VAR for VaR: Measuring Tail Dependence Using Multivariate Regression Quantiles*

Halbert White† Tae-Hwan Kim‡ Simone Manganelli§

August 8, 2012

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 equity returns data to analyse spillovers in the values at risk (VaR) between a market index and financial institutions. We construct impulse-response functions for the quantiles 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. We show how our methodology can successfully identify both in-sample and out-of-sample the set of financial institutions whose risk is most sentitive to market wide shocks in situations of financial distress, and can prove a valuable addition to the traditional toolkit of policy makers and supervisors.

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

JEL classification: C13, C14, C32.

*We would like to thank Peter Christoffersen, Rob Engle, Rossen Valkanov, as well seminar participants at the ECB, the Fourth Annual SoFiE Conference in Chicago, the Cass Business School, the London School of Business, the Fourth Tremblant Risk Management Conference, the Vienna Graduate School of Finance and the Fourteenth International Korean Economic Association Conference. Francesca Fabbri and Thomas Kostka provided data support. 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 authors and do not necessarily reflect those of the European Central Bank or the Eurosystem. Tae-Hwan Kim is grateful for the financial support from the National Research Foundation of Korea – a grant funded by the Korean Government (NRF-2009-327-B00088).

†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

This paper suggests a multivariate regression quantile model to directly study the degree of tail interdependence among different random variables. Our theoretical framework allows the quantiles of several random variables to depend on (lagged) quantiles, as well as past innovations and other covariates of interest. This modeling strategy has at least three advantages over the more traditional approaches that rely on the parameterization of the entire multivariate distribution. First, regression quantile estimates are known to be robust to outliers, a desirable feature in general and for applications to financial data in particular. Second, regression quantile is a semi-parametric technique and as such imposes minimal distributional assumptions on the underlying data generating process (DGP). Third, our multivariate framework allows researchers to directly measure the tail dependence among the random variables of interest, rather than recovering it indirectly via models of time-varying first and second moments.

To illustrate our approach and its usefulness, consider a simple set-up with two random variables, Y_{1t} and Y_{2t}. All information available at time t is represented by the information set $\mathcal{F}_{t-1}$. For a given level of confidence $\theta \in (0,1)$, the quantile q_{it} at time t for random variables Y_{it} $i = 1, 2$ conditional on $\mathcal{F}_{t-1}$ is

$$\Pr[Y_{it} \leq q_{it} | \mathcal{F}_{t-1}] = \theta, \qquad i = 1, 2. \tag{1}$$

A simple version of our proposed structure relates the conditional quantiles of the two random variables according to a vector autoregressive (VAR) structure:

$$\begin{aligned} q_{1t} &= X_t'\beta_1 + b_{11}q_{1t-1} + b_{12}q_{2t-1}, \\ q_{2t} &= X_t'\beta_2 + b_{21}q_{1t-1} + b_{22}q_{2t-1}, \end{aligned}$$

where X_t represents predictors belonging to $\mathcal{F}_{t-1}$ and typically includes lagged values of Y_{it}. If $b_{12} = b_{21} = 0$, the above model reduces to the univariate CAViaR model of Engle and Manganelli (2004), and the two specifications can be estimated independently from each other. If, however, b_{12} and/or b_{21} are different from zero, the model requires the joint estimation of all of the parameters in the system. The off-diagonal coefficients b_{12} and b_{21} represent the measure of tail codependence between the two random variables, thus the hypothesis of no tail codependence can be assessed by testing $H_0 : b_{12} = b_{21} = 0$.

The first part of this paper develops the consistency and asymptotic theory for the multivariate regression quantile model. Our fully general model is much richer than the above example, as we can accommodate: (i) more than two random variables; (ii) multiple lags of q_{it}; and (iii) multiple confidence levels, say $(\theta_1, ..., \theta_p)$.

In the second part of this paper, as an empirical illustration of the model, we estimate a series of bivariate VAR models for the conditional quantiles of the return distributions of individual financial institutions from around the world. Since quantiles represent one of the key inputs for the computation of the Value at Risk (VaR) for financial assets, we call this model VAR for VaR, that is, a vector autoregressive (VAR) model where the dependent variables are the VaR of the financial institutions, which are dependent on (lagged) VaR and past shocks.

Our modeling framework appears particularly suitable to develop sound measures of financial spillover, the importance of which has been brought to the forefront by the recent financial crisis. In the current policy debate, great emphasis has been put on how to measure the additional capital needed by financial institutions in a situation of generalized market distress. The logic is that if the negative externality associated with the spillover of risks within the system is not properly internalised, banks may find themselves in need of additional capital at exactly the worst time, such as when it is most difficult and expensive to raise fresh new capital. If the stability of the whole system is threatened, taxpayer money has to be used to backstop the financial system, to avoid systemic bank failures that may bring the whole economic system to a collapse.

Adrian and Brunnermeier (2009), Acharya et al. (2009), and Brownlees and Engle (2010) have recently proposed to classify financial institutions according to the sensitivity of their VaR to shocks to the whole financial system. The empirical section of this paper illustrates how the multivariate regression quantile model provides an ideal framework to estimate directly the sensitivity of VaR of a given financial institution to system-wide shocks. A useful by-product of our modeling strategy is the ability to compute quantile impulse-response functions. These are obtained by computing the long run quantiles and then applying a suitable identified shock to the multivariate quantile model. Using the quantile impulse-response functions, we can assess the resilience of financial institutions to shocks to the overall index, as well as their persistence.

The model is estimated on a sample of 230 financial institutions from around the world. For each of these equity return series, we estimate a bivariate VAR for VaR where one variable is the return on a portfolio of financial institutions and the other variable is the return on the single financial institution. We find strong evidence of significant tail codependence for a large fraction of the financial institutions in our sample. When aggregating the impulse response functions at the sectoral and geographic level no striking differences are revealed. We, however, find significant cross-sectional differences. By aggregating the 20 stocks with the strongest and weakest tail codependence to market shocks (thus, forming two portfolios), we find that, while in tranquil times, the two portfolios have comparable risk. In times of severe financial distress, the risk of the first portfolio increases disproportionately relative to the second. This result holds for both in-sample and out-of-sample.

The plan of this paper is as follows. In Section 2, we set forth the multivariate and multi-quantile CAViaR framework, a generalization of Engle and Manganelli's original CAViaR (2004) model. Section 3 provides conditions ensuring the consistency and asymptotic normality of the estimator, as well as the results which provide a consistent asymptotic covariance matrix estimator. Section 4 contains our empirical study. Section 5 provides a summary and some concluding remarks. The appendix contains all of the technical proofs of the theorems in the text.

2 The Multivariate and Multi-Quantile Process and Its Model

We consider a sequence of random variables denoted by $\{(Y_t', X_t') : t = 1, 2, ..., T\}$ where Y_t is a finitely dimensioned $n \times 1$ vector and X_t is also a countably dimensioned vector whose first element is one. To fix ideas, Y_t can be considered as the dependent variables and X_t as the explanatory variables in a typical regression framework. In this sense, the proposed model which will be developed below is sufficiently general enough to handle multiple dependent variables. We specify the data generating process as follows.

Assumption 1 The sequence $\{(Y_t', X_t')\}$ is a stationary and ergodic stochastic process on the complete probability space $(\Omega, \mathcal{F}, P_0)$, where Ω is the sample space, $\mathcal{F}$ is a suitably chosen σ-field, and P_0 is the probability measure providing a complete description of the stochastic behavior of the sequence of $\{(Y_t', X_t')\}$.

We define $\mathcal{F}_{t-1}$ to be the σ-algebra generated by $Z^{t-1} := \{X_t, (Y_{t-1}, X_{t-1}), ...\}$, i.e. $\mathcal{F}_{t-1} := \sigma(Z^{t-1})$. For $i = 1, ..., n$, we also define $F_{it}(y) := P_0[Y_{it} < y \mid \mathcal{F}_{t-1}]$ which is the cumulative distribution function (CDF) of Y_{it} conditional on $\mathcal{F}_{t-1}$. In the quantile regression literature, it is typical to focus on a specific quantile index; for example, $\theta \in (0, 1)$. In this paper, we will develop a more general quantile model where multiple quantile indexes can be accounted for jointly. To be more specific, we consider p quantile indexes denoted by $\theta_{i1}, \theta_{i2}, ..., \theta_{ip}$ for the ith element (denoted by Y_{it}) of Y_t. The p quantile indexes do not need to be the same for all of the elements of Y_t, which explains the double indexing of θ_{ij}. Moreover, we note that we specify the same number (p) of quantile indexes for each $i = 1, ..., n$. 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.

To formalize our argument, we assume that the quantile indexes are ordered such that $0 < \theta_{i1} < ... < \theta_{ip} < 1$. For $j = 1, ..., p$, the $\theta_{ij}th$ quantile of Y_{it} conditional on $\mathcal{F}_{t-1}$, denoted $q^*_{i,j,t}$, is

$$q^*_{i,j,t} := \inf\{y : F_{it}(y) = \theta_{ij}\}, \tag{2}$$

and if F_{it} is strictly increasing,

$$q^*_{i,j,t} = F^{-1}_{it}(\theta_{ij}).$$

Alternatively, $q^*_{i,j,t}$ can be represented as

$$\int_{-\infty}^{q^*_{i,j,t}} dF_{it}(y) = E[1_{[Y_{it} \leq q^*_{i,j,t}]} \mid \mathcal{F}_{t-1}] = \theta_{ij}, \tag{3}$$

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}.

Our objective is to jointly estimate the conditional quantile functions $q^*_{i,j,t}$ for $i = 1, ..., n$ and $j = 1, 2, ..., p$. For this, we write $q^*_t := (q^{*\prime}_{1,t}, q^{*\prime}_{2,t}, ..., q^{*\prime}_{n,t})'$ with $q^*_{i,t} := (q^*_{i,1,t}, q^*_{i,2,t}, ..., q^*_{i,p,t})'$ and impose an additional appropriate structure. First, we ensure that the conditional distributions of Y_{it} are everywhere continuous, with positive densities at each of the conditional quantiles of interest, $q^*_{i,j,t}$. We let f_{it} denote the conditional probability density function (PDF) which corresponds 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 $f_{i,t}(y)$. The 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 (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 $\mathbb{R}$, $t = 1, 2, ..., T$; (ii) For the given $0 < \theta_{i1} < ... < \theta_{ip} < 1$ and $\{q^*_{i,j,t}\}$ as defined above, we suppose the following: (a) for each i, j, t, and ω, $f_{it}(\omega, q^*_{i,j,t}(\omega)) > 0$; and (b) for the given finite integers k and m, there exist a stationary ergodic sequence of random $k \times 1$ vectors $\{\Psi_t\}$, with Ψ_t measurable$-\mathcal{F}_{t-1}$, and real vectors $\beta^*_{ij} := (\beta^*_{i,j,1}, ..., \beta^*_{i,j,k})'$ and $\gamma^*_{i,j,\tau} := (\gamma^{*\prime}_{i,j,\tau,1}, ..., \gamma^{*\prime}_{i,j,\tau,n})'$, where each $\gamma^*_{i,j,\tau,k}$ is a $p \times 1$ vector, such that for $i = 1, ..., n$, $j = 1, ..., p$, and all t,

$$q^*_{i,j,t} = \Psi'_t \beta^*_{ij} + \sum_{\tau=1}^{m} q^{*\prime}_{t-\tau} \gamma^*_{i,j,\tau} \ . \tag{4}$$

The structure of equation in (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 $\gamma^*_{i,j,\tau}$, 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 the single quantiles of the (lags of) other elements of Y_t. Thus, we call any process that satisfies our structure "Multivariate MQ-CAViaR" (MVMQ-CAViaR) processes or naively "VAR for VaR."

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 the given j, elements of $\gamma^*_{i,j,\tau}$ may be zero for values of τ greater than some given integer. For notational simplicity, we do not represent m as being dependent 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 Ψ_t may contain lagged values of Y_t, as well as measurable functions of X_t and lagged X_t. In particular, Ψ_t may contain Stinchcombe and White's (1998) GCR transformations, as discussed in White (2006).

For a particular quantile, say θ_{ij}, the coefficients to be estimated are β^*_{ij} and $\gamma^*_{ij} := (\gamma^{*\prime}_{i,j,1}, ..., \gamma^{*\prime}_{i,j,m})'$. Let $\alpha^{*\prime}_{ij} := (\beta^{*\prime}_{ij}, \gamma^{*\prime}_{ij})$, and write $\alpha^* = (\alpha^{*\prime}_{11}, ..., \alpha^{*\prime}_{1p}, ..., \alpha^{*\prime}_{n1}, ...,$
$\alpha^{*\prime}_{np})'$, an $\ell \times 1$ vector, where $\ell := np(k + npm)$. We call α^* the "MVMQ-CAViaR coefficient vector." We estimate α^* using a correctly specified model for the MVMQ-CAViaR process. First, we specify our model.

Assumption 3 (i) Let $\mathbb{A}$ be a compact subset of $\mathbb{R}^\ell$. For $i = 1, ..., n$, and $j = 1, ..., p$, we suppose the following: (a) the sequence of functions $\{q_{i,j,t} : \Omega \times \mathbb{A} \rightarrow \mathbb{R}^{p_i}\}$ is such that for each t and each $\alpha \in \mathbb{A}, q_{i,j,t}(\cdot, \alpha)$ is measurable$-\mathcal{F}_{t-1}$; (b) for each t and each $\omega \in \Omega$, $q_{i,j,t}(\omega, \cdot)$ is continuous on $\mathbb{A}$; and (c) for each i, j, and t, $q_{i,j,t}(\cdot, \alpha)$ is specified as follows:

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

Next, we impose the correct specification assumption together with an identification condition. Assumption 4(i.a) below delivers the correct specification by ensuring that the MVMQ-CAViaR coefficient vector α^* belongs to the parameter space, $\mathbb{A}$. This ensures that α^* optimizes the estimation objective function asymptotically. Assumption 4(i.b) delivers the identification by ensuring that α^* is the only 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 $\alpha^* \in \mathbb{A}$ such that for all i, j, t,

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

(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 $\alpha \in \mathbb{A}$ 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 the estimation of a sufficient number ($\#\mathcal{I}$) of the 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 μ alone, so there is only one "degree of freedom" for the quantile variation. Similarly, the quantiles of the 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 Asymptotic Theory

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

$$\min_{\alpha\in\mathbb{A}} \bar{S}_T(\alpha) := T^{-1}\sum_{t=1}^{T}\{\sum_{i=1}^{n}\sum_{j=1}^{p}\rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot,\alpha))\}, \tag{6}$$

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

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

as the quasi log-likelihood for the 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 can be regarded as a *quasi* maximum likelihood.

We establish consistency and asymptotic normality for $\hat{\alpha}_T$ through methods analogous to those of White, Kim, and Manganelli (2008). For conciseness, we place the remaining regularity conditions (i.e., Assumptions 5,6 and 7) and technical discussions in the appendix.

Theorem 1 Suppose that Assumptions 1, 2(i,ii), 3(i), 4(i) and 5(i,ii) old. Then, we have

$$\hat{\alpha}_T \overset{a.s.}{\rightarrow} \alpha^*.$$

With Q^* and V^* as given below, the asymptotic normality result is provided in the following theorem.

Theorem 2 Suppose that Assumptions 1-6 hold. Then, the asymptotic distribution of the QMLE estimator $\hat{\alpha}_T$ obtain from (6) is given by:

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

where

$$\begin{aligned} Q^* \quad &: = \sum_{i=1}^{n}\sum_{j=1}^{p} E[f_{i,j,t}(0)\nabla q_{i,j,t}(\cdot,\alpha^*)\nabla' q_{i,j,t}(\cdot,\alpha^*)], \\ V^* \quad &: = E(\eta_t^*\eta_t^{*\prime}), \\ \eta_t^* \quad &: = \sum_{i=1}^{n}\sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot,\alpha^*)\psi_{\theta_{ij}}(\varepsilon_{i,j,t}), \\ \varepsilon_{i,j,t} \quad &: = Y_{it} - q_{i,j,t}(\cdot,\alpha^*) \end{aligned}$$

We note that the transformed error term of $\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) = \theta_{ij} - 1_{[\varepsilon_{i,j,t}\leq 0]}$ appearing in Theorem 2 can be viewed as a generalized residual. To test restrictions on α^* or to obtain confidence intervals, we require a consistent estimator of the asymptotic covariance matrix $C^* := Q^{*-1}V^*Q^{*-1}$. First, we provide a consistent estimator $\hat{V}_T$ for V^*; then we propose a consistent estimator $\hat{Q}_T$ for Q^*. Once $\hat{V}_T$ and $\hat{Q}_T$ are proved to be consistent for V^* and Q^* respectively, then it follows by the continuous mapping theorem that $\hat{C}_T := \hat{Q}_T^{-1}\hat{V}_T\hat{Q}_T^{-1}$ is a consistent estimator for C^*.

A straightforward plug-in estimator of V^* is constructed as follows:

$$\begin{aligned} \hat{V}_T \quad &:= T^{-1}\sum_{t=1}^{T}\hat{\eta}_t\hat{\eta}_t', \\ \hat{\eta}_t \quad &:= \sum_{i=1}^{n}\sum_{j=1}^{p}\nabla q_{i,j,t}(\cdot,\hat{\alpha}_T)\psi_{\theta_{ij}}(\hat{\varepsilon}_{i,j,t}), \\ \hat{\varepsilon}_{i,j,t} \quad &:= Y_{it} - q_{i,j,t}(\cdot,\hat{\alpha}_T). \end{aligned}$$

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

Theorem 3 Suppose that Assumptions 1-6 hold. Then, we have the following result:

$$\hat{V}_T \overset{p}{\rightarrow} V^*.$$

Next, we provide a consistent estimator of Q^*. We follow Powell's (1984) suggestion of estimating $f_{i,j,t}(0)$ with $1_{[-\hat{c}_T \leq \hat{\varepsilon}_{i,j,t} \leq \hat{c}_T]}/2\hat{c}_T$ for a suitably chosen sequence $\{\hat{c}_T\}$. 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

$$\hat{Q}_T = (2\hat{c}_T T)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} \sum_{j=1}^{p} 1_{[-\hat{c}_T \leq \hat{\varepsilon}_{i,j,t} \leq \hat{c}_T]} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \nabla' q_{i,j,t}(\cdot, \hat{\alpha}_T).$$

Theorem 4 Suppose that Assumptions 1-7 hold. Then, we obtain the consistency result for $\hat{Q}_T$ as follows:

$$\hat{Q}_T \xrightarrow{p} Q^*.$$

There is no guarantee that $\hat{\alpha}_T$ is asymptotically efficient. There is now considerable literature that investigates the efficient estimation in quantile models; see, for example, Newey and Powell (1990), Otsu (2003), Komunjer and Vuong (2006, 2007a, 2007b). Thus far, this literature has only considered single quantile models. It is not obvious how the results for the single quantile models extend to multivariate and 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 equation (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 equation in (6) will improve the estimator efficiency for p and/or n greater than 1. We leave the study of the asymptotically efficient multivariate and multi-quantile estimator for future work.

4 Assessing Tail Reactions of Financial Institutions to System Wide Shocks

The financial crisis which started in 2007 has had a deep impact on the conceptual thinking of systemic risk among both academics and policy makers. There has been a recognition of the shortcomings of the measures that are tailored to dealing with institution-level risks. In particular, institution-level Value at Risk measures miss important externalities associated with the need to bail out systemically important banks in order to contain potentially devastating spillovers to the rest of the economy. Therefore, government and supervisory authorities may find themselves compelled to save ex post systemically important financial institutions, while these ignore ex ante any negative externalities associated with their behavior. There exists many contributions, both theoretical and empirical, as summarised, for instance, in Brunnermeier (2012) or Bisias et al. (2012). For the purpose of the application we have in mind, it is useful to structure the material around two early contributions, the CoVaR of Adrian and Brunnermeier (2009) and the systemic expected shortfall (SES) of Acharya et al. (2010).

Both measures aim to capture the risk of a financial institution conditional on a significant negative shock hitting another financial institution or the whole financial system. Formally, the $CoVaR_\theta^{j|i}$ is the VaR of financial institution j conditional on the event C that hits the financial institution i (denoted by C^i):

$$\Pr(y^j < CoVaR_\theta^{j|i} | C^i) = \theta.$$

The systemic expected shortfall is shown to be proportional to the marginal expected shortfall, which is analogously defined as:

$$MES_\theta^{j|i} = E(y^j | C^i).$$

The main difference is that the expectation of the whole left tail, rather than just the quantile, is considered. In practice, loss distributions in the tail are extremely hard to estimate. One strategy is to standardize the returns by the estimated volatility or quantiles, and then apply non-parametric techniques, as done for instance in Manganelli and Engle (2002) or Brownlees and Engle (2010). An alternative is to use extreme value theory to impose a parametric structure on the tail behavior, as done in Hartmann et al. (2004).

As we will show in the rest of this section, the theoretical framework developed in this paper lends itself to a coherent modeling of the dynamics of the tail interdependence implicit in both the CoVaR and systemic expected shortfall measures. Unlike standard GARCH based approaches, which require the modeling of the entire multivariate distribution, the advantage of our multivariate regression quantiles framework - besides providing a robust, semi-parametric technique which does not rely on strong distributional assumptions - is that it is tailored to directly model the object of interest.

In this section, we apply our model to study the spillover that occur in the equity return quantiles of a sample of 230 financial institution around the world. We first describe our empirical model and show how to compute impulse-response functions within the multivariate and multi-quantile framework. We next present the data and the optimization strategy. Finally, we discuss the empirical findings.

4.1 Empirical specification

The specification we use in our empirical analysis is the following simple bivariate quantile model:

$$q_t = c + A|Y_{t-1}| + Bq_{t-1}, \tag{7}$$

where q_t, Y_{t-1}, and c are 2-dimensional vectors, and A, B are (2,2)-matrices. The parameters can be consistently estimated by minimizing the multivariate regression quantile objective function (6). It is straightforward to derive an estimate of the CoVaR from this model. For instance, if the conditioning event C^i is defined as $Y_{2,t-1} = q_{2,t-1}$, that is financial institution 2 is hit by a shock equal to its quantile, the associated CoVaR for financial institution 1 is given by $q_{1,t} = c_1 + a_{11}|Y_{1t-1}| + a_{12}|q_{2,t-1}| + b_{11}q_{1,t-1} + b_{12}q_{2,t-1}$.[1]

A DGP consistent with (7) can be obtained assuming that the data are structurally generated as:

$$Y_t = L_t \varepsilon_t, \tag{8}$$

where $L_t := L_t(Z^{t-1})$ is an $\mathcal{F}_t-$measurable lower triangular matrix and the elements of $\varepsilon_t := [\varepsilon_{1t}, \varepsilon_{2t}]'$ are mutually independent with $\{\varepsilon_t|\mathcal{F}_t\}$ being a martingale difference sequence.

A suitable choice of L_t ensuring that the conditional quantiles of Y_t obey (4) and (7) is the following bivariate model:

$$\begin{bmatrix} Y_{1t} \\ Y_{2t} \end{bmatrix} = \begin{bmatrix} \alpha_t & 0 \\ \beta_t & \gamma_t \end{bmatrix} \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{bmatrix},$$

where ε_t is the bivariate standard normal random variable. Note that the standard deviations of Y_{1t} and Y_{2t} are given by $\sigma_t(Y_{1t}) = \alpha_t$ and $\sigma_t(Y_{2t}) = \sqrt{\beta_t^2 + \gamma_t^2}$ respectively. Suppose that α_t, β_t and γ_t are specified in such a way to satisfy the following:

$$\begin{aligned} \sigma_t(Y_{1t}) &= \tilde{c}_1 + \tilde{a}_{11}|Y_{1t-1}| + \tilde{a}_{12}|Y_{2t-1}| + b_{11}\sigma_{1t-1} + b_{12}\sigma_{2t-1}, \\ \sigma_t(Y_{2t}) &= \tilde{c}_2 + \tilde{a}_{21}|Y_{1t-1}| + \tilde{a}_{22}|Y_{2t-1}| + b_{21}\sigma_{1t-1} + b_{22}\sigma_{2t-1}. \end{aligned}$$

The respective quantile processes associated with this DGP are given by:

$$\begin{aligned} q_{1t} &= k\tilde{c}_1 + k\tilde{a}_{11}|Y_{1t-1}| + k\tilde{a}_{12}|Y_{2t-1}| + b_{11}q_{1t-1} + b_{12}q_{2t-1}, \\ q_{2t} &= k\tilde{c}_2 + k\tilde{a}_{21}|Y_{1t-1}| + k\tilde{a}_{22}|Y_{2t-1}| + b_{21}q_{1t-1} + b_{22}q_{2t-1}, \end{aligned}$$

where k is the θ^{th}-quantile of the standard normal distribution. In matrix form, the above system can be rewritten as in (7), where $c_i = k\tilde{c}_i$ and $a_{ij} = k\tilde{a}_{ij}$.

---

[1] $Y_{i,t}$ denotes the i-th element of the vector Y_t, while a_{ij} is the element of the i-th row and j-th column of the matrix A. Similar notation is used to denote the individual elements of the vectors and matrices q_t, c and B.

In the empirical application, we impose the following identification assumption:

Identification Assumption: The first element of Y_t in (8) denotes the per-period return on a financial index and the second element is the per-period return on a specific financial institution within the index.

The identification assumption behind this decomposition is that shocks to the financial 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 index. Here, we limit ourselves to a bivariate system, as we are interested in the interaction between a financial index and an individual financial institution. The theoretical framework of this paper can accommodate higher dimensional models, although at the cost of increasing the computational burden.

Incidentally, this identification scheme illustrates the potential pitfalls of choosing appropriate conditioning events for the CoVaR measures. Defining the conditioning event C^i as $Y_{2,t-1} = q_{2,t-1}$, as done before, neglects the fact that shock to the financial institution 2 may be correlated with that of other financial institutions, therefore producing a potentially misleading classification of the systemic importance of financial institutions. In the following, based on the identification scheme which is implicit in the triangular structure of L_t, our attention will focus mainly on the impact of market shocks on the VaR of individual financial institutions.

4.1.1 The Quantile Impuse Response Function (QIRF)

This modeling framework allows us to proceed a step further beyond the scope of the static analysis implicit in the CoVaR framework because we can introduce the concept of impulse-response functions for quantiles. To see how quantile impulse-response functions can be defined, assume that in the DGP (8) there is a shock δ (or intervention) to ε_{1t} only at time t so that $\tilde{\varepsilon}_{1t} \equiv \varepsilon_{1t} + \delta$ (since ε_t is assumed to be the standard normal, it is actually $\delta-$standard deviation). In all other times there is no intervention. In other words, the time path of the error terms without the intervention would be

$$\{..., \varepsilon_{1t-2}, \varepsilon_{1t-1}, \varepsilon_{1t}, \varepsilon_{1t+1}, \varepsilon_{1t+2}, ...\}$$

while the time path with the intervention would be

$$\{..., \varepsilon_{1t-2}, \varepsilon_{1t-1}, \tilde{\varepsilon}_{1t}, \varepsilon_{1t+1}, \varepsilon_{1t+2}, ...\}.$$

Our objective is the measure the impact of the one-off intervention at time t on the quantile dynamics.

We first consider the time path of Y_t without and with the intervention. The affected Y_t will be denoted as $\tilde{Y}_t$. Note that the intervention δ at time

t will only affect Y_t, but not the future realizations of Y_{t+s} $(s = 1, 2, 3, ...)$, due to the way we specify our model. Hence, the time path of $\tilde{Y}_t$ with the intervention is given by

$$\{..., Y_{t-2}, Y_{t-1}, \tilde{Y}_t, Y_{t+1}, Y_{t+2}, ...\}.$$

In fact, the difference between the origainal series and the affected series is zero for $n \geq 1$. Furthermore note that $\tilde{Y}_t = L_t[\tilde{\varepsilon}_{1t,}\varepsilon_{2t}]'$ because of (8).

The θ^{th} quantile impulse-response function (QIRF) for the i^{th} variable (Y_{it}) denoted as $\Delta_{i,s}(\tilde{\varepsilon}_{1t})$ is defined as

$$\Delta_{i,s}(\tilde{\varepsilon}_{1t}) = \tilde{q}_{i,t+s} - q_{i,t+s} \qquad\qquad s = 1, 2, 3, ...$$

where $\tilde{q}_{i,t+s}$ is the θ^{th} conditional quantile of the affected series $(\tilde{Y}_{it+s})$ and $q_{i,t+s}$ is the θ^{th} conditional quantile of the unaffected series (Y_{it+s}).

First, we consider the case for $i = 1$, i.e. $\Delta_{1,s}(\tilde{\varepsilon}_{1t})$. When $s = 1$, the QIRF is given by

$$\Delta_{1,1}(\tilde{\varepsilon}_{1t}) = a_{11}(|\tilde{Y}_{1t}| - |Y_{1t}|) + a_{12}(|\tilde{Y}_{2t}| - |Y_{2t}|).$$

For $s > 1$, the QIRF is given by

$$\Delta_{1,s}(\tilde{\varepsilon}_{1t}) = b_{11}\Delta_{1,s-1}(\tilde{\varepsilon}_{1t}) + b_{12}\Delta_{2,s-1}(\tilde{\varepsilon}_{1t}).$$

The case for $i = 2$ is similarly obtained as follows. For $s = 1$,

$$\Delta_{2,1}(\tilde{\varepsilon}_{1t}) = a_{21}(|\tilde{Y}_{1t}| - |Y_{1t}|) + a_{22}(|\tilde{Y}_{2t}| - |Y_{2t}|),$$

while for $s > 1$,

$$\Delta_{2,s}(\tilde{\varepsilon}_{1t}) = b_{21}\Delta_{1,s-1}(\tilde{\varepsilon}_{1t}) + b_{22}\Delta_{2,s-1}(\tilde{\varepsilon}_{1t}).$$

Now, let us define

$$\Delta_s(\tilde{\varepsilon}_{1t}) := \begin{bmatrix} \Delta_{1,s}(\tilde{\varepsilon}_{1t}) \\ \Delta_{2,s}(\tilde{\varepsilon}_{1t}) \end{bmatrix},$$

and

$$\begin{aligned} D_t &= |\tilde{Y}_t| - |Y_t| \\ &= |L_t\tilde{\varepsilon}_t^1| - |L_t\varepsilon_t|. \end{aligned}$$

where $\tilde{\varepsilon}_t^1 = [\tilde{\varepsilon}_{1t}, \varepsilon_{2t}]'$. Then, we can show that the QIRF is compactly expressed as follows:

$$\begin{aligned} \Delta_s(\tilde{\varepsilon}_{1t}) &= AD_t \\ \Delta_s(\tilde{\varepsilon}_{1t}) &= B^{(s-1)}AD_t \quad \text{for } s > 1. \end{aligned}$$

The QIRF when there is a shock (or intervention) to ε_{2t} only at time t can be analogously obtained.

In the empirical application, the matrix L_t - needed to compute the impact of $\tilde{\varepsilon}_{1t}$ on Y_t - is estimated using a standard Cholesky decomposition.

4.2 Data and Optimization Strategy

The data used in this section have been 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. Prices were transformed into continuously compounded log returns, giving an estimation sample size of 2765 observations. We use 453 additional observations up to 2 May 2012, for the out-of-sample exercises. We eliminated all the stocks whose times series started later than 1 January 2000, or which stopped after this date. At the end of this process, we were left with 230 stocks.

Table 1 reports the names of the financial institutions in our sample, together with the country of origin and the sector they are associated with, as from Datastream classification. Table 2 shows the breakdown of the stocks by sector and by geographic area. There are twice as many financial institutions classified as banks in our sample relative to those classified as financial services or insurances. The distribution across geographic areas is more balanced, with a greater number of EU financial institutions and a slightly lower Asian representation. The proxy for the market index used in each bivariate quantile estimation is the equally weighted average of all the financial institutions in the same geographic area, in order to avoid asynchronicity issues.

We estimated 230 bivariate 1% quantile models between the market index and each of the 230 financial institutions in our sample. Each model is estimated using, as starting values in the optimization routine, 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 40 initial conditions, we minimized the regression quantile objective function (6) using the *fminsearch* optimization function in Matlab, which is based on the Nelder-Mead simplex algorithm. Finally, among the resulting 40 vectors of the parameter estimates, we chose the vector yielding the lowest value for the function (6). We adopt this strategy because we have found that parameter estimates are sometimes sensitive to the choices of the initial conditions (possibly due to a flat likelihood near the optimum). Such an optimization strategy is more time consuming, but delivers more reliable results. In calculating the standard errors, we have set the bandwidth to 1 throughout the sample.

4.3 Results

Table 3 reports, as an example, the estimation results for four well-known financial institutions: Barclays, Deutsche Bank, Citigroup and Goldman Sachs. The diagonal autoregressive coefficients for the B matrix are around 0.90 and all of them are statistically significant, which indicates the VaR processess are significantly autocorrelated. These findings are consistent with what is typically found in the literature using CAViaR models. Notice, however, that some of the non-diagonal coefficients for the A and B matrices 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 5% level for 142 financial institutions out of the 230 in our sample. The resulting estimated 1% quantiles for Barclays, Deutsche Bank, Citigroup and Goldman Sachs are reported in Figure 1. The quantile plots clearly reveal the generalized sharp increase in risk following the Lehman bankruptcy. Careful inspection of the plots also reveals a noticeable cross-sectional difference, with the risk for Goldman Sachs being contained to less than half the risk of Citigroup at the height of the crisis.

The methodology introduced in this paper, however, allows us to go beyond the analysis of the univariate quantiles, and directly looks at the tail codependence between financial institutions and the market index. Figure 2 displays the impulse response of the risks of the four financial institutions to a 2 standard deviation shock to the market index (see the discussion in the previous sub-section for a detailed explanation of how the impulse-response functions are computed). The horizontal axis measures the time (expressed in days), while the vertical axis measures the change in the 1% quantiles of the individual financial institutions (expressed in percentage returns) as a reaction to the market shock. The impulse response functions track how this shock propagates through the system and how long it takes to absorb it. The shock is completely reabsorbed after the impulse response function has converged again to zero.

Looking more closely at the impulse response functions of the four selected financial institutions reveals a few differences in how their long run risks react to shocks. For instance, Deutsche Bank and Barclays have a sharp initial reaction. However, while the shock to Deutsche Bank's VaR is entirely absorbed after around 35 days, the shock to Barclays' risk appear to be more persistent, with its effect not being completely absorbed after more than 50 days. Similarly, risk shocks on Citigroup's VaR appears to be long lasting, while Goldman Sachs quantiles overall exhibit very little tail correlation with the market.

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 non trivial 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, Figure 3 plots the average impulse-response functions $\Delta_{1,s}(\tilde{\varepsilon}_{2t})$ and $\Delta_{2,s}(\tilde{\varepsilon}_{1t})$ 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 the left column, the average is taken with respect to the geographical distribution. That is, the average impulse-response for Europe, for example, is obtained by averaging all the impulse-response functions for the European financial institutions. We notice two things. First, the impact of a shock to the index (charts in the top row) is much stronger than the impact of a shock to the individual financial institution (charts in the bottom row). 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 somewhat less sensitive to global shocks than their European and North American counterparts.

The charts on the right column of Figure 3 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.

Overall, however, it is fair to say that no big differences can be noted among the impulse responses by aggregating over the geographic or sectorial dimension. To highlight the still sizeable cross-sectional difference, and to get an idea of the orders of magnitude involved, we ranked the financial institutions by their overall risk impact by integrating out all the individual impulse-responses. Figure 4 plots the average impulse-responses which correspond to the 20 financial institutions whose risk is most and least sensitive to market shocks, together with those of the largest and smallest impact on the risk of the market risk index. It is clear that the shocks to the index have an impact of an order of magnitude greater than that of the shocks to the individual financial institutions. A two standard deviation shock to the index produces an average initial increase in the daily VaR of the most sensitive financial institutions at more than 3%. The shock is also quite persistent, as it is not yet completely absorbed after 50 days. 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 1%, which is then entirely absorbed after the third week.

To gauge what extent the model correctly identifies the financial institutions whose risks are most exposed to market shocks, Figure 5 plots the average quantiles of the two sets of financial institutions identified in Figure 4. Specifically, the charts in the top panels of the figure, track the estimated in-sample quantiles and stock price developments of the 20 financial institutions which have been identified in Figure 4 as being most and least exposed to market shocks. For comparison, we have added the risk and price developments of the market index. Prices have been normalized to 100 at the beginning of the sample. The charts in the bottom panels replicate the same exercise with the out-of-sample data.

The figure presents two striking facts. First, during normal times, i.e. between 2004 and mid-2007, the quantiles of the most and least sensitive groups of financial institutions is roughly equal. Actually, there are some periods in 2003 in which the quantiles of the least sensitive financial institutions exceeded the quantiles of the most sensitive ones. The second striking fact is that the situation changes abruptly in periods of market turbulence. For instance, at the beginning of the sample, in 2001-2003, the quantiles of the most sensitive financial institutions increased significantly more than that of the least sensitive ones. The change in behavior during crisis periods is even more striking from 2008 onwards, showing a greater exposure to common shocks. The chart on price developments on the right hand side confirms that the group of financial institutions identified as the most sensitive to market shocks are those whose stock market value dropped the most during the Lehman crisis, with their values dropping on 18 February 2009 by more than 90% with respect to the beginning of the sample. In contrast, the values of the least sensitive financial institutions have remained relatively stable throughout the sample, and in particular, they suffer only minor losses at the height of the crisis. The bottom panels reveal that the same result holds for the out-of-sample period. Of particular notice is the sharp drop in the out-of-sample quantile for the group of the most sensitive financial institutions which occurred on 12 August, 2011, the beginning of the second phase of the euro area sovereign debt crisis.

This application illustrates how the proposed methodology can usefully inform policy makers by helping identify the set of financial institutions which may be most exposed to common shocks, especially in times of crisis. Of course, this should only be considered as a partial model-based screening device for the identification of the most systemic banks. Further analysis, market intelligence and sound judgment are other necessary elements to produce a reliable risk assessment method for the larger and more complex financial groups.

Again, we emphasize that the results presented in these figures merely summarize the pattern of the results found in the bivariate analysis of our

230 financial institutions. Cross-comparisons could be improved by estimating for instance a 3- or 4- or $n-$variate system using a common information set, or adopting an appropriate factor structure which would minimize the number of parameters to be estimated. 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 the computational tractability.

5 Conclusion

We have developed a theory ensuring the consistency and asymptotic normality of multivariate and 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 conducted an empirical analysis in which we estimate a vector autoregressive model for the Value at Risk – VAR for VaR – 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 the financial institutions' long run risk reactions to shocks to the overall index. Judging from our bivariate models, we found that the risk of Asian financial institutions tend to be less sensitive to system wide shocks, whereas insurance companies exhibit a greater sensitivity to global shocks. By looking at the integral of all the individual impulse-responses, we found wide differences on how financial institutions react to different shocks. Both in-sample and out-of-sample analyses reveal that financial institutions with the strongest impulse-responses to global shocks are those which suffer the most in periods of market turbulence.

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

[1] Acharya, V., Pedersen, L., Philippe, T., and Richardson, M. (2010). Measuring Systemic Risk. Technical report, Department of Finance, NYU.

[2] Adrian, T. and Brunnermeier, M. (2009). CoVaR. Manuscript, Princeton University.

[3] Andersen, T. G., Bollerslev, T., Diebold, F. X., and Labys, P. (2003). Modeling and Forecasting Realized Volatility. *Econometrica* 71, 579–625.

[4] Andrews, D.W.K. (1988). Laws of Large Numbers for Dependent Non-Identically Distributed Random Variables. *Econometric Theory* 4, 458-467.

[5] Bartle, R. (1966). *The Elements of Integration.* New York: Wiley.

[6] Bisias, D., Flood, M., Lo, A.W., and Valavanis, S. (2012). A Survey of Systemic Risk Analyics. Office of Financial Research Working Paper.

[7] Brownlees, C.T. and Engle, R.F. (2011). Volatility, Correlation and Tails for Systemic Risk Measurement. Manuscript, Stern School of Business, New York University.

[8] Brunnermeier, M. and Oehmke, M. (2012). Bubbles, Financial Crises, and Systemic Risk. Handbook of Finance and Economics, forthcoming.

[9] Engle, R.F., Hendry, D.F., and Richard, J.-F. (1983). Exogeneity. *Econometrica* 51, 277–304.

[10] Engle, R.F. and Manganelli, S. (2004). CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business & Economic Statistics* 22, 367-381.

[11] Hartmann, P., Straetmans, S., and de Vries, C.G. (2004). Asset Market Linkages in Crisis Periods. *Review of Economics and Statistics* 86, 313-326.

[12] Huber, P.J. (1967). The Behavior of Maximum Likelihood Estimates under Nonstandard Conditions. *Proceedings of the Fifth Berkeley Symposium on Mathematical Statistics and Probability.* Berkeley: University of California Press, pp. 221–233.

[13] Kim, T.H. and White, H. (2003). Estimation, Inference, and Specification Testing for Possibly Misspecified Quantile Regression. In T. Fomby and C. Hill, eds., *Maximum Likelihood Estimation of Misspecified Models: Twenty Years Later.* New York: Elsevier, pp. 107-132.

[14] Koenker, R. and Bassett, G. (1978). Regression Quantiles. *Econometrica* 46, 33–50.

[15] Komunjer, I. (2005). Quasi-Maximum Likelihood Estimation for Conditional Quantiles. *Journal of Econometrics* 128, 127-164.

[16] Komunjer, I. and Vuong, Q. (2006). Efficient Conditional Quantile Estimation: the Time Series Case. University of California, San Diego Department of Economics Discussion Paper 2006-10, .

[17] Komunjer, I. and Vuong, Q. (2007a). Semiparametric Efficiency Bound and M-Estimation in Time-Series Models for Conditional Quantiles. University of California, San Diego Department of Economics Discussion Paper.

[18] Komunjer, I. and Vuong, Q. (2007b). Efficient Estimation in Dynamic Conditional Quantile Models. University of California, San Diego Department of Economics Discussion Paper.

[19] Manganelli, S. and Engle, R.F. (2002). Value at Risk Models in Finance. ECB Working Paper No. 75.

[20] Newey, W.K. and Powell, J.L. (1990). Efficient Estimation of Linear and Type I Censored Regression Models under Conditional Quantile Restrictions. *Econometric Theory* 6, 295-317.

[21] Otsu, T. (2003). Empirical Likelihood for Quantile Regression. University of Wisconsin, Madison Department of Economics Discussion Paper.

[22] Powell, J. (1984). Least Absolute Deviations Estimators for the Censored Regression Model. *Journal of Econometrics* 25, 303-325.

[23] Stinchcombe, M. and White, H. (1998). Consistent Specification Testing with Nuisance Parameters Present only under the Alternative. *Econometric Theory* 14, 295-324.

[24] Weiss, A. (1991). Estimating Nonlinear Dynamic Models Using Least Absolute Error Estimation. *Econometric Theory* 7, 46-68.

[25] White, H. (1994). *Estimation, Inference and Specification Analysis.* New York: Cambridge University Press.

[26] White, H. (2001). *Asymptotic Theory for Econometricians.* San Diego: Academic Press.

[27] White, H. (2006). Approximate Nonlinear Forecasting Methods. In G. Elliott, C.W.J. Granger, and A. Timmermann, eds., *Handbook of Economic Forecasting.* New York: Elsevier, pp. 460-512.

[28] White, H., Kim, T.H., and Manganelli, S. (2008). Modeling Autoregressive Conditional Skewness and Kurtosis with Multi-Quantile CAViaR. In J. Russell and M. Watson, eds., *Volatility and Time Series Econometrics: A Festschrift in Honor of Robert F. Engle*.

Appendix

We establish the consistency of $\hat{\alpha}_T$ by applying the 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 $i = 1, ..., n$, $E|Y_{it}| < \infty$; (ii) let $D_{0,t} := \max_{i=1,...,n} \max_{j=1,...,p} \sup_{\alpha \in \mathbb{A}} |q_{i,j,t}(\cdot, \alpha)|$. 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 $\varphi_t(Y_t, q_t(\cdot, \alpha)) := -\{\sum_{i=1}^{n} \sum_{j=1}^{p} \rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))\}$. Assumption 1 ensures White's Assumption 2.1. Assumption 3(i) ensures White's Assumption 5.1. Our choice of $\rho_{\theta_{ij}}$ 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 α 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

$$\begin{aligned}
|\varphi_t(Y_t, q_t(\cdot, \alpha))| &\leq \sum_{i=1}^{n} \sum_{j=1}^{p} |\rho_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))| \\
&= \sum_{i=1}^{n} \sum_{j=1}^{p} |(Y_{it} - q_{i,j,t}(\cdot, \alpha))(\theta_{ij} - 1_{[Y_{it} - q_{i,j,t}(\cdot, \alpha) \leq 0]})| \\
&\leq 2 \sum_{i=1}^{n} \sum_{j=1}^{p} (|Y_{it}| + |q_{i,j,t}(\cdot, \alpha)|) \\
&\leq 2p \sum_{i=1}^{n} |Y_{it}| + 2np|D_{0,t}|,
\end{aligned}$$

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).

White's Assumption 3.2 remains to be verified; here, this is the condition that α^* is the unique maximizer of $E(\varphi_t(Y_t, q_t(\cdot, \alpha))$. Given Assumptions 2(ii.b) and 4(i), it follows through the argument that directly parallels to that of the proof by 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 $\mathcal{F}_{t-1}$. Noting that $\delta_{i,j,t}(\alpha, \alpha^*) := q_{i,j,t}(\cdot, \alpha) - q_{i,j,t}(\cdot, \alpha^*)$, we next can show through some algebra and Assumption 2(ii.a) that

$$\begin{aligned}
E(\Delta_{i,j,t}(\alpha)) &= E[\int_0^{\delta_{i,j,t}(\alpha,\alpha^*)} (\delta_{i,j,t}(\alpha, \alpha^*) - s)\ f_{i,j,t}(s) ds] \\
&\geq E[\frac{1}{2}\delta_\epsilon^2 1_{[|\delta_{i,j,t}(\alpha,\alpha^*)| > \delta_\epsilon]} + \frac{1}{2}\delta_{i,j,t}(\alpha, \alpha^*)^2 1_{[|\delta_{i,j,t}(\alpha,\alpha^*)| \leq \delta_\epsilon]})] \\
&\geq \frac{1}{2}\delta_\epsilon^2 E[1_{[|\delta_{i,j,t}(\alpha,\alpha^*)| > \delta_\epsilon]}].
\end{aligned}$$

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,

$$\begin{aligned}
\Delta(\alpha) &:= \sum_{i=1}^{n} \sum_{j=1}^{p} E(\Delta_{i,j,t}(\alpha)) \geq \frac{1}{2}\delta_\epsilon^2 \sum_{i=1}^{n} \sum_{j=1}^{p} E[1_{[|\delta_{i,j,t}(\alpha,\alpha^*)| > \delta_\epsilon]}] \\
&= \frac{1}{2}\delta_\epsilon^2 \sum_{i=1}^{n} \sum_{j=1}^{p} P[|\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon] \geq \frac{1}{2}\delta_\epsilon^2 \sum_{(i,j) \in \mathcal{I}} P[|\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon] \\
&\geq \frac{1}{2}\delta_\epsilon^2 P[\cup_{(i,j) \in \mathcal{I}} \{|\delta_{i,j,t}(\alpha, \alpha^*)| > \delta_\epsilon\}] > 0,
\end{aligned}$$

where the final inequality follows from Assumption 4(i.b). As α 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

$$T^{-1} \sum_{t=1}^{T} \{\sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T)\ \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T))\} = o_p(T^{1/2}), \qquad (9)$$

where $\nabla q_{i,j,t}(\cdot,\alpha)$ is the $\ell\times 1$ gradient vector with elements $(\partial/\partial\alpha_s)q_{i,j,t}(\cdot,\alpha)$, $s = 1,...,\ell$, 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 equation (9) 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 α^* interior to $\mathbb{A}$, we can apply the mean value theorem to obtain

$$\lambda(\alpha) = \lambda(\alpha^*) + Q_0(\alpha - \alpha^*), \tag{10}$$

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

$$Q_0 = -Q^* + O(||\alpha - \alpha^*||), \tag{11}$$

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 $f_{i,j,t}(0)$ representing the value at zero of the density $f_{i,j,t}$ of $\varepsilon_{i,j,t} := Y_{it} - q_{i,j,t}(\cdot,\alpha^*)$, conditional on $\mathcal{F}_{t-1}$. Combining equations (10) and (11) and putting $\lambda(\alpha^*) = 0$, we obtain

$$\lambda(\alpha) = -Q^*(\alpha - \alpha^*) + O(||\alpha - \alpha^*||^2). \tag{12}$$

The next step is to show that

$$T^{1/2}\lambda(\hat{\alpha}_T) + H_T = o_p(1), \tag{13}$$

where $H_T := T^{-1/2}\sum_{t=1}^{T}\eta_t^*$, with $\eta_t^* := \sum_{i=1}^{n}\sum_{j=1}^{p}\nabla q_{i,j,t}(\cdot,\alpha^*)\psi_{\theta_{ij}}(\varepsilon_{i,j,t})$. Equations (12) and (13) then yield the following asymptotic representation of our estimator $\hat{\alpha}_T$:

$$T^{1/2}(\hat{\alpha}_T - \alpha^*) = Q^{*-1}T^{-1/2}\sum_{t=1}^{T}\eta_t^* + o_p(1). \tag{14}$$

As we impose conditions sufficient to ensure that $\{\eta_t^*, \mathcal{F}_t\}$ is a martingale difference sequence (MDS), a suitable central limit theorem (e.g., theorem 5.24 in White, 2001) is applied to equation (14) to yield the desired asymptotic normality of $\hat{\alpha}_T$:

$$T^{1/2}(\hat{\alpha}_T - \alpha^*) \xrightarrow{d} N(0, Q^{*-1}V^*Q^{*-1}), \tag{15}$$

where $V^* := E(\eta_t^*\eta_t^{*\prime})$.

We now strengthen the conditions given in the text 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 α.

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

To exploit the mean value theorem, we require that α^* belongs to $int(\mathbb{A})$, the interior of $\mathbb{A}$.

Assumption 4 (ii) $\alpha^* \in int(\mathbb{A})$.

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

Assumption 5 (iii) Let $D_{1,t} := \max_{i=1,\ldots,n} \max_{j=1,\ldots,p} \max_{s=1,\ldots,\ell} \sup_{\alpha \in \mathbb{A}} |(\partial/\partial\alpha_s)q_{i,j,t}(\cdot, \alpha)|$. Then (a) $E(D_{1,t}) < \infty$; (b) $E(D_{1,t}^2) < \infty$; (iv) Let us define

$$D_{2,t} := \max_{i=1,\ldots,n} \max_{j=1,\ldots,p} \max_{s=1,\ldots,\ell} \max_{h=1,\ldots,\ell} \sup_{\alpha \in \mathbb{A}} |(\partial^2/\partial\alpha_s\partial\alpha_h)q_{i,j,t}(\cdot, \alpha)|.$$

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) $V^* := E(\eta_t^* \eta_t^{*\prime})$ is positive definite.

Assumptions 3(ii) and 5(iii.a) are additional assumptions that help to ensure that equation (9) holds. Further imposing Assumptions 2(iii), 3(iii.a), 4(ii), and 5(iv.a) suffices to ensure that equation (12) holds. The additional regularity provided by Assumptions 5(iii.b), 5(iv.b), and 6(i) ensures that equation (13) holds. Assumptions 5(iii.b) and 6(ii) help ensure the availability of the MDS central limit theorem. We now have conditions that are sufficient to prove the asymptotic normality of our MVMQ-CAViaR estimator.

Proof of Theorem 2 As outlined above, we first prove

$$T^{-1} \sum_{t=1}^{T} \{ \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T) \; \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)) \} = o_p(1). \qquad (16)$$

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

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

where the last inequality follows from 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 (16) is proven.

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

$$\begin{aligned} \lambda(\alpha) &:= \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha) \int_{\delta_{i,j,t}(\alpha, \alpha^*)}^{0} f_{i,j,t}(s) ds], \end{aligned}$$

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

$$\lambda(\alpha) = \lambda(\alpha^*) + Q_0(\alpha - \alpha^*), \tag{17}$$

for α in a convex compact neighborhood of α^*, where Q_0 is an $\ell \times \ell$ matrix with $(1 \times \ell)$ rows $Q_s(\bar{\alpha}_{(s)}) = \nabla' \lambda(\bar{\alpha}_{(s)})$, where $\bar{\alpha}_{(s)}$ is a mean value (different for each s) lying on the segment connecting α and α^* with $s = 1, ..., \ell$. 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

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

where

$$\begin{aligned} A_s(\alpha) &:= \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla_s \nabla' q_{i,j,t}(\cdot, \alpha) \int_{\delta_{i,j,t}(\alpha,\alpha^*)}^{0} f_{i,j,t}(s)ds] \\ B_s(\alpha) &:= \sum_{i=1}^{n} \sum_{j=1}^{p} E[f_{i,j,t}(\delta_{i,j,t}(\alpha, \alpha^*)) \nabla_s q_{i,j,t}(\cdot, \alpha) \nabla' q_{i,j,t}(\cdot, \alpha)]. \end{aligned}$$

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

$$\begin{aligned} A_s(\bar{\alpha}_{(s)}) &= O(||\alpha - \alpha^*||) \\ B_s(\bar{\alpha}_{(s)}) &= Q_s^* + O(||\alpha - \alpha^*||), \end{aligned}$$

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

$$Q_0 = -Q^* + O(||\alpha - \alpha^*||). \tag{18}$$

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

$$\begin{aligned} \lambda(\alpha^*) &= \sum_{i=1}^{n} \sum_{j=1}^{p} E[\nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*))] \\ &= \sum_{i=1}^{n} \sum_{j=1}^{p} E(E[\nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*)) \mid \mathcal{F}_{t-1}]) \end{aligned}$$

$$\begin{aligned} &= \sum_{i=1}^{n} \sum_{j=1}^{p} E(\nabla q_{i,j,t}(\cdot, \alpha^*) E[\psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*)) \mid \mathcal{F}_{t-1}]) \\ &= \sum_{i=1}^{n} \sum_{j=1}^{p} E(\nabla q_{i,j,t}(\cdot, \alpha^*) E[\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}]) \\ &= 0, \end{aligned}$$

as $E[\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}] = \theta_{ij} - E[1_{[Y_{it} \leq q^*_{i,j,t}]} \mid \mathcal{F}_{t-1}] = 0$, by definition of $q^*_{i,j,t}$ for $i = 1, ..., n$ and $j = 1, ..., p$ (see equation (3)). Combining $\lambda(\alpha^*) = 0$ with equations (17) and (18), we obtain

$$\lambda(\alpha) = -Q^*(\alpha - \alpha^*) + O(||\alpha - \alpha^*||^2). \tag{19}$$

The next step is to show that

$$T^{1/2}\lambda(\hat{\alpha}_T) + H_T = o_p(1) \tag{20}$$

where $H_T := T^{-1/2}\sum_{t=1}^{T} \eta^*_t$, with $\eta^*_t := \eta_t(\alpha^*)$ and $\eta_t(\alpha) := \sum_{i=1}^{n}\sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \alpha)$ $\psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha))$. 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 (20) 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 the given (small) $d > 0$

$$\begin{aligned}
&u_t(\alpha, d) \\
&\leq \sup_{\{\tau:||\tau-\alpha||\leq d\}} \sum_{i=1}^{n}\sum_{j=1}^{p} ||\nabla q_{i,j,t}(\cdot,\tau)\psi_{\theta_{ij}}(Y_{it}-q_{i,j,t}(\cdot,\tau)) - \nabla q_{i,j,t}(\cdot,\alpha)\psi_{\theta_{ij}}(Y_{it}-q_{i,j,t}(\cdot,\alpha))|| \\
&\leq \sum_{i=1}^{n}\sum_{j=1}^{p} \sup_{\{\tau:||\tau-\alpha||\leq d\}} ||\psi_{\theta_{ij}}(Y_{it}-q_{i,j,t}(\cdot,\tau))|| \times \sup_{\{\tau:||\tau-\alpha||\leq d\}} ||\nabla q_{i,j,t}(\cdot,\tau) - \nabla q_{i,j,t}(\cdot,\alpha)|| \\
&\quad + \sum_{i=1}^{n}\sum_{j=1}^{p} \sup_{\{\tau:||\tau-\alpha||\leq d\}} ||\psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot,\alpha)) - \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot,\tau))|| \\
&\qquad\qquad \times \sup_{\{\tau:||\tau-\alpha||\leq d\}} ||\nabla q_{i,j,t}(\cdot,\alpha)|| \\
&\leq npD_{2,t}d + D_{1,t}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[|Y_{it}-q_{i,j,t}(\cdot,\alpha)|<D_{1,t}d]}
\end{aligned}$$

using the following: (i) $||\psi_{\theta_{ij}}(Y_{it}-q_{i,j,t}(\cdot,\tau))|| \leq 1$; (ii) $||\psi_{\theta_{ij}}(Y_{it}-q_{i,j,t}(\cdot,\alpha)) - \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot,\tau))|| \leq 1_{[|Y_{it}-q_{i,j,t}(\cdot,\alpha)|<|q_{i,j,t}(\cdot,\tau)-q_{i,j,t}(\cdot,\alpha)|]}$; 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

$$E[u_t(\alpha, d)] \leq npC_0 d + 2npC_1 f_0 d$$

for some constants C_0 and C_1, given Assumptions 2(iii.a), 5(iii.a), and 5(iv.a). Hence, (ii) holds for $b = npC_0 + 2npC_1 f_0$ and $d_0 = 2d$. The last condition (iii) can be similarly verified by applying the c_r-inequality to equation (**??**) with $d < 1$ (so that $d^2 < d$) and using Assumptions 2(iii.a), 5(iii.b), and 5(iv.b). As a result, equation (20) is verified.

Combining equations (19) and (20) yields

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

However, $\{\eta_t^*, \mathcal{F}_t\}$ is a stationary ergodic martingale difference sequence (MDS). In particular, η_t^* is measurable$-\mathcal{F}_t$, and $E(\eta_t^* | \mathcal{F}_{t-1}) = E(\sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \alpha^*) \psi_{\theta_{ij}}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}) = \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}(\cdot, \alpha^*) E(\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}) = 0$, as $E[\psi_{\theta_{ij}}(\varepsilon_{i,j,t}) \mid \mathcal{F}_{t-1}] = 0$ for all $i = 1, ..., n$ and $j = 1, ..., p$. Assumption 5(iii.b) ensures that $V^* := E(\eta_t^* \eta_t^{*\prime})$ 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 $T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t^{*\prime} = V^* + o_p(1)$, 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

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

where $\hat{\eta}_t := \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla \hat{q}_{i,j,t} \hat{\psi}_{i,j,t}$ and $\eta_t^* := \sum_{i=1}^{n} \sum_{j=1}^{p} \nabla q_{i,j,t}^* \psi_{i,j,t}^*$, with $\nabla \hat{q}_{i,j,t} := \nabla q_{i,j,t}(\cdot, \hat{\alpha}_T), \hat{\psi}_{i,j,t} := \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \hat{\alpha}_T)), \nabla q_{i,j,t}^* := \nabla q_{i,j,t}(\cdot, \alpha^*)$, and $\psi_{i,j,t}^* := \psi_{\theta_{ij}}(Y_{it} - q_{i,j,t}(\cdot, \alpha^*))$. Assumptions 1 and 2(i,ii) ensure that $\{\eta_t^* \eta_t^{*\prime}\}$ is a stationary ergodic sequence. Assumptions 3(i,ii), 4(i.a), and 5(iii) ensure that $E[\eta_t^* \eta_t^{*\prime}] < \infty$. It follows by the ergodic theorem that $T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t^{*\prime} - E[\eta_t^* \eta_t^{*\prime}] = o_p(1)$. Thus, it suffices to prove $T^{-1} \sum_{t=1}^{T} \hat{\eta}_t \hat{\eta}_t' - T^{-1} \sum_{t=1}^{T} \eta_t^* \eta_t^{*\prime} = o_p(1)$.

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

$$T^{-1} \sum_{t=1}^{T} \{ \sum_{i=1}^{n} \sum_{j=1}^{p} \sum_{l=1}^{n} \sum_{k=1}^{p} (\hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla_h q_{i,j,t}^* \nabla_s q_{l,k,t}^*) \}.$$

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

$$T^{-1} \sum_{t=1}^{T} \{ \hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla_h q_{i,j,t}^* \nabla_s q_{l,k,t}^* \} = o_p(1).$$

By the triangle inequality,

$$|T^{-1} \sum_{t=1}^{T} \{ \hat{\psi}_{i,j,t} \hat{\psi}_{l,k,t} \nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{l,k,t} - \psi_{i,j,t}^* \psi_{l,k,t}^* \nabla_h q_{i,j,t}^* \nabla_s q_{l,k,t}^* \}| \ \leq A_T + B_T,$$

where

$$
\begin{aligned}
A_T &= T^{-1}\sum_{t=1}^{T}|\hat{\psi}_{i,j,t}\hat{\psi}_{l,k,t}\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}-\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}| \\
B_T &= T^{-1}\sum_{t=1}^{T}|\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q^*_{i,j,t}\nabla_s q^*_{l,k,t}-\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}|.
\end{aligned}
$$

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

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

where

$$
\begin{aligned}
A_{1T} &= T^{-1}\sum_{t=1}^{T}\theta_{ij}|1_{[\varepsilon_{i,j,t}\le 0]}-1_{[\hat{\varepsilon}_{i,j,t}\le 0]}||\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}| \\
A_{2T} &= T^{-1}\sum_{t=1}^{T}\theta_{lk}|1_{[\varepsilon_{l,k,t}\le 0]}-1_{[\hat{\varepsilon}_{l,k,t}\le 0]}||\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}| \\
A_{3T} &= T^{-1}\sum_{t=1}^{T}|1_{[\varepsilon_{i,j,t}\le 0]}1_{[\varepsilon_{l,k,t}\le 0]}-1_{[\hat{\varepsilon}_{i,j,t}\le 0]}1_{[\hat{\varepsilon}_{l,k,t}\le 0]}||\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}|.
\end{aligned}
$$

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,

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

where

$$
\begin{aligned}
B_{1T} &= T^{-1}\sum_{t=1}^{T}|\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q^*_{i,j,t}\nabla_s q^*_{l,k,t}-E[\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q^*_{i,j,t}\nabla_s q^*_{l,k,t}]| \\
B_{2T} &= T^{-1}\sum_{t=1}^{T}|\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h\hat{q}_{i,j,t}\nabla_s\hat{q}_{l,k,t}-E[\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q^*_{i,j,t}\nabla_s q^*_{l,k,t}]|.
\end{aligned}
$$

Assumptions 1, 2(i,ii), 3(i,ii), 4(i.a), and 5(iii) ensure that the ergodic theorem applies to the stochastic sequence of $\{\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q^*_{i,j,t}\nabla_s q^*_{l,k,t}\}$, so $B_{1T} = o_p(1)$. Next, Assumptions 1, 3(i,ii), and 5(iii) ensure that the stationary ergodic ULLN applies to $\{\psi^*_{i,j,t}\psi^*_{l,k,t}\nabla_h q_{i,j,t}(\cdot,\alpha)$ $\nabla_s q_{l,k,t}(\cdot,\alpha)\}$. 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 the consistency of $\hat{Q}_T$, we strengthen the domination condition on $\nabla q_{i,j,t}$ and impose conditions on $\{\hat{c}_T\}$.

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

Assumption 7 $\{\hat{c}_T\}$ is a stochastic sequence and $\{c_T\}$ is a non-stochastic sequence such that (i) $\hat{c}_T/c_T \xrightarrow{p} 1$; (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

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

and then we will show the following:

$$Q^* - E(Q_T) \xrightarrow{p} 0, \tag{21}$$

$$E(Q_T) - Q_T \xrightarrow{p} 0, \tag{22}$$

$$Q_T - \hat{Q}_T \xrightarrow{p} 0. \tag{23}$$

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

For (21), one can show by applying the mean value theorem to $F_{i,j,t}(c_T) - F_{i,j,t}(-c_T)$, where $F_{i,j,t}(c) := \int 1_{\{s \le c\}} f_{i,j,t}(s) ds$, 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

$$\begin{aligned}
& |\sum_{i=1}^{n} \sum_{j=1}^{p} E\{f_{i,j,t}(\xi_{i,j,T}) - f_{i,j,t}(0) \nabla_h q_{i,j,t}^* \nabla_s q_{i,j,t}^*\}| \\
\le & \sum_{i=1}^{n} \sum_{j=1}^{p} E\{|f_{i,j,t}(\xi_{i,j,T}) - f_{i,j,t}(0)||\nabla_h q_{i,j,t}^* \nabla_s q_{i,j,t}^*|\} \\
\le & \sum_{i=1}^{n} \sum_{j=1}^{p} L_0 E\{|\xi_{i,j,T}||\nabla_h q_{i,j,t}^* \nabla_s q_{i,j,t}^*|\} \\
\le & \; npL_0 c_T E[D_{1,t}^2],
\end{aligned}$$

which converges to zero as $c_T \to 0$. The second inequality follows from Assumption 2(iii.b), and the last inequality follows by Assumption 5(iii.b). Therefore, we have the result shown in equation (21).

To show (22), it suffices to simply apply an LLN for double arrays, e.g. theorem 2 in Andrews (1988). Finally, for (23), we consider the (h, s) element of $|\hat{Q}_T - Q_T|$, given by

$$
\begin{aligned}
&|\frac{1}{2\hat{c}_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-\hat{c}_T \le \hat{\varepsilon}_{i,j,t} \le \hat{c}_T]} \nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t} \\
&\qquad - \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} \nabla_h q^*_{i,j,t} \nabla_s q^*_{i,j,t}| \\
&= \frac{c_T}{\hat{c}_T} \times |\frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} (1_{[-\hat{c}_T \le \hat{\varepsilon}_{i,j,t} \le \hat{c}_T]} - 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]}) \nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t} \\
&\qquad + \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} (\nabla_h \hat{q}_{i,j,t} - \nabla_h q^*_{i,j,t}) \nabla_s \hat{q}_{i,j,t} \\
&\qquad + \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} \nabla_h q^*_{i,j,t} (\nabla_s \hat{q}_{i,j,t} - \nabla_s q^*_{i,j,t}) \\
&\qquad + \frac{1}{2c_T T}(1 - \frac{\hat{c}_T}{c_T})\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} \nabla_h q^*_{i,j,t} \nabla_s q^*_{i,j,t}| \\
&\le \frac{c_T}{\hat{c}_T}[A_{1T} + A_{2T} + A_{3T} + (1 - \frac{\hat{c}_T}{c_T})A_{4T}],
\end{aligned}
$$

where

$$
\begin{aligned}
A_{1T} &:= \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} |1_{[-\hat{c}_T \le \hat{\varepsilon}_{i,j,t} \le \hat{c}_T]} - 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]}| \times |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}| \\
A_{2T} &:= \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} |\nabla_h \hat{q}_{i,j,t} - \nabla_h q^*_{i,j,t}| \times |\nabla_s \hat{q}_{i,j,t}| \\
A_{3T} &:= \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} |\nabla_h q^*_{i,j,t}| \times |\nabla_s \hat{q}_{i,j,t} - \nabla_s q^*_{i,j,t}| \\
A_{4T} &:= \frac{1}{2c_T T}\sum_{t=1}^{T}\sum_{i=1}^{n}\sum_{j=1}^{p} 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]} |\nabla_h q^*_{i,j,t} \nabla_s q^*_{i,j,t}|.
\end{aligned}
$$

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,

$$
\frac{1}{2c_T T}\sum_{t=1}^{T} |1_{[-\hat{c}_T \le \hat{\varepsilon}_{i,j,t} \le \hat{c}_T]} - 1_{[-c_T \le \varepsilon_{i,j,t} \le c_T]}| \times |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}| = o_p(1).
$$

Let α_T lie between $\hat{\alpha}_T$ and α^*, 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

$$(2c_T T)^{-1} \sum_{t=1}^{T} |1_{[-\hat{c}_T \leq \hat{\varepsilon}_{i,j,t} \leq \hat{c}_T]} - 1_{[-c_T \leq \varepsilon_{i,j,t} \leq c_T]}| \times |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}| \ \leq U_T + V_T,$$

where

$$\begin{aligned} U_T &:= (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|\varepsilon_{i,j,t} - c_T| < d_{i,j,t,T}]} |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}| \\ V_T &:= (2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|\varepsilon_{i,j,t} + c_T| < d_{i,j,t,T}]} |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}|. \end{aligned}$$

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 $\eta > 0$,

$$\begin{aligned} P(U_T &> \eta) \leq P((2c_T T)^{-1} \sum_{t=1}^{T} 1_{[|\varepsilon_{i,j,t} - c_T| < (||\nabla q_{i,j,t}(\cdot, \alpha_T)|| + 1) z c_T]}) |\nabla_h \hat{q}_{i,j,t} \nabla_s \hat{q}_{i,j,t}| > \eta) \\ &\leq \frac{z f_0}{\eta T} \sum_{t=1}^{T} E\left\{(||\nabla q_{i,j,t}(\cdot, \alpha_T)|| + 1) |\nabla_h \hat{q}_{j,t} \nabla_s \hat{q}_{j,t}|\right\} \\ &\leq z f_0 \{E|D_{1,t}^3| + E|D_{1,t}^2|\}/\eta, \end{aligned}$$

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 $U_T \xrightarrow{p} 0$. The same argument is used to show $V_T \xrightarrow{p} 0$. Hence, $A_{1T} = o_p(1)$ is proved.

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

$$\begin{aligned} A_{2T,i,j} &\leq \frac{1}{2c_T T} \sum_{t=1}^{T} |\nabla_h \hat{q}_{i,j,t} - \nabla_h q^*_{i,j,t}| \times |\nabla_s \hat{q}_{i,j,t}| \\ &\leq \frac{1}{2c_T T} \sum_{t=1}^{T} ||\nabla_h^2 q_{i,j,t}(\cdot, \tilde{\alpha})|| \times ||\hat{\alpha}_T - \alpha^*|| \times |\nabla_s \hat{q}_{i,j,t}| \\ &\leq \frac{1}{2c_T} ||\hat{\alpha}_T - \alpha^*|| \frac{1}{T} \sum_{t=1}^{T} D_{2,t} D_{1,t} \\ &= \frac{1}{2c_T T^{1/2}} T^{1/2} ||\hat{\alpha}_T - \alpha^*|| \frac{1}{T} \sum_{t=1}^{T} D_{2,t} D_{1,t}, \end{aligned}$$

where $\tilde{\alpha}$ is between $\hat{\alpha}_T$ and α^*, and $\nabla_h^2 q_{j,t}(\cdot, \tilde{\alpha})$ is the first derivative of $\nabla_h \hat{q}_{j,t}$ with respect to α, 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 $Q_T - \hat{Q}_T \xrightarrow{p} 0$, which, together with (21) and (22), implies that $\hat{Q}_T - Q^* \xrightarrow{p} 0$. As a result, the proof is complete. ■

Table 1 – Financial institutions included in the sample

	NAME	COUNTRY	SECTOR
1	77 BANK	JP	BK
2	ALLIED IRISH BANKS	IE	BK
3	ALPHA BANK	GR	BK
4	AUS.AND NZ.BANKING GP.	AU	BK
5	AWA BANK	JP	BK
6	BANK OF IRELAND	IE	BK
7	BANKINTER 'R'	ES	BK
8	BARCLAYS	GB	BK
9	BB&T	US	BK
10	BANCA CARIGE	IT	BK
11	BANCA MONTE DEI PASCHI	IT	BK
12	BANCA POPOLARE DI MILANO	IT	BK
13	BANCA PPO.DI SONDRIO	IT	BK
14	BANCA PPO.EMILIA ROM.	IT	BK
15	BBV.ARGENTARIA	ES	BK
16	BANCO COMR.PORTUGUES 'R'	PT	BK
17	BANCO DE VALENCIA	ES	BK
18	BANCO ESPIRITO SANTO	PT	BK
19	BANCO POPOLARE	IT	BK
20	BANCO POPULAR ESPANOL	ES	BK
21	BANCO SANTANDER	ES	BK
22	BNP PARIBAS	FR	BK
23	BANK OF AMERICA	US	BK
24	BANK OF EAST ASIA	HK	BK
25	BANK OF KYOTO	JP	BK
26	BANK OF MONTREAL	CA	BK
27	BK.OF NOVA SCOTIA	CA	BK
28	BANK OF QLND.	AU	BK
29	BANK OF YOKOHAMA	JP	BK
30	BENDIGO & ADELAIDE BANK	AU	BK
31	COMMERZBANK (XET)	DE	BK
32	CREDIT SUISSE GROUP N	CH	BK
33	CREDITO VALTELLINES	IT	BK
34	CANADIAN IMP.BK.COM.	CA	BK
35	CHIBA BANK	JP	BK
36	CHUGOKU BANK	JP	BK
37	CHUO MITSUI TST.HDG.	JP	BK
38	CITIGROUP	US	BK
39	COMERICA	US	BK
40	COMMONWEALTH BK.OF AUS.	AU	BK
41	DANSKE BANK	DK	BK
42	DBS GROUP HOLDINGS	SG	BK
43	DEUTSCHE BANK (XET)	DE	BK
44	DEXIA	BE	BK
45	DNB NOR	NO	BK
46	DAISHI BANK	JP	BK
47	EFG EUROBANK ERGASIAS	GR	BK
48	ERSTE GROUP BANK	AT	BK
49	FIFTH THIRD BANCORP	US	BK
50	FUKUOKA FINANCIAL GP.	JP	BK
51	SOCIETE GENERALE	FR	BK
52	GUNMA BANK	JP	BK
53	HSBC HOLDINGS	HK	BK
54	HACHIJUNI BANK	JP	BK
55	HANG SENG BANK	HK	BK
56	HIGO BANK	JP	BK
57	HIROSHIMA BANK	JP	BK
58	HOKUHOKU FINL. GP.	JP	BK
59	HUDSON CITY BANC.	US	BK
60	HUNTINGTON BCSH.	US	BK
61	HYAKUGO BANK	JP	BK
62	HYAKUJUSHI BANK	JP	BK
63	INTESA SANPAOLO	IT	BK
64	IYO BANK	JP	BK
65	JP MORGAN CHASE & CO.	US	BK
66	JYSKE BANK	DK	BK
67	JOYO BANK	JP	BK
68	JUROKU BANK	JP	BK
69	KBC GROUP	BE	BK
70	KAGOSHIMA BANK	JP	BK
71	KEIYO BANK	JP	BK
72	KEYCORP	US	BK
73	LLOYDS BANKING GROUP	GB	BK
74	M&T BK.	US	BK
75	MEDIOBANCA	IT	BK
76	MARSHALL & ILSLEY	US	BK
77	MIZUHO TST.& BKG.	JP	BK
78	NATIONAL BK.OF GREECE	GR	BK
79	NATIXIS	FR	BK
80	NORDEA BANK	SE	BK
81	NANTO BANK	JP	BK
82	NATIONAL AUS.BANK	AU	BK
83	NAT.BK.OF CANADA	CA	BK
84	NY.CMTY.BANC.	US	BK
85	NISHI-NIPPON CITY BANK	JP	BK
86	NORTHERN TRUST	US	BK
87	OGAKI KYORITSU BANK	JP	BK
88	OVERSEA-CHINESE BKG.	SG	BK
89	BANK OF PIRAEUS	GR	BK
90	PNC FINL.SVS.GP.	US	BK
91	POHJOLA PANKKI A	FI	BK
92	PEOPLES UNITED FINANCIAL	US	BK
93	ROYAL BANK OF SCTL.GP.	GB	BK
94	REGIONS FINL.NEW	US	BK
95	RESONA HOLDINGS	JP	BK
96	ROYAL BANK CANADA	CA	BK
97	SEB 'A'	SE	BK
98	STANDARD CHARTERED	GB	BK
99	SVENSKA HANDBKN.'A'	SE	BK
100	SWEDBANK 'A'	SE	BK
101	SYDBANK	DK	BK
102	SAN-IN GODO BANK	JP	BK
103	SHIGA BANK	JP	BK
104	SHINKIN CENTRAL BANK PF.	JP	BK
105	SUMITOMO MITSUI FINL.GP.	JP	BK
106	SUMITOMO TRUST & BANK.	JP	BK
107	SUNTRUST BANKS	US	BK
108	SUNCORP-METWAY	AU	BK
109	SURUGA BANK	JP	BK
110	TORONTO-DOMINION BANK	CA	BK
111	US BANCORP	US	BK
112	UBS 'R'	CH	BK
113	UNICREDIT	IT	BK
114	UNITED OVERSEAS BANK	SG	BK
115	VALIANT 'R'	CH	BK
116	WELLS FARGO & CO	US	BK
117	WESTPAC BANKING	AU	BK
118	WING HANG BANK	HK	BK
119	YAMAGUCHI FINL.GP.	JP	BK
120	3I GROUP	GB	FS
121	ABERDEEN ASSET MAN.	GB	FS
122	ACKERMANS & VAN HAAREN	BE	FS
123	AMP	AU	FS
124	ASX	AU	FS
125	ACOM	JP	FS
126	AMERICAN EXPRESS	US	FS
127	BANK OF NEW YORK MELLON	US	FS
128	BLACKROCK	US	FS
129	CI FINANCIAL	CA	FS
130	CLOSE BROTHERS GROUP	GB	FS
131	CIE.NALE.A PTF.	BE	FS
132	CRITERIA CAIXACORP	ES	FS
133	CHALLENGER FINL.SVS.GP.	AU	FS
134	CHARLES SCHWAB	US	FS
135	CHINA EVERBRIGHT	HK	FS
136	COMPUTERSHARE	AU	FS
137	CREDIT SAISON	JP	FS
138	DAIWA SECURITIES GROUP	JP	FS
139	EURAZEO	FR	FS
140	EATON VANCE NV.	US	FS
141	EQUIFAX	US	FS
142	FRANKLIN RESOURCES	US	FS
143	GAM HOLDING	CH	FS
144	GBL NEW	BE	FS
145	GOLDMAN SACHS GP.	US	FS
146	ICAP	GB	FS
147	IGM FINL.	CA	FS

	NAME	COUNTRY	SECTOR
148	INDUSTRIVARDEN 'A'	SE	FS
149	INTERMEDIATE CAPITAL GP.	GB	FS
150	KINNEVIK 'B'	SE	FS
151	INVESTOR 'B'	SE	FS
152	LEGG MASON	US	FS
153	MAN GROUP	GB	FS
154	MARFIN INV.GP.HDG.	GR	FS
155	MACQUARIE GROUP	AU	FS
156	MITSUB.UFJ LSE.& FINANCE	JP	FS
157	MIZUHO SECURITIES	JP	FS
158	MOODY'S	US	FS
159	MORGAN STANLEY	US	FS
160	NOMURA HDG.	JP	FS
161	ORIX	JP	FS
162	PARGESA 'B'	CH	FS
163	PROVIDENT FINANCIAL	GB	FS
164	PERPETUAL	AU	FS
165	RATOS 'B'	SE	FS
166	SCHRODERS	GB	FS
167	SLM	US	FS
168	SOFINA	BE	FS
169	STATE STREET	US	FS
170	T ROWE PRICE GP.	US	FS
171	TD AMERITRADE HOLDING	US	FS
172	WENDEL	FR	FS
173	ACE	US	IN
174	AEGON	NL	IN
175	AFLAC	US	IN

	NAME	COUNTRY	SECTOR
176	AGEAS (EX-FORTIS)	BE	IN
177	ALLIANZ (XET)	DE	IN
178	AMLIN	GB	IN
179	AON	US	IN
180	GENERALI	IT	IN
181	AVIVA	GB	IN
182	AXA ASIA PACIFIC HDG.	AU	IN
183	AXA	FR	IN
184	ALLSTATE	US	IN
185	AMERICAN INTL.GP.	US	IN
186	ARCH CAP.GP.	US	IN
187	BALOISE-HOLDING AG	CH	IN
188	BERKSHIRE HATHAWAY 'B'	US	IN
189	CNP ASSURANCES	FR	IN
190	CHUBB	US	IN
191	CINCINNATI FINL.	US	IN
192	EVEREST RE GP.	US	IN
193	FAIRFAX FINL.HDG.	CA	IN
194	GREAT WEST LIFECO	CA	IN
195	HANNOVER RUCK. (XET)	DE	IN
196	HELVETIA HOLDING N	CH	IN
197	HARTFORD FINL.SVS.GP.	US	IN
198	ING GROEP	NL	IN
199	JARDINE LLOYD THOMPSON	GB	IN
200	LEGAL & GENERAL	GB	IN
201	LINCOLN NAT.	US	IN
202	LOEWS	US	IN
203	MAPFRE	ES	IN

	NAME	COUNTRY	SECTOR
204	MS&AD INSURANCE GP.HDG.	JP	IN
205	MUENCHENER RUCK. (XET)	DE	IN
206	MANULIFE FINANCIAL	CA	IN
207	MARKEL	US	IN
208	MARSH & MCLENNAN	US	IN
209	OLD MUTUAL	GB	IN
210	PRUDENTIAL	GB	IN
211	PARTNERRE	US	IN
212	POWER CORP.CANADA	CA	IN
213	POWER FINL.	CA	IN
214	PROGRESSIVE OHIO	US	IN
215	QBE INSURANCE GROUP	AU	IN
216	RSA INSURANCE GROUP	GB	IN
217	RENAISSANCERE HDG.	US	IN
218	SAMPO 'A'	FI	IN
219	SCOR SE	FR	IN
220	STOREBRAND	NO	IN
221	SWISS LIFE HOLDING	CH	IN
222	SWISS RE 'R'	CH	IN
223	TOPDANMARK	DK	IN
224	TORCHMARK	US	IN
225	TRAVELERS COS.	US	IN
226	UNUM GROUP	US	IN
227	VIENNA INSURANCE GROUP A	AT	IN
228	W R BERKLEY	US	IN
229	XL GROUP	US	IN
230	ZURICH FINANCIAL SVS.	CH	IN

Note: The abbreviation for the sector classification are as follows: BK = Bank, FS = Financial Services, IN = Insurance. Classification as provided by Datastream.

Table 2 – Breakdown of financial institutions by sector and by geographic area

	Banks	*Financial Services*	*Insurances*	
EU	47	22	27	**96**
North America	25	17	28	**70**
Asia	47	14	3	**64**
	119	**53**	**58**	**230**

Note: Swiss and Norwegian financial institutions have been classified as EU. Asia includes Australian financial institutions.

Table 3 – Estimates and standard errors for selected financial institutions

Barclays

	c_1	a_{11}	a_{12}	b_{11}	b_{12}
	-0.19	**-0.45**	-0.12	**0.77**	-0.01
s.e.	***0.07***	***0.14***	*0.09*	***0.06***	*0.01*
	c_2	a_{21}	a_{22}	b_{21}	b_{22}
	-0.17	**-0.40**	-0.22	**-0.21**	**0.96**
s.e.	***0.09***	***0.13***	*0.12*	***0.08***	***0.01***

Deutsche Bank

	c_1	a_{11}	a_{12}	b_{11}	b_{12}
	-0.12	**-0.37**	-0.07	**0.87**	-0.03
s.e.	***0.05***	***0.15***	*0.04*	***0.05***	*0.02*
	c_2	a_{21}	a_{22}	b_{21}	b_{22}
	-0.16	-0.11	-0.32	-0.03	**0.87**
s.e.	*0.11*	*0.34*	*0.22*	*0.11*	***0.08***

Citigroup

	c_1	a_{11}	a_{12}	b_{11}	b_{12}
	-0.13	-0.08	**-0.15**	**0.83**	0.02
s.e.	*0.08*	*0.08*	***0.05***	***0.09***	*0.02*
	c_2	a_{21}	a_{22}	b_{21}	b_{22}
	-0.02	-0.10	-0.17	-0.06	**0.97**
s.e.	*0.10*	*0.13*	*0.11*	*0.12*	***0.03***

Goldman Sachs

	c_1	a_{11}	a_{12}	b_{11}	b_{12}
	-0.03	**-0.13**	**-0.11**	**0.95**	-0.04
s.e.	*0.02*	***0.05***	***0.05***	***0.02***	*0.02*
	c_2	a_{21}	a_{22}	b_{21}	b_{22}
	-0.04	-0.03	**-0.17**	0.00	**0.93**
s.e.	*0.03*	*0.11*	***0.07***	*0.04*	***0.03***

Note: Coefficients significant at the 5% level formatted in bold.

Figure 1 – 1% quantile for selected financial institutions

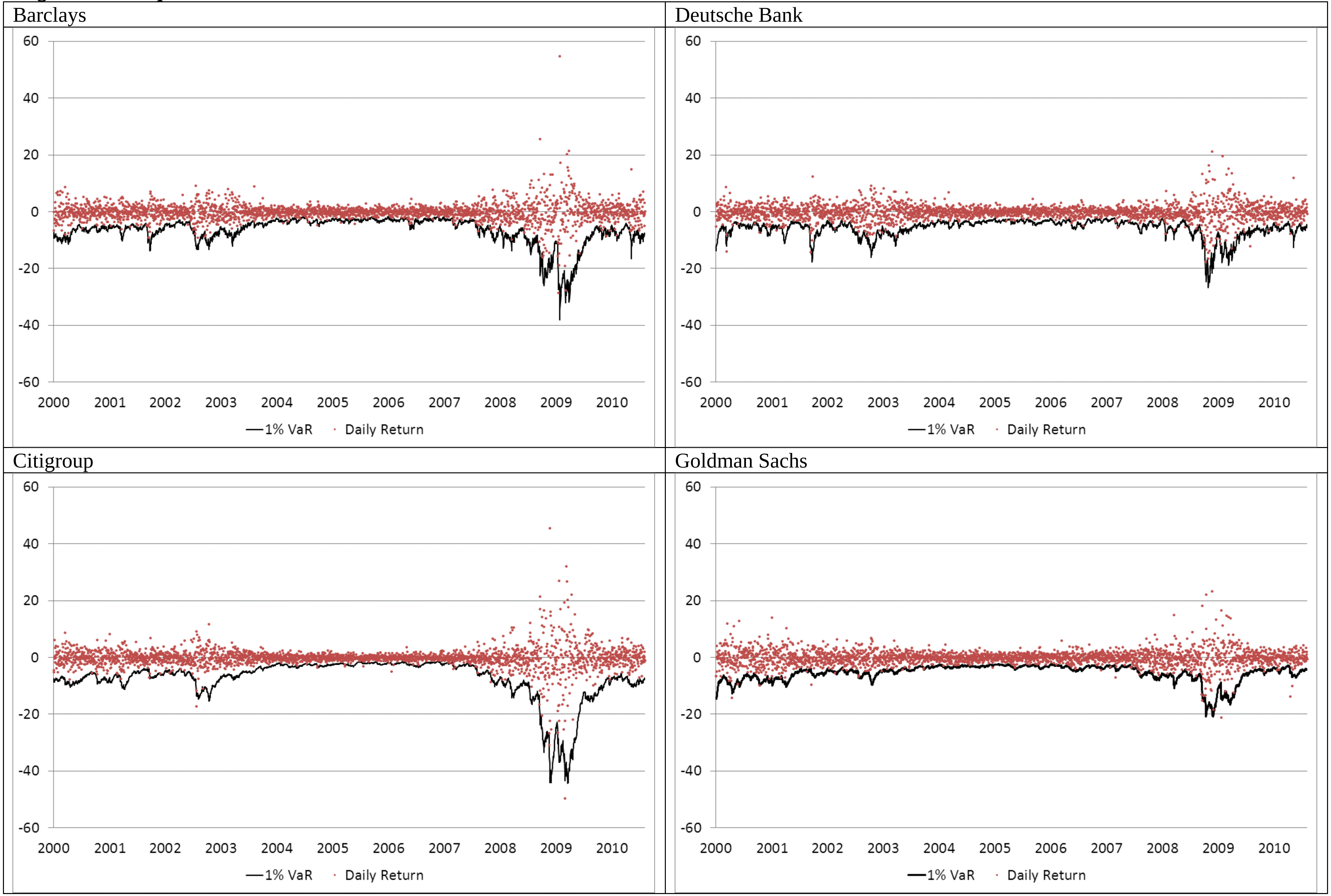


Figure 2 – Impulse-response functions to a shock to the market for selected financial institutions

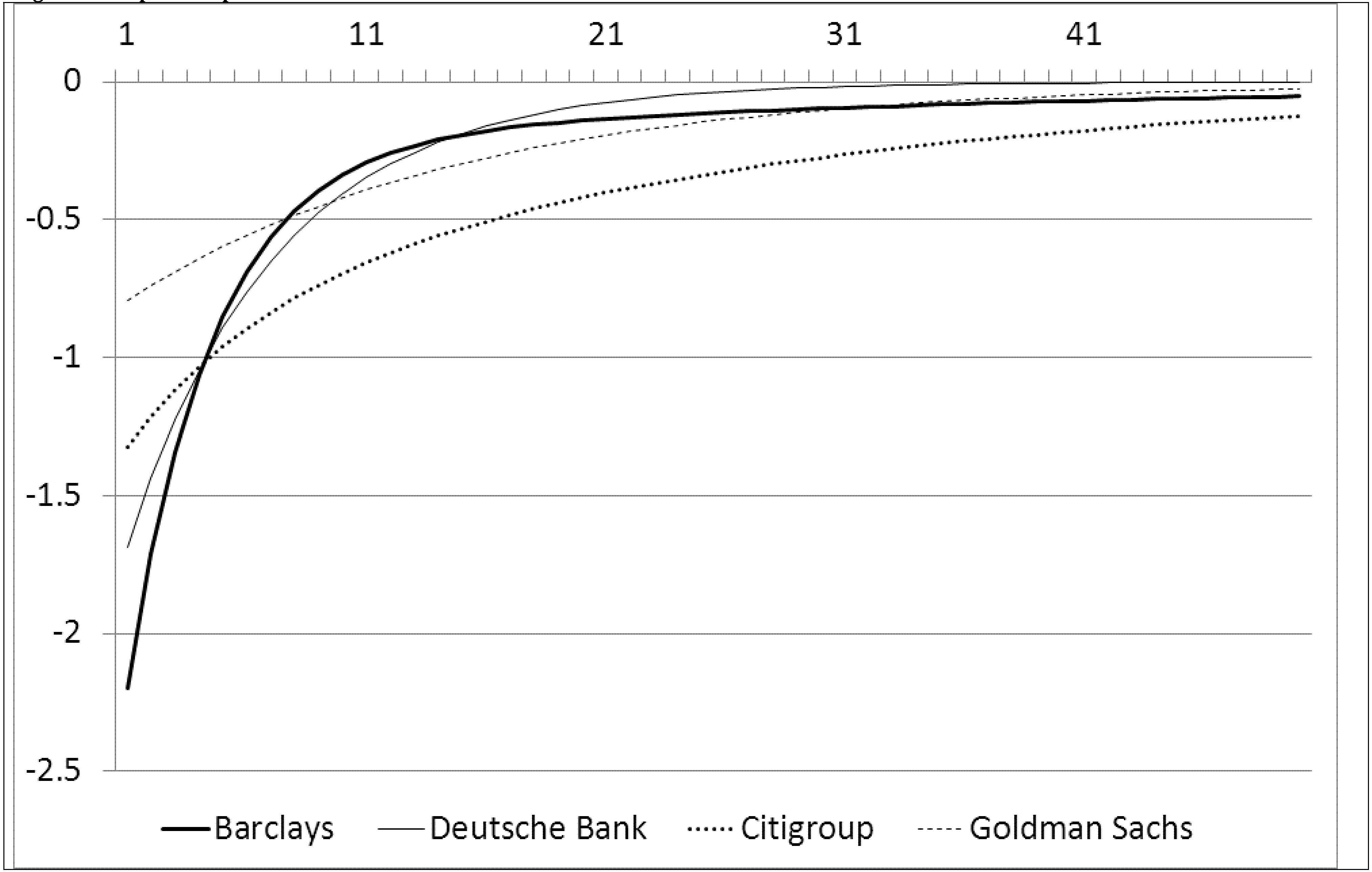


Figure 3 – Impulse-response functions by sectoral and geographic aggregation

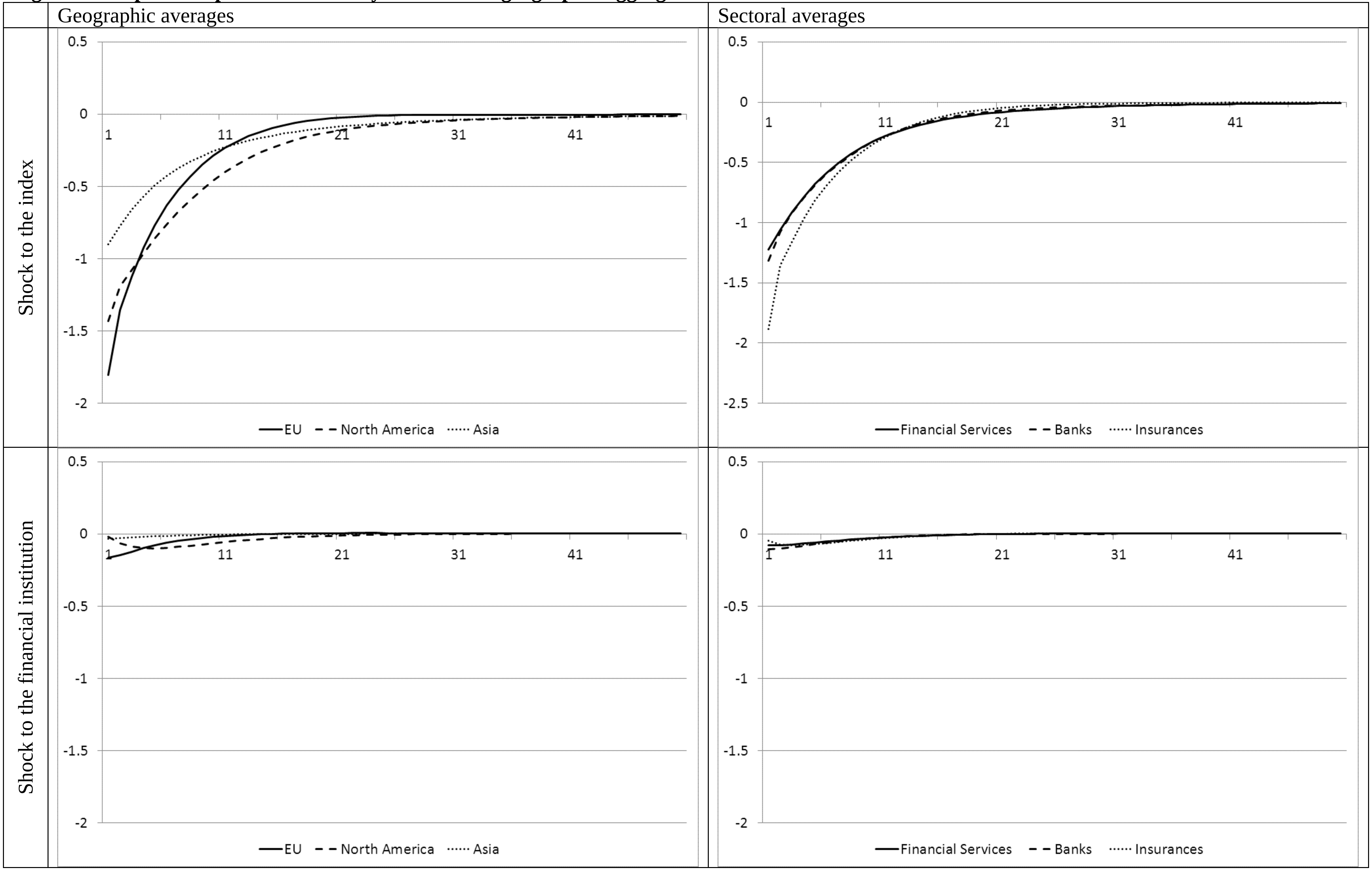
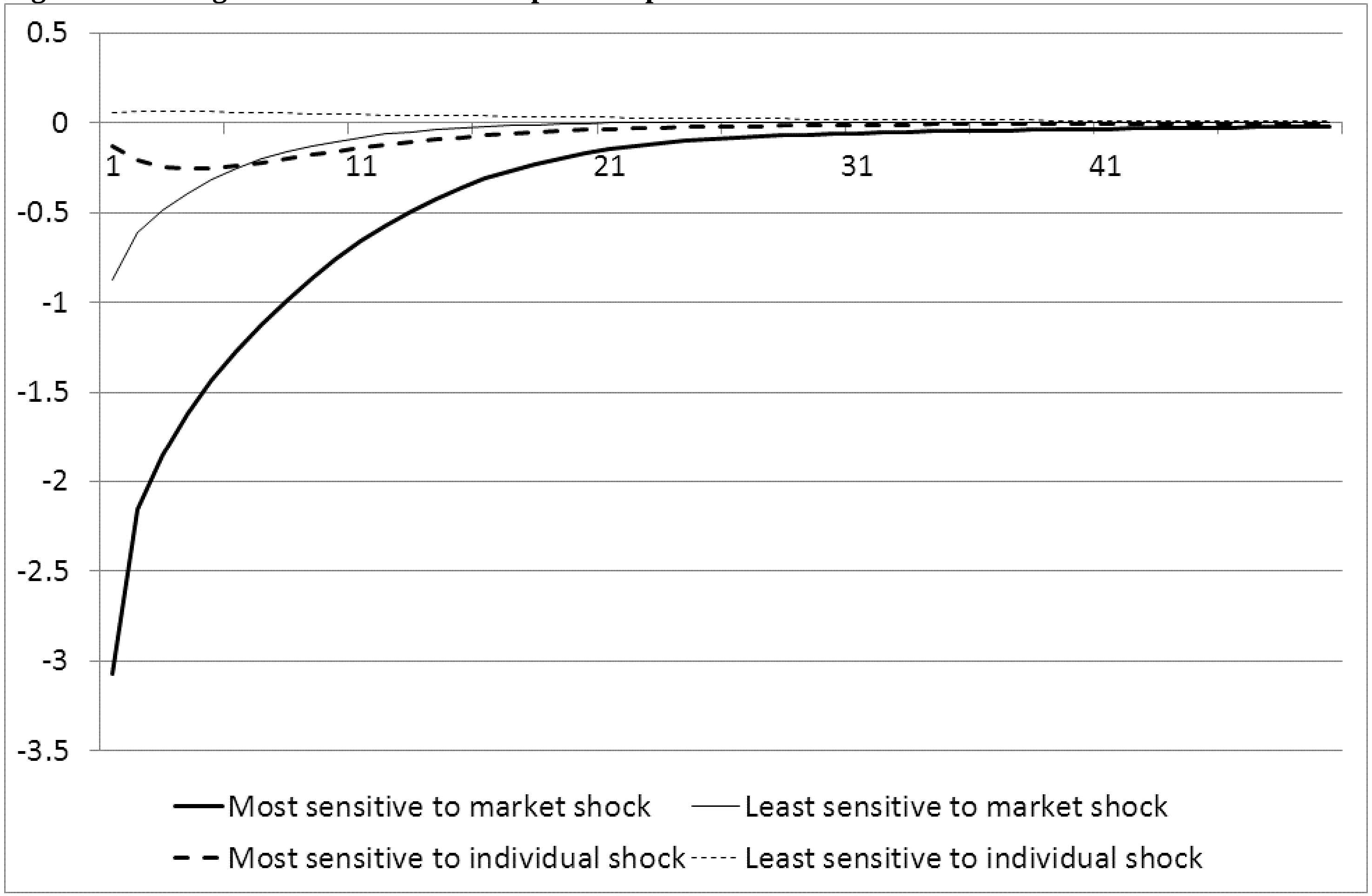


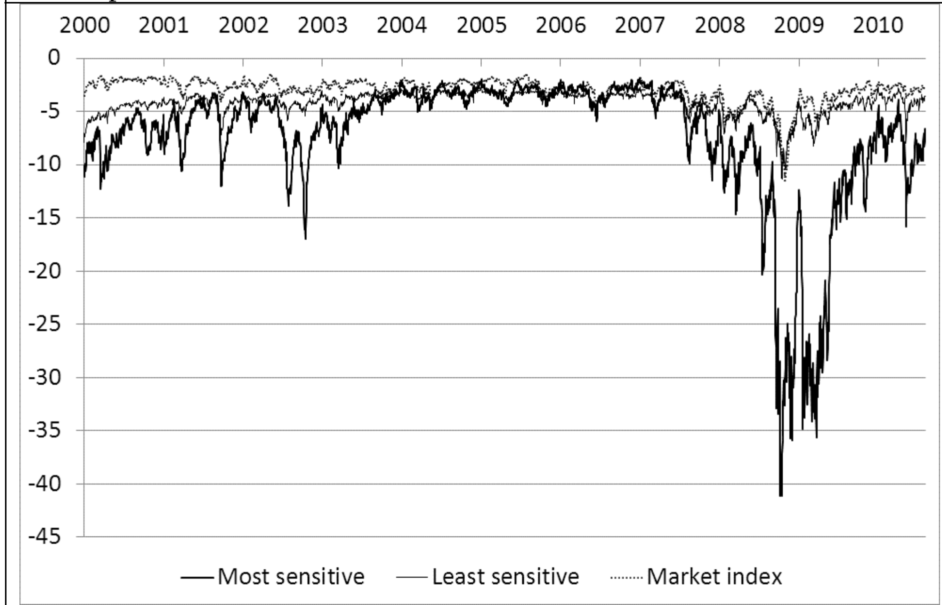
Figure 4 – Strongest and weakest VaR impulse-responses



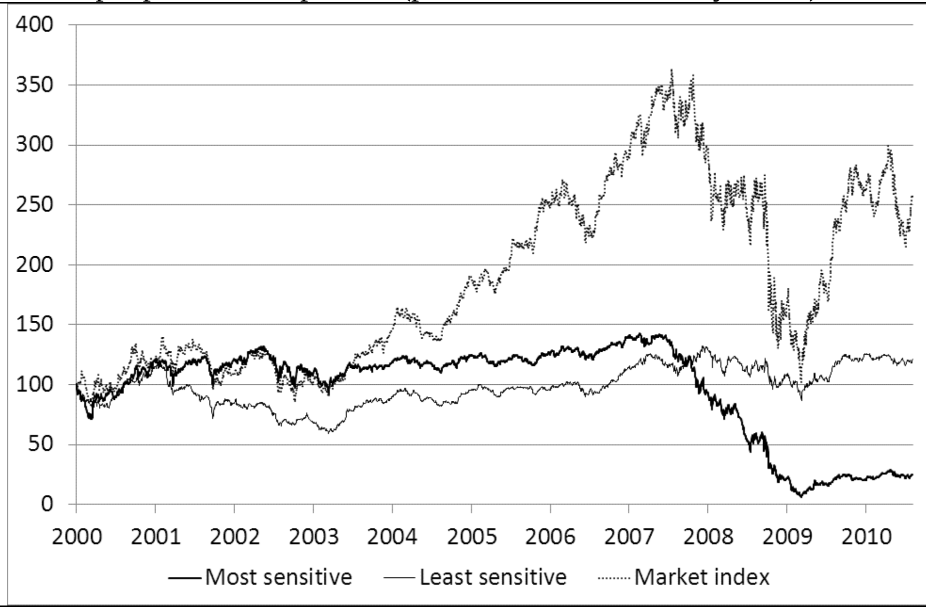
Note: The figure reports the average impulse-response function of the 20 financial institutions with the strongest and weakest impact, as measured by the area below the impulse-response function.

Figure 5 – In-sample and out-of-sample average VaR and price developments of the 20 financial institutions with the strongest and weakest VaR impulse-responses to a market shock

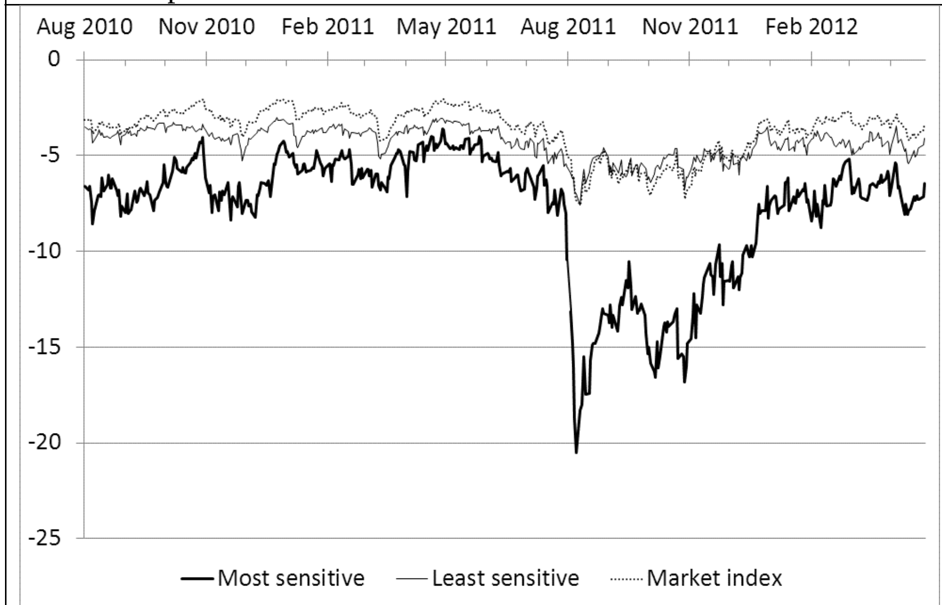
In-sample VaR



In-sample price developments (price = 100 on 2 January, 2000)



Out-of-sample VaR



Out-of-sample price developments (price = 100 on 6 August, 2010)