VAR for VaR and CoVaR

Julien Hambuckers[1], Alain Hecq[2], Stefan Straetmans[2], and Li Sun[*1]

[1]University of Liège - HEC Liège
[2]Maastricht University - School of Business and Economics

March 5, 2022

Abstract

This paper generalizes multivariate multi-quantile CAViaR models (MVMQ-CAViaR, see White et al., 2015) by incorporating CoVaR specification (see Adrian and Brunnermeier, 2011) into the model specification. The proposed model presents a vector-autoregression (VAR) of financial institutions' value-at-risk (VaR) as well as their CoVaR. This model generalization is able to capture contemporaneous tail dependence of financial institutions and market indexes so that we can interpret the systemic risks of the institutions more timely. We provide consistency and asymptotic normality of the general model estimator as well as some relevant inference tests. For tracing the transmission of a single shock to a financial institution in the financial system, we also construct quantile impulse response functions (QIRF) accordingly using the local projection idea (Jordà, 2005) and score vectors. Applications to real data shows strong evidence of contemporaneous effects of big banks on the market index S&P500, and supports this methodology.

JEL Codes: C22
Keywords: MVMQ CAViaRs, SVAR, VaR, specification testing, systemic risk, CoVaR.

1 Introduction

Value-at-Risk (VaR) is a standard risk measure for market risk management with defining risk as loss on a fixed asset over a fixed time horizon. It is widely employed in the financial industry for both internal control and regulatory reporting. Among many popular approaches for VaR estimation, quantile regressions stand out for the advantages in semi-parametric specification and numerical efficiency. The quantile regression family working for this measure has been extended from static quantile regression models (QR, see Koenker and Bassett Jr, 1978) to quantile autoregression models (QAR, see Koenker and Xiao, 2006), to conditional autoregressive value-at-risk models (CAViaR, see Engle and Manganelli, 2004), to multivariate multi-quantile CAViaR models (MVMQ-CAViaR, see White et al., 2015). This evolution can be hinted by the specification testing summarized by Chernozhukov and Umantsev (2001) with regard to checking the conditionality and the functional form validity of quantile regression models, which enlightens us to see that each model generation in the quantile regression evolution can be viewed as a consequence from the inability of its predecessors in the conditional information sufficiency or in the functional form validity.

MVMQ-CAViaR is capable of measuring the individual tail risk as well as the tail dependence of financial institutions by modelling the VaR of their stock returns in a vector-autoregressive way. We notice that only predetermined (i.e., lagged and exogenous) information is accounted in the MVMQ-CAViaR specification (White et al., 2015) which cannot measure the contemporaneous tail dependence among financial assets or cover the CoVaR specification (Adrian and Brunnermeier, 2011) for measuring systemic risk of financial institutions. As we always see clear comovement pattern between market portfolios and some big financial institutions, we question if MVMQ-CAViaRs are sufficient to explain this pattern.

[*]Corresponding author, Email: l.sun@uliege.be.

In this paper we propose to generalize the vector-autoregressive VaR in the generic MVMQ-CAViaR specification by incorporating the CoVaR specification of a financial market portfolio. In this paper, we would like to see if we can find significant contemporaneous tail dependence between the financial market portfolio and some financial institutions so as to measure their systemic risk. We will also provide the estimation consistency and asymptotic normality proofs of this generalized model along with some testing methods to infer the significance of the contemporaneous tail dependence between a market portfolio/index and some big financial institutions. This model generalization also can allow us to study the links across the whole financial market network and to trace the transmission of a single shock to a financial institution in the financial system by using quantile impulse response functions (QIRF) which we will construct in this paper accordingly with the use of the local projection idea (Jordà, 2005) and expansion of estimated terms.

The remainder of this paper is organized as follows. In Section 2, we introduce the generic MVMQ-CAViaR model specification first and propose to generalize it to the vector autoregressive model of VaR and CoVaR along with the proofs on the estimation consistency and asymptotic normality of this generalized model. We also call this generalized model as *systemic MVMQ-CAViaR model*. It follows that in Section 2.3 some inference tests are given in order to infer the significance of contemporaneous terms in the CoVaR specification. In Section 3, we illustrate on applying CoVaR returned by our generalized model to measure the systemic risk of financial institutions. In Section 4, we construct quantile impulse response functions in correspondence to our model and apply the local projection with use of expansion of estimated terms for QIRF estimation. Some results of Monte Carlo simulations regarding systemic MVMQ-CAViaR models are presented in Section 5. Section 6 presents an empirical application to real data. Section 7 concludes this paper.

2 Model Generalization to Systemic MVMQ-CAViaR

In Section 2.1, we first introduce the generic MVMQ-CAViaR model specification which is also refered to as VAR for VaR (see White et al., 2015). Being motivated to measure the systemic risk of some financial institutions and to measure the contemporaneous tail dependence among financial assets, we propose to generalize MVMQ-CAViaR models to the vector autoregressive model of VaR and CoVaR in Section 2.2. In Section 2.3 some inference tests are proposed in order to infer the significance of the contemporaneous tail dependence.

2.1 MVMQ-CAViaRs (VAR for VaR)

White et al. (2015) proposed the framework of multivariate multi-quantile CAViaR (MVMQ-CAViaR) models and established the theoretical validity in its application by proving its estimation consistency and asymptotic normality under some regularity conditions. In essense, this framework is based on autoregressing the VaR of multiple random variables onto their lags so as to measure their tail dependence, which can be regarded as an extension to CAViaR models which are autoregressions of univariate VaR. The generic MVMQ-CAViAR specification given by White et al. (2015) is shown below.

Suppose $\boldsymbol{Y} := [Y_1, Y_2, \ldots, Y_n]'$ is a vector of n random variables of interest, with its multivariate time series $\{\boldsymbol{Y}_t\}_{t=1}^T$. We consider a vector of explanatory exogenous variables denoted by $\boldsymbol{X}$ whose first element is one, with time series $\{\boldsymbol{X}_t\}_{t=1}^T$. We consider p quantile indexes denoted by $\theta_{i,1}, \ldots, \theta_{i,p}$ and $0 < \theta_{i,1} < \ldots, < \theta_{i,p} < 1$ for each $Y_i, i = 1, \ldots, n$. Define the information set $\mathcal{F}_t$ until time t to be the σ-algebra generated by $Z^{(t)} := \{\boldsymbol{X}_{t+1}, \boldsymbol{X}_t, \boldsymbol{Y}_t, \boldsymbol{X}_{t-1}, \boldsymbol{Y}_{t-1}, \ldots\}$, i.e., $\mathcal{F}_t := \sigma(Z^{(t)}), t = 1, \ldots, T$. Denote the cumulative distribution function of Y_{it} conditional on $\mathcal{F}_{t-1}$ by $F_{it}(\cdot)$[1]. The θ_{ij}-th quantile of Y_{it} conditional on $\mathcal{F}_{t-1}$, denoted by $q_{i,j,t} := \inf\{y : F_{it}(y) \geq \theta_{ij}\}$, can be specified by a generic MVMQ-CAViaR as follows:

$$q_{i,j,t} = \Psi_t' \boldsymbol{\beta}_{ij} + \sum_{\tau=1}^{m} \boldsymbol{q}_{t-\tau}' \gamma_{i,j,\tau}, \qquad i = 1, \ldots, n; \ j = 1, \ldots, p, \tag{1}$$

[1]White et al. (2015) specified the dependence of $F_{it}(\cdot)$ on each $\omega \in \Omega$ by denoting $F_{it}(\cdot)$ more specifically as $F_{it}(\omega, \cdot)$. In this paper, we do not specify explicitly the role of each $\omega \in \Omega$ as $\omega \in \Omega$ cannot be extracted explicitly to formulate its role and its influence can be exerted partially or fully through the conditional set $\mathcal{F}_{t-1}$ each time.

where $\{\Psi_t\}$ is a sequence of $k \times 1$ variables with Ψ_t predetermined to $\boldsymbol{Y}_t$ and being $\mathcal{F}_{t-1}$-measurable, $\boldsymbol{\beta}_{ij} := (\beta_{i,j,1}, \ldots, \beta_{i,j,k})'$ is a $k \times 1$ real vector, and $\boldsymbol{\gamma}_{i,j,\tau} := (\boldsymbol{\gamma}'_{i,j,\tau,1}, \ldots, \boldsymbol{\gamma}'_{i,j,\tau,n})$ with each $\boldsymbol{\gamma}_{i,j,\tau,k}$ being a $p \times 1$ real vector. Let $\boldsymbol{\gamma}_{ij} := (\boldsymbol{\gamma}'_{i,j,1}, \ldots, \boldsymbol{\gamma}'_{i,j,m}), \boldsymbol{\alpha}_{ij} := (\boldsymbol{\beta}'_{ij}, \boldsymbol{\gamma}'_{ij})$ and $\boldsymbol{\alpha} := (\boldsymbol{\alpha}'_{11}, \ldots, \boldsymbol{\alpha}'_{1p}, \ldots, \boldsymbol{\alpha}'_{n1}, \ldots, \boldsymbol{\alpha}'_{np})$, where $\boldsymbol{\alpha} \in \mathbb{A}$, a compact set of $\mathbb{R}^{l_r}$ with $l_r := np(k+m)$.

MVMQ-CAViaR is capable of measuring the individual tail risk as well as the tail dependence of financial institutions. However, due to the conditional limitation in MVMQ-CAViaR, MVMQ-CAViaR models cannot measure the contemporaneous tail dependence among financial assets or cover the CoVaR specification (Adrian and Brunnermeier, 2011) for measuring systemic risk of financial institutions. As we always see clear contemporaneous comovement patterns between market portfolios and some big financial institutions, we question if MVMQ-CAViaRs are sufficient to explain this pattern. This question makes sense as it is often believed that information is rapidly reflected into stock prices. Neglecting contemporaneous return spillovers probably leads to underestimation of systemic risk contributions or exposures. Appealed to the systemic feature of financial systems, in the following we are going to generalize MVMQ-CAViaR models by incorporating the CoVaR specification on a financial market portfolio return.

2.2 Systemic MVMQ-CAViaR (VAR for VaR and CoVaR)

The financial market can react to news rapidly and extensively. One type of news deemed to be influential on the market is regarding 'too big to fail' financial institutions due to their systemic risk. Systemic risk in the financial market is defined as the risk that an event at the company level triggers severe instability or collapse of an entire industry or even the economy. The 'too big to fail' financial institutions are highly interconnected with the financial market both directly and indirectly. The direct links can happen through contractual commits and counterparty credit risk. The indirect links are, for instance, price effects and liquidity spirals. Such interconnection the 'too big to fail' financial institutions possess in the financial market is also accompanied with their systemic risk to the whole market. With the objective to measure the systemic risk and contemporaneous tail dependence of financial institutions, we would like to generalize MVMQ-CAViaR models by incorporating the CoVaR specification (see Adrian and Brunnermeier, 2011) into the autoregressive model (1) as follows which presents to us a vector-autoregressive model of financial institutions' VaR as well as their CoVaR.

A generic *systemic MVMQ-CAViAR* model specification given below is based on the set-up of the generic MVMQ-CAViaR model specification (1) with generalization to measure the contemporaneous tail-dependence of response variables. It is worth mentioning that the contemporaneous dependence direction is predetermined by our expertise and belief.[2]

$$\begin{cases} q_{i,j,t} = \Phi'_t \boldsymbol{\beta}_{ij} + \sum_{\tau=1}^{m} \boldsymbol{q}'_{t-\tau} \gamma_{i,j,\tau} & i = 1, j = 1, \ldots, p, \\ q_{i,j,t} = \Phi'_t \boldsymbol{\beta}_{ij} + \sum_{\tau=1}^{m} \boldsymbol{q}'_{t-\tau} \gamma_{i,j,\tau} + g_{i,j}\left(\boldsymbol{s}_{i,j}, \boldsymbol{u}_{i,t}\right), & i = 2, \ldots, n, \ j = 1, \ldots, p, \end{cases} \tag{2}$$

where

$$\boldsymbol{u}_{i,t} := \begin{cases} \left[y_{1,t} - q_{1,1,t}, \ldots, y_{1,t} - q_{1,p,t}\right]', & \text{when} \quad i = 2; \\ \left[y_{1,t} - q_{1,1,t}, \ldots, y_{1,t} - q_{1,p,t}, \ldots, y_{i-1,t} - q_{i-1,1,t}, \ldots, y_{i-1,t} - q_{i-1,p,t}\right]', & \text{when} \quad i = 3, \ldots, n. \end{cases}$$

And parameters $\boldsymbol{s}_{i,j} \in \mathbb{S}_{ij}$ which are compact sets of $\mathbb{R}^{d_{ij}}$ for $i = 2, \ldots, n$ and $j = 1, \ldots, p$ with d_{ij} being nonnegative integers. The mapping $g_{i,j} : \mathbb{S}_{ij} \times \mathbb{R}^{(i-1)p} \to \mathbb{R}$, accounts for the contemporaneous effect on $q_{i,j,t}$ due to the realization $\boldsymbol{u}_{i,t}$. Intuitively speaking, $g_{i,j}$ intends to capture the influence of contemporaneous news due to $y_{1,t}, \ldots, y_{i-1,t}$ onto the conditional quantiles of $y_{i,t}$ so as to measure their contemporaneous tail dependence and the systemic risk of $y_{1,t}, \ldots, y_{i-1,t}$ if Y_i is a market index return variable. Some functional form examples of $g_{i,j}$ are given below. For instance,

$$g_{i,j}\left(\boldsymbol{s}_{i,j}, \boldsymbol{u}_{i,t}\right) = \boldsymbol{s}'_{i,j}\, \boldsymbol{u}_{i,t} = \sum_{k_i, k_j = 1}^{k_i = i-1, k_j = p} s_{k_i,k_j}(y_{k_i,t} - q_{k_i,k_j,t})$$

[2] Identifying the contemporaneous dependence direction is beyond the scope of this paper.

is intended to explain the comovement of $q_{i,j,t}$ with $y_{1,t}, \ldots, y_{i-1,t}$ and $\boldsymbol{u}_{i,t}$, where $\boldsymbol{s}_{i,j} = [s_{1,1}, \ldots, s_{i-1,p}] \in \mathbb{R}^{(i-1)p}$. Another example is that

$$g_{i,j}\left(\boldsymbol{s}_{i,j}, \boldsymbol{u}_{i,t}\right) = \sum_{k_i,k_j=1}^{k_i=i-1,k_j=p} s_{k_i,k_j,1} \left\{ \left[1+\exp(s_{k_i,k_j,2}[y_{k_i,t} - q_{k_i,k_j,t}])\right]^{-1} - \theta_{k_i,k_j} \right\}$$

is intended to explain the θ_{k_i,k_j}-th quantile violation of $y_{k_i,t}$ shifting $q_{i,j,t}$ so as to affect y_{it}.

For the ease of notations, we stack the parameters of (2) into $\boldsymbol{\alpha} := \left[\boldsymbol{\alpha}'_{11}, \ldots, \boldsymbol{\alpha}'_{1p}, \ldots, \boldsymbol{\alpha}'_{n1}, \ldots, \boldsymbol{\alpha}'_{np}\right]$ where $\boldsymbol{\alpha} \in \Theta := \mathbb{A} \times \mathbb{S}_{21} \times \ldots \times \mathbb{S}_{np}$, a compact set of $\mathbb{R}^{l_s}$ with $l_s := np(k+m) + \sum_{j=1}^{p}\sum_{i=2}^{n} d_{ij}$, $\boldsymbol{\gamma}_{ij} := [\boldsymbol{\gamma}'_{i,j,1}, \ldots, \boldsymbol{\gamma}'_{i,j,m}]', \boldsymbol{\alpha}_{ij} := [\boldsymbol{\beta}'_{ij}, \boldsymbol{\gamma}'_{ij}, \boldsymbol{s}'_{ij}]'$. To estimate true parameters $\boldsymbol{\alpha}^o$, we apply the quasi-maximum likelihood method by optimizing the objective function $\bar{S}_T(\boldsymbol{\alpha})$ and obtain the quasi-maximum likelihood estimator (QMLE) $\widehat{\boldsymbol{\alpha}}$ as shown below.

$$\begin{aligned} \widehat{\boldsymbol{\alpha}} &= \underset{\boldsymbol{\alpha}\in\Theta}{\arg\min}\, \bar{S}_T(\boldsymbol{\alpha}), \\ \bar{S}_T(\boldsymbol{\alpha}) &:= T^{-1}\sum_{t=1}^{T}\left\{\sum_{i=1}^{n}\sum_{j=1}^{p} \rho_{\theta_{ij}}(y_{it} - q_{i,j,t}(\boldsymbol{\alpha}))\right\}, \end{aligned} \tag{3}$$

where $\rho_\theta(\epsilon) = \epsilon(\theta - \mathbf{I}_{[\epsilon<0]})$ is known as check function in quantile regressions. Denote $\psi_\theta := (\theta - \mathbf{I}_{[\epsilon<0]})$.

To prove the consistency and asymptotic normality of the generic systemic MVMQ-CAviaR (2), we impose some assumptions below on the contemporaneous terms $g_{i,j}\left(\boldsymbol{s}_{i,j}, \boldsymbol{u}_{i,t}\right)$ in addition to all the assumptions on MVMQ-CAViaRs given by White et al. (2015).

Assumption 1 (contemporaneous terms)

G1: For each $\boldsymbol{s}_{i,j} \in \mathbb{S}_{ij}$, a compact sets of $\mathbb{R}^{d_{ij}}$ $(i = 2, \ldots, n,\ j = 1, \ldots, p)$, $g_{i,j}\left(\boldsymbol{s}_{i,j}, \cdot\right)$ is measurable with respect to an updated information set $\mathcal{F}_{t,i-1}$ which is the σ-algebra generated by $Z^{(t,i-1)} := \{\boldsymbol{Y}_t[1:(i-1)], \boldsymbol{X}_{t+1}, \boldsymbol{X}_t, \boldsymbol{Y}_t, \boldsymbol{X}_{t-1}, \boldsymbol{Y}_{t-1}, \ldots\}$, i.e., $\mathcal{F}_{t,i-1} := \sigma(Z^{(t,i-1)}), t = 1, \ldots, T$. [3]

G2: For any $\omega \in \mathcal{F}_{t,i-1}$, $g_{i,j}\left(\cdot, \omega\right)$ is continuous on $\mathbb{S}_{ij}$, $i = 2, \ldots, n$, $j = 1, \ldots, p$.

G3: For any $1 \le t \le T$, $\mathbb{E}\left[\sup_{\boldsymbol{s}_{ij}\in\mathbb{S}_{ij}} g_{i,j}\left(\boldsymbol{s}_{i,j}, \boldsymbol{u}_{i,j,t}\right)\right] < \infty$, $i = 2, \ldots, n,\ j = 1, \ldots, p$.

G4: For any $\omega \in \mathcal{F}_{t,i-1}$, $g_{i,j}\left(\cdot, \omega\right)$ is differentiable in $\mathbb{S}_{ij}$, $i = 2, \ldots, n;\ j = 1, \ldots, p$.

Theorem 2 (Consistency) *Suppose that the assumptions of Theorem 1 of White et al. (2015) and Assumptions G1-3 hold. Then we have*

$$\widehat{\boldsymbol{\alpha}} \overset{a.s.}{\to} \boldsymbol{\alpha}^o, \tag{4}$$

where $\widehat{\boldsymbol{\alpha}}$ is the quasi-likelihood estimator (QML) obtained in (3) *for estimating the true parameter α^o in the underlying systemic MVMQ-CAViaR process $\{\boldsymbol{Y}_t\}$.*

Proof. See Appendix B. ■

Theorem 3 (Asymptotic Normality) *Suppose that the assumptions of Theorem 2 of White et al. (2015) and Assumptions G1-4 hold. Then the asymptotic distribution of the QML estimator $\widehat{\boldsymbol{\alpha}}$ obtained from* (3) *is as follows:*

$$\sqrt{T}\left(\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^o\right) \overset{\mathcal{D}}{\sim} N\left(\mathbf{0}, Q^{-1} V Q^{-1}\right), \tag{5}$$

[3] $\boldsymbol{Y}_t[1:(i-1)]$ are the vector containing the first $(i-1)$ elements of $\boldsymbol{Y}_t$, $i = 1, 2, \ldots, n$. And when $i = 1$, $\boldsymbol{Y}_t[1:(i-1)] = 0$.

where

$$
\begin{aligned}
Q &:= \sum_{i=1}^{n}\sum_{j=1}^{p} \mathbb{E}\left[f_{i,j,t}(0)\nabla q_{i,j,t}(\boldsymbol{\alpha}^o)\nabla' q_{i,j,t}(\boldsymbol{\alpha}^o)\right], \\
V &:= \mathbb{E}\left[\boldsymbol{\eta}_t^o\,(\boldsymbol{\eta}_t^o)'\right] \\
\boldsymbol{\eta}_t^o &:= \sum_{i=1}^{n}\sum_{j=1}^{p} \mathbb{E}\left[\nabla q_{i,j,t}(\boldsymbol{\alpha}^o)\psi(\epsilon_{i,j,t})\right], \\
\epsilon_{i,j,t} &:= Y_{it} - q_{i,j,t}(\boldsymbol{\alpha}^o),
\end{aligned}
\tag{6}
$$

and $f_{i,j,t}(\cdot)$ is the continuous density of $\epsilon_{i,j,t}$ conditional on $F_{t,i-1}$, and α^o is the true parameter in the underlying systemic MVMQ-CAViaR process $\{\mathbf{Y}_t\}$.

Proof. See Appendix B. ■

Theorem 4 *Suppose that the assumptions of Theorem 3 of White et al. (2015) and Assumptions G1-4 hold. To estimate V in Theorem 3, $\widehat{V}_T$ is obtained by plugging the QMLE $\widehat{\alpha}$ into* (6) *as follows:*

$$
\begin{aligned}
\widehat{V}_T &= \frac{1}{T}\sum_{t=1}^{T}\widehat{\boldsymbol{\eta}}_t\widehat{\boldsymbol{\eta}}_t' \\
\widehat{\boldsymbol{\eta}}_t &= \sum_{i=1}^{n}\sum_{j=1}^{p} \nabla q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\psi(\widehat{\epsilon_{i,j,t}}), \\
\widehat{\epsilon}_{i,j,t} &= Y_{it} - q_{i,j,t}(\widehat{\boldsymbol{\alpha}}).
\end{aligned}
\tag{7}
$$

Then we have

$$
\widehat{V}_T \xrightarrow{p} V. \tag{8}
$$

Proof. See Appendix B. ■

Denote the conditional probability density of $y_{i,t}$ at its conditional j-th quantile $q_{i,j,t}$ as $f_{i,j,t}(0)$, $i = 1,\ldots,n, j = 1,\ldots,p, t = 1,\ldots,T$. We apply the adaptive random bandwidth method (Hecq and Sun, 2020) to estimate $f_{i,j,t}(0)$ $(i = 1,\ldots,n, j = 1,\ldots,p, t = 1,\ldots,T)$ as follows.

Theorem 5 (Adaptive Random Bandwidth Method)
Given the conditions and the asymptotic normality result in Theorem 3 and assuming the condition that

$$
\sqrt{T}\,(\boldsymbol{\alpha}_z - \widehat{\boldsymbol{\alpha}}) \overset{\mathcal{D}}{\sim} N(\mathbf{0}, \boldsymbol{I}_{l_s\times l_s}), \qquad z = 1,\ldots,N,
$$

with $l_s := np(k+m) + \sum_{j=1}^{p}\sum_{i=2}^{n} d_{ij}$ and the exclusion of $\boldsymbol{\alpha}_z$ such that

$$
\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_z - \widehat{\boldsymbol{\alpha}}) = 0,
$$

we can get the following estimator of $f_{i,j,t}(0)$:

$$
\begin{aligned}
\widehat{f}_{i,j,t}(0) &= \frac{1}{N}\sum_{i=1}^{N}\frac{\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})+\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i-\widehat{\boldsymbol{\alpha}})\}} - \mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \\
&\xrightarrow{p} f_{i,j,t}(0),
\end{aligned}
\tag{9}
$$

as $N \to \infty$ for $i = 1,\ldots,n, j = 1,\ldots,p, t = 1,\ldots,T$, where $\mathbf{I}_{\{S\}}$ is the indicator function on a set S or a logical statement S, and $\mathbf{I}_{\{S\}} = 1$ if S is non-empty or true, otherwise $\mathbf{I}_{\{S\}} = 0$.

Theorem 6 *Suppose that the assumptions of Theorem 3 of White et al. (2015) and Assumptions G1-4 hold. To estimate Q in Theorem 3, $\widehat{Q}_T$ is obtained by plugging the QMLE $\widehat{\alpha}$ into* (6) *as follows:*

$$\widehat{Q}_T := \frac{1}{T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p}\widehat{f}_{i,j,t}(0)\nabla q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}}), \tag{10}$$

where $\widehat{f}_{i,j,t}(0)$ is obtained by the adaptive random bandwidth method (Hecq and Sun, 2020) in Theorem 5. Then we have

$$\widehat{Q}_T \xrightarrow{p} Q. \tag{11}$$

Proof. See Appendix B. ■

Here we would like to show a very basic data generating process (DGP) example on the proposed systemic MVMQ-CAViaR (2). We want to use this example to compare MVMQ-CAViaR with the systemic one in their conditions on ensuring the estimation consistency as well as to see their linkage. Suppose a bivariate time series $\{\boldsymbol{Y}_t := (y_{1,t}, y_{2,t})\}_{t=1}^{T}$ follows a systemic bivariate CAViaR DGP specified as follows:

$$\begin{bmatrix} 1 & 0 \\ s^o_{2,1} & 1 \end{bmatrix}\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \beta^o_{10} \\ \beta^o_{20} \end{bmatrix} + \begin{bmatrix} \beta^o_{11} & \beta^o_{12} \\ \beta^o_{21} & \beta^o_{22} \end{bmatrix}\begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \gamma^o_{11} & \gamma^o_{12} \\ \gamma^o_{21} & \gamma^o_{22} \end{bmatrix}\begin{bmatrix} q_{1,t-1} \\ q_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}, \tag{12}$$

where $\{\epsilon_{1,t}\}$ and $\{\epsilon_{2,t}\}$ are independently distributed to each other as well as within their own processes with $[P\{\epsilon_{1,t} \le 0|\mathcal{F}_{t,0}\}, P\{\epsilon_{2,t} \le 0|\mathcal{F}_{t,1}\}] = [\theta, \theta]$. $q_{1,t}$ and $q_{2,t}$ are the θ-th quantiles of $y_{1,t}$ and $y_{2,t}$ conditional on $\mathcal{F}_{t,0}$ and $\mathcal{F}_{t,1}$ respectively, i.e., $[P\{y_{1,t} \le q_{1,t}|\mathcal{F}_{t,0}\}, P\{y_{2,t} \le q_{2,t}|\mathcal{F}_{t,1}\}] = [\theta, \theta]$. Under the assumptions of Theorem 2, we can get its QML estimator $\widehat{\boldsymbol{\alpha}}^{(s)}$ consistent to its true parameters denoted as $\boldsymbol{\alpha}^{o\,(s)}$.[4]

If we ignore the contemporaneous effect and just regress $\{\boldsymbol{Y}_t := (y_{1,t}, y_{2,t})\}_{t=1}^{T}$ onto the bivariate CAViaR without contemporaneous terms, a correct quantile specification condition has to be imposed further as follows in order to ensure the estimation consistency in this regression.

$$\begin{bmatrix} P\{\epsilon_{1,t} \le 0|\mathcal{F}_{t,0}\} \\ P\{\epsilon_{2,t} - s_{2,1}\cdot\epsilon_{1,t} \le 0|\mathcal{F}_{t,0}\} \end{bmatrix} = \begin{bmatrix} \theta \\ \theta \end{bmatrix} \tag{13}$$

With this correct quantile specification condition, we also have $[P\{y_{1,t} \le q_{1,t}|\mathcal{F}_{t,0}\}, P\{y_{2,t} \le q_{2,t}|\mathcal{F}_{t,0}\}] = [\theta, \theta]$ as well as get consistent estimator $\widehat{\boldsymbol{\alpha}}^{(r)}$ to the true reduced-form one denoted as $\boldsymbol{\alpha}^{o\,(r)}$ consisting of the following elements.

$$\begin{aligned} \boldsymbol{\beta}^{o\,(r)} &:= \begin{bmatrix} 1 & 0 \\ -s^o_{2,1} & 1 \end{bmatrix}\boldsymbol{\beta}^{o\,(s)} = \begin{bmatrix} 1 & 0 \\ -s^o_{2,1} & 1 \end{bmatrix}\begin{bmatrix} \beta^o_{10} & \beta^o_{11} & \beta^o_{12} \\ \beta^o_{20} & \beta^o_{21} & \beta^o_{22} \end{bmatrix}, \\ \boldsymbol{\gamma}^{o\,(r)} &= \begin{bmatrix} 1 & 0 \\ -s^o_{2,1} & 1 \end{bmatrix}\boldsymbol{\gamma}^{o\,(s)} = \begin{bmatrix} 1 & 0 \\ -s^o_{2,1} & 1 \end{bmatrix}\begin{bmatrix} \gamma^o_{11} & \gamma^o_{12} \\ \gamma^o_{21} & \gamma^o_{22} \end{bmatrix}. \end{aligned} \tag{14}$$

From this example, we can see the condition (13) is a linkage which realizes the transition between the bivariate CAViaR estimator and the systemic one. On the other hand, the benefits of the systemic estimator are that we can check if there is significant contemporaneous effect to be accounted for and measure the systemic risk of Y_1 if Y_2 is a market index return variable. For inferring the significance of the contemporaneous terms in systemic MVMQ CAViaR models, we are going to provide some testing tools in the following.

2.3 Inference Testing

In this subsection, some inference tests are proposed in order to indicate if the contemporaneous terms in systemic MVMQ CAViaR models are significant enough so as to favour MVMQ CAViaR models or the systemic ones. There are some ways to design such tests. If we run MVMQ CAViaR regression first, it is rigours to consider if one response variable has some contemporaneous explanatory power on the conditional quantile of another response variable. Considering that, Section 2.3.1 takes the way to test if the quantile coverage of the latter response variable has a significant difference between with conditioning on the contemporaneous quantile violation of the first response variable and without the contemporaneous conditioning. We can also

[4] The subscript '(s)' is used to distinguish the parameters for systemic MVMQ-CAViaR models with the ones for MVMQ-CAViaR models which can be deemed in the reduced-form so use subscript '(r)'.

test if the first response variable or its functional forms, like its disturbance term, has significant contemporaneous explanatory power on the latter response variable. This path is taken in Section 2.3.2 based on dynamic quantile (DQ) tests developed by Engle and Manganelli (2004). If we run systemic MVMQ CAViaR regression first, it is also rigours to see if involved contemporaneous terms are significant enough. Wald tests typically can fulfil this testing role and are presented in Section 2.3.3.

2.3.1 Conditional Coverage Method

The conditional coverage method is commonly used for VaR backtesting. Unlike the unconditional coverage method which only focuses on the the frequency of VaR exceedance, the conditional coverage method (Christoffersen, 1998) is to perform tests on the conditional VaR coverage which is estimated by the conditional frequency of the VaR exceedance of a univariate time series. The conditional coverage in the tests is set to equal the unconditional one under the null hypothesis so that we can see if the conditional variables can significantly affect the probability of the VaR exceedance.

We adopt this method to test if the occurrence of one asset loss rate ($-y_{1t}$) exceeding its VaR$_\alpha$ ($-q_{1t}$) significantly affects the probability of a financial market index loss rate ($-y_{2t}$) exceeding its VaR$_\alpha$ ($-q_{2t}$) at the same time. To serve this purpose, the testing hypothesis is stated as follows:

$$\begin{cases} H_o: & \mathbb{E}\left[\mathbf{I}_{\{y_{2t}\le q_{2t}\}}\middle|\mathbf{I}_{\{y_{1t}\le q_{1t}\}=1}\right] = \alpha, \quad \text{for all } t; \\ H_A: & \mathbb{E}\left[\mathbf{I}_{\{y_{2t}\le q_{2t}\}}\middle|\mathbf{I}_{\{y_{1t}\le q_{1t}\}=1}\right] \neq \alpha, \quad \text{for all } t, \end{cases} \tag{15}$$

where $\mathbf{I}_{\{S\}}$ is the indicator function on a set S or a logical statement S, and $\mathbf{I}_{\{S\}} = 1$ if S is non-empty or true, otherwise $\mathbf{I}_{\{S\}} = 0$.

Count the frequencies of four possible outcomes of $\mathbf{I}_{\{y_{2t}\le q_{2t}\}} \times \mathbf{I}_{\{y_{1t}\le q_{1t}\}}$ over $t = 1, 2, \ldots, T$, and summarize it into the following table:

frequency in $\{y_t\}_{t=1}^T$	$\mathbf{I}_{\{y_{1t}\le q_{1t}\}} = 0$	$\mathbf{I}_{\{y_{1t}\le q_{1t}\}} = 1$
$\mathbf{I}_{\{y_{2t}\le q_{2t}\}} = 0$	n_{00}	n_{10}
$\mathbf{I}_{\{y_{2t}\le q_{2t}\}} = 1$	n_{01}	n_{11}

For testing the above hypothesis in VaR backtesting, the conditional coverage method provides a likelihood ratio statistic as follows:

$$\begin{aligned} LR_{cc} &= -2\,\log\left(\frac{L(\alpha, T)}{L(n_{00}, n_{01}, n_{10}, n_{11}, T)}\right) \\ &= -2\,\log\left(\frac{(1-\alpha)^{n_{00}+n_{10}}\alpha^{n_{01}+n_{11}}}{(1-\frac{n_{01}}{n_{00}+n_{01}})^{n_{00}}(\frac{n_{01}}{n_{00}+n_{01}})^{n_{01}}(1-\frac{n_{11}}{n_{10}+n_{11}})^{n_{10}}(\frac{n_{11}}{n_{10}+n_{11}})^{n_{11}}}\right) \\ &\overset{\mathcal{D}}{\sim} \chi^2(2). \end{aligned} \tag{16}$$

2.3.2 DQ Tests in bootstrap method to mitigate the heterogeneity effect of y_t

Dynamic quantile (DQ) tests proposed by Engle and Manganelli (2004) are intended to test if there is significant explanatory power of some omitted variables on the conditional quantile of a response variable over time. It can also serve for testing the hypothesis (15). DQ tests are analogous to the specification test illustrated by Chernozhukov and Umantsev (2001) for quantile regressions. Compared with the conditional coverage method, DQ tests are more flexible in testing omitted variables and can also perform in-sample tests by taking the asymptotic distribution of QLM estimator $\widehat{\boldsymbol{\alpha}}$ into account.

To test the contemporaneous explanatory power of $y_{i_c,t}$ on the $\theta_{i,j}$-th quantile $q_{i,j,t}$ of $y_{i,t}$ conditional on $\mathcal{F}_{t,0}$ with $1 \le i_c < i < n$ and $1 \le j \le p$, the testing hypothesis in our DQ tests is stated as follows:

$$\begin{cases} H_o: & P\left\{\epsilon_{i,j,t} \le 0 | y_{i_c,t}\right\} = \theta \quad \text{for all } t, \\ H_A: & P\left\{\epsilon_{i,j,t} \le 0 | y_{i_c,t}\right\} \neq \theta \quad \text{for all } t, \end{cases} \tag{17}$$

where $\epsilon_{i,j,t} := y_{i,t} - q_{i,j,t}$.

For testing the above hypothesis, the in-sample and the out-of-sample DQ test statistics are given in Theorem 7 and 8 respectively along with their asymptotic distributions.

Theorem 7 (In-sample DQ test statistic) *Suppose that the assumptions of Theorem 3 of White et al. (2015) and the H_o hypothesis in* (17) *hold. The in-sample DQ test statistic denoted as DQ_{IS} is given below with its asymptotic distribution.*

$$\begin{aligned} &DQ_{IS} \\ &= \frac{1}{T\theta(1-\theta)}\left(\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}-\theta_{i,j}\right)\right)^2\left(\widehat{G}_T'\widehat{Q}_T^{-1}\sum_{i=1,i\neq 2}^{n}\frac{\partial\widehat{q}_{i,j,t}}{\partial\widehat{\boldsymbol{\alpha}}}\frac{\partial\widehat{q}_{i,j,t}}{\partial\widehat{\boldsymbol{\alpha}}'}\widehat{Q}_T^{-1}\widehat{G}_T+\frac{1}{T}\sum_{t=1}^{T}\left(y_{i_c,t}-\frac{\partial q_{i,j,t}(\widehat{\boldsymbol{\alpha}})}{\partial\widehat{\boldsymbol{\alpha}}'}\widehat{Q}_T^{-1}\widehat{G}_T\right)^2\right) \\ &\overset{\mathcal{D}}{\sim}\chi^2(1), \end{aligned} \tag{18}$$

where

$$\widehat{G}_T := \frac{1}{T}\sum_{t=1}^{T} y_{i_c,t}\,\widehat{f}_{i,j,t}(0)\frac{\partial\widehat{q}_{i,j,t}}{\partial\widehat{\boldsymbol{\alpha}}}, \tag{19}$$

$\widehat{f}_{i,j,t}(0)$ is the estimate of the probability density function $f_{i,j,t}$ of $\epsilon_{i,j,t}$ at 0 *conditional on $\mathcal{F}_{t,i-1}$, obtained by the adaptive bandwidth method, and $\widehat{\boldsymbol{\alpha}}$, $\frac{\partial\widehat{q}_{i,j,t}}{\partial\widehat{\boldsymbol{\alpha}}}$ and $\widehat{Q}_T$ are obtained as instructed in the above theorems.*

Proof. See Appendix B. ■

Theorem 8 (Out-of-sample DQ test statistic) *Suppose that the assumptions of Theorem 3 of White et al. (2015) and the H_o hypothesis in* (17) *hold. Denote the number of in-sample observations as T_R and the number of out-of-sample observations as N_R. The subscript R indicates the dependence of T_R and N_R on R with the following properties:*

$$\begin{cases} \lim_{R\to\infty}\dfrac{N_R}{T_R}=0, \\ \lim_{R\to\infty}T_R=\infty, \\ \lim_{R\to\infty}N_R=\infty. \end{cases} \tag{20}$$

Then the out-of-sample DQ test statistic denoted as DQ_{OOS} is given below with its asymptotic distribution.

$$\begin{aligned} &DQ_{OOS} \\ &= \frac{1}{N_R\theta(1-\theta)}\left(\sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}-\theta\right)\right)^2\left(\frac{1}{N_R}\sum_{t=1}^{N_R}y_{i_c,t}^2\right)^{-1} \\ &\overset{\mathcal{D}}{\sim}\chi^2(1), \end{aligned} \tag{21}$$

where $q_{i,j,t}(\widehat{\boldsymbol{\alpha}})$ is obtained by plugging the QLME $\widehat{\alpha}$ of the in-sample $\{\mathbf{Y}_t\}_{t=1}^{T_R}$ into $q_{i,j,t}(\cdot)$.

Proof. See Appendix B. ■

2.3.3 Wald Tests

If we run systemic MVMQ CAViaR regression first and have the concern on the significance of the involved contemporaneous terms, Wald tests are adopted here to check if the contemporaneous terms are significant enough to remain in the model.

To test the significance of the contemporaneous term $g_{i,j}(\boldsymbol{s}_{i,j},\boldsymbol{u}_{i,t})$ in the generic model (2), the testing hypothesis in our Wald tests is stated as follows:

$$\begin{cases} H_o: & \boldsymbol{s}_{i,j}=\mathbf{0}, \\ H_A: & \boldsymbol{s}_{i,j}\neq\mathbf{0}. \end{cases} \tag{22}$$

For testing the above hypothesis, the Wald test statistic denoted by W_T is given as follows:

$$W_T = T\widehat{\boldsymbol{s}}_{i,j}'\left[R\widehat{Q}_T^{-1}\widehat{V}_T\,\widehat{Q}_T^{-1}R'\right]^{-1}\widehat{\boldsymbol{s}}_{i,j}\overset{\mathcal{D}}{\sim}\chi^2(d_{ij}) \tag{23}$$

where R is a $d_{ij} \times n$ matrix indicating the location of each element of $\boldsymbol{s}_{i,j}$ in $\boldsymbol{\alpha}$ such that $R\boldsymbol{\alpha} = \boldsymbol{s}'_{i,j}$, and $\widehat{\boldsymbol{s}}_{i,j}$, $\widehat{Q}_T$ and $\widehat{V}_T$ are the estimates of $\boldsymbol{s}_{i,j}$, Q and V in (6) respectively.

3 Measuring Systemic Risk via CoVaR

The European Central Bank (ECB) (2010) defines systemic risk in the financial market as a risk of financial instability so widespread that it impairs the functioning of a financial system to the point where economic growth and welfare suffer materially. CoVaR, a systemic risk measure proposed by Adrian and Brunnermeier (2011), is defined as the VaR of the financial system conditional on institutions being under distress. We generalized MVMQ-CAViaR models into the proposed systemic specification in order to incorporate the CoVaR specification on financial index return variables. Therefore, we can use systemic MVMQ-CAViaR models to measure the systemic risk of financial institutions of interest, and this application is elaborated below.

Given two asset return variables Y_1 and Y_2, $\text{CoVaR}_\theta^{2|\mathbb{C}(Y_1)}$ is formulated by Adrian and Brunnermeier (2011) as the θ-th quantile of the institution 2 (or the financial system) conditional on some event $\mathbb{C}(Y_1)$ of the institution 1, which fits in the following property:

$$Pr\left\{Y_2 \leq \text{CoVaR}_\theta^{2|\mathbb{C}(Y_1)} \middle| \mathbb{C}(Y_1)\right\} = \theta. \tag{24}$$

Suppose the bivariate time series $\{\boldsymbol{Y}_t = (y_{1,t}, y_{2,t})\}_{t=1}^T$ follows the DGP as (12) and we run the following systemic bivariate CAViaR regression:

$$\boldsymbol{Y}_t = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Y}_{t-1} + \boldsymbol{\gamma}\boldsymbol{q}_{t-1} + \begin{bmatrix} 0 \\ s \end{bmatrix} y_{1,t}, \tag{25}$$

where $\boldsymbol{\beta}_0$, $\boldsymbol{\beta}_1$, $\boldsymbol{\gamma}$ are 2×1, 2×2 and 2×2 parameter matrices. After the regression, we estimate $\text{CoVaR}_\theta^{2|Y_1}$ in use of the estimates $\widehat{\boldsymbol{\beta}}_0, \widehat{\boldsymbol{\beta}}_1, \widehat{\boldsymbol{\gamma}}, \widehat{s}$ as follows:

$$\widehat{\text{CoVaR}}_{\theta,t}^{2|Y_1} = \widehat{\beta}_{20} + \widehat{\beta}_{21}\, y_{1,t-1} + \widehat{\beta}_{22}\, y_{2,t-1} + \widehat{\gamma}_{21}\, q_{1,t-1} + \widehat{\gamma}_{22}\, q_{2,t-1} + \widehat{s}\, y_{1,t}. \tag{26}$$

And

$$\begin{aligned} &\widehat{\text{CoVaR}}_{\theta,t}^{2|y_{1,t}=\widehat{q}_{1,t}} \\ &= \left(\widehat{\beta}_{20} + \widehat{s}\,\widehat{\beta}_{10}\right) + \left(\widehat{\beta}_{21} + \widehat{s}\,\widehat{\beta}_{11}\right) y_{1,t-1} + \left(\widehat{\beta}_{22} + \widehat{s}\,\widehat{\beta}_{12}\right) y_{2,t-1} + \left(\widehat{\gamma}_{21} + \widehat{s}\,\widehat{\gamma}_{11}\right) q_{1,t-1} + \left(\widehat{\gamma}_{22} + \widehat{s}\,\widehat{\gamma}_{12}\right) q_{2,t-1}. \end{aligned} \tag{27}$$

If we also model the conditional 50%-th quantile $q_{1,t}^{(50\%)}$ of $y_{1,t}$ into the above systemic bivariate CAViaR model and get estimate $\widehat{q}_{1,t}^{(50\%)}$. $\Delta\text{CoVaR}_{\theta,t}^{2|1}$ defined by Adrian and Brunnermeier (2011) as the part of institution 1's systemic risk that can be attributed to Y_2 can be estimated as follows:

$$\Delta\text{CoVaR}_{\theta,t}^{2|1} := \widehat{\text{CoVaR}}_{\theta,t}^{2|y_{1,t}=\widehat{q}_{1,t}} - \widehat{\text{CoVaR}}_{\theta,t}^{2|y_{1,t}=\widehat{q}_{1,t}^{(50\%)}}. \tag{28}$$

Unlike the situation above which we know everything for certain, in reality we do not know the true model specification on an underlying DGP or precise contemporaneous terms to be involved. Considering that, we would like to give a rigorous application procedure here for studying contemporaneous tail dependence and CoVaR in use of systemic MVMQ-CAViaRs. In general, at first we would like to run an MVMQ-CAViaR regression based upon our knowledge on the multivariate time series of our interest. If we consider over some possible contemporaneous terms on explaining the conditional quantile of a response variate, DQ tests are implemented to check the significance of their explanatory power so that we can convincingly implement the systemic MVMQ-CAViaR regression when seeing the significance. After running the systemic MVMQ-CAViaR regression, Wald tests are applied to check if all the explanatory terms in the model are significant enough to be kept. After the Wald tests, the confirmed model specification is used to measure the systemic risk of involved financial institutions via their CoVaR as well as to measure contemporaneous tail dependence of involved financial assets. We will implement the above procedure in Section 5 with results presented correspondingly to each step.

4 Quantile Impulse Response Functions

The literature on quantile impulse response functions (QIRF) is scarce. We have a brief review here. White et al. (2015) presented a concept called pseudo quantile impulse response function in order to study how a shock to the present variable y_t influences the quantile (denoted as $q_{t+h|t}$) of its future variable y_{t+h} at h-th ($h \geq 1$) step ahead given the current information set $\mathcal{F}_t$. Actually, pseudo quantile impulse functions derived by White et al. (2015) strongly assume that the intermediate future variables $(y_{t+1}, \ldots, y_{t+h-1})$ right before the h-th step are fixed and not affected by the shock. Instead of fixing the intermediate future values $(y_{t+1}, \ldots, y_{t+h-1})$, Montes-Rojas (2019) considered quantile paths of $y_{t+1}, \ldots, y_{t+h-1}$ for forecasting q_{t+h}. However, the way that Montes-Rojas (2019) tackles the randomness of future quantile paths in forecasting $q_{t+h|t}$ is by fixing a specific future quantile path, such as assuming all median occurrences in the path. Although a future quantile path can be freely chosen to match some senario, the way of Montes-Rojas (2019) in forecasting $q_{t+h|t}$ can still not adapt to distributional characteristics of $y_{t+h}|\mathcal{F}_t$ (short for y_{t+h} conditional on $\mathcal{F}_t$), let alone $q_{t+h|t}$. The local projection method proposed by Jordà (2005) for estimating mean impluse response functions is also touched upon by Montes-Rojas (2019) to linearly regress $q_{t+h|t+h-1}$ on a specific quantile path of $y_{t+1}, \ldots, y_{t+h-1}$ and variables measurable to $\mathcal{F}_t$. Chavleishvili and Manganelli (2019) still used a quantile specification of $y_t|\mathcal{F}_{t-1}$ to represent the specification of the quantile of $y_{t+h}|\mathcal{F}_t$, and obtained the QIRF by manipulating the part of intermediate disturbances into zeros. Analogously to the fixed-intermediate (White et al., 2015) or specific future quantile path (Montes-Rojas, 2019) ideas, Han et al. (2019) and Jung and Lee (2019) used expectation of intermediate variables to define quantile impulse response functions, and adopted the local projection Jordà (2005) for estimation.

We aim in this section to define quantile impulse response function in a general way which can adapt to distributional characteristics of $y_{t+h}|\mathcal{F}_t$, and then to adopt the local projection idea with expansion of estimated terms to estimate quantile impulse response functions.

Considering a multivariate time series $\{\boldsymbol{Y}_t\}$ in a DGP as (2), let us discuss on how to forecast $\boldsymbol{Y}_{t+h}|\mathcal{F}_t$. Without loss of generality, we take a bivariate time series $\{\boldsymbol{Y}_t := (y_{1,t}, y_{2,t})\}_{t=1}^T$ with its model specification as follows:

$$\begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} q_{1,t-1} \\ q_{2,t-1} \end{bmatrix} + \begin{bmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{bmatrix}, \tag{29}$$

where

$$\boldsymbol{q}_t := \begin{bmatrix} q_{1,t} \\ q_{2,t} \end{bmatrix}$$

is the $\theta-$th quantile of $\boldsymbol{Y}_t := \begin{bmatrix} y_{1,t} \\ y_{2,t} \end{bmatrix}$ conditional on $\mathcal{F}_{t-1}$ which is the σ-algebra generated by $\{\boldsymbol{Y}_{t-1}, \boldsymbol{q}_{t-1}, \boldsymbol{Y}_{t-2}, \boldsymbol{q}_{t-2}, \ldots\}$ with

$$\begin{bmatrix} P\{\epsilon_{1,t} \leq 0|\mathcal{F}_{t-1}\} \\ P\{\epsilon_{2,t} \leq 0|\mathcal{F}_{t-1}\} \end{bmatrix} = \begin{bmatrix} \theta \\ \theta \end{bmatrix}. \tag{30}$$

Therefore,

$$\begin{bmatrix} q_{1,t} \\ q_{2,t} \end{bmatrix} = \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} + \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \begin{bmatrix} y_{1,t-1} \\ y_{2,t-1} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} q_{1,t-1} \\ q_{2,t-1} \end{bmatrix}. \tag{31}$$

Denote

$$\boldsymbol{\beta}_0 := \begin{bmatrix} \beta_{10} \\ \beta_{20} \end{bmatrix} \quad \boldsymbol{\beta}_1 := \begin{bmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{bmatrix} \quad \boldsymbol{\gamma} := \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix}. \tag{32}$$

Suppose the above model specification is known for $\{\boldsymbol{Y}_t\}$. And we want to forecast the quantile (denoted as $q_{t+h|t}$) of y_{t+h} ($h \geq 1$) given $\mathcal{F}_t$ now. At the first step, we need to rewrite the specification (29) of $\boldsymbol{Y}_{t+h}$ by iteratively substitutions until manifesting $\boldsymbol{\epsilon}_t$ and $\boldsymbol{q}_t$ as follows:

$$\begin{aligned}
\boldsymbol{Y}_{t+h} &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Y}_{t+h-1} + \boldsymbol{\gamma} \boldsymbol{q}_{t+h-1} + \boldsymbol{\epsilon}_{t+h} \\
&= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 (\boldsymbol{q}_{t+h-1} + \boldsymbol{\epsilon}_{t+h-1}) + \boldsymbol{\gamma} \boldsymbol{q}_{t+h-1} + \boldsymbol{\epsilon}_{t+h} \\
&= \boldsymbol{\beta}_0 + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})\, \boldsymbol{q}_{t+h-1} + (\boldsymbol{\epsilon}_{t+h} + \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+h-1}) \\
&= \boldsymbol{\beta}_0 + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})\, (\boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Y}_{t+h-2} + \boldsymbol{\gamma} \boldsymbol{q}_{t+h-2}) + (\boldsymbol{\epsilon}_{t+h} + \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+h-1}) \\
&= \sum_{i=1}^{h} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_0 + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{h-1} \boldsymbol{\beta}_1 \boldsymbol{Y}_t + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{h-1} \boldsymbol{\gamma} \boldsymbol{q}_t + \boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{\max\{h-1,1\}} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+h-i} \\
&= \sum_{i=1}^{h} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_0 + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{h} \boldsymbol{q}_t + \boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{h} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+h-i},
\end{aligned} \tag{33}$$

where $h \in \{1, 2, \ldots\}$.

It is worth mentioning that based on (33), there is an alternative way to rewrite $\boldsymbol{Y}_{t+h}$ into a function of $\{\boldsymbol{\epsilon}_{t+h}, \boldsymbol{\epsilon}_{t+h-1}, \ldots\}$:

$$\begin{aligned}
(\boldsymbol{I} - (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})\, L)\, \boldsymbol{Y}_{t+h} &= \boldsymbol{\beta}_0 + (\boldsymbol{I} - \boldsymbol{\gamma} L)\, \boldsymbol{\epsilon}_{t+h}, \\
\iff \boldsymbol{Y}_{t+h} &= (\boldsymbol{I} - (\boldsymbol{\beta}_1 + \boldsymbol{\gamma}))^{-1} \boldsymbol{\beta}_0 + \left(\sum_{i=0}^{\infty} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i} L^{i} \right) (\boldsymbol{I} - \boldsymbol{\gamma} L)\, \boldsymbol{\epsilon}_{t+h} \\
&= (\boldsymbol{I} - (\boldsymbol{\beta}_1 + \boldsymbol{\gamma}))^{-1} \boldsymbol{\beta}_0 + \boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{\infty} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+h-i}
\end{aligned} \tag{34}$$

where L is the lag operator, and it holds under the condition that the spectral radius of $(\boldsymbol{\beta}_1 + \boldsymbol{\gamma})$ is less than one . However, the first rewriting way (33) is more generally applicable to systemic MVMQ-CAViaR DGPs, which also reduces the number of explanatory variables in the consideration for forecasting $q_{t+h|t}$.

Following the result (33) from the first step, we now can get the preliminary predetermined part in $q_{t+h|t}$ as follows:

$$\begin{aligned}
q_{t+h|t} &:= \text{Quant}_\theta \left(\boldsymbol{Y}_{t+h} | \mathcal{F}_t \right) \\
&= \sum_{i=1}^{h-1} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_0 + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{h} \boldsymbol{q}_t + (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{h-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_t + \text{Quant}_\theta \left(\boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{h-1} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+i} \middle| \mathcal{F}_t \right).
\end{aligned} \tag{35}$$

Based on our assumptions before, $\left(\boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{h-1} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+i} \right)$ is not necessary to be independent of $\boldsymbol{q}_t$ or $\boldsymbol{\epsilon}_t$ unless $\{\boldsymbol{\epsilon}_t\}$ is independently distributed. The distribution of $\left(\boldsymbol{\epsilon}_{t+h} + \sum_{i=1}^{h-1} (\boldsymbol{\beta}_1 + \boldsymbol{\gamma})^{i-1} \boldsymbol{\beta}_1 \boldsymbol{\epsilon}_{t+i} \right)$ can vary with the information on $\{\boldsymbol{\epsilon}_t, \boldsymbol{p}_t, \boldsymbol{\epsilon}_{t-1}, \;\; \boldsymbol{p}_{t-1}, \ldots, \}$ so as to influence the distributional characteristic $q_{t+h|t}$. In order to forecast $q_{t+h|t}$ and study the effect of $\boldsymbol{\epsilon}_t$, we draw on the local projection idea of Jordà (2005) and consider to run the θ-th quantile regression of $\boldsymbol{Y}_{t+h}$ onto some explanatory variables measurable to $\mathcal{F}_t$. Based on the result (35), we know at least we should use explanatory variables $\boldsymbol{q}_t$ and $\boldsymbol{Y}_t$. We do not use $\boldsymbol{\epsilon}_t$ directly because it is not observed and has the same role as $\boldsymbol{Y}_t$ in $q_{t+h|t}$ when we include $\boldsymbol{q}_t$ as another explanatory variable. However, we do not observe $\boldsymbol{q}_t$ either. Using estimated $\boldsymbol{q}_t$ brings its estimation error in the quantile regression and make the regression result unreliable. So to mitigate the effect of its estimation error on the regression, we expand $\boldsymbol{q}_t$ to have more observed terms and use its expansion terms along with $\boldsymbol{Y}_t$ into the local quantile regression. Specifically, we expand $\boldsymbol{q}_t$ as follows:

$$\boldsymbol{q}_t = \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1 \boldsymbol{Y}_{t-1} + \boldsymbol{\gamma} \boldsymbol{q}_{t-1}, \tag{36}$$

$$= (\boldsymbol{\beta}_0 + \boldsymbol{\gamma} \boldsymbol{\beta}_0) + \boldsymbol{\beta}_1 \boldsymbol{Y}_{t-1} + \boldsymbol{\gamma} \boldsymbol{\beta}_1 \boldsymbol{Y}_{t-2} + \boldsymbol{\gamma}^2 \boldsymbol{q}_{t-2}. \tag{37}$$

So we can use explanatory variables $\{\boldsymbol{Y}_{t-1}, \boldsymbol{q}_{t-1}\}$ in replacement of $\boldsymbol{q}_t$ to mitigate the estimation error effect of $\boldsymbol{q}_t$, or use variables $\{\boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \boldsymbol{q}_{t-2}\}$ to further mitigate the estimation error effect as long as the spectral radius of $\boldsymbol{\gamma}$ is smaller than one because any v (denoted as an estimation error) will get vanished by $\lim_{n\to\infty}\boldsymbol{\gamma}^n\boldsymbol{v}=\mathbf{0}$. Simulation results in next section show that the local quantile regression result on the coefficient of $\boldsymbol{Y}_t$ become much more reliable when we replace the explanatory variable $\boldsymbol{q}_t$ with $\{\boldsymbol{Y}_{t-1},\boldsymbol{q}_{t-1}\}$ or $\{\boldsymbol{Y}_{t-1},\boldsymbol{Y}_{t-2},\boldsymbol{q}_{t-2}\}$.[5]

Now we come to defining *θ-th quantile response function* of $\{\boldsymbol{Y}_t\}$ given a shock $\boldsymbol{\delta}$ to $\boldsymbol{\epsilon}_t$ by taking the difference between $\text{Quant}_\theta\left(\boldsymbol{Y}^*_{t+h}\middle|\boldsymbol{\epsilon}^*_t:=\boldsymbol{\epsilon}_t+\boldsymbol{\delta},\mathcal{F}_{t-1}\right)$ and $\text{Quant}_\theta\left(\boldsymbol{Y}_{t+h}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)$ as follows:

$$\begin{aligned}&\text{QIRF}_h\left(\theta,\boldsymbol{\delta}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)\\&=\text{Quant}_\theta\left(\boldsymbol{Y}^*_{t+h}\middle|\boldsymbol{\epsilon}^*_t:=\boldsymbol{\epsilon}_t+\boldsymbol{\delta},\mathcal{F}_{t-1}\right)-\text{Quant}_\theta\left(\boldsymbol{Y}_{t+h}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)\\&=\text{Quant}_\theta\left(\boldsymbol{Y}^*_{t+h}\middle|\boldsymbol{Y}^*_t:=\boldsymbol{Y}_t+\boldsymbol{\delta},\mathcal{F}_{t-1}\right)-\text{Quant}_\theta\left(\boldsymbol{Y}_{t+h}|\boldsymbol{Y}_t,\mathcal{F}_{t-1}\right),\end{aligned}\tag{38}$$

where we can notice that $\text{Quant}_\theta\left(\boldsymbol{\epsilon}^*_t|\mathcal{F}_{t-1}\right)=\delta$ due to the shock, but $\text{Quant}_\theta\left(\boldsymbol{\epsilon}^*_{t+i}\middle|\boldsymbol{\epsilon}^*_{t+i-1},\ldots,\boldsymbol{\epsilon}^*_t,\mathcal{F}_{t-1}\right)=0$ $(i=1,\ldots,h)$ according to the correct specification assumption (30).

$\text{QIRF}_h\left(\theta,\boldsymbol{\delta}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)$ can be obtained by the local θ-quantile regression of Y_{t+h} onto $\{\boldsymbol{Y}_t,\boldsymbol{Y}_{t-1},\boldsymbol{q}_{t-1}\}$ or onto $\{\boldsymbol{Y}_t,\boldsymbol{Y}_{t-1},\boldsymbol{Y}_{t-2},\boldsymbol{q}_{t-2}\}$ as follows:

$$\text{QIRF}_h\left(\theta,\boldsymbol{\delta}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)=\lambda\delta,\tag{39}$$

where λ is the coefficient of Y_t in the local θ-quantile regression of Y_{t+h} onto $\{\boldsymbol{Y}_t,\boldsymbol{Y}_{t-1},\boldsymbol{q}_{t-1}\}$ or onto $\{\boldsymbol{Y}_t,\boldsymbol{Y}_{t-1},\boldsymbol{Y}_{t-w},\boldsymbol{q}_{t-2}\}$. We can also use higher moments of $\boldsymbol{Y}_t$ as explantory variables in the local quantile regression, and $\text{QIRF}_h\left(\theta,\boldsymbol{\delta}|\boldsymbol{\epsilon}_t,\mathcal{F}_{t-1}\right)$ can easily obtained by plugging the local quantile regression result into the definition (38).

5 Simulations

In this section, we are going to generate multivariate time series in an MVMQ-CAViaR DGP and a systemic MVMQ-CAViaR DGP, and implement the application procedure proposed in Section 3 on both DGPs to study their performances.

The MVMQ-CAViaR DGP that we simulate in this section is the bivariate CAViaR DGP specified below in which a bivariate time series sample is denoted as $\{\boldsymbol{Y}^{(r)}_t\}$.

$$\boldsymbol{Y}^{(r)}_t:=\begin{bmatrix}y^{(r)}_{1,t}\\y^{(r)}_{2,t}\end{bmatrix}=\begin{bmatrix}F^{-1}_{t(3)}(0.3)\\F^{-1}_{t(3)}(0.3)\end{bmatrix}+\begin{bmatrix}0.3&0.2\\0.2&0.3\end{bmatrix}\begin{bmatrix}y^{(r)}_{1,t-1}\\y^{(r)}_{2,t-1}\end{bmatrix}+\begin{bmatrix}0.2&0\\0&0.1\end{bmatrix}\begin{bmatrix}q^{(r)}_{1,t-1}\\q^{(r)}_{2,t-1}\end{bmatrix}+\begin{bmatrix}\epsilon^{(r)}_{1,t}\\\epsilon^{(r)}_{2,t}\end{bmatrix},\tag{40}$$

or equivalently

$$\boldsymbol{Y}^{(r)}_t=A_i+A_y\boldsymbol{Y}^{(r)}_{t-1}+A_q\boldsymbol{q}^{(r)}_{t-1}+\boldsymbol{\epsilon}^{(r)}_t,\tag{41}$$

where

$$A_i:=\begin{bmatrix}F^{-1}_{t(3)}(0.3)\\F^{-1}_{t(3)}(0.3)\end{bmatrix},\qquad A_y:=\begin{bmatrix}0.3&0.2\\0.2&0.3\end{bmatrix},\qquad A_q:=\begin{bmatrix}0.2&0\\0&0.1\end{bmatrix},\tag{42}$$

$\boldsymbol{q}^{(r)}_t:=\text{Quant}_{0.3}(\boldsymbol{Y}^{(r)}_t|\mathcal{F}_{t-1})$, and $\{\boldsymbol{\epsilon}^{(r)}_t-A_i\}$ is i.i.d. in Student's t-distribution with 3 degrees of freedom ($t(3)$ as the shorthand notation thereafter) with $\text{Quant}_{0.3}(\boldsymbol{\epsilon}^{(r)}_t-A_i|\mathcal{F}_{t-1})=[0,0]'$ for all t and $F^{-1}_{t(3)}(\cdot)$ denoted as the inverse probability distribution function of $t(3)$.

The systemic MVMQ-CAViaR DGP that we simulate in this section is the systemic bivariate CAViaR DGP specified below in which a bivariate time series sample is denoted as $\{\boldsymbol{Y}^{(s)}_t\}$.

$$\begin{bmatrix}1&0\\-0.5&1\end{bmatrix}\begin{bmatrix}y^{(s)}_{1,t}\\y^{(s)}_{2,t}\end{bmatrix}=\begin{bmatrix}F^{-1}_{t(3)}(0.3)\\F^{-1}_{t(3)}(0.3)\end{bmatrix}+\begin{bmatrix}0.3&0.2\\0.2&0.3\end{bmatrix}\begin{bmatrix}y^{(s)}_{1,t-1}\\y^{(s)}_{2,t-1}\end{bmatrix}+\begin{bmatrix}0.2&0\\0&0.1\end{bmatrix}\begin{bmatrix}q^{(s)}_{1,t-1}\\q^{(s)}_{2,t-1}\end{bmatrix}+\begin{bmatrix}\epsilon^{(s)}_{1,t}\\\epsilon^{(s)}_{2,t}\end{bmatrix},\tag{43}$$

or equivalently

$$A_s\,\boldsymbol{Y}^{(s)}_t=A_i+A_y\boldsymbol{Y}^{(s)}_{t-1}+A_q\boldsymbol{q}^{(s)}_{t-1}+\boldsymbol{\epsilon}^{(s)}_t,\tag{44}$$

[5]The optimal number of explanatory variables in replacing q_t is out of scope of this paper.

where

$$A_s = \begin{bmatrix} 1 & 0 \\ -0.5 & 1 \end{bmatrix}, \tag{45}$$

with A_i, A_y, A_q as defined in (42), $\boldsymbol{q}_t^{(s)} := [\text{Quant}_{0.3}(y_{1,t}^{(s)}|\mathcal{F}_{t,0}), \text{Quant}_{0.3}(y_{2,t}^{(s)}|\mathcal{F}_{t,1})]'$, and $\{\boldsymbol{\epsilon}_t^{(s)} - A_i\}$ is i.i.d. in $t(3)$ with $\text{Quant}_{0.3}(\boldsymbol{\epsilon}_t^{(s)} - A_i|\mathcal{F}_{t-1}) = [0,0]'$ for all t.

It is easy to simulate samples from these two DGPs. Specifically on simulating a sample in the bivariate DGP (41) with its sample size denoted as T, we first generate a sample of $\{\boldsymbol{\epsilon}_t^{(r)} - A_i\} \overset{\mathcal{D}}{\sim} t(3)$ in sample size $T+200$. And simulate $\left\{\boldsymbol{Y}_t^{(r)}\right\}_{t=1}^{T+200}$ based on the following equation:

$$\boldsymbol{Y}_t^{(r)} = \sum_{i=1}^{t-1}(A_y + A_q)^{i-1}A_i + (A_y + A_q)^{t-1}\,\boldsymbol{q}_1 + \boldsymbol{\epsilon}_t + \sum_{i=1}^{t-1}(A_y + A_q)^{i-1}A_y\boldsymbol{\epsilon}_{t-i}^{(r)} \tag{46}$$

with setting the initial value $\boldsymbol{q}_1^{(r)} = [0,0]'$. Delete the first 200 observations due to the burn-in effect of the initial value $\boldsymbol{q}_1^{(r)} = [0,0]'$ in the simulation, and return $\left\{\boldsymbol{Y}_t^{(r)}\right\}_{t=201}^{T+200}$ as the generated sample in the DGP (41). Analogously, we simulate a sample in the systemic bivariate DGP (44) by the following equation:

$$\boldsymbol{Y}_t^{(s)} = \sum_{i=1}^{t-1}(A_s^{-1}A_y + A_s^{-1}A_q)^{i-1}A_s^{-1}A_i + \left(A_s^{-1}A_y + A_s^{-1}A_q\right)^{t-1}\boldsymbol{q}_1^{(s)} + \boldsymbol{\epsilon}_t + \sum_{i=1}^{t-1}(A_s^{-1}A_y + A_s^{-1}A_q)^{i-1}A_s^{-1}A_yA_s^{-1}\boldsymbol{\epsilon}_{t-i}^{(s)} \tag{47}$$

with the same set-up that $\boldsymbol{q}_1^{(s)} = [0,0]'$ and the burn-in period of 200 observations.[6]

We can visually compare $\boldsymbol{Y}^{(r)}$ and $\boldsymbol{Y}^{(s)}$ by a plot of their samples as shown in Figure 1. As can be seen in Figure 1, $\{\boldsymbol{q}_t^{(r)}\}$ is quite smooth , not as bumpy or comovement-like as $\{\boldsymbol{q}^{(s)}\}$ which is due to the fact that in the systemic DGP the movement of $y_{1,t}$ immediately influences the conditional distribution of $y_{2,t}$ so as to be reflected in $y_{2,t}$.

After obtaining samples $\{\boldsymbol{Y}_t^{(r)}\}$ and $\{\boldsymbol{Y}_t^{(s)}\}$ of sample size $T = 5000$ from the DGPs (41) and (44) respectively, we regress both $\{\boldsymbol{Y}_t^{(r)}\}$ and $\{\boldsymbol{Y}_t^{(s)}\}$ onto the bivariate CAViaR model specification (41) of quantile index 0.3. After regressions, we run the DQ tests to check if $y_{1,t}$ still has significant contemporaneous explanatory power on the conditional 0.3-th quantile of $y_{2,t}$ with the hypothesis statement as in (17). We use two methods to estimate the asymptotic covariance matrix of the bivariate CAViaR model for $\{\boldsymbol{Y}_t^{(r)}\}$, namely the adaptive random bandwidth method (Hecq and Sun, 2020) and the kernel method with the optimal bandwidth used by White et al. (2015). From the size performances of these two methods in the DQ test of $\{\boldsymbol{Y}_t^{(r)}\}$ (see Table 1), we found that the adaptive random bandwidth (ARB) can well adapt to multivariate CAViaR models and robust in estimating the model asymptotic covariance matrix for various inference tests. Hereafter we only show test results in use of the adaptive random bandwidth method. The size performances of the DQ tests of $\{\boldsymbol{Y}_t^{(r)}\}$ and $\{\boldsymbol{Y}_t^{(s)}\}$ are shown in Table 1. We can see that the DQ test works robustly in indicating if some contemporaneous terms are significant to be involved into the modelling.

Table 1: Rejection rates (Size performances) of DQ tests (17) after regressing both $\boldsymbol{Y}^{(r)}$ and $\boldsymbol{Y}^{(s)}$ onto the bivariate CAViaR model specification (41)

DQ Tests	significance level: $\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$
$\boldsymbol{Y}^{(r)}$ (ARB)	0.010	0.0470	0.0790	0.1900	0.2980
$\boldsymbol{Y}^{(r)}$ (kernel)	0.005	0.0420	0.0730	0.1790	0.2790
$\boldsymbol{Y}^{(s)}$ (ARB)	1.000	1.000	1.000	1.000	1.000

Now we regress both $\{\boldsymbol{Y}_t^{(r)}\}$ and $\{\boldsymbol{Y}_t^{(s)}\}$ onto the systemic bivariate CAViaR model specification (44) of quantile index 0.3. After regressions, we run the Wald tests to check if $y_{1,t}$ is significant enough in explaining

[6] The length of burn-in periods chosen in this paper is based on our experience, and it can be adjusted for each specific DGP based on readers' expertise.

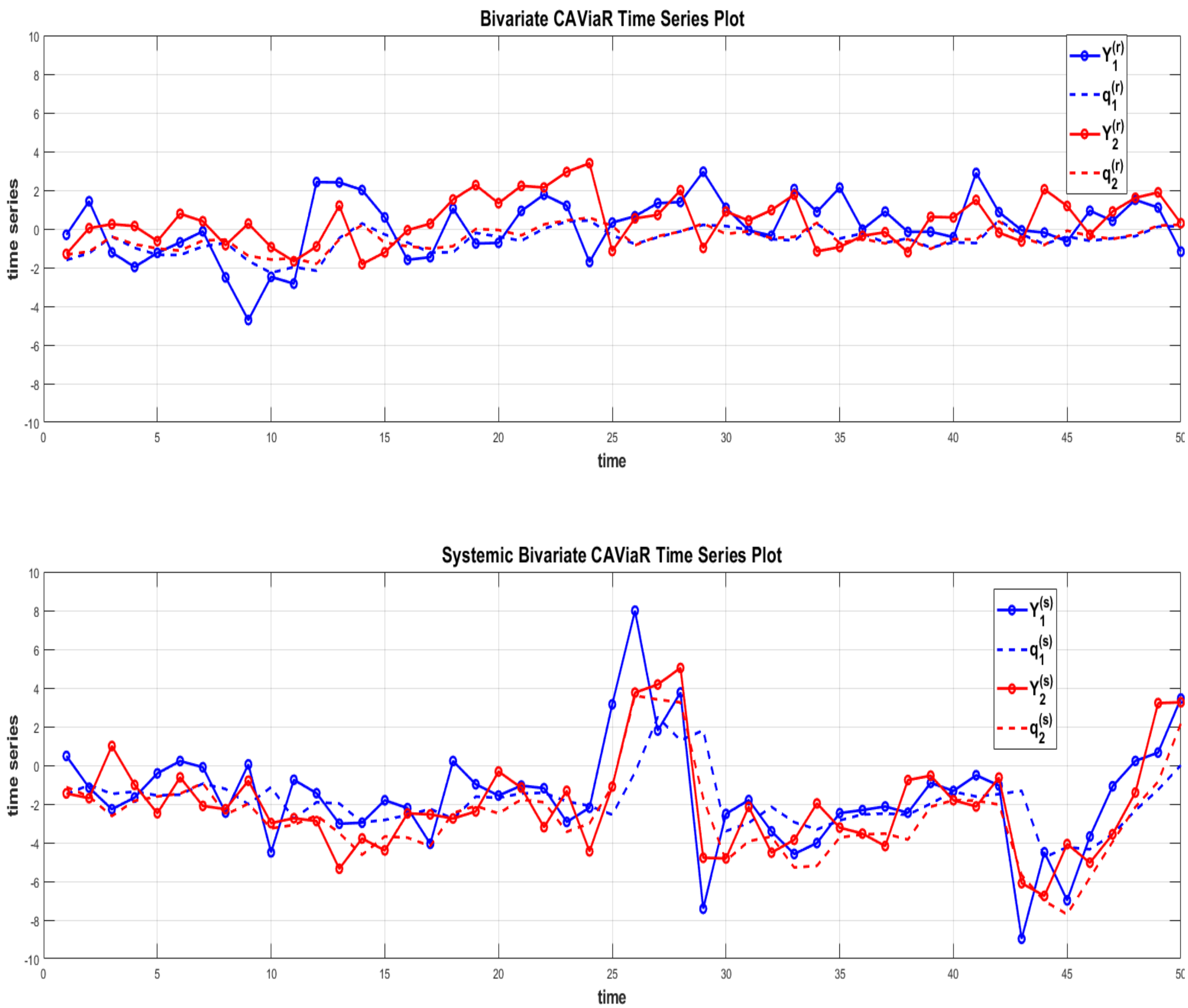

Figure 1: Compare $\boldsymbol{Y}^{(r)}$ and $\boldsymbol{Y}^{(s)}$.

Table 2: Rejection rates (Size performances) of Wald tests (22) after regressing both $\boldsymbol{Y}^{(r)}$ and $\boldsymbol{Y}^{(s)}$ onto the systemic bivariate CAViaR model specification (44)

Wald Tests (22)	significance level: $\alpha = 0.01$	$\alpha = 0.05$	$\alpha = 0.1$	$\alpha = 0.2$	$\alpha = 0.3$
$\boldsymbol{Y}^{(r)}$ (ARB)	0.0260	0.0680	0.1120	0.1980	0.2860
$\boldsymbol{Y}^{(s)}$ (ARB)	1.000	1.000	1.000	1.000	1.000

the conditional 0.3-th quantile of $y_{2,t}$ with the hypothesis statement as in (22). The Wald test performances are shown in Table 2. We can see that the Wald test works robustly on confirming if some contemporaneous terms are significant enough to be kept in the model.

After confirming the systemic model specification for $\boldsymbol{Y}^{(s)}$, we can use the systemic model regression result to measure the systemic risk of Y_1 by estimating $\text{CoVaR}_{0.3,t}^{2|y_{1,t}=q_{1,t}}$ as instructed in Section 3. Figure 2 shows a sample of $\boldsymbol{Y}^{(s)}$ with its $\widehat{\boldsymbol{q}}^{(s)}$, $-\widehat{\text{CoVaR}}_{0.3,t}^{2|y_{1,t}=\widehat{q}_{1,t}}$. We plot $-\widehat{\text{CoVaR}}_{0.3,t}^{2|y_{1,t}=\widehat{q}_{1,t}}$ not $\widehat{\text{CoVaR}}_{0.3,t}^{2|y_{1,t}=\widehat{q}_{1,t}}$ with $\boldsymbol{Y}^{(s)}$ because we regard $\boldsymbol{Y}^{(s)}$ as return variables so that we can comparatively view $\boldsymbol{Y}^{(s)}$ with its $\widehat{\boldsymbol{q}}^{(s)}$ and $-\widehat{\text{CoVaR}}_{0.3,t}^{2|y_{1,t}=\widehat{q}_{1,t}}$ in one plot.

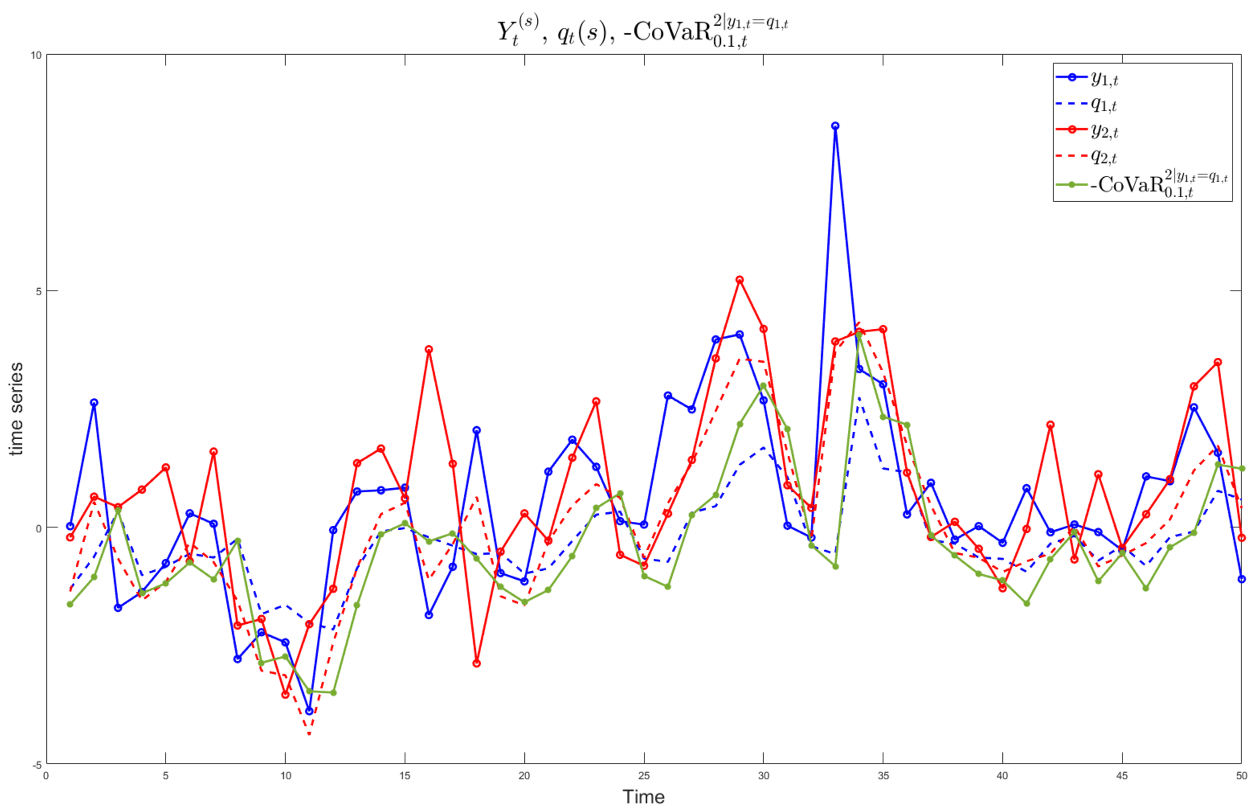

Figure 2: time series plot of sample $\boldsymbol{Y}^{(s)}$ with its $\widehat{\boldsymbol{q}}^{(s)}$, $-\widehat{\text{CoVaR}}_{0.3,t}^{2|y_{1,t}=\widehat{q}_{1,t}}$

Now we confirm the bivariate CAViaR model for $\boldsymbol{Y}^{(r)}$ and the systemic model for $\boldsymbol{Y}^{(s)}$. And we study their 0.3-th quantile impulse response functions (QIRFs) by the local 0.3-th quantile regression of $\boldsymbol{Y}_{t+h}$ $(h \leq 1)$ onto vector regressors $\{\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \widehat{\boldsymbol{q}}_{t-2}\}$. Figure 3 and 4 compare the QIRF results among using explanatory variables $\{\boldsymbol{Y}_t, \widehat{\boldsymbol{q}}_t\}$ and $\{\boldsymbol{Y}_t, \boldsymbol{Y}_{t-1}, \boldsymbol{Y}_{t-2}, \widehat{\boldsymbol{q}}_{t-2}\}$ with the true QIRF of $\boldsymbol{Y}$, in which we can see that using the expansion terms of $\widehat{\boldsymbol{q}}_t$ is more robust over using $\widehat{\boldsymbol{q}}_t$ directly. In fact, the outperformance of using $\widehat{\boldsymbol{q}}_t$ is more obvious for tail quantile indexes like 0.1 with the sample size being relatively large enough compared to the number of coefficients to be estimated in a local quantile regression.

In next section, we are going to implement the above application procedure on some empirical data and

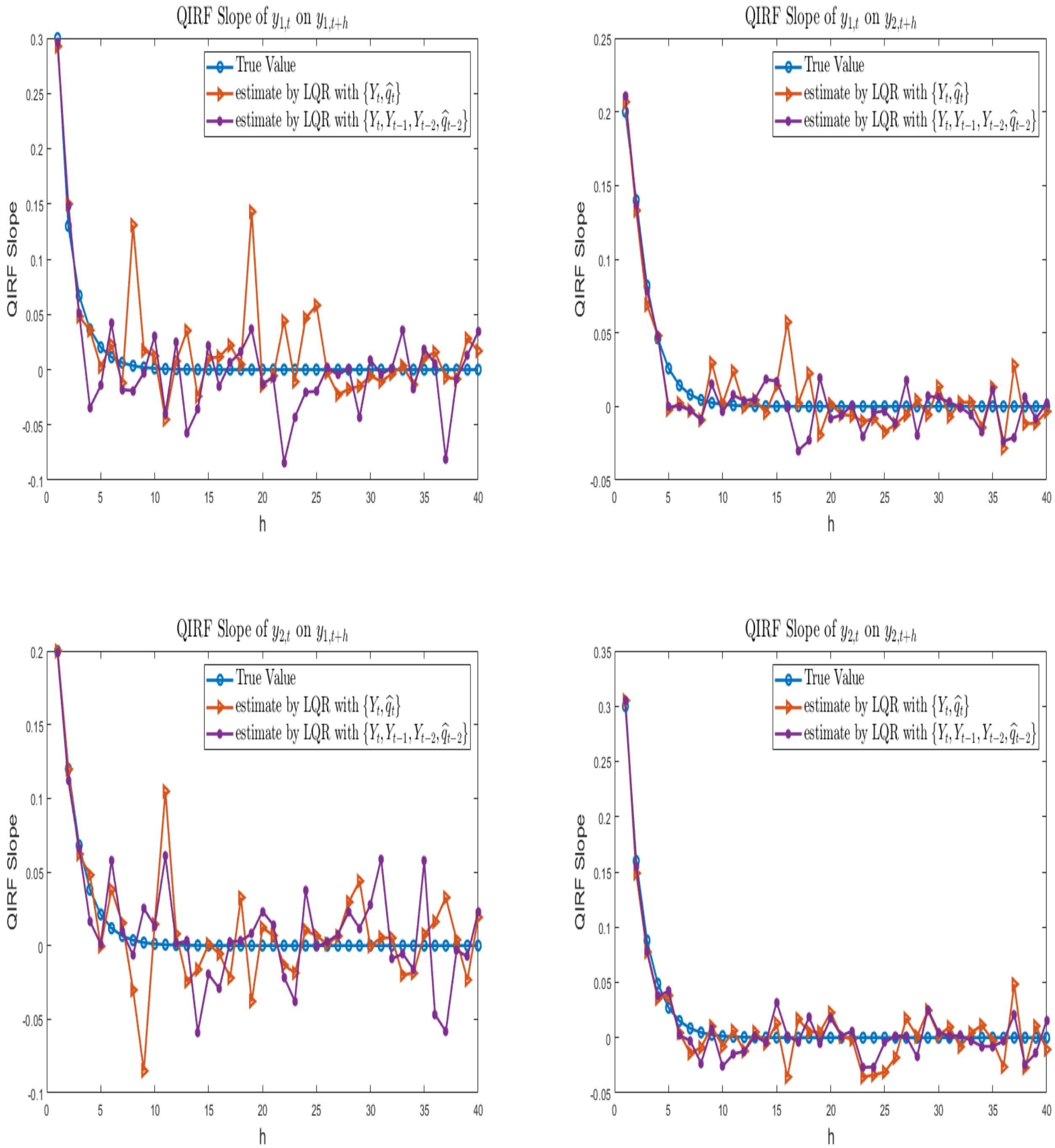

Figure 3: QIRF of $\boldsymbol{Y}^{(r)}$ estimated by the local 0.1-th quantile regression

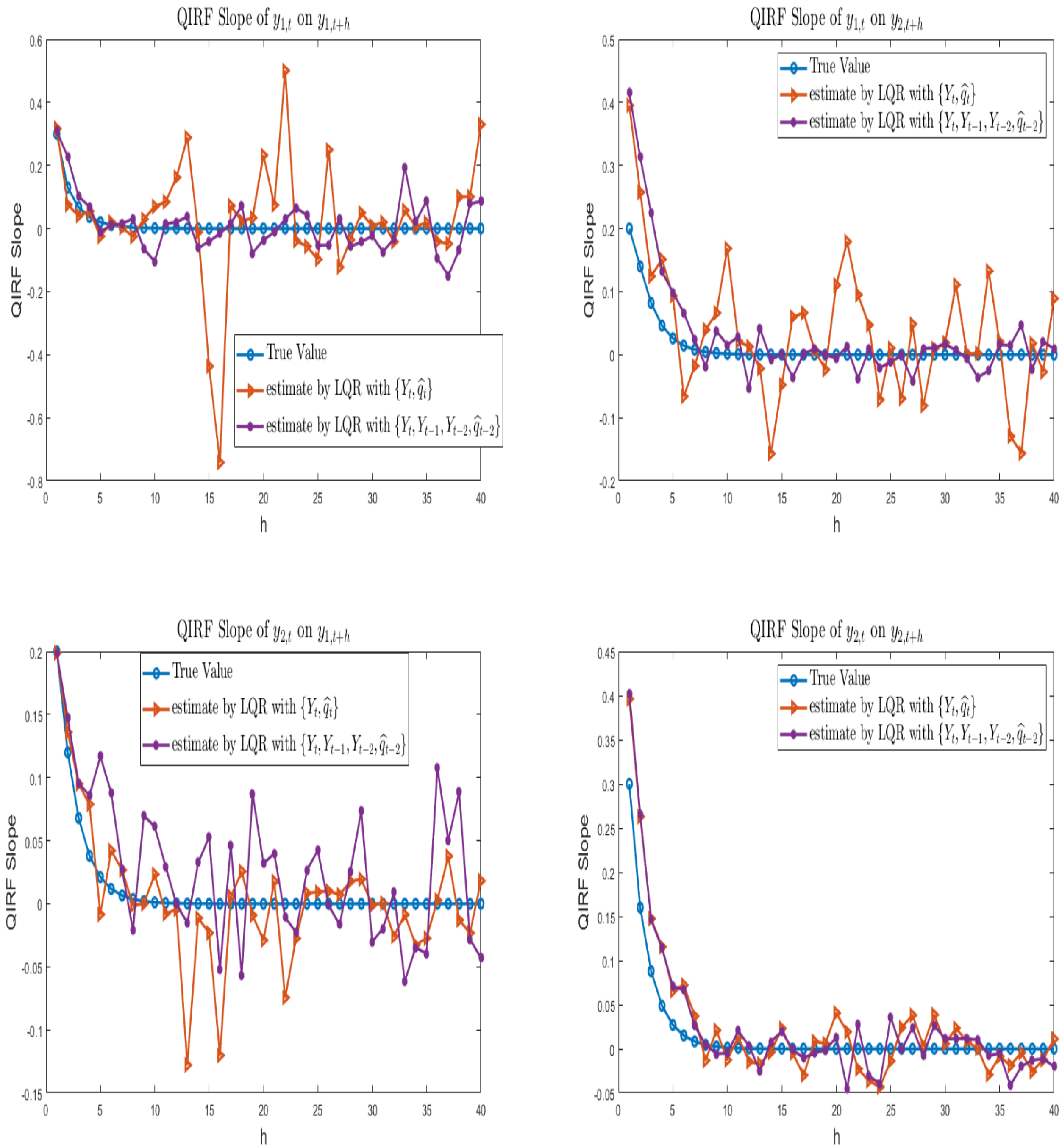

Figure 4: QIRF of $\boldsymbol{Y}^{(s)}$ estimated by the local 0.1-th quantile regression

analyse the results for our empirical knowledge.

6 Empirical Applications

We apply the above systemic modelling procedure to study the systemic risks of JPMorgan Chase & Co., the Bank of America Corporation, Wells Fargo & Company, the Goldman Sachs Group, Inc., Citigroup Inc. and Morgan Stanley which are the six largest bank holding companies in the United States ranked by total assets of March 31, 2020 per the Federal Financial Institutions Examination Council. Systemic risk of large financial intuitions can be as scary as the Great Recession that occurred between 2007 – 2009 in national economies globally. Before Lehman Brothers Holdings Inc. filed for bankruptcy in 2008, it was the fourth-largest investment bank in the United States. After Lehman Brothers filed for bankruptcy, global markets immediately plummeted and investors lost confidence, which caused bank runs, funding liquidity shortage, high haircuts, fire sales of assets and high counterpart credit risk in financial markets. The distress was spreading over financial institutions globally, and triggered the financial crisis of 2007—2008 which sparked the Great Recession (2007 – 2009), the most severe global recession since the Great Depression (1929-1933). Therefore, we are concerned about systemic risk of big financial intuitions and would like to measure the systemic risk of the six largest bank holding companies by our proposed method.

We use the S&P500 as the market index of interest which is deemed vulnerable to the systemic risks of those big banks. We downloaded daily adjusted closing stock prices of these six banks and the S&P500 from *Yahoo! Finance*, and each stock price time series has 3189 daily prices, ranging from 31-Dec-2006 to 01-Sep-2019. The price data were converted to percentage returns by multiplying 100 with the difference of the natural logarithm of the daily prices. The obtained return time series of each stock contains 3188 observations in the period of 01-Jan-2007 to 01-Sep-2019. In each return time series, the last 300 observations are used for the out-of-sample testing after the first 2888 observations are used to estimate the model.

We measure the systemic risk of each of these six banks individually by bivariate CAViaR models with the S&P500 daily returns. The bivariate CAViaR models used for estimating the conditional 5% quantiles of a big bank and the S&P500 daily returns are (29) and (12) for bivariate CAViaR and systemic bivariate CAViaR regressions respectively.

We set up inference tests in the same way as we did in the preceding sections. The inference testing results based on our systemic modelling procedure for the conditional 5% quantiles of the six banks' and the S&P500 daily returns are presented in Table 3 for in-sample tests and Table 4 for out-of-sample tests. The DQ p-value (MVMQ) columns of Table 3 and 4 provide strong evidence of contemporaneous (daily) spillovers of financial distress at individual financial institutions to the S&P500 index return. There is further evidence that the involved contemporaneous terms are significant in the systemic bivariate CAViaR models according to the Wald p-value (SMVMQ) columns. From the results of the out-of-sample tests in Table 4 for $\text{VaR}_{5\%}$ backtesting, we see that we can not reject the systemic model of Goldman Sachs and the S&P500 which even has $\text{VaR}_{5\%}$ exceedance rates close to the risk level 5%. Other systemic models are rejected by the out-of-sample DQ tests which means that those banks still have significant explanatory power on conditional 0.5-th quantiles of the market index which is not revealed by the in-sample estimation. There are many possible reasons behind those model rejections. One reason can be the inappropriate functional form of the contemporaneous terms in banks' returns we considered in those models. Another possible reason is that those rejected models omitted some other significant contemporaneous terms such as Goldman Sachs. The results in Table 3 and 4 let us confirm the systemic model of Goldman Sachs and the S&P500 first so as to measure the systemic risk of Goldman Sachs, and also gives us some clues to explore the functional forms of the contemporaneous effect of the banks on the market index. For example, we can run the systemic model on one of the five banks in model rejections with Goldman Sachs so that we count out the common contemporaneous effect of that bank on the market index with Goldman Sachs and focus on their idiosyncratic parts for the contemporaneous effects and the proper functional forms. We are not going to measure the systemic risks of these six banks together by a seven-variate CAViaR model in this paper. Since systemic MVMQ-CAViaR models are directional, we have to decide on the contemporaneous influence direction among these seven stocks. However, the results above of the (systemic) bivariate CAViaR models can be the starting point to build a proper seven-variate CAViaR model and then study their systemic risks in a whole system, which involves a enumeration of model estimations and inference tests and is left for

Table 3: In-sample test results on the empirical data (quantile index=0.05)

tests	VaR exceedance rates (MVMQ) [bank, the S&P500]	DQ p-value (MVMQ)	VaR exceedance rates (SMVMQ) [bank, the S&P500]	Wald p-value (SMVMQ)
BAC	[0.0502, 0.0499]	0	[0.0495, 0.0495]	0
C	[0.0506, 0.0502]	0	[0.0502, 0.0506]	0
GS	[0.0516, 0.0509]	0	[0.0519, 0.0492]	0
JPM	[0.0502, 0.0533]	0	[0.0499, 0.0488]	0
MS	[0.0512, 0.0506]	0	[0.0509, 0.0519]	0
WFC	[0.0499, 0.0561]	0	[0.0495, 0.0495]	0

Table 4: Out-of-sample test results on the empirical data (quantile index=0.05)

tests	VaR exceedance rates (MVMQ) [bank, the S&P500]	DQ p-value (MVMQ)	VaR exceedance rates (SMVMQ) [bank, the S&P500]	DQ p-value (SMVMQ)
BAC	[0.0733, 0.0367]	0	[0.0833, 0.0733]	0
C	[0.0667, 0.0367]	0	[0.0867, 0.0800]	0
GS	[0.0500, 0.0400]	0	[0.0567, 0.0700]	0.4576
JPM	[0.0167, 0.0400]	0	[0.0167, 0.0767]	0.0005
MS	[0.0267, 0.0433]	0	[0.0467, 0.0700]	0
WFC	[0.0367, 0.0433]	0	[0.0467, 0.0833]	0.0015

future research.

After confirming the systemic model specification for the conditional 5%-th quantiles of Goldman Sachs and the S&P500, we can use the systemic model regression result to measure the systemic risk of Goldman Sachs by estimating $\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(5\%)}$ as instructed in Section 3.

$$\Delta\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}} := \text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(5\%)} - \text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(50\%)}$$

is defined by Adrian and Brunnermeier (2011) as the part of Goldman Sachs' systemic risk that can be attributed to the S&P500. To view $\Delta\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}}$, we also need to model and estimate the conditional 50%-th quantiles of Goldman Sachs' returns. Analogously, we regress the returns of Goldman Sachs and the S&P500 onto the bivariate CAViaR (29) and the systemic bivariate CAViaR model (12) of quantile index 0.50 respectively, and perform the inference tests as we did before which results in Table 5 and 6.

Table 5: In-sample test results on the returns of Goldman Sachs and the S&P500 (quantile index=0.50)

tests	VaR exceedance rates (MVMQ) [bank, the S&P500]	DQ p-value (MVMQ)	VaR exceedance rates (SMVMQ) [bank, the S&P500]	Wald p-value (SMVMQ)
GS	[0.5000, 0.4997]	0	[0.4997, 0.4997]	0

As we can see that the systemic model is not rejected for the conditional 0.5-th quantiles of Goldman Sachs' and the S&P500 returns so that we can use its estimated conditional 0.5-th quantiles of Goldman Sachs' returns to calculate $\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(50\%)}$ so for us to view $\Delta\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}}$. The estimated conditional 0.05-th and 0.5-th quantiles of Goldman Sachs' and the S&P500 returns are plotted in Figure 8 and 9 respectively, see Appendix A. Figure 5 shows the returns of the S&P500 index, $-\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(5\%)}$ and $-\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}=q_{GS,t}(50\%)}$. We plot $-\text{CoVaR}_{0.05,t}^{SP500|y_{GS,t}=q_{GS,t}}$ not $\text{CoVaR}_{0.05,t}^{SP500|y_{GS,t}=q_{GS,t}}$ with the returns so that we can comparatively view the returns of the S&P500 index with $-\text{CoVaR}_{0.05,t}^{SP500|y_{GS,t}=q_{GS,t}}$ in one plot. Figure 6 reflects the part of Goldman Sachs' systemic risk attributed to the S&P500 index over time by plotting $\{\Delta\text{CoVaR}_{5\%,t}^{SP500|y_{GS,t}}\}$. We can see from Figure 6 that the part of Goldman Sachs' systemic risk attributed to the S&P500 index was cumulating after Lehman Brothers filed for bankruptcy on September 15, 2008, and reached unprecedentedly high until mid-2009. Figure 6 is quite story-telling and

Table 6: Out-of-sample test results on the returns of Goldman Sachs and the S&P500 (quantile index=0.05)

tests	VaR exceedance rates (MVMQ) [bank, the S&P500]	DQ p-value (MVMQ)	VaR exceedance rates (SMVMQ) [bank, the S&P500]	DQ p-value (SMVMQ)
GS	[0.5133, 0.5033]	0	[0.5233, 0.4700]	0.5607

links almost every peak in the figure to a distress event on Goldman Sachs, which can be informative for financial market regulators in systemic risk management.

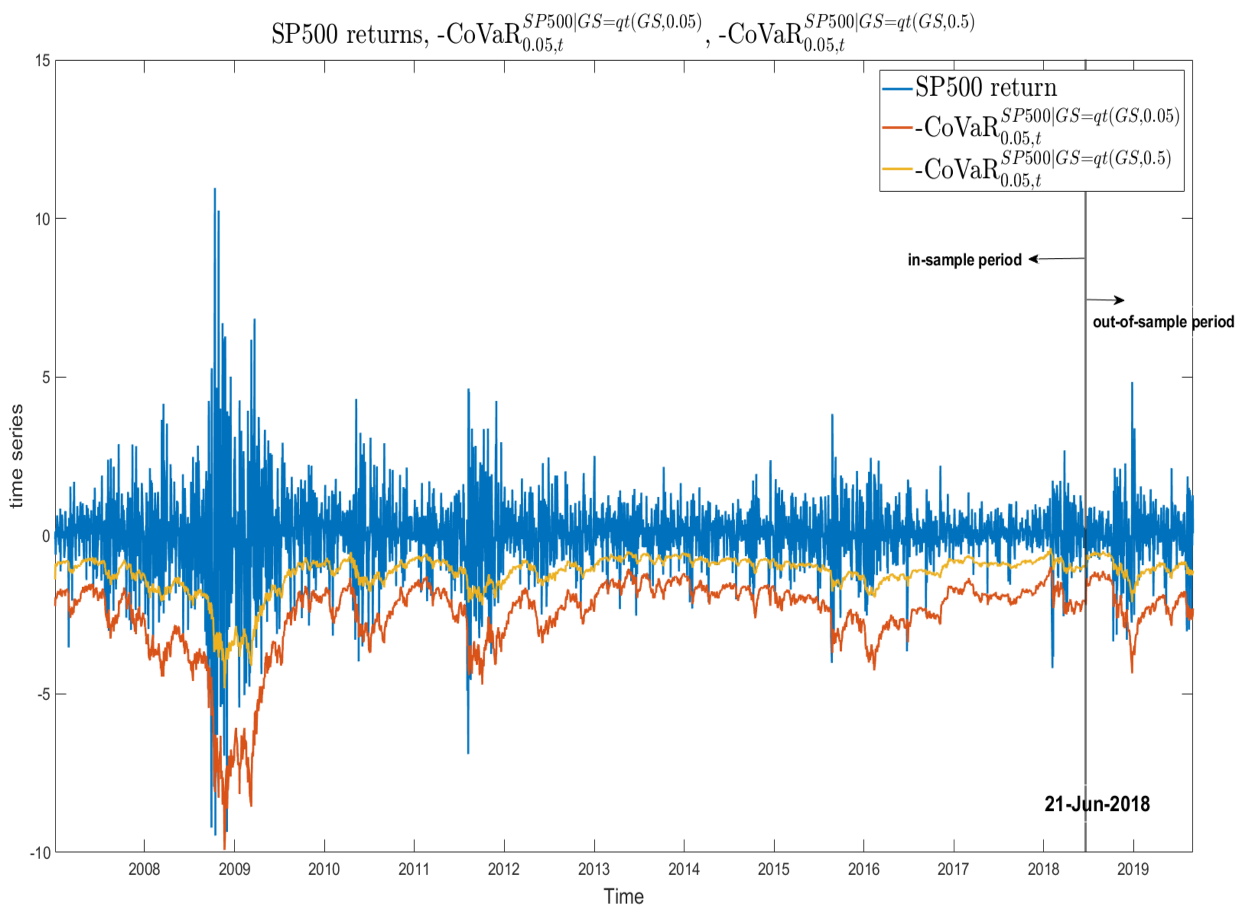

Figure 5: CoVaR plot of Goldman Sachs on the SP500

The 0.05-th QIRF coefficients of Goldman Sachs on the SP500 is obtained by the local 0.05-th quantile regressions (see Section 4) in use of the expansion terms up to 3-step lagged return vector variables, which is plotted in Figure 7. The 0.05-th QIRF coefficient drawn at $h = 0$ in Figure 7 is the estimated coefficient of the contemporaneous term in Goldman Sachs' return on the conditional 0.05-th quantile of the SP500 return by the systemic model regression 12.[7] Figure 7 says that a shock to Goldman Sachs' return at $h = 0$ contemporaneously shifts the conditional 0.05-th quantile of the S&P500 return in tandem considerably, and in the rest of days the conditional 0.05-th quantile of the S&P500 returns are less memorable of this shock.

From the empirical application above, we have seen that the contemporaneous effects of the big banks' returns are significant on conditional quantiles of the S&P500 returns, and it is informative to use systemic MVMQ CAViaR models proposed in this paper with the systemic modelling procedure (see Section 3) to monitor and analyse the systemic risks of big financial institutions.

[7]When we used the expansion terms up to 4-step lagged return vector variables in the local 0.05-th quantile regression, we got a similar result to the 0.05-th QIRF coefficientsin Figure 7.

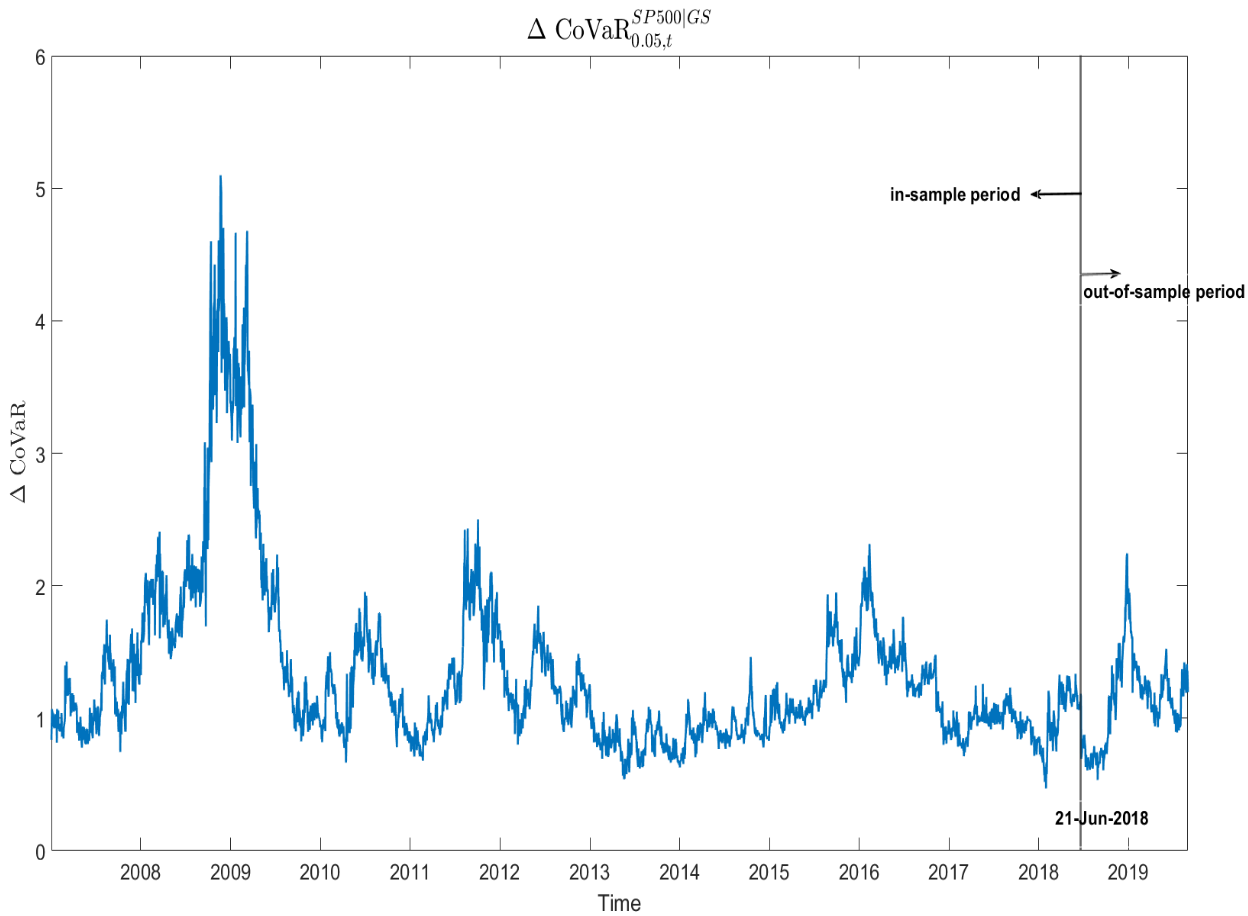

Figure 6: $\Delta\text{CoVaR}^{SP500|y_{GS,t}}_{5\%,t}$ plot over time

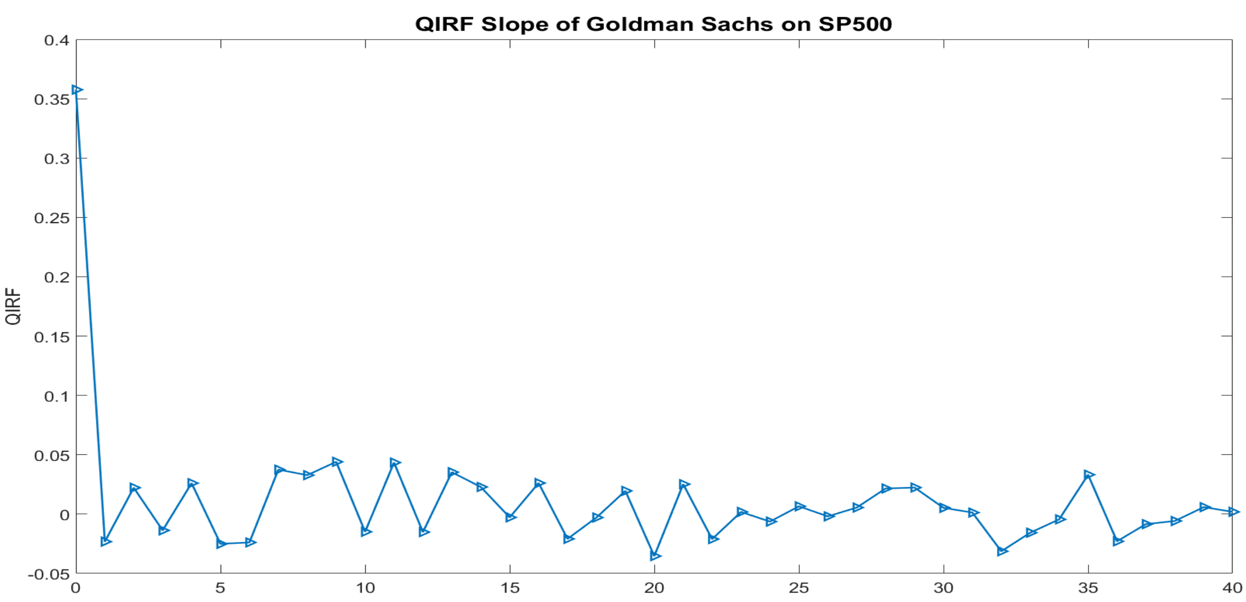

Figure 7: 0.05-th QIRF coefficient of Goldman Sachs on the SP500

7 Conclusions

We generalized multivariate multi-quantile CAViaR models (MVMQ-CAViaR, see White et al., 2015) by incorporating CoVaR specification (see Adrian and Brunnermeier, 2011) into the model specification in this paper. Our proposed model presents a vector-autoregressive (VAR) specification of financial institutions' value-at-risk (VaR) as well as their CoVaR. This model generalization is able to capture contemporaneous tail dependence of financial institutions and market indexes so that we can interpret the systemic risks of the institutions over time. The consistency and asymptotic normality proofs of this generalized model are provided in this paper along with some relevant inference tests, for which we implemented simulation tests and showed robust model performances. For tracing the transmission of a single shock to a financial institution in the financial system, we also constructed quantile impulse response functions (QIRF) accordingly in use of the local projection idea (Jordà, 2005) and expansion of estimated terms. Based on our simulation results, we can see that using the expansion terms of $\widehat{\boldsymbol{q}}_t$ is more robust than directly using $\widehat{\boldsymbol{q}}_t$ in the local quantile regression for QIRF estimation. Applications to real data provided empirical support to this methodology.

References

Adrian, T. and Brunnermeier, M. K. (2011). Covar. Technical report, National Bureau of Economic Research.

Chavleishvili, S. and Manganelli, S. (2019). Forecasting and stress testing with quantile vector autoregression.

Chernozhukov, V. and Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics*, 26(1):271–292.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International economic review*, pages 841–862.

Engle, R. F. and Manganelli, S. (2004). Caviar: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics*, 22(4):367–381.

Han, H., Jung, W., and Lee, J. H. (2019). Estimation and inference of quantile impulse response functions by local projections: With applications to var dynamics. *Available at SSRN 3466198*.

Hecq, A. and Sun, L. (2020). Adaptive random bandwidth for caviar models.

Jordà, Ò. (2005). Estimation and inference of impulse responses by local projections. *American economic review*, 95(1):161–182.

Jung, W. and Lee, J. H. (2019). Quantile impulse response analysis with applications in macroeconomics and finance.

Koenker, R. and Bassett Jr, G. (1978). Regression quantiles. *Econometrica: journal of the Econometric Society*, pages 33–50.

Koenker, R. and Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association*, 101(475):980–990.

Montes-Rojas, G. (2019). Multivariate quantile impulse response functions. *Journal of Time Series Analysis*, 40(5):739–752.

White, H. (2001). *Asymptotic Theory for Econometricians. 2001.* Academic Press, New York.

White, H., Kim, T. H., and Manganelli, S. (2015). Var for var: Measuring tail dependence using multivariate regression quantiles. *Journal of Econometrics*, 187(1):169–188.

A Appendix: Extra Figures

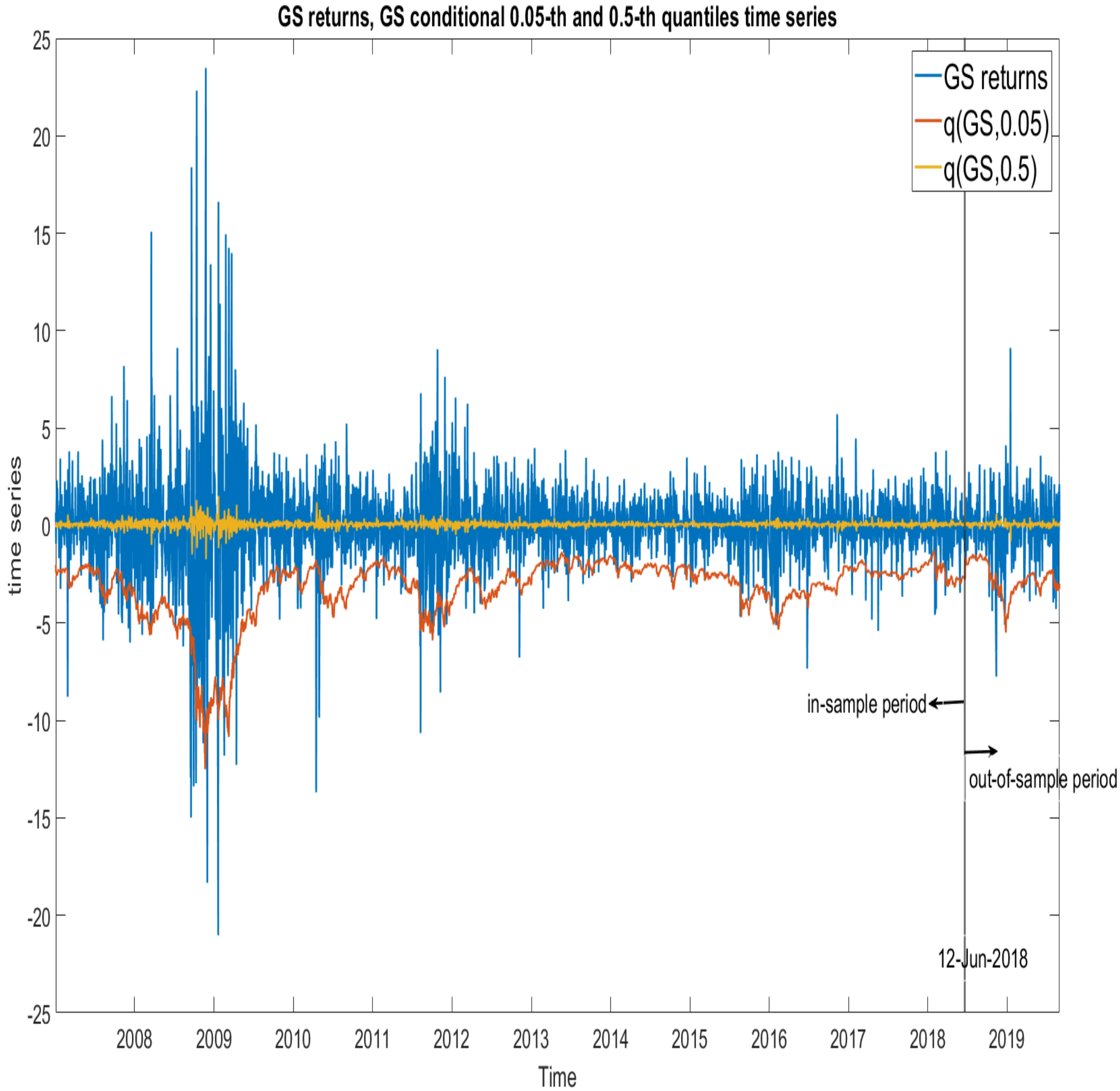

Figure 8: time series plot of the returns of Goldman Sachs with its fitted conditional 0.05-th and 0.5-th quantiles by the systemic MVMQ CAViaR regression with the SP500

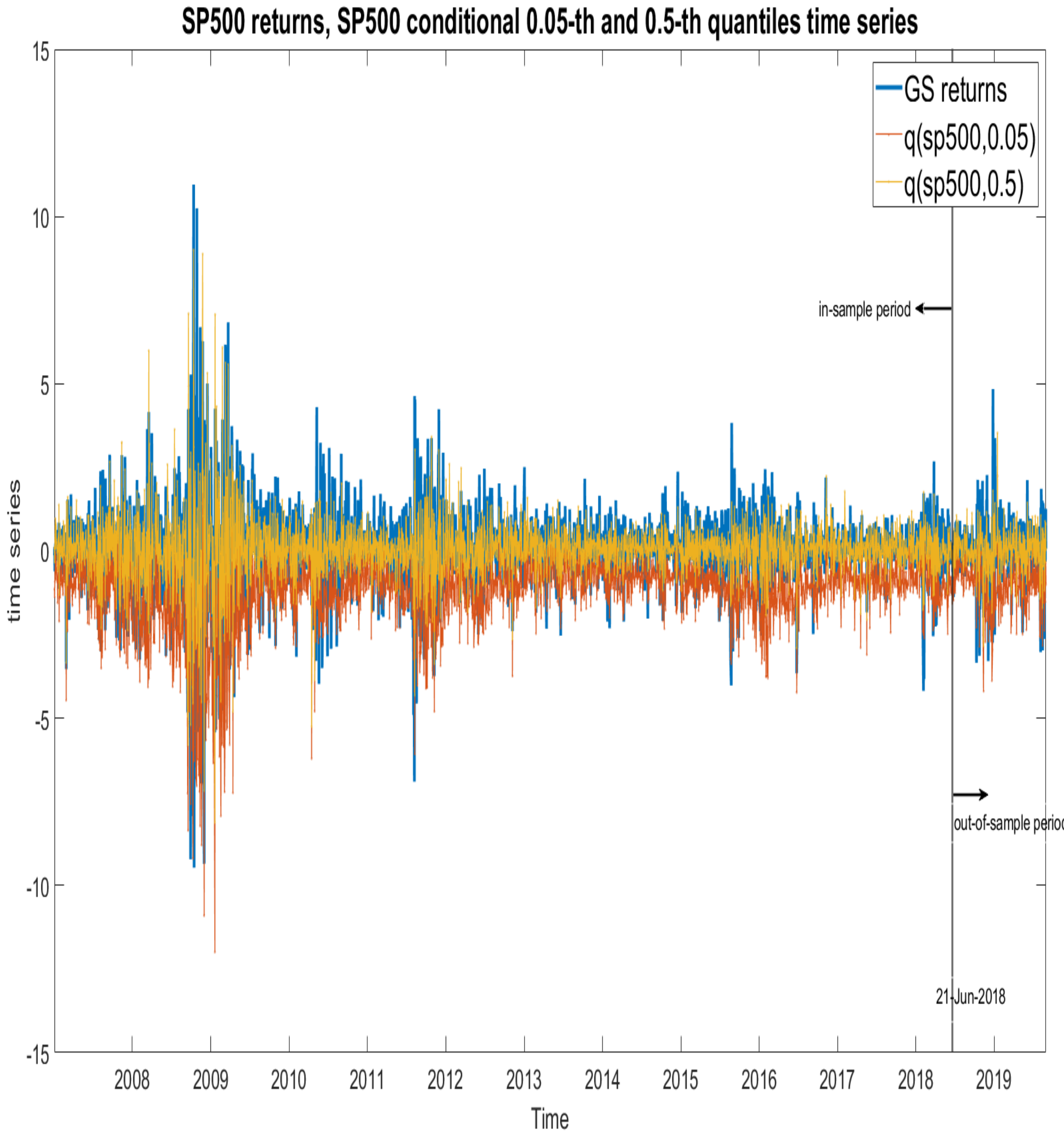

Figure 9: time series plot of the returns of the SP500 with its fitted conditional 0.05-th and 0.5-th quantiles by the systemic MVMQ CAViaR regression with Goldman Sachs

B Appendix: Proofs

Proof of Theorem 2. The proof builds on Engle and Manganelli (2004)'s Thoerem 1 and White et al. (2015)'s Theorem 1, and can be obtained immediately by following the proof of White et al. (2015)'s Theorem 1 on pp. 184. ■

Proof of Theorem 3. The proof builds on Engle and Manganelli (2004)'s Thoerem 2 and White et al. (2015)'s Theorem 2, and can be obtained by following the proof of White et al. (2015)'s Theorem 2 on pp. 186. ■

Proof of Theorem 4. The proof builds on Engle and Manganelli (2004)'s Thoerem 3 and White et al. (2015)'s Theorem 3, and can be obtained immediately by following the proof of White et al. (2015)'s Theorem 3 on pp. 186. ■

Proof of Theorem 5.
Since $\{(\boldsymbol{\alpha}_z - \widehat{\boldsymbol{\alpha}})\}_{z=1}^{N}$ is i.i.d in $N(\mathbf{0}, \boldsymbol{I}_{l_s \times l_s})$, we can get that for each $t \in \{1, \ldots, T\}$,

$$\left\{ \frac{\mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) + \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\}} - \mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \right\}_{z=1}^{N}$$

is a sequence of independent random variables with finite second moments by Assumption 5(iii) of White et al. (2015) that

$$\begin{cases} D_{1,t} := \max\limits_{i=1,\ldots,n} \max\limits_{j=1,\ldots,p} \max\limits_{s=1,\ldots,l_s} \sup\limits_{\alpha \in \Theta} \left| \dfrac{\partial q_{i,j,t}(\boldsymbol{\alpha})}{\partial \alpha_s} \right|, \\ \mathbb{E}[D_{1,t}] < \infty, \\ \mathbb{E}[D_{1,t}^2] < \infty. \end{cases} \tag{48}$$

Hence, we can use the Law of Large Number to get that

$$\widehat{f}_{i,j,t}(0) \xrightarrow{p} \mathbb{E}\left[\frac{\mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) + \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\}} - \mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \middle| \mathcal{F}_{t-1} \right],$$

as $N \to \infty$. Denote $F_{i,t}(\cdot)$ and $f_{i,t}(\cdot)$ as the probability distribution function and the probability density function of $y_{i,t}$ conditional on $\mathcal{F}_{t,i-1}$ respectively. By Assumption 2(i) and 3(ii) of White et al. (2015) that $F_{i,t}(\cdot)$ and $f_{i,t}(\cdot)$ are continuous in $\mathbb{R}$ and $q_{i,j,t}(\cdot)$ is continuously differential on Θ with the conditional probability density of $y_{i,t}$ at its conditional $\theta_{i,j}$-th quantile $q_{i,j,t}(\boldsymbol{\alpha}^o)$ being $f_{i,j,t}(0)$, we can get that

$$\begin{aligned} &\mathbb{E}\left[\frac{\mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) + \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\}} - \mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \middle| \mathcal{F}_{t-1} \right] \\ &= \mathbb{E}\left[\mathbb{E}\left[\frac{\mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) + \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\}} - \mathbf{I}_{\{y_{i,t} \le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \middle| \boldsymbol{\alpha}_i, \mathcal{F}_{t-1} \right] \middle| \mathcal{F}_{t-1} \right] \\ &= \mathbb{E}\left[\frac{F_{i,t}\left(q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) + \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\right) - F_{i,t}\left(q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\right)}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \middle| \mathcal{F}_{t-1} \right] \\ &= \mathbb{E}\left[\frac{f_{i,t}(q_{i,j,t}(\widehat{\boldsymbol{\alpha}}))\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}}) + \mathcal{O}_p\left((\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})' \nabla q_{i,j,t}(\widehat{\boldsymbol{\alpha}}) \nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\right)}{\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\,(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})} \middle| \mathcal{F}_{t-1} \right] \\ &= \mathbb{E}\left[f_{i,t}(q_{i,j,t}(\widehat{\boldsymbol{\alpha}})) + \mathcal{O}_p\left(\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\right) | \mathcal{F}_{t-1} \right] \\ &= \mathbb{E}\left[f_{i,j,t}(0) + \mathcal{O}_p\left(\nabla' q_{i,j,t}(\widehat{\boldsymbol{\alpha}})(\boldsymbol{\alpha}_i - \widehat{\boldsymbol{\alpha}})\right) | \mathcal{F}_{t-1} \right] \\ &= f_{i,j,t}(0), \end{aligned} \tag{49}$$

where the second to last line is obtained by applying the mean value theorem, and the last line is obtained by applying the dominated convergence theorem as $\mathbb{E}[D_{1,t}] < \infty$.

Therefore, we have that $\widehat{f}_{i,j,t}(0) \xrightarrow{p} f_{i,j,t}(0)$ for $t \in \{1, 2, \ldots, T\}$ as $N \to \infty$, and conclude this proof. ■

Proof of Theorem 6. Denote that

$$\bar{Q}_T := \frac{1}{T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \widehat{f}_{i,j,t}(0) \nabla q_{i,j,t}(\boldsymbol{\alpha}^o) \nabla' q_{i,j,t}(\boldsymbol{\alpha}^o),. \tag{50}$$

Note that

$$\widehat{Q}_T - Q = \widehat{Q}_T - \bar{Q}_T + \bar{Q}_T - Q. \tag{51}$$

It is straightforward to get that

$$\widehat{Q}_T - \bar{Q}_T = o_p\left(1\right), \tag{52}$$

because that $\widehat{f}_{i,j,t}(0)$ converges in probability to $f_{i,j,t}(0)$, i.e, $\widehat{f}_{i,j,t}(0) - f_{i,j,t}(0) = o_p\left(1\right)$ for $t \in 1, 2, \ldots, T$ which is proved in Theorem 5, and by Assumption 5(iii) of White et al. (2015) as shown in (48).
And we can get that

$$\bar{Q}_T - Q = o_p\left(1\right), \tag{53}$$

since that $\widehat{V}_T \xrightarrow{p} V$ in Theorem 4 and $q_{i,j,t}(\cdot)$ is bounded so as for $f_{i,t}(q_{i,j,t}(\cdot))$.

Therefore, we have that $\widehat{Q}_T - Q = o_p\left(1\right)$ and conclude this proof. ■

Proof of Theorem 7. To prove this theorem, we need to find out the limiting distribution of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta_{i,j}\right)$. First we define

$$\begin{cases} \mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}) := \mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\boldsymbol{\alpha})\}} - \theta_{i,j}, \\ \mathbf{Hit}_{i,t}(\boldsymbol{\alpha}) := [\mathrm{Hit}_{i,1,t}(\boldsymbol{\alpha}), \ldots, \mathrm{Hit}_{i,p,t}(\boldsymbol{\alpha})]', \\ \mathbf{Hit}_{t}(\boldsymbol{\alpha}) := [\mathbf{Hit}'_{1,t}(\boldsymbol{\alpha}), \ldots, \mathbf{Hit}'_{n,t}(\boldsymbol{\alpha})]', \\ \boldsymbol{q}_{i,t}(\boldsymbol{\alpha}) := [q_{i,1,t}(\boldsymbol{\alpha}), \ldots, q_{i,p,t}(\boldsymbol{\alpha})]', \\ \boldsymbol{q}_{t}(\boldsymbol{\alpha}) := [\boldsymbol{q}'_{1,t}(\boldsymbol{\alpha}), \ldots, \boldsymbol{q}'_{n,t}(\boldsymbol{\alpha})]', \end{cases} \tag{54}$$

and start to derive the the limiting behaviour of $\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta_{i,j}\right)$ as follows:

$$\begin{aligned} &\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta_{i,j}\right) \\ &\quad = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathrm{Hit}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) - \mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o) + \mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)\right) \\ &\quad = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathrm{Hit}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) - \mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)\right) + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o). \end{aligned} \tag{55}$$

Use the result proved by Engle and Manganelli (2004) that

$$\begin{cases} \mathrm{Hit}^{\oplus}_{i,j,t}(\boldsymbol{\alpha}) := \left(1 + exp\left(\dfrac{y_{i,t} - q_{i,j,t}(\boldsymbol{\alpha})}{c_T}\right)\right)^{-1} - \theta_{i,j}, \\ \dfrac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}^{\oplus}_{i,j,t}(\boldsymbol{\alpha}^o) - \dfrac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o) = o_p(1), \\ \dfrac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}^{\oplus}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) - \dfrac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) = o_p(1), \end{cases} \tag{56}$$

where c_T is a nonstochastic sequence such that $\lim_{T\to\infty} c_T = 0$. We can approximate the first term in the last line of (55) as follows:

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathrm{Hit}_{i,j,t}(\widehat{\boldsymbol{\alpha}})-\mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)\right)\\
&=\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathrm{Hit}^{\oplus}_{i,j,t}(\widehat{\boldsymbol{\alpha}})-\mathrm{Hit}^{\oplus}_{i,j,t}(\boldsymbol{\alpha}^o)\right)+o_p(1)\\
&=\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t} f_{i,j,t}(0)\frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^o)+o_p(1) \qquad (57)\\
&=\sqrt{T}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^o)'\,\frac{1}{T}\sum_{t=1}^{T} y_{i_c,t} f_{i,j,t}(0)\frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}+o_p(1)\\
&=-\frac{1}{\sqrt{T}}\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbf{Hit}'_t(\boldsymbol{\alpha}^o)\frac{\partial \boldsymbol{q}_t(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}\right)Q^{-1}G+o_p(1),
\end{aligned}
$$

where $G := \frac{1}{T}\sum_{t=1}^{T} y_{i_c,t} f_{i,j,t}(0)\frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}$, and the third and last lines are obtained by respectively applying (B.5) of Engle and Manganelli (2004) on pp. 379 and

$$
\begin{aligned}
\sqrt{T}(\widehat{\boldsymbol{\alpha}}-\boldsymbol{\alpha}^o)' &= -\frac{1}{\sqrt{T}}\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbf{Hit}'_t(\boldsymbol{\alpha}^o)\frac{\partial \boldsymbol{q}_t(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}\right)Q^{-1}+o_p(1)\\
&\overset{\mathcal{D}}{\sim} N(\mathbf{0}, Q^{-1}VQ^{-1})
\end{aligned} \qquad (58)
$$

which is in the proof of White et al. (2015) on pp.187. So we substitute(57) back into (55) and get

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}-\theta_{i,j}\right)\\
&\quad=-\frac{1}{\sqrt{T}}\left(\sum_{t=1}^{T}\sum_{i=1}^{n}\mathbf{Hit}'_t(\boldsymbol{\alpha}^o)\frac{\partial \boldsymbol{q}_t(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}}\right)Q^{-1}G+\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\mathrm{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)+o_p(1).
\end{aligned} \qquad (59)
$$

Apply Assumption 5(i)-(iii) of White et al. (2015), the ergodic theorem and the martingale difference central limit theorem (see Theorem 3.35 and 5.24 of White (2001)) on (59) and obtain that

$$
\begin{aligned}
&\frac{1}{\sqrt{T}}\sum_{t=1}^{T} y_{i_c,t}\left(\mathbf{I}_{\{y_{i,t}\le q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}}-\theta_{i,j}\right)\\
&\quad\overset{\mathcal{D}}{\sim} N(\mathbf{0}, \frac{1}{T}\theta_{i,j}(1-\theta_{i,j}))\left[G'Q^{-1'}\sum_{t=1}^{T}\sum_{i\neq i_c}^{N}\frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\boldsymbol{\alpha}^{o}}\frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\boldsymbol{\alpha}^{o\prime}}Q^{-1}G\right].
\end{aligned} \qquad (60)
$$

And we know that

$$
\begin{aligned}
\widehat{Q}_T &\xrightarrow{p} Q,\\
\widehat{f}_{i,j,t}(0)-f_{i,j,t}(0) &= o_p(1),\\
\frac{\partial q_{i,j,t}(\widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}'} &\xrightarrow{p} \frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}},
\end{aligned} \qquad (61)
$$

where $\widehat{Q}_T$ is the estimator given in Theorem 6, $\widehat{f}_{i,j,t}(0)$ is the estimator given in Theorem 5, and the last equality is obtained because that $\frac{\partial q_{i,j,t}(\boldsymbol{\alpha})}{\partial \boldsymbol{\alpha}}$ is continuous in Θ and $\widehat{\boldsymbol{\alpha}} \xrightarrow{p} \boldsymbol{\alpha}^o$. So we also have that

$$
\widehat{G}_T := \frac{1}{T}\sum_{t=1}^{T} y_{i_c,t}\widehat{f}_{i,j,t}(0)\frac{\partial q_{i,j,t}(\widehat{\boldsymbol{\alpha}})}{\partial \widehat{\boldsymbol{\alpha}}'} \xrightarrow{p} G, \qquad (62)
$$

and

$$\text{DQ}_{\text{IS}} \overset{\mathcal{D}}{\sim} \chi^2(1),$$

which concludes this proof. ■

Proof of Theorem 8. As in the proof of Theorem 7, we apply the approximation result (56) to derive the limiting distribution of $\frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \left(\mathbf{I}_{\{y_{i,t} \leq q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta\right)$ as follows:

$$\begin{aligned}
&\lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \left(\mathbf{I}_{\{y_{i,t} \leq q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta\right) \\
&\quad = \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \left(\text{Hit}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) - \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o) + \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)\right) \\
&\quad = \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \left(\text{Hit}^{\oplus}_{i,j,t}(\widehat{\boldsymbol{\alpha}}) - \text{Hit}^{\oplus}_{i,j,t}(\boldsymbol{\alpha}^o) + \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o)\right) \\
&\quad = \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} f_{i,j,t}(0) \frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^{o\prime}} (\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^o) + \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o), \\
&\quad = \lim_{R\to\infty} \sqrt{\frac{N_R}{T_R}} \sqrt{T_R} (\widehat{\boldsymbol{\alpha}} - \boldsymbol{\alpha}^o)' \frac{1}{N_R} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} f_{i,j,t}(0) \frac{\partial q_{i,j,t}(\boldsymbol{\alpha}^o)}{\partial \boldsymbol{\alpha}^o} + \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o), \\
&\quad = \lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \text{Hit}_{i,j,t}(\boldsymbol{\alpha}^o).
\end{aligned} \tag{63}$$

Apply Assumption 5(i)-(iii) of White et al. (2015), the ergodic theorem and the martingale difference central limit theorem (see Theorem 3.35 and 5.24 of White (2001)) on (63) and obtain that

$$\begin{aligned}
&\lim_{R\to\infty} \frac{1}{\sqrt{N_R}} \sum_{t=T_R+1}^{T_R+N_R} y_{i_c,t} \left(\mathbf{I}_{\{y_{i,t} \leq q_{i,j,t}(\widehat{\boldsymbol{\alpha}})\}} - \theta\right) \\
&\overset{\mathcal{D}}{\sim} N\left(\mathbf{0}, \frac{\theta_{i,j}(1-\theta_{i,j})}{N_R} \sum_{t=1}^{N_R} y^2_{i_c,t}\right)
\end{aligned} \tag{64}$$

which leads to

$$\text{DQ}_{\text{OOS}} \overset{\mathcal{D}}{\sim} \chi^2(1), \tag{65}$$

and concludes this proof. ■