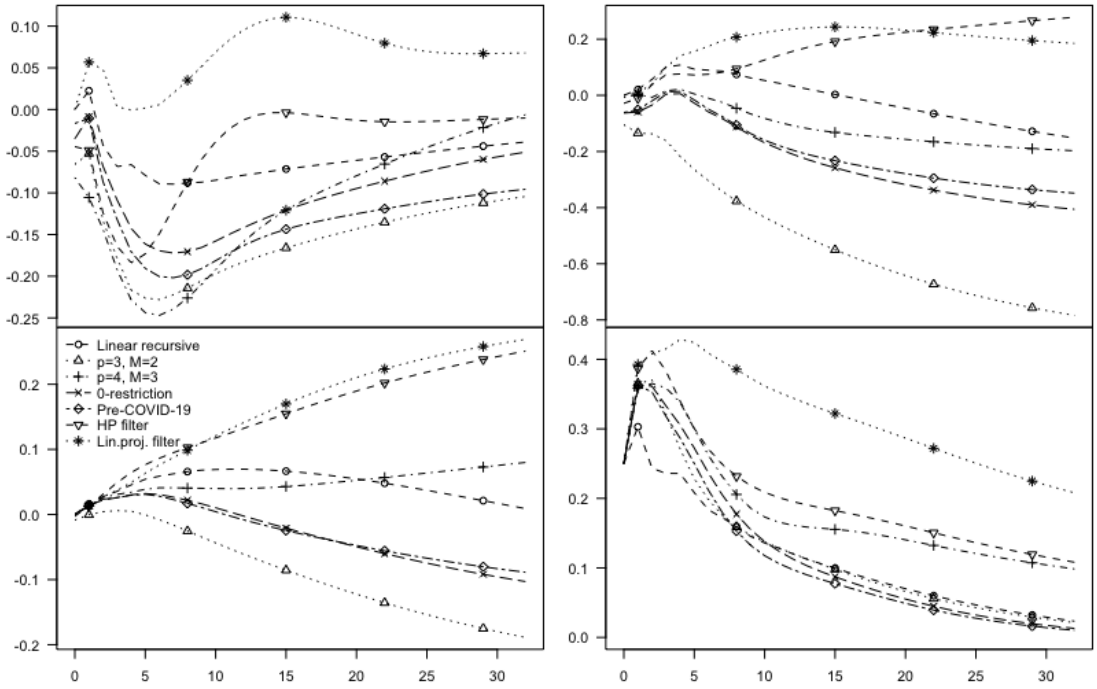

Figure 6: The IRFs of the identified monetary policy shock for the alternative specifications considered. The IRFs of GDP (top left), GDP deflator (bottom left), and produced price index (top right) are in $100\times$log-levels, whereas the IRFs of the interest rate variable (bottom right) is presented in the original scale.

C.2 Robustness checks

As robustness checks, we consider several alternative model specifications. In each of the model specifications, the monetary policy shock is identified as the shock that significantly moves the interest rate variable. In the presence of several alternative such shocks, the one that moves output or prices to the opposite direction from the interest rate variable deemed as the monetary policy shock. In each case where the unrestricted estimate of the impact response of output or prices is positive in response to a contractionary monetary policy shock, a zero restriction is imposed on the impact response. The impulse response functions estimated for the alternative model specifications are presented in Figure 6. The IRFs are scaled to correspond to a 25 basis point increase of the interest rate variable. The IRFs of the GDP (top left), GDP deflator (bottom left), and produced price index (top right) are presented in $100\times$log-levels, whereas the IRFs of the interest rate variable (bottom right) is presented in the original scale.

First, we consider the standard recursively identified linear Gaussian SVAR model with the monetary policy shock ordered last (dashed line with circles). The same autoregressive order $p = 4$ as in our benchmark model is preferred over the order $p = 3$ suggested by AIC, because the linear model with $p = 3$ induces a long-run price puzzle, while the $p = 4$ model produces only a medium-run price puzzle (prices finally decrease 9 years after the impact).

To investigate whether the results are sensitive to the choice of p in our SVAR with two volatility regimes, we estimate IRFs for the model with the autoregressive order $p = 3$ (dotted line with up-pointing triangles). The IRFs are very similar to our benchmark specification, but the response of the GDP deflator is slightly positive from the second to the fourth quarter after the impact. To study how robust our findings are to the number of volatility regimes, we estimate the IRFs for our SVAR model with three volatility regimes (dot-dashed line with plus signs), and we find a significant decrease in output but also a long-run price puzzle.

We check how the IRFs change when our identification is strengthened by placing a zero restriction on the instantaneous response of the GDP deflator to the monetary policy shock (dashed line with x symbols). Assuming the model is identified also without the zero restriction, it obtains the p-value 0.428 from a likelihood ratio test, so the restriction is not rejected. The IRFs are quite similar to our benchmark specification with the exception that the zero constraint seems to induce a short-term price puzzle. We also check whether our results are robust to excluding the COVID-19 period and fit our two-regime SVAR model to the sub-sample that ends in 2019Q4 (dot-dashed line with tilted squares). The resulting IRFs are otherwise quite similar to the benchmark specification except that there is a short-term price puzzle.

In order to see how robust the results are to using alternative methods for detrending the log of the GDP, we consider the backward-looking Hodrick-Prescott (HP) filter with the standard smoothing parameter value of 1600 (dashed line with down-pointing triangles) and the linear projection filter proposed by Hamilton (2018) (dotted line with stars). In both of these specifications, there is a severe long-run price puzzle. The specification employing the linear projection filter also displays

Figure 7: The data and time series of the identified monetary policy shocks covering the period from 1979Q3 to 1982Q4.

a severe production puzzle, as the GDP increases in response to a supposedly contractionary monetary policy shock. With the HP filter, the response of the GDP is negative and humped shaped but relatively short-lived.

Finally, we study whether the monetary policy shock behaves consistently to its real history around the Volcker recession in the early 1980's. Since our model assumes linear autoregressive dynamics, the reduced form errors and thereby also structural shocks can be recovered from the fitted model. Figure 7 presents the time series of the identified monetary policy shock together with the four variables of the data for the period from 1979Q3 to 1982Q4. The figure illustrates that prior to the Volcker recession in 1979Q3 and 1979Q4 the monetary policy shocks were contractionary. During the interest rate cuts in 1980 there were expansionary monetary policy shocks, which were followed by a large contractionary monetary policy shock when the interest rate was vastly increased again. Throughout the remainder of the Volcker recession, there were alternating expansionary and contractionary monetary policy shocks.

References

Gouriéroux C., Monfort A., Renne J.-P. (2017). "Statistical inference for independent component analysis: Application to structural VAR models." *Journal of Econometrics*, **196**(1), 111–126.

Hamilton J. D. (2018). "WHY YOU SHOULD NEVER USE THE HODRICK-PRESCOTT FILTER." *The Review of Economics and Statistics*, **100**(5), 831–843.

Kalliovirta L. (2012). "Misspecification tests based on quantile residuals." *The Econometrics Journal*, **15**(2), 358–393.

Kalliovirta L., Saikkonen P. (2010). "Reliable Residuals for Multivariate Nonlinear Time Series Models." *Unpublished revision of HECER discussion paper No. 247*.

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

Muirhead R. (1982). *Aspects of Multivariate Statistical Theory*. 1st edition. John Wiley & Sons, Hoboken, New Jersey.

Virolainen S. (2018). *gmvarkit: Estimate Gaussian and Student's t Mixture Vector Autoregressive Models*. R package version 2.1.0 available at CRAN: `https://CRAN.R-project.org/package=gmvarkit`.