2019-08

Working paper. Economics

ISSN 2340-5031

CO-INTEGRATION AND COMMON TRENDS ANALYSIS WITH SCORE-DRIVEN MODELS: AN APPLICATION TO THE FEDERAL FUNDS EFFECTIVE RATE AND US INFLATION RATE

Szabolcs Blazsek[a], Alvaro Escribano[b], and Adrian Licht[a]

[a]School of Business, Universidad Francisco Marroquín, Guatemala City, Guatemala
[b]Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain

uc3m | Universidad Carlos III de Madrid

Serie disponible en http://hdl.handle.net/10016/11

Web: http://economia.uc3m.es/

Correo electrónico: departamento.economia@eco.uc3m.es

Co-integration and common trends analysis with score-driven models: an application to the federal funds effective rate and US inflation rate

Szabolcs Blazsek[a], Alvaro Escribano[b,*], and Adrian Licht[a]

[a]School of Business, Universidad Francisco Marroquín, Guatemala City, Guatemala

[b]Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain

Abstract: Co-integration and common trends are studied for time series variables, by introducing the new t-QVARMA (quasi-vector autoregressive moving average) model. t-QVARMA is an outlier-robust nonlinear score-driven model for the multivariate t-distribution. In t-QVARMA, the $I(0)$ and $I(1)$ components of the variables are separated in a way that is similar to the Granger-representation of VAR models. The relationship between the co-integrated federal funds effective rate and United States (US) inflation rate variables is studied for the period of July 1954 to January 2019. The in-sample statistical and out-of-sample forecasting performances of t-QVARMA are superior to those of the classical Gaussian-VAR model.

Keywords: Dynamic conditional score (DCS) models; Quasi-vector autoregressive moving average (QVARMA) model; Outliers; Co-integration; Predictive accuracy.

JEL classification codes: C32; C51; C52

*Corresponding author. Department of Economics, Universidad Carlos III de Madrid, Calle Madrid 126, Getafe (Madrid) 28903, Spain. E-mail: alvaroe@eco.uc3m.es

1. Introduction

In this paper, the quasi-vector autoregressive moving average location model is introduced for the multivariate t-distribution (hereinafter, t-QVARMA), in order to study dynamic relationships among $I(0)$ and co-integrated $I(1)$ macroeconomic variables in an outlier-robust way. t-QVARMA is a dynamic conditional score (DCS) model (Harvey, 2013), which is also named generalized autoregressive score (GAS) model (Creal et al., 2011, 2013). DCS models are non-linear state-space models, which are updated by using the conditional score of the log-likelihood (LL). Due to the score-driven updating mechanism, the information gain in the filters is optimal, according to the Kullback–Leibler divergence measure (Blasques et al., 2015). t-QVARMA is robust to outliers, due to the outlier-discounting property of the score function. As a consequence, outliers appear in the irregular component, and the measurements of dynamic interaction effects are less distorted by outliers than the same measurements for the Gaussian linear alternatives. t-QVARMA includes $I(0)$ and co-integrated $I(1)$ variables, and the structure of its formulation is similar to that of the Granger-representation of VAR models (e.g. Johansen, 1995). A limiting special case of t-QVARMA is the linear Gaussian-QVARMA model, in which the $I(0)$ and $I(1)$ components have Gaussian-VARMA forms.

For t-QVARMA, the reduced-form representation, the structural-form representation and the impulse response functions (IRFs) are presented for an application case study, in which monthly data on federal funds effective rate and the United States (US) inflation rate are used for the period of July 1954 to January 2019. Classical co-integration and common trends tests (e.g. Engle and Granger, 1987) suggest that those two variables are co-integrated. The t-QVARMA, Gaussian-QVARMA and Gaussian-VAR (e.g. Lütkepohl, 2005) models are estimated by using the maximum likelihood (ML) method. For t-QVARMA, the conditions of the asymptotic properties of ML are presented. The statistical performance of t-QVARMA is superior to that of Gaussian-VAR, which support the practical use of the IRF estimates for t-QVARMA. The out-of-sample multi-step ahead predictive accuracies of t-QVARMA and Gaussian-VAR are also compared for the period of January 2010 to January 2019 (Diebold and Mariano, 1995; Giacomini and Rossi, 2010). According to the results, the multi-step ahead forecast performance of t-QVARMA is superior to the multi-step ahead forecast performance of Gaussian-VAR.

The remainder of this paper is organized as follows: Section 2 reviews the relevant literature. Section 3 describes the US macroeconomic dataset. Section 4 presents the reduced-form representation of t-QVARMA and its statistical inference, and reports on the ML estimation results for the dataset. Section 5 presents the structural-form representation of t-QVARMA and the corresponding tools of IRF analysis, and reports on the IRF results for the dataset. Section 6 presents the Gaussian-QVARMA model that is a limiting linear special case of t-QVARMA and reports on the ML estimation and IRF results for the dataset. Section 7 presents the ML estimation and IRF results for the most important classical multivariate alternative from the literature: the Gaussian-VAR model. Section 8 presents out-of-sample multi-step ahead predictive accuracy comparisons for the t-QVARMA and Gaussian-VAR models. Section 9 concludes.

2. Review of the literature

Score-driven time series models were introduced in the works of Creal et al. (2011, 2013) and Harvey (2013). Those models are observation-driven (Cox et al., 1981) state space models of univariate or multivariate time series variables, which are estimated by using the ML method. An important property of DCS models is that they are more robust to extreme observations than classical time series models (e.g. Harvey, 2013). DCS models can be applied to study variables with different orders of integration, for example, $I(0)$ variables (e.g. financial return; economic growth) or $I(1)$ variables (e.g. currency exchange rate; GDP level). An example of DCS models is the quasi-AR (QAR) model (Harvey, 2013), which is an outlier-robust alternative to the ARMA model (Box and Jenkins, 1970). Another example of DCS models is the Beta-t-EGARCH (exponential generalized autoregressive conditional heteroscedasticity) model (Harvey and Chakravarty, 2008), which is an outlier-robust alternative to the GARCH (Engle, 1982; Bollerslev, 1986) and EGARCH (Nelson, 1991) models. Both QAR and Beta-t-EGARCH are univariate time series models that are robust to extreme observations.

In this paper new multivariate models of location are used, which are extensions of the 'DCS model for the multivariate t-distribution' (Harvey, 2013). The model of Harvey (2013) is an alternative to the VARMA model (Tiao and Tsay, 1989). The multivariate DCS model for the t-distribution, also named t-QVAR(1), is extended in the work of Blazsek et al. (2018a) to the t-QVAR(p) model. Moreover, Blazsek et al. (2018b) suggest the use of score-driven multivariate seasonality dynamics in the QVAR model and name the new model Seasonal-t-QVAR(p). The works of Blazsek et al. (2018a-b) estimate t-QVAR(p) and Seasonal-t-QVAR(p), respectively, for $I(0)$ variables and present the conditions of the asymptotic properties of the ML estimator for both models. The multivariate DCS models of location can also be applied to the study of $I(1)$ variables. Harvey (2013) presents different DCS models for the multivariate t-distribution for $I(0)$ variables or co-integrated $I(1)$ variables (Granger, 1981; Engle and Granger, 1987).

In the present work, the use of the new t-QVARMA model is suggested for a combination of $I(0)$ and co-integrated $I(1)$ variables. An important property of t-QVARMA models is that they are robust to outliers, as opposed to the classical linear Gaussian vector error correction models (VECMs) which are not. In the body of literature, several works suggest outlier-robust methods for integrated and co-integrated time series. For example, the pseudo-likelihood ratio (PLR) test of Lucas (1997) is an outlier-robust test of co-integration for $I(1)$ time series. See also the related works of Lucas (1995a-b, 1998), Franses and Lucas (1998), and de Jong et al. (2007). Based on those works, Bosco et al. (2010) suggest outlier-robust co-integration tests with the Student's t-distribution. Furthermore, Escribano et al. (2011) suggest the use of median filters for outlier-robust tests of co-integration. The t-QVARMA model of the present paper is an alternative to those outlier-robust frameworks.

3. Data

In this paper, monthly time series data are used from the non-seasonally adjusted effective federal funds effective rate $y_{1,t}$, and the annual percentage change in the non-seasonally adjusted US GNP (gross national product) implicit price deflector index $y_{2,t}$, for the period of June 1954 to January 2019 (source of data: Federal Reserve Economic Data). The use of these variables is motivated by the works of Inoue and Kilian (2016) and Kilian and Lütkepohl (2017). The evolution of the dependent variables is shown in Fig. 1. Descriptive statistics and co-integration test results are reported in Table 1. The co-integration test results are summarized as follows:

Firstly, the Engle–Granger co-integration test (Engle and Granger, 1987) is performed, by using the augmented Dickey–Fuller (ADF) test (Dickey and Fuller, 1979) and the generalized least squares (GLS) estimator (Elliott et al., 1996). The GLS estimator is used to increase the robustness of the Engle–Granger test. In Table 1, results for the two steps of the Engle–Granger test are reported. For the first step, the ADF-GLS test results are reported for $y_{1,t}$ and $y_{2,t}$, which suggest that both variables are $I(1)$. For the second step, the existence of a co-integration relationship between $y_{1,t}$ and $y_{2,t}$ is equivalent to the rejection of the unit root null hypothesis of the ADF-GLS test for the residuals of the linear regression $y_{2t} = \beta_0 + \beta_1 y_{1t} + \epsilon_t$. The test results indicate that federal funds effective rate and US inflation rate are co-integrated.

Secondly, the common trends test of Nyblom and Harvey (2000) is applied:

$$y_t = \mu_t + \eta_{1,t} \text{ where } \eta_{1,t} \sim N(0_{K\times 1}, \Sigma_{1,\eta}) \text{ is serially correlated} \tag{3.1}$$

$$\mu_t = \mu_{t-1} + \eta_{2,t} \text{ where } \eta_{2,t} \sim N(0_{K\times 1}, \Sigma_{2,\eta}) \text{ is serially independent} \tag{3.2}$$

where $K = 2$ for our dataset. Under H_0, $\mathrm{rank}(\Sigma_{2,\eta}) = 1$, i.e. y_t has one common trend. Test statistics for lag-orders $m = 30, 50, 100, 200$ and critical values are reported in Table 1. According to those results, there is a common trend in federal funds effective rate and US inflation rate.

Thirdly, the "maximum eigenvalue" co-integration test of Johansen (1988, 1991, 1995) is applied, which assumes that the reduced-form error term v_t has a multivariate Gaussian distribution in the VECM representation of the VAR(p) model:

$$\Delta y_t = \Pi y_{t-1} + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + v_t \tag{3.3}$$

where Π $(K \times K)$, $\Gamma_1, \ldots, \Gamma_{p-1}$ (each $K \times K$) are time-invariant parameters. H_0 and H_1 for this test, the value of the test statistic and the corresponding p-value are reported in Table 1. The Johansen test suggests that rank(Π) = 1, i.e. federal funds effective rate and US inflation rate are co-integrated.

Fourthly, the outlier-robust co-integration test of Lucas (1997) is performed, by using the multivariate t-distribution for the error term in the previous VECM of Eq. (3.3). H_0 and H_1 for this test are presented in Table 1. For the co-integration test of Lucas, the PLR test statistic for the multivariate t-distribution is reported. According to the results, rank(Π) = 1. Thus, federal funds effective rate and US inflation rate are co-integrated. These co-integration test results motivate the t-QVARMA specification of the following section.

4. The t-QVARMA model

4.1. Reduced-form representation

For the dependent variables y_t $(K \times 1)$ with $t = 1 \ldots, T$, it is assumed that K^* variables are $I(0)$ and $K^\dagger$ variables are $I(1)$ and co-integrated. The $K = K^* + K^\dagger$ dependent variables are ordered in such a way that the first K^* variables are $I(0)$ and the remaining $K^\dagger$ variables are $I(1)$. The reduced-form representation of t-QVARMA(p,q,r) is:

$$y_t = c^* + \mu_t + v_t = c^* + \mu_t^* + \mu_t^\dagger + v_t \tag{4.1}$$

$$\mu_t^* = \sum_{i=1}^{p} \Phi_i^* \mu_{t-i}^* + \sum_{j=1}^{q} \Psi_j^* u_{t-j} \tag{4.2}$$

$$\mu_t^\dagger = \mu_{t-1}^\dagger + \sum_{l=1}^{r} \Psi_l^\dagger u_{t-l} \tag{4.3}$$

$$v_t \sim t_K(0_{K\times 1}, \Sigma, \nu) = t_K[0, \Omega^{-1}(\Omega^{-1})', \nu] \text{ i.i.d.} \tag{4.4}$$

where c^* $(K \times 1)$, $\Phi_1^*, \ldots, \Phi_p^*$ (all $K \times K$), $\Psi_1^*, \ldots, \Psi_q^*$ (all $K \times K$), $\Psi_1^\dagger, \ldots, \Psi_r^\dagger$ (all $K \times K$), Ω^{-1} $(K \times K)$, and ν are constant parameters. Variable u_t $(K \times 1)$ is the scaled score function and it is defined in the next section. Vector μ_t^* includes $I(0)$ variables and vector $\mu_t^\dagger$ includes co-integrated $I(1)$ variables. In this way, y_t is decomposed into $I(0)$ and $I(1)$ components in a way that is similar to the Granger-representation of the VAR model (e.g. Johansen, 1995). The reduced-form error term v_t represents the unexpected part of y_t, since the conditional mean of the dependent variables is: $E(y_t|y_1,\ldots,y_{t-1}) = c^* + \mu_t = c^* + \mu_t^* + \mu_t^\dagger$.

For the first $\max(p,q)$ observations of the sample, μ_t^* is initialized by using $0_{K\times 1}$, in order to effectively separate μ_t^* and $\mu_t^\dagger$. For the first r observations of the sample, $\mu_t^\dagger$ is initialized by estimating the parameter vectors $\mu_{0,1}^\dagger,\ldots,\mu_{0,r}^\dagger$ (each $K\times 1$). For each of those vectors, the first K^* elements are zero and the last $K^\dagger$ elements are estimated. The first K^* elements of c^* are estimated, and the last $K^\dagger$ elements of c^* are restricted to zeros due to reasons of parameter identification (i.e. the random walk without drift specification is used for the $I(1)$ variables). All of the elements of $\Phi_1^*,\ldots,\Phi_p^*$ and $\Psi_1^*,\ldots,\Psi_q^*$ are unrestricted real numbers. For each of the matrices $\Psi_1^\dagger,\ldots,\Psi_r^\dagger$, all of the elements in the first K^* rows and all of the elements in the first K^* columns are restricted to zeros. For each of the matrices $\Psi_1^\dagger,\ldots,\Psi_r^\dagger$, the rank of the quadratic submatrix that is formed by the elements of the last $K^\dagger$ rows and the last $K^\dagger$ columns is $1 \leq R < K^\dagger$. The positive-definite matrix Σ is estimated by using the decomposition $\Omega^{-1}(\Omega^{-1})'$. In this paper, it is assumed that $\Sigma = \Omega^{-1}(\Omega^{-1})'$ is a Cholesky decomposition (i.e. Ω^{-1} is a $K\times K$ lower-triangular matrix with positive elements in the diagonal) (see for example Lütkepohl, 2005). The latter assumption can be extended to allow for more general decompositions of Σ. The degrees of freedom parameter ν of the multivariate t distribution is jointly estimated with the rest of the parameters. It is assumed that $\nu > 2$ to ensure that the covariance matrix of v_t is well-defined.

4.2. Scaled score function

In this section, the updating term u_t that updates the $I(0)$ and $I(1)$ components of t-QVARMA is defined. The log of the conditional density of y_t given $\mathcal{F}_{t-1} = (y_1,\ldots,y_{t-1})$ is

$$\ln f(y_t|\mathcal{F}_{t-1};\Theta) = \ln\Gamma\left(\frac{\nu+K}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{K}{2}\ln(\pi\nu) \tag{4.5}$$

$$-\frac{1}{2}\ln|\Sigma| - \frac{\nu+K}{2}\ln\left(1+\frac{v_t'\Sigma^{-1}v_t}{\nu}\right)$$

where $\Theta = (\Theta_1, \ldots, \Theta_S)'$ is the vector of time-invariant parameters. The partial derivative of the log of the conditional density with respect to μ_t is

$$\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} = \frac{\nu+K}{\nu}\Sigma^{-1} \times \left(1 + \frac{v_t'\Sigma^{-1}v_t}{\nu}\right)^{-1} v_t = \frac{\nu+K}{\nu}\Sigma^{-1} \times u_t \tag{4.6}$$

(Harvey, 2013). The latter equality defines the scaled score function u_t, by using the reduced-form error term. In the definition of u_t, v_t is multiplied by $[1+(v_t'\Sigma_v^{-1}v_t)/\nu]^{-1} = \nu/(\nu+v_t'\Sigma_v^{-1}v_t) \in (0,1)$. Therefore, the scaled score function is bounded by the reduced-form error term: $|u_t| < |v_t|$. The scaled score function u_t is multivariate i.i.d. with mean zero and covariance matrix:

$$\text{Var}(u_t) = E\left[\frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t} \times \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial \mu_t'}\right] = \frac{\nu+K}{\nu+K+2}\Sigma^{-1}. \tag{4.7}$$

Under the definition of u_t, t-QVARMA(p,q,r) can be written as:

$$y_t = c^* + \mu_t^* + \mu_t^\dagger + v_t \tag{4.8}$$

$$\mu_t^* = \sum_{i=1}^{p} \Phi_i^* \mu_{t-i}^* + \sum_{j=1}^{q} \Psi_j^* \left(1 + \frac{v_{t-j}'\Sigma^{-1}v_{t-j}}{\nu}\right)^{-1} v_{t-j} \tag{4.9}$$

$$\mu_t^\dagger = \mu_{t-1}^\dagger + \sum_{l=1}^{r} \Psi_l^\dagger \left(1 + \frac{v_{t-l}'\Sigma^{-1}v_{t-l}}{\nu}\right)^{-1} v_{t-l} \tag{4.10}$$

$$v_t \sim t_K(0_{K\times 1}, \Sigma, \nu) = t_K[0, \Omega^{-1}(\Omega^{-1})', \nu] \text{ i.i.d.} \tag{4.11}$$

where the nonlinear dependence on v_t is shown in Eqs. (4.9) and (4.10).

4.3. First-order representation of the reduced form

The first-order representation of t-QVARMA(p, q, r) is:

$$Y_t = C^* + M_t^* + M_t^\dagger + V_t \tag{4.12}$$

$$M_t^* = \Phi^* M_{t-1}^* + \Psi^* U_{t-1}^* \tag{4.13}$$

$$M_t^{\dagger} = M_{t-1}^{\dagger} + \Psi^{\dagger} U_{t-1}^{\dagger} \tag{4.14}$$

where

$$Y_t = \begin{bmatrix} y_t \\ y_{t-1} \\ \vdots \\ y_{t-p+1} \end{bmatrix}_{(Kp\times 1)} \qquad C^* = \begin{bmatrix} c^* \\ c^* \\ \vdots \\ c^* \end{bmatrix}_{(Kp\times 1)} \qquad M_t^* = \begin{bmatrix} \mu_t^* \\ \mu_{t-1}^* \\ \vdots \\ \mu_{t-p+1}^* \end{bmatrix}_{(Kp\times 1)}$$

$$M_t^{\dagger} = \begin{bmatrix} \mu_t^{\dagger} \\ \mu_{t-1}^{\dagger} \\ \vdots \\ \mu_{t-p+1}^{\dagger} \end{bmatrix}_{(Kp\times 1)} \qquad V_t = \begin{bmatrix} v_t \\ v_{t-1} \\ \vdots \\ v_{t-p+1} \end{bmatrix}_{(Kp\times 1)} \qquad U_t^* = \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-q+1} \end{bmatrix}_{(Kq\times 1)}$$

$$U_t^{\dagger} = \begin{bmatrix} u_t \\ u_{t-1} \\ \vdots \\ u_{t-r+1} \end{bmatrix}_{(Kr\times 1)} \qquad \Phi^* = \begin{bmatrix} \Phi_1^* & \Phi_2^* & \cdots & \Phi_{p-1}^* & \Phi_p^* \\ I_K & 0_{K\times K} & \cdots & \cdots & 0_{K\times K} \\ 0_{K\times K} & I_K & 0_{K\times K} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0_{K\times K} & \cdots & 0_{K\times K} & I_K & 0_{K\times K} \end{bmatrix}_{(Kp\times Kp)}$$

$$\Psi^* = \begin{bmatrix} \Psi_1^* & \Psi_2^* & \cdots & \Psi_q^* \\ 0_{K\times K} & 0_{K\times K} & \cdots & 0_{K\times K} \\ \cdots & \cdots & \cdots & \cdots \\ 0_{K\times K} & \cdots & \cdots & 0_{K\times K} \end{bmatrix}_{(Kp\times Kq)} \qquad \Psi^{\dagger} = \begin{bmatrix} \Psi_1^{\dagger} & \Psi_2^{\dagger} & \cdots & \Psi_r^{\dagger} \\ 0_{K\times K} & 0_{K\times K} & \cdots & 0_{K\times K} \\ \cdots & \cdots & \cdots & \cdots \\ 0_{K\times K} & \cdots & \cdots & 0_{K\times K} \end{bmatrix}_{(Kp\times Kr)}$$

4.4. Maximum likelihood estimator

The parameters of t-QVARMA are estimated by using the ML method, as follows:

$$\hat{\Theta}_{\mathrm{ML}} = \arg\max_{\Theta} \mathrm{LL}(y_1, \ldots, y_T; \Theta) = \arg\max_{\Theta} \sum_{t=1}^{T} \ln f(y_t | \mathcal{F}_{t-1}; \Theta) \tag{4.15}$$

The following assumptions are used: (A1) $f(y_t|\mathcal{F}_{t-1};\Theta) = p_0(y_t|\mathcal{F}_{t-1};\Theta_0)$ for some Θ from the parameter set $\tilde{\Theta}$, where p_0 is the true conditional density and Θ_0 represents the true values of Θ. (A2) $\int_{\mathbb{R}} f(y_t|\mathcal{F}_{t-1};\Theta)dy_t = 1$ for all y_t and Θ. (A3) $\tilde{\Theta} \in \mathbb{R}^S$ is compact. (A4) $\hat{\Theta}_{\text{ML}}$ is a unique solution to the problem of Eq. (4.15). (A5) $\text{LL}(\cdot;\Theta)$ is a Borel measurable function on $\mathbb{R}^T$. (A6) For each $(y_1,\ldots,y_T) \in \mathbb{R}^T$, $\text{LL}(y_1,\ldots,y_T;\cdot)$ is a continuous function on $\tilde{\Theta}$. (A7) $|\text{LL}(y_1,\ldots,y_T;\Theta)| < b(y_1,\ldots,y_T)$ for all Θ and $E[b(y_1,\ldots,y_T)] < \infty$. Under assumptions (A1) to (A7), the ML estimator of parameters is consistent: $\hat{\Theta}_{\text{ML}} \to_p \Theta_0$ as $T \to \infty$.

The following additional assumptions are also used: (A8) Θ_0 is an interior point within $\tilde{\Theta} \in \mathbb{R}^K$. (A9) $\text{LL}(y_1,\ldots,y_T;\Theta)$ is twice continuously differentiable on all of the interior points of $\tilde{\Theta}$. (A10) $\partial[\int_{\mathbb{R}} f(y_t|\mathcal{F}_{t-1};\Theta)dy_t]/\partial\Theta = \int_{\mathbb{R}}[\partial f(y_t|\mathcal{F}_{t-1};\Theta)/\partial\Theta]dy_t$.

The $T \times S$ matrix of contributions to the gradient $G(y_1,\ldots,y_T,\Theta)$ are defined by its elements:

$$G_{ti}(\Theta) = \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial\Theta_i} \tag{4.16}$$

for period $t = 1,\ldots,T$ and parameter $i = 1,\ldots,S$. The t-th row of $G(y_1,\ldots,y_T,\Theta)$ is denoted by using $G_t(\Theta)$, which is the score vector for the t-th observation. Under (A1) to (A10), the ML estimator of Eq. (4.15) is equivalent to the representation:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} G_{t1}(\hat{\Theta}_{\text{ML}}) \\ \vdots \\ G_{tS}(\hat{\Theta}_{\text{ML}}) \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T}\begin{bmatrix} \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};p_0,\hat{\Theta}_{\text{ML}})}{\partial\Theta_1} \\ \vdots \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1};p_0,\hat{\Theta}_{\text{ML}})}{\partial\Theta_S} \end{bmatrix} = 0_{S\times 1} \tag{4.17}$$

According to the mean-value expansion about Θ_0:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' = \frac{1}{T}\sum_{t=1}^{T} G_t(\Theta_0)' + \frac{1}{T}\left[\sum_{t=1}^{T} H_t(\bar{\Theta})\right](\hat{\Theta}_{\text{ML}} - \Theta_0) \tag{4.18}$$

where each row of the $S \times S$ Hessian matrix

$$H_t(\Theta) = \frac{\partial^2 \ln f(y_t|\mathcal{F}_{t-1};\Theta)}{\partial\Theta\Theta'} \tag{4.19}$$

which is evaluated at S different mean values, indicated by $\bar{\Theta}$. Each $\bar{\Theta}$ is located between Θ_0 and $\hat{\Theta}_{\text{ML}}$ that is expressed as: $||\bar{\Theta} - \Theta_0|| \leq ||\hat{\Theta}_{\text{ML}} - \Theta_0||$, where $||\cdot||$ is the Euclidean norm.

The following additional assumptions are also used: (A11) $\partial[\int_{\mathbb{R}} G_t(\Theta)' f(y_t|\mathcal{F}_{t-1};\Theta)dy_t]/\partial\Theta = \int_{\mathbb{R}}[\partial G_t(\Theta)' f(y_t|\mathcal{F}_{t-1};\Theta)/\partial\Theta]dy_t$. (A12) The information matrix $\mathcal{I}(\Theta_0) \equiv -E[H_t(\Theta_0)]$ is positive definite. (A13) The elements of $\mathcal{I}(\Theta_0)$ are bounded in absolute value by function $b(y_1,\ldots,y_T)$ for all Θ and $E[b(y_1,\ldots,y_T)] < \infty$. The conditions of (A13) are studied in the next subsection. Under (A1) to (A13), the $S \times S$ contribution to the information matrix for period t is given by:

$$I_t(\Theta_0) = -E[H_t(\Theta_0)|\mathcal{F}_{t-1}] = E[G_t(\Theta_0)' G_t(\Theta_0)|\mathcal{F}_{t-1}] \tag{4.20}$$

which is evaluated at the true values of parameters. From Eqs. (4.17) and (4.18):

$$\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) = \left[-\frac{1}{T}\sum_{t=1}^{T} H_t(\bar{\Theta})\right]^{-1} \left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] \tag{4.21}$$

$$\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) = \mathcal{I}^{-1}(\Theta_0)\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] + \text{o}_p(1) \tag{4.22}$$

Finally, the following assumptions are also used: (A14) The conditions of a central limit theorem are satisfied for Eq. (4.22). (A15) t-QVARMA is invertible (see the invertibility of VARMA models in Lütkepohl, 2005). Under (A1) to (A15), the asymptotic result is:

$$\sqrt{T}(\hat{\Theta}_{\text{ML}} - \Theta_0) \rightarrow_d N_S\left[0_{S\times 1}, \mathcal{I}^{-1}(\Theta_0)\right] \quad \text{as} \quad T \rightarrow \infty \tag{4.23}$$

The asymptotic covariance matrix of $\hat{\Theta}_{\text{ML}}$ is estimated by using $[\sum_{t=1}^{T} G_t(\hat{\Theta}_{\text{ML}})' G_t(\hat{\Theta}_{\text{ML}})]^{-1}$ (e.g. Harvey, 2013; Creal et al., 2013). If Eq. (4.23) holds, then: $E(u_t) = 0_{K\times 1}$, each element of u_t forms a martingale difference sequence (MDS), u_t discounts the effects of extreme unexpected observations in the dynamic equations. These properties of the scaled score function motivate its use as an updating term in the t-QVARMA model.

4.5. Finiteness of the information matrix

In this section, the conditions of (A13) are studied. Those conditions extend the results of Harvey (2013, Section 2.4). The majority of this section has a focus on the parameters of M_t^*. At the end of this section, the ML conditions related to the parameters of $M_t^\dagger$ are also presented.

With respect to M_t^*, the first condition is related to the covariance stationarity of μ_t^*. For the estimates of Φ^*, the maximum modulus of the eigenvalues is denoted by C_1. If $C_1 < 1$, then μ_t^* is covariance stationary.

The second condition is the application of a condition from the work of Harvey (2013):

$$E\left[\left(U_{j,t}^*\right)^{2-i}\left(\frac{\partial U_{k,t}^*}{\partial M_{l,t}^*}\right)^i\right] < \infty \tag{4.24}$$

for $i = 0, 1, 2$, where $j, k = 1, \ldots, Kq$ and $l = 1, \ldots, Kp$. Condition 2 is denoted by using C_2.

For the third condition, the following derivatives are considered from Eq. (4.13):

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} = \Phi^* \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + \Psi^* \frac{\partial U_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \tag{4.25}$$

for $i = 1, \ldots, Kp$ and $j = 1, \ldots, Kq$; $W_{i,j}$ is a $(Kp \times Kq)$ matrix, in which element (i, j) is one and the rest of the elements are zero. The derivatives of Eq. (4.25) are in $G_t(\Theta_0)'$ and also in $I_t(\Theta_0)$. By using the chain rule, from Eq. (4.25):

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} = \left[\Phi^* + \Psi^* \frac{\partial U_{t-1}^*}{\partial (M_{t-1}^*)'}\right] \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \equiv X_t \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \tag{4.26}$$

Condition 3 is that all eigenvalues of $E(X_t)$ are within the unit circle; under Condition 2, $E(X_t)$ is finite. The maximum modulus of the eigenvalues of the estimates of $E(X_t)$ is denoted by using C_3. Conditions 1 to 3 are among the conditions of (A14).

For the fourth condition, the following derivatives are within the information matrix:

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} \frac{\partial (M_t^*)'}{\partial \Psi_{k,l}^*} = X_t \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} \frac{\partial (M_{t-1}^*)'}{\partial \Psi_{k,l}^*} X_t' + X_t \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} (U_{t-1}^*)' W_{i,j}' \tag{4.27}$$

$$+W_{k,l}U^*_{t-1}\frac{\partial(M^*_{t-1})'}{\partial\Psi^*_{k,l}}X'_t+W_{i,j}U^*_{t-1}(U^*_{t-1})'W'_{k,l}$$

which can be written as:

$$\text{vec}\left[\frac{\partial M^*_t}{\partial\Psi^*_{i,j}}\frac{\partial(M^*_t)'}{\partial\Psi^*_{k,l}}\right]=(X_t\otimes X_t)\text{vec}\left[\frac{\partial M^*_{t-1}}{\partial\Psi^*_{i,j}}\frac{\partial(M^*_{t-1})'}{\partial\Psi^*_{k,l}}\right]+\text{vec}\left[X_t\frac{\partial M^*_{t-1}}{\partial\Psi^*_{i,j}}(U^*_{t-1})'W'_{i,j}\right] \quad (4.28)$$

$$+\text{vec}\left[W_{k,l}U^*_{t-1}\frac{\partial(M^*_{t-1})'}{\partial\Psi^*_{k,l}}X'_t\right]+\text{vec}\left[W_{i,j}U^*_{t-1}(U^*_{t-1})'W'_{k,l}\right]$$

Condition 4 is that all eigenvalues of $E(X_t\otimes X_t)$ are within the unit circle; under Condition 2, the expectations within $E(X_t\otimes X_t)$ are finite. The maximum modulus of the eigenvalues of the estimates of $E(X_t\otimes X_t)$ is denoted by using C_4. (A13) holds under Conditions 1 to 4.

With respect to the conditions of ML for $M^\dagger_t$, we refer to the work of Harvey (2013, p. 49). According to Harvey, provided that the first-order dynamic parameter matrix is fixed to the identity matrix in Eq. (4.14), the asymptotic theory for the ML estimator holds.

4.6. Empirical application

The t-QVARMA(1,1,1) and t-QVARMA(1,1,4) specifications are estimated for the federal funds effective rate $y_{1,t}$ and US inflation rate $y_{2,t}$ variables (the QVARMA lag-orders are chosen based on LL-based performance metrics, by comparing several alternative specifications). For these co-integrated $I(1)$ variables, t-QVARMA(1,1,1) is formulated as follows:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix}=\begin{bmatrix} \mu^*_{1,t} \\ \mu^*_{2,t} \end{bmatrix}+\begin{bmatrix} \mu^\dagger_{1,t} \\ \mu^\dagger_{2,t} \end{bmatrix}+\begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} \quad (4.29)$$

$$\begin{bmatrix} \mu^*_{1,t} \\ \mu^*_{2,t} \end{bmatrix}=\begin{bmatrix} \Phi^*_{1,11} & \Phi^*_{1,12} \\ \Phi^*_{1,21} & \Phi^*_{1,22} \end{bmatrix}\begin{bmatrix} \mu^*_{1,t-1} \\ \mu^*_{2,t-1} \end{bmatrix}+\begin{bmatrix} \Psi^*_{1,11} & \Psi^*_{1,12} \\ \Psi^*_{1,21} & \Psi^*_{1,22} \end{bmatrix}\begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} \quad (4.30)$$

$$\begin{bmatrix} \mu^\dagger_{1,t} \\ \mu^\dagger_{2,t} \end{bmatrix}=\begin{bmatrix} \mu^\dagger_{1,t-1} \\ \mu^\dagger_{2,t-1} \end{bmatrix}+\begin{bmatrix} \Psi^\dagger_{1,11} & \Psi^\dagger_{1,12} \\ \kappa\Psi^\dagger_{1,11} & \kappa\Psi^\dagger_{1,12} \end{bmatrix}\begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} \quad (4.31)$$

$$v_t \sim t_2 \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega_{11}^{-1} & 0 \\ \Omega_{21}^{-1} & \Omega_{22}^{-1} \end{bmatrix} \times \begin{bmatrix} \Omega_{11}^{-1} & \Omega_{21}^{-1} \\ 0 & \Omega_{22}^{-1} \end{bmatrix}, \nu \right\} \text{ i.i.d.} \tag{4.32}$$

In Eq. (4.31), parameter κ (1×1) is used and under that specification $\text{rank}(\Psi_1^\dagger) = 1$. An alternative way of ensuring the same rank is to use a parametrization that is used in the classical VECM representation of VAR models for co-integrated $I(1)$ variables: $\Psi_1^\dagger = \alpha\beta'$, where α and β (both $K \times R$) are parameter matrices (e.g. Lütkepohl, 2005). The first R rows of β are restricted to the identity matrix I_R, due to parameter identification in the decomposition $\Psi_1^\dagger = \alpha\beta'$. For both parametrizations similar results were obtained, hence only the results for Eq. (4.31) are reported in this paper. The μ_t^* component is initialized by using a 2×1 vector of zeros. The $\mu_t^\dagger$ component is initialized by using parameter vector $(\mu_{0,11}^\dagger, \kappa\mu_{0,11}^\dagger)'$, which maintains the co-integration relation. For the t-QVARMA(1,1,4) specification, Eq. (4.31) is replaced by:

$$\begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \end{bmatrix} = \begin{bmatrix} \mu_{1,t-1}^\dagger \\ \mu_{2,t-1}^\dagger \end{bmatrix} + \begin{bmatrix} \Psi_{1,11}^\dagger & \Psi_{1,12}^\dagger \\ \kappa\Psi_{1,11}^\dagger & \kappa\Psi_{1,12}^\dagger \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \end{bmatrix} + \ldots + \begin{bmatrix} \Psi_{4,11}^\dagger & \Psi_{4,12}^\dagger \\ \kappa\Psi_{4,11}^\dagger & \kappa\Psi_{4,12}^\dagger \end{bmatrix} \begin{bmatrix} u_{1,t-4} \\ u_{2,t-4} \end{bmatrix} \tag{4.33}$$

where parameter κ is used for all the lags to ensure the co-integration relationship between federal funds effective rate and US inflation rate. For both t-QVARMA specifications, the co-integration vector is $(-\kappa, 1)'$ and $\mu_{1,t}^\dagger$ defines the common factor.

ML parameter estimates and model diagnostics for t-QVARMA(1,1,1) and t-QVARMA(1,1,4) are presented in Table 2. There are statistically significant coefficients within the Φ_i^*, Ψ_j^* and $\Psi_i^\dagger$ parameter matrices, which indicate that both short-run and long-run dynamics exist in the federal funds effective rate and US inflation rate times series. Moreover, κ is significant for both specifications, indicating a significant co-integration relation between the two dependent variables. The degrees of freedom parameter estimates for the t-QVARMA(1,1,1) and t-QVARMA(1,1,4) models are 9.8545 and 9.1904, respectively. Thus, (i) the probability distribution of v_t is not a multivariate Gaussian distribution, which is a special case of the multivariate

t distribution for large values of ν; (ii) the first two moments of v_t are well-defined, which supports the asymptotic properties of the ML estimator.

According to the conditions of the ML estimator, $C_1 < 1$ for both t-QVARMA(1,1,1) and t-QVARMA(1,1,4). Thus, Condition 1 is supported. Moreover, the ADF tests support the ML estimator for all variables of Condition 2. With respect to Conditions 3 and 4, the results show that t-QVARMA(1,1,1) is not supported, but $C_3 < 1$ and $C_4 < 1$ for t-QVARMA(1,1,4). As a consequence, Conditions 1 to 4 are supported for t-QVARMA(1,1,4). These results indicate that it is necessary to include several lags of the score function into the $I(1)$ component of t-QVARMA for the dataset of this paper. This motivates the formulation of the $I(1)$ component according to Eq. (4.3). To compare the statistical performances of alternative models, the mean LL (LL/T), Akaike information criterion (AIC), Bayesian information criterion (BIC), and Hannan–Quinn criterion (HQC) metrics are used. The results indicate that the statistical performance of t-QVARMA(1,1,4) is superior to the statistical performance of t-QVARMA(1,1,1).

The evolution of the time series components μ_t^*, $\mu_t^\dagger$, v_t and the evolution of the score function u_t for the t-QVARMA(1,1,4) model are presented in Fig. 2. The discounting property of u_t is studied for t-QVARMA(1,1,4) in Fig. 3, in which both elements of u_t are presented, as functions of $\epsilon_{1,t}$ and $\epsilon_{2,t}$. The figure indicates that $u_{1,t} \to 0$ and $u_{2,t} \to 0$, where $|\epsilon_{1,t}| \to \infty$ and $|\epsilon_{2,t}| \to \infty$. This can be described as an asymptotic trimming of extreme values.

As an additional example, a second specification of t-QVARMA for three variables is provided in the Appendix, for which the first variable is $I(0)$ and the second and third variables are $I(1)$ and they are co-integrated. That specification may be applied to the US gross national product (GNP) growth $y_{1,t}$, federal funds effective rate $y_{2,t}$, and US inflation rate $y_{3,t}$ variables.

5. Dynamic interaction effects

5.1. Structural-form representation

Firstly, the variance of the reduced-form error term $v_t \sim t_K(0, \Sigma, \nu)$ is factorized, as follows:

$$\text{Var}(v_t) = \Sigma \times \frac{\nu}{\nu - 2} = \left(\frac{\nu}{\nu - 2}\right)^{1/2} \times \Omega^{-1}(\Omega^{-1})' \times \left(\frac{\nu}{\nu - 2}\right)^{1/2} \tag{5.1}$$

Based on that, the following multivariate i.i.d. structural-form error term ϵ_t is introduced:

$$v_t = \left(\frac{\nu}{\nu-2}\right)^{1/2} \Omega^{-1} \times \epsilon_t \tag{5.2}$$

where $E(\epsilon_t) = 0$, $\text{Var}(\epsilon_t) = I_K$ and $\epsilon_t \sim t_K[0, I_K \times (\nu-2)/\nu, \nu]$. This gives the following structural-form representation of the t-QVARMA model:

$$\left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega y_t = \tag{5.3}$$

$$= \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega c^* + \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega \mu_t^* + \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega \mu_t^\dagger + \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega v_t =$$

$$= \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega c^* + \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega \mu_t^* + \left(\frac{\nu}{\nu-2}\right)^{-1/2} \Omega \mu_t^\dagger + \epsilon_t$$

By using the structural-form error term, the scaled score function is expressed as:

$$u_t = [(\nu-2)\nu]^{1/2} \Omega^{-1} \times \frac{\epsilon_t}{\nu - 2 + \epsilon_t' \epsilon_t} \tag{5.4}$$

This latter representation is useful for the IRF analysis of t-QVARMA estimates.

5.2. IRF analysis

The dynamic interaction effects of y_t are studied for the nonlinear t-QVARMA model:

$$\text{IRF}_{j,t} = \frac{\partial y_{t+j}}{\partial \epsilon_t} = \frac{\partial \mu_{t+j}^*}{\partial \epsilon_t} + \frac{\partial \mu_{t+j}^\dagger}{\partial \epsilon_t} + \frac{\partial v_{t+j}}{\partial \epsilon_t} \quad \text{for} \quad j = 0, 1, \ldots, \infty \tag{5.5}$$

The first and second terms of Eq. (5.5) are zero for $j = 0$, since both μ_t^* and $\mu_t^\dagger$ include lags of ϵ_t. Those terms represent the short-run dynamic interaction effects $\text{IRF}_{j,t}^*$ and the long-run dynamic interaction effects $\text{IRF}_{j,t}^\dagger$, respectively. The third term of Eq. (5.5) is zero for $j > 0$, and it represents the contemporaneous interaction effects $\text{IRF}_{j,t}^v$. Under this notation,

$$\text{IRF}_{j,t} = \text{IRF}_{j,t}^* + \text{IRF}_{j,t}^\dagger + \text{IRF}_{j,t}^v \quad \text{for} \quad j = 0, 1, \ldots, \infty \tag{5.6}$$

With respect to $\mathrm{IRF}^*_{j,t}$, if $C_1 < 1$, then

$$M_t^* = (1 - \Phi^* L)^{-1} \Psi^* U^*_{t-1} \tag{5.7}$$

$$M_t^* = \sum_{j=0}^{\infty} (\Phi^*)^j \Psi^* U^*_{t-1-j} \tag{5.8}$$

$$J_1 M_t^* = \sum_{j=0}^{\infty} J_1 (\Phi^*)^j \Psi^* J_2' J_2 U^*_{t-1-j} \tag{5.9}$$

where $J_1 = (I_K, 0_{K\times K}, \cdots, 0_{K\times K})$ is $(K \times Kp)$, $J_2 = (I_K, 0_{K\times K}, \cdots, 0_{K\times K})$ is $(K \times Kq)$. Then,

$$\mu_t^* = \sum_{j=0}^{\infty} J_1 (\Phi^*)^j \Psi^* J_2' u_{t-1-j} = \sum_{j=0}^{\infty} J_1 (\Phi^*)^j \Psi^* J_2' [(\nu - 2)\nu]^{1/2} \Omega^{-1} \times \frac{\epsilon_{t-1-j}}{\nu - 2 + \epsilon'_{t-1-j} \epsilon_{t-1-j}} \tag{5.10}$$

Therefore, the IRF for short-run interaction effects is given by:

$$\mathrm{IRF}^*_{j,t} = \partial \mu^*_{t+j} / \partial \epsilon_t = J_1 (\Phi^*)^{j-1} \Psi^* J_2' [(\nu - 2)\nu]^{1/2} \Omega^{-1} D_t \quad \text{for} \quad j = 1, \ldots, \infty \tag{5.11}$$

$$D_t = \frac{\partial \frac{\epsilon_t}{\nu - 2 + \epsilon_t' \epsilon_t}}{\partial \epsilon_t} = \begin{bmatrix} d_{11,t} & \cdots & d_{1K,t} \\ \cdots & \cdots & \cdots \\ d_{K1,t} & \cdots & d_{KK,t} \end{bmatrix} = \tag{5.12}$$

$$= \begin{bmatrix} \frac{\nu-2+\epsilon_t'\epsilon_t - 2\epsilon_{1,t}^2}{(\nu-2+\epsilon_t'\epsilon_t)^2} & \frac{-2\epsilon_{1,t}\epsilon_{2,t}}{(\nu-2+\epsilon_t'\epsilon_t)^2} & \cdots & \frac{-2\epsilon_{1,t}\epsilon_{K,t}}{(\nu-2+\epsilon_t'\epsilon_t)^2} \\ \frac{-2\epsilon_{2,t}\epsilon_{1,t}}{(\nu-2+\epsilon_t'\epsilon_t)^2} & \frac{\nu-2+\epsilon_t'\epsilon_t - 2\epsilon_{2,t}^2}{(\nu-2+\epsilon_t'\epsilon_t)^2} & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots \\ \frac{-2\epsilon_{K,t}\epsilon_{1,t}}{(\nu-2+\epsilon_t'\epsilon_t)^2} & \cdots & \cdots & \frac{\nu-2+\epsilon_t'\epsilon_t - 2\epsilon_{K,t}^2}{(\nu-2+\epsilon_t'\epsilon_t)^2} \end{bmatrix}$$

With respect to $\mathrm{IRF}^\dagger_{j,t}$, component $\mu_t^\dagger$ is expressed from Eq. (4.3) by recursive substitution as:

$$\begin{aligned} \mu_t^\dagger = \; & \Psi_1^\dagger u_{t-1} + (\Psi_1^\dagger + \Psi_2^\dagger) u_{t-2} + \ldots + (\Psi_1^\dagger + \ldots + \Psi_K^\dagger) u_{t-K} \\ & + (\Psi_1^\dagger + \ldots + \Psi_K^\dagger) u_{t-K-1} + \ldots + (\Psi_1^\dagger + \ldots + \Psi_K^\dagger) u_K \\ & + (\Psi_1^\dagger + \ldots + \Psi_{K-1}^\dagger) u_{K-1} + \ldots + (\Psi_1^\dagger + \Psi_2^\dagger) u_2 + \Psi_1^\dagger u_1 \end{aligned} \tag{5.13}$$

Therefore, the IRF for long-run interaction effects is given by:

$$
\begin{aligned}
\mathrm{IRF}^{\dagger}_{1,t} &= \partial\mu^{\dagger}_{t+1}/\partial\epsilon_t = \Psi^{\dagger}_1[(\nu-2)\nu]^{1/2}\Omega^{-1}D_t \\
\mathrm{IRF}^{\dagger}_{2,t} &= \partial\mu^{\dagger}_{t+2}/\partial\epsilon_t = (\Psi^{\dagger}_1+\Psi^{\dagger}_2)[(\nu-2)\nu]^{1/2}\Omega^{-1}D_t \\
&\vdots \\
\mathrm{IRF}^{\dagger}_{K-1,t} &= \partial\mu^{\dagger}_{t+K-1}/\partial\epsilon_t = (\Psi^{\dagger}_1+\ldots+\Psi^{\dagger}_{K-1})[(\nu-2)\nu]^{1/2}\Omega^{-1}D_t \\
\mathrm{IRF}^{\dagger}_{j,t} &= \partial\mu^{\dagger}_{t+j}/\partial\epsilon_t = (\Psi^{\dagger}_1+\ldots+\Psi^{\dagger}_{K})[(\nu-2)\nu]^{1/2}\Omega^{-1}D_t \quad \text{for} \quad j \geq K
\end{aligned}
\tag{5.14}
$$

The IRFs in Eqs. (5.11) and (5.14) depend on D_t, i.e. those IRFs are time-varying. In the QVARMA specification, v_t and ϵ_t are multivariate i.i.d. error terms; hence, they are strictly stationary and ergodic. According to White (1984, Theorem 3.35), a possibly nonlinear measurable function transforms strictly stationary and ergodic variables to new strictly stationary and ergodic variables. Therefore, all elements of D_t are strictly stationary and ergodic, and $E(D_t)$ can be estimated by using the sample average due to the ergodic theorem. The use of the sample average estimator after the ML estimation is also supported, by using the ADF unit root test for all the elements of D_t. It is noteworthy that an alternative to the use of $E(\mathrm{IRF}^{*}_{j,t})$ and $E(\mathrm{IRF}^{\dagger}_{j,t})$ is the period-by-period estimation of $\mathrm{IRF}^{*}_{j,t}$ and $\mathrm{IRF}^{\dagger}_{j,t}$, respectively. In those applications, both $\mathrm{IRF}^{*}_{j,t}$ and $\mathrm{IRF}^{\dagger}_{j,t}$ may be averaged for several sub-sample periods and the resulting IRF estimates may be compared. Finally, with respect to $\mathrm{IRF}^{v}_{j,t}$:

$$
\mathrm{IRF}^{v}_{j,t} = \left(\frac{\nu}{\nu-2}\right)^{1/2}\Omega^{-1} \quad \text{for} \quad j=0. \tag{5.15}
$$

5.3. Empirical application

In Figs. 4 and 5, the IRFs for the t-QVARMA(1,1,1) and t-QVARMA(1,1,4) specifications, respectively, are presented, for which the signs of the dynamic interaction effects are similar. For t-QVARMA(1,1,4) a more flexible convergence to the long-run effect is allowed than for t-QVARMA(1,1,1). Both figures indicate that federal funds effective rate shocks have negative short-run effects on the US inflation rate, while the same shocks have positive long-run effects on the US inflation rate. The total effects of federal funds effective rate shocks on the US inflation rate are positive, after the first year of the future, approximately. US inflation rate shocks have negative short-run effects on the federal funds effective rate, while the same shocks have positive long-run effects on the federal funds effective rate. The total effects of US inflation rate shocks on the federal funds effective rate are positive, after the second month of the future.

6. Limiting special case

6.1. The Gaussian-QVARMA model

If $\nu \to \infty$ for the t-QVARMA(p,q,r) model, then $v_t \sim t_K(0, \Sigma, \nu) \to_d N_K(0, \Sigma)$ and $u_t = v_t[1+(v_t'\Sigma_v^{-1}v_t)/\nu]^{-1} \to_p v_t$. This provides the following linear Gaussian-QVARMA(p,q,r) model:

$$y_t = c^* + \mu_t^* + \mu_t^\dagger + v_t \tag{6.1}$$

$$\mu_t^* = \sum_{i=1}^{p} \Phi_i^* \mu_{t-i}^* + \sum_{j=1}^{q} \Psi_j^* v_{t-j} \tag{6.2}$$

$$\mu_t^\dagger = \mu_{t-1}^\dagger + \sum_{l=1}^{r} \Psi_l^\dagger v_{t-l} \tag{6.3}$$

$$v_t \sim N_S(0_{S\times 1}, \Sigma) = N_S[0, \Omega^{-1}(\Omega^{-1})'] \text{ i.i.d.} \tag{6.4}$$

where K^* variables are $I(0)$ and $K^\dagger$ co-integrated variables are $I(1)$. The IRFs of the Gaussian-QVARMA model are: $\text{IRF}_{j,t} = \text{IRF}_{j,t}^* + \text{IRF}_{j,t}^\dagger + \text{IRF}_{j,t}^v$, where the short-run effects are $\text{IRF}_{j,t}^* = J_1(\Phi^*)^{j-1}\Psi^* J_2'\Omega^{-1}$ for $j = 1, \ldots, \infty$, the long-run effects are given by:

$$\begin{aligned}
\text{IRF}_{1,t}^\dagger &= \partial\mu_{t+1}^\dagger/\partial\epsilon_t = \Psi_1^\dagger\Omega^{-1} \\
\text{IRF}_{2,t}^\dagger &= \partial\mu_{t+2}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \Psi_2^\dagger)\Omega^{-1} \\
&\vdots \\
\text{IRF}_{K-1,t}^\dagger &= \partial\mu_{t+K-1}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \ldots + \Psi_{K-1}^\dagger)\Omega^{-1} \\
\text{IRF}_{j,t}^\dagger &= \partial\mu_{t+j}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \ldots + \Psi_K^\dagger)\Omega^{-1} \quad \text{for} \quad j \geq K
\end{aligned} \tag{6.5}$$

and the contemporaneous effects are $\text{IRF}_{j,t}^v = \Omega^{-1}$ for $j = 0$.

6.2. Empirical application

As an example, the following Gaussian-QVARMA(1,1,1) model is estimated:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \mu^*_{1,t} \\ \mu^*_{2,t} \end{bmatrix} + \begin{bmatrix} \mu^\dagger_{1,t} \\ \mu^\dagger_{2,t} \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \end{bmatrix} \tag{6.6}$$

$$\begin{bmatrix} \mu^*_{1,t} \\ \mu^*_{2,t} \end{bmatrix} = \begin{bmatrix} \Phi^*_{1,11} & \Phi^*_{1,12} \\ \Phi^*_{1,21} & \Phi^*_{1,22} \end{bmatrix} \begin{bmatrix} \mu^*_{1,t-1} \\ \mu^*_{2,t-1} \end{bmatrix} + \begin{bmatrix} \Psi^*_{1,11} & \Psi^*_{1,12} \\ \Psi^*_{1,21} & \Psi^*_{1,22} \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \end{bmatrix} \tag{6.7}$$

$$\begin{bmatrix} \mu^\dagger_{1,t} \\ \mu^\dagger_{2,t} \end{bmatrix} = \begin{bmatrix} \mu^\dagger_{1,t-1} \\ \mu^\dagger_{2,t-1} \end{bmatrix} + \begin{bmatrix} \Psi^\dagger_{1,11} & \Psi^\dagger_{1,12} \\ \kappa\Psi^\dagger_{1,11} & \kappa\Psi^\dagger_{1,12} \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \end{bmatrix} \tag{6.8}$$

$$v_t \sim N_2 \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega^{-1}_{11} & 0 \\ \Omega^{-1}_{21} & \Omega^{-1}_{22} \end{bmatrix} \times \begin{bmatrix} \Omega^{-1}_{11} & \Omega^{-1}_{21} \\ 0 & \Omega^{-1}_{22} \end{bmatrix} \right\} \text{ i.i.d.} \tag{6.9}$$

The specification of $\Psi^\dagger_1$ ensures that $R = 1$. For the first observation of the sample, μ^*_t is initialized by using a 2×1 vector of zeros. For the first observation of the sample, $\mu^\dagger_t$ is initialized by using parameter vector $(\mu^\dagger_{0,11}, \kappa\mu^\dagger_{0,11})'$. This initialization of $\mu^\dagger_t$ maintains the co-integration relation for the initialization. The co-integration vector is given by $(-\kappa, 1)$ and $\mu^\dagger_{1,t}$ defines a common factor for the federal funds effective rate and US inflation rate variables.

The ML parameter estimates for Gaussian-QVARMA(1,1,1) are presented in Table 2. Those estimates indicate that both short-run and long-run dynamic effects are significant, as several elements of Φ^*_1, Ψ^*_1 and $\Psi^\dagger_1$ are significant. Parameter κ is significantly different from zero, indicating a significant co-integration relationship between federal funds effective rate and US inflation rate. In Table 2, the LL, AIC, BIC and HQC model selection criteria are also presented, which indicate that the likelihood-based statistical performances of the t-QVARMA specifications are superior to the statistical performance of the Gaussian-QVARMA model.

In addition, the IRF estimates for the Gaussian-QVARMA(1,1,1) specification are studied in Fig. 6. With respect to the signs of the short-run and long-run dynamic interaction effects between federal funds effective rate and US inflation rate, the same results as for t-QVARMA are found (see Panels (a), (c), (f) and (g) in Figs. 4 to 6). With respect to total effects between federal funds effective rate and US inflation rate, no significant effects are found. The most important conclusion from Fig. 6 is that the IRF estimates are much less precise for Gaussian-QVARMA than for t-QVARMA. This is probably due to the distortions of the extreme observations of the dataset in the measurements of dynamic interaction effects.

7. Classical alternative

7.1. The Gaussian-VAR(p) model

The reduced-form representation of the classical Gaussian-VAR(p) model is:

$$y_t = c + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + v_t \tag{7.1}$$

where c ($K \times 1$), $\Phi_1, \ldots, \Phi_p$ (each $K \times K$) are time-invariant parameters. For the reduced-form error term, $v_t \sim N_K(0_{K\times 1}, \Sigma) = N_K[0, \Omega^{-1}(\Omega^{-1})']$ i.i.d. is considered, where Ω^{-1} ($K \times K$) includes time-invariant parameters. It is assumed that the dependent variables are either $I(0)$ or co-integrated $I(1)$. Co-integration is measured by using the following VECM representation:

$$\Delta y_t = \Pi(y_{t-1} - c) + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + v_t \tag{7.2}$$

where Π ($K \times K$), c ($K \times 1$), $\Gamma_1, \ldots, \Gamma_{p-1}$ (each $K \times K$) are constant parameters (e.g. Lütkepohl, 2005). For the first p observations, y_t is initialized by using zeros. The co-integration relationships are defined by imposing a rank restriction on matrix Π. Denote rank(Π) $= R$, where $1 \leq R < K$. It is assumed that rank(Π) $= R$ within the following VECM form:

$$\Delta y_t = \alpha\beta'(y_{t-1} - c) + \Gamma_1 \Delta y_{t-1} + \ldots + \Gamma_{p-1} \Delta y_{t-p+1} + v_t \tag{7.3}$$

where α and β (each $K \times R$) are time-invariant parameter matrices, and where β includes the co-integration vectors and the first R rows of β to the identity matrix I_R. The Gaussian-VECM model is estimated by using the ML method (Johansen, 1988, 1991, 1995).

Given the VECM parameter estimates, the VAR parameters are expressed as follows:

$$
\begin{aligned}
\Phi_1 &= \Pi + I_K + \Gamma_1 \\
\Phi_i &= \Gamma_i - \Gamma_{i-1} \quad \text{for} \quad i = 2, \dots, p-1 \\
\Phi_p &= -\Gamma_{p-1}
\end{aligned}
\tag{7.4}
$$

By using those VAR parameters, the following matrix is defined:

$$
\Phi = \begin{bmatrix}
\Phi_1 & \Phi_2 & \cdots & \Phi_{p-1} & \Phi_p \\
I_K & 0_{K\times K} & \cdots & \cdots & 0_{K\times K} \\
0_{K\times K} & I_K & 0_{K\times K} & \cdots & \cdots \\
\cdots & \cdots & \cdots & \cdots & \cdots \\
0_{K\times K} & \cdots & 0_{K\times K} & I_K & 0_{K\times K}
\end{bmatrix}_{(Kp\times Kp)}
\tag{7.5}
$$

for which the maximum modulus of the eigenvalues is C_1. Variable y_t is covariance stationary for $C_1 < 1$. Furthermore, the IRF for the VAR model is given by: $\text{IRF}_{j,t} = \partial y_{t+j}/\partial \epsilon_t = J_1 \Phi^j J_1' \Omega^{-1}$ for $j = 0, \dots, \infty$ where $J_1 = (I_K, 0_{K\times K}, \cdots, 0_{K\times K})$ is $(K \times Kp)$ (e.g. Lütkepohl, 2005). This IRF of total effects is decomposed to short-run, long-run and contemporaneous effects, as follows:

$$
\text{IRF}_{j,t} = \text{IRF}^*_{j,t} + \text{IRF}^\dagger_{j,t} + \text{IRF}^v_{j,t} \quad \text{for} \quad j = 0, 1, \dots, \infty \tag{7.6}
$$

which is computed according to the Granger-representation theorem (e.g. Lütkepohl, 2005).

Finally, the use of the multivariate t-distribution in the VAR framework is also studied in this paper, because t-VECM is more robust to outliers than Gaussian-VECM (e.g. Lucas, 1997). Nevertheless, the estimation results for federal funds rate and US inflation rate indicate that the IRF confidence intervals are significantly wider for t-VECM than for Gaussian-VECM, which is due to the low degrees of freedom parameter estimates for t-VECM. As a consequence, only the Gaussian-VECM estimation results are reported in this paper.

7.2. Empirical application

In this section, the ML estimates for Gaussian-VAR(1) and Gaussian-VAR(4) are presented in Table 2 (the VAR lag-orders are chosen based on LL-based performance metrics). Parameters α and β are significant. Hence, a co-integration relationship is suggested for federal funds effective rate and US inflation rate. Several elements of Γ_i are also significant. The LL, AIC, BIC and HQC metrics indicate that Gaussian-VAR(4) is superior to Gaussian-VAR(1), and that the outlier-sensitive Gaussian-VAR models are inferior to t-QVARMA. IRFs for Gaussian-VAR(4) are presented in Fig. 7. There are negative short-run, positive long-run and positive total effects between federal fund effective rate and US inflation rate, similar to t-QVARMA(1,1,4). The IRF estimates for Gaussian-VAR(4) are much less precise than the same estimates for t-QVARMA(1,1,4), which motivates the practical use of the outlier-robust score-driven model.

8. Predictive accuracy

In this section, the predictive accuracies of t-QVARMA and Gaussian-VAR are compared, by performing out-of-sample predictions of federal funds effective rate and US inflation rate for the period of January 2010 to January 2019 ($P = 109$ periods; forecasting window). Two specifications are used for each of those models: (i) t-QVARMA(1,1,1) and t-QVARMA(1,1,4); (ii) Gaussian-VAR(1) and Gaussian-VAR(4). In the forecasting exercise, 103 rolling windows are used for parameter estimation. Those windows are from the period of July 1954 to December 2009 until the period of February 1963 to July 2018 ($Q = 666$ periods are included in each rolling window). After each estimation, $h = 1, \ldots, 6$ month ahead out-of-sample forecasts are computed. For t-QVARMA, the h-step ahead forecast μ_{t+h} is given by the following formulas:

$$\mu^*_{t+h} = \sum_{i=1}^{p} \Phi^*_i \mu^*_{t+h-i} + \sum_{j=1}^{q} \Psi^*_j u_{t+h-j} \tag{8.1}$$

$$\mu^\dagger_{t+h} = \mu^\dagger_{t+h-1} + \sum_{l=1}^{r} \Psi^\dagger_l u_{t+h-l} \tag{8.2}$$

$$\mu_{t+h} = \mu^*_{t+h} + \mu^\dagger_{t+h} \tag{8.3}$$

In these equations, the estimates of μ^*_{t+h-i}, u_{t+h-j}, $\mu^\dagger_{t+h-1}$ and u_{t+h-l} from the rolling window are used for forecasting if those are available from the rolling window. If the estimates of μ^*_{t+h-i} and $\mu^\dagger_{t+h-1}$ are not available from the rolling window, then those terms are replaced by their forecasts. If the estimates of u_{t+h-j} and u_{t+h-l} are not available from the rolling window, then those terms are omitted from the forecasting formulas (Harvey, 2013). For Gaussian-VAR, the h-step ahead forecast μ_{t+h} is given by:

$$\mu_{t+h} = c + \sum_{i=1}^{p} \Phi_i y_{t+h-i} \tag{8.4}$$

In this equation, y_{t+h-i} is used for forecasting whether its observations are available from the rolling window. Otherwise, y_{t+h-i} is replaced by its forecasts.

Given the out-of-sample forecasts, the loss function is computed for each month of the forecasting window. In this paper, the loss function is defined as: $L_{t,h} = (y_{t+h} - \mu_{t+h})^2$ for $h = 1, \ldots, 6$. The mean of $L_{t,h}$ for the forecasting window is presented in Table 3, where the lower loss function estimate for the t-QVARMA specifications and the lower loss function estimate for the Gaussian-VAR specifications are highlighted. For the highlighted specifications, the difference of the loss functions is computed for each period of the forecasting window: $\Delta L_{t,h} = L_{1,t,h} - L_{2,t,h}$, where $L_{1,t,h}$ denotes the loss function of the highlighted t-QVARMA specification and $L_{2,t,h}$ denotes the loss function of the highlighted Gaussian-VAR specification. The statistical significance of $\Delta L_{t,h}$ is studied by using the following tests: (i) the Diebold–Mariano test (Diebold and Mariano, 1995) of average predictive accuracy; (ii) the Giacomini–Rossi fluctuation test (Giacomini and Rossi, 2010) of dynamic predictive accuracy.

(i) The null hypothesis of the Diebold–Mariano test is the equal average predictive accuracy of both models for the forecasting window. This test is performed by using the OLS-HAC (ordinary least squares, heteroscedasticity and autocorrelation consistent) estimator (Newey and West, 1987) for the linear regression model $\Delta L_{t,h} = c_h + \epsilon_t$ for the forecasting window. If the estimate of c_h is significantly negative, then the predictive accuracy of t-QVARMA, on average, is superior to the predictive accuracy of Gaussian-VAR. The t-ratios of the estimates of c_h are reported in Table 3. Those results suggest that the forecast performance of t-QVARMA, on average, is superior to the forecast performance of Gaussian-VAR for all cases; with the exception of $h = 1$ for the US inflation rate, for which there is no significance difference.

(ii) The null hypothesis of the Giacomini–Rossi fluctuation test is $E(\Delta L_{t,h}) = 0$ for all $t = Q + h, \ldots, T$. For each period, the following test statistic is used:

$$\hat{\sigma}^{-1} m^{-1/2} \sum_{j=t-m/2}^{t+m/2-1} \Delta L_{j,h} \tag{8.5}$$

for $h = 1, \ldots, 6$, where $\hat{\sigma}$ is the HAC estimator of the standard deviation of residuals in the linear regression $\Delta L_{t,h} = c_h + \epsilon_t$, and $m = 12$ months in the application of the present paper. For $\Delta L_{j,h}$, the loss functions of the same specifications are used as for the Diebold–Mariano test. For this choice, $m/P = 0.1101$, which is a parameter for the 10% critical value for the one-sided test; that critical value is 2.8762. For the approximation of the critical value, a quadratic interpolation was used from the critical values that are reported in the work of Giacomini and Rossi (2010). The Giacomini–Rossi predictive accuracy test is performed for each $t = Q + h, \ldots, T$ within the forecasting window. The evolution of the test statistic of Eq. (8.5) is compared with the critical value in Figs. 8 and 9. For the federal funds effective rate, the predictive accuracy of the t-QVARMA specification is superior to that of the Gaussian-VAR specification for the periods of December 2010-November 2011 and March 2012-December 2014 (these periods are approximate, since the exact period of forecast superiority depends on h; see Figs. 8 and 9). For the US inflation rate, the predictive accuracy of the t-QVARMA specification is superior to that of the Gaussian-VAR specification for the period of December 2010-December 2011 (this period is approximate, since the exact period of forecast superiority depends on h; see Figs. 8 and 9).

In summary, according to the Giacomini-Rossi fluctuation test results, the predictive accuracy of t-QVARMA is significantly superior to the predictive accuracy of Gaussian-VAR for several months within the forecasting window. On the other hand, the predictive accuracy of

Gaussian-VAR is never significantly superior to the predictive accuracy of t-QVARMA. These fluctuation test results, combined with the results for the Diebold–Mariano test, provide an additional support for the practical use of the t-QVARMA model.

9. Conclusions

In this paper, the t-QVARMA model has been introduced, which is a new outlier-robust framework for co-integrated time series variables. The measurements of dynamic interaction effects for t-QVARMA are less distorted by outliers than the same measurements for the classical Gaussian multivariate alternatives. The t-QVARMA model has included both $I(0)$ and co-integrated $I(1)$ variables. The reduced-form and the structural-form representations of t-QVARMA have been presented and tools have been provided for IRF analysis. For t-QVARMA, the conditions of the asymptotic properties of ML are presented. Furthermore, it has been shown that a limiting special case of t-QVARMA is the linear Gaussian-QVARMA model.

An application case study to US macroeconomic data has been performed, by using monthly time series data on the federal funds effective rate and the US inflation rate variables for the period of July 1954 to January 2019. Those variables are integrated of order one and are co-integrated. The t-QVARMA model has been estimated by using the ML method, and the conditions of the asymptotic properties of the ML estimator have been studied. The estimates of different t-QVARMA specifications have been compared with the estimates of multivariate linear Gaussian alternatives. Our results suggest that the statistical performance of the t-QVARMA model is superior to the statistical performance of the classical Gaussian-VAR model. The out-of-sample multi-step ahead predictive accuracies of t-QVARMA and Gaussian-VAR have been compared for the period of January 2010 to January 2019. According to the results, the multi-step ahead forecast performance of t-QVARMA is superior to that of Gaussian-VAR. These results may motivate the practical use of the t-QVARMA model in applications, which involve macroeconomic time series data with common stochastic trends and outliers.

Acknowledgements

A previous version of this paper was presented at the GESG Research Seminar of Universidad Francisco Marroquín (20 June 2019, Guatemala City). The authors are thankful for the helpful comments of Juan Carlos Arriaza Herrera, Astrid Ayala, Matthew Copley, Carla Hess, Søren Johansen and seminar participants. Blazsek and Licht acknowledge funding from Universidad Francisco Marroquín. Escribano acknowledges funding from Ministerio de Economía, Industria y Competitividad (ECO2016-00105-001 and MDM 2014-0431) and Comunidad de Madrid (MadEco-CM S2015/HUM-3444).

References

Blasques, F., Koopman, S.J., Lucas, A., 2015. Information-theoretic optimality of observation-driven time series models for continuous responses. Biometrika 102 (2): 325–343. doi: 10.1093/biomet/asu076.

Blazsek, S., Escribano, A., Licht, A., 2018a. Seasonality detection in small samples using score-driven nonlinear multivariate dynamic location models. University Carlos III of Madrid, Department of Economics, Working Paper 18-09. https://orff.uc3m.es/bitstream/handle/10016/27483/we1809.pdf.

Blazsek, S., Escribano, A., Licht, A., 2018b. Seasonal quasi-vector autoregressive models with an application to crude oil production and economic activity in the United States and Canada. University Carlos III of Madrid, Department of Economics, Working Paper 18-10. https://e-archivo.uc3m.es/bitstream/handle/10016/27484/we1810.pdf.

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. Journal of Econometrics 31 (3): 307–327. doi: 10.1016/0304-4076(86)90063-1.

Bosco, B., Parisio, L., Pelagatti, M., Baldi, F., 2010. Long-run relations in European electricity prices. Journal of Applied Econometrics 25 (5): 805–832. doi: 10.1002/jae.1095.

Box, G.E.P., Jenkins, G.M., 1970. Time Series Analysis, Forecasting and Control. Holden-Day, San Francisco.

Cox, D.R., Gudmundsson, G., Lindgren, G., Bondesson, L., Harsaae, E., Laake, P., Juselius, K., Lauritzen, S.L., 1981. Statistical analysis of time series: some recent developments. Scandinavian Journal of Statistics 8 (2): 93–115. https://www.jstor.org/stable/4615819.

Creal, D., Koopman, S.J., Lucas, A., 2011. A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. Journal of Business and Economic Statistics, 29 (4), 552–563. doi: 10.1198/jbes.2011.10070.

Creal, D., Koopman, S.J., Lucas, A., 2013. Generalized autoregressive score models with applications. Journal of Applied Econometrics 28 (5): 777–795. doi: 10.1002/jae.1279.

de Jong, R.M., Amsler, C., Schmidt, P., 2007. A robust version of the KPSS test based on indicators. Journal Econometrics 137 (2): 311–333. doi: 10.1016/j.jeconom.2006.01.001.

Dickey, D.A., Fuller, W.A., 1979. Distribution of the estimators for autoregressive time series with a unit root. Journal of the American Statistical Association 74 (366): 427–431. doi: 10.2307/2286348.

Diebold, F.X., Mariano, R.S., 1995. Predictive accuracy. Journal of Business & Economic Statistics 13 (3): 253–263. doi: 10.1080/07350015.1995.10524599.

Elliott, G., Rothenberg, T.J., Stock, J.H., 1996. Efficient tests for an autoregressive unit root. Econometrica 64 (4): 813–836. doi: 10.2307/2171846.

Engle, R.F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. Econometrica 50 (4): 987–1007. doi: 10.2307/1912773.

Engle, R.F., Granger, C.W.J., 1987. Co-integration and error correction: representation, estimation, and testing. Econometrica 55 (2): 251–276. doi: 10.2307/1913236.

Escribano, A., Peña, J.L., Villaplana, P., 2011. Modeling electricity prices: international evidence. Oxford Bulletin of Economics and Statistics 73 (5): 622–650. doi: 10.1111/j.1468-0084.2011.00632.x.

Franses, P.H., Lucas, A., 1998. Outlier detection in cointegration analysis. Journal of Business and Economic Statistics 16 (4): 459–468. doi: 10.1080/07350015.1998.10524785.

Giacomini, R., Rossi, B., 2010. Forecast comparisons in unstable environments. Journal of Applied Econometrics 25 (4): 595–620. doi: 10.1002/jae.1177.

Granger, C.W.J., 1981. Some properties of time series data and their use in econometric model specification. Journal of Econometrics 16 (1): 121–130. doi: 10.1016/0304-4076(81)90079-8.

Harvey, A.C., 2013. Dynamic Models for Volatility and Heavy Tails. Cambridge University Press, Cambridge.

Harvey, A.C., Chakravarty, T., 2008. Beta-t-(E)GARCH. Cambridge Working Papers in Economics 0840, Faculty of Economics, University of Cambridge, Cambridge.
http://www.econ.cam.ac.uk/research/repec/cam/pdf/cwpe0840.pdf.

Inoue, A., Kilian, L., 2016. Joint confidence sets for structural impulse responses. Journal of Econometrics 192 (2): 421–432. doi: 10.1016/j.jeconom.2016.02.008.

Johansen, S., 1988. Statistical analysis of cointegration vectors. Journal of Economic Dynamics and Control 12 (2-3): 231–254. doi: 10.1016/0165-1889(88)90041-3.

Johansen, S., 1991. Estimation and hypothesis testing of cointegration vectors in Gaussian vector autoregressive models. Econometrica 59 (6): 1551–1580. doi: 10.2307/2938278.

Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, Oxford.

Kilian, L., Lütkepohl, H., 2017. Structural Vector Autoregressive Analysis. Cambridge University Press, Cambridge.

Lucas, A., 1995a. An outlier robust unit root test with an application to the extended Nelson-Plosser data. Journal of Econometrics 66 (1-2): 153–173. doi: 10.1016/0304-4076(94)01613-5.

Lucas, A., 1995b. Unit root test based on M-estimators. Econometric Theory 11 (2): 331–346. doi: 10.1017/S0266466600009191.

Lucas, A., 1997. Cointegration testing using pseudolikelihood ratio tests. Econometric Theory 13 (2): 149–169. doi: 10.1017/S0266466600005703.

Lucas, A., 1998. Inference on cointegrating ranks using LR and LM tests based on pseudo-likelihoods. Econometric Reviews 17 (2): 185–214. doi: 10.1080/07474939808800411.

Lütkepohl, H., 2005. New Introduction to Multivariate Time Series Analysis. Springer-Verlag, Berlin Heidelberg.

Nelson, D.B., 1991. Conditonal heteroskedasticity in asset returns: a new approach. Econometrica 59 (2): 347–370. doi: 10.2307/2938260.

Newey, W.K., West, K.D., 1987. A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. Econometrica 55 (3): 703–708. doi: 10.2307/1913610.

Nyblom, J., Harvey, A., 2000. Tests of common stochastic trends. Econometric Theory 16 (2): 176–199. doi: 10.1017/S0266466600162024.

Tiao, G.C., Tsay, R.S., 1989. Model specification in multivariate time series. Journal of the Royal Statistical Society 51 (2): 157–213. http://www.jstor.org/stable/2345602.

White, H., 1984. Asymptotic Theory for Econometricians. Academic Press, San Diego.

Appendix

For the t-QVARMA model of this example, $K = 3$, $R = 1$, $p = 1$, $q = 1$ and $r = 1$. With respect to the order of the variables, the first variable is $I(0)$ and the rest of the variables are $I(1)$ and co-integrated. The model is formulated as follows:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} c_1^* \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^* \\ \mu_{2,t}^* \\ \mu_{3,t}^* \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \\ \mu_{3,t}^\dagger \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix} \tag{A.1}$$

$$\begin{bmatrix} \mu_{1,t}^* \\ \mu_{2,t}^* \\ \mu_{3,t}^* \end{bmatrix} = \begin{bmatrix} \Phi_{1,11}^* & \Phi_{1,12}^* & \Phi_{1,13}^* \\ \Phi_{1,21}^* & \Phi_{1,22}^* & \Phi_{1,23}^* \\ \Phi_{1,31}^* & \Phi_{1,32}^* & \Phi_{1,33}^* \end{bmatrix} \begin{bmatrix} \mu_{1,t-1}^* \\ \mu_{2,t-1}^* \\ \mu_{3,t-1}^* \end{bmatrix} + \begin{bmatrix} \Psi_{1,11}^* & \Psi_{1,12}^* & \Psi_{1,13}^* \\ \Psi_{1,21}^* & \Psi_{1,22}^* & \Psi_{1,23}^* \\ \Psi_{1,31}^* & \Psi_{1,32}^* & \Psi_{1,33}^* \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \end{bmatrix} \tag{A.2}$$

$$\begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \\ \mu_{3,t}^\dagger \end{bmatrix} = \begin{bmatrix} \mu_{1,t-1}^\dagger \\ \mu_{2,t-1}^\dagger \\ \mu_{3,t-1}^\dagger \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Psi_{1,22}^\dagger & \Psi_{1,23}^\dagger \\ 0 & \kappa\Psi_{1,22}^\dagger & \kappa\Psi_{1,23}^\dagger \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \end{bmatrix} \tag{A.3}$$

$$v_t \sim t_3 \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega_{11}^{-1} & 0 & 0 \\ \Omega_{21}^{-1} & \Omega_{22}^{-1} & 0 \\ \Omega_{31}^{-1} & \Omega_{32}^{-1} & \Omega_{33}^{-1} \end{bmatrix} \times \begin{bmatrix} \Omega_{11}^{-1} & \Omega_{21}^{-1} & \Omega_{31}^{-1} \\ 0 & \Omega_{22}^{-1} & \Omega_{32}^{-1} \\ 0 & 0 & \Omega_{33}^{-1} \end{bmatrix}, \nu \right\} \text{ i.i.d.} \tag{A.4}$$

The specification of $\Psi_1^\dagger$ ensures that $R = 1$. Variable μ_t^* is initialized by using a 3×1 vector of zeros. Variable $\mu_t^\dagger$ is initialized by using parameter vector $(0, \mu_{0,12}^\dagger, \kappa\mu_{0,12}^\dagger)'$, which maintains the co-integration relation. The co-integration vector for $y_{2,t}$ and $y_{3,t}$ is given by $(-\kappa, 1)'$.

Table 1. Descriptive statistics.

	Federal funds effective rate $y_{1,t}$	US inflation rate $y_{2,t}$
Start date	July 1954	July 1954
End date	January 2019	January 2019
Sample size	775	775
Minimum	0.0700	-2.1195
Maximum	19.1000	13.7642
Mean	4.8105	3.4612
Standard deviation	3.6044	2.6763
Skewness	1.0360	1.4888
Excess kurtosis	1.5328	2.4955
Engle–Granger cointegration test:		
ADF-GLS test with constant on y_t	$-1.4645(0.1340)$	$-1.3683(0.1593)$
ADF-GLS test with constant on residuals	$-2.4570^{**}(0.0136)$	
Nyblom–Harvey common trends test:		
Lag-order $m = 30$	0.1362	
Lag-order $m = 50$	0.1003	
Lag-order $m = 100$	0.0756	
Lag-order $m = 200$	0.0894	
Johansen cointegration test (maximum eigenvalue):		
H_0: rank(Π)=0; H_1: rank(Π)=1	$17.926^{**}(0.0212)$	
H_0: rank(Π)=1; H_1: rank(Π)=2	5.7819(0.2153)	
Lucas outlier-robust co-integration test (t-distribution):		
H_0: rank(Π)=1; H_1: rank(Π)=2	0.0354	

Source of data: Federal Reserve Economic Data, https://fred.stlouisfed.org.

Notes: United States (US); augmented Dickey–Fuller (ADF); generalized least squares (GLS). All variables are measured in percentage points. ADF-GLS with constant on y_t indicates the ADF-GLS test statistic with p-value in parentheses for each dependent variable. ADF-GLS with constant on residuals indicates the ADF-GLS test statistic with p-value in parentheses for the residuals of the linear regression $y_{2t} = \beta_0 + \beta_1 y_{1t} + \epsilon_t$. For the Nyblom–Harvey test the test statistic is reported, by considering the possibility of serial correlation in the error term. According to H_0 of the Nyblom–Harvey test, y_t has one common trend. Results for different lag-orders m are presented. The 90%, 95% and 99% critical values of the Nyblom–Harvey test are 0.1620, 0.2180 and 0.3830, respectively. These critical values indicate that H_0 cannot be rejected for the Nyblom–Harvey test. For the Johansen maximum eigenvalue tests, test statistics and the p-values are reported in parentheses. The Johansen tests suggest that rank(Π) = 1. For the outlier-robust co-integration test of Lucas, the pseudo-likelihood-ratio (PLR) test statistic is reported for the multivariate t-distribution. The estimated degrees of freedom parameter is $\hat{\nu} = 2.23$, for which the critical value is obtained by using a quadratic approximation from the critical values of Lucas (1997, p. 159): $\text{CV}(90\%, \nu = 2.23) = 4.27$. This indicates that H_0 cannot be rejected for the PLR test. ** indicates significance at the 5% level.

Table 2. Parameter estimates and model diagnostics.

	t-QVARMA(1,1,1)	t-QVARMA(1,1,4)	Gaussian-QVARMA(1,1,1)		Gaussian-VAR(1)	Gaussian-VAR(4)
$\Phi^*_{1,11}$	0.5162***(0.0453)	0.5132***(0.1540)	0.8407***(0.1207)	c_1	0.2911(0.4135)	0.6452*(0.3295)
$\Phi^*_{1,12}$	−0.1097*(0.0641)	−0.1285**(0.0581)	−0.0975(0.0998)	c_2	0.1255(0.5412)	0.4489(0.4368)
$\Phi^*_{1,21}$	−0.3416***(0.0532)	−0.4033***(0.0380)	−0.6617**(0.3202)	α_1	−0.0136^{+}(0.0083)	−0.0181**(0.0080)
$\Phi^*_{1,22}$	0.8789***(0.0501)	0.8141***(0.0573)	0.3145^{+}(0.1962)	α_2	0.0162***(0.0053)	0.0153***(0.0048)
$\Psi^*_{1,11}$	−0.6608***(0.1949)	1.3881^{+}(0.8681)	0.4696***(0.1540)	β_1	−1.3486***(0.2812)	−1.3858***(0.2467)
$\Psi^*_{1,12}$	−0.5907***(0.1476)	−3.1473^{+}(1.9832)	−0.8159***(0.2347)	β_2	1.0000	1.0000
$\Psi^*_{1,21}$	−1.4040***(0.1817)	−0.3083(0.5625)	−0.6148***(0.1104)	$\Gamma_{1,11}$	NA	0.4395***(0.0333)
$\Psi^*_{1,22}$	1.2610***(0.1076)	−0.3061(1.2477)	0.5840***(0.1727)	$\Gamma_{1,12}$	NA	0.0796*(0.0430)
$\Psi^\dagger_{1,11}$	2.7844***(0.2398)	0.7974(0.8530)	0.9381***(0.1567)	$\Gamma_{1,21}$	NA	0.0519**(0.0239)
$\Psi^\dagger_{1,12}$	0.6299***(0.1579)	3.0949^{+}(1.9810)	0.9293***(0.2334)	$\Gamma_{1,22}$	NA	0.2619***(0.0331)
$\Psi^\dagger_{2,11}$	NA	1.1755*(0.6946)	NA	$\Gamma_{2,11}$	NA	−0.1371***(0.0358)
$\Psi^\dagger_{2,12}$	NA	−1.6167(1.5833)	NA	$\Gamma_{2,12}$	NA	−0.0398(0.0446)
$\Psi^\dagger_{3,11}$	NA	0.4578**(0.1971)	NA	$\Gamma_{2,21}$	NA	0.0526**(0.0262)
$\Psi^\dagger_{3,12}$	NA	−0.6375^{+}(0.4086)	NA	$\Gamma_{2,22}$	NA	0.1001***(0.0335)
$\Psi^\dagger_{4,11}$	NA	0.1070*(0.0636)	NA	$\Gamma_{3,11}$	NA	−0.0221(0.0340)
$\Psi^\dagger_{4,12}$	NA	−0.1985**(0.0900)	NA	$\Gamma_{3,12}$	NA	−0.0180(0.0428)
κ	0.5805***(0.0504)	0.6499***(0.0391)	0.7199***(0.0527)	$\Gamma_{3,21}$	NA	0.0201(0.0248)
$\mu_{0,1,1}$	0.6217(0.4802)	0.7175***(0.2322)	0.6378(2.9691)	$\Gamma_{3,22}$	NA	−0.0819**(0.0332)
$\mu_{0,2,1}$	NA	0.6369^{+}(0.4009)	NA	Ω^{-1}_{11}	0.4997***(0.0124)	0.4540***(0.0113)
$\mu_{0,3,1}$	NA	−0.4379(0.8954)	NA	Ω^{-1}_{21}	0.0479***(0.0125)	0.0326***(0.0111)
$\mu_{0,4,1}$	NA	0.1702(0.3860)	NA	Ω^{-1}_{22}	0.3593***(0.0091)	0.3381***(0.0080)
Ω^{-1}_{11}	0.3448***(0.0069)	0.3355***(0.0083)	0.4547***(0.0039)			
Ω^{-1}_{21}	0.0152(0.0163)	0.0276**(0.0111)	0.0311**(0.0144)			
Ω^{-1}_{22}	0.3535***(0.0089)	0.3437***(0.0090)	0.3391***(0.0062)			
ν	9.8545***(0.2913)	9.1904***(0.5548)	NA			
C_1	0.9628	0.9365	0.9433	C_1	1.0000	1.0000
C_2	ADF✓	ADF✓	NA	C_2	NA	NA
C_3	1.4446	0.9645	NA	C_3	NA	NA
C_4	2.1104	0.9657	NA	C_4	NA	NA
LL	−0.9362	**−0.8954**	−0.9684	LL	−1.1207	−0.9638
AIC	1.9137	**1.8552**	1.9756	AIC	2.2620	1.9793
BIC	2.0097	**2.0053**	2.0656	BIC	2.3101	2.0995
HQC	1.9506	**1.9130**	2.0102	HQC	2.2805	2.0255

Notes: Quasi-vector autoregressive moving average (QVARMA); not available (NA); log-likelihood (LL); Akaike information criterion (AIC); Bayesian information criterion (BIC); Hannan-Quinn criterion (HQC). Bold LL, AIC, BIC and HQC values indicate the best performing model. Standard errors are in parentheses. ADF✓indicates that for the ADF test with constant, the unit root null hypothesis is rejected for all variables of C_2. $^{+}$, *, ** and *** indicate significance at the 15%, 10%, 5% and 1% levels, respectively.

Table 3. Diebold–Mariano test and average loss functions for the period of January 2010 to January 2019.

Federal funds effective rate, average loss function for $h = 1$		US inflation rate, average loss function for $h = 1$	
t-QVARMA(1,1,1)	0.0038	t-QVARMA(1,1,1)	0.0943
✓t-QVARMA(1,1,4)	0.0033	✓t-QVARMA(1,1,4)	0.0919
✓Gaussian-VAR(1)	0.0415	Gaussian-VAR(1)	0.1258
Gaussian-VAR(4)	0.0551	✓Gaussian-VAR(4)	0.1032
Diebold–Mariano test statistic	$-5.5597^{***}(0.0000)$	Diebold–Mariano test statistic	$-0.8872(0.3770)$

Federal funds effective rate, average loss function for $h = 2$		US inflation rate, average loss function for $h = 2$	
✓t-QVARMA(1,1,1)	0.0117	✓t-QVARMA(1,1,1)	0.2207
t-QVARMA(1,1,4)	0.0129	t-QVARMA(1,1,4)	0.2324
✓Gaussian-VAR(1)	0.1577	Gaussian-VAR(1)	0.3828
Gaussian-VAR(4)	0.3003	✓Gaussian-VAR(4)	0.3335
Diebold–Mariano test statistic	$-5.5861^{***}(0.0000)$	Diebold–Mariano test statistic	$-3.0751^{***}(0.0027)$

Federal funds effective rate, average loss function for $h = 3$		US inflation rate, average loss function for $h = 3$	
✓t-QVARMA(1,1,1)	0.0188	✓t-QVARMA(1,1,1)	0.3829
t-QVARMA(1,1,4)	0.0216	t-QVARMA(1,1,4)	0.3921
✓Gaussian-VAR(1)	0.3414	Gaussian-VAR(1)	0.7280
Gaussian-VAR(4)	0.7432	✓Gaussian-VAR(4)	0.6380
Diebold–Mariano test statistic	$-5.6797^{***}(0.0000)$	Diebold–Mariano test statistic	$-2.8541^{***}(0.0052)$

Federal funds effective rate, average loss function for $h = 4$		US inflation rate, average loss function for $h = 4$	
✓t-QVARMA(1,1,1)	0.0254	✓t-QVARMA(1,1,1)	0.5124
t-QVARMA(1,1,4)	0.0310	t-QVARMA(1,1,4)	0.5326
✓Gaussian-VAR(1)	0.5848	Gaussian-VAR(1)	1.1267
Gaussian-VAR(4)	1.3264	✓Gaussian-VAR(4)	0.9453
Diebold–Mariano test statistic	$-5.5460^{***}(0.0000)$	Diebold–Mariano test statistic	$-2.8062^{***}(0.0060)$

Federal funds effective rate, average loss function for $h = 5$		US inflation rate, average loss function for $h = 5$	
✓t-QVARMA(1,1,1)	0.0372	✓t-QVARMA(1,1,1)	0.6446
t-QVARMA(1,1,4)	0.0464	t-QVARMA(1,1,4)	0.6694
✓Gaussian-VAR(1)	0.8825	Gaussian-VAR(1)	1.6042
Gaussian-VAR(4)	2.0221	✓Gaussian-VAR(4)	1.2873
Diebold–Mariano test statistic	$-5.6782^{***}(0.0000)$	Diebold–Mariano test statistic	$-2.7808^{***}(0.0064)$

Federal funds effective rate, average loss function for $h = 6$		US inflation rate, average loss function for $h = 6$	
✓t-QVARMA(1,1,1)	0.0470	✓t-QVARMA(1,1,1)	0.7950
t-QVARMA(1,1,4)	0.0596	t-QVARMA(1,1,4)	0.8393
✓Gaussian-VAR(1)	1.2270	Gaussian-VAR(1)	2.1674
Gaussian-VAR(4)	2.8133	✓Gaussian-VAR(4)	1.7070
Diebold–Mariano test statistic	$-5.6783^{***}(0.0000)$	Diebold–Mariano test statistic	$-2.9010^{***}(0.0045)$

Notes: *** indicates significance at the 1% level. p-values of the Diebold-Mariano test statistic are reported in parentheses. For each model, the average loss function is defined as the average of $L_{t,h} = (y_{t+h} - \mu_{t+h})^2$ for $t = Q + h, \ldots, T$ within the forecasting window, where $h = 1, \ldots, 6$ and μ_{t+h} represents the h-step ahead forecast of y_{t+h}. ✓indicates the specification with lower loss function estimate from each pair of the specifications that are estimated for t-QVARMA and Gaussian-VAR. The predictive accuracy of the specifications that are marked with ✓are compared by using the Diebold–Mariano test and the Giacomini–Rossi test.

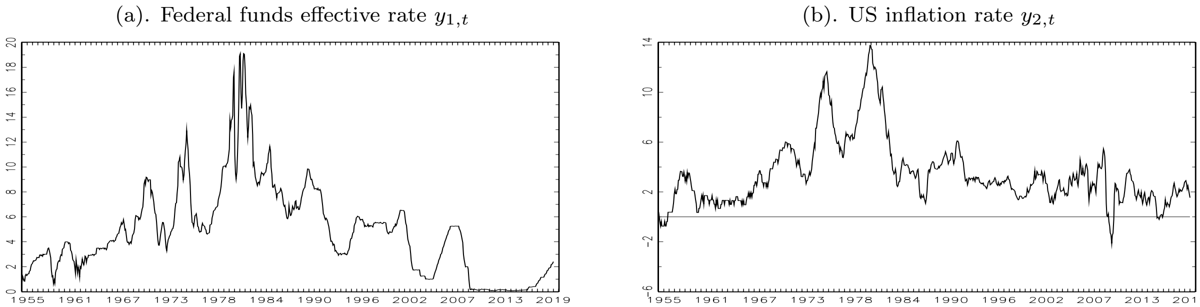

Fig. 1. Evolution of federal funds effective rate and US inflation rate for the period of July 1954 to January 2019.

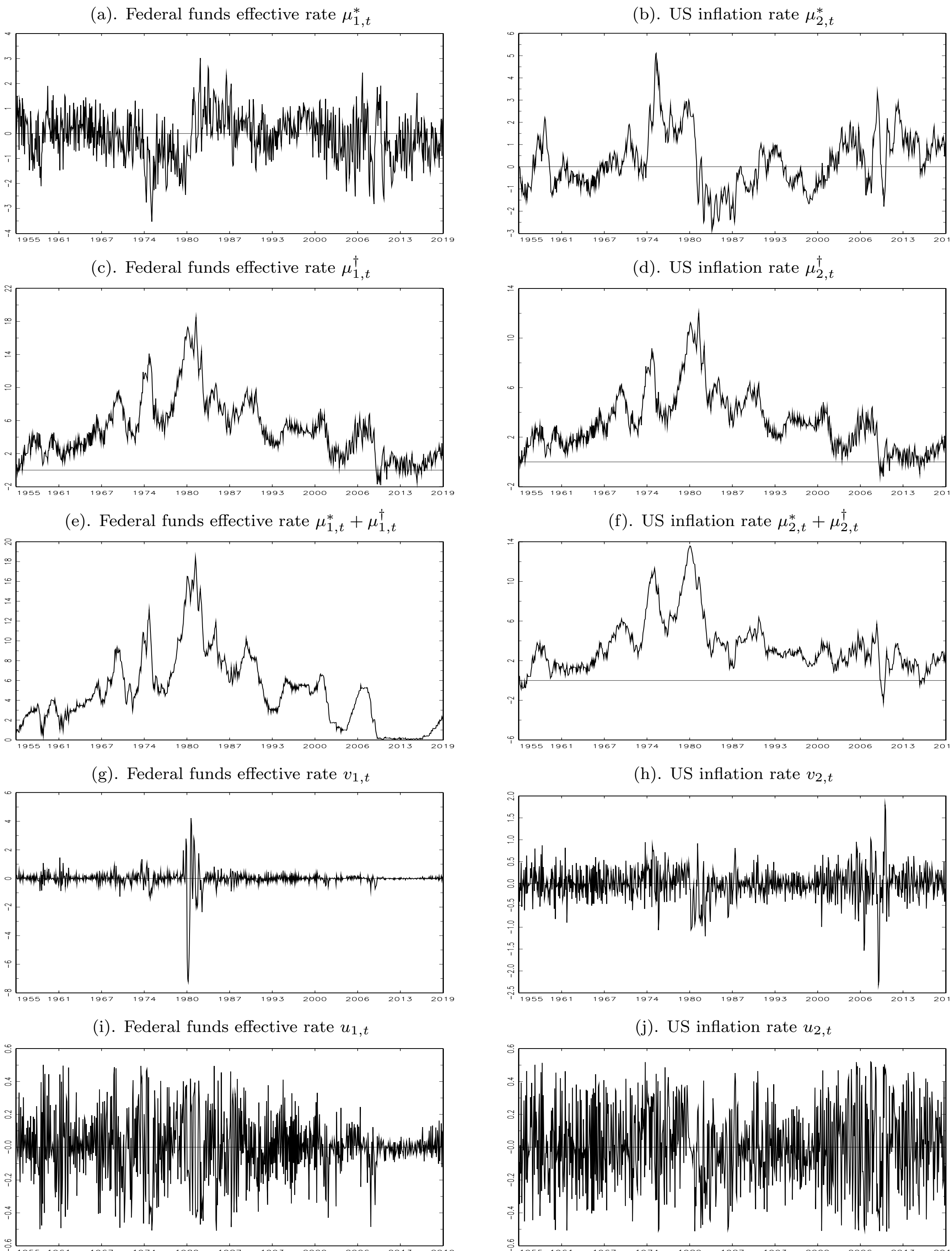

Fig. 2. Location components μ^*_t and $\mu^\dagger_t$, irregular component v_t, and score function u_t for the t-QVARMA(1,1,4) model.

(a). $u_{1,t}$ as a function of $\epsilon_{1,t}$ and $\epsilon_{2,t}$ (b). $u_{2,t}$ as a function of $\epsilon_{1,t}$ and $\epsilon_{2,t}$

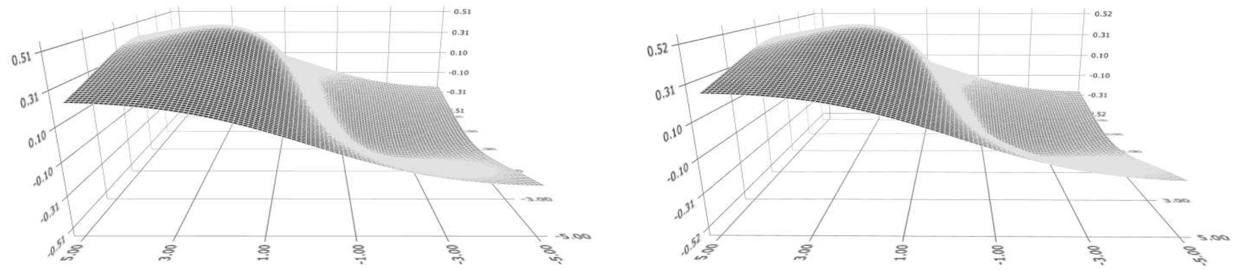

Fig. 3. Robustness to extreme values in the noise for the t-QVARMA(1,1,4) model.

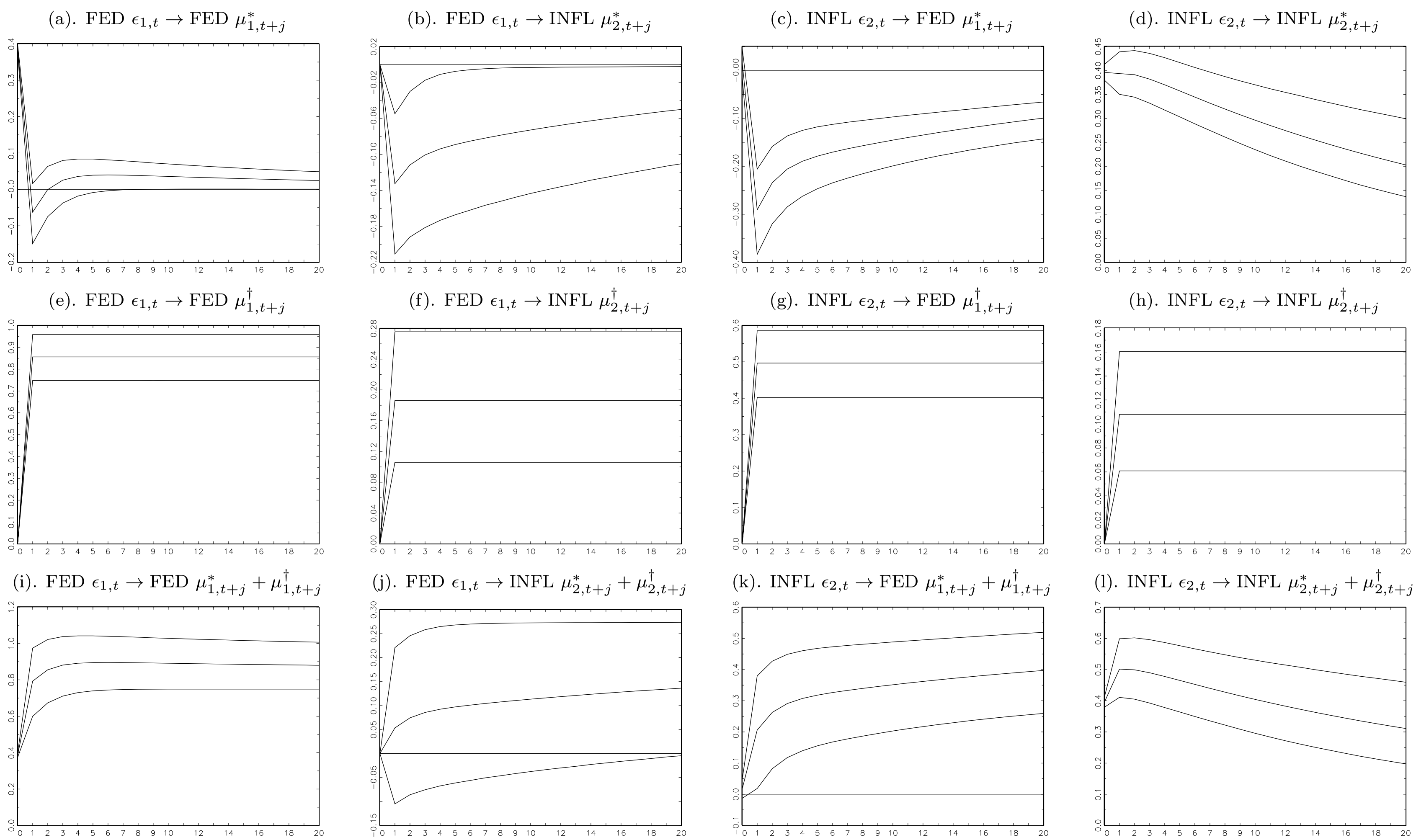

Fig. 4. Impulse response function with 90% confidence interval for the t-QVARMA(1,1,1) model. *Notes*: The IRF confidence interval is estimated by using 10,000 Monte Carlo simulations from the ML estimates. Note that in this figure, FED and INFL refer to federal funds effective rate and US inflation rate, respectively. Panels (a) to (d) present $\text{IRF}^*_{j,t} + \text{IRF}^v_{j,t}$. Panels (e) to (h) present $\text{IRF}^\dagger_{j,t}$. Panels (i) to (l) present $\text{IRF}^*_{j,t} + \text{IRF}^\dagger_{j,t} + \text{IRF}^v_{j,t}$.

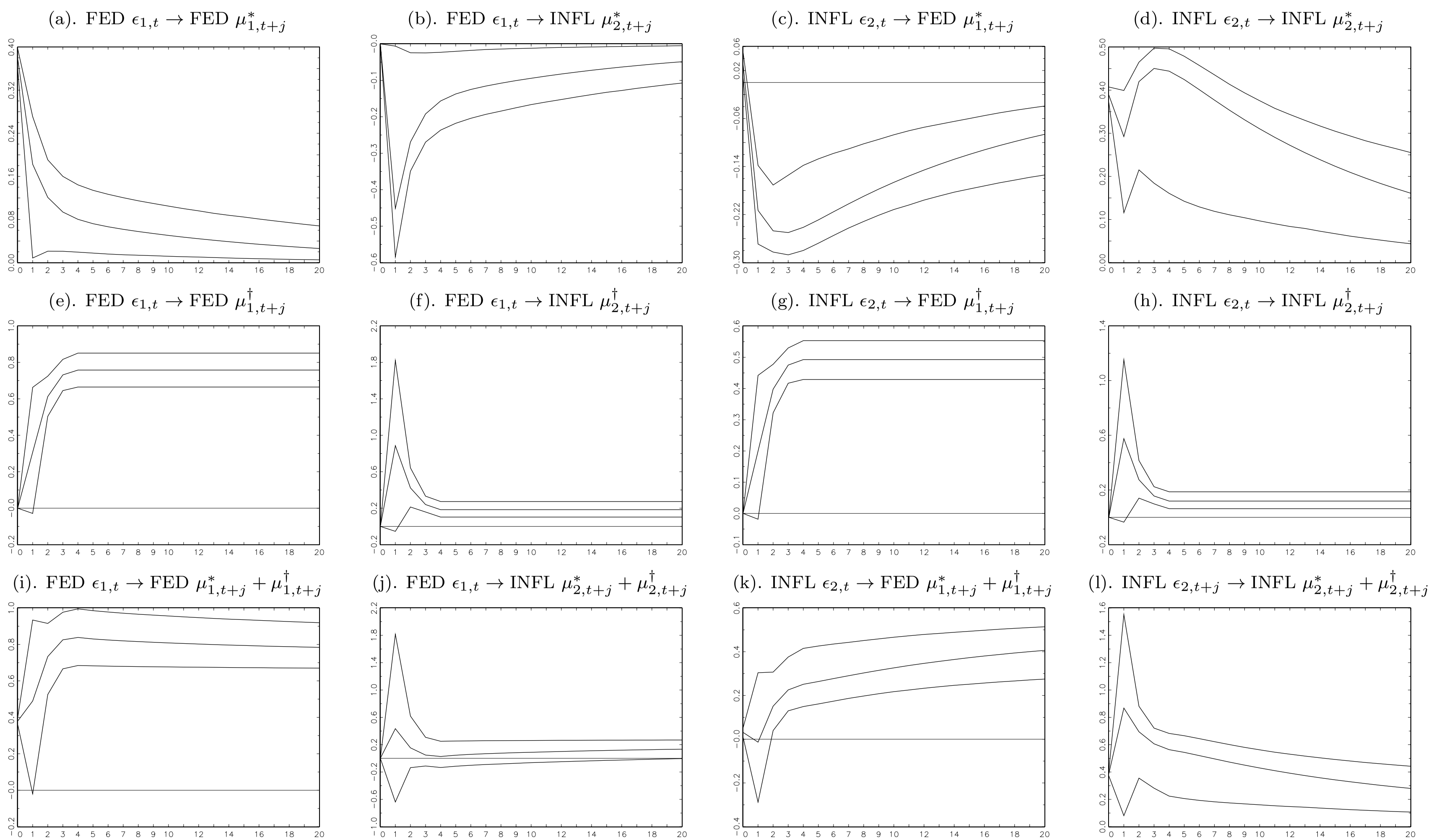

Fig. 5. Impulse response function with 90% confidence interval for the t-QVARMA(1,1,4) model. *Notes*: The IRF confidence interval is estimated by using 10,000 Monte Carlo simulations from the ML estimates. Note that in this figure, FED and INFL refer to federal funds effective rate and US inflation rate, respectively. Panels (a) to (d) present $\text{IRF}^*_{j,t} + \text{IRF}^v_{j,t}$. Panels (e) to (h) present $\text{IRF}^\dagger_{j,t}$. Panels (i) to (l) present $\text{IRF}^*_{j,t} + \text{IRF}^\dagger_{j,t} + \text{IRF}^v_{j,t}$.

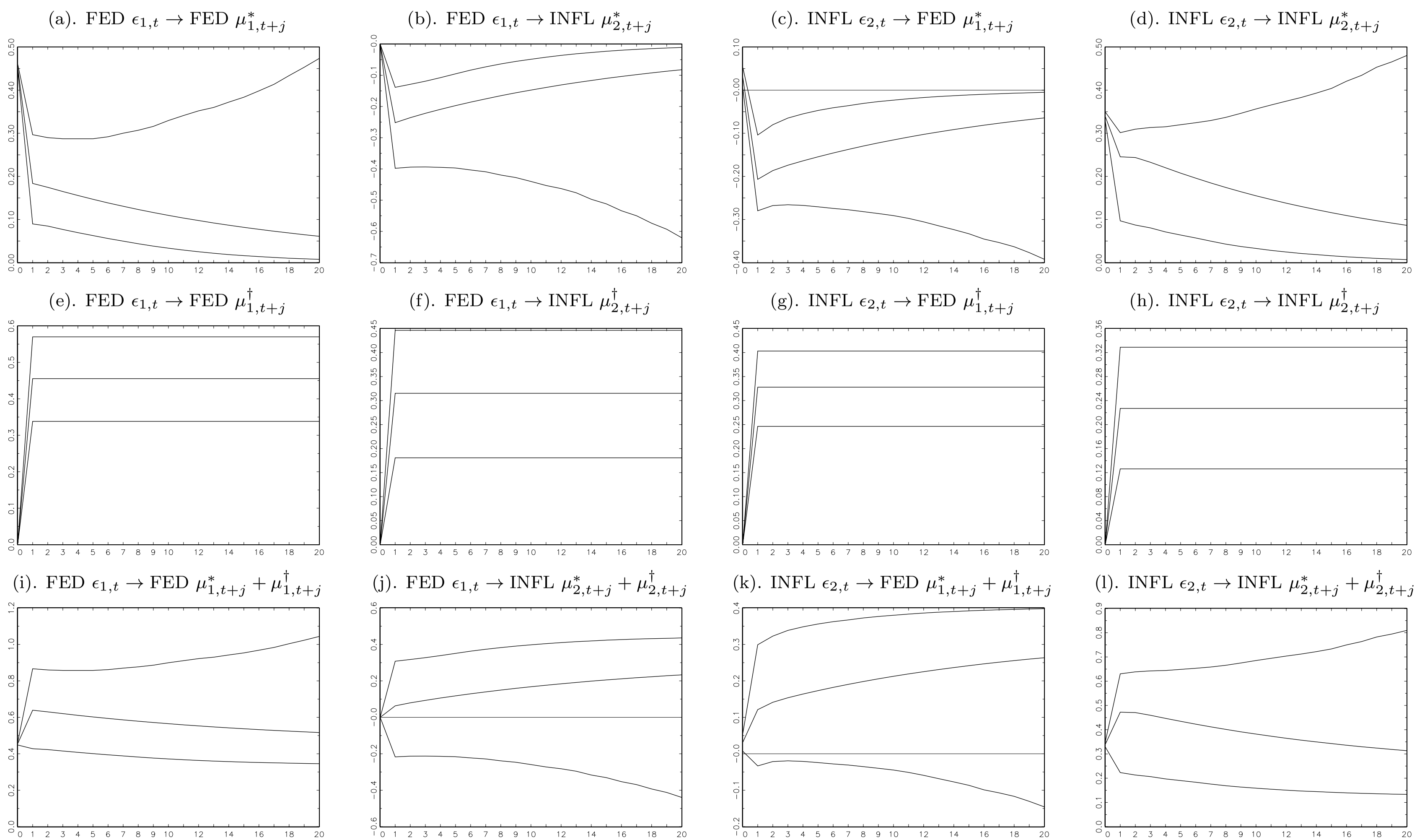

Fig. 6. Impulse response function with 90% confidence interval for the Gaussian-QVARMA(1,1,1) model. *Notes*: The IRF confidence interval is estimated by using 10,000 Monte Carlo simulations from the ML estimates. Note that in this figure, FED and INFL refer to federal funds effective rate and US inflation rate, respectively. Panels (a) to (d) present $\text{IRF}^*_{j,t} + \text{IRF}^v_{j,t}$. Panels (e) to (h) present $\text{IRF}^\dagger_{j,t}$. Panels (i) to (l) present $\text{IRF}^*_{j,t} + \text{IRF}^\dagger_{j,t} + \text{IRF}^v_{j,t}$.

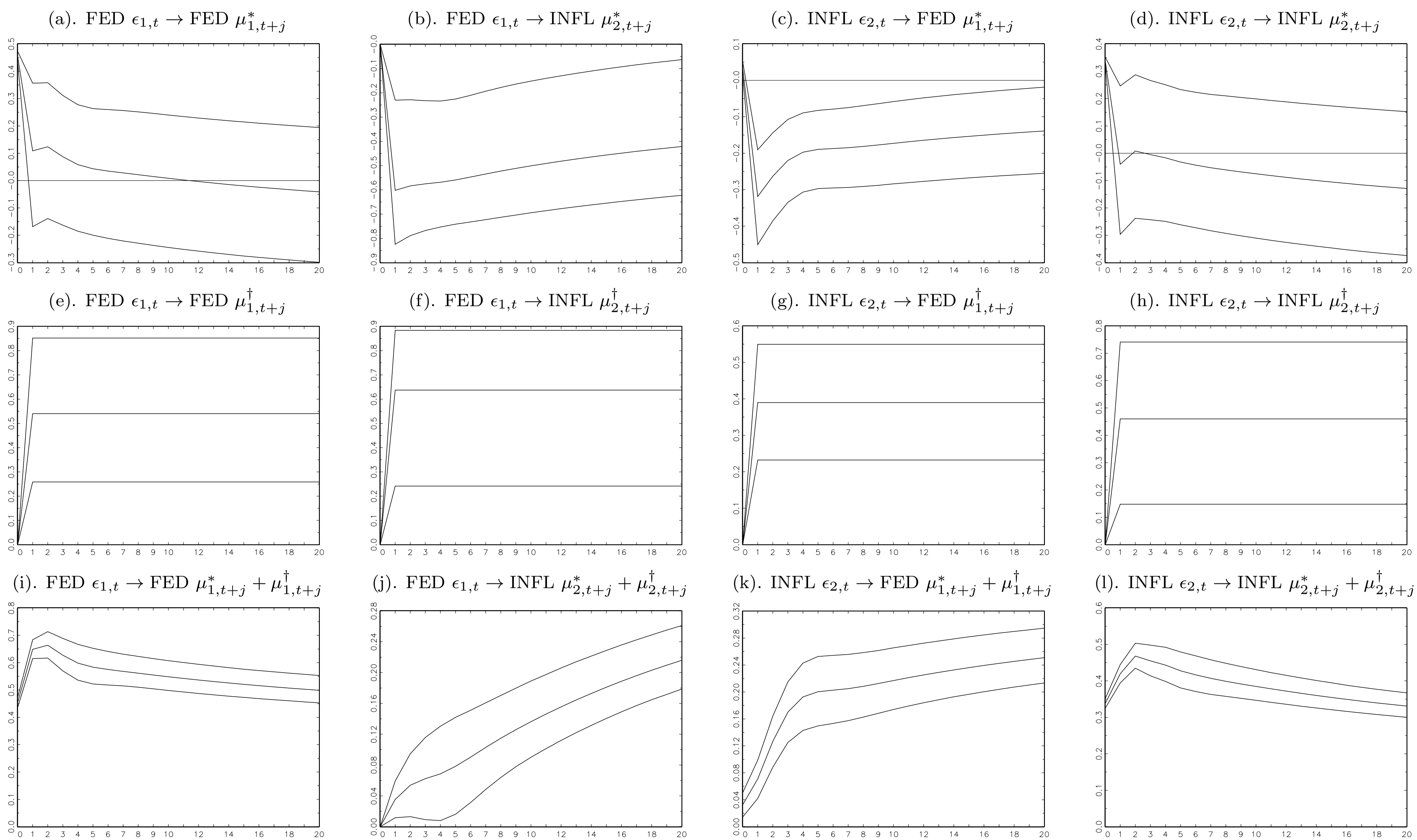

Fig. 7. Impulse response function with 90% confidence interval for the Gaussian-VAR(4) model. *Notes*: The IRF confidence interval is estimated by using 10,000 Monte Carlo simulations from the ML estimates. Note that in this figure, FED and INFL refer to federal funds effective rate and US inflation rate, respectively. Panels (a) to (d) present $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$. Panels (e) to (h) present $\mathrm{IRF}^\dagger_{j,t}$. Panels (i) to (l) present $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^\dagger_{j,t} + \mathrm{IRF}^v_{j,t}$.

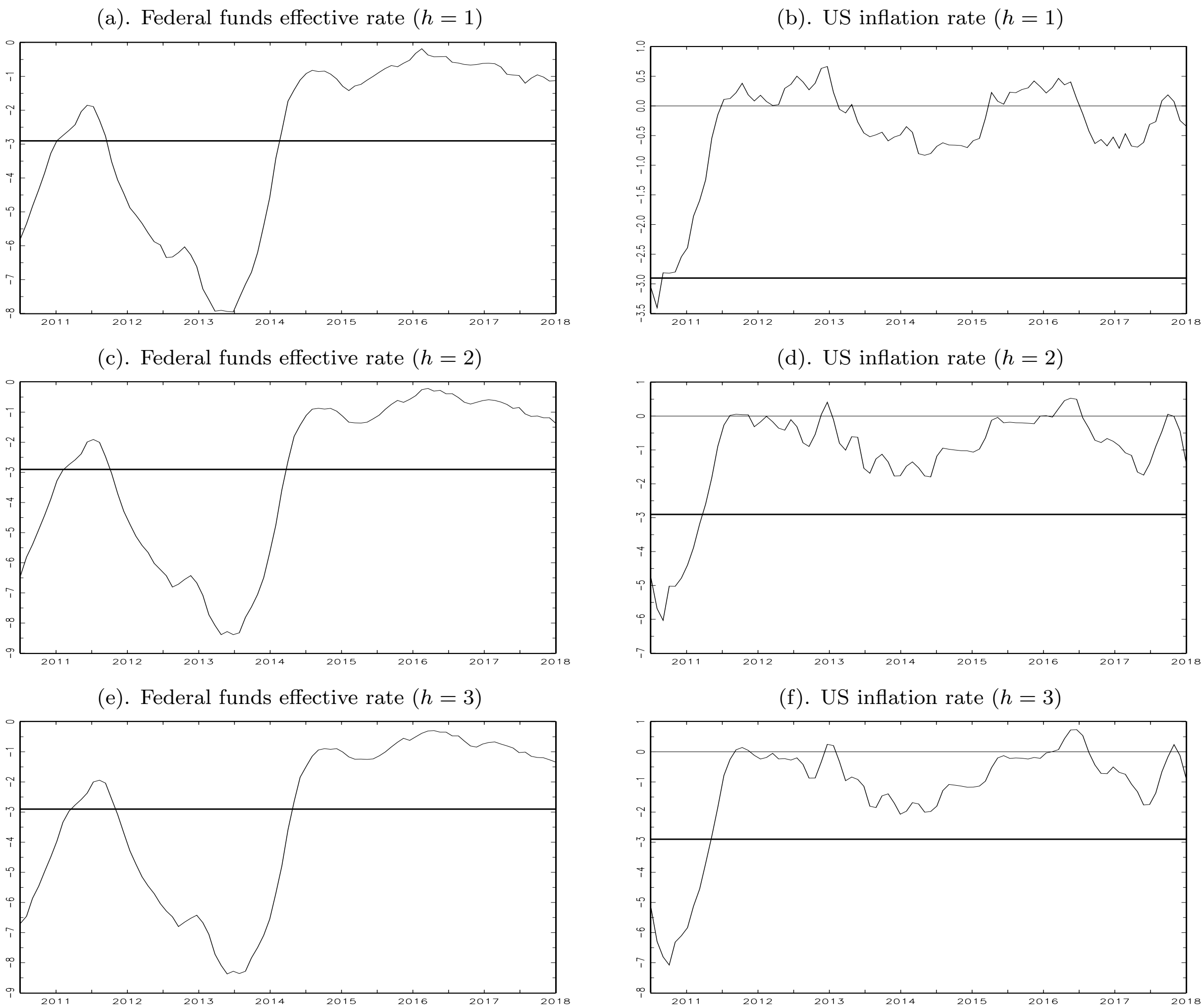

Fig. 8. Giacomini–Rossi fluctuation test for the period of January 2010 to January 2019 ($h = 1, 2, 3$). *Notes*: If the fluctuation test statistic (solid thin) is lower than the 10% critical value for one-sided test (solid thick), then the null hypothesis, $E(\Delta L_{t,h}) = 0$ for all $t = Q + h, \dots, T$, is rejected. A negative sign of the fluctuation test statistic suggests that the predictive accuracy of t-QVARMA is superior to the predictive accuracy of Gaussian-VAR.

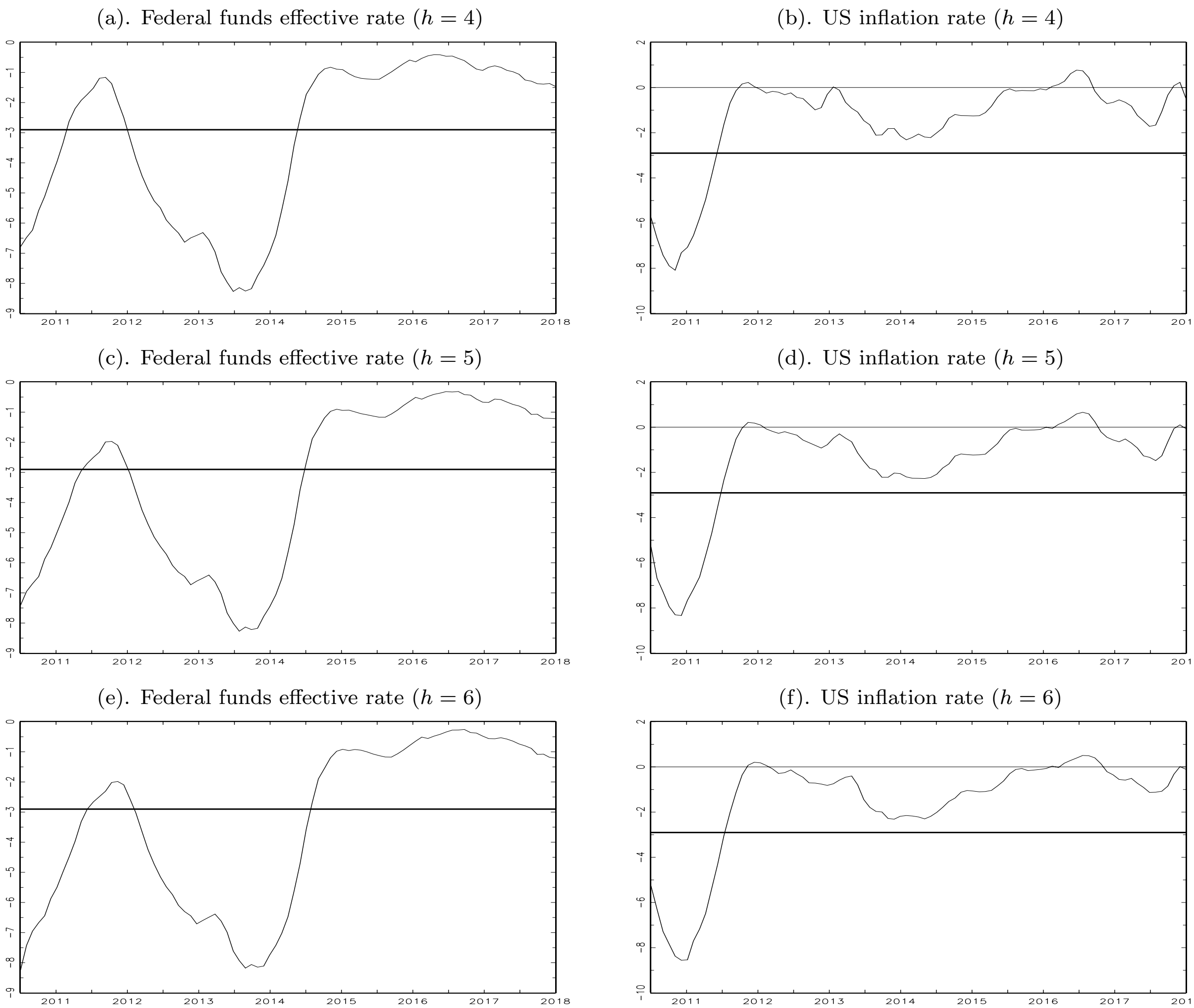

Fig. 9. Giacomini–Rossi fluctuation test for the period of January 2010 to January 2019 ($h = 4, 5, 6$). *Notes*: If the fluctuation test statistic (solid thin) is lower than the 10% critical value for one-sided test (solid thick), then the null hypothesis, $E(\Delta L_{t,h}) = 0$ for all $t = Q + h, \ldots, T$, is rejected. A negative sign of the fluctuation test statistic suggests that the predictive accuracy of t-QVARMA is superior to the predictive accuracy of Gaussian-VAR.