Supplementary material:

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

Szabolcs Blazsek*, Alvaro Escribano**, and Adrian Licht*

*School of Business, Universidad Francisco Marroquín, Guatemala City (Guatemala), Guatemala

**Department of Economics, Universidad Carlos III de Madrid, Getafe (Madrid), Spain

Supplementary Material A: Statistical inference of t-QVARMA(p,q,r)

1. Maximum likelihood (ML) estimator

The parameters of t-QVARMA (quasi-vector autoregressive moving average) are estimated by using the ML method, as follows:

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

where LL is log-likelihood, $\mathcal{F}_{t-1} \equiv (y_1, \ldots, y_{t-1}, \mu_1^*, \mu_1^\dagger)$, and Θ is the $S \times 1$ vector of parameters.

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

$$G_{t,i}(\Theta) = -\frac{\partial \ln f(y_t|\mathcal{F}_{t-1}; \Theta)}{\partial \Theta_i} \tag{A.2}$$

for period $t = 1, \ldots, T$ and parameter $i = 1, \ldots, S$ (Wooldridge 1994). 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 the assumptions of this section, the ML estimator of (A.1) is equivalent to the representation:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\mathrm{ML}})' = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} G_{t,1}(\hat{\Theta}_{\mathrm{ML}}) \\ \vdots \\ G_{t,S}(\hat{\Theta}_{\mathrm{ML}}) \end{bmatrix} = \frac{1}{T}\sum_{t=1}^{T} \begin{bmatrix} \frac{\partial \ln f(y_t|\mathcal{F}_{t-1}; \hat{\Theta}_{\mathrm{ML}})}{\partial \Theta_1} \\ \vdots \\ \frac{\partial \ln f(y_t|\mathcal{F}_{t-1}; \hat{\Theta}_{\mathrm{ML}})}{\partial \Theta_S} \end{bmatrix} = 0_{S\times 1} \tag{A.3}$$

According to the mean-value expansion about Θ_0:

$$\frac{1}{T}\sum_{t=1}^{T} G_t(\hat{\Theta}_{\mathrm{ML}})' = \frac{1}{T}\sum_{t=1}^{T} G_t(\Theta_0)' + \frac{1}{T}\left[\sum_{t=1}^{T} H_t(\bar{\Theta})\right](\hat{\Theta}_{\mathrm{ML}} - \Theta_0) \tag{A.4}$$

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

$$H_t(\Theta) = -\frac{\partial^2 \ln f(y_t|\mathcal{F}_{t-1}; \Theta)}{\partial \Theta \partial \Theta'} \tag{A.5}$$

is evaluated at S different mean values, indicated by $\bar{\Theta}$ in (A.4) (Wooldridge 1994). Each $\bar{\Theta}$ is located between Θ_0 and $\hat{\Theta}_{\mathrm{ML}}$ that is expressed as: $||\bar{\Theta} - \Theta_0|| \leq ||\hat{\Theta}_{\mathrm{ML}} - \Theta_0||$, where $||\cdot||$ is the Euclidean norm.

From (A.3) and (A.4):

$$\sqrt{T}(\hat{\Theta}_{\mathrm{ML}} - \Theta_0) = \left[-\frac{1}{T} \sum_{t=1}^{T} H_t(\bar{\Theta}) \right]^{-1} \left[\frac{1}{\sqrt{T}} \sum_{t=1}^{T} G_t(\Theta_0)' \right] \tag{A.6}$$

The asymptotic covariance matrix of $\hat{\Theta}_{\mathrm{ML}}$ is estimated according to the inverse information matrix: $\{(1/T)\sum_{t=1}^{T}[G_t(\hat{\Theta}_{\mathrm{ML}})' G_t(\hat{\Theta}_{\mathrm{ML}})]\}^{-1}$ (Creal et al. 2013; Harvey 2013; Blasques et al. 2017).

In the following sections, we present the assumptions required for the asymptotic properties of the ML estimator and for the use of the inverse information matrix for the ML standard errors.

2. ML assumptions

The following properties are important for the asymptotic properties of ML: (i) scaled score function u_t, and its first and second derivatives with respect to μ_t, are bounded functions of v_t. (ii) The dynamic parameter matrix of $\mu_t^\dagger$ is set to the identity matrix (hence, it is not an estimated parameter matrix).

For the asymptotic properties of ML, results from the works of Wooldridge (1994), Creal et al. (2013), Harvey (2013), and Blasques et al. (2017) are used. We use the following assumptions:

(S1) Parameter set $\tilde{\Theta} \in \mathbb{R}^S$ is compact.

(S2) Θ_0 is an interior point within $\tilde{\Theta} \subset \mathbb{R}^S$, where Θ_0 represents the true values of Θ.

(S3) For each $\Theta \in \tilde{\Theta}$, $\ln f(\cdot|\mathcal{F}_{t-1};\Theta)$ is a Borel measurable function on $\mathbb{R}$.

(S4) For each $y_t \in \mathbb{R}$, $\ln f(y_t|\mathcal{F}_{t-1};\cdot)$ is a continuous function on $\tilde{\Theta}$.

(S5) For each $y_t \in \mathbb{R}$, $\ln f(y_t|\mathcal{F}_{t-1};\cdot)$ is twice continuously differentiable on all interior points of $\tilde{\Theta}$.

(S6) $E[|\ln f(y_t|\mathcal{F}_{t-1};\Theta)|] < \infty$ for all $\theta \in \tilde{\Theta}$.

(S7) $\lim_{T\to\infty}(1/T)\sum_{t=1}^{T} E[\ln f(y_t|\mathcal{F}_{t-1};\Theta)]$ exists for all $\theta \in \tilde{\Theta}$.

(S8) $\max_{\theta\in\tilde{\Theta}}|(1/T)\sum_{t=1}^{T} \ln f(y_t|\mathcal{F}_{t-1};\Theta) - E[\ln f(y_t|\mathcal{F}_{t-1};\Theta)]| \to_p 0$ if $T \to \infty$.

(S9) $\int_{\mathbb{R}} f(y_t|\mathcal{F}_{t-1};\Theta)dy_t = 1$ for all Θ.

(S10) $f(y_t|\mathcal{F}_{t-1};\Theta_0) = p_0(y_t|\mathcal{F}_{t-1};\Theta_0)$, where p_0 is the true conditional density.

(S11) $f(y_t|\mathcal{F}_{t-1};\Theta_0) = p_0(y_t|\mathcal{F}_{t-1};\Theta_0)$ for Θ_0 is a dynamically complete density.

(S12) $\partial[\int_{\rm I\!R} f(y_t|\mathcal{F}_{t-1};\Theta)dy_t]/\partial\Theta = \int_{\rm I\!R}[\partial f(y_t|\mathcal{F}_{t-1};\Theta)/\partial\Theta]dy_t$.

(S13) $\partial[\int_{\rm I\!R} G_t(\Theta)'f(y_t|\mathcal{F}_{t-1};\Theta)dy_t]/\partial\Theta = \int_{\rm I\!R}[\partial G_t(\Theta)'f(y_t|\mathcal{F}_{t-1};\Theta)/\partial\Theta]dy_t$.

(S14) Θ_0 is a unique solution to:

$$\max_{\Theta\in\tilde{\Theta}} \lim_{T\to\infty} \sum_{t=1}^{T} \ln f(y_t|\mathcal{F}_{t-1};\Theta) \tag{A.7}$$

(S15) M_t^* converges exponentially fast almost surely (e.a.s.) to a unique strictly stationary and ergodic sequence.

(S16) For each $\Theta \in \tilde{\Theta}$, each element of $H_t(\Theta)$ is a Borel measurable function on ${\rm I\!R}$.

(S17) For each $y_t \in {\rm I\!R}$, each element of $H_t(\Theta)$ is a continuous function on $\tilde{\Theta}$.

(S18) For each element of $H_t(\Theta)$, $|H_{i,j,t}(\Theta)| < \infty$ for all Θ.

(S19) Matrix

$$A_0 = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[H_t(\Theta_0)] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} \text{Var}[G_t(\Theta_0)'] = \lim_{T\to\infty} \frac{1}{T}\sum_{t=1}^{T} E[G_t(\Theta_0)'G_t(\Theta_0)] \tag{A.8}$$

is positive definite.

(S20) $H_t(\Theta_0)$ converges e.a.s. to a unique strictly stationary and ergodic sequence.

(S21) $E[G_t(\Theta_0)G_t(\Theta_0)'] < \infty$ for all t.

(S22) $(1/\sqrt{T})\sum_{t=1}^{T} E[G_t(\Theta_0)'] \to 0_{S\times 1}$ for $T\to\infty$.

(S23) $(1/\sqrt{T})\sum_{t=1}^{T} G_t(\Theta_0)' \to_d N_S(0, B_0)$ for $T\to\infty$, where:

$$B_0 = \lim_{T\to\infty} \text{Var}\left[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} G_t(\Theta_0)'\right] \tag{A.9}$$

(S24) $G_t(\Theta_0)'$ converges e.a.s. to a unique strictly stationary and ergodic sequence.

Conditions (S1) to (S9), (S12) to (S14), and (S16) to (S18) are regularity conditions, which hold for the t-QVARMA models of this paper. Conditions (S10) and (S11) are maintained assumptions.

Condition (S15) is proven in Proposition 5. Conditions (S16) to (S20) imply the law of large numbers for $H_t(\bar{\Theta})$ in (A.6) (White 2001, Theorem 3.34). Condition (S19) is proven in Proposition 4, which uses Proposition 2. Condition (S20) is proven in Proposition 7.

Conditions (S21), (S22), (S23), and (S24) imply the central limit theorem (CLT) for $G_t(\Theta_0)'$ in (A.6). For (S21), we use the following result: $E[H_t(\Theta_0)] = \text{Var}[G_t(\Theta_0)')] = E[G_t(\Theta_0)'G_t(\Theta_0)] < \infty$, where the equalities hold due to (S13) (Wooldridge 1994, p. 2675) and (S12) (Wooldridge 1994, p. 2674), respectively. Inequality $E[G_t(\Theta_0)'G_t(\Theta_0)] < \infty$ is shown in Proposition 4, which implies $E[G_t(\Theta_0)G_t(\Theta_0)'] < \infty$, because the terms of the sum defined by $G_t(\Theta_0)G_t(\Theta_0)'$ are in the diagonal of $G_t(\Theta_0)'G_t(\Theta_0)$. For condition (S22), we use the following result: $E[G_t(\Theta_0)'] = 0_{S\times 1}$, which holds under (S12) (Wooldridge 1994, p. 2674). Moreover, the time-invariant $E[G_t(\Theta_0)']$ is proven in Proposition 3, which uses Proposition 2. For condition (S23), we use White (2001, Theorem 5.16): (i) $G_t(\Theta_0)'$ is a martingale difference sequence (MDS), which holds under (S11) (Wooldridge 1994, p. 2677). Therefore, $G_t(\Theta_0)$ is an adapted mixingale (White 2001, Definition 5.15, p. 125). (ii) Condition (S24) that is proven in Proposition 6. (i) and (ii) imply (S23).

3. Propositions

Proposition 1 presents the consistency and asymptotic normality of the ML estimates.

Proposition 1: Under (S1) to (S24):

$$\sqrt{T}(\hat{\Theta} - \Theta_0) \to_d N_S\left(0_{S\times 1}, A_0^{-1}B_0A_0^{-1}\right) = N_S\left(0_{S\times 1}, A_0^{-1}\right) \quad \text{as} \quad T \to \infty \tag{A.10}$$

Proof: It follows from the works of Wooldridge (1994), Creal et al. (2013), Harvey (2013), and Blasques et al. (2017), hence it is omitted.

For Propositions 2 to 4, arguments from the work of Harvey (2013) are used. For Propositions 5 to 7, the proofs of the works of Brandt (1986), Elton (1990), and Straumann and Mikosch (2006) are used. For the proofs of Propositions 2 to 4, we use:

$$E\left[\left(U_{j,t}^*\right)^{2-i}\left(\frac{\partial U_{k,t}^*}{\partial M_{l,t}^*}\right)^i\right] < \infty \tag{A.11}$$

for $i = 0, 1, 2,\ j, k, l = 1, \ldots, K$ (Harvey 2013), which imply finite variances and covariances for the score functions and their derivatives, and which hold for the t-QVARMA model of this paper.

Proposition 2: If the maximum modulus of eigenvalues, denoted C_1, of Φ^* is less than one, and Ψ^* is non-zero, then μ_t^* is covariance stationary. **Proof:** For the equation $M_t^* = \Phi^* M_{t-1}^* + \Psi^* U_{t-1}^*$, U_t^* is i.i.d. with zero mean and a well-defined covariance matrix for $\nu > 2$. If the maximum modulus of eigenvalues of Φ^* is less than one and Ψ^* is non-zero, then μ_t^* is covariance stationary. QED

Proposition 3: The expected value of $G_t(\Theta_0)'$ is time-invariant if the maximum modulus of eigenvalues of matrix $E_2 = E(X_{t-1}^*)$ is less than one, where

$$X_{t-1}^* \equiv \Phi^* + \Psi^* \frac{\partial U_{t-1}^*}{\partial (M_{t-1}^*)'} \tag{A.12}$$

Proof: First, we focus on the following derivatives of μ_t^* with respect to $\Psi_{i,j}^*$, where $\Psi_{i,j}^*$ is an element of Ψ^*. Those derivatives by the chain rule are in $G_t(\Theta_0)'$ and $I_t(\Theta_0)$, and are given by:

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} = \Phi^* \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + \Psi^* \frac{\partial U_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \tag{A.13}$$

for $i, j = 1, \ldots, K(p+q-1)$; $W_{i,j}$ is a $[K(p+q-1) \times K(p+q-1)]$ matrix, in which element (i,j) is one and the rest of the elements are zero. Therefore, $W_{i,j} U_{t-1}^*$ is the j-th element of U_{t-1}^*. By using the chain rule, from (A.13):

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} = \left[\Phi^* + \Psi^* \frac{\partial U_{t-1}^*}{\partial (M_{t-1}^*)'} \right] \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* = X_{t-1}^* \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \tag{A.14}$$

The conditional expectation of the latter equation, conditional on $\mathcal{F}_{t-2}$, is:

$$E\left(\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} | \mathcal{F}_{t-2} \right) = E(X_{t-1}^* | \mathcal{F}_{t-2}) \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} E(U_{t-1}^* | \mathcal{F}_{t-2}) \tag{A.15}$$

where $\partial M_{t-1}^* / \partial \Psi_{i,j}^*$ is outside the conditional expectation, because it is determined by $\mathcal{F}_{t-2}$. The unconditional expectation of the latter equation is:

$$E\left(\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} \right) = E(X_{t-1}^*) E\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} \right) + W_{i,j} E(U_{t-1}^*) \tag{A.16}$$

because $E(X_{t-1}^* | \mathcal{F}_{t-2})$ and $\partial M_{t-1}^* / \partial \Psi_{i,j}^*$ are independent. The last term of (A.16) is an i.i.d. time series with zero unconditional mean. Therefore, $E(\partial M_t^* / \partial \Psi_{i,j}^*)$ is finite if the maximum modulus of eigenvalues of $E(X^*_{t-1})$ is less than one.

Second, we focus on the following derivatives of μ^*_t with respect to $\Phi^*_{i,j}$, where $\Phi^*_{i,j}$ is an element of Φ^*. Those derivatives by the chain rule are in $G_t(\Theta_0)'$ and $I_t(\Theta_0)$, and are given by:

$$\frac{\partial M^*_t}{\partial \Phi^*_{i,j}} = \left[\Phi^* + \Psi^* \frac{\partial U^*_{t-1}}{\partial (M^*_{t-1})'}\right] \frac{\partial M^*_{t-1}}{\partial \Phi^*_{i,j}} + W_{i,j} M^*_{t-1} = X^*_{t-1} \frac{\partial M^*_{t-1}}{\partial \Phi^*_{i,j}} + W_{i,j} M^*_{t-1} \tag{A.17}$$

for $i, j = 1, \ldots, K(p+q-1)$. Therefore, $W_{i,j}M^*_{t-1}$ is the j-th element of M^*_{t-1}. The conditional expectation of the latter equation, conditional on $\mathcal{F}_{t-2}$, is:

$$E\left(\frac{\partial M^*_t}{\partial \Phi^*_{i,j}} | \mathcal{F}_{t-2}\right) = E(X^*_{t-1}|\mathcal{F}_{t-2}) \frac{\partial M^*_{t-1}}{\partial \Phi^*_{i,j}} + W_{i,j} E(M^*_{t-1}|\mathcal{F}_{t-2}) \tag{A.18}$$

where $\partial M^*_{t-1}/\partial \Phi^*_{i,j}$ is outside the conditional expectation, because it is determined by $\mathcal{F}_{t-2}$. The unconditional expectation of the latter equation is:

$$E\left(\frac{\partial M^*_t}{\partial \Phi^*_{i,j}}\right) = E(X^*_{t-1}) E\left(\frac{\partial M^*_{t-1}}{\partial \Phi^*_{i,j}}\right) + W_{i,j} E(M^*_{t-1}) \tag{A.19}$$

because $E(X^*_{t-1}|\mathcal{F}_{t-2})$ and $\partial M^*_{t-1}/\partial \Phi^*_{i,j}$ are independent. Under the assumption that the maximum modulus of eigenvalues of Φ^* is less than one, $W_{i,j}E(M^*_{t-1}) = 0$. Therefore, $E(\partial M^*_t/\partial \Phi^*_{i,j})$ is finite if the maximum modulus of eigenvalues of $E(X^*_{t-1})$ is less than one. QED

Proposition 4: Matrix $E[G_t(\Theta_0)'G_t(\Theta_0)]$ is time-invariant if the maximum modulus of eigenvalues of matrix $E_3 = E(X^*_{t-1} \otimes X^*_{t-1})$ is less than one, where $\otimes$ is the Kronecker product. **Proof:** First, we focus on a particular derivative, which contributes to the information matrix. From (A.14), the following derivative is expressed:

$$\frac{\partial M^*_t}{\partial \Psi^*_{i,j}} \left(\frac{\partial M^*_t}{\partial \Psi^*_{k,l}}\right)' = X^*_{t-1} \frac{\partial M^*_{t-1}}{\partial \Psi^*_{i,j}} \left(\frac{\partial M^*_{t-1}}{\partial \Psi^*_{k,l}}\right)' (X^*_{t-1})' + X^*_{t-1} \frac{\partial M^*_{t-1}}{\partial \Psi^*_{i,j}} (U^*_{t-1})' W'_{k,l} \tag{A.20}$$

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

which, by using the equation $\text{vec}(ABC) = (C' \otimes A)\text{vec}(B)$, where $\text{vec}(x)$ indicates that the columns of the matrix are being stacked one upon the other, (A.20) can be written as:

$$\text{vec}\left[\frac{\partial M_t^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Psi_{k,l}^*}\right)'\right] = (X_{t-1}^* \otimes X_{t-1}^*)\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'\right] \tag{A.21}$$

$$+[(W_{k,l}U_{t-1}^*) \otimes X_{t-1}^*]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right) + [X_{t-1}^* \otimes W_{i,j}U_{t-1}^*]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'$$

$$+\text{vec}\left[W_{i,j}U_{t-1}^*(U_{t-1}^*)'W_{k,l}'\right]$$

The conditional expectation of the latter equation, conditional on $\mathcal{F}_{t-2}$, is:

$$E\left\{\text{vec}\left[\frac{\partial M_t^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Psi_{k,l}^*}\right)'\right]|\mathcal{F}_{t-2}\right\} = E(X_{t-1}^* \otimes X_{t-1}^*|\mathcal{F}_{t-2})\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'\right] \tag{A.22}$$

$$+E[(W_{k,l}U_{t-1}^*) \otimes X_{t-1}^*|\mathcal{F}_{t-2}]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right) + E[X_{t-1}^* \otimes W_{i,j}U_{t-1}^*|\mathcal{F}_{t-2}]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'$$

$$+\text{vec}\left\{W_{i,j}E\left[U_{t-1}^*(U_{t-1}^*)'|\mathcal{F}_{t-2}\right]W_{k,l}'\right\}$$

where $\text{vec}[(\partial M_{t-1}^*/\partial \Psi_{i,j}^*)(\partial M_{t-1}^*/\partial \Psi_{k,l}^*)']$, $\text{vec}[(\partial M_{t-1}^*/\partial \Psi_{i,j}^*)]$, and $\text{vec}[(\partial M_{t-1}^*/\partial \Psi_{k,l}^*)']$ are outside the conditional expectations, because they are determined by $\mathcal{F}_{t-2}$. The unconditional expectation of (A.22) is:

$$E\left\{\text{vec}\left[\frac{\partial M_t^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Psi_{k,l}^*}\right)'\right]\right\} = E(X_{t-1}^* \otimes X_{t-1}^*)E\left\{\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'\right]\right\}+ \tag{A.23}$$

$$E\left[(W_{k,l}U_{t-1}^*) \otimes X_{t-1}^*\right]E\left\{\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\right)\right\} + E\left[X_{t-1}^* \otimes (W_{i,j}U_{t-1}^*)\right]E\left\{\text{vec}\left[\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'\right]\right\}$$

$$+\text{vec}\left\{W_{i,j}E\left[U_{t-1}^*(U_{t-1}^*)'\right]W_{k,l}'\right\}$$

because

(i) $E(X_{t-1}^* \otimes X_{t-1}^*|\mathcal{F}_{t-2})$ and $\text{vec}[(M_{t-1}^*/\partial \Psi_{i,j}^*)(\partial M_{t-1}^*/\partial \Psi_{k,l}^*)']$ are independent;

(ii) $E[(W_{k,l}U_{t-1}^*) \otimes X_{t-1}^*|\mathcal{F}_{t-2}]$ and $\text{vec}(\partial M_{t-1}^*/\partial \Psi_{i,j}^*)$ are independent;

(iii) $E[X_{t-1}^* \otimes (W_{i,j}U_{t-1}^*)|\mathcal{F}_{t-2}]$ and $\text{vec}[(\partial M_{t-1}^*/\partial \Psi_{k,l}^*)']$ are independent.

The unconditional means in the second and third terms of (A.23) are finite because of (A.11).

The unconditional mean of the last term of (A.23) is finite, because $E[U_{t-1}^*(U_{t-1}^*)']$ is well-defined for $\nu > 2$. Therefore, $E[(\partial M_t^*/\partial \Psi_{i,j}^*)(\partial M_t^*/\partial \Psi_{k,l}^*)']$ is finite if the maximum modulus of eigenvalues of $E(X_{t-1}^* \otimes X_{t-1}^*)$ is less than one.

Second, from (A.17) we express:

$$\text{vec}\left[\frac{\partial M_t^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Phi_{k,l}^*}\right)'\right] = (X_{t-1}^* \otimes X_{t-1}^*)\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right] \tag{A.24}$$

$$[(W_{k,l}M_{t-1}^*) \otimes X_{t-1}^*]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\right) + [X_{t-1}^* \otimes (W_{i,j}M_{t-1}^*)]\text{vec}\left[\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right]$$

$$+\text{vec}\left[W_{i,j}M_{t-1}^*(M_{t-1}^*)'W_{k,l}'\right]$$

The conditional expectation of the latter equation, conditional on $\mathcal{F}_{t-2}$, is:

$$E\left\{\text{vec}\left[\frac{\partial M_t^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Phi_{k,l}^*}\right)'\right]|\mathcal{F}_{t-2}\right\} = E(X_{t-1}^* \otimes X_{t-1}^*|\mathcal{F}_{t-2})\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right] \tag{A.25}$$

$$+E[(W_{k,l}M_{t-1}^*) \otimes X_{t-1}^*|\mathcal{F}_{t-2}]\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\right) + E[X_{t-1}^* \otimes (W_{i,j}M_{t-1}^*)|\mathcal{F}_{t-2}]\text{vec}\left[\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right]$$

$$+\text{vec}\left\{W_{i,j}E\left[M_{t-1}^*(M_{t-1}^*)'|\mathcal{F}_{t-2}\right]W_{k,l}'\right\}$$

The unconditional expectation of (A.25) is:

$$E\left\{\text{vec}\left[\frac{\partial M_t^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Phi_{k,l}^*}\right)'\right]\right\} = E(X_{t-1}^* \otimes X_{t-1}^*)E\left\{\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right]\right\} \tag{A.26}$$

$$+E[(W_{k,l}M_{t-1}^*) \otimes X_{t-1}^*]E\left[\text{vec}\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\right)\right] + E[X_{t-1}^* \otimes (W_{i,j}M_{t-1}^*)]E\left\{\text{vec}\left[\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{k,l}^*}\right)'\right]\right\}$$

$$+\text{vec}\left\{W_{i,j}E\left[M_{t-1}^*(M_{t-1}^*)'\right]W_{k,l}'\right\}$$

because

(i) $E(X_{t-1}^* \otimes X_{t-1}^*|\mathcal{F}_{t-2})$ and $\text{vec}[(M_{t-1}^*/\partial \Phi_{i,j}^*)(\partial M_{t-1}^*/\partial \Phi_{k,l}^*)']$ are independent;

(ii) $E[(W_{k,l}M_{t-1}^*) \otimes X_{t-1}^*|\mathcal{F}_{t-2}]$ and $\text{vec}(\partial M_{t-1}^*/\partial \Phi_{i,j}^*)$ are independent;

(iii) $E[X_{t-1}^* \otimes (W_{i,j}M_{t-1}^*)|\mathcal{F}_{t-2}]$ and $\text{vec}[(\partial M_{t-1}^*/\partial \Phi_{k,l}^*)']$ are independent.

Under the assumption that the maximum modulus of eigenvalues of Φ^* is less than one and under (A.11), (i) the unconditional means in the second and third terms of (A.26) are finite, and (ii) the unconditional mean of the last term in (A.26) is finite because $E[M^*_{t-1}(M^*_{t-1})']$ is well-defined for $\nu > 2$. Therefore, $E[(\partial M^*_t/\partial \Phi^*_{i,j})(\partial M^*_t/\partial \Phi^*_{k,l})']$ is finite if the maximum modulus of eigenvalues of $E(X^*_{t-1} \otimes X^*_{t-1})$ is less than one. QED

For the conditions of ML estimation of filter $M^\dagger_t$, Propositions 3 and 4 are modified, by using the parameter matrices of filter $M^\dagger_t$ (Harvey 2013):

$$M^\dagger_t = \Phi^\dagger M^\dagger_{t-1} + \Psi^\dagger U^\dagger_{t-1} \tag{A.27}$$

Moreover, for the ML estimation of $M^\dagger_t$, the dynamic parameter of $\mu^\dagger_t$ is set to the identity matrix (see the definition of $\Phi^\dagger$ after (9)), and it is not estimated (Harvey 2013, pp. 45-46 and pp. 210-212). Therefore, for the modified versions of Propositions 3 and 4 with respect to the parameters of filter $M^\dagger_t$, the asymptotic properties of ML hold (Harvey 2013).

Proposition 5: Filter $M^*_t = \Phi^* M^*_{t-1} + \Psi^* U^*_{t-1}$ converges e.a.s. to a unique strictly stationary and ergodic sequence if: (i) the following Lyapunov exponent is negative:

$$\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\frac{\partial M^*_t}{\partial (M^*_{t-1})'}\right|\right|_1\right]\right\} = \inf_{n\geq 1}\left\{n^{-1}E\left(\ln\left|\left|\prod_{t=1}^{n} X^*_{t-1}\right|\right|_1\right)\right\} < 0 \tag{A.28}$$

(ii) $E(\ln^+ ||X^*||_1) < \infty$, where $||X^*||_1 \equiv \sup\{||X^*_1 x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$, and $\ln^+(x) = 0$ if $0 \leq x \leq 1$ and $\ln^+(x) = \ln(x)$ if $x > 1$. (iii) $E(\ln^+ ||\Psi^* U^*||_1) < \infty$, where $||\Psi^* U^*||_1 \equiv \sup\{||\Psi^* U^*_1 x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$. (iv) $\Psi^* U^*_t$ is strictly stationary and ergodic. (v) X^*_t is strictly stationary and ergodic. For conditions (i) to (iii), the matrix norm $||A||_1 = \max_{1\leq j\leq Kp}\sum_{i=1}^{Kp}|a_{i,j}|$ is used, where $A = \{a_{i,j}\}$ for $i, j = 1, \ldots, Kp$. **Proof:** It follows from the proofs of the works of Brandt (1986), Elton (1990), and Straumann and Mikosch (2006), hence it is omitted.

$G_t(\Theta_0)'$ converges e.a.s. to a unique strictly stationary and ergodic sequence, if the first derivatives of M^*_t and $M^\dagger_t$ with respect to Θ, at the true parameter values, converge e.a.s. to unique strictly stationary and ergodic sequences. In Proposition 6, we present the conditions for M^*_t, which can be modified for $M^\dagger_t$ by using the parameter matrices of filter $M^\dagger_t$.

Proposition 6: Filter $\partial M_t^*/\partial\Theta'$ from Proposition 3, can be represented by the following equations:

$$\frac{\partial M_t^*}{\partial \Psi_{i,j}^*} = X_{t-1}^* \frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*} + W_{i,j} U_{t-1}^* \tag{A.29}$$

$$\frac{\partial M_t^*}{\partial \Phi_{i,j}^*} = X_{t-1}^* \frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*} + W_{i,j} M_{t-1}^* \tag{A.30}$$

for $i, j = 1, \ldots, K(p+q-1)$; $W_{i,j}$ is a $[K(p+q-1) \times K(p+q-1)]$ matrix, in which element (i, j) is one and the rest of the elements are zero. Filter $\partial M_t^*/\partial\Theta'$ converges e.a.s. to a unique strictly stationary and ergodic sequence if: (i) the following Lyapunov exponents are negative:

$$\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\frac{\partial \frac{\partial M_t^*}{\partial \Psi_{i,j}^*}}{\partial\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\right)'}\right|\right|_1\right]\right\} \equiv \inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\tilde{\Psi}_{t-1}^{(1)}\right|\right|_1\right]\right\} < 0 \tag{A.31}$$

$$\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\frac{\partial \frac{\partial M_t^*}{\partial \Phi_{i,j}^*}}{\partial\left(\frac{\partial M_{t-1}^*}{\partial \Phi_{i,j}^*}\right)'}\right|\right|_1\right]\right\} \equiv \inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\tilde{\Phi}_{t-1}^{(1)}\right|\right|_1\right]\right\} < 0 \tag{A.32}$$

where we define $\tilde{\Psi}_{t-1}^{(1)}$ and $\tilde{\Phi}_{t-1}^{(1)}$, respectively. (ii) $E(\ln^+ ||\tilde{\Psi}^{(1)}||_1) < \infty$ and $E(\ln^+ ||\tilde{\Phi}^{(1)}||_1) < \infty$, where $||\tilde{\Psi}^{(1)}||_1 \equiv \sup\{||\tilde{\Psi}_1^{(1)}x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$ and $||\tilde{\Phi}^{(1)}||_1 \equiv \sup\{||\tilde{\Phi}_1^{(1)}x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$, respectively. (iii) $E(\ln^+ ||W_{i,j}U^*||_1) < \infty$ and $E(\ln^+ ||W_{i,j}M^*||_1) < \infty$, where $||W_{i,j}U^*||_1 \equiv \sup\{||W_{i,j}U_1^* x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$ and $||W_{i,j}M^*||_1 \equiv \sup\{||W_{i,j}M_1^* x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$, respectively. (iv) $W_{i,j}U_t^*$ and $W_{i,j}M_t^*$ are strictly stationary and ergodic. (v) X_t^* is strictly stationary and ergodic. **Proof:** It follows from the proofs of the works of Brandt (1986), Elton (1990), and Straumann and Mikosch (2006), hence it is omitted.

$G_t(\Theta_0)'G_t(\Theta_0) = H_t(\Theta_0)$ converges e.a.s. to a unique strictly stationary and ergodic sequence, if $(\partial M_t^*/\partial\Theta') \times (\partial M_t^*/\partial\Theta')'$ and $(\partial M_t^\dagger/\partial\Theta') \times (\partial M_t^\dagger/\partial\Theta')'$, at the true parameter values, converge e.a.s. to unique strictly stationary and ergodic sequences. In Proposition 7, we present the conditions for M_t^*, which can be modified for $M_t^\dagger$ by using the parameter matrices of filter $M_t^\dagger$.

Proposition 7: Filter $(\partial M_t^*/\partial\Theta') \times (\partial M_t^*/\partial\Theta')'$ from Proposition 4, can be represented by the following equations:

$$\text{vec}\left[\frac{\partial M_t^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_t^*}{\partial \Psi_{k,l}^*}\right)'\right] = (X_{t-1}^* \otimes X_{t-1}^*)\text{vec}\left[\frac{\partial M_{t-1}^*}{\partial \Psi_{i,j}^*}\left(\frac{\partial M_{t-1}^*}{\partial \Psi_{k,l}^*}\right)'\right] \tag{A.33}$$

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

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

and

$$\text{vec}\left[\frac{\partial M^*_{t}}{\partial \Phi^*_{i,j}}\left(\frac{\partial M^*_{t}}{\partial \Phi^*_{k,l}}\right)'\right]=(X^*_{t-1}\otimes X^*_{t-1})\text{vec}\left[\frac{\partial M^*_{t-1}}{\partial \Phi^*_{i,j}}\left(\frac{\partial M^*_{t-1}}{\partial \Phi^*_{k,l}}\right)'\right] \tag{A.34}$$

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

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

where we define $\tilde{U}^*_{t-1}$ and $\tilde{M}^*_{t-1}$, respectively, for ease of notation.

Filter $(\partial M^*_t/\partial\Theta')\times(\partial M^*_t/\partial\Theta')'$ converges e.a.s. to a unique strictly stationary and ergodic sequence if: (i) the following Lyapunov exponents are negative:

$$\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\left[\frac{\partial\frac{\partial M^*_t}{\partial\Psi^*_{i,j}}}{\partial\left(\frac{\partial M^*_{t-1}}{\partial\Psi^*_{i,j}}\right)'}\right]\times\left[\frac{\partial\frac{\partial M^*_t}{\partial\Psi^*_{i,j}}}{\partial\left(\frac{\partial M^*_{t-1}}{\partial\Psi^*_{i,j}}\right)'}\right]'\right|\right|_1\right]\right\} \tag{A.35}$$

$$\equiv\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\tilde{\Psi}^{(2)}_{t-1}\right|\right|_1\right]\right\}<0$$

and

$$\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\left[\frac{\partial\frac{\partial M^*_t}{\partial\Phi^*_{i,j}}}{\partial\left(\frac{\partial M^*_{t-1}}{\partial\Phi^*_{i,j}}\right)'}\right]\times\left[\frac{\partial\frac{\partial M^*_t}{\partial\Phi^*_{i,j}}}{\partial\left(\frac{\partial M^*_{t-1}}{\partial\Phi^*_{i,j}}\right)'}\right]'\right|\right|_1\right]\right\} \tag{A.36}$$

$$\equiv\inf_{n\geq 1}\left\{n^{-1}E\left[\ln\left|\left|\prod_{t=1}^{n}\tilde{\Phi}^{(2)}_{t-1}\right|\right|_1\right]\right\}<0$$

where we define $\tilde{\Psi}^{(2)}_{t-1}$ and $\tilde{\Phi}^{(2)}_{t-1}$, respectively. (ii) $E(\ln^+||\tilde{\Psi}^{(2)}||_1)<\infty$ and $E(\ln^+||\tilde{\Phi}^{(2)}||_1)<\infty$, where $||\tilde{\Psi}^{(2)}||_1\equiv\sup\{||\tilde{\Psi}^{(2)}_1 x||_1 : x\in\mathbb{R}^{Kp}, ||x||_1\leq 1\}$ and $||\tilde{\Phi}^{(2)}||_1\equiv\sup\{||\tilde{\Phi}^{(2)}_1 x||_1 : x\in\mathbb{R}^{Kp}, ||x||_1\leq 1\}$, respectively. (iii) $E(\ln^+||\tilde{U}^*||_1)<\infty$ and $E(\ln^+||\tilde{M}^*||_1)<\infty$, where $||\tilde{U}^*||_1\equiv\sup\{||\tilde{U}^*_1 x||_1 : x\in$

$\mathbb{R}^{Kp}, ||x||_1 \leq 1\}$ and $||\tilde{M}^*||_1 \equiv \sup\{||\tilde{M}_1^* x||_1 : x \in \mathbb{R}^{Kp}, ||x||_1 \leq 1\}$, respectively. (iv) $\tilde{U}_t^*$ and $\tilde{M}_t^*$ are strictly stationary and ergodic. (v) $(X_{t-1}^* \otimes X_{t-1}^*)$ is strictly stationary and ergodic. **Proof:** It follows from the proofs of the works of Brandt (1986), Elton (1990), and Straumann and Mikosch (2006), hence it is omitted.

Supplementary Material B: t-QVARMA(1,1,1) and t-QVARMA(3,1,1)

The IRF estimates for t-QVARMA(1,1,1) are presented in Figures B1 to B3. The estimates of $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B1. The estimates of $\mathrm{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B2. The estimates of $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t} + \mathrm{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B3.

The IRF estimates for t-QVARMA(3,1,1) are presented in Figures B4 to B6. The estimates of $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B4. The estimates of $\mathrm{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B5. The estimates of $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t} + \mathrm{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure B6.

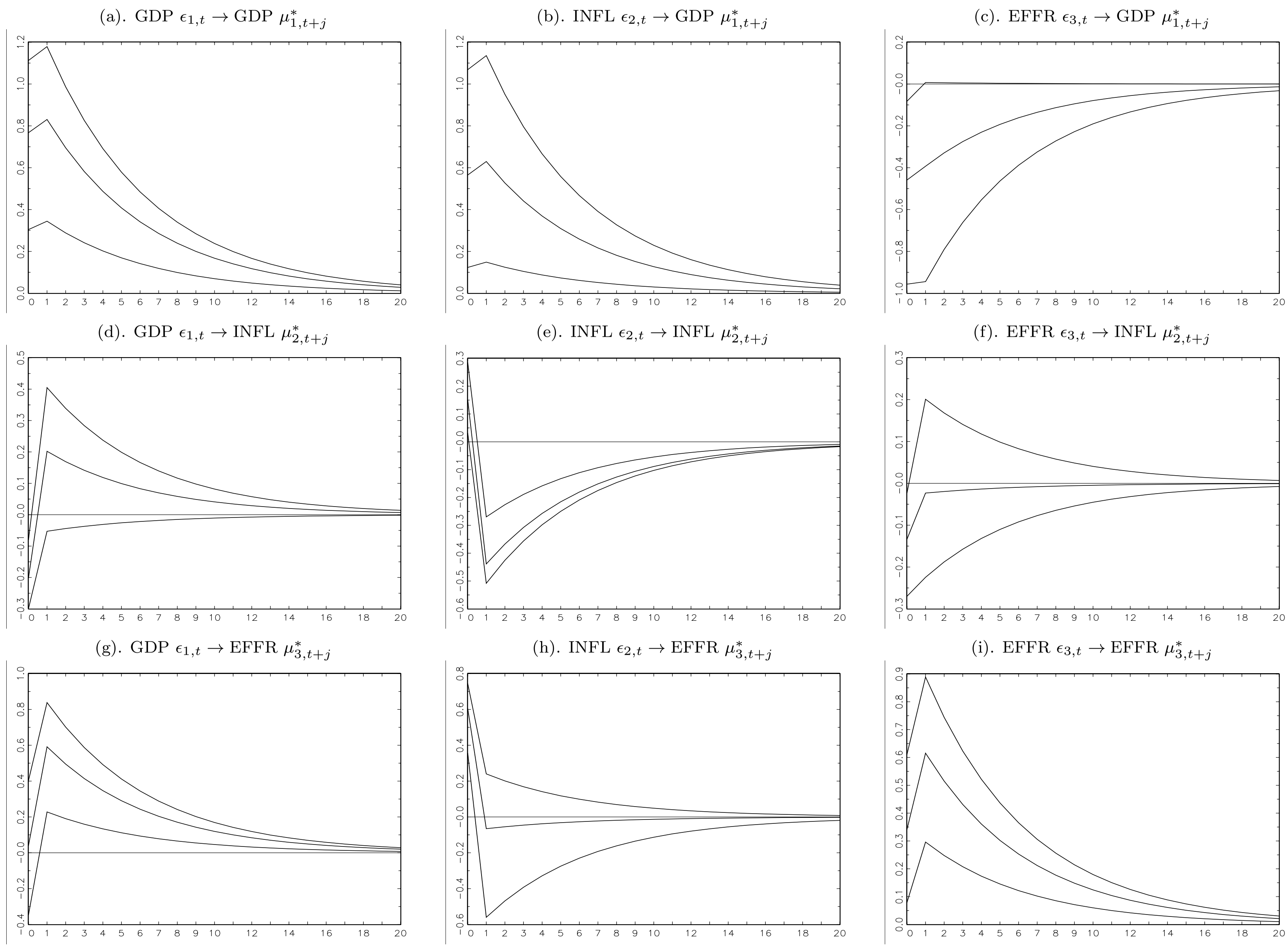

Figure B1: Short-run $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(1,1,1). *Notes*: The confidence interval is for 8,975 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

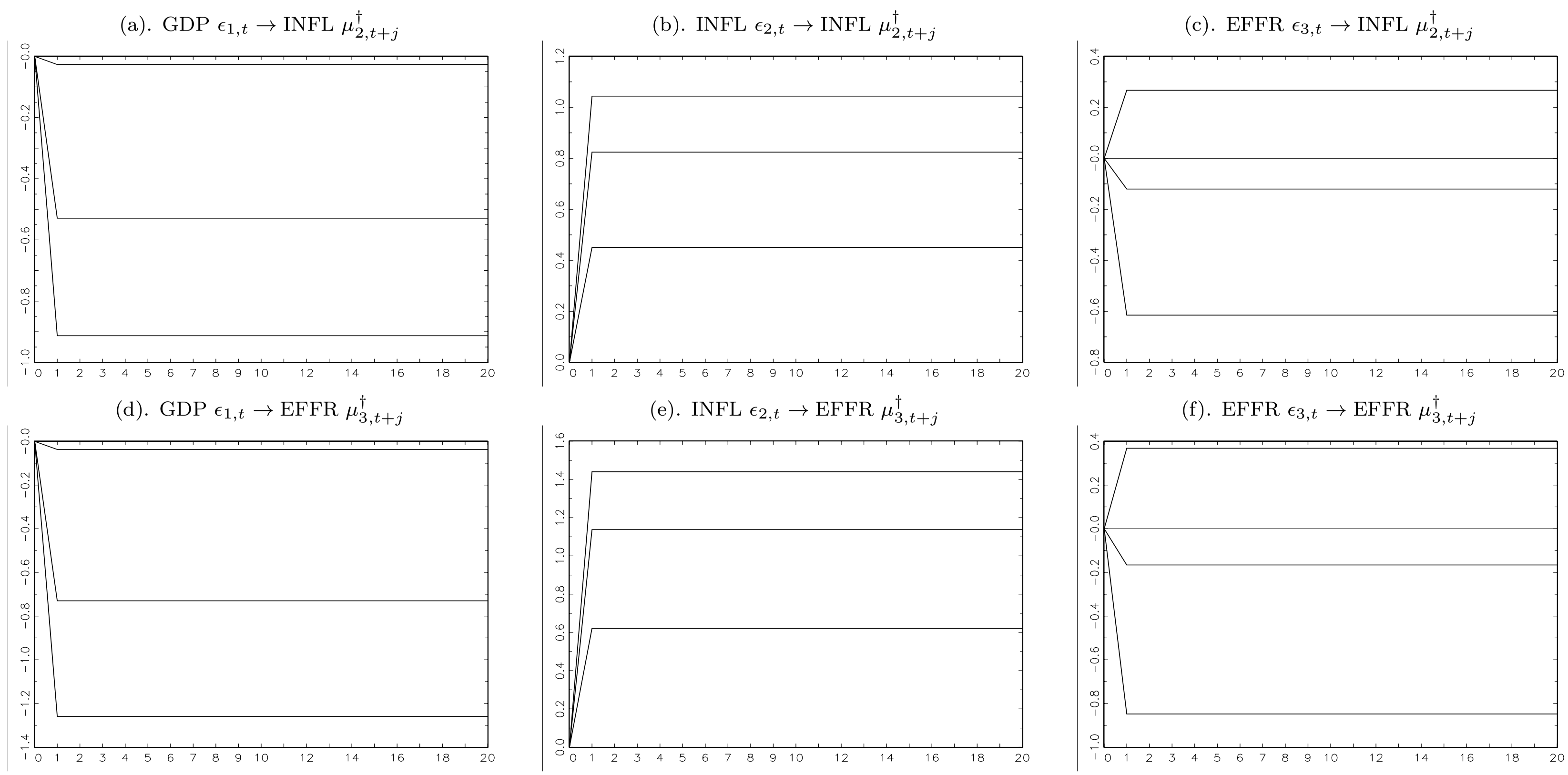

Figure B2: Long-run $\text{IRF}^{\dagger}_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(1,1,1). *Notes*: The confidence interval is for 8,975 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

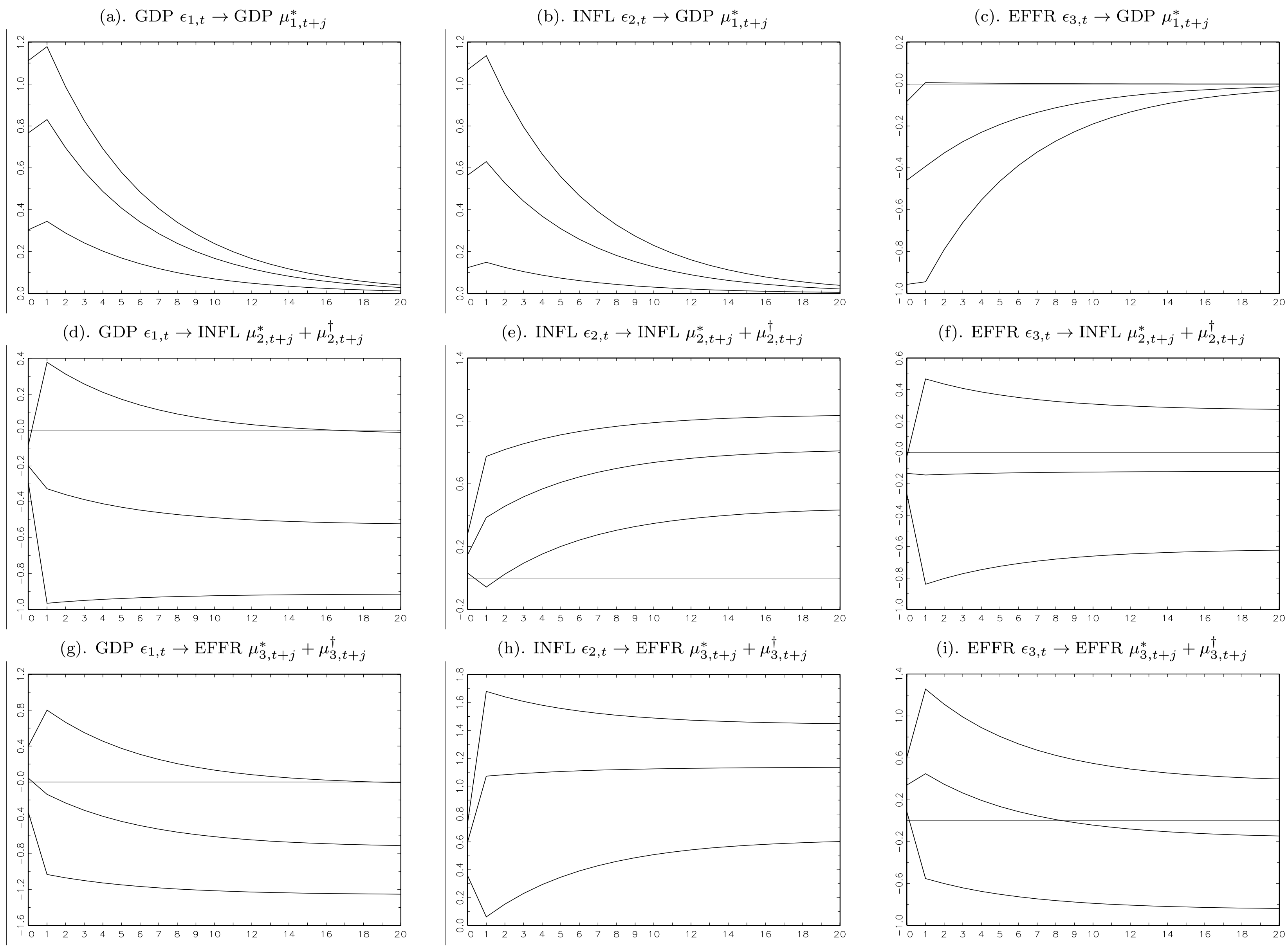

Figure B3: Total $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^\dagger_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(1,1,1). *Notes*: The confidence interval is for 8,975 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

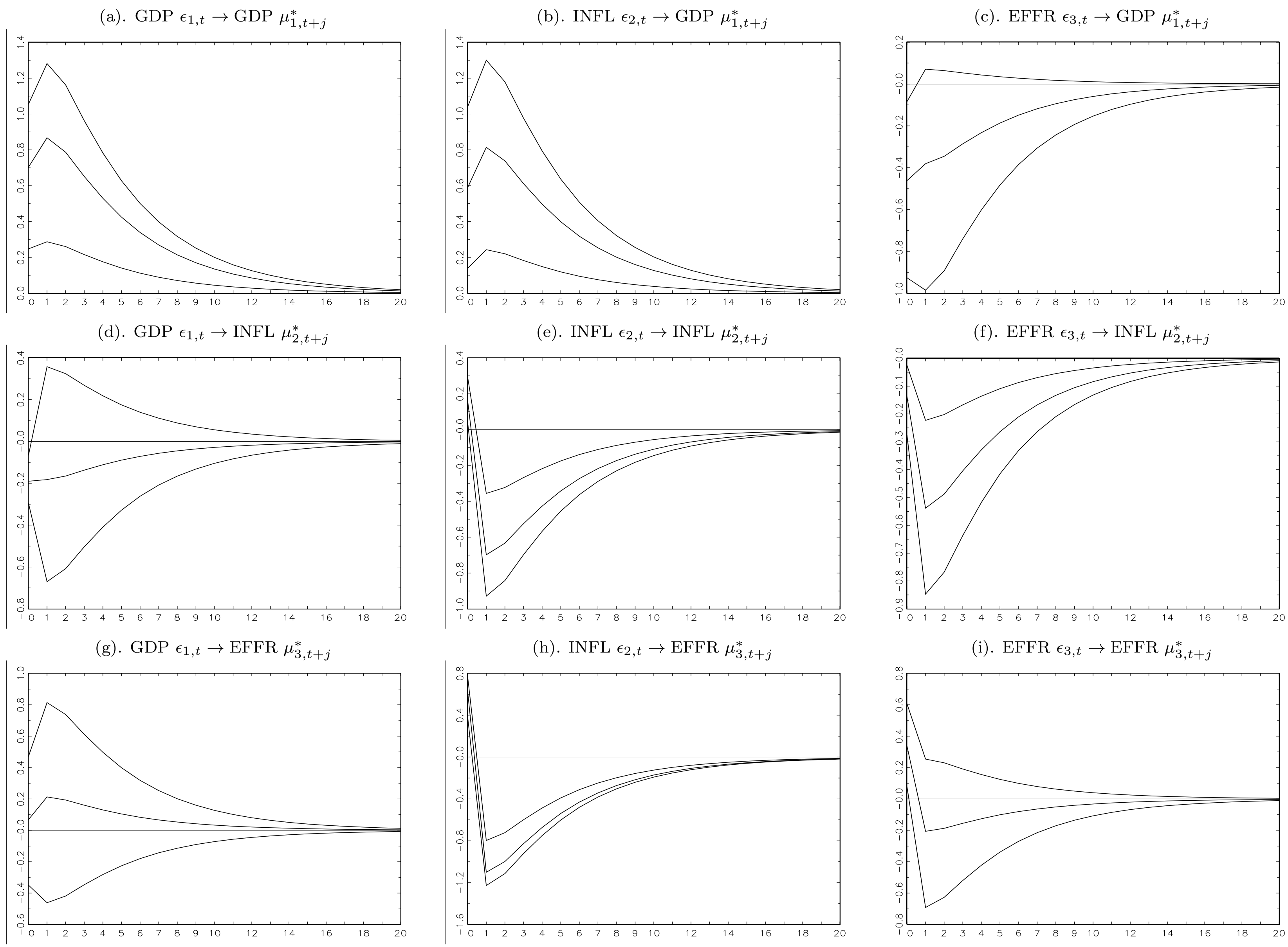

Figure B4: Short-run $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(3,1,1). *Notes*: The confidence interval is for 9,903 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

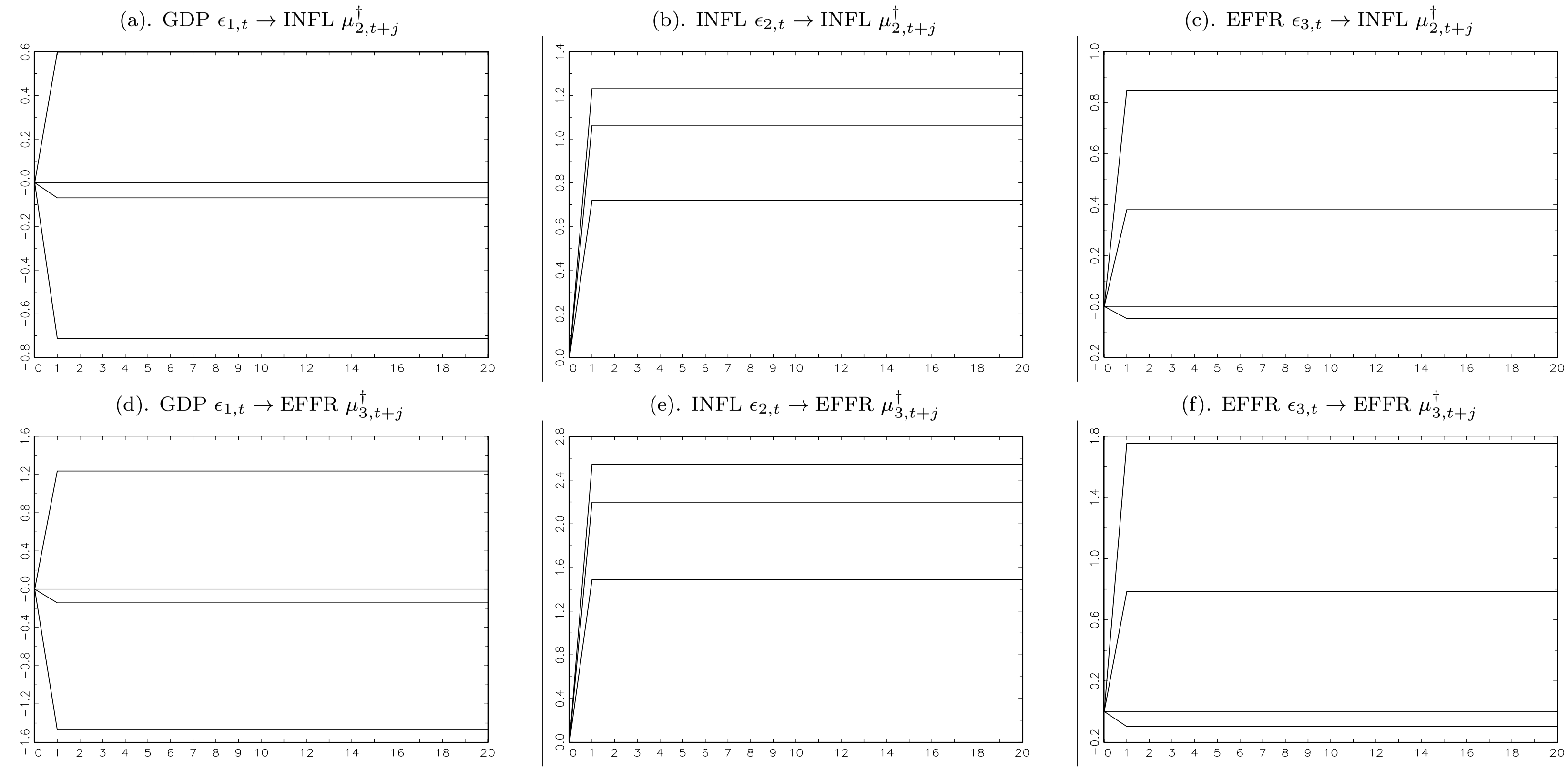

Figure B5: Long-run $\text{IRF}^{\dagger}_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(3,1,1). *Notes*: The confidence interval is for 9,903 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

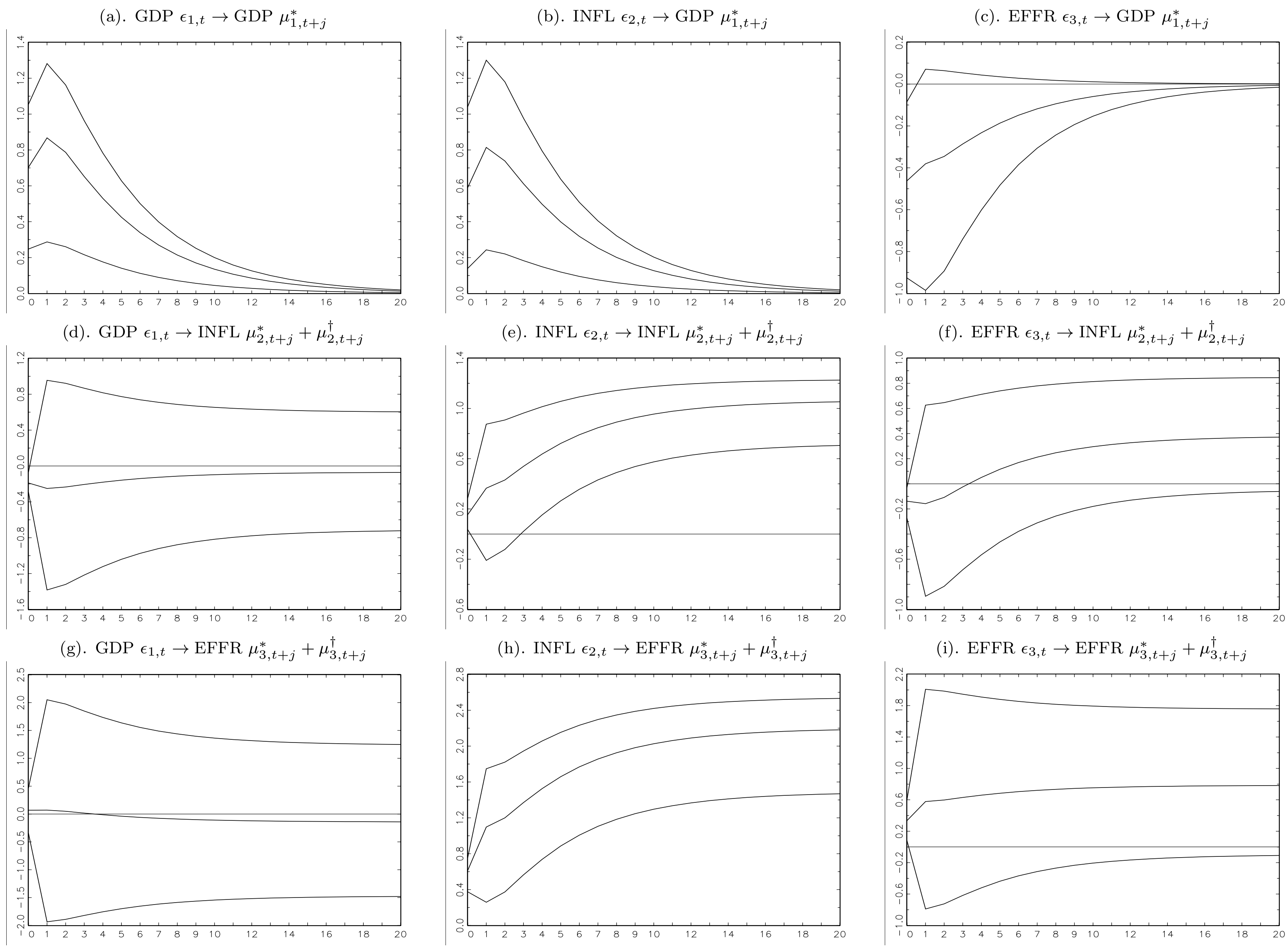

Figure B6: Total $\text{IRF}^*_{j,t} + \text{IRF}^\dagger_{j,t} + \text{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(3,1,1). *Notes*: The confidence interval is for 9,903 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

Supplementary Material C: t-QVARMA(2,1,1) for $\beta_2 = 1$

The IRF estimates for t-QVARMA(2,1,1) for $\beta_2 = 1$ are presented in Figures C1 to C3. The estimates of $\mathrm{IRF}^{*}_{j,t} + \mathrm{IRF}^{v}_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure C1. The estimates of $\mathrm{IRF}^{\dagger}_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure C2. The estimates of $\mathrm{IRF}^{*}_{j,t} + \mathrm{IRF}^{v}_{j,t} + \mathrm{IRF}^{\dagger}_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure C3.

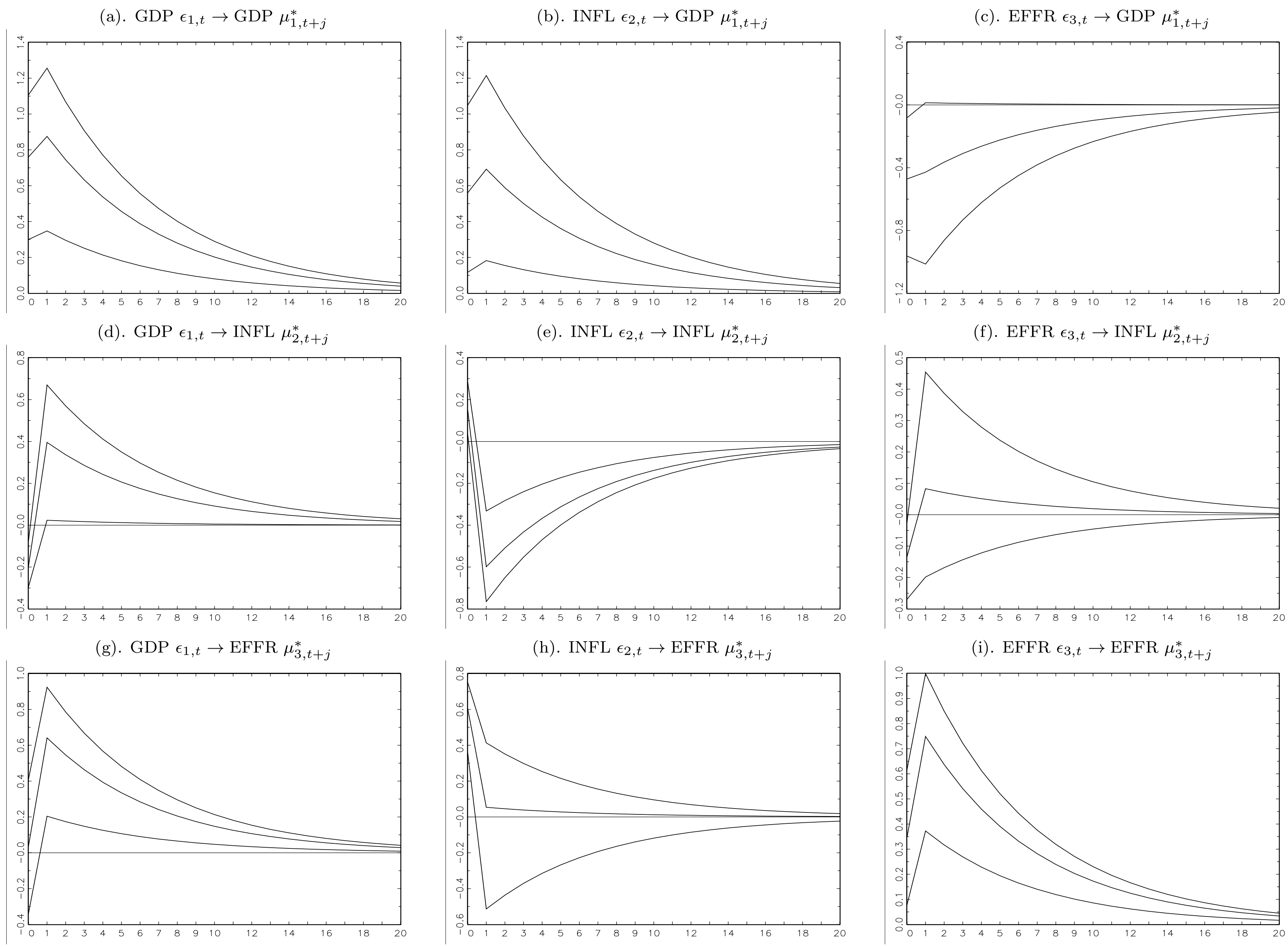

Figure C1: Short-run $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(2,1,1) with $\beta_2 = 1$. *Notes*: The confidence interval is for 9,402 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

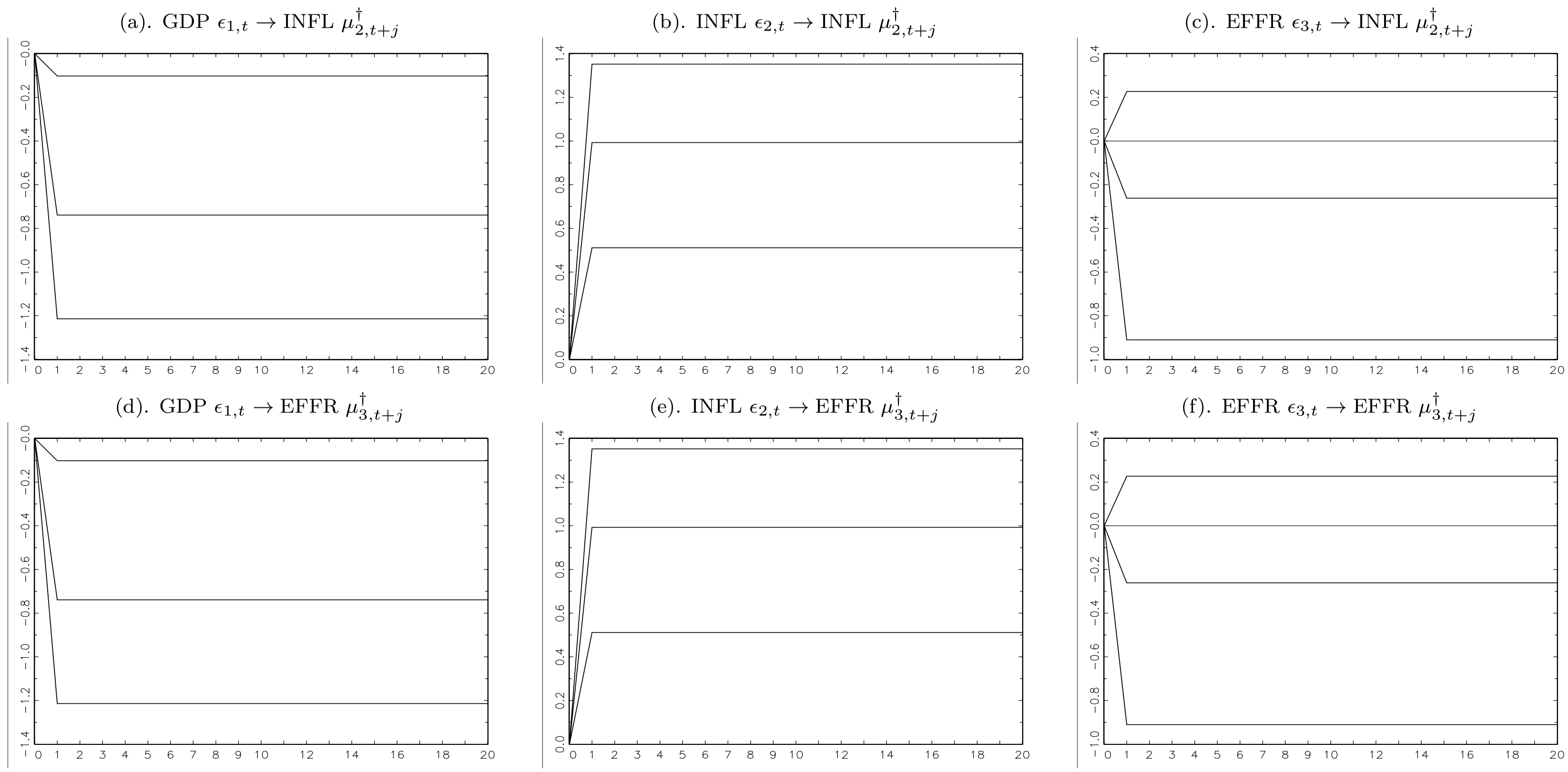

Figure C2: Long-run $\text{IRF}^{\dagger}_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(2,1,1) with $\beta_2 = 1$. *Notes*: The confidence interval is for 9,402 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

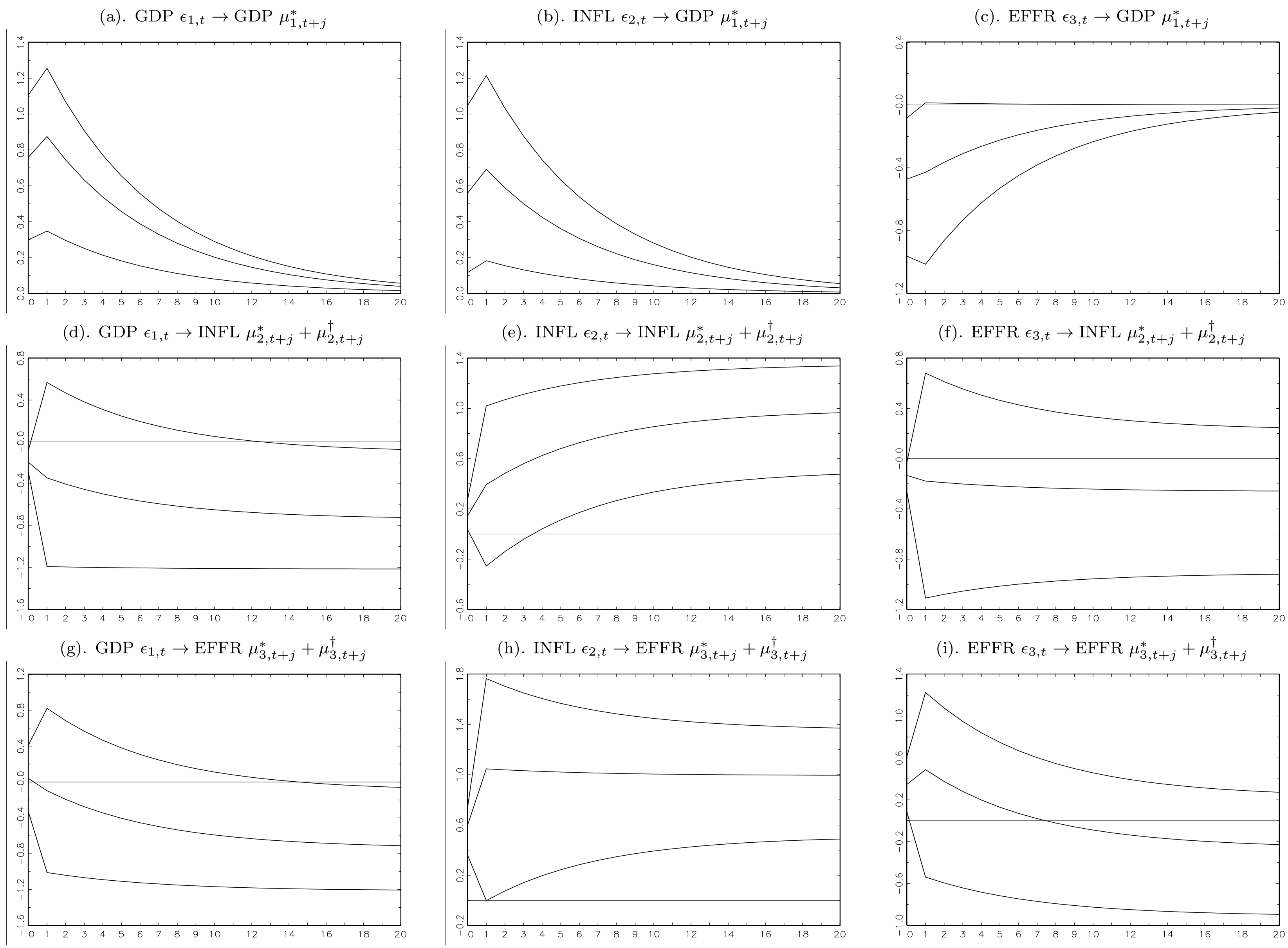

Figure C3: Total $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^\dagger_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for t-QVARMA(2,1,1) with $\beta_2 = 1$. *Notes*: The confidence interval is for 9,402 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

Supplementary Material D: Limiting special case of t-QVARMA

Gaussian-QVARMA

If $\nu \to \infty$ for t-QVARMA(p,q,r), 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):

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

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

$$\mu_t^\dagger = \mu_{t-1}^\dagger + \sum_{l=1}^{r} \Psi_l^\dagger v_{t-l} \tag{D.3}$$

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

where K^* variables are $I(0)$ and $K^\dagger$ co-integrated variables are $I(1)$. By using the notation of the first-order representation (8) and the notation of the IRFs for t-QVARMA, the IRFs of Gaussian-QVARMA can be written as: $\text{IRF}_{j,t} = \text{IRF}_{j,t}^* + \text{IRF}_{j,t}^\dagger + \text{IRF}_{j,t}^v$, where the short-run effects are $\text{IRF}_{j,t}^* = J(\Phi^*)^{j-1}\Psi^* J'\Omega^{-1}$ for $j = 1, \ldots, \infty$, the long-run effects are given by:

$$\begin{aligned}
\text{IRF}_{1,t}^\dagger &= \partial\mu_{t+1}^\dagger/\partial\epsilon_t = \Psi_1^\dagger\Omega^{-1} \\
\text{IRF}_{2,t}^\dagger &= \partial\mu_{t+2}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \Psi_2^\dagger)\Omega^{-1} \\
&\ \vdots \\
\text{IRF}_{K-1,t}^\dagger &= \partial\mu_{t+K-1}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \ldots + \Psi_{K-1}^\dagger)\Omega^{-1} \\
\text{IRF}_{j,t}^\dagger &= \partial\mu_{t+j}^\dagger/\partial\epsilon_t = (\Psi_1^\dagger + \ldots + \Psi_K^\dagger)\Omega^{-1} \quad \text{for} \quad j \geq K
\end{aligned} \tag{D.5}$$

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

Empirical application

In the empirical application, the following Gaussian-QVARMA(2,1,1) is estimated (the VAR lag-order is chosen by using the AIC, BIC, and HQC metrics):

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \\ y_{3,t} \end{bmatrix} = \begin{bmatrix} c_1^* \\ c_2^* \\ c_3^* \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^* \\ \mu_{2,t}^* \\ \mu_{3,t}^* \end{bmatrix} + \begin{bmatrix} \mu_{1,t}^\dagger \\ \mu_{2,t}^\dagger \\ \mu_{3,t}^\dagger \end{bmatrix} + \begin{bmatrix} v_{1,t} \\ v_{2,t} \\ v_{3,t} \end{bmatrix} \tag{D.6}$$

$$\begin{bmatrix} \mu^*_{1,t} \\ \mu^*_{2,t} \\ \mu^*_{3,t} \end{bmatrix} = \phi^*_1 \begin{bmatrix} \mu^*_{1,t-1} \\ \mu^*_{2,t-1} \\ \mu^*_{3,t-1} \end{bmatrix} + \phi^*_2 \begin{bmatrix} \mu^*_{1,t-2} \\ \mu^*_{2,t-2} \\ \mu^*_{3,t-2} \end{bmatrix} + \begin{bmatrix} \Psi^*_{1,11} & \Psi^*_{1,12} & \Psi^*_{1,13} \\ \Psi^*_{1,21} & \Psi^*_{1,22} & \Psi^*_{1,23} \\ \Psi^*_{1,31} & \Psi^*_{1,32} & \Psi^*_{1,33} \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \end{bmatrix} \tag{D.7}$$

$$\begin{bmatrix} \mu^\dagger_{1,t} \\ \mu^\dagger_{2,t} \\ \mu^\dagger_{3,t} \end{bmatrix} = \begin{bmatrix} \mu^\dagger_{1,t-1} \\ \mu^\dagger_{2,t-1} \\ \mu^\dagger_{3,t-1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & \Psi^\dagger_{1,22} & \Psi^\dagger_{1,23} \\ 0 & \beta_2 \Psi^\dagger_{1,22} & \beta_2 \Psi^\dagger_{1,23} \end{bmatrix} \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \end{bmatrix} \tag{D.8}$$

$$v_t \sim t_3 \left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Omega^{-1}_{11} & 0 & 0 \\ \Omega^{-1}_{21} & \Omega^{-1}_{22} & 0 \\ \Omega^{-1}_{31} & \Omega^{-1}_{32} & \Omega^{-1}_{33} \end{bmatrix} \times \begin{bmatrix} \Omega^{-1}_{11} & \Omega^{-1}_{21} & \Omega^{-1}_{31} \\ 0 & \Omega^{-1}_{22} & \Omega^{-1}_{32} \\ 0 & 0 & \Omega^{-1}_{33} \end{bmatrix}, \nu \right\} \text{ i.i.d.} \tag{D.9}$$

The specification of $\Psi^\dagger_1$ ensures that $R = 1$. Variables μ^*_t and $\mu^\dagger_t$ are initialized by using 3×1 vectors of zeros. The co-integration vector for $y_{2,t}$ and $y_{3,t}$ is given by $(-\beta_2, 1)'$.

The ML parameter estimates for Gaussian-QVARMA(2,1,1) are presented in Table 3. Those estimates indicate that both short-run and long-run dynamic effects are significant, as several elements of ϕ^*_1, ϕ^*_2, Ψ^*_1, and $\Psi^\dagger_1$ are significant. In Table 3, the AIC, BIC, and HQC model selection criteria are also presented, which indicate that the likelihood-based statistical performances of the t-QVARMA(2,1,1) specifications are superior to the statistical performance of Gaussian-QVARMA(2,1,1).

The IRF estimates for Gaussian-QVARMA(2,1,1) are presented in Figures D1 to D3. The estimates of $\text{IRF}^*_{j,t} + \text{IRF}^v_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure D1. The estimates of $\text{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure D2. The estimates of $\text{IRF}^*_{j,t} + \text{IRF}^v_{j,t} + \text{IRF}^\dagger_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure D3.

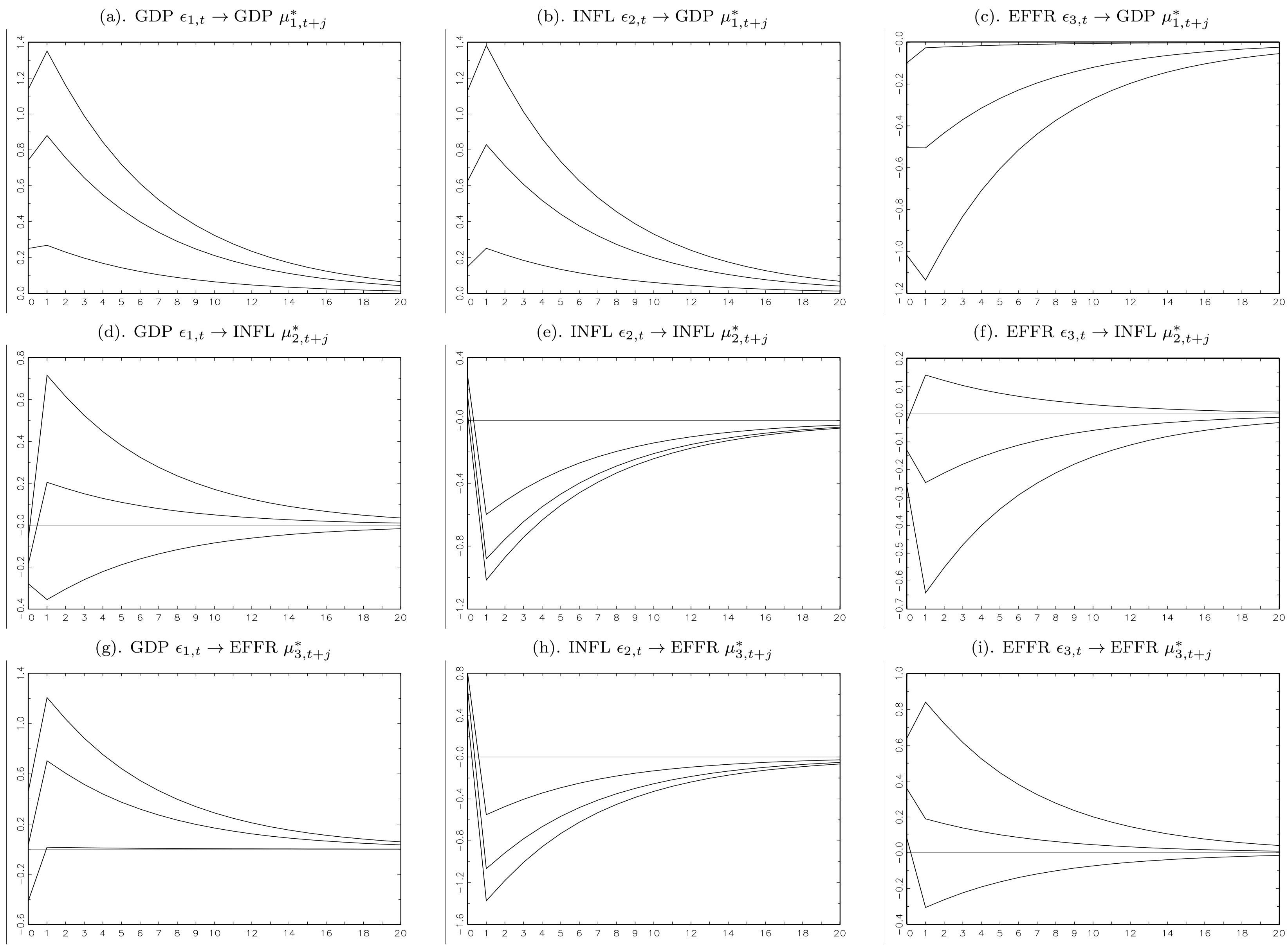

Figure D1: Short-run $\text{IRF}^*_{j,t}+\text{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for Gaussian-QVARMA(2,1,1). *Notes*: The confidence interval is for 9,960 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

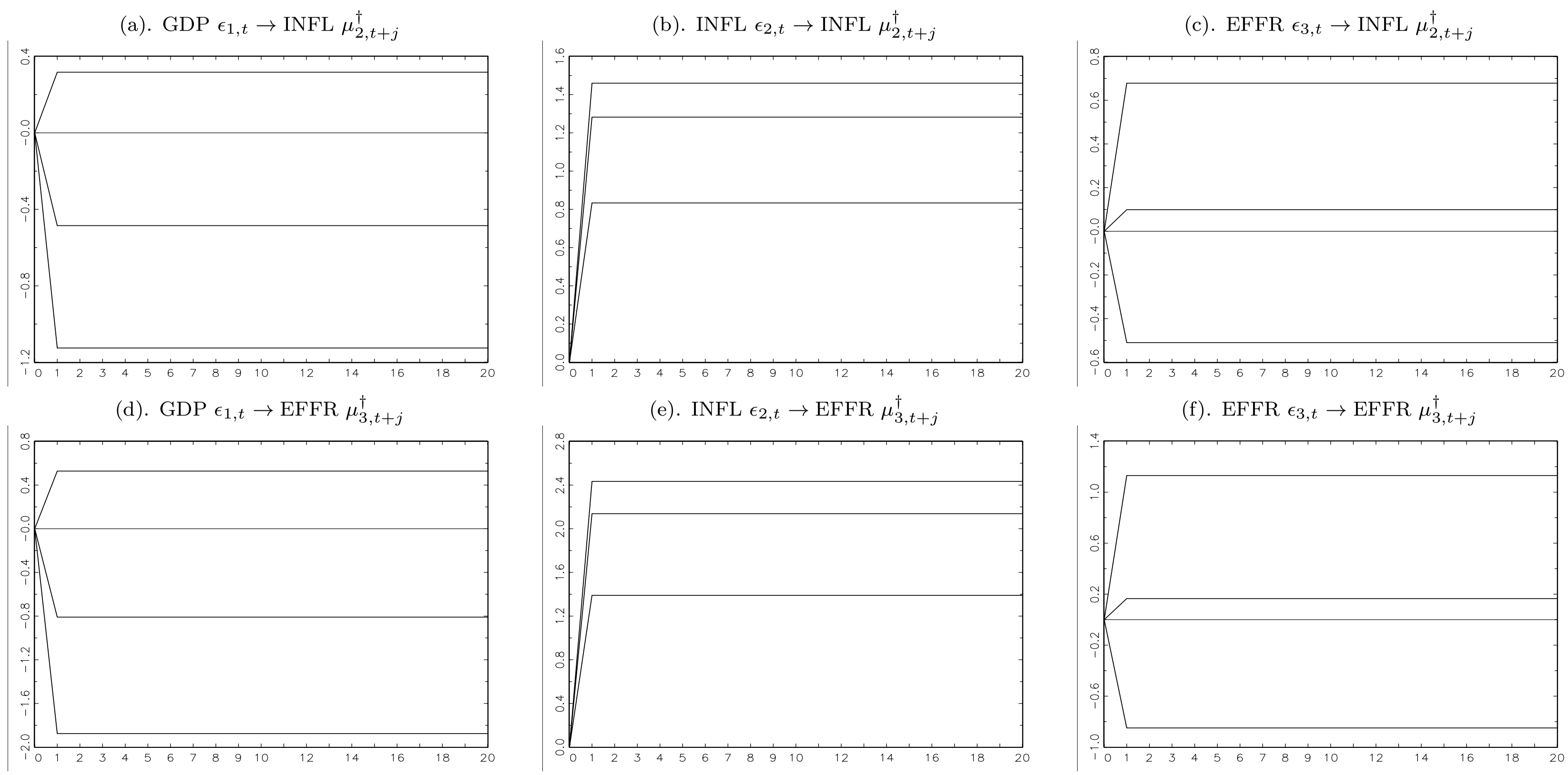

Figure D2: Long-run $\text{IRF}^{\dagger}_{j,t}$ (10%, 50%, and 90% percentiles) for Gaussian-QVARMA(2,1,1). *Notes*: The confidence interval is for 9,960 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

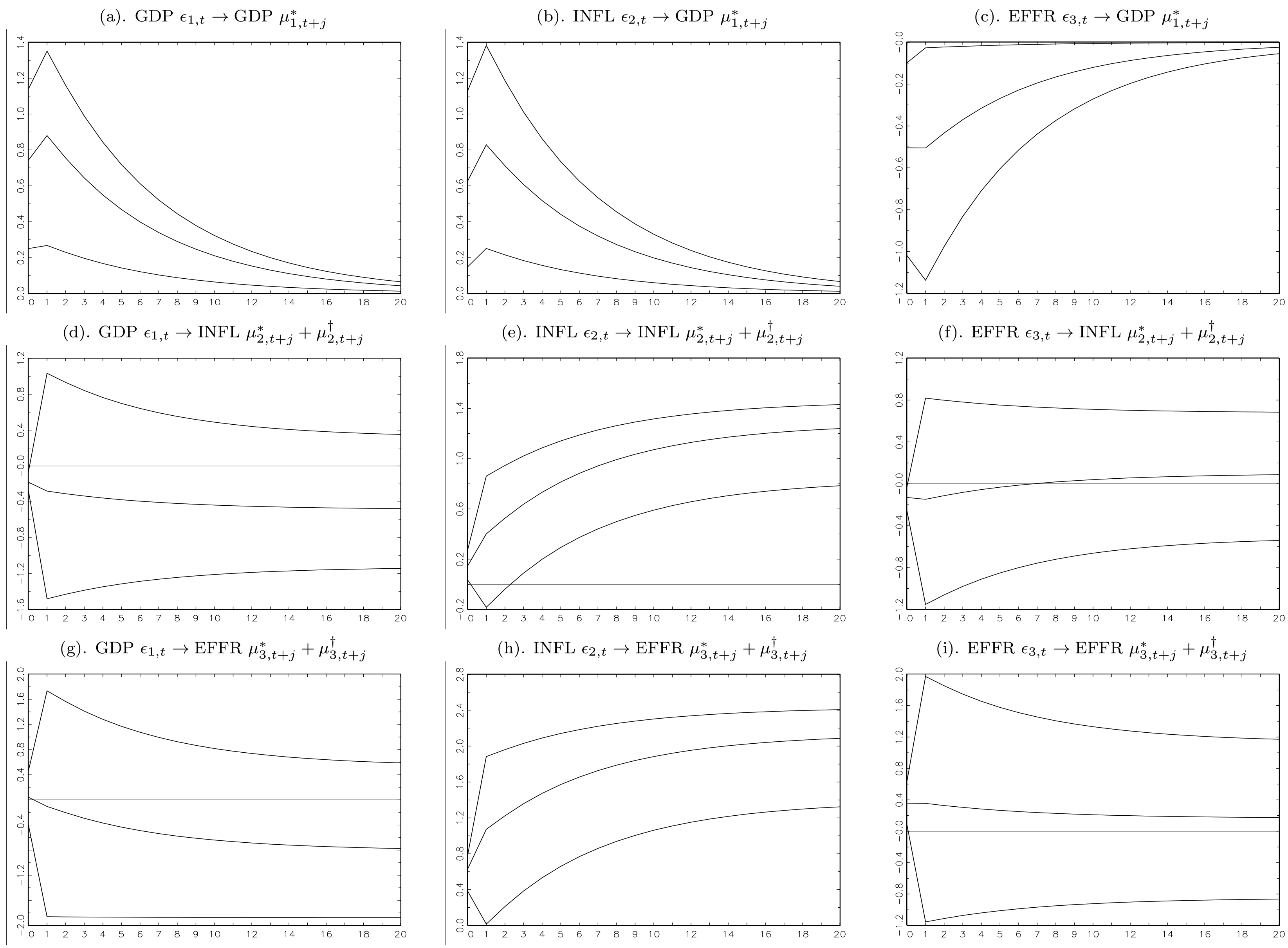

Figure D3: Total $\mathrm{IRF}^*_{j,t} + \mathrm{IRF}^\dagger_{j,t} + \mathrm{IRF}^v_{j,t}$ (10%, 50%, and 90% percentiles) for Gaussian-QVARMA(2,1,1). *Notes*: The confidence interval is for 9,960 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

Supplementary Material E: Classical alternatives to t-QVARMA

The ML estimates for Gaussian-VAR(2) are presented in Table 3 (the VAR lag-order is chosen by using the AIC, BIC, and HQC metrics). For Gaussian-VAR(2), the estimates of $\text{IRF}_{j,t}$ for $j = 0, \ldots, 20$ are presented in Figure E1.

For Gaussian-VAR(2)-VECM, the following specification is estimated, where $y_{1,t}$ is $I(0)$, and $y_{2,t}$ and $y_{3,t}$ are $I(1)$ and co-integrated:

$$\begin{bmatrix} y_{1,t} - c_1 \\ \Delta y_{2,t} \\ \Delta y_{3,t} \end{bmatrix} = \alpha\beta' \begin{bmatrix} y_{1,t-1} - c_1 \\ y_{2,t-1} \\ y_{3,t-1} \end{bmatrix} + \begin{bmatrix} \Gamma_{1,11} & \Gamma_{1,12} & \Gamma_{1,13} \\ \Gamma_{1,21} & \Gamma_{1,22} & \Gamma_{1,23} \\ \Gamma_{1,31} & \Gamma_{1,32} & \Gamma_{1,33} \end{bmatrix} \begin{bmatrix} y_{1,t-1} - c_1 \\ \Delta y_{2,t-1} \\ \Delta y_{3,t-1} \end{bmatrix} + \begin{bmatrix} v_{1,t-1} \\ v_{2,t-1} \\ v_{3,t-1} \end{bmatrix} \tag{E.1}$$

where $\alpha = (0, \alpha_2, \alpha_3)'$ and $\beta = (0, \beta_2, 1)'$.

The ML estimates for Gaussian-VAR(2)-VECM are presented in Table 3. With respect to the IRFs of Gaussian-VAR(2)-VECM, the estimation results indicate that the IRF confidence intervals are significantly wider for Gaussian-VAR(2)-VECM than for Gaussian-VAR(2). Therefore, the Gaussian-VAR(2)-VECM IRF results are not reported in this paper.

We also studied the use of the multivariate t-distribution in the VAR framework, because t-VECM is more robust to outliers than Gaussian-VECM (Lucas 1997). The estimation results for t-VECM indicate that the IRF confidence intervals are significantly wider for t-VECM than for Gaussian-VECM. As a consequence, t-VAR and t-VECM results are not reported in this paper.

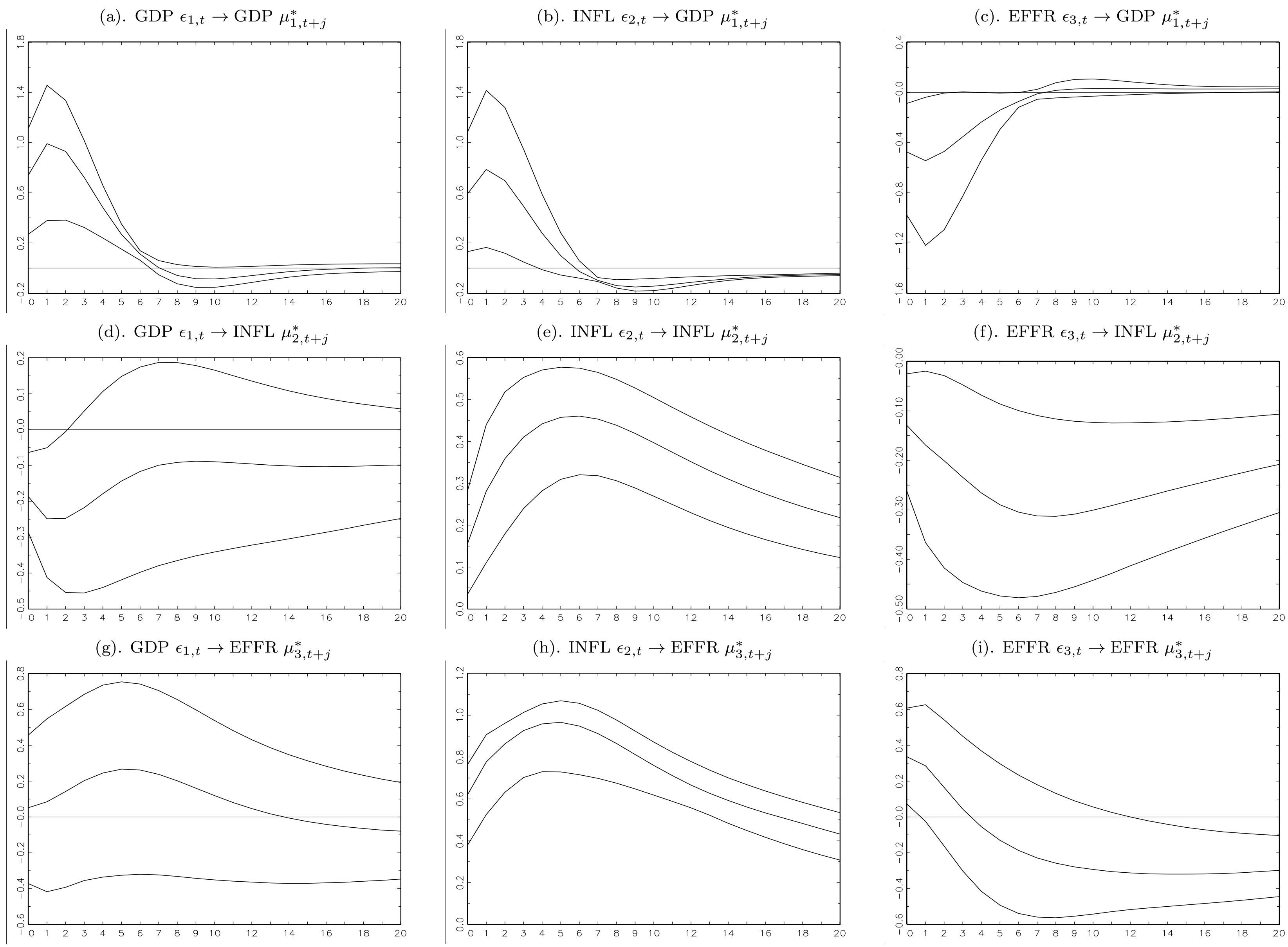

Figure E1: Total $\text{IRF}_{j,t}$ (10%, 50%, and 90% percentiles) for Gaussian-VAR(2). *Notes*: The confidence interval is for 9,410 out of the 1 million simulations, for which the restrictions of Table 1 are satisfied. GDP, INFL, and EFFR are US real GDP growth, US inflation rate, and effective federal funds rate, respectively.

References

Blasques, Francisco, Siem J. Koopman, and Andre Lucas (2017) Maximum likelihood estimation for score-driven models. TI 2014-029/III Tinbergen Institute Discussion Paper. https://papers.tinbergen.nl/14029.pdf.

Brandt, Andreas (1986) The stochastic equation $Y_{n+1} = A_n Y_n + B_n$ with stationary coefficients. *Advances in Applied Probability* 18(1), 211–220.

Creal, Drew, Siem J. Koopman, and Andre Lucas (2013) Generalized autoregressive score models with applications. *Journal of Applied Econometrics* 28(5), 777–795.

Elton, John H. (1990) A multiplicative ergodic theorem for Lipschitz maps. *Stochastic Processes and their Applications* 34(1), 39–47.

Harvey, Andrew C. (2013) *Dynamic Models for Volatility and Heavy Tails*. Cambridge: Cambridge University Press.

Lucas, Andre (1997) Cointegration testing using pseudolikelihood ratio tests. *Econometric Theory* 13(2), 149–169.

Lütkepohl, Helmut (2005) *New Introduction to Multivariate Time Series Analysis*. Berlin Heidelberg: Springer-Verlag.

Straumann, Daniel, Thomas Mikosch (2006) Quasi-maximum-likelihood estimation in conditionally heteroscedastic time series: a stochastic recurrence equations approach. *The Annals of Statistics* 34(1), 2449–2495.

White, H. (2001) *Asymptotic Theory for Econometricians*, revised edition. San Diego: Academic Press.

Wooldridge, Jeffrey M. (1994) Estimation and inference for dependent processes. In: Engle, Robert F., Daniel L. McFadden (editors), *Handbook of Econometrics*, Volume 4, 2639–2738. Amsterdam: North-Holland.