VAR for VaR: Measuring Systemic Risk Using Multivariate Regression Quantiles

Halbert White† Tae-Hwan Kim‡ Simone Manganelli§

December 29, 2010

Abstract

This paper proposes methods for estimation and inference in multivariate, multi-quantile models. The theory can simultaneously accommodate models with multiple random variables, multiple confidence levels, and multiple lags of the associated quantiles. The proposed framework can be conveniently thought of as a vector autoregressive (VAR) extension to quantile models. We estimate a simple version of the model using market returns data to analyse spillovers in the values at risk (VaR) of different financial institutions. We construct impulse-response functions for the quantile processes of a sample of 230 financial institutions around the world and study how financial institution-specific and system-wide shocks are absorbed by the system.

Keywords: Quantile impulse-responses, spillover, codependence, CAViaR

JEL classification: C13, C14, C32.

*We would like to thank Francesca Fabbri and Thomas Kostka for excellent research assistance. The phrase "VAR for VaR" was first used by Andersen et al. (2003), in the title of their section 6.4. The views expressed in this paper are those of the author and do not necessarily reflect those of the European Central Bank or the Eurosystem.

†Department of Economics, University of California, San Diego.
‡School of Economics, Yonsei University.
§Corresponding author. European Central Bank, DG-Research, Kaiserstrasse 29, 60433 Frankfurt am Main, Germany. Email: simone.manganelli@ecb.int.
1 Introduction

The recent financial crisis has brought to the forefront the importance of having sound measures of financial spillover. In the current debate, great emphasis has been put on how to measure whether an institution is of systemic importance. In particular, it has been argued that since the failure of a systemically important financial institutions could produce severe negative externalities on the whole financial system, the supervision of financial institutions should, among other things, take into account the spillover of risks within the system. The regulatory constraints imposed on firms should therefore also reflect their overall systemic impact.

A popular measure to assess the systemic importance of a financial institution is to look at the sensitivity of its Value at Risk (VaR) to shocks to the whole financial system (see, for instance, Adrian and Brunnermeier 2009, Acharya et al. 2009, Engle and Brownlees 2010). This paper proposes a novel method to estimate the sensitivity of VaR of a given financial institution to system-wide shocks and, vice versa, the sensitivity of the financial system VaR to shocks to individual financial institutions. We develop the econometric framework to estimate and make inferences in a “VAR for VaR” model, that is, a vector autoregressive (VAR) model where the dependent variables are the VaR of financial institutions, which depend on (lagged) VaR and past shocks. This allows us to study the spillover and feedback effects among the variables of the system. In addition, from the estimated parameters, we can compute the long run VaR equilibria, as well as impulse-response functions.

To illustrate our approach and its usefulness, consider a simple set-up with two financial institutions. Let $Y_{1t}$ and $Y_{2t}$ denote the returns at time $t$ for the two institutions. All information available in both markets at time $t$ is represented by the information set $\mathcal{F}_t$. As is standard, define VaR as the worst monetary loss over
a relevant holding period and with a given level of confidence $\theta \in (0, 1)$. Assuming that the total money value of each market is $1$, $VaR_{it}$ for market $i = 1, 2$ at time $t$ is

$$Pr[Y_{it} \leq -VaR_{it}|F_{t-1}] = \theta. \quad (1)$$

Hence, $-VaR_{it}$ is the $\theta^{th}$ quantile of $Y_{it}$ conditional on $F_{t-1}$; we will find it convenient to denote this as $q_{i,t}^\theta$ in the analysis to follow.

A simple version of our proposed structure relates Value at Risk in two country-wide markets according to

$$VaR_{1t} = X'_t \beta_{1}^* + b_{11}^* VaR_{1t-1} + b_{12}^* VaR_{2t-1}$$

$$VaR_{2t} = X'_t \beta_{2}^* + b_{21}^* VaR_{1t-1} + b_{22}^* VaR_{2t-1},$$

where $X_t$ represents predictors belonging to $F_{t-1}$. The codependence between the two markets is measured by the off-diagonal coefficients $b_{12}^*$ and $b_{21}^*$, and the hypothesis of no codependence can be tested by testing $H_0 : b_{12}^* = b_{21}^* = 0$. The direction of risk spread also can be detected by examining these two coefficients. For example, if $b_{12}^* = 0$ and $b_{21}^* \neq 0$, then the direction of risk spread is from country 1 to country 2, not the other way around. Our fully general model, explained in the next sections, is much richer than the above in that, among other things: (i) we can accommodate more than two markets; (ii) multiple lags of $VaR_{it}$ can be included; and (iii) we can simultaneously consider multiple confidence levels, say $(\theta_1, ..., \theta_p)$.

In our empirical analysis, we estimate the VAR for VaR model using returns of individual financial institutions from around the world and a global financial sector index. By constructing the impulse-response functions, we can assess the financial institutions’ resilience to shocks to the overall index and the impact they have on the VaR of the financial sector index.
The plan of the paper is as follows. In Section 2, we set forth the multivariate multi-quantile CAViaR (MVMQ-CAViaR) framework, a generalization of White, Kim, and Manganelli’s (2008) MQ-CAViaR extension of Engle and Manganelli’s original CAViaR (2004) framework. Section 3 provides conditions ensuring the consistency and asymptotic normality of the MVMQ-CAViaR estimator, as well as results providing a consistent asymptotic covariance matrix estimator. Section 4 contains our empirical study. Section 5 provides a summary and concluding remarks. The appendix contains the remaining regularity conditions and the proofs of the theorems in the text.

2 The MVMQ-CAViaR Process and Model

We consider data generated as a realization of the following stochastic process.

**Assumption 1** The sequence \( \{(Y_t', X_t') : t = 0, \pm 1, \pm 2, \ldots, \} \) is a stationary and ergodic stochastic process on the complete probability space \((\Omega, \mathcal{F}, P_0)\), where \( Y_t \) is a finitely dimensioned \( n \times 1 \) vector and \( X_t \) is a countably dimensioned vector whose first element is one.

Let \( \mathcal{F}_{t-1} \) be the \( \sigma \)-algebra generated by \( Z_t^{-1} := \{X_t, (Y_{t-1}, X_{t-1}), \ldots\} \), i.e. \( \mathcal{F}_{t-1} := \sigma(Z_t^{-1}) \). For \( i = 1, \ldots, n \), we let \( F_{it}(y) := P_0[Y_{it} < y \mid \mathcal{F}_{t-1}] \) define the cumulative distribution function (CDF) of \( Y_{it} \) conditional on \( \mathcal{F}_{t-1} \).

Let \( 0 < \theta_{it} < \ldots < \theta_{ip} < 1 \). For \( j = 1, \ldots, p \), the \( \theta_{ij} \)th quantile of \( Y_{it} \) conditional on \( \mathcal{F}_{t-1} \), denoted \( q_{i,j,t}^* \), is

\[
q_{i,j,t}^* := \inf\{y : F_{it}(y) = \theta_{ij}\},
\]

and if \( F_{it} \) is strictly increasing,

\[
q_{i,j,t}^* = F_{it}^{-1}(\theta_{ij}).
\]
Alternatively, $q_{i,j,t}^*$ can be represented as

$$
\int_{-\infty}^{q_{i,j,t}^*} dF_{it}(y) = E[1_{Y_{it} \leq q_{i,j,t}^*} \mid \mathcal{F}_{t-1}] = \theta_{ij},
$$

where $dF_{it}(\cdot)$ is the Lebesgue-Stieltjes probability density function (PDF) of $Y_{it}$ conditional on $\mathcal{F}_{t-1}$, corresponding to $F_{it}$. Note that we specify the same number ($p$) of quantile indexes for each $i = 1, \ldots, n$; however, this is just for notational simplicity. Our theory easily applies to the case in which the number of quantile indexes differs with $i$; i.e., $p$ can be replaced with $p_i$.

Our objective is to jointly estimate the conditional quantile functions $q_{i,j,t}^*$, $j = 1, 2, \ldots, p$, $i = 1, \ldots, n$. For this we write $q^*_t := (q^*_{1,t}, q^*_{2,t}, \ldots, q^*_{n,t})'$ with $q^*_{i,t} := (q^*_{i,1,t}, q^*_{i,2,t}, \ldots, q^*_{i,p,t})'$ and impose additional appropriate structure.

First, we ensure that the conditional distributions of $Y_{it}$ are everywhere continuous, with positive density at each conditional quantile of interest, $q^*_{i,j,t}$. We let $f_{it}$ denote the conditional probability density function (PDF) corresponding to $F_{it}$. In stating our next condition (and where helpful elsewhere), we make explicit the dependence of the conditional CDF $F_{it}$ on $\omega \in \Omega$ by writing $F_{it}(\omega, y)$ in place of $F_{it}(y)$. Similarly, we may write $f_{i,t}(\omega, y)$ in place of $f_{i,t}(y)$. Realized values of the conditional quantiles are correspondingly denoted $q_{i,j,t}^*(\omega)$.

Our next assumption ensures the desired continuity and imposes specific structure on the quantiles of interest.

**Assumption 2** For $i = 1, \ldots, n$, (i) $Y_{it}$ is continuously distributed such that for each $\omega \in \Omega$, $F_{it}(\omega, \cdot)$ and $f_{it}(\omega, \cdot)$ are continuous on $\mathbb{R}$, $t = 1, 2, \ldots$; (ii) For given $0 < \theta_1 < \ldots < \theta_p < 1$ and $\{q_{i,j,t}^*\}$ as defined above, suppose: (a) For each $i, j, t$, and $\omega$, $f_{it}(\omega, q_{i,j,t}^*(\omega)) > 0$; and (b) For given finite integers $k$ and $m$, there exist a stationary ergodic sequence of random $k \times 1$ vectors $\{\Psi_t\}$, with $\Psi_t$ measurable $- \mathcal{F}_{t-1}$, and real vectors $\beta_{ij}^* := (\beta_{i,j,1}^*, \ldots, \beta_{i,j,k}^*)'$ and $\gamma_{i,j,t}^* := (\gamma_{i,j,t,1}^*, \ldots, \gamma_{i,j,t,n}^*)'$, where each
\( \gamma^*_{i,j,\tau,k} \) is a \( p \times 1 \) vector, such that for \( j = 1, \ldots, p \), \( i = 1, \ldots, n \), and all \( t \),

\[
q^*_{i,j,t} = \Psi_t \beta^*_{ij} + \sum_{\tau=1}^{m} q^*_{t-\tau} \gamma^*_{i,j,\tau}. \tag{4}
\]

The structure of eq. (4) is a multivariate version of the MQ-CAViaR process of White, Kim, and Manganelli (2008), itself a multi-quantile version of the CAViaR process introduced by Engle and Manganelli (2004). Under suitable restrictions on the \( \gamma^*_{i,j,\tau} \)'s, we obtain as special cases (1) separate MQ-CAViaR processes for each element of \( Y_t \); (2) standard (single quantile) CAViaR processes for each element of \( Y_t \); or (3) multivariate CAViaR processes, in which a single quantile of each element of \( Y_t \) is related dynamically to single quantiles of the (lags of) other elements of \( Y_t \). Thus, we call processes satisfying our structure “Multivariate MQ-CAViaR” (MVMQ-CAViaR) processes.

For MVMQ-CAViaR, the number of relevant lags can differ across the elements of \( Y_t \) and the conditional quantiles; this is reflected in the possibility that for given \( j \), elements of \( \gamma^*_{i,j,\tau} \) may be zero for values of \( \tau \) greater than some given integer. For notational simplicity, we do not represent \( \mu \) as depending on \( i \) or \( j \). Nevertheless, by convention, for no \( \tau \leq m \) does \( \gamma^*_{i,j,\tau} \) equal the zero vector for all \( i \) and \( j \).

The finitely dimensioned random vectors \( \Psi_t \) may contain lagged values of \( Y_t \), as well as measurable functions of \( X_t \) and lagged \( X_t \) or \( Y_t \). In particular, \( \Psi_t \) may contain Stinchcombe and White’s (1998) GCR transformations, as discussed in White (2006).

For a particular quantile, say \( \theta_{ij} \), the coefficients to be estimated are \( \beta^*_{ij} \) and \( \gamma^*_{ij} := (\gamma^*_{i1,j,1}, \ldots, \gamma^*_{i,n,j,m})' \). Let \( \alpha^{*'}_{ij} := (\beta^*_{ij}, \gamma^*_{ij}) \), and write \( \alpha^{*'} = (\alpha^{*'}_{11}, \ldots, \alpha^{*'}_{1p}, \ldots, \alpha^{*'}_{m1}, \ldots, \alpha^{*'}_{np})' \), an \( \ell \times 1 \) vector, where \( \ell := np(k + npm) \). We call \( \alpha^{*'} \) the "MVMQ-CAViaR coefficient vector." We estimate \( \alpha^{*'} \) using a correctly specified model of the MVMQ-CAViaR process.

First, we specify our model.
Assumption 3 Let $A$ be a compact subset of $\mathbb{R}^\ell$. For $i = 1, ..., n$, and $j = 1, ..., p$, (i) the sequence of functions $\{q_{i,j,t} : \Omega \times A \to \mathbb{R}^p\}$ is such that for each $t$ and each $\alpha \in A$, $q_{i,j,t}(\cdot, \alpha)$ is measurable--$\mathcal{F}_{t-1}$; for each $t$ and each $\omega \in \Omega$, $q_{i,j,t}(\omega, \cdot)$ is continuous on $A$; and for each $i, j$, and $t$,

$$q_{i,j,t}(\cdot, \alpha) = \Psi_t' \beta_{ij} + \sum_{\tau=1}^{m} q_{t-\tau}(\cdot, \alpha)' \gamma_{i,j,\tau}.$$ 

Next, we impose correct specification and an identification condition. Assumption 4(i.a) delivers correct specification by ensuring that the MVMQ-CAViaR coefficient vector $\alpha^*$ belongs to the parameter space, $A$. This ensures that $\alpha^*$ optimizes the estimation objective function asymptotically. Assumption 4(i.b) delivers identification by ensuring that $\alpha^*$ is the only such optimizer. In stating the identification condition, we define $\delta_{i,j,t}(\alpha, \alpha^*) := q_{i,j,t}(\cdot, \alpha) - q_{i,j,t}(\cdot, \alpha^*)$ and use the norm $||\alpha|| := \max_{s=1,...,\ell} |\alpha_s|$, where for convenience we also write $\alpha = (\alpha_1, ..., \alpha_{\ell})'$.

Assumption 4 (i)(a) There exists $\alpha^* \in A$ such that for all $i, j, t$

$$q_{i,j,t}(\cdot, \alpha^*) = q_{i,j,t}^*; \quad (5)$$

(b) There is a non-empty index set $\mathcal{I} \subseteq \{(1, 1), ..., (1, p), ..., (n, 1), ..., (n, p)\}$ such that for each $\epsilon > 0$ there exists $\delta_\epsilon > 0$ such that for all $\alpha \in A$ with $||\alpha - \alpha^*|| > \epsilon$,

$$P[\cup_{(i,j) \in \mathcal{I}} \{|\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon\}] > 0.$$ 

Among other things, this identification condition ensures that there is sufficient variation in the shape of the conditional distribution to support estimation of a sufficient number ($\#\mathcal{I}$) of variation-free conditional quantiles. As in the case of MQ-CAViaR, distributions that depend on a given finite number of variation-free parameters, say $r$, will generally be able to support $r$ variation-free quantiles. For
example, the quantiles of the \( N(\mu, 1) \) distribution all depend on \( \mu \) alone, so there is only one "degree of freedom" for the quantile variation. Similarly the quantiles of scaled and shifted \( t \)–distributions depend on three parameters (location, scale, and kurtosis), so there are only three "degrees of freedom" for the quantile variation.

### 3 MVMQ-CAViaR: Asymptotic Theory

We estimate \( \alpha^* \) by the method of quasi-maximum likelihood. Specifically, we construct a quasi-maximum likelihood estimator (QMLE) \( \hat{\alpha}_T \) as the solution to the optimization problem

\[
\min_{\alpha \in \mathcal{K}} S_T(\alpha) := T^{-1} \sum_{t=1}^{T} \left\{ \sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha)) \right\},
\]

where \( \rho_{\theta}(e) = e \psi_{\theta}(e) \) is the standard "check function," defined using the usual quantile step function, \( \psi_{\theta}(e) = \theta - 1_{e \leq 0} \).

We thus view

\[
S_t(\alpha) := -\left\{ \sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha)) \right\}
\]

as the quasi log-likelihood for observation \( t \). In particular, \( S_t(\alpha) \) is the log-likelihood of a vector of \( np \) independent asymmetric double exponential random variables (see White, 1994, ch. 5.3; Kim and White, 2003; Komunjer, 2005). Because \( Y_{it} - q_{i,j,t}(\cdot, \alpha) \) does not need to actually have this distribution, the method is quasi maximum likelihood.

We establish consistency and asymptotic normality for \( \hat{\alpha}_T \) by methods analogous to those of White, Kim, and Manganelli (2008). For conciseness, we place the remaining regularity conditions and technical discussion in the appendix.

**Theorem 1** Suppose that Assumptions 1, 2(i, ii), 3(i), 4(i), and 5(i, ii) hold. Then
With \( Q^* \) and \( V^* \) as given below, the asymptotic normality result is

**Theorem 2** Suppose that Assumptions 1-6 hold. Then

\[
T^{1/2}(\hat{\alpha}_T - \alpha^*) \xrightarrow{d} N(0, Q^* V^* Q^*-1).
\]

To test restrictions on \( \alpha^* \) or to obtain confidence intervals, we require a consistent estimator of the asymptotic covariance matrix \( C^* := Q^* V^* Q^*-1 \). First, we provide a consistent estimator \( \hat{V}_T \) for \( V^* \); then we give a consistent estimator \( \hat{Q}_T \) for \( Q^* \). It follows that \( \hat{Q}_T := \hat{Q}_T^{-1} \hat{V}_T \hat{Q}_T^{-1} \) is a consistent estimator for \( C^* \).

We have \( V^* := E(\eta_t^* \eta_t^*) \) with \( \eta_t^* := \sum_{i=1}^n \sum_{j=1}^p \nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{\theta_{ij}}(\epsilon_{i,j,t}) \), where \( \psi_{\theta_{ij}}(\epsilon_{i,j,t}) \equiv \theta_{ij} - 1_{[\epsilon_{i,j,t} \leq 0]} \) is a generalized residual. A straightforward plug-in estimator of \( V^* \) is

\[
\hat{V}_T := T^{-1} \sum_{t=1}^T \hat{\eta}_t \hat{\eta}^\prime_t, \quad \text{with}
\]

\[
\hat{\eta}_t := \sum_{i=1}^n \sum_{j=1}^p \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \psi_{\theta_{ij}}(\hat{\epsilon}_{i,j,t}),
\]

\[
\hat{\epsilon}_{i,j,t} := \hat{Y}_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T).
\]

The next result establishes the consistency of \( \hat{V}_T \) for \( V^* \).

**Theorem 3** Suppose that Assumptions 1-6 hold. Then \( \hat{V}_T \xrightarrow{p} V^* \).

Next, we provide a consistent estimator of

\[
Q^* := \sum_{i=1}^n \sum_{j=1}^p E[f_{i,j,t}(0) \nabla q_{i,j,t}(\cdot, \alpha^*) \nabla q_{i,j,t}(\cdot, \alpha^*)].
\]
We follow Powell’s (1984) suggestion of estimating \( f_{i,j,t}(0) \) with \( 1_{[-\hat{\epsilon}_T \leq \hat{\epsilon}_{i,j,t} \leq \hat{\epsilon}_T]} / 2\hat{\epsilon}_T \) for a suitably chosen sequence \( \{\hat{\epsilon}_T\} \). This is also the approach taken in Kim and White (2003), Engle and Manganelli (2004), and White, Kim, and Manganelli (2008). Accordingly, our proposed estimator is

\[
\hat{Q}_T = (2\hat{\epsilon}_T T)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{j=1}^{P} 1_{[-\hat{\epsilon}_T \leq \hat{\epsilon}_{i,j,t} \leq \hat{\epsilon}_T]} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T).
\]

**Theorem 4** Suppose that Assumptions 1-7 hold. Then \( \hat{Q}_T \xrightarrow{p} Q^* \).

There is no guarantee that \( \hat{\alpha}_T \) is asymptotically efficient. There is now a considerable literature investigating efficient estimation in quantile models; see, for example, Newey and Powell (1990), Otsu (2003), Komunjer and Vuong (2006, 2007a, 2007b). So far, this literature has only considered single quantile models. It is not obvious how the results for single quantile models extend to multivariate multi-quantile models. Nevertheless, Komunjer and Vuong (2007a) show that the class of QML estimators is not large enough to include an efficient estimator, and that the class of \( M \)-estimators, which strictly includes the QMLE class, yields an estimator that attains the efficiency bound. Specifically, when \( p = n = 1 \), they show that replacing the usual quantile check function \( \rho_{\theta_{ij}}(\cdot) \) in eq.(6) with

\[
\rho^*_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha)) = (\theta_{ij} - 1|Y_{it} - q_{i,j,t}(\cdot, \alpha)\leq 0)(F_{it}(Y_{it}) - F_{it}(q_{i,j,t}(\cdot, \alpha)))
\]

will deliver an asymptotically efficient quantile estimator. We conjecture that replacing \( \rho_{\theta_{ij}} \) with \( \rho^*_{\theta_{ij}} \) in eq.(6) will improve estimator efficiency for \( p \) and/or \( n \) not equal to 1. We leave the study of the asymptotically efficient multivariate multi-quantile estimator for future work.
4 Assessing the Systemic Importance of Financial Institutions

We apply the model developed in the previous sections to study spillovers in the returns quantiles of a sample of 230 financial institutions from around the world. In this section we first show how to compute impulse-response functions within the multivariate, multiquantile framework. We choose a particular quantile specification for our empirical analysis, linking it to the DGP of more familiar multivariate mean-variance models. We next describe the data and the optimization strategy. Finally, we present the results of an empirical application to market returns of financial institutions.

4.1 Impulse-response Functions for Multivariate CAViaR

Suppose data are structurally generated as:

\[ Y_t = L_t u_t, \]

where \( Y_t \) is a 2-dimensional vector, \( L_t := L_t(Z^t) \) is an \( \mathcal{F}_{t-1} \)-measurable lower triangular matrix and the elements of \( u_t \) are mutually independent with \( \{ u_t | \mathcal{F}_{t-1} \} \) a martingale difference sequence. By convention, we let the first element of \( Y_t \) denote the per-period return on a financial sector index and the second element the per-period return on a specific financial institution. The identification assumption behind this decomposition is that shocks to the financial sector index are allowed to have a direct impact on the return of the specific financial institution, but shocks to the specific financial institution do not have a direct impact on the financial sector index. Here we limit ourselves to a bivariate system, as we are interested in the interaction between a financial sector index and an individual financial institution. The theoretical framework of this paper can accommodate higher
dimensional models, although at the cost of a rapidly increasing computational burden.

For suitable choices of $L_t$, the conditional quantiles of $Y_t$ obey (4). For our purposes here, suppose

$$q_{t,t} = c + A|Y_{t-1}| + Bq_{t-1}$$

(7)

where $q_{t,t}$, $Y_{t-1}$, and $c$ are 2-dimensional vectors, and $A$, $B$ are $(2,2)$-matrices. See Appendix 1 for an example of how this representation corresponds to a bivariate GARCH model.

Let the long run matrix $L$ be defined as:

$$L := \lim_{t \to \infty} L_t(Z_{t-1})|_{Z_{t-1}=0}.$$ 

If we set $y_{t+n} = 0$ for $n > 0$, the system converges to $\tilde{q} = (I - B)^{-1}c$. Assuming the system is at its long run equilibrium, if we denote by $t^{1}$ a one standard deviation shock to the first element of the orthogonal error $u$ at time $t$, such a shock implies the following quantile response:

$$q_{1,t+1} \equiv c + A|L_{t^{1}}| + B\tilde{q}$$

$$q_{1,t+n} \equiv c + Bq_{1,t+n-1} \quad n > 1$$

The impulse-responses for a shock to the second element of $u$ can be computed analogously.

We can compute four types of impulse-responses:

1. $\partial q_{1,t+n}/\partial u_{1,t}$ is the reaction of the system’s risk to a system shock;

2. $\partial q_{1,t+n}/\partial u_{2,t}$ is the reaction of the system’s risk to an individual bank’s shock;

3. $\partial q_{2,t+n}/\partial u_{1,t}$ is the reaction of the bank’s risk to a system shock;
4. \( \frac{\partial q_{2,t+n}}{\partial u_{2,t}} \) is the reaction of the bank’s risk to its own shock.

4.2 Data and Optimization Strategy

The data were downloaded from Datastream. We considered three main global sub-indices: banks, financial services, and insurances. The sample includes daily closing prices from 1 January 2000 to 6 August 2010. We eliminated all the stocks whose times series started later than 1 January 2000. At the end of this process, we were left with 230 stocks.

Table 1 reports the breakdown of the stocks by sector and by geographic area. There are twice as many financial institutions classified as banks in our sample as there are those classified as financial services or insurances. The distribution across geographic areas is more balanced, with a greater number of EU financial institutions and slightly lower Asian representation.

To cope with asynchronicity issues due to differing time zones, the data were transformed to weekly frequency. Weekly returns were computed as the log difference of weekly closing prices and expressed in percentage terms. The proxy for the overall index used in each bivariate quantile estimation is the World Financials price index, as provided by Datastream.

We estimated 230 bivariate 1% quantile models between the index and each of the 230 financial institutions in our sample. Each model is estimated using as starting values for optimization the univariate CAViaR estimates and initializing the remaining parameters at zero. We also generated 40 additional initial conditions by adding a normally distributed noise to this vector. For each of these 41 initial conditions, we minimized the regression quantile objective function (6) using the \texttt{fminsearch} optimization function in Matlab, which is based on the Nelder-Mead simplex algorithm. Finally, among the resulting 41 vectors of parameter estimates, we chose the vector yielding the lowest value for the function.
We adopt this strategy because we find that parameter estimates are sensitive to the choice of initial conditions (possibly due to a very flat likelihood near the optimum). Such an optimization strategy is more time consuming, but delivers more reliable results. Still, for some time series, either we did not get convergence or the parameter estimates were associated with an explosive impulse-response function. This happened in about 20% of the sample. In this case, we restricted to zero the coefficients associated with the second lagged quantile (i.e. the quantile associated with the single financial institutions) in the process (7).

In calculating the standard errors, we have set the bandwidth to 1.

4.3 Results

Table 2 reports as an example the estimation results for the Sumitomo Mitsui Financial Group, the second largest bank in Japan by market value (as of November 2009). Notice that the non-diagonal coefficients for the $B$ matrix are significantly different from zero, illustrating how the multivariate quantile model can uncover dynamics that cannot be detected by estimating univariate quantile models. In general, we reject the joint null hypothesis that all off-diagonal coefficients of the matrices $A$ and $B$ are equal to zero at the .05 level for about 85% of the firms in our sample. The resulting estimated 1% quantiles for the index and the Sumitomo Mitsui Financial Group are reported in Figure 1. The quantile of the global index is generally much smaller in absolute terms than the quantile of Sumitomo, reflecting the portfolio diversification effect of the index. Only around the Lehman bankruptcy (September 2008) is the situation briefly inverted, with the estimates indicating a higher risk associated with the global index than with the Sumitomo bank.

It should be borne in mind that each of the 230 bivariate models is estimated using a different information set (as the time series of the index and of a different
financial institution is used for each estimation). Therefore, each pair produces a different estimate of the VaR of the index, simply because we condition on a different information set. Moreover, the coefficients and any quantities derived from them, such as impulse responses are information set-specific. This means that naive comparisons across bivariate pairs can be misleading and are generally unwarranted. The proper context for comparing sensitivities and impulse responses is in a multivariate setting using a common information set. Because of the considerable computational challenges involved, we leave this for future study.

Nevertheless, averaging across the bivariate results can still provide useful summary information and suggest general features of the results. Accordingly, Figures 2-5 plot the average impulse-response functions $\frac{\partial q_{1,t+n}}{\partial u_{2,t}}$ and $\frac{\partial q_{2,t+n}}{\partial u_{1,t}}$ measuring the impact of a two standard deviation individual financial institution shock on the index and the impact of a two standard deviation shock to the index on the individual financial institution’s risk. In Figures 2 and 3, the average is taken with respect to the geographical distribution. That is, the average impulse-response for Europe, say, is obtained by averaging all the impulse-response functions for European financial institutions. We notice two things. First, the impact of a shock to the index is much stronger than the impact of a shock to the individual financial institution. This result is partly driven by our identification assumption that shocks to the index have a contemporaneous impact on the return of the single financial institutions, while the institution’s specific shocks have only a lagged impact on the global financial index. Second, we notice that the risk of Asian financial institutions appears to be on average much less sensitive to global shocks than their European and North American counterparts.

Figures 4 and 5 plot the average impulse-response functions for the financial institutions grouped by line of business, i.e. banks, financial services, and insurances. We see that a shock to the index has a stronger initial impact on the group of insurance companies. Regarding the impact of shocks to the individual financial
institutions on the risk of the global index, banks have on average a lower initial impact, but the shock appears to be more persistent than for financial services and insurance companies.

To get an idea of the orders of magnitude involved, we compute the overall risk impact by integrating out all the individual impulse-responses. Figure 6 plots the average impulse-responses corresponding to the 20 financial institutions whose risk is most and least sensitive to a shock to the index, as well as those with the largest and smallest impact on the risk of the global index. It is clear that the shocks to the index have an impact of an order of magnitude greater than the shocks to the individual financial institutions. A two standard deviation shock to the index produces an average initial increase in the VaR of the most sensitive financial institutions of more than 20%. The shock is also quite persistent, as it is not yet completely absorbed after 15 weeks. On the other hand, for the least sensitive financial institutions, a shock to the index produces an average immediate increase in the VaR of less than 3%, which is then entirely absorbed after the second week. The figure also shows that the shocks to the individual financial institutions have a significantly lower impact on the risk of the index, in line with the results shown in Figures 3 and 5.

Again, we emphasize that the results presented in these figures merely summarize the pattern of results found in the bivariate analysis of our 230 financial institutions. To perform cross-comparisons, one could estimate a 3- or 4- or \( n \)-variate system using a common information set. Alternatively, one could impose that the \( B \) matrix in (7) is diagonal, which would be equivalent to assuming that the parameters of the system are variation free, as in Engle et al. (1983). This assumption would have the added advantage of allowing a separate estimation of each quantile. That is, for an \( n \)-variate system, the optimization problem in (6) can be broken down into \( n \) independent optimization problems, which in turn would considerably increase computational tractability.
5 Conclusion

We have developed theory ensuring the consistency and asymptotic normality of multivariate multi-quantile models. Our theory is general enough to comprehensively cover models with multiple random variables, multiple confidence levels and multiple lags of the quantiles.

We conduct an empirical analysis in which we estimate a VAR for VaR model using returns of individual financial institutions from around the world and a global financial sector index. By examining the impulse-response functions, we can study financial institutions’ resilience to shocks to the overall index and the impact they have on the VaR of the financial sector index. Judging from our bivariate models, we find that the risk of Asian financial institutions tends to be less sensitive to systemic shocks, whereas insurance companies exhibit a greater sensitivity to global shocks. By looking at the integral of all the individual impulse-responses, we find wide differences about how financial institutions react to different shocks. It is therefore of interest to see if these apparent differences persist, once we move from bivariate analyses to a multivariate analysis with a common information set.

The methods developed in this paper can be applied to many other contexts. For instance, many stress-test models are built from vector autoregressive models on credit risk indicators and macroeconomic variables. Starting from the estimated mean and adding assumptions on the multivariate distribution of the error terms, one can deduce the impact of a macro shock on the quantile of the credit risk variables. Our methodology provides a convenient alternative for stress testing, by allowing researchers to estimate vector autoregressive processes directly on the quantiles of interest, rather than on the mean.
References


6 Appendix 1 - Multi-quantile representation of a bivariate GARCH process with zero mean

Consider the following data generating process:

\[
\begin{bmatrix}
Y_{1t} \\
Y_{2t}
\end{bmatrix} = \begin{bmatrix}
\alpha_t & 0 \\
\beta_t & \gamma_t
\end{bmatrix} \begin{bmatrix}
\varepsilon_{1t} \\
\varepsilon_{2t}
\end{bmatrix} \quad \varepsilon_t \sim N(0, I)
\]

\[
\sigma_t(Y_{1t}) = \alpha_t
\]
\[
\sigma_t(Y_{2t}) = \sqrt{\beta_t^2 + \gamma_t^2}
\]

The respective quantile processes are:

\[
q_{1t} = kc_1 + ka_{11}|Y_{1t-1}| + ka_{12}|Y_{2t-1}| + b_{11}q_{1t-1} + b_{12}q_{2t-1}
\]
\[
q_{2t} = kc_2 + ka_{21}|Y_{1t-1}| + ka_{22}|Y_{2t-1}| + b_{21}q_{1t-1} + b_{22}q_{2t-1}
\]

where \(k\) is the \(\alpha\)-quantile of the standard normal distribution. In this case, setting the shocks equal to zero is equivalent to setting \(Y_t = 0\). The long-run quantiles are:

\[
\bar{q} = \begin{bmatrix}
1 - b_{11} & b_{12} \\
b_{21} & 1 - b_{22}
\end{bmatrix}^{-1} \begin{bmatrix}
kc_1 \\
kc_2
\end{bmatrix}
\]

7 Appendix 2 - Proofs

The proofs that follow are straightforward modifications of those in White, Kim, and Manganelli (2008).
We establish the consistency of \( \hat{\alpha}_T \) by applying results of White (1994). For this we impose the following moment and domination conditions. In stating this next condition and where convenient elsewhere, we exploit stationarity to omit explicit reference to all values of \( t \).

**Assumption 5** (i) For \( i = 1, \ldots, n \), \( E|Y_{it}| < \infty \); (ii) let \( D_{0,t} := \max_{i=1,\ldots,n} \max_{j=1,\ldots,p} \sup_{\alpha \in \mathbb{A}} |q_{i,j,t}(\cdot, \alpha)| \). Then \( E(D_{0,t}) < \infty \).

**Proof of Theorem 1**: We verify the conditions of corollary 5.11 of White (1994), which delivers \( \hat{\alpha}_T \to \alpha^* \), where

\[
\hat{\alpha}_T := \arg \max_{\alpha \in \mathbb{A}} T^{-1} \sum_{t=1}^{T} \varphi_t(Y_t, q_t(\cdot, \alpha)),
\]

and \( \varphi_t(Y_t, q_t(\cdot, \alpha)) := -\{\sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{ii,j}(Y_{it} - q_{i,j,t}(\cdot, \alpha))\} \). Assumption 1 ensures White’s Assumption 2.1. Assumption 3(i) ensures White’s Assumption 5.1. Our choice of \( \rho_{ii,j} \) satisfies White’s Assumption 5.4. To verify White’s Assumption 3.1, it suffices that \( \varphi_t(Y_t, q_t(\cdot, \alpha)) \) is dominated on \( \mathbb{A} \) by an integrable function (ensuring White’s Assumption 3.1(a,b)) and that for each \( \alpha \) in \( \mathbb{A} \), \( \{\varphi_t(Y_t, q_t(\cdot, \alpha))\} \) is stationary and ergodic (ensuring White’s Assumption 3.1(c), the strong uniform law of large numbers (ULLN)). Stationarity and ergodicity is ensured by Assumptions 1 and 3(i). To show domination, we write

\[
|\varphi_t(Y_t, q_t(\cdot, \alpha))| \leq \sum_{i=1}^{n} \sum_{j=1}^{p} |\rho_{ii,j}(Y_{it} - q_{i,j,t}(\cdot, \alpha))|
= \sum_{i=1}^{n} \sum_{j=1}^{p} |(Y_{it} - q_{i,j,t}(\cdot, \alpha))(\theta_{ij} - 1|Y_{it} - q_{i,j,t}(\cdot, \alpha)| \leq 0)|
\leq 2 \sum_{i=1}^{n} \sum_{j=1}^{p} (|Y_{it}| + |q_{i,j,t}(\cdot, \alpha)|)
\leq 2p \sum_{i=1}^{n} |Y_{it}| + 2np|D_{0,t}|,
\]
so that
\[
\sup_{\alpha \in \mathbb{A}} |\varphi_t(Y_t, q_t(\cdot, \alpha))| \leq 2p \sum_{i=1}^{n} |Y_{it}| + 2np|D_{0,t}|.
\]
Thus, \(2p \sum_{i=1}^{n} |Y_{it}| + 2np|D_{0,t}|\) dominates \(|\varphi_t(Y_t, q_t(\cdot, \alpha))|\); this has finite expectation by Assumption 5(i,ii).

It remains to verify White’s Assumption 3.2; here this is the condition that \(\alpha^*\) is the unique maximizer of \(E(\varphi_t(Y_t, q_t(\cdot, \alpha)))\). Given Assumptions 2(ii,b) and 4(i), it follows by argument directly parallel to that in the proof of White (1994, corollary 5.11) that for all \(\alpha \in \mathbb{A},\)

\[
E(\varphi_t(Y_t, q_t(\cdot, \alpha))) \leq E(\varphi_t(Y_t, q_t(\cdot, \alpha^*))).
\]

Thus, it suffices to show that the above inequality is strict for \(\alpha \neq \alpha^*\). Consider \(\alpha \neq \alpha^*\) such that \(||\alpha - \alpha^*|| > \epsilon\) and let \(\Delta(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E(\Delta_{i,j,t}(\alpha))\) with \(\Delta_{i,j,t}(\alpha) := \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha)) - \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*))\). It will suffice to show that \(\Delta(\alpha) > 0\). First, we define the "error" \(\varepsilon_{i,j,t} := Y_{it} - q_{i,j,t}(\cdot, \alpha^*)\) and let \(f_{i,j,t}(\cdot)\) be the density of \(\varepsilon_{i,j,t}\) conditional on \(F_{t-1}\). Noting that \(\delta_{i,j,t}(\alpha, \alpha^*) := q_{i,j,t}(\cdot, \alpha) - q_{i,j,t}(\cdot, \alpha^*)\), we next can show by some algebra and Assumption 2(ii.a) that

\[
E(\Delta_{i,j,t}(\alpha)) = E\left[ \int_{0}^{\delta_{i,j,t}(\alpha, \alpha^*)} (\delta_{i,j,t}(\alpha, \alpha^*) - s) f_{i,j,t}(s) ds \right] \\
\quad \geq E\left[ \frac{1}{2} \delta_{i,j,t}^2 |\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon \right] + \frac{1}{2} \delta_{i,j,t}(\alpha, \alpha^*)^2 1_{|\delta_{i,j,t}(\alpha, \alpha^*)| \leq \delta_\epsilon} \\
\quad \geq \frac{1}{2} \delta_{i,j,t}^2 E[1_{|\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon}].
\]

The first inequality above comes from the fact that Assumption 2(ii.a) implies
that for any $\delta > 0$ sufficiently small, we have $f_{i,j,t}(s) > \delta$ for $|s| < \delta$. Thus,

$$
\Delta(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E(\Delta_{i,j,t}(\alpha)) \geq \frac{1}{2} \delta^2 \sum_{i=1}^{n} \sum_{j=1}^{p} E[1[|\delta_{i,j,t}(\alpha) > \delta_i]]
$$

$$
= \frac{1}{2} \delta^2 \sum_{i=1}^{n} \sum_{j=1}^{p} P[|\delta_{i,j,t}(\alpha) > \delta_e] \geq \frac{1}{2} \delta^2 \sum_{(i,j) \in I} P[|\delta_{i,j,t}(\alpha, \alpha^*) > \delta_e]
$$

$$
\geq \frac{1}{2} \delta^2 \mathbb{P}[\bigcup_{(i,j) \in I} \{|\delta_{i,j,t}(\alpha, \alpha^*) > \delta_e\}] > 0,
$$

where the final inequality follows from Assumption 4(i.b). As $\alpha$ is arbitrary, the result follows. ■

Next, we establish the asymptotic normality of $T^{1/2}(\hat{\alpha}_T - \alpha^*)$. We use a method originally proposed by Huber (1967) and later extended by Weiss (1991). We first sketch the method before providing formal conditions and the proof.

Huber’s method applies to our estimator $\hat{\alpha}_T$, provided that $\hat{\alpha}_T$ satisfies the asymptotic first order conditions

$$
T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)) = o_p(T^{1/2}),
$$

where $\nabla q_{i,j,t}(\cdot, \alpha)$ is the $\ell \times 1$ gradient vector with elements $(\partial / \partial \alpha_s)q_{i,j,t}(\cdot, \alpha), s = 1, \ldots, \ell$, and $\psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T))$ is a generalized residual. Our first task is thus to ensure that eq. (8) holds.

Next, we define

$$
\lambda(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha)\psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))].
$$

With $\lambda(\alpha)$ continuously differentiable at $\alpha^*$ interior to $\mathcal{A}$, we can apply the mean value theorem to obtain

$$
\lambda(\alpha) = \lambda(\alpha^*) + Q_0(\alpha - \alpha^*),
$$

(9)
where $Q_0$ is an $\ell \times \ell$ matrix with $(1 \times \ell)$ rows $Q_{0,s} = \nabla' \lambda(\tilde{\alpha}_{(s)})$, where $\tilde{\alpha}_{(s)}$ is a mean value (different for each $s$) lying on the segment connecting $\alpha$ and $\alpha^*, s = 1, \ldots, \ell$. It is straightforward to show that correct specification ensures that $\lambda(\alpha^*)$ is zero.

We will also show that

$$Q_0 = -Q^* + O(||\alpha - \alpha^*||), \quad (10)$$

where $Q^* := \sum_{i=1}^n \sum_{j=1}^p E[f_{i,j,t}(0) \nabla q_{i,j,t}(\cdot, \alpha^*) \nabla' q_{i,j,t}(\cdot, \alpha^*)]$ with $f_{i,j,t}(0)$ the value at zero of the density $f_{i,j,t}$ of $\varepsilon_{i,j,t} := Y_t - q_{i,j,t}(\cdot, \alpha^*)$, conditional on $F_{t-1}$. Combining eqs. (9) and (10) and putting $\lambda(\alpha^*) = 0$, we obtain

$$\lambda(\alpha) = -Q^*(\alpha - \alpha^*) + O(||\alpha - \alpha^*||^2). \quad (11)$$

The next step is to show that

$$T^{1/2} \lambda(\hat{\alpha}_T) + H_T = o_p(1), \quad (12)$$

where $H_T := T^{-1/2} \sum_{t=1}^T \eta_t^*, \text{ with } \eta_t^* := \sum_{i=1}^n \sum_{j=1}^p \nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{\theta_{ij}}(\varepsilon_{i,j,t})$. Eqs. (11) and (12) then yield the following asymptotic representation of our estimator $\hat{\alpha}_T$:

$$T^{1/2}(\hat{\alpha}_T - \alpha^*) = Q^{*-1}T^{-1/2} \sum_{t=1}^T \eta_t^* + o_p(1). \quad (13)$$

As we impose conditions sufficient to ensure that $\{\eta_t^*, F_t\}$ is a martingale difference sequence (MDS), a suitable central limit theorem (e.g., theorem 5.24 in White, 2001) applies to eq. (13) to yield the desired asymptotic normality of $\hat{\alpha}_T$:

$$T^{1/2}(\hat{\alpha}_T - \alpha^*) \overset{d}{\to} N(0, Q^{*-1}V^*Q^{*-1}), \quad (14)$$

where $V^* := E(\eta_t^* \eta_t'^*)$.

We now strengthen the conditions above to ensure that each step of the above
argument is valid.

**Assumption 2** (iii) (a) There exists a finite positive constant \( f_0 \) such that for each \( i \) and \( t \), each \( \omega \in \Omega \), and each \( y \in \mathbb{R} \), \( f_{it}(\omega, y) \leq f_0 < \infty \); (b) There exists a finite positive constant \( L_0 \) such that for each \( i \) and \( t \), each \( \omega \in \Omega \), and each \( y_1, y_2 \in \mathbb{R} \), \(|f_{it}(\omega, y_1) - f_{it}(\omega, y_2)| \leq L_0|y_1 - y_2|\).

Next we impose sufficient differentiability of \( q_t \) with respect to \( \alpha \).

**Assumption 3** (ii) For each \( t \) and each \( \omega \in \Omega \), \( q_t(\omega, \cdot) \) is continuously differentiable on \( A \); (iii) For each \( t \) and each \( \omega \in \Omega \), \( q_t(\omega, \cdot) \) is twice continuously differentiable on \( A \); (iv) For each \( t \) and each \( \omega \in \Omega \), \( q_t(\omega, \cdot) \) is continuously differentiable on \( A \).

To exploit the mean value theorem, we require that \( \alpha^* \) belongs to \( int(A) \), the interior of \( A \).

**Assumption 4** (ii) \( \alpha^* \in int(A) \).

Next, we place domination conditions on the derivatives of \( q_t \).

**Assumption 5** (iii) Let \( D_{1,t} := \max_{i=1,...,n} \max_{j=1,...,p} \max_{s=1,...,l} \sup_{\alpha \in \mathbb{R}} |(\partial/\partial \alpha_s) q_{i,j,t}(\cdot, \alpha)| \). Then (a) \( E(D_{1,t}) < \infty \); (b) \( E(D_{1,t}^2) < \infty \); (iv) Let \( D_{2,t} := \max_{i=1,...,n} \max_{j=1,...,p} \max_{s=1,...,l} \max_{h=1,...,l} \sup_{\alpha \in \mathbb{R}} |(\partial^2/\partial \alpha_s \partial \alpha_h) q_{i,j,t}(\cdot, \alpha)| \). Then (a) \( E(D_{2,t}) < \infty \); (b) \( E(D_{2,t}^2) < \infty \).

**Assumption 6** (i) \( Q^* := \sum_{i=1}^n \sum_{j=1}^p E[f_{i,j,t}(0) \nabla q_{i,j,t}(\cdot, \alpha^*) \nabla' q_{i,j,t}(\cdot, \alpha^*)] \) is positive definite; (ii) \( V^* := E(\eta^*_t \eta^*_t) \) is positive definite.

Assumptions 3(ii) and 5(iii.a) are additional assumptions helping to ensure that eq. (8) holds. Further imposing Assumptions 2(iii), 3(iii.a), 4(ii), and 5(iv.a) suffices to ensure that eq. (11) holds. The additional regularity provided by Assumptions 5(iii.b), 5(iv.b), and 6(i) ensures that eq. (12) holds. Assumptions 5(iii.b) and 6(ii) help ensure the availability of the MDS central limit theorem.
We now have conditions sufficient to prove asymptotic normality of our MVMQ-
CAViaR estimator.

**Proof of Theorem 2**: As outlined above, we first prove

\[
T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)) = o_p(1). \quad (15)
\]

The existence of \( \nabla q_{i,j,t} \) is ensured by Assumption 3(ii). Let \( e_i \) be the \( \ell \times 1 \) unit vector with \( i^{th} \) element equal to one and the rest zero, and let

\[
G_s(c) := T^{-1/2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T + ce_s)),
\]

for any real number \( c \). Then by the definition of \( \hat{\alpha}_T \), \( G_s(c) \) is minimized at \( c = 0 \). Let \( H_s(c) \) be the derivative of \( G_s(c) \) with respect to \( c \) from the right. Then

\[
H_s(c) = -T^{-1/2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla_s q_{i,j,t}(\cdot, \hat{\alpha}_T + ce_s) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T + ce_s)),
\]

where \( \nabla_s q_{i,j,t}(\cdot, \hat{\alpha}_T + ce_s) \) is the \( s^{th} \) element of \( \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T + ce_s) \). Using the facts that (i) \( H_s(c) \) is non-decreasing in \( c \) and (ii) for any \( \epsilon > 0 \), \( H_s(-\epsilon) \leq 0 \) and \( H_s(\epsilon) \geq 0 \), we have

\[
|H_s(0)| \leq H_s(\epsilon) - H_s(-\epsilon) \\
\leq T^{-1/2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} |\nabla_s q_{i,j,t}(\cdot, \hat{\alpha}_T)| 1_{|Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)| = 0} \\
\leq T^{-1/2} \max_{1 \leq t \leq T} D_{1,t} \sum_{i=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{|Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)| = 0},
\]

where the last inequality follows by the domination condition imposed in Assumption 5(iii.a). Because \( D_{1,t} \) is stationary, \( T^{-1/2} \max_{1 \leq t \leq T} D_{1,t} = o_p(1) \). The second term is bounded in probability: \( \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{|Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)| = 0} = O_p(1) \) given
Assumption 2(i,ii.a) (see Koenker and Bassett, 1978, for details). Since $H_s(0)$ is the $s^{th}$ element of $T^{-1/2} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T))$, the claim in (15) is proved.

Next, for each $\alpha \in \mathbb{A}$, Assumptions 3(ii) and 5(iii.a) ensure the existence and finiteness of the $\ell \times 1$ vector

$$\lambda(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha) \int_{\delta_{i,j,t}(\alpha, \alpha^*)}^{0} f_{i,j,t}(s) ds],$$

where $\delta_{i,j,t}(\alpha, \alpha^*) := q_{i,j,t}(\cdot, \alpha) - q_{i,j,t}(\cdot, \alpha^*)$ and $f_{i,j,t}(s) = (d/ds)F_{it}(s + q_{i,j,t}(\cdot, \alpha^*))$ represents the conditional density of $\varepsilon_{i,j,t} := Y_{it} - q_{i,j,t}(\cdot, \alpha^*)$ with respect to Lebesgue measure. The differentiability and domination conditions provided by Assumptions 3(iii) and 5(iv.a) ensure (e.g., by Bartle, 1966, corollary 5.9) the continuous differentiability of $\lambda(\alpha)$ on $\mathbb{A}$, with

$$\nabla \lambda(\alpha) = \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla \nabla q_{i,j,t}(\cdot, \alpha) \int_{\delta_{i,j,t}(\alpha, \alpha^*)}^{0} f_{i,j,t}(s) ds].$$

Since $\alpha^*$ is interior to $\mathbb{A}$ by Assumption 4(ii), the mean value theorem applies to each element of $\lambda(\alpha)$ to yield

$$\lambda(\alpha) = \lambda(\alpha^*) + Q_0(\alpha - \alpha^*), \quad (16)$$

for $\alpha$ in a convex compact neighborhood of $\alpha^*$ where $Q_0$ is an $\ell \times \ell$ matrix with $(1 \times \ell)$ rows $Q_s(\bar{\alpha}(s)) = \nabla^T \lambda(\bar{\alpha}(s))$, where $\bar{\alpha}(s)$ is a mean value (different for each $s$) lying on the segment connecting $\alpha$ and $\alpha^*$ with $s = 1, \ldots, \ell$. The chain rule and an application of the Leibniz rule to $\int_{\delta_{i,j,t}(\alpha, \alpha^*)}^{0} f_{i,j,t}(s) ds$ then give

$$Q_s(\alpha) = A_s(\alpha) - B_s(\alpha),$$

28
where

\[
A_s(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla_s \nabla' q_{i,j,t}(\cdot, \alpha) \int_{0}^{T} f_{i,j,t}(s) ds]
\]

\[
B_s(\alpha) := \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(\delta_{i,j,t}(\alpha, \alpha^*)) \nabla_s q_{i,j,t}(\cdot, \alpha) \nabla' q_{i,j,t}(\cdot, \alpha)].
\]

Assumption 2(iii) and the other domination conditions (those of Assumption 5) then ensure that

\[
A_s(\bar{\alpha}_t) = O(||\alpha - \alpha^*||)
\]

\[
B_s(\bar{\alpha}_t) = Q_s^* + O(||\alpha - \alpha^*||),
\]

where \(Q_s^* := \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(0) \nabla q_{i,j,t}(\cdot, \alpha^*) \nabla' q_{i,j,t}(\cdot, \alpha^*)].\) Letting \(Q_s := \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(0) \nabla q_{i,j,t}(\cdot, \alpha^*) \nabla' q_{i,j,t}(\cdot, \alpha^*)],\) we obtain

\[
Q_0 = -Q_s + O(||\alpha - \alpha^*||). \tag{17}
\]

Next, we have that \(\lambda(\alpha^*) = 0.\) To show this, we write

\[
\lambda(\alpha^*) = \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{t_i} (Y_{it} - q_{i,j,t}(\cdot, \alpha^*))]
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{p} E(E[\nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{t_i} (Y_{it} - q_{i,j,t}(\cdot, \alpha^*)) \mid \mathcal{F}_{t-1}])
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{p} E(\nabla q_{i,j,t}(\cdot, \alpha^*) E[\psi_{t_i} (Y_{it} - q_{i,j,t}(\cdot, \alpha^*)) \mid \mathcal{F}_{t-1}])
\]

\[
= \sum_{i=1}^{n} \sum_{j=1}^{p} E(\nabla q_{i,j,t}(\cdot, \alpha^*) E[\psi_{t_i} (\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}]) = 0,
\]
as $E[\psi_{\theta,j}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}] = \theta_{ij} - E[1_{Y_{at} \leq q_{i,j,t}^*} \mid \mathcal{F}_{t-1}] = 0$, by definition of $q_{i,j,t}^*$ for $i = 1, \ldots, n$ and $j = 1, \ldots, p$ (see eq. (3)). Combining $\lambda(\alpha^*) = 0$ with eqs. (16) and (17), we obtain

$$\lambda(\alpha) = -Q^*(\alpha - \alpha^*) + O(||\alpha - \alpha^*||^2).$$

(18)

The next step is to show that

$$T^{1/2}\lambda(\hat{\alpha}_T) + H_T = o_p(1)$$

(19)

where $H_T := T^{-1/2} \sum_{t=1}^T \eta_t^*$, with $\eta_t^* := \eta_t(\alpha^*)$ and $\eta_t(\alpha) := \sum_{i=1}^n \sum_{j=1}^p \nabla q_{i,j,t}(\cdot, \alpha) \psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \alpha))$. Let $u_t(\alpha, d) := \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} ||\eta_t(\tau) - \eta_t(\alpha)||$. By the results of Huber (1967) and Weiss (1991), to prove (19) it suffices to show the following: (i) there exist $a > 0$ and $d_0 > 0$ such that $||\lambda(\alpha)|| \geq a ||\alpha - \alpha^*||$ for $||\alpha - \alpha^*|| \leq d_0$; (ii) there exist $b > 0$, $d_0 > 0$, and $d \geq 0$ such that $E[u_t(\alpha, d)] \leq bd$ for $||\alpha - \alpha^*|| + d \leq d_0$; and (iii) there exist $c > 0$, $d_0 > 0$, and $d \geq 0$ such that $E[u_t(\alpha, d)^2] \leq cd$ for $||\alpha - \alpha^*|| + d \leq d_0$.

The condition that $Q^*$ is positive-definite in Assumption 6(i) is sufficient for (i). For (ii), we have that for given (small) $d > 0$

$$u_t(\alpha, d) \leq \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} \sum_{i=1}^n \sum_{j=1}^p \|\nabla q_{i,j,t}(\cdot, \tau) \psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \tau)) - \nabla q_{i,j,t}(\cdot, \alpha) \psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \alpha))\|$$

$$\leq \sum_{i=1}^n \sum_{j=1}^p \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} \|\psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \tau))\| \times \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} \|\nabla q_{i,j,t}(\cdot, \tau) - \nabla q_{i,j,t}(\cdot, \alpha)\|$$

$$+ \sum_{i=1}^n \sum_{j=1}^p \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} \|\psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \alpha)) - \psi_{\theta,j} (Y_{at} - q_{i,j,t}(\cdot, \tau))\|$$

$$\times \sup_{\{\tau: ||\tau - \alpha|| \leq d\}} \|\nabla q_{i,j,t}(\cdot, \alpha)\|$$

$$\leq npD_{2,t}d + D_{1,t} \sum_{i=1}^n \sum_{j=1}^p 1_{|Y_{at} - q_{i,j,t}(\cdot, \alpha)| < D_{1,t}d}$$
using the following: (i) \[||\psi_{\theta_i}(Y_{it} - q_{i,j,t}(\cdot, \tau))|| \leq 1; \]
(ii) \[||\psi_{\theta_i}(Y_{it} - q_{i,j,t}(\cdot, \alpha)) - \psi_{\theta_i}(Y_{it} - q_{i,j,t}(\cdot, \tau))|| \leq 1||Y_{it} - q_{i,j,t}(\cdot, \alpha) < |q_{i,j,t}(\cdot, \tau) - q_{i,j,t}(\cdot, \alpha)||; \]
and (iii) the mean value theorem applied to \(\nabla q_{i,j,t}(\cdot, \tau)\) and \(q_{i,j,t}(\cdot, \alpha)\). Hence, we have

\[E[u_t(\alpha, d)] \leq npC_0d + 2npC_1f_0d\]

for some constants \(C_0\) and \(C_1\), given Assumptions 2(iii.a), 5(iii.a), and 5(iv.a). Hence, (ii) holds for \(b = npC_0 + 2npC_1f_0\) and \(d_0 = 2d\). The last condition (iii) can be similarly verified by applying the \(c_r\)--inequality to eq. (??) with \(d < 1\) (so that \(d^2 < d\)) and using Assumptions 2(iii.a), 5(iii.b), and 5(iv.b). Thus, eq. (19) is verified.

Combining eqs. (18) and (19) thus yields

\[Q^*T^{1/2}(\hat{\alpha}_T - \alpha^*) = T^{-1/2} \sum_{t=1}^{T} \eta_t^* + o_p(1).\]

But \(\{\eta_t^*, \mathcal{F}_t\}\) is a stationary ergodic martingale difference sequence (MDS). In particular, \(\eta_t^*\) is measurable--\(\mathcal{F}_t\), and \(E(\eta_t^*|\mathcal{F}_{t-1}) = E(\sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \alpha^*)\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) | \mathcal{F}_{t-1}) = 0,\) as \(E[\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) | \mathcal{F}_{t-1}] = 0\) for all \(i = 1, ..., n\) and \(j = 1, ..., p\). Assumption 5(iii.b) ensures that \(V^* := E(\eta_t^*\eta_t'^*)\) is finite. The MDS central limit theorem (e.g., theorem 5.24 of White, 2001) applies, provided \(V^*\) is positive definite (as ensured by Assumption 6(ii)) and that \(T^{-1} \sum_{t=1}^{T} \eta_t^*\eta_t'^* = V^* + o_p(1),\) which is ensured by the ergodic theorem. The standard argument now gives

\[V^{* -1/2}Q^*T^{1/2}(\hat{\alpha}_T - \alpha^*) \xrightarrow{d} N(0, I),\]

which completes the proof. 

Proof of Theorem 3: We have

\[ \hat{V}_T - V^* = (T^{-1} \sum_{t=1}^{T} \hat{\eta}_t \hat{\eta}_t' - T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t') + (T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t' - E[\eta_t^* \eta_t']), \]

where \( \hat{\eta}_t := \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t} \hat{\psi}_{i,j,t} \) and \( \eta_t^* := \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}^* \psi_{i,j,t}^* \), with \( \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \), \( \hat{\psi}_{i,j,t} := \psi_{\hat{\theta}_{i,j}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)) \), \( \nabla q_{i,j,t}^* := \nabla q_{i,j,t}(\cdot, \alpha^*) \), and \( \psi_{i,j,t}^* := \psi_{\theta_{i,j}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*)) \). Assumptions 1 and 2(i,ii) ensure that \( \{\eta_t^* \eta_t'\} \) is a stationary ergodic sequence. Assumptions 3(i,ii), 4(i,a), and 5(iii) ensure that \( E[\eta_t^* \eta_t'] < \infty \).

It follows by the ergodic theorem that \( T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t' - E[\eta_t^* \eta_t'] = o_p(1) \). Thus, it suffices to prove \( T^{-1} \sum_{t=1}^{T} \hat{\eta}_t \hat{\eta}_t' - T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t' = o_p(1) \).

The \((h, s)\) element of \( T^{-1} \sum_{t=1}^{T} \hat{\eta}_t \hat{\eta}_t' - T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t' \) is

\[ T^{-1} \sum_{t=1}^{T} \{ \sum_{i=1}^{n} \sum_{j=1}^{p} \sum_{k=1}^{p} (\hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla h q_{i,j,t} \nabla s q_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t}^* \nabla s q_{l,k,t}^*) \}. \]

Thus, it will suffice to show that for each \((h, s)\) and \((i, j, l, k)\),

\[ T^{-1} \sum_{t=1}^{T} \{ \hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla h q_{i,j,t} \nabla s q_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t}^* \nabla s q_{l,k,t}^* \} = o_p(1). \]

By the triangle inequality,

\[ |T^{-1} \sum_{t=1}^{T} \{ \hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla h q_{i,j,t} \nabla s q_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t}^* \nabla s q_{l,k,t}^* \}| \leq A_T + B_T, \]

where

\[ A_T = T^{-1} \sum_{t=1}^{T} |\hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla h q_{i,j,t} \nabla s q_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t}^* \nabla s q_{l,k,t}^*| \]

\[ B_T = T^{-1} \sum_{t=1}^{T} |\psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t}^* \nabla s q_{l,k,t}^* - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla h q_{i,j,t} \nabla s q_{l,k,t}^*|. \]
We now show that $A_T = o_p(1)$ and $B_T = o_p(1)$, delivering the desired result. For $A_T$, the triangle inequality gives

$$A_T \leq A_{1T} + A_{2T} + A_{3T},$$

where

$$A_{1T} = T^{-1} \sum_{t=1}^{T} \theta_{ij} |1[\varepsilon_{i,j,t} \leq 0] - 1|1[\varepsilon_{i,j,t} \leq 0]| \| \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} |,$$

$$A_{2T} = T^{-1} \sum_{t=1}^{T} \theta_{ik} |1[\varepsilon_{i,k,t} \leq 0] - 1|1[\varepsilon_{i,k,t} \leq 0]| \| \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} |,$$

$$A_{3T} = T^{-1} \sum_{t=1}^{T} |1[\varepsilon_{i,j,t} \leq 0] 1[\varepsilon_{i,k,t} \leq 0] 1[\varepsilon_{i,k,t} \leq 0]| \| \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} |.$$

Theorem 2, ensured by Assumptions 1 - 6, implies that $T^{1/2} |\hat{\alpha}_T - \alpha^*| = O_p(1)$. This, together with Assumptions 2(iii,iv) and 5(iii,b), enables us to apply the same techniques used in Kim and White (2003) to show $A_{1T} = o_p(1)$, $A_{2T} = o_p(1)$, and $A_{3T} = o_p(1)$, implying $A_T = o_p(1)$.

It remains to show $B_T = o_p(1)$. By the triangle inequality,

$$B_T \leq B_{1T} + B_{2T},$$

where

$$B_{1T} = T^{-1} \sum_{t=1}^{T} |\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} - E[\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t}]|,$$

$$B_{2T} = T^{-1} \sum_{t=1}^{T} |\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} - E[\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t}]|.$$

Assumptions 1, 2(i,ii), 3(i,ii), 4(i,a), and 5(iii) ensure that the ergodic theorem applies to $\{\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,k,t} \}$, so $B_{1T} = o_p(1)$. Next, Assumptions 1, 3(i,ii), and 5(iii) ensure that the stationary ergodic ULLN applies to $\{\psi_{i,j,t}^* \psi_{i,k,t}^* \nabla h \hat{q}_{i,j,t} (\cdot, \alpha) \}$.
\( \nabla_{\alpha} q_{i,k,t}(\cdot, \alpha) \). This and the result of Theorem 1 \((\hat{\alpha}_T - \alpha^* = o_p(1))\) ensure that \( B_{2T} = o_p(1) \) by e.g., White (1994, corollary 3.8), and the proof is complete. ■

To establish consistency of \( \hat{Q}_T \), we strengthen the domination condition on \( \nabla q_{i,j,t} \) and impose conditions on \( \{\hat{\epsilon}_T\} \).

**Assumption 5** (iii.c) \( E(D_{1,t}^3) < \infty \).

**Assumption 7** \( \{\hat{\epsilon}_T\} \) is a stochastic sequence and \( \{c_T\} \) is a non-stochastic sequence such that (i) \( \hat{\epsilon}_T/c_T \xrightarrow{p} 1 \); (ii) \( c_T = o(1) \); and (iii) \( c_T^{-1} = o(T^{1/2}) \).

**Proof of Theorem 4**: We begin by sketching the proof. We first define

\[
Q_T := (2c_T T)^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{[-c_T \leq \epsilon_{i,j,t} \leq c_T]} \nabla q_{i,j,t}^* \nabla' q_{i,j,t}^* ,
\]

and then we will show the following:

\[
Q^* - E(Q_T) \xrightarrow{p} 0, \tag{20}
\]

\[
E(Q_T) - Q_T \xrightarrow{p} 0, \tag{21}
\]

\[
Q_T - \hat{Q}_T \xrightarrow{p} 0. \tag{22}
\]

Combining the results above will deliver the desired outcome: \( \hat{Q}_T - Q^* \xrightarrow{p} 0 \).

For (20), one can show by applying the mean value theorem to \( F_{i,j,t}(c_T) - F_{i,j,t}(-c_T) \), where \( F_{i,j,t}(c) := \int 1_{s \leq c} f_{i,j,t}(s) \, ds \), that

\[
E(Q_T) = T^{-1} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(\xi_{i,j,T}) \nabla q_{i,j,t}^* \nabla' q_{i,j,t}^*] = \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(\xi_{i,j,T}) \nabla q_{i,j,t}^* \nabla' q_{i,j,t}^*],
\]

where \( \xi_{i,j,T} \) is a mean value lying between \(-c_T\) and \(c_T\), and the second equality
follows by stationarity. Therefore, the \((h, s)\) element of \(|E(Q_T) - Q^*|\) satisfies

\[
\frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ \frac{1}{2c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1[|\hat{\theta}_T - \hat{\theta}_T|] \nabla_{\theta} \hat{Q}_{i,j,t} \nabla_{\rho} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \right] 
\leq \frac{c_T}{c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ |\hat{\theta}_T - \hat{\theta}_T| \nabla_{\theta} \hat{Q}_{i,j,t} \nabla_{\rho} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \right] 
\leq \frac{c_T}{c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ A_{1T} + A_{2T} + A_{3T} + (1 - \frac{\hat{\theta}_T}{c_T})A_{4T} \right],
\]

which converges to zero as \(c_T \rightarrow 0\). The second inequality follows by Assumption 2(iii.b), and the last inequality follows by Assumption 5(iii.b). Therefore, we have the result in eq.(20).

To show (21), it suffices simply to apply a LLN for double arrays, e.g. theorem 2 in Andrews (1988).

Finally, for (22), we consider the \((h, s)\) element of \(|\hat{Q}_T - Q_T|\), given by

\[
\frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ \frac{1}{2c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1[|\hat{\theta}_T - \hat{\theta}_T|] \nabla_{\theta} \hat{Q}_{i,j,t} \nabla_{\rho} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \right] 
\leq \frac{c_T}{c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ |\hat{\theta}_T - \hat{\theta}_T| \nabla_{\theta} \hat{Q}_{i,j,t} \nabla_{\rho} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \nabla_{\sigma} \hat{Q}_{i,j,t} \right] 
\leq \frac{c_T}{c_T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left[ A_{1T} + A_{2T} + A_{3T} + (1 - \frac{\hat{\theta}_T}{c_T})A_{4T} \right],
\]
where

\[ A_{1T} := \frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} \left| 1_{[-\hat{c}_T \leq \hat{e}_{i,j,t} \leq \hat{c}_T]} - 1_{[-c_T \leq e_{i,j,t} \leq c_T]} \right| \times |\nabla_h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}| \]

\[ A_{2T} := \frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{[-c_T \leq e_{i,j,t} \leq c_T]} |\nabla_h \hat{q}_{i,j,t} - \nabla_h \hat{q}_{i,j,t}^*| \times |\nabla s \hat{q}_{i,j,t}| \]

\[ A_{3T} := \frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{[-c_T \leq e_{i,j,t} \leq c_T]} |\nabla_h \hat{q}_{i,j,t}^* \times |\nabla s \hat{q}_{i,j,t} - \nabla s \hat{q}_{i,j,t}^*| \]

\[ A_{4T} := \frac{1}{2c_T T} \sum_{t=1}^{T} \sum_{i=1}^{n} \sum_{j=1}^{p} 1_{[-c_T \leq e_{i,j,t} \leq c_T]} |\nabla_h \hat{q}_{i,j,t}^* \nabla s \hat{q}_{i,j,t}^*| \].

It will suffice to show that \( A_{1T} = o_p(1), A_{2T} = o_p(1), A_{3T} = o_p(1), \) and \( A_{4T} = O_p(1). \) Then, because \( \hat{c}_T/c_T \xrightarrow{p} 1, \) we obtain the desired result: \( \hat{Q}_T - Q^* \xrightarrow{p} 0. \)

We first show \( A_{1T} = o_p(1). \) It will suffice to show that for each \( i \) and \( j, \)

\[ \frac{1}{2c_T T} \sum_{t=1}^{T} \left| 1_{[-\hat{c}_T \leq \hat{e}_{i,j,t} \leq \hat{c}_T]} - 1_{[-c_T \leq e_{i,j,t} \leq c_T]} \right| \times |\nabla_h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}| = o_p(1). \]

Let \( \alpha_T \) lie between \( \hat{\alpha}_T \) and \( \alpha^* \), and put \( d_{i,j,t,T} := ||\nabla q_{i,j,t}(\cdot, \alpha_T)|| \times ||\hat{\alpha}_T - \alpha^*|| + |\hat{c}_T - c_T|. \) Then

\[ (2c_T T)^{-1} \sum_{t=1}^{T} \left| 1_{[-\hat{c}_T \leq \hat{e}_{i,j,t} \leq \hat{c}_T]} - 1_{[-c_T \leq e_{i,j,t} \leq c_T]} \right| \times |\nabla_h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}| \leq U_T + V_T, \]

where

\[ U_T := (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_{i,j,t} - c_T| < d_{i,j,t,T}]} |\nabla_h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}| \]

\[ V_T := (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|e_{i,j,t} + c_T| < d_{i,j,t,T}]} |\nabla_h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}|. \]

It will suffice to show that \( U_T \xrightarrow{p} 0 \) and \( V_T \xrightarrow{p} 0. \) Let \( \eta > 0 \) and let \( z \) be an arbitrary positive number. Then, using reasoning similar to that of Kim and White (2003,
lemma 5), one can show that for any $\eta > 0$,

$$P(U_T > \eta) \leq P((2c_T T)^{-1} \sum_{t=1}^{T} 1_{|e_{i,j,t} - c_T| < (||\nabla q_{i,j,t}(\cdot, \alpha_T)|| + 1) \epsilon_T}) |\nabla h \hat{q}_{i,j,t} \nabla s \hat{q}_{i,j,t}| > \eta)$$

$$\leq \frac{z f_0}{\eta T} \sum_{t=1}^{T} E \{(||\nabla q_{i,j,t}(\cdot, \alpha_T)|| + 1) |\nabla h \hat{q}_{j,t} \nabla s \hat{q}_{j,t}|\}$$

$$\leq z f_0 \{E|D_{1,t}^3| + E|D_{2,t}^2|\}/\eta,$$

where the second inequality is due to the Markov inequality and Assumption 2(iii.a), and the third is due to Assumption 5(iii.c). As $z$ can be chosen arbitrarily small and the remaining terms are finite by assumption, we have $U_T \to 0$. The same argument is used to show $V_T \to 0$. Hence, $A_{1T} = o_p(1)$ is proved.

Next, we show $A_{2T} = o_p(1)$. For this, it suffices to show $A_{2T,i,j} := \frac{1}{2c_T T} \sum_{t=1}^{T} 1_{[-c_T \leq e_{i,j,t} \leq c_T]} |\nabla h \hat{q}_{i,j,t} - \nabla h q^*_{i,j,t}| \times |\nabla s \hat{q}_{i,j,t}| = o_p(1)$ for each $i$ and $j$. Note that

$$A_{2T,i,j} \leq \frac{1}{2c_T T} \sum_{t=1}^{T} |\nabla h \hat{q}_{i,j,t} - \nabla h q^*_{i,j,t}| \times |\nabla s \hat{q}_{i,j,t}|$$

$$\leq \frac{1}{2c_T T} \sum_{t=1}^{T} ||\nabla^2 h q_{i,j,t}(\cdot, \alpha)|| \times ||\hat{\alpha}_T - \alpha^*|| \times |\nabla s \hat{q}_{i,j,t}|$$

$$\leq \frac{1}{2c_T} ||\hat{\alpha}_T - \alpha^*|| \frac{1}{T} \sum_{t=1}^{T} D_{2,t} D_{1,t}$$

$$= \frac{1}{2c_T T^{1/2}} T^{1/2} ||\hat{\alpha}_T - \alpha^*|| \frac{1}{T} \sum_{t=1}^{T} D_{2,t} D_{1,t},$$

where $\hat{\alpha}$ is between $\hat{\alpha}_T$ and $\alpha^*$, and $\nabla^2 h q_{i,j,t}(\cdot, \tilde{\alpha})$ is the first derivative of $\nabla h \hat{q}_{j,t}$ with respect to $\alpha$, which is evaluated at $\tilde{\alpha}$. The last expression above is $o_p(1)$ because: (i) $T^{1/2} ||\hat{\alpha}_T - \alpha^*|| = O_p(1)$ by Theorem 2; (ii) $T^{-1} \sum_{t=1}^{T} D_{2,t} D_{1,t} = O_p(1)$ by the ergodic theorem; and (iii) $1/(c_T T^{1/2}) = o_p(1)$ by Assumption 7(iii). Hence, $A_{2T} = o_p(1)$. The other claims $A_{3T} = o_p(1)$ and $A_{4T} = O_p(1)$ can be analogously and more easily proven. Hence, they are omitted. Therefore, we finally have
$Q_T - \dot{Q}_T \xrightarrow{p} 0$, which, together with (20) and (21), implies that $\dot{Q}_T - Q^* \xrightarrow{p} 0$.

The proof is complete. ■
Table 1 – Breakdown of financial institutions by sector and by geographic area

<table>
<thead>
<tr>
<th></th>
<th>Banks</th>
<th>Financial Services</th>
<th>Insurances</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>EU</strong></td>
<td>47</td>
<td>22</td>
<td>27</td>
</tr>
<tr>
<td><strong>North America</strong></td>
<td>25</td>
<td>17</td>
<td>28</td>
</tr>
<tr>
<td><strong>Asia</strong></td>
<td>47</td>
<td>14</td>
<td>3</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>119</td>
<td>53</td>
<td>58</td>
</tr>
</tbody>
</table>

*Note: Classification as provided by Datastream. Swiss and Norwegian financial institutions have been classified as EU. Asia includes Australian financial institutions.*

Table 2 – Estimates and standard errors for the Sumitomo Mitsui Financial Group

<table>
<thead>
<tr>
<th></th>
<th>$c_1$</th>
<th>$a_{11}$</th>
<th>$a_{12}$</th>
<th>$b_{11}$</th>
<th>$b_{12}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>s.e.</td>
<td>-2.71</td>
<td>-0.89</td>
<td>-0.13</td>
<td>0.61</td>
<td>-0.11</td>
</tr>
<tr>
<td></td>
<td>1.48</td>
<td>0.37</td>
<td>0.13</td>
<td>0.12</td>
<td>0.06</td>
</tr>
<tr>
<td></td>
<td>$c_2$</td>
<td>$a_{21}$</td>
<td>$a_{22}$</td>
<td>$b_{21}$</td>
<td>$b_{22}$</td>
</tr>
<tr>
<td>s.e.</td>
<td>-1.05</td>
<td>0.07</td>
<td>-0.12</td>
<td>-0.08</td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>0.48</td>
<td>0.10</td>
<td>0.05</td>
<td>0.04</td>
<td>0.03</td>
</tr>
</tbody>
</table>

*Note: Coefficients significant at the 5% level formatted in bold.*
Figure 1 – 1% quantile for the overall index and Sumitomo Mitsui Financial Group

Note: The figure reports the average impulse-response function of European, North American and Asian financial institutions to a shock to the index.
Figure 3 – Geographical breakdown of VaR impulse-responses: shock to the financial institution

Note: The figure reports the average impulse-response function of the index to a shock to European, North American and Asian financial institutions.

Figure 4 – Breakdown by institution of VaR impulse-responses: shock to the index

Note: The figure reports the average impulse-response function by sector to a shock to the index.
Figure 5 – Breakdown by institution of VaR impulse-responses: shock to the bank

Note: The figure reports the average impulse-response function of the index to a shock to the financial institutions classified by sector.

Figure 6 – Strongest and weakest VaR impulse-responses

Note: The figure reports the average impulse-response function of the 20 financial institutions with the strongest and weakest impact.