

ARTICLE

# Co-integration with score-driven models: an application to US real GDP growth, US inflation rate, and effective federal funds rate

Szabolcs Blazsek<sup>1</sup>, Alvaro Escribano<sup>2,\*</sup> and Adrian Licht<sup>1</sup>

<sup>1</sup>School of Business, Universidad Francisco Marroquín, Guatemala City (Guatemala), Guatemala and <sup>2</sup>Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain

\*Corresponding author. Email: [alvaroe@eco.uc3m.es](mailto:alvaroe@eco.uc3m.es).

## Abstract

Nonlinear co-integration is studied for score-driven models, using a new multivariate dynamic conditional score/generalized autoregressive score model. The model is named  $t$ -QVARMA (quasi-vector autoregressive moving average model), which is a location model for the multivariate  $t$ -distribution. In  $t$ -QVARMA,  $I(0)$  and co-integrated  $I(1)$  components of the dependent variables are included. For  $t$ -QVARMA, the conditions of the maximum likelihood estimator and impulse response functions (IRFs) are presented. A limiting special case of  $t$ -QVARMA, named Gaussian-QVARMA, is a Gaussian-VARMA specification with  $I(0)$  and  $I(1)$  components. As an empirical application, the US real gross domestic product growth, US inflation rate, and effective federal funds rate are studied for the period of 1954 Q3 to 2020 Q2. Statistical performance and predictive accuracy of  $t$ -QVARMA are superior to those of Gaussian-VAR. Estimates of the short-run IRF, long-run IRF, and total IRF impacts for the US data are reported.

**Keywords:** Nonlinear co-integration, nonlinear common trends, dynamic conditional score (DCS), generalized autoregressive score (GAS), quasi-vector autoregressive moving average (QVARMA)

**JEL Classifications:** 62M10, 91B84

## 1. Introduction

In this paper, the quasi-vector autoregressive moving average location model (hereinafter,  $t$ -QVARMA) is presented for the multivariate  $t$ -distribution, to study relationships among  $I(0)$  and co-integrated  $I(1)$  macroeconomic variables. The  $t$ -QVARMA model is a dynamic conditional score (DCS) model (Harvey and Chakravarty (2008), Harvey (2013)) or generalized autoregressive score (GAS) model (Creal et al. (2008, 2011, 2013)). Score-driven models are nonlinear observation-driven state space models, which are updated using the conditional score of the log-likelihood (LL) (hereinafter, score function). The score functions are nonlinear transformations of the dependent variables. Thus, for the co-integrated  $I(1)$  variables, a new form of nonlinear common trends (e.g. Escribano (2004), Escribano et al. (2006), (2008), Escanciano and Escribano (2008)) is considered in this paper.

The  $t$ -QVARMA( $p, q, r$ ) model of the present paper is in relation to the works of Harvey (2013) and Creal et al. (2014). The  $t$ -QVARMA( $p, q, r$ ) model is an extension of the DCS model for the multivariate  $t$ -distribution (Harvey (2013)), which is abbreviated as  $t$ -QVAR(1). The  $t$ -QVARMA( $p, q, r$ ) model is also an extension of the  $t$ -QVARMA( $p, q$ ) factor model of Creal et al. (2014), because the latter is a model for  $I(0)$  dependent variables, while the former is a model for  $I(0)$  and co-integrated  $I(1)$  dependent variables (Granger (1981), Engle and Granger (1987)),

© The Author(s), 2021. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted re-use, distribution, and reproduction in any medium, provided the original work is properly cited.

and because in  $t$ -QVARMA( $p, q, r$ ) less restrictions are imposed on the parameters of the multivariate score-driven filter than in the work of Creal et al. (2014). Thus, dynamic interaction effects among variables are more completely measured in our work than in the score-driven factor model of Creal et al. (2014).

As further contributions to the literature, for  $t$ -QVARMA( $p, q, r$ ), technical details of the model formulation for  $I(0)$  and co-integrated  $I(1)$  dependent variables, first-order representation, impulse response analysis, and statistical inference procedures are presented. We present conditions of the asymptotic properties of the maximum likelihood (ML) estimator for  $t$ -QVARMA( $p, q, r$ ). We provide conditions for the convergence of the multivariate  $I(0)$  filter of  $t$ -QVARMA to a stationary and ergodic sequence, using the results from the works of Brandt (1986), Elton (1990), and Straumann and Mikosch (2006). We also provide conditions for the identification of the parameters of  $t$ -QVARMA, using the results of Lütkepohl (2005). As a further empirical contribution, we present a macroeconomic application of  $t$ -QVARMA with identified structural shocks for US macroeconomic data.

The  $t$ -QVARMA( $p, q, r$ ) model includes  $I(0)$  and co-integrated  $I(1)$  filters, which is similar to the Granger representation of VAR models. For  $t$ -QVARMA( $p, q, r$ ), technical details of reduced-form and structural-form representations, and sign restrictions-based impulse response functions (IRFs) are presented in our paper. If the degrees of freedom parameter  $\nu \rightarrow \infty$ , then the  $t$ -QVARMA model will converge to a Gaussian-QVARMA specification with  $I(0)$  and  $I(1)$  components.

The empirical application uses quarterly data for the US real gross domestic product (GDP) growth, US inflation rate, and effective federal funds rate for period: 1954 Q3–2020 Q2. Estimation and forecasting results for  $t$ -QVARMA, Gaussian-VAR, and co-integrated Gaussian-VAR for vector error correction model (VECM) representation are presented. For the empirical application, (i) the statistical performance of  $t$ -QVARMA is superior to the statistical performance of Gaussian-VAR, and (ii) the out-of-sample multistep-ahead predictive accuracies of  $t$ -QVARMA are superior to those of Gaussian-VAR. We report the short-run IRF, long-run IRF, and total IRF estimates for the US data.

The remainder of this paper is organized as follows: Section 2 reviews the literature. Section 3 presents the reduced-form representation of  $t$ -QVARMA and its statistical inference. Section 4 presents the structural-form representation of  $t$ -QVARMA and the tools of IRF analysis. Section 5 presents an illustrative empirical application of  $t$ -QVARMA. Section 6 concludes. Technical details of the statistical inference and model specifications, and IRF estimates for  $t$ -QVARMA, Gaussian-QVARMA, and Gaussian-VAR are presented in Supplementary Material.

## 2. Review of the literature

Score-driven time series models are observation-driven (Cox (1981)) state space models of univariate or multivariate time series variables, which are estimated using the ML method. Score-driven models can be applied to the study of variables with different orders of integration, for example,  $I(0)$  variables (e.g. financial return; GDP growth) or  $I(1)$  variables (e.g. currency exchange rates). An example of univariate score-driven models is the quasi-AR (QAR) model (Harvey (2013)), which is a nonlinear outlier-robust alternative to the ARMA model (Box and Jenkins (1970)). Another example of univariate score-driven models is the Beta- $t$ -EGARCH (exponential generalized autoregressive conditional heteroskedasticity) model (Harvey and Chakravarty (2008)), which is an alternative to the GARCH (Engle (1982), Bollerslev (1986, 1987)) and EGARCH (Nelson (1991)) models. Due to the score-driven updating mechanism, the information gain in the filters is optimal in Beta- $t$ -EGARCH,

according to the Kullback—Leibler divergence in favor of the true data generating process (Blasques et al. (2015)).

The DCS model for the multivariate  $t$ -distribution (Harvey (2013, p. 210)) which is an alternative to the VARMA model (Tiao and Tsay (1989), see also Lütkepohl (2005)), and the  $t$ -QVARMA( $p,q$ ) model of Creal et al. (2014) which uses  $I(0)$  variables, are extended in the present paper. The  $t$ -QVARMA( $p,q,r$ ) model is for  $I(0)$  and co-integrated  $I(1)$  variables, in which an outlier in any variable influences the other variables through the covariance matrix of the errors. The measurements of dynamic interaction effects are less distorted by outliers than the same measurements for the Gaussian linear alternatives.

In the 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), (1995b), (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 using 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 of the present paper is an alternative to those outlier-robust frameworks.

### 3. $t$ -QVARMA

#### 3.1 Reduced-form representation

For the dependent variables  $y_t$  ( $K \times 1$ ) with  $t = 1 \dots, T$ , 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

$$y_t = c^* + \mu_t + v_t = c^* + \mu_t^* + \mu_t^\dagger + v_t, \tag{1}$$

$$\mu_t^* = \sum_{i=1}^p \Phi_i^* \mu_{t-i}^* + \sum_{j=1}^q \Psi_j^* u_{t-j}, \tag{2}$$

$$\mu_t^\dagger = \mu_{t-1}^\dagger + \sum_{l=1}^r \Psi_l^\dagger u_{t-l}, \tag{3}$$

$$v_t \sim t_K(0_{K \times 1}, \Sigma, v) = t_K[0, \Omega^{-1}(\Omega^{-1})', v] \text{ i.i.d.}, \tag{4}$$

where  $c^*$  ( $K \times 1$ ),  $\Phi_1^*, \dots, \Phi_p^*$  (all  $K \times K$ ),  $\Psi_1^*, \dots, \Psi_q^*$  (all  $K \times K$ ),  $\Psi_1^\dagger, \dots, \Psi_r^\dagger$  (all  $K \times K$ ),  $\Omega^{-1}$  ( $K \times K$ ), and  $v$  are constant parameters. In the work of Creal et al. (2014),  $\Phi_1^*, \dots, \Phi_p^*$  and  $\Psi_1^*, \dots, \Psi_q^*$  are diagonal matrices. In the present paper, the “final equations form” restriction for  $\Phi_1^*, \dots, \Phi_p^*$  (Lütkepohl (2005)) is used, according to which the  $K \times K$  parameter matrices  $\Phi_1^*, \dots, \Phi_p^*$  in (2) are replaced by the scalar parameters  $\phi_1^*, \dots, \phi_p^*$ . The final equations form ensures the identification of parameters in  $\mu_t^*$ , and full matrices for  $\Psi_1^*, \dots, \Psi_q^*$  are used.

Moreover,  $\mu_t^\dagger$  is in the final equations form of representation.

Variable  $u_t$  ( $K \times 1$ ) is the scaled score function, and it is defined in Section 3.2. Vector  $\mu_t^*$  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 VAR (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  $E(y_t | y_1, \dots, y_{t-1}, \mu_1^*, \mu_1^\dagger) = c^* + \mu_t = c^* + \mu_t^* + \mu_t^\dagger$ .

Filters  $\mu_t^*$  and  $\mu_t^\dagger$  are initialized using their unconditional means. For the first  $\max(p, q)$  observations of the sample,  $\mu_t^*$  is initialized using  $0_{K \times 1}$ . For the first  $r$  observations of the sample,  $\mu_t^\dagger$  is initialized using  $0_{K^\dagger \times 1}$ . All elements of  $\Phi_1^*, \dots, \Phi_p^*$  and  $\Psi_1^*, \dots, \Psi_q^*$  are real numbers. For each of the matrices  $\Psi_1^\dagger, \dots, \Psi_r^\dagger$ , all elements in the first  $K^*$  rows and all elements in the first  $K^*$  columns are restricted to zeros. For each of the matrices  $\Psi_1^\dagger, \dots, \Psi_r^\dagger$ , the rank of the quadratic submatrix 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  $\Sigma$  is estimated using the Cholesky decomposition  $\Omega^{-1}(\Omega^{-1})'$ . The degrees of freedom parameter  $\nu$  is jointly estimated using the rest of the parameters, where  $\nu > 2$  to ensure that the covariance matrix of  $\nu_t$  is well defined.

**3.2 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  $\mathcal{F}_{t-1} \equiv (y_1, \dots, y_{t-1}, \mu_1^*, \mu_1^\dagger)$  is

$$\ln f(y_t | \mathcal{F}_{t-1}; \Theta) = \ln \Gamma \left( \frac{\nu + K}{2} \right) - \ln \Gamma \left( \frac{\nu}{2} \right) - \frac{K}{2} \ln(\pi \nu), \tag{5}$$

$$-\frac{1}{2} \ln |\Sigma| - \frac{\nu + K}{2} \ln \left[ 1 + \frac{(y_t - c^* - \mu_t)' \Sigma^{-1} (y_t - c^* - \mu_t)}{\nu} \right],$$

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

$$\frac{\partial \ln f(y_t | \mathcal{F}_{t-1}; \Theta)}{\partial \mu_t} = \frac{\nu + K}{\nu} \Sigma^{-1} \times \left( 1 + \frac{\nu_t' \Sigma^{-1} \nu_t}{\nu} \right)^{-1} \nu_t \equiv \frac{\nu + K}{\nu} \Sigma^{-1} \times u_t, \tag{6}$$

(Harvey (2013)), where  $\nu_t = y_t - c^* - \mu_t$ . The scaled score function  $u_t$  is defined in (6) using the reduced-form error term, where  $\nu_t$  is multiplied by  $[1 + (\nu_t' \Sigma^{-1} \nu_t) / \nu]^{-1} = \nu / (\nu + \nu_t' \Sigma^{-1} \nu_t) \in (0, 1)$ . Therefore, the scaled score function is bounded by the reduced-form error term:  $|u_t| < |\nu_t|$ . The scaled score function  $u_t$  is multivariate i.i.d. with mean zero and a covariance matrix (Harvey (2013)):

$$\text{Var}(u_t) = E \left[ \frac{\partial \ln f(y_t | \mathcal{F}_{t-1}; \Theta)}{\partial \mu_t} \times \frac{\partial \ln f(y_t | \mathcal{F}_{t-1}; \Theta)}{\partial \mu_t'} \right] = \frac{\nu + K}{\nu + K + 2} \Sigma^{-1}. \tag{7}$$

**3.3 First-order representations of the score-driven filters**

The first-order representation of filter  $\mu_t^*$  is

$$M_t^* = \Phi^* M_{t-1}^* + \Psi^* U_{t-1}^*, \tag{8}$$

where the dimensions of the matrices of (8) are  $M_t^* [K(p + q - 1) \times 1]$ ,  $\Phi^* [K(p + q - 1) \times K(p + q - 1)]$ ,  $\Psi^* [K(p + q - 1) \times K(p + q - 1)]$ , and  $U_{t-1}^* [K(p + q - 1) \times 1]$ , and the elements of

those matrices are defined as follows:

$$\begin{bmatrix} \mu_t^* \\ \mu_{t-1}^* \\ \vdots \\ \mu_{t-p+1}^* \\ \dots \\ u_{t-1} \\ \vdots \\ u_{t-q+1} \end{bmatrix} = \begin{bmatrix} \Phi_1^* & \Phi_2^* & \dots & \Phi_{p-1}^* & \Phi_p^* & \Psi_2^* & \dots & \Psi_{q-1}^* & \Psi_q^* \\ I_K & 0_{K \times K} & \dots & \dots & 0_{K \times K} & 0_{K \times K} & \dots & \dots & 0_{K \times K} \\ 0_{K \times K} & I_K & 0_{K \times K} & \dots & \dots & \dots & \dots & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & \dots & 0_{K \times K} & I_K & 0_{K \times K} & 0_{K \times K} & \dots & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & \dots & \dots & \dots & 0_{K \times K} & 0_{K \times K} & \dots & \dots & 0_{K \times K} \\ 0_{K \times K} & \dots & \dots & \dots & 0_{K \times K} & I_K & \dots & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & \dots & \dots & \dots & 0_{K \times K} & 0_{K \times K} & \dots & I_K & 0_{K \times K} \end{bmatrix}$$

$$\begin{bmatrix} \mu_{t-1}^* \\ \mu_{t-2}^* \\ \vdots \\ \mu_{t-p}^* \\ \dots \\ u_{t-2} \\ \vdots \\ u_{t-q} \end{bmatrix} + \begin{bmatrix} \Psi_1^* & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ I_K & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} & 0_{K \times K} & \dots & 0_{K \times K} \end{bmatrix} \begin{bmatrix} u_{t-1} \\ 0_{K \times 1} \\ \vdots \\ 0_{K \times 1} \\ \dots \\ 0_{K \times 1} \\ 0_{K \times 1} \\ \vdots \\ 0_{K \times 1} \end{bmatrix}$$

If the maximum modulus of eigenvalues, denoted  $C_1$ , of  $\Phi^*$  is less than 1, and  $\Psi^*$  is non-zero, then  $M_t^*$  is covariance stationary.

The first-order representation of filter  $\mu_t^\dagger$  is

$$M_t^\dagger = \Phi^\dagger M_{t-1}^\dagger + \Psi^\dagger U_{t-1}^\dagger, \tag{9}$$

where the dimensions of the matrices of (9) are  $M_t^\dagger$  ( $Kr \times 1$ ),  $\Phi^\dagger$  ( $Kr \times Kr$ ),  $\Psi^\dagger$  ( $Kr \times Kr$ ), and  $U_{t-1}^\dagger$  ( $Kr \times 1$ ), and the elements of those matrices are defined as follows:

$$\begin{bmatrix} \mu_t^\dagger \\ \dots \\ u_{t-1} \\ \vdots \\ u_{t-r+1} \end{bmatrix} = \begin{bmatrix} I_K & \Psi_2^\dagger & \dots & \Psi_{r-1}^\dagger & \Psi_r^\dagger \\ \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & 0_{K \times K} & \dots & \dots & 0_{K \times K} \\ 0_{K \times K} & I_K & 0_{K \times K} & \dots & 0_{K \times K} \\ \dots & 0_{K \times K} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0_{K \times K} & 0_{K \times K} & \dots & I_K & 0_{K \times K} \end{bmatrix} \begin{bmatrix} \mu_{t-1}^\dagger \\ \dots \\ u_{t-2} \\ \vdots \\ u_{t-r} \end{bmatrix} +$$

$$+ \begin{bmatrix} \Psi_1^\dagger & 0_{K \times K} & \cdots & 0_{K \times K} \\ \vdots & \vdots & \ddots & \vdots \\ I_K & 0_{K \times K} & \cdots & 0_{K \times K} \\ 0_{K \times K} & 0_{K \times K} & \cdots & 0_{K \times K} \\ \vdots & \vdots & \ddots & \vdots \\ 0_{K \times K} & 0_{K \times K} & \cdots & 0_{K \times K} \end{bmatrix} \begin{bmatrix} u_{t-1} \\ \vdots \\ 0_{K \times 1} \\ 0_{K \times 1} \\ \vdots \\ 0_{K \times 1} \end{bmatrix}.$$

**3.4 ML estimator**

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

$$\hat{\Theta}_{ML} = \arg \max_{\Theta} LL(y_1, \dots, y_T; \Theta) = \arg \max_{\Theta} \sum_{t=1}^T \ln f(y_t | \mathcal{F}_{t-1}; \Theta), \tag{10}$$

where LL is log-likelihood. The asymptotic covariance matrix of parameters  $\hat{\Theta}_{ML}$  is estimated using the inverse information matrix (Creal et al. (2013), Harvey (2013), Blasques et al. (2017)). The conditions of the asymptotic properties of ML are presented in Supplementary Material A.

**4. Dynamic interaction effects**

**4.1 Structural-form representation**

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

$$\text{Var}(v_t) = \Sigma \times \frac{\nu}{\nu - 2} = \left( \frac{\nu}{\nu - 2} \right)^{1/2} \times \Omega^{-1} (\Omega^{-1})' \times \left( \frac{\nu}{\nu - 2} \right)^{1/2}, \tag{11}$$

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

$$v_t = \left( \frac{\nu}{\nu - 2} \right)^{1/2} \Omega^{-1} \times \epsilon_t, \tag{12}$$

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  $t$ -QVARMA:

$$\begin{aligned} \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega y_t &= \\ &= \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega c^* + \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega \mu_t^* + \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega \mu_t^\dagger + \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega v_t = \\ &= \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega c^* + \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega \mu_t^* + \left( \frac{\nu}{\nu - 2} \right)^{-1/2} \Omega \mu_t^\dagger + \epsilon_t. \end{aligned} \tag{13}$$

**4.2 IRF analysis**

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

$$\text{IRF}_{j,t} = \frac{\partial y_{t+j}}{\partial \epsilon_t} = \frac{\partial \mu_{t+j}^*}{\partial \epsilon_t} + \frac{\partial \mu_{t+j}^\dagger}{\partial \epsilon_t} + \frac{\partial v_{t+j}}{\partial \epsilon_t} \text{ for } j = 0, 1, \dots, \infty. \tag{14}$$

The first and second terms of (14) are zero for  $j = 0$ , since both  $\mu_t^*$  and  $\mu_t^\dagger$  include lags of  $\epsilon_t$ . 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 (14) is zero for  $j > 0$ , and it represents the contemporaneous interaction effects  $\text{IRF}_{j,t}^\nu$ . Under this notation,

$$\text{IRF}_{j,t} = \text{IRF}_{j,t}^* + \text{IRF}_{j,t}^\dagger + \text{IRF}_{j,t}^\nu \quad \text{for } j = 0, 1, \dots, \infty. \tag{15}$$

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

$$M_t^* = [I_{K(p+q-1)} - \Phi^*L]^{-1}\Psi^*U_{t-1}^*, \tag{16}$$

$$M_t^* = \sum_{j=0}^{\infty} (\Phi^*)^j\Psi^*U_{t-1-j}^*, \tag{17}$$

$$JM_t^* = \sum_{j=0}^{\infty} J(\Phi^*)^j\Psi^*J'JU_{t-1-j}^*, \tag{18}$$

where  $J = (I_K, 0_{K \times K}, \dots, 0_{K \times K})$  is  $[K \times K(p + q - 1)]$ . Then,

$$\mu_t^* = \sum_{j=0}^{\infty} J(\Phi^*)^j\Psi^*J'u_{t-1-j} = \sum_{j=0}^{\infty} J(\Phi^*)^j\Psi^*J'[(\nu - 2)\nu]^{1/2}\Omega^{-1} \times \frac{\epsilon_{t-1-j}}{\nu - 2 + \epsilon'_{t-1-j}\epsilon_{t-1-j}}. \tag{19}$$

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

$$\text{IRF}_{j,t}^* = \partial\mu_{t+j}^*/\partial\epsilon_t = J(\Phi^*)^{j-1}\Psi^*J'[(\nu - 2)\nu]^{1/2}\Omega^{-1}D_t \quad \text{for } j = 1, \dots, \infty, \tag{20}$$

$$D_t = \frac{\partial \frac{\epsilon_t}{\nu - 2 + \epsilon'_t\epsilon_t}}{\partial\epsilon_t} = \begin{bmatrix} d_{11,t} & \dots & d_{1K,t} \\ \dots & \dots & \dots \\ d_{K1,t} & \dots & d_{KK,t} \end{bmatrix} = \tag{21}$$

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

With respect to  $\text{IRF}_{j,t}^\dagger$ , component  $\mu_t^\dagger$  is expressed from (3) by recursive substitution as:

$$\begin{aligned} \mu_t^\dagger &= \Psi_1^\dagger u_{t-1} + (\Psi_1^\dagger + \Psi_2^\dagger)u_{t-2} + \dots + (\Psi_1^\dagger + \dots + \Psi_K^\dagger)u_{t-K} \\ &+ (\Psi_1^\dagger + \dots + \Psi_K^\dagger)u_{t-K-1} + \dots + (\Psi_1^\dagger + \dots + \Psi_K^\dagger)u_K \\ &+ (\Psi_1^\dagger + \dots + \Psi_{K-1}^\dagger)u_{K-1} + \dots + (\Psi_1^\dagger + \Psi_2^\dagger)u_2 + \Psi_1^\dagger u_1. \end{aligned} \tag{22}$$

**Table 1.** Sign restrictions on impact responses

	Supply shock	Demand shock	Monetary policy shock
US real GDP growth $y_{1,t}$	+	+	–
US inflation rate $y_{2,t}$	–	+	–
Effective federal funds rate $y_{3,t}$	unrestricted	+	+

Notes: Gross domestic product (GDP).

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

$$\begin{aligned}
 \text{IRF}_{1,t}^\dagger &= \partial \mu_{t+1}^\dagger / \partial \epsilon_t = \Psi_1^\dagger [(\nu - 2)\nu]^{1/2} \Omega^{-1} D_t \\
 \text{IRF}_{2,t}^\dagger &= \partial \mu_{t+2}^\dagger / \partial \epsilon_t = (\Psi_1^\dagger + \Psi_2^\dagger) [(\nu - 2)\nu]^{1/2} \Omega^{-1} D_t \\
 &\vdots \\
 \text{IRF}_{K-1,t}^\dagger &= \partial \mu_{t+K-1}^\dagger / \partial \epsilon_t = (\Psi_1^\dagger + \dots + \Psi_{K-1}^\dagger) [(\nu - 2)\nu]^{1/2} \Omega^{-1} D_t \\
 \text{IRF}_{j,t}^\dagger &= \partial \mu_{t+j}^\dagger / \partial \epsilon_t = (\Psi_1^\dagger + \dots + \Psi_K^\dagger) [(\nu - 2)\nu]^{1/2} \Omega^{-1} D_t \quad \text{for } j \geq K.
 \end{aligned}
 \tag{23}$$

The IRFs in (20) and (23) depend on  $D_t$ , that is, those IRFs are time-varying. Motivated by the estimation of IRF for a nonlinear model in the work of Herwartz and Lütkepohl (2000), in the present paper, the sample average of the IRFs is used. In  $t$ -QVARMA,  $\nu_t$  and  $\epsilon_t$  are multivariate i.i.d. error terms; hence, they are strictly stationary and ergodic. According to White (1984, Theorem 3.35), a nonlinear measurable function transforms strictly stationary and ergodic variables into new strictly stationary and ergodic variables. Therefore, all elements of  $D_t$  are strictly stationary and ergodic, and  $E(D_t)$  can be estimated using the sample average due to the ergodic theorem. An alternative to the use of  $E(\text{IRF}_{j,t}^*)$  and  $E(\text{IRF}_{j,t}^\dagger)$  is the period-by-period estimation of  $\text{IRF}_{j,t}^*$  and  $\text{IRF}_{j,t}^\dagger$ , respectively. In those applications, both  $\text{IRF}_{j,t}^*$  and  $\text{IRF}_{j,t}^\dagger$  may be averaged for several subsample periods, and the resulting IRF estimates may be compared. Finally, with respect to  $\text{IRF}_{j,t}^\nu$ :

$$\text{IRF}_{j,t}^\nu = \left( \frac{\nu}{\nu - 2} \right)^{1/2} \Omega^{-1} \quad \text{for } j = 0.
 \tag{24}$$

**4.3 Identification of the IRFs using sign restrictions**

IRFs are identified using the sign restrictions, which is based on simulations of matrix  $\Omega^{-1}$ , according to the following procedure: first, the ML estimates of  $\Omega^{-1}$  are used. Second, a  $K \times K$  matrix  $\tilde{K}$  of independent  $N(0, 1)$  numbers is simulated. Third, the QR decomposition of  $\tilde{K}$  is performed (Rubio-Ramirez et al. (2010)), and the resulting matrices are denoted as  $\tilde{Q}$  and  $\tilde{R}$ . Fourth, we define  $\tilde{\Omega}^{-1} \equiv \Omega^{-1} \times \tilde{Q}$ . For each simulation of  $\tilde{\Omega}^{-1}$ , sign restrictions are used in accordance with Table 1 (i.e. in this paper, we impose sign restrictions on the contemporaneous effects of structural shocks). For the IRFs, 1 million simulations of  $\tilde{K}$  are generated, and only those simulations are used that satisfy the sign restrictions, and for each simulation  $\Omega^{-1}$  is replaced by  $\tilde{\Omega}^{-1}$  in equations (14) to (23). For the simulations that satisfy the sign restrictions, we report the 10%, 50%, and 90% percentiles of the IRFs.

**5. Empirical application**

**5.1 Data**

Quarterly data from seasonally adjusted (SA) US real GDP growth  $y_{1,t}$ , SA US inflation rate  $y_{2,t}$ , and non-SA effective federal funds rate  $y_{3,t}$  are used for 1954 Q3 to 2020 Q2 (data source:



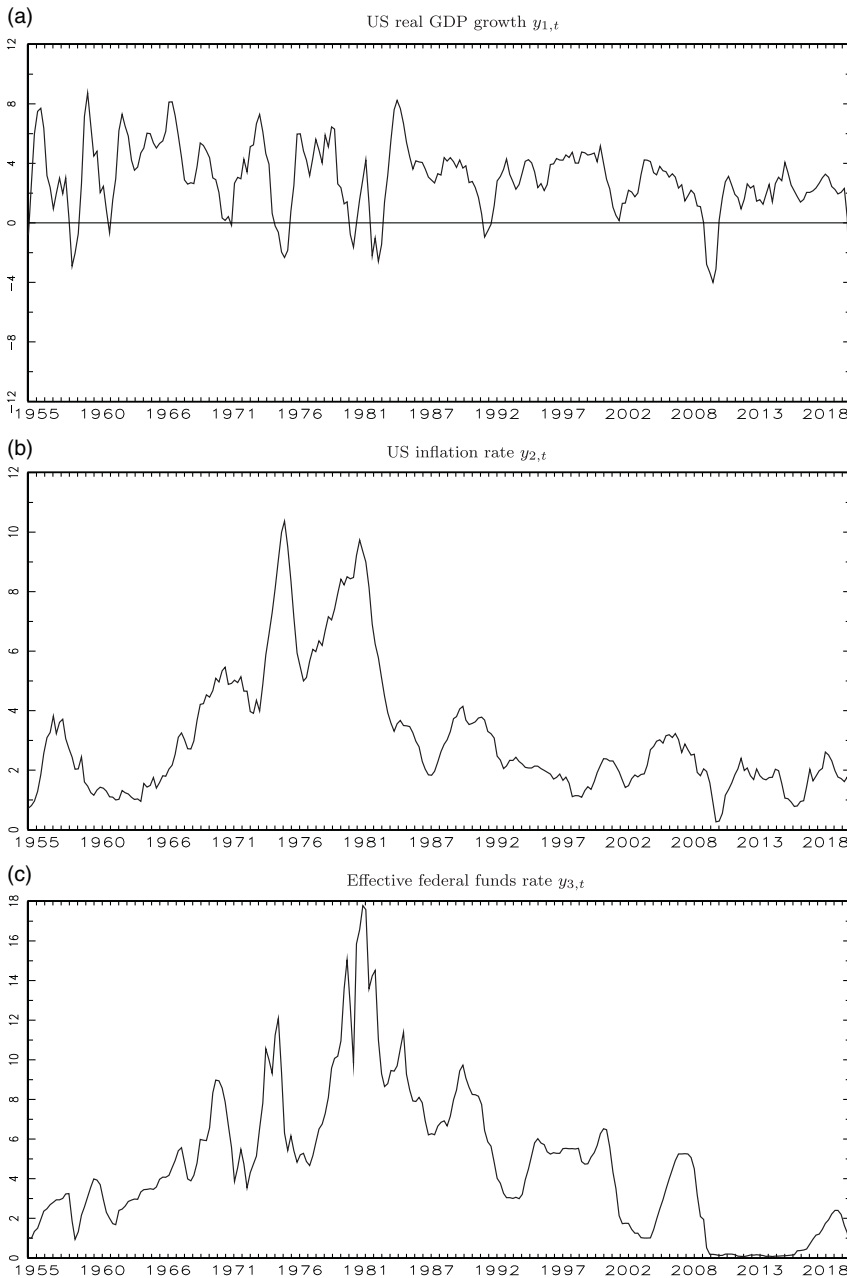


Figure 1. Evolution of US real GDP growth, US inflation rate, and effective federal funds rate from 1954 Q3 to 2020 Q2.

Federal Reserve Economic Data). The use of those variables is motivated by the work of Christiano et al. (1999) for structural VAR (SVAR) models. The evolution of  $y_{1,t}$ ,  $y_{2,t}$ , and  $y_{3,t}$  is shown in Figure 1. The co-integration test results for  $y_{2,t}$  and  $y_{3,t}$  are summarized as follows: First, the Engle–Granger co-integration test (Engle and Granger (1987)) is applied for  $\tilde{y}_t = (y_{2,t}, y_{3,t})'$  using the augmented Dickey–Fuller (ADF) test (Dickey and Fuller (1979)). In Table 2, results for the two steps of the Engle–Granger test are reported: (i) the ADF test results are reported for  $y_{2,t}$  and  $y_{3,t}$  (Panel A of Table 2), which suggest that both variables are  $I(1)$ . (ii) The existence of a co-integration relationship between  $y_{2,t}$  and  $y_{3,t}$  is equivalent to the rejection of the unit root null

Table 2. Descriptive statistics

A. Summary statistics	US real GDP growth $y_{1,t}$	US inflation rate $y_{2,t}$	Effective federal funds rate $y_{3,t}$
Start date	1954 Q3	1954 Q3	1954 Q3
End date	2020 Q2	2020 Q2	2020 Q2
Sample size	264	264	264
Minimum	-10.0275	0.2616	0.0600
Maximum	8.7343	10.3687	17.7800
Mean	2.9455	3.1082	4.7419
Standard deviation	2.4139	2.1252	3.5797
Skewness	-0.7946	1.4679	0.9984
Excess kurtosis	3.0183	1.6377	1.2248
ADF with constant	-3.1656** (0.0221)	-1.2512 (0.6543)	-1.8336 (0.3645)
B. Optimal lag-order, VAR( $p$ )	AIC	BIC	HQC
$p = 1$	6.5328	6.7028	6.6012
$p = 2$	6.0619	<b>6.3594</b>	6.1816
$p = 3$	6.0040	6.4290	6.1750
$p = 4$	6.0151	6.5676	6.2375
$p = 5$	5.8707	6.5507	6.1445
$p = 6$	5.8009	6.6084	<b>6.1260</b>
$p = 7$	5.7743	6.7094	6.1507
$p = 8$	5.7758	6.8383	6.2035
$p = 9$	5.6729	6.8630	6.1520
$p = 10$	<b>5.6614</b>	6.9790	6.1918
$p = 11$	5.6708	7.1159	6.2525
$p = 12$	5.7189	7.2914	6.3519
C. Co-integration tests for $\tilde{y}_t = (y_{2,t}, y_{3,t})'$			
Engle–Granger co-integration test statistic ( $p$ -value):			
ADF with constant on residuals	-2.7756* (0.0618)		
Nyblom–Harvey common trends test statistic:			
Lags $m = 30$	0.0934		
Lags $m = 40$	0.0882		
Lags $m = 50$	0.0889		
Lags $m = 100$	0.1377		
Lucas outlier-robust ( $t$ -distribution) co-integration test statistic:			
$H_0: \text{rank}(\Pi)=1; H_1: \text{rank}(\Pi)=2$	0.2182		

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

Notes: Gross domestic product (GDP); vector autoregression (VAR); Akaike information criterion (AIC); Bayesian information criterion (BIC); Hannan–Quinn criterion (HQC); augmented Dickey–Fuller (ADF); not available (NA). All variables are measured in percentage points. Bold numbers indicate the optimal lag-order, which is  $p = 2$  according to BIC. For all ADF tests of the table, the lag-orders are selected by using BIC, and  $p$ -values are presented in parentheses. \* and \*\* indicate significance at the 10% and 5% levels, respectively. For the Engle–Granger test, the significant co-integration is found. For the Nyblom–Harvey test, the test statistic for the possibility of serial correlation in the error term is reported, the 90%, 95%, and 99% level critical values are 0.162, 0.218, and 0.383, respectively, and results for different bandwidths  $m$  are presented. According to  $H_0$  of the Nyblom–Harvey test,  $\tilde{y}_t$  has one common trend for  $m = 30, 40, 50,$  and  $100$ . For the outlier-robust co-integration test of Lucas, the pseudo-likelihood ratio (PLR) test statistic is reported for the multivariate  $t$ -distribution, and the 90%, 95%, and 99% level critical values are 4.3114, 5.7948, and 10.5151, respectively. The estimated degrees of freedom parameter is  $\hat{\nu} = 2.0691$ , for which the critical values are obtained by using a linear approximation from the critical values of Lucas (1997, p. 159), which indicate that  $H_0$  cannot be rejected for the PLR test.

hypothesis of the ADF test for the residuals of the linear regression  $y_{3,t} = \beta_0 + \beta_1 y_{2,t} + \epsilon_t$ . The test results indicate that US inflation rate and effective federal funds rate are co-integrated (Panel C of Table 2).

Second, the common trends test of Nyblom and Harvey (2000) is applied for  $\tilde{y}_t = (y_{2,t}, y_{3,t})'$ :

$$\tilde{y}_t = \tilde{\mu}_t + \eta_{1,t} \text{ where } \eta_{1,t} \sim N(0_{2 \times 1}, \Sigma_{1,\eta}) \text{ is serially correlated,} \tag{25}$$

$$\tilde{\mu}_t = \tilde{\mu}_{t-1} + \eta_{2,t} \text{ where } \eta_{2,t} \sim N(0_{2 \times 1}, \Sigma_{2,\eta}) \text{ is serially independent.} \tag{26}$$

Under  $H_0$ ,  $\text{rank}(\Sigma_{2,\eta}) = 1$ , that is,  $y_{2,t}$  and  $y_{3,t}$  have one common trend. Test statistics for bandwidths  $m = 30, 40, 50, 100$  and critical values are reported in Table 2. According to those results, there is a common trend in US inflation rate and effective federal funds rate (Panel C of Table 2).

Third, the outlier-robust co-integration test of Lucas (1997) is applied for  $\tilde{y}_t = (y_{2,t}, y_{3,t})'$ , using the multivariate  $t$ -distribution for the error term in the following model:

$$\Delta \tilde{y}_t = \Pi \tilde{y}_{t-1} + \Gamma_1 \Delta \tilde{y}_{t-1} + \dots + \Gamma_{p-1} \Delta \tilde{y}_{t-p+1} + \tilde{v}_t, \tag{27}$$

where  $\Pi$  ( $2 \times 2$ ),  $\Gamma_1, \dots, \Gamma_{p-1}$  (each  $2 \times 2$ ) are time-invariant parameters.  $H_0$  and  $H_1$  for this test are presented in Table 2. For the co-integration test of Lucas, the PLR test statistic for the multivariate  $t$ -distribution is reported. According to the results,  $\text{rank}(\Pi) = 1$  (Panel C of Table 2). Thus, US inflation rate and effective federal funds rate are co-integrated.

**5.2 In-sample ML estimates**

Results for the  $t$ -QVARMA(1,1,1),  $t$ -QVARMA(2,1,1), and  $t$ -QVARMA(3,1,1) specifications are presented in this section, where the  $t$ -QVARMA(2,1,1) specification is formulated as follows:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} c_1^* \\ c_2^* \\ c_3^* \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^* \\ \mu_{2,t}^* \\ \mu_{3,t}^* \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \\ \mu_{3,t}^\dagger \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix}. \tag{28}$$

$$\begin{bmatrix} \mu_{1,t}^* \\ \mu_{2,t}^* \\ \mu_{3,t}^* \end{bmatrix} = \phi_1^* \begin{bmatrix} \mu_{1,t-1}^* \\ \mu_{2,t-1}^* \\ \mu_{3,t-1}^* \end{bmatrix} + \phi_2^* \begin{bmatrix} \mu_{1,t-2}^* \\ \mu_{2,t-2}^* \\ \mu_{3,t-2}^* \end{bmatrix} + \begin{bmatrix} \Psi_{1,11}^* & \Psi_{1,12}^* & \Psi_{1,13}^* \\ \Psi_{1,21}^* & \Psi_{1,22}^* & \Psi_{1,23}^* \\ \Psi_{1,31}^* & \Psi_{1,32}^* & \Psi_{1,33}^* \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \end{bmatrix}, \tag{29}$$

$$\begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \\ \mu_{3,t}^\dagger \end{bmatrix} = \begin{bmatrix} \mu_{1,t-1}^\dagger \\ \mu_{2,t-1}^\dagger \\ \mu_{3,t-1}^\dagger \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Psi_{1,22}^\dagger & \Psi_{1,23}^\dagger \\ 0 & \beta_2 \Psi_{1,22}^\dagger & \beta_2 \Psi_{1,23}^\dagger \end{bmatrix} \begin{bmatrix} u_{1,t-1} \\ u_{2,t-1} \\ u_{3,t-1} \end{bmatrix}, \tag{30}$$

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

The specification of  $\Psi_1^\dagger$  ensures that  $R = 1$ . Variables  $\mu_t^*$  and  $\mu_t^\dagger$  are initialized by using  $3 \times 1$  vectors of zeros. The co-integrating vector for  $y_{2,t}$  and  $y_{3,t}$  is given by  $(-\beta_2, 1)'$ . An alternative way of ensuring the same rank is to use the parametrization of the VECM representation of VAR models:  $\Psi_1^\dagger = \alpha\beta'$ , where  $\alpha$  and  $\beta$  (both  $K \times R$ ) are parameter matrices (Lütkepohl (2005)). For both parametrizations identical results were obtained, hence only the results for (30) are reported in this paper.

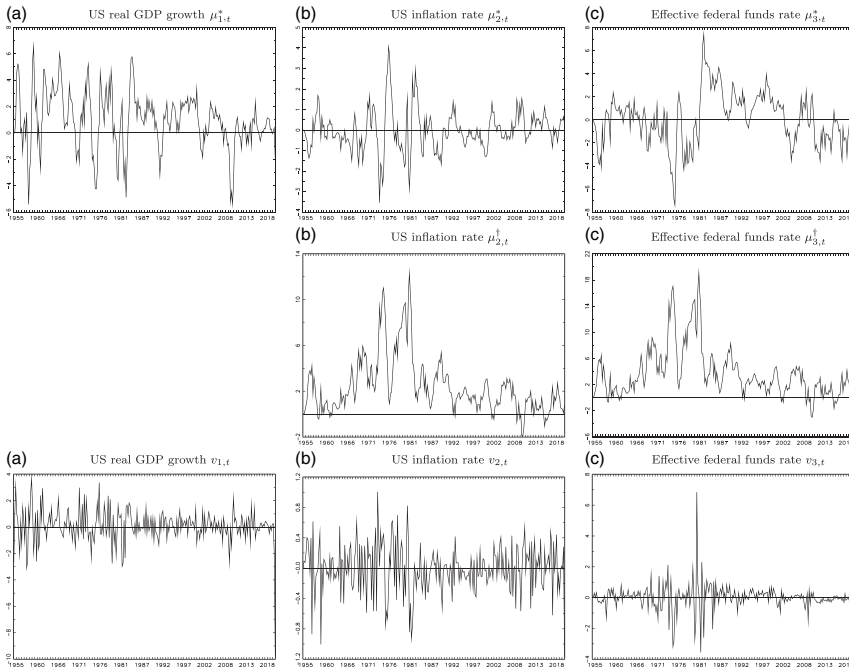
**Table 3.** Parameter estimates and model diagnostics

	<i>t</i> -QVARMA(1,1,1)	<i>t</i> -QVARMA(2,1,1)	<i>t</i> -QVARMA(3,1,1)	<i>t</i> -QVARMA(2,1,1) $\beta_2 = 1$	Gaussian-QVARMA(2,1,1)	Gaussian-VAR(2)	Gaussian-VAR(2) VECM
$c_1$	2.1301*** (0.4497)	1.9066*** (0.4519)	2.5447*** (0.3275)	1.6628*** (0.4477)	1.6107*** (0.5309)	$c_1$ 0.6430*** (0.1987)	$c_1$ 2.9519*** (0.1118)
$c_2$	0.7334 (5.7496)	0.7181 (0.5320)	0.9513** (0.4060)	0.6101 (0.4995)	0.7303+ (0.4863)	$c_2$ -0.0166 (0.0576)	$\alpha_2$ 0.0057 (0.0079)
$c_3$	1.4490 (7.9636)	0.9254 (0.7833)	0.7923 (0.7625)	1.4791* (0.7840)	0.7828 (0.7508)	$c_3$ -0.2405 (0.1983)	$\alpha_3$ -0.0920*** (0.0201)
$\phi_1^*$	0.8371*** (0.0291)	0.8849*** (0.0258)	0.9518*** (0.0183)	0.8797*** (0.0243)	0.9297*** (0.0177)	$\Phi_{1,11}$ 1.2739*** (0.0704)	$\beta_2$ -1.5346*** (0.1410)
$\phi_2^*$	NA	-0.0531** (0.0259)	-0.0428* (0.0236)	-0.0257 (0.0212)	-0.0661*** (0.0238)	$\Phi_{1,12}$ -0.2008 (0.2621)	$\Gamma_{1,11}$ 0.8356*** (0.0984)
$\phi_3^*$	NA	NA	-0.0656*** (0.0231)	NA	NA	$\Phi_{1,13}$ 0.0855 (0.1562)	$\Gamma_{1,12}$ 0.5526+ (0.3360)
$\Psi_{1,11}^*$	1.3447*** (0.0719)	1.2618*** (0.0733)	1.2573*** (0.0794)	1.3003*** (0.0697)	1.2382*** (0.0730)	$\Phi_{1,21}$ 0.0192 (0.0224)	$\Gamma_{1,13}$ 0.6616*** (0.1413)
$\Psi_{1,12}^*$	-0.2580 (0.2985)	-0.2216 (0.2848)	-0.2596 (0.2876)	-0.1958 (0.2840)	-0.0666 (0.3101)	$\Phi_{1,22}$ 1.4194*** (0.0646)	$\Gamma_{1,21}$ 0.0104 (0.0160)
$\Psi_{1,13}^*$	0.1619 (0.1437)	0.2415* (0.1367)	0.3009** (0.1343)	0.1925 (0.1362)	0.1813+ (0.1127)	$\Phi_{1,23}$ 0.0701** (0.0323)	$\Gamma_{1,22}$ 0.4581*** (0.0516)
$\Psi_{1,21}^*$	0.0091 (0.0200)	0.0176 (0.0187)	0.0284 (0.0214)	0.0150 (0.0186)	0.0261+ (0.0180)	$\Phi_{1,31}$ 0.0883+ (0.0598)	$\Gamma_{1,23}$ 0.0809*** (0.0231)
$\Psi_{1,22}^*$	-1.3387*** (0.4649)	-1.1919** (0.5239)	0.6081+ (0.3926)	-2.3649*** (0.7122)	-1.2749* (0.6586)	$\Phi_{1,32}$ 0.2331+ (0.1566)	$\Gamma_{1,31}$ 0.0374 (0.0390)
$\Psi_{1,23}^*$	-0.5969*** (0.1692)	-0.7391*** (0.2271)	-1.4523*** (0.3397)	-0.5653*** (0.2330)	-1.1929*** (0.3348)	$\Phi_{1,33}$ 1.0955*** (0.0614)	$\Gamma_{1,32}$ 0.1928+ (0.1304)
$\Psi_{1,31}^*$	-0.0151 (0.0510)	-0.0194 (0.0457)	-0.0269 (0.0466)	-0.0129 (0.0420)	0.0229 (0.0484)	$\Phi_{2,11}$ -0.4696*** (0.0733)	$\Gamma_{1,33}$ 0.2534*** (0.0608)
$\Psi_{1,32}^*$	-3.7277*** (0.7721)	-3.9209*** (0.8832)	-1.7727** (0.7786)	-3.6147*** (0.7273)	-4.0340*** (1.1356)	$\Phi_{2,12}$ 0.1932 (0.2585)	$\Omega_{11}^{-1}$ 2.0448*** (0.0877)
$\Psi_{1,33}^*$	0.7504*** (0.2739)	0.3416 (0.3805)	-1.4886* (0.7966)	0.9644*** (0.2293)	-0.8029 (0.5962)	$\Phi_{2,13}$ -0.1023 (0.1478)	$\Omega_{21}^{-1}$ 0.0537*** (0.0199)
$\Psi_{1,22}^\dagger$	2.9390*** (0.4883)	2.7393*** (0.5485)	0.9092** (0.4025)	3.9639*** (0.7387)	2.7345*** (0.6783)	$\Phi_{2,21}$ 0.0031 (0.0204)	$\Omega_{22}^{-1}$ 0.3120*** (0.0111)
$\Psi_{1,23}^\dagger$	0.7201*** (0.1797)	0.8764*** (0.2408)	1.5909*** (0.3538)	0.6896*** (0.2489)	1.3077*** (0.3416)	$\Phi_{2,22}$ -0.4385*** (0.0626)	$\Omega_{31}^{-1}$ 0.3101*** (0.0481)
$\beta_2$	1.3793*** (0.1606)	1.5457*** (0.1614)	2.0668*** (0.2144)	NA	1.6673*** (0.1649)	$\Phi_{2,23}$ -0.0688** (0.0305)	$\Omega_{32}^{-1}$ 0.1430*** (0.0399)
$\Omega_{11}^{-1}$	1.1674*** (0.0433)	1.1645*** (0.0430)	1.1386*** (0.0450)	1.1765*** (0.0418)	1.2424*** (0.0410)	$\Phi_{2,31}$ 0.0074 (0.0549)	$\Omega_{33}^{-1}$ 0.7417*** (0.0254)
$\Omega_{21}^{-1}$	0.0021 (0.0189)	0.0095 (0.0189)	0.0219 (0.0217)	0.0057 (0.0185)	0.0250 (0.0190)	$\Phi_{2,32}$ -0.1123 (0.1772)	
$\Omega_{22}^{-1}$	0.3146*** (0.0168)	0.3084*** (0.0158)	0.3173*** (0.0174)	0.3151*** (0.0143)	0.3065*** (0.0163)	$\Phi_{2,33}$ -0.1854*** (0.0596)	
$\Omega_{31}^{-1}$	0.2000*** (0.0581)	0.2105*** (0.0602)	0.2526*** (0.0662)	0.1974*** (0.0560)	0.2332*** (0.0642)	$\Omega_{11}^{-1}$ 1.2082*** (0.0450)	
$\Omega_{32}^{-1}$	0.1359*** (0.0438)	0.1363*** (0.0446)	0.1129** (0.0442)	0.1288*** (0.0477)	0.1445*** (0.0471)	$\Omega_{21}^{-1}$ 0.0188 (0.0204)	
$\Omega_{33}^{-1}$	0.7098*** (0.0204)	0.7333*** (0.0194)	0.7244*** (0.0225)	0.7281*** (0.0190)	0.7722*** (0.0189)	$\Omega_{22}^{-1}$ 0.3142*** (0.0173)	
$\nu$	44.0938*** (3.3988)	96.6340*** (12.3807)	78.3488*** (8.9625)	105.1299*** (13.8536)	NA	$\Omega_{31}^{-1}$ 0.2402*** (0.0616)	
						$\Omega_{32}^{-1}$ 0.1547*** (0.0493)	
						$\Omega_{33}^{-1}$ 0.7389*** (0.0219)	

**Table 3.** Continued

	<i>t</i> -QVARMA(1,1,1)	<i>t</i> -QVARMA(2,1,1)	<i>t</i> -QVARMA(3,1,1)	<i>t</i> -QVARMA(2,1,1) $\beta_2 = 1$	Gaussian-QVARMA(2,1,1)	Gaussian-VAR(2)	Gaussian-VAR(2) VECM		
LL	-2.9802	-2.9537	-2.9543	-2.9755	-3.0327	LL	-2.9855	LL	-3.5085
AIC	6.1346	<b>6.0892</b>	6.0980	6.1252	6.2396	AIC	6.1756	AIC	7.1609
BIC	6.4462	<b>6.4142</b>	6.4366	6.4368	6.5511	BIC	6.5413	BIC	7.4183
HQC	6.2598	<b>6.2198</b>	6.2340	6.2504	6.3648	HQC	6.3226	HQC	7.2643
$C_1$	0.8371	0.8849	0.9518	0.8797	0.9297	$C_1$	0.9430	$C_1$	1.0000

Notes: Quasi-vector autoregressive moving average (QVARMA); vector error correction model (VECM); not available (NA); log-likelihood (LL); Akaike information criterion (AIC); Bayesian information criterion (BIC); Hannan–Quinn criterion (HQC). Bold numbers indicate superior statistical performance. For the empirical estimation,  $C_1 < 1$  suggests covariance stationarity for  $\mu_t^*$ . Notice that *t*-QVARMA(2,1,1) is not a special case of *t*-QVARMA(3,1,1), because different initializations for  $\mu_t^*$  and  $\mu_t^\dagger$  are used. Hence, the LL of *t*-QVARMA(3,1,1) is not necessarily superior to the LL of *t*-QVARMA(2,1,1). Standard errors are in parentheses. †, \*, \*\*, and \*\*\* indicate significance at the 15%, 10%, 5%, and 1% levels, respectively.



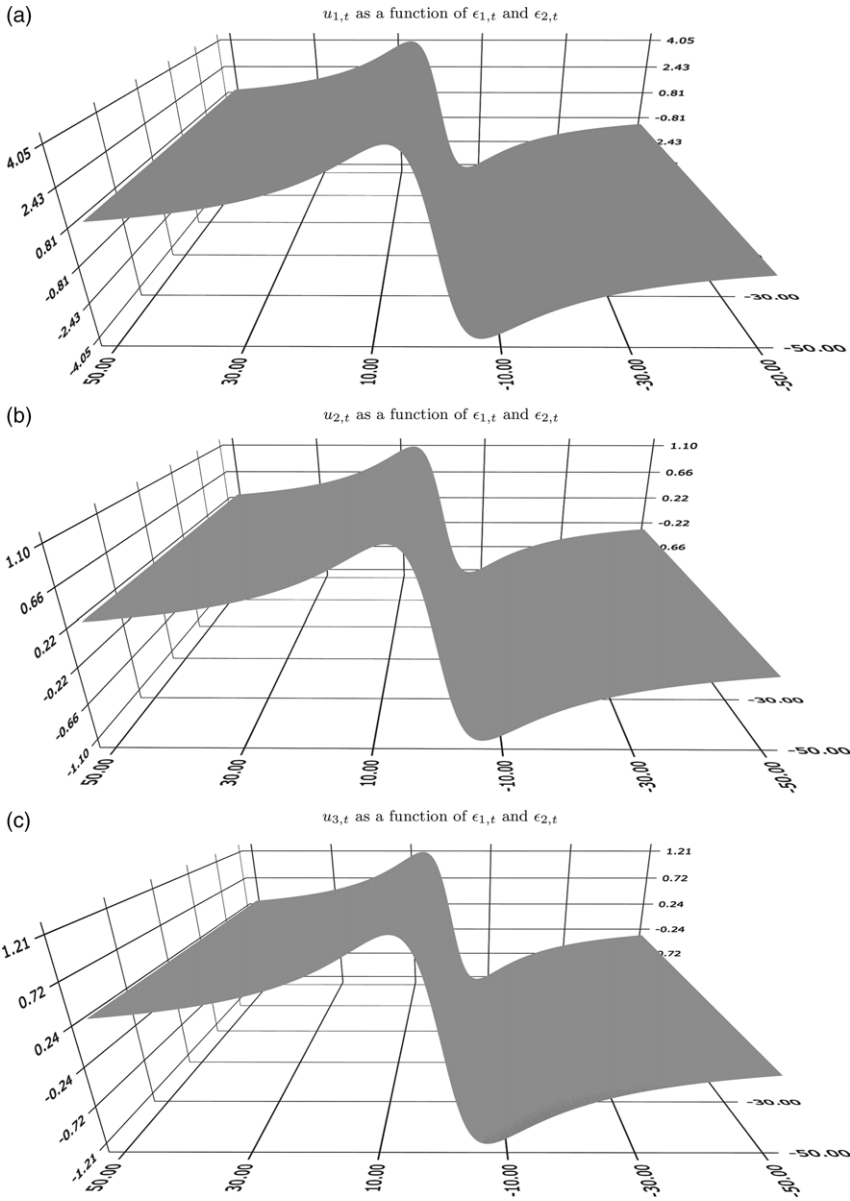
**Figure 2.** Location components  $\mu_t^*$  and  $\mu_t^\dagger$ , and irregular component  $v_t$  for  $t$ -QVARMA(2,1,1) from 1954 Q3 to 2020 Q2. Notes: US real GDP growth  $\mu_{1,t}^\dagger$  is not presented, since  $y_{1,t}$  does not include an  $I(1)$  component.

Parameter estimates for the  $t$ -QVARMA(1,1,1),  $t$ -QVARMA(2,1,1), and  $t$ -QVARMA(3,1,1) specifications are presented in Table 3. Several parameters for  $\phi_t^*$ ,  $\Psi_t^*$ , and  $\Psi_t^\dagger$  are statistically significant, which indicates that there are significant short-run dynamics in US real GDP, and there are significant short-run and long-run dynamics in US inflation rate and effective federal funds rate.

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. For the lag-order selection of  $t$ -QVARMA and Gaussian-VAR, the specific-to-general procedure is applied, by using the AIC, BIC, and HQC metrics. The use of the likelihood-based model selection metrics for the selection of lag-orders of score-driven models is motivated in the work of (Harvey, 2013, p.75). The results indicate that the statistical performance of  $t$ -QVARMA(2,1,1) is superior to  $t$ -QVARMA(1,1,1) and  $t$ -QVARMA(3,1,1).

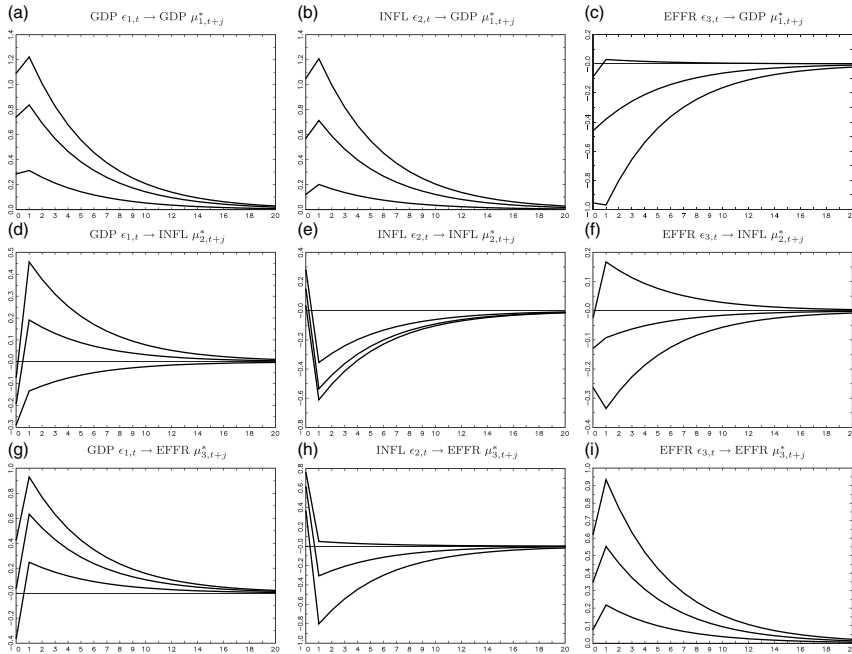
The evolution of the time series components  $\mu_t^*$ ,  $\mu_t^\dagger$ , and  $v_t$  for  $t$ -QVARMA(2,1,1) are presented in Figure 2, according to which the time variation in  $\mu_{2,t}^*$  and  $\mu_{3,t}^*$  indicates how the common factor captures structural changes for the US inflation rate and effective federal funds rate. For example,  $\mu_{2,t}^*$  show high positive values during the two oil crises, 1973 Q4 to 1974 Q1, and 1979 Q1 to 1981 Q4, respectively. This indicates that  $t$ -QVARMA is robust to structural changes in the dependent variables.

The discounting of extreme observations by the nonlinear transformation of the scaled score function  $u_t$  is studied for  $t$ -QVARMA(2,1,1) in Figure 3, in which all elements of  $u_t$  are presented, as functions of  $\epsilon_{1,t}$  and  $\epsilon_{2,t}$  ( $\epsilon_{3,t}$  is assumed to be zero in the figure). The figure indicates that  $u_{j,t} \rightarrow_p 0$  for  $j = 1, 2, 3$ , as  $|\epsilon_{1,t}| \rightarrow \infty$  or  $|\epsilon_{2,t}| \rightarrow \infty$ , which can be described as an asymptotic trimming of extreme values (Caivano and Harvey (2014)). Thus,  $t$ -QVARMA is robust to extreme observations in the variables.

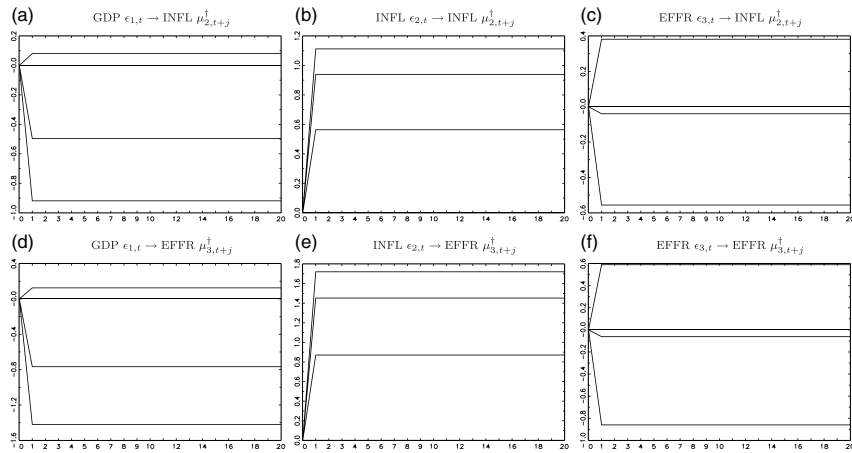


**Figure 3.** Robustness to extreme values in the noise for  $t$ -QVARMA(2,1,1).  
 Notes:  $\epsilon_{3,t} = 0$  is assumed for this figure.

IRFs for short-run, long-run, and total interaction effects for  $t$ -QVARMA(2,1,1) are presented in Figures 4, 5, and 6, respectively, and similar IRFs for  $t$ -QVARMA(1,1,1) and  $t$ -QVARMA(3,1,1) are presented in Supplementary Material B. US inflation rate shocks have positive short-run, long-run, and total effects on the US real GDP growth (Figures 4, 5, and 6, respectively). US inflation rate shocks have positive long-run and total effects on the effective federal funds rate (Figures 5 and 6, respectively). Effective federal funds rate shocks have negative short-run and total effects on the US real GDP growth (Figures 4 and 6, respectively). Effective federal funds rate shocks have a negative contemporaneous effect and a negative median estimate for short-run, long-run, and total interaction effects for all lags on the US inflation rate (Figures 4, 5, and 6, respectively).



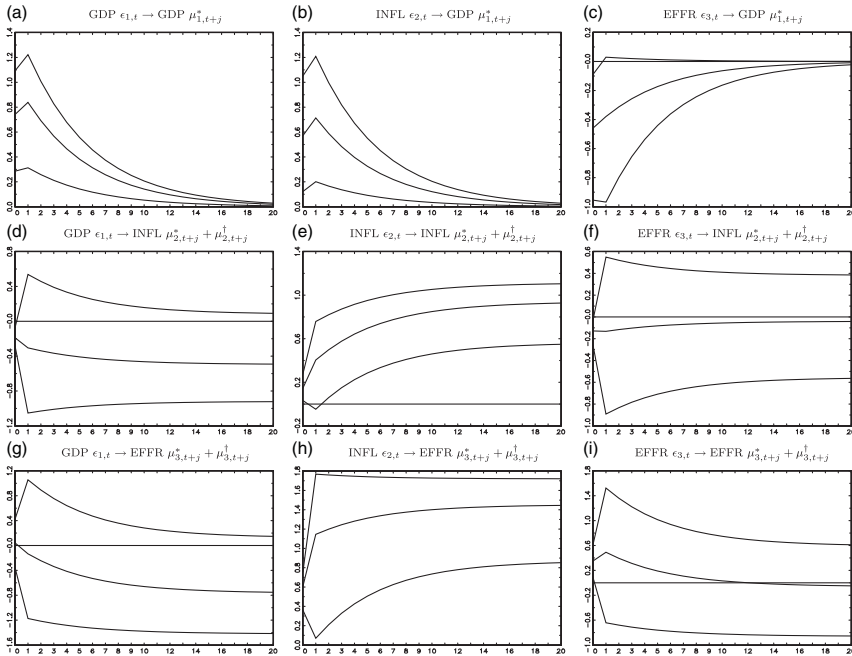
**Figure 4.** Short-run  $IRF_{j,t}^* + IRF_{j,t}^v$  with 10%, 50%, and 90% percentiles for  $t$ -QVARMA(2,1,1).  
 Notes: The confidence interval is for 9561 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.



**Figure 5.** Long-run  $IRF_{j,t}^{\dagger}$  with 10%, 50%, and 90% percentiles for  $t$ -QVARMA(2,1,1).  
 Notes: The confidence interval is for 9561 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

Estimation results and IRFs for  $t$ -QVARMA(2,1,1) under the restriction  $\beta_2 = 1$  are presented in Table 3 and Supplementary Material C, respectively. For the latter restriction, the co-integrating vector  $(-1, 1)$ , that is, the real interest rate is modeled. Estimation results and IRFs for a limiting special case of  $t$ -QVARMA(2,1,1) for  $\nu \rightarrow \infty$ , that is, Gaussian-QVARMA(2,1,1), are presented in Table 3 and Supplementary Material D, respectively. Estimation results for Gaussian-VAR(2), that is, unrestricted Gaussian-VAR(2) and co-integrated Gaussian-VAR(2) that is estimated by





**Figure 6.** Total  $IRF_{j,t}^* + IRF_{j,t}^\dagger + IRF_{j,t}^v$ , with 10%, 50%, and 90% percentiles for  $t$ -QVARMA(2,1,1).  
 Notes: The confidence interval is for 9561 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

using the VECM representation, are presented in Table 3. The IRFs for Gaussian-VAR(2) are presented in Supplementary Material E. The lag-orders of all competing models are selected using the AIC, BIC, and HQC criteria. The likelihood-based statistical performance of  $t$ -QVARMA(2,1,1) is superior to the statistical performances of all alternatives (Table 3).

**5.3 Predictive accuracy**

The predictive accuracies of  $t$ -QVARMA(2,1,1), Gaussian-VAR(2), and Gaussian-VAR(2)-VECM are compared by performing out-of-sample predictions of US real GDP growth, US inflation rate, and effective federal funds rate for the period of 2010 Q1 to 2020 Q2 (42 quarters). In the forecasting exercise, 42 rolling windows are used for parameter estimation, which are from the period of 1954 Q3 to 2009 Q4 until the period of 1964 Q4 to 2020 Q1 (there are  $Q = 222$  observations within each rolling window). For each rolling window,  $h = 1, \dots, 6$ -step ahead out-of-sample forecasts are computed.

For  $t$ -QVARMA, the  $h$ -step ahead forecast  $\mu_{t+h}$  is given by the following formulas:

$$\mu_{t+h}^* = \sum_{i=1}^p \Phi_i^* \mu_{t+h-i}^* + \sum_{j=1}^q \Psi_j^* u_{t+h-j} \tag{32}$$

$$\mu_{t+h}^\dagger = \mu_{t+h-1}^\dagger + \sum_{l=1}^r \Psi_l^\dagger u_{t+h-l} \tag{33}$$

$$\mu_{t+h} = \mu_{t+h}^* + \mu_{t+h}^\dagger \tag{34}$$

**Table 4.** RMSE and Giacomini–White out-of-sample forecast performance test results for the period of 2010 Q1 to 2020 Q2

Panel A.	t-QVARMA(2,1,1)			Gaussian-VAR(2)			Gaussian-VAR(2)-VECM		
	RMSE	GDP	INFL	EFFR	GDP	INFL	EFFR	GDP	INFL
1	<b>1.8082</b>	<b>0.3401</b>	0.2937	1.8277	0.3463	<b>0.2776</b>	3.3510	0.3418	0.3218
2	<b>2.2607</b>	0.4436	<b>0.4580</b>	2.3974	0.4556	0.5093	5.2463	<b>0.4378</b>	0.5981
3	<b>2.3398</b>	<b>0.5413</b>	<b>0.6725</b>	2.5769	0.5751	0.7916	6.7044	0.5455	0.9344
4	<b>2.4800</b>	0.6707	<b>0.9082</b>	2.7796	0.6981	1.0989	7.8704	<b>0.6453</b>	1.2898
5	<b>2.5807</b>	0.7536	<b>1.0703</b>	2.9058	0.7996	1.4211	8.8020	<b>0.6979</b>	1.6365
6	<b>2.6271</b>	0.9497	<b>1.2739</b>	2.9759	0.9331	1.7441	9.5672	<b>0.7990</b>	1.9849
Panel B.	Gaussian-VAR(2) minus t-QVARMA(2,1,1)								
GW test	GDP			INFL			EFFR		
1	1.0763(0.0658)			0.4479(0.0095)			−0.7047(0.0130)		
2	3.1780*** (0.2003)			0.4140(0.0261)			1.2660(0.0392)		
3	3.4192*** (0.3409)			0.6884(0.0549)			1.6942* (0.1029)		
4	3.0744*** (0.5124)			0.3823(0.0981)			1.8412* (0.2079)		
5	2.6859** (0.6640)			0.4298(0.1662)			2.5239** (0.3462)		
6	2.2675** (0.8619)			−0.1137(0.2743)			2.6426** (0.5370)		
Panel C.	Gaussian-VAR(2)-VECM minus t-QVARMA(2,1,1)								
GW test	GDP			INFL			EFFR		
1	4.6690*** (1.7048)			0.1396(0.0083)			1.1470(0.0151)		
2	9.1757*** (2.4426)			−0.3881(0.0132)			2.2009** (0.0672)		
3	10.7714*** (3.6647)			0.1866(0.0246)			2.2997** (0.1830)		
4	11.2813*** (4.9456)			−0.7946(0.0421)			2.1234** (0.3950)		
5	10.8058*** (6.5535)			−1.2200(0.0663)			2.2545** (0.6797)		
6	9.7992*** (8.6363)			−1.9932* (0.1322)			2.1801** (1.0627)		

Notes: Quasi-vector autoregressive moving average (QVARMA); vector error correction model (VECM); root mean squared error (RMSE); Giacomini–White (GW). In this table, GDP, INFL, and EFFR refer to US real GDP growth, US inflation rate, and effective federal funds rate, respectively. The loss function is defined as  $L_{t,h} = (y_{t+h} - \mu_{t+h})^2$  for  $h = 1, \dots, 6$  and  $t = Q + h, \dots, T$  for the forecasting window, where  $\mu_{t+h}$  represents an  $h$ -step ahead forecast. For each variable and forecast horizon, a bold number indicates the lowest RMSE estimate. Significantly positive GW test statistic indicates that the forecast performance of t-QVARMA(2,1,1) is superior to the forecast performance of Gaussian-VAR(2) or Gaussian-VAR(2)-VECM. Robust standard errors of the GW test statistic are reported in parentheses. \*, \*\*, and \*\*\* indicate significance at the 10%, 5%, and 1% levels, respectively.

In (32)–(34), the estimates of  $\mu_{t+h-i}^*$ ,  $u_{t+h-j}$ ,  $\mu_{t+h-1}^\dagger$ , 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  $\mu_{t+h-i}^*$  and  $\mu_{t+h-1}^\dagger$  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 and Gaussian-VAR(2)-VECM, the  $h$ -step ahead forecast  $\mu_{t+h}$  is given by:

$$\mu_{t+h} = c + \sum_{i=1}^p \Phi_i y_{t+h-i} \tag{35}$$

In (35),  $y_{t+h-i}$  is used for forecasting when its observations are available from the rolling window. Otherwise,  $y_{t+h-i}$  is replaced by its forecasts.

The root mean squared error (RMSE) of the prediction and a loss function are computed for each forecast, where the loss function is defined as:  $L_{t,h} = (y_{t+h} - \mu_{t+h})^2$  for  $h = 1, \dots, 6$ . The

RMSE for the forecasting window is presented in Panel A of Table 4, where for each forecast horizon  $h$  and for each variable, the lowest loss function estimate is highlighted. 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 Gaussian-VAR(2) (Panel B of Table 4) or Gaussian-VAR(2)-VECM (Panel C of Table 4), and  $L_{2,t,h}$  denotes the loss function of  $t$ -QVARMA(2,1,1). The significance of  $\Delta L_{t,h}$  is studied using the predictive accuracy test of the work of Giacomini and White (2006). The Giacomini–White test statistics are reported in Table 4. According to the results,  $t$ -QVARMA(2,1,1) is superior to Gaussian-VAR(2) and Gaussian-VAR(2)-VECM for the US real GDP growth and effective federal funds rate, and the predictive accuracies do not differ significantly for the US inflation rate.

## 6. Conclusions

In this paper,  $t$ -QVARMA has been introduced for the analysis of dynamic interactions effects among  $I(0)$  and co-integrated  $I(1)$  time series variables. The reduced-form and the structural-form representations of  $t$ -QVARMA have been presented, and tools have been provided for IRF analysis. Conditions of the asymptotic properties of ML have been presented, and it has also been shown that a limiting special case of  $t$ -QVARMA is the linear Gaussian-QVARMA.

An empirical application of  $t$ -QVARMA to quarterly time series data on US real GDP growth, US inflation rate, and effective federal funds rate for the period of 1954 Q3 to 2020 Q2 has been performed. The estimates of different  $t$ -QVARMA specifications have been compared with the estimates of multivariate linear Gaussian alternatives. For the empirical application, the results have suggested that the statistical performance of  $t$ -QVARMA is superior to the statistical performances of Gaussian-QVARMA, Gaussian-VAR, and Gaussian-VAR-VECM. The out-of-sample predictive accuracies of  $t$ -QVARMA, Gaussian-VAR, and Gaussian-VAR-VECM have been compared for the period of 2010 Q1 to 2020 Q2. For the empirical application, the results have suggested that the predictive accuracy of  $t$ -QVARMA is superior to the predictive accuracies of multivariate linear Gaussian alternatives. These illustrative results of  $t$ -QVARMA for the macroeconomic control data may motivate the practical consideration of the new model for  $I(0)$  and co-integrated  $I(1)$  time series variables.

**Acknowledgements.** The paper was presented at the GESG Research Seminar of Universidad Francisco Marroquín (June 20, 2019, Guatemala City). The authors greatly appreciate the comments of the anonymous reviewers of the journal, Juan Carlos Arriaza Herrera, Astrid Ayala, Matthew Copley, Carla Hess, and Søren Johansen. All remaining errors are our own. No potential conflict of interest is reported by the authors. Computer codes are available from the authors upon request. 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), Comunidad de Madrid (MadEco-CM S2015/HUM-3444), and Agencia Estatal de Investigación (2019/00419/001).

**Supplementary materials.** To view supplementary material for this article, please visit <http://dx.doi.org/10.1017/S1365100521000365>.

## References

- Blasques, F., S. J. Koopman, and A. Lucas (2015) Information-theoretic optimality of observation-driven time series models for continuous responses. *Biometrika* 102(2), 325–343.
- Blasques, F., S. J. Koopman, and A. Lucas (2017) Maximum likelihood estimation for score-driven models. *TI 2014-029/III Tinbergen Institute Discussion Paper*. <https://papers.tinbergen.nl/14029.pdf>.
- Bollerslev, T. (1986) Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31(3), 307–327.
- Bollerslev, T. (1987) A conditionally heteroskedastic time series model for security prices and rates of return data. *Review of Economics and Statistics* 69(3), 542–547.
- Bosco, B., L. Parisio, M. Pelagatti, and F. Baldi (2010) Long-run relations in European electricity prices. *Journal of Applied Econometrics* 25(5), 805–832.
- Box, G. E. P. and G. M. Jenkins (1970) *Time Series Analysis, Forecasting and Control*. San Francisco: Holden-Day.

- Brandt, A. (1986) The stochastic equation  $Y_{n+1} = A_n Y_n + B_n$  with stationary coefficients. *Advances in Applied Probability* 18(1), 211–220.
- Caivano, M. and A. Harvey (2014) Time series models with an EGB2 conditional distribution. *Journal of Time Series Analysis* 35(6), 558–571.
- Christiano, L. J., M. Eichenbaum, and C. L. Evans. Monetary policy shocks: what have we learned and to what end? *Handbook of Macroeconomics* 1(A), 65–148.
- Cox, D. R. (1981) Statistical analysis of time series: some recent developments. *Scandinavian Journal of Statistics* 8(2), 93–115.
- Creal, D., S. J. Koopman, and A. Lucas (2008) A general framework for observation driven time-varying parameter models. *Tinbergen Institute Discussion Paper* 08-108/4. <https://www.tinbergen.nl/discussion-paper/2649/08-108-4-a-general-framework-for-observation-driven-time-varying-parameter-models>
- Creal, D., S. J. Koopman, and A. Lucas (2011) A dynamic multivariate heavy-tailed model for time-varying volatilities and correlations. *Journal of Business and Economic Statistics* 29(4), 552–563.
- Creal, D., S. J. Koopman, and A. Lucas (2013) Generalized autoregressive score models with applications. *Journal of Applied Econometrics* 28(5), 777–795.
- Creal, D., B. Schwaab, S. J. Koopman, and A. Lucas (2014) Observation-driven mixed-measurement dynamic factor models with an application to credit risk. *The Review of Economics and Statistics* 96(5), 898–915.
- de Jong, R. M., C. Amsler, and P. Schmidt (2007) A robust version of the KPSS test based on indicators. *Journal Econometrics* 137(2), 311–333.
- Dickey, D. A. and W. A. Fuller (1979) Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association* 74(366), 427–431.
- Elton, J. H. (1990) A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Processes and their Applications* 34(1), 39–47.
- Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50(4), 987–1007.
- Engle, R. F. and C. W. J. Granger (1987) Co-integration and error correction: representation, estimation, and testing. *Econometrica* 55(2), 251–276.
- Escanciano, J.-C. and A. Escribano (2008) Econometrics: Nonlinear cointegration. In R. A. Meyers (ed.), *Complex Systems in Finance and Econometrics*, pp. 203–215. New York: Springer.
- Escribano, A. (2004) Nonlinear error correction: The case of money demand in the United Kingdom (1878–2000). *Macroeconomic Dynamics* 8(1), 76–116.
- Escribano, A., J. I. Peña, and P. Villaplana (2011) Modeling electricity prices: international evidence. *Oxford Bulletin of Economics and Statistics* 73(5), 622–650.
- Escribano, A., A. E. Sipols, and F. Aparicio (2006) Nonlinear cointegration and nonlinear error correction: Record counting cointegration tests. *Communications in Statistics—Simulation and Computation* 35(4), 939–956.
- Escribano, A., A. E. Sipols, and M. Teresa Santos (2008) Testing for cointegration using induced-order statistics. *Computational Statistics* 23, 131–151.
- Franses, P. H. and A. Lucas (1998) Outlier detection in cointegration analysis. *Journal of Business and Economic Statistics* 16(4), 459–468.
- Giacomini, R. and H. White (2006) Tests of conditional predictive ability. *Econometrica* 74(6), 1545–1578.
- 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.
- Harvey, A. C. (2013) *Dynamic Models for Volatility and Heavy Tails*. Cambridge: Cambridge University Press.
- Harvey, A. C. and T. Chakravarty (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>
- Herwartz, H. and H. Lütkepohl (2000) Multivariate volatility analysis of VW stock prices. *International Journal of Intelligent Systems in Accounting, Finance & Management* 9(1), 35–54.
- Johansen, S. (1995) *Likelihood-Based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.
- 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.
- Lucas, A. (1995b) Unit root test based on M-estimators. *Econometric Theory* 11(2), 331–346.
- Lucas, A. (1997) Cointegration testing using pseudolikelihood ratio tests. *Econometric Theory* 13(2), 149–169.
- Lucas, A. (1998) Inference on cointegrating ranks using LR and LM tests based on pseudo-likelihoods. *Econometric Reviews* 17(2), 185–214.
- Lütkepohl, H. (2005) *New Introduction to Multivariate Time Series Analysis*. Berlin Heidelberg: Springer-Verlag.
- Nelson, D. B. (1991) Conditional heteroskedasticity in asset returns: A new approach. *Econometrica* 59(2), 347–370.
- Nyblom, J. and A. Harvey (2000) Tests of common stochastic trends. *Econometric Theory* 16(2), 176–199.
- Peersman, G. (2005) What caused the early millennium slowdown? Evidence based on vector autoregressions. *Journal of Applied Econometrics* 20(2), 185–207.

- Rubio-Ramirez, J. F., D. F. Waggoner, and T. Zha (2010) Structural vector autoregressions: Theory for identification and algorithms for inference. *Review of Economic Studies* 77(2), 665–696.
- Straumann, D. and T. Mikosch (2006) Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: A stochastic recurrence equations approach. *The Annals of Statistics* 34(1), 2449–2495.
- Tiao, G. C. and R. S. Tsay (1989) Model specification in multivariate time series. *Journal of the Royal Statistical Society* 51(2), 157–213.
- White, H. (1984) *Asymptotic Theory for Econometricians*. San Diego: Academic Press.

---

**Cite this article:** Blazsek S, Escibano A and Licht A (2023). “Co-Integration with Score-Driven Models: An Application to US Real GDP Growth, US Inflation Rate, and Effective Federal Funds Rate.” *Macroeconomic Dynamics* 27, 203–223. <https://doi.org/10.1017/S1365100521000365>