Journal of Econometrics 193 (2016) 251–270

Contents lists available at ScienceDirect

Journal of Econometrics

journal homepage: www.elsevier.com/locate/jeconom

The cross-quantilogram: Measuring quantile dependence and testing directional predictability between time series[☆]

Heejoon Han[a], Oliver Linton[b], Tatsushi Oka[c], Yoon-Jae Whang[d,*]

[a] Department of Economics, Sungkyunkwan University, Seoul, Republic of Korea
[b] Faculty of Economics, University of Cambridge, Cambridge, UK
[c] Department of Economics, National University of Singapore, Singapore
[d] Department of Economics, Seoul National University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 2 October 2014
Received in revised form
7 October 2015
Accepted 13 March 2016
Available online 30 March 2016

Keywords:
Quantile
Correlogram
Dependence
Predictability
Systemic risk

ABSTRACT

This paper proposes the cross-quantilogram to measure the quantile dependence between two time series. We apply it to test the hypothesis that one time series has no directional predictability to another time series. We establish the asymptotic distribution of the cross-quantilogram and the corresponding test statistic. The limiting distributions depend on nuisance parameters. To construct consistent confidence intervals we employ a stationary bootstrap procedure; we establish consistency of this bootstrap. Also, we consider a self-normalized approach, which yields an asymptotically pivotal statistic under the null hypothesis of no predictability. We provide simulation studies and two empirical applications. First, we use the cross-quantilogram to detect predictability from stock variance to excess stock return. Compared to existing tools used in the literature of stock return predictability, our method provides a more complete relationship between a predictor and stock return. Second, we investigate the systemic risk of individual financial institutions, such as JP Morgan Chase, Morgan Stanley and AIG.

1. Introduction

Linton and Whang (2007) introduced the quantilogram to measure predictability in different parts of the distribution of a stationary time series based on the correlogram of "quantile hits". They applied it to test the hypothesis that a given time series has no directional predictability. More specifically, their null hypothesis was that the past information set of the stationary time series $\{y_t\}$ does not improve the prediction about whether y_t will be above or below the unconditional quantile. The test is based on comparing the quantilogram to a pointwise confidence band. This contribution fits into a long literature of testing predictability using signs or rank statistics, including the papers of Cowles and Jones (1937), Dufour et al. (1998), and Christoffersen and Diebold (2002). The quantilogram has several advantages compared to other test statistics for directional predictability. It is conceptually appealing and simple to interpret. Since the method is based on quantile hits it does not require moment conditions like the ordinary correlogram and statistics like the variance ratio that are derived from it, Mikosch and Starica (2000), and so it works well for heavy tailed series. Many financial time series have heavy tails, see, e.g., Mandelbrot (1963), Fama (1965), Rachev and Mittnik (2000), Embrechts et al. (1997), Ibragimov et al. (2009), and Ibragimov (2009), and so this is an important consideration in practice. Additionally, this type of method allows researchers to consider very long lags in comparison with regression type methods, such as Engle and Manganelli (2004).

There have been a number of recent works either extending or applying this methodology. Davis and Mikosch (2009) have introduced the extremogram, which is essentially the quantilogram for extreme quantiles, and Davis et al. (2012) has provided the inference methods based on bootstrap and permutation for the extremogram. See also Davis et al. (2013). Li (2008, 2012) has introduced a Fourier domain version of the quantilogram while Hong (2000) has used a Fourier domain approach for test statistics based on distributions. Further development in the Fourier domain approach has been made by Hagemann (2013) and Dette et al. (2015).

☆ We thank a Co-Editor, Jianqing Fan, an Associate Editor and three anonymous referees for constructive comments. Han's work was supported by the National Research Foundation of Korea (NRF-2013S1A5A8021502). Linton's work was supported by Cambridge INET and the ERC (NAMSEF). Oka's work was supported by Singapore Academic Research Fund (FY2013-FRC2-003). Whang's work was supported by the National Research Foundation of Korea (NRF-2011-342-B00004) and Seoul National University.

* Corresponding author.
E-mail address: whang@snu.ac.kr (Y.-J. Whang).

http://dx.doi.org/10.1016/j.jeconom.2016.03.001
0304-4076/

See also Li (2014) and Kley et al. (2016). The quantilogram has recently been applied to stock returns and exchange rates, Laurini et al. (2008) and Chang and Shie (2011).

Our paper addresses three outstanding issues with regard to the quantilogram. First, the construction of confidence intervals that are valid under general dependence structures. Linton and Whang (2007) derived the limiting distribution of the sample quantilogram under the null hypothesis that the quantilogram itself is zero, in fact under a special case of that where the process has a type of conditional heteroskedasticity structure. Even in that very special case, the limiting distribution depends on model specific quantities. They derived a bound on the asymptotic variance that allows one to test the null hypothesis of the absence of predictability (or rather the special case of this that they work with). Even when this model structure is appropriate, the bounds can be quite large especially when one looks into the tails of the distribution. The quantilogram is also useful in cases where the null hypothesis of no predictability is not thought to be true—one can be interested in measuring the degree of predictability of a series across different quantiles. We provide a more complete solution to the issue of inference for the quantilogram. Specifically, we derive the asymptotic distribution of the quantilogram under general weak dependence conditions, specifically strong mixing. The limiting distribution is quite complicated and depends on the long run variance of the quantile hits. To conduct inference we propose the stationary bootstrap method of Politis and Romano (1994) and prove that it provides asymptotically valid confidence intervals. We investigate the finite sample performance of this procedure and show that it works well. We also provide R code that carries out the computations efficiently.[1] We also define a self-normalized version of the statistic for testing the null hypothesis that the quantilogram is zero, following Lobato (2001). This statistic has an asymptotically pivotal distribution, under the null hypothesis, whose critical values have been tabulated so that there is no need for long run variance estimation or even bootstrap.

Second, we develop our methodology inside a multivariate setting and explicitly consider the cross-quantilogram. Linton and Whang (2007) briefly mentioned such a multivariate version of the quantilogram but they provided neither theoretical results nor empirical results. In fact, the cross-correlogram is a vitally important measure of dependence between time series: Campbell et al. (1997), for example, use the cross autocorrelation function to describe lead lag relations between large stocks and small stocks. We apply the cross-quantilogram to the study of stock return predictability; our method provides a more complete picture of the predictability structure. We also apply the cross-quantilogram to the question of systemic risk. Our theoretical results described in the previous paragraph are all derived for the multivariate case.

Third, we explicitly allow the cross-quantilogram to be based on conditional (or regression) quantiles (Koenker and Bassett, 1978). Using conditional quantiles rather than unconditional quantiles, we measure directional dependence between two time-series after parsimoniously controlling for the information at the time of prediction.[2] Moreover, we derive the asymptotic distribution of the cross-quantilogram that are valid uniformly over a range of quantiles.

The remainder of the paper is as follows: Section 2 introduces the cross-quantilogram and Section 3 discusses its asymptotic properties. For consistent confidence intervals and hypothesis tests, we define the bootstrap procedure and introduce the self normalized test statistic. Section 4 considers the partial cross-quantilogram and gives a full treatment of its behavior in large samples. In Section 5 we report results of some Monte Carlo simulations to evaluate the finite sample properties of our procedures. In Section 6 we give two applications: we investigate stock return predictability and system risk using our methodology. Appendix contains all the proofs.

We use the following notation: The norm $\|\cdot\|$ denotes the Euclidean norm, i.e., $\|z\| = (\sum_{j=1}^{d} z_j^2)^{1/2}$ for $z = (z_1, \ldots, z_d)^\top \in \mathbb{R}^d$ and the norm $\|\cdot\|_p$ indicates the L^p norm of a $d \times 1$ random vector z, given by $\|z\|_p = (\sum_{j=1}^{d} E|z_j|^p)^{1/p}$ for $p > 0$. Let $1[\cdot]$ be the indicator function taking the value one when its argument is true, and zero otherwise. We use $\mathbb{R}$, $\mathbb{Z}$ and $\mathbb{N}$ to denote the set of all real numbers, all integers and all positive integers, respectively. Let $\mathbb{Z}_+ = \mathbb{N} \cup \{0\}$.

2. The cross-quantilogram

Let $\{(\mathbf{y}_t, \mathbf{x}_t) : t \in \mathbb{Z}\}$ be a strictly stationary time series with $\mathbf{y}_t = (y_{1t}, y_{2t})^\top \in \mathbb{R}^2$ and $\mathbf{x}_t = (x_{1t}, x_{2t}) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, where $x_{it} = [x_{it}^{(1)}, \ldots, x_{it}^{(d_i)}]^\top \in \mathbb{R}^{d_i}$ with $d_i \in \mathbb{N}$ for $i = 1, 2$. We use $F_{y_i|x_i}(\cdot|x_{it})$ to denote the conditional distribution function of the series y_{it} given x_{it} with density function $f_{y_i|x_i}(\cdot|x_{it})$, and the corresponding conditional quantile function is defined as $q_{i,t}(\tau_i) = \inf\{v : F_{y_i|x_i}(v|x_{it}) \geq \tau_i\}$ for $\tau_i \in (0, 1)$, for $i = 1, 2$. Let $\mathcal{T}$ be the range of quantiles we are interested in evaluating the directional predictability. For simplicity, we assume that $\mathcal{T}$ is a Cartesian product of two closed intervals in $(0, 1)$, that is $\mathcal{T} \equiv \mathcal{T}_1 \times \mathcal{T}_2$, where $\mathcal{T}_i = [\underline{\tau}_i, \overline{\tau}_i]$ for some $0 < \underline{\tau}_i < \overline{\tau}_i < 1$.[3]

We consider a measure of serial dependence between two events $\{y_{1t} \leq q_{1,t}(\tau_1)\}$ and $\{y_{2,t-k} \leq q_{2,t-k}(\tau_2)\}$ for an arbitrary pair of $\tau = (\tau_1, \tau_2)^\top \in \mathcal{T}$ and for an integer k. In the literature, $\{1[y_{it} \leq q_{i,t}(\cdot)]\}$ is called the quantile-hit or quantile-exceedance process for $i = 1, 2$. The cross-quantilogram is defined as the cross-correlation of the quantile-hit processes

$$\rho_\tau(k) = \frac{E\left[\psi_{\tau_1}(y_{1t} - q_{1,t}(\tau_1))\psi_{\tau_2}(y_{2,t-k} - q_{2,t-k}(\tau_2))\right]}{\sqrt{E\left[\psi_{\tau_1}^2(y_{1t} - q_{1,t}(\tau_1))\right]}\sqrt{E\left[\psi_{\tau_2}^2(y_{2,t-k} - q_{2,t-k}(\tau_2))\right]}}, \quad (1)$$

for $k = 0, \pm 1, \pm 2, \ldots$, where $\psi_a(u) \equiv 1[u < 0] - a$. The cross-quantilogram captures serial dependence between the two series at different conditional quantile levels. In the special case of a single time series, the cross-quantilogram becomes the quantilogram proposed by Linton and Whang (2007). Note that it is well-defined even for processes $\{(y_{1t}, y_{2t})\}_{t \in \mathbb{N}}$ with infinite moments. Like the quantilogram, the cross-quantilogram is invariant to any strictly monotonic transformation applied to both series, such as the logarithmic transformation.[4]

[1] This can be found at http://www.oliverlinton.me.uk/research/software.

[2] Our analysis includes the cross-quantilogram based on unconditional quantiles as a special case. In this case, the cross-quantilogram is shown to be a functional of the empirical copula introduced by Ruschendorf (1976) and Deheuvels (1979) as some nonparametric measures of dependence, such as Spearman's rho and Kendall's tau. In this special case, the asymptotic results for the empirical copula, which are found in Stute (1984), Fermanian et al. (2004) and Segers (2012) among others, can apply for the cross-quantilogram. Generally, however, the cross-quantilogram here differs from the empirical copula process and needs different treatment for analyzing its properties.

[3] It is straightforward to extend the results to a more general case, e.g. the case for which $\mathcal{T}$ is the union of a finite number of disjoint closed subsets of $(0, 1)^2$.

[4] When one is interested in measuring serial dependence between two events $\{q_{1,t}(\tau_1^l) \leq y_{1t} \leq q_{1,t}(\tau_1^h)\}$ and $\{q_{2,t-k}(\tau_2^l) \leq y_{2,t-k} \leq q_{2,t-k}(\tau_2^h)\}$ for arbitrary $[\tau_1^l, \tau_1^h]$ and $[\tau_2^l, \tau_2^h]$, one can use an alternative version of the cross-quantilogram that is defined by replacing $\psi_{\tau_i}(y_{it} - q_{i,t}(\tau_i))$ in (1) with

$$\psi_{[\tau_i^l, \tau_i^h]}\left(y_{it} - q_{i,t}\left([\tau_i^l, \tau_i^h]\right)\right) = 1\left[q_{i,t}(\tau_i^l) < y_{it} < q_{i,t}\left(\tau_i^h\right)\right] - \left(\tau_i^h - \tau_i^l\right).$$

To construct the sample analogue of the cross-quantilogram based on observations $\{(\mathbf{y}_t, \mathbf{x}_t)\}_{t=1}^{T}$, we first estimate conditional quantile functions. In this paper, we consider the linear quantile regression model proposed by Koenker and Bassett (1978) for simplicity and let $q_{i,t}(\tau_i) = x_{it}^{\top}\beta_i(\tau_i)$ with a $d_i \times 1$ vector of unknown parameters $\beta_i(\tau_i)$ for $i = 1, 2$. To estimate the parameters $\beta(\tau) \equiv [\beta_1(\tau_1)^{\top}, \beta_2(\tau_2)^{\top}]^{\top}$, we separately solve the following minimization problems:

$$\hat{\beta}_i(\tau_i) = \arg\min_{\beta_i \in \mathbb{R}^{d_i}} \sum_{t=1}^{T} \varrho_{\tau_i}\left(y_{it} - x_{it}^{\top}\beta_i\right),$$

where $\varrho_a(u) \equiv u(a - 1[u < 0])$. Let $\hat{\beta}(\tau) \equiv [\hat{\beta}_1(\tau_1)^{\top}, \hat{\beta}_2(\tau_2)^{\top}]^{\top}$ and $\hat{q}_{i,t}(\tau_i) = x_{it}^{\top}\hat{\beta}_i(\tau_i)$ for $i = 1, 2$. The sample cross-quantilogram is defined by

$$\hat{\rho}_{\tau}(k) = \frac{\sum_{t=k+1}^{T} \psi_{\tau_1}(y_{1t} - \hat{q}_{1,t}(\tau_1))\psi_{\tau_2}(y_{2,t-k} - \hat{q}_{2,t-k}(\tau_2))}{\sqrt{\sum_{t=k+1}^{T} \psi_{\tau_1}^2(y_{1t} - \hat{q}_{1,t}(\tau_1))}\sqrt{\sum_{t=k+1}^{T} \psi_{\tau_2}^2(y_{2,t-k} - \hat{q}_{2,t-k}(\tau_2))}}, \quad (2)$$

for $k = 0, \pm 1, \pm 2, \ldots$. Given a set of conditional quantiles, the cross-quantilogram considers dependence in terms of the direction of deviation from conditional quantiles and thus measures the directional predictability from one series to another. This can be a useful descriptive device. By construction, $\hat{\rho}_{\tau}(k) \in [-1, 1]$ with $\hat{\rho}_{\tau}(k) = 0$ corresponding to the case of no directional predictability. The form of the statistic generalizes to the l dimensional multivariate case and the (i, j)th entry of the corresponding cross-correlation matrices $\Gamma_{\bar{\tau}}(k)$ is given by applying (2) for a pair of variables (y_{it}, x_{it}) and (y_{jt-k}, x_{jt-k}) and a pair of conditional quantiles $(\hat{q}_{i,t}(\tau_i), \hat{q}_{j,t-k}(\tau_j))$ for $\bar{\tau} = (\tau_1, \ldots, \tau_l)^{\top}$. The cross-correlation matrices possess the usual symmetry property $\Gamma_{\bar{\tau}}(k) = \Gamma_{\bar{\tau}}(-k)^{\top}$ when $\tau_1 = \cdots = \tau_d$.

Suppose that $\tau \in \mathcal{T}$ and p are given. One may be interested in testing the null hypothesis $H_0 : \rho_{\tau}(1) = \cdots = \rho_{\tau}(p) = 0$ against the alternative hypothesis that $\rho_{\tau}(k) \neq 0$ for some $k \in \{1, \ldots, p\}$. This is a test for the directional predictability of events up to p lags $\{y_{2,t-k} \leq q_{2,t-k}(\tau_2) : k = 1, \ldots, p\}$ for $\{y_{1t} \leq q_{1,t}(\tau_1)\}$. For this hypothesis, we can use the Box–Pierce type statistic $\hat{Q}_{\tau}^{(p)} = T\sum_{k=1}^{p}\hat{\rho}_{\tau}^2(k)$. In practice, we recommend to use the Box–Ljung version $\check{Q}_{\tau}^{(p)} \equiv T(T+2)\sum_{k=1}^{p}\hat{\rho}_{\tau}^2(k)/(T-k)$ which had small sample improvements in our simulations.

On the other hand, one may be interested in testing a stronger null hypothesis, i.e. the absence of directional predictability over a set of quantiles: $H_0 : \rho_{\tau}(1) = \cdots = \rho_{\tau}(p) = 0, \forall \tau \in \mathcal{T}$, against the alternative hypothesis that $\rho_{\tau}(k) \neq 0$ for some $(k, \tau) \in \{1, \ldots, p\} \times \mathcal{T}$ with p fixed. In this case, we can use the sup-version test statistic

$$\sup_{\tau \in \mathcal{T}} \hat{Q}_{\tau}^{(p)} = \sup_{\tau \in \mathcal{T}} T\sum_{k=1}^{p}\hat{\rho}_{\tau}^2(k).$$

Note that the portmanteau test statistic $\hat{Q}_{\tau}^{(p)}$ for a specific quantile is a special case of the sup-version test statistic.

For example, if $\tau_1 = [0.9, 1.0]$ and $\tau_2 = [0.4, 0.6]$, the alternative version measures dependence between an event that y_{1t} is in a high range and an event that $y_{2,t-k}$ is in a mid-range. In some cases, such an alternative version could be easier to interpret and therefore be useful. The inference procedure provided in this paper is also valid for the alternative version of the cross-quantilogram. See the working paper version of this paper for an empirical application using the alternative version.

3. Asymptotic properties

We next present the asymptotic properties of the sample cross-quantilogram and related test statistics. Since these quantities contain non-smooth functions, we employ techniques widely used in the literature on quantile regression, see Koenker and Bassett (1978) and Pollard (1991) among others.

Define $\mathbf{y}_{t,k} = (y_{1t}, y_{2,t-k})^{\top}$, $\mathbf{x}_{t,k} = (x_{1t}, x_{2,t-k})$, $\mathbf{q}_{t,k}(\tau) = [q_{1,t}(\tau_1), q_{2,t-k}(\tau_2)]^{\top}$ and $\hat{\mathbf{q}}_{t,k}(\tau) = [\hat{q}_{1,t}(\tau_1), \hat{q}_{2,t-k}(\tau_2)]^{\top}$ and let $\{\mathbf{y}_{t,k} \leq \mathbf{q}_{t,k}(\tau)\} = \{y_{1t} \leq q_1(\tau_1|x_{1t}), y_{2,t-k} \leq q_2(\tau_2|x_{2t-k})\}$ and $F_{\mathbf{y}|\mathbf{x}}^{(k)}(\cdot|\mathbf{x}_{t,k}) = P(\mathbf{y}_{t,k} \leq \cdot|\mathbf{x}_{t,k})$ for $t = k+1, \ldots, T$ and for some finite integer $k > 0$. We use $\nabla G^{(k)}(\tau)$ to denote $\partial/\partial\mathbf{v}E[F_{\mathbf{y}|\mathbf{x}}^{(k)}(\mathbf{v}_{t,k}|\mathbf{x}_{t,k})]$ evaluated at $\mathbf{v}_{t,k} = \mathbf{q}_{t,k}(\tau)$, where $\mathbf{v}_{t,k} = [x_{1t}^{\top}v_1, x_{2,t-k}^{\top}v_2]^{\top}$ for $v_i \in \mathbb{R}^{d_i}$ $(i = 1, 2)$. Let $d_0 = 1 + d_1 + d_2$.

Assumption

A1. $\{(\mathbf{y}_t, \mathbf{x}_t)\}_{t\in\mathbb{Z}}$ is strictly stationary and strong mixing with coefficients $\{\alpha_j\}_{j\in\mathbb{Z}_+}$ that satisfy $\sum_{j=0}^{\infty}(j+1)^{2s-2}\alpha_j^{\nu/(2s+\nu)} < \infty$ for some integer $s \geq 3$ and $\nu \in (0, 1)$. For each $i = 1, 2$, $E|x_{it}^{(j)}|^{2s+\nu} < \infty$ for all $j = 1, \ldots, d_i$, given $x_{it} = [x_{it}^{(1)}, \ldots, x_{it}^{(d_i)}]^{\top}$.

A2. The conditional distribution function $F_{y_i|x_i}(\cdot|x_{it})$ has continuous densities $f_{y_i|x_i}(\cdot|x_{it})$, which is uniformly bounded away from 0 and ∞ at $q_{i,t}(\tau_i)$ uniformly over $\tau_i \in \mathcal{T}_i$, for $i = 1, 2$ and for all $t \in \mathbb{Z}$.

A3. For any $\epsilon > 0$ there exists a $\nu(\epsilon)$ such that $\sup_{\tau_i\in\mathcal{T}_i}\sup_{s:|s|\leq\nu(\epsilon)} |f_{y_i|x_i}(q_{i,t}(\tau_i)|x_{it}) - f_{y_i|x_i}(q_{i,t}(\tau_i) + s|x_{it})| < \epsilon$ for $i = 1, 2$ and for all $t \in \mathbb{Z}$.

A4. For every $k \in \{1, \ldots, p\}$, the conditional joint distribution $F_{\mathbf{y}|\mathbf{x}}^{(k)}(\cdot|\mathbf{x}_{t,k})$ has the conditional density $f_{\mathbf{y}|\mathbf{x}}^{(k)}(\cdot|\mathbf{x}_{\mathbf{t},\mathbf{k}})$, which is bounded uniformly in the neighborhood of quantiles of interest, and also has a bounded, continuous first derivative for each argument uniformly in the neighborhood of quantiles of interest and thus $\nabla G^{(k)}(\tau)$ exists over $\tau \in \mathcal{T}$.

A5. For each $i = 1, 2$, there exist positive definite matrices M_i and $D_i(\tau_i)$ such that (a) $\text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^{T}x_{it}x_{it}^{\top} = M_i$ and (b) $\text{plim}_{T\to\infty}T^{-1}\sum_{t=1}^{T}f_{y_i|x_i}(q_{i,t}(\tau_i)|x_{it})x_{it}x_{it}^{\top} = D_i(\tau_i)$ uniformly in $\tau_i \in \mathcal{T}_i$.

Assumption A1 imposes the mixing rate used in Andrews and Pollard (1994) and a moment condition on regressors, while allowing for the dependent variables to be processes with infinite moments. For a strong mixing process, $\rho_{\tau}(k) \to 0$ as $k \to \infty$ for all $\tau \in (0, 1)$. Assumption A2 ensures that the conditional quantile function given x_{it} is uniquely defined while allowing for dynamic misspecification, or $P(y_{it} \leq q_{i,t}(\tau_i)|\mathcal{F}_{it}) \neq \tau_i$ given some information set $\mathcal{F}_{it}$ containing all "relevant" information available at t for $i = 1, 2$. In the absence of dynamic misspecification, which is assumed in Hong et al. (2009) under their null hypothesis, the analysis becomes substantially simple because each hit-process $\{\psi_{\tau_i}(y - q_{i,t}(\tau_i))\}$ is a sequence of iid Bernoulli random variables. As Corradi and Swanson (2006) discuss, however, results under correct dynamic specification crucially rely on an appropriate choice of the information set; specification search for the information set based on pre-testing may have a nontrivial impact on inference. Thus, Assumption A2 is appropriate for the purpose of testing directional predictability given a particular information set x_{it}. Assumption A3 implies that the densities are smooth in some neighborhood of the quantiles of interest. Assumption A4 ensures that the joint distribution of (x_{1t}, x_{2t-k}) is continuously differentiable. Assumption A5 is standard in the quantile regression literature.

To describe the asymptotic behavior of the cross-quantilogram, we define a set of d_0-dimensional mean-zero Gaussian process

$\{\mathbb{B}_k(\tau) : \tau \in [0,1]^2\}_{k=1}^p$ with covariance-matrix function for $k, k' \in \{1, \dots, p\}$ and for $\tau, \tau' \in \mathcal{T}$, given by

$$\Xi_{kk'}(\tau,\tau') \equiv E[\mathbb{B}_k(\tau)\mathbb{B}_{k'}^\top(\tau')] = \sum_{l=-\infty}^{\infty} \mathrm{cov}\left(\xi_{l,k}(\tau), \xi_{0,k'}^\top(\tau')\right),$$

where $\xi_{t,k}(\tau) = (1[\mathbf{y_{t,k}} \le \mathbf{q_{t,k}}(\tau)], x_{1t}^\top 1[y_{1t} \le q_{1,t}(\tau_1)], x_{2t}^\top 1[y_{2t} \le q_{2,t}(\tau_2)])^\top$ for $t \in \mathbb{Z}$. Define $\mathbb{B}^{(p)}(\tau) = [\mathbb{B}_1(\tau)^\top, \dots, \mathbb{B}_p(\tau)^\top]^\top$ as the d_0p-dimensional zero-mean Gaussian process with the covariance-matrix function denoted by $\Xi^{(p)}(\tau,\tau')$ for $\tau, \tau' \in \mathcal{T}$. We use $\ell^\infty(\mathcal{T})$ to denote the space of all bounded functions on $\mathcal{T}$ equipped with the uniform topology and $(\ell^\infty(\mathcal{T}))^p$ to denote the p-product space of $\ell^\infty(\mathcal{T})$ equipped with the product topology. Let the notation "$\Rightarrow$" denote the weak convergence due to Hoffman-Jorgensen in order to handle the measurability issues, although outer probabilities and expectations are not used explicitly in this paper for notational simplicity. See Chapter 1 of van der Vaart and Wellner (1996) for a comprehensive treatment of weak convergence in non-separable metric spaces.

The next theorem establishes the asymptotic properties of the cross-quantilogram.

Theorem 1. *Suppose that Assumptions* A1–A5 *hold for some finite integer $p > 0$. Then, in the sense of weak convergence of the stochastic process in $(\ell^\infty(\mathcal{T}))^p$ we have:*

$$\sqrt{T}\left(\hat{\rho}_\tau^{(p)} - \rho_\tau^{(p)}\right) \Rightarrow \Lambda_\tau^{(p)}\mathbb{B}^{(p)}(\tau), \tag{3}$$

where $\hat{\rho}_\tau^{(p)} \equiv [\hat{\rho}_\tau(1), \dots, \hat{\rho}_\tau(p)]^\top$ *and* $\Lambda_\tau^{(p)} = \mathrm{diag}(\lambda_{\tau 1}^\top, \dots, \lambda_{\tau p}^\top)$ *with*

$$\lambda_{\tau,k} = \frac{1}{\sqrt{\tau_1(1-\tau_1)\tau_2(1-\tau_2)}} \times \begin{bmatrix} 1 \\ -\nabla G^{(k)}(\tau)[D_1^{-1}(\tau_1), D_2^{-1}(\tau_2)]^\top \end{bmatrix}. \tag{4}$$

Under the null hypothesis that $\rho_\tau(1) = \cdots = \rho_\tau(p) = 0$ for every $\tau \in \mathcal{T}$, it follows that

$$\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)} \Rightarrow \sup_{\tau\in\mathcal{T}} \|\Lambda_\tau^{(p)}\mathbb{B}^{(p)}(\tau)\|^2, \tag{5}$$

by the continuous mapping theorem.

3.1. Inference methods

3.1.1. The stationary bootstrap

The asymptotic null distribution presented in Theorem 1 depends on nuisance parameters. We suggest to estimate the critical values by the stationary bootstrap of Politis and Romano (1994). The stationary bootstrap is a block bootstrap method with blocks of random lengths. The stationary bootstrap resample is strictly stationary conditional on the original sample.

Let $\{L_i\}_{i\in\mathbb{N}}$ denote a sequence of iid random block lengths having the geometric distribution with a scalar parameter $\gamma \equiv \gamma_T \in (0,1)$: $P^*(L_i = l) = \gamma(1-\gamma)^{l-1}$ for each positive integer l, where P^* denotes the conditional probability given the original sample. We assume that the parameter γ satisfies the following growth condition:

Assumption A6. $T^{\nu/2(2s+\nu)(s-1)}\gamma + (\sqrt{T}\gamma)^{-1} \to 0$ as $T \to \infty$, where s and ν are defined in Assumption A1.

We need the condition that $\gamma = o(T^{-\nu/2(2s+\nu)(s-1)})$ for the purpose of establishing uniform convergence over the subset $\mathcal{T}$ of $[0,1]^2$, given the moment conditions on regressors under Assumption A1. This condition can be relaxed when regressors are uniformly bounded because $\gamma = o(1)$ when $s = \infty$.

Let $\{K_i\}_{i\in\mathbb{N}}$ be a sequence of iid random variables, which have the discrete uniform distribution on $\{k+1, \dots, T\}$ and are independent of both the original data and $\{L_i\}_{i\in\mathbb{N}}$. We set $B_{K_i,L_i} = \{(\mathbf{y}_{t,k}, \mathbf{x}_{t,k})\}_{t=K_i}^{K_i+L_i-1}$ representing the blocks of length L_i starting with the K_ith pair of observations. The stationary bootstrap procedure generates the bootstrap samples $\{(\mathbf{y}_{t,k}^*, \mathbf{x}_{t,k}^*)\}_{t=k+1}^T$ by taking the first $(T-k)$ observations from a sequence of the resampled blocks $\{B_{K_i,L_i}\}_{i\in\mathbb{N}}$. In this notation, when $t > T$, $(\mathbf{y}_{t,k}, \mathbf{x}_{t,k})$ is set to be $(\mathbf{y}_{j,k}, \mathbf{x}_{j,k})$, where $j = k + (t \bmod (T-k))$ and $(\mathbf{y}_{k,k}, \mathbf{x}_{k,k}) = (\mathbf{y}_{t,k}, \mathbf{x}_{t,k})$, where mod denotes the modulo operator.[5]

Using the stationary bootstrap resample, we estimate the parameter $\beta(\tau)$ by solving the minimization problem:

$$\hat{\beta}_1^*(\tau_1) = \arg\min_{\beta_1\in\mathbb{R}^{d_1}} \sum_{t=k+1}^{T} \varrho_{\tau_1}(y_{1t}^* - x_{1t}^{*\top}\beta_1) \quad \text{and}$$

$$\hat{\beta}_2^*(\tau_2) = \arg\min_{\beta_2\in\mathbb{R}^{d_2}} \sum_{t=1}^{T-k} \varrho_{\tau_2}(y_{2t}^* - x_{2t}^{*\top}\beta_2).$$

Then the conditional quantile function given the stationary bootstrap resample, $q_{i,t}^*(\tau_i) \equiv x_{it}^{*\top}\beta_i(\tau_i)$, is estimated by $\hat{q}_{i,t}^*(\tau_i) \equiv x_{it}^{*\top}\hat{\beta}_i^*(\tau_i)$ for each $i = 1, 2$. Define $\hat{\beta}^*(\tau) = [\hat{\beta}_1^{*\top}(\tau_1), \hat{\beta}_2^{*\top}(\tau_2)]^\top$ and let $\hat{\mathbf{q}}_{t,k}^*(\tau) = [\hat{q}_{1,t}^*(\tau_1), \hat{q}_{2,t-k}^*(\tau_2)]^\top$ and $\mathbf{q}_{t,k}^*(\tau) = [q_{1,t}^*(\tau_1), q_{2,t-k}^*(\tau_2)]^\top$. We construct $\hat{\beta}^*(\tau)$ by using $(T-k)$ bootstrap observations, while $\hat{\beta}(\tau)$ is based on T observations, but the difference of sample sizes is asymptotically negligible given the finite lag order k.

The cross-quantilogram based on the stationary bootstrap resample is defined as follows:

$$\hat{\rho}_\tau^*(k) = \frac{\sum_{t=k+1}^{T} \psi_{\tau_1}(y_{1t}^* - \hat{q}_{1,t}^*(\tau_1))\psi_{\tau_2}(y_{2,t-k}^* - \hat{q}_{2,t-k}^*(\tau_2))}{\sqrt{\sum_{t=k+1}^{T} \psi_{\tau_1}^2(y_{1t}^* - \hat{q}_{1,t}^*(\tau_1))}\sqrt{\sum_{t=k+1}^{T} \psi_{\tau_2}^2(y_{2,t-k}^* - \hat{q}_{2,t-k}^*(\tau_2))}}.$$

We consider the stationary bootstrap to construct a confidence interval for each statistic of p cross-quantilograms $\{\hat{\rho}_\tau(1), \dots, \hat{\rho}_\tau(p)\}$ for a finite positive integer p and subsequently construct a confidence interval for the omnibus test based on the p statistics. To maintain the original dependence structure, we use $(T-p)$ pairs of observations $\{[(\mathbf{y}_{t,1}, \mathbf{x}_{t,1}), \dots, (\mathbf{y}_{t,p}, \mathbf{x}_{t,p})]\}_{t=p+1}^T$ to resample the blocks of random lengths.

Given a vector cross-quantilogram $\hat{\rho}_\tau^{(p)*}$, we define the omnibus test based on the stationary bootstrap resample as $\hat{Q}_\tau^{(p)*} = T(\hat{\rho}_\tau^{(p)*} - \hat{\rho}_\tau^{(p)})^\top(\hat{\rho}_\tau^{(p)*} - \hat{\rho}_\tau^{(p)})$. The following theorem shows the validity of the stationary bootstrap procedure for the cross-quantilogram. We use the concept of weak convergence in probability conditional on the original sample, which is denoted by "$\Rightarrow^*$", see van der Vaart and Wellner (1996, p. 181).

Theorem 2. *Suppose that Assumptions* A1–A6 *hold. Then, in the sense of weak convergence conditional on the sample we have:*

(a) $\sqrt{T}\left(\hat{\rho}_\tau^{(p)*} - \hat{\rho}_\tau^{(p)}\right) \Rightarrow^* \Lambda_\tau^{(p)}\mathbb{B}^{(p)}(\tau)$ *in probability;*

(b) *Under the null hypothesis that* $\rho_\tau(1) = \cdots = \rho_\tau(p) = 0$ *for every* $\tau \in \mathcal{T}$,

$$\sup_{z\in\mathbb{R}}\left|P^*\left(\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)*} \le z\right) - P\left(\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)} \le z\right)\right| \to^p 0.$$

[5] For any positive integers a and b, the modulo operation $a \bmod b$ is equal to the remainder, on division of a by b.

In practice, repeating the stationary bootstrap procedure B times, we obtain B sets of cross-quantilograms and $\{\hat{\rho}^{(p)*}_{\tau,b} = [\hat{\rho}^{*}_{\tau,b}(1), \ldots, \hat{\rho}^{*}_{\tau,b}(p)]^{\top}\}_{b=1}^{B}$ and B sets of omnibus tests $\{\hat{Q}^{(p)*}_{\tau,b}\}_{b=1}^{B}$ with $\hat{Q}^{(p)*}_{\tau,b} = T(\hat{\rho}^{(p)*}_{\tau,b} - \hat{\rho}^{(p)}_{\tau})^{\top}(\hat{\rho}^{(p)*}_{\tau,b} - \hat{\rho}^{(p)}_{\tau})$. For testing jointly the null of no directional predictability, a critical value, $c^{*}_{Q,\alpha}$, corresponding to a significance level α is given by the $(1-\alpha)100\%$ percentile of B test statistics $\{\sup_{\alpha\in\mathcal{T}} \hat{Q}^{(p)*}_{\alpha,b}\}_{b=1}^{B}$, that is,

$$c^{*}_{Q,\alpha} = \inf\left\{c : P^{*}\left(\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)*}_{\tau,b} \le c\right) \ge 1-\alpha\right\}.$$

For the individual cross-quantilogram, we pick up percentiles $(c^{*}_{1k,\alpha}, c^{*}_{2k,\alpha})$ of the bootstrap distribution of $\{\sqrt{T}(\hat{\rho}^{*}_{\tau,b}(k) - \hat{\rho}_{\tau}(k))\}_{b=1}^{B}$ such that $P^{*}(c^{*}_{1k,\alpha} \le \sqrt{T}(\hat{\rho}^{*}_{\tau,b}(k) - \hat{\rho}_{\tau}(k)) \le c^{*}_{2k,\alpha}) = 1-\alpha$, in order to obtain a $100(1-\alpha)\%$ confidence interval for $\rho_{\tau}(k)$ given by $[\hat{\rho}_{\tau}(k) + T^{-1/2}c^{*}_{1k,\alpha},\ \hat{\rho}_{\tau}(k) + T^{-1/2}c^{*}_{2k,\alpha}]$.

In the following theorem, we provide a power analysis of the omnibus test statistic $\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)}_{\tau}$ when we use a critical value $c^{*}_{Q,\alpha}$. We consider fixed and local alternatives. The fixed alternative hypothesis against the null of no directional predictability is

$$H_1 : \rho_{\tau}(k) \neq 0 \quad \text{for some } (\tau, k) \in \mathcal{T} \times \{1, \ldots, p\}, \tag{6}$$

and the local alternative hypothesis is given by

$$H_{1T} : \rho_{\tau}(k) = \zeta/\sqrt{T} \quad \text{for some } (\tau, k) \in \mathcal{T} \times \{1, \ldots, p\}, \tag{7}$$

where ζ is a finite non-zero constant. Thus, under the local alternative, there exists a $p \times 1$ vector $\zeta^{(p)}_{\tau}$ such that $\rho^{(p)}_{\tau} = T^{-1/2}\zeta^{(p)}_{\tau}$ with $\zeta^{(p)}_{\tau}$ having at least one non-zero element for some $\tau \in \mathcal{T}$.

We consider the asymptotic power of a test for the directional predictability over a range of quantiles with multiple lags in the following theorem; however, the results can be applied to test for a specific quantile or a specific lag order. The following theorem shows that the cross-quantilogram process has non-trivial local power against the $\sqrt{T}$-local alternatives.

Theorem 3. *Suppose that Assumptions* A1–A6 *hold. Then:* (a) *Under the fixed alternative in* (6),

$$\lim_{T\to\infty} P\left(\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)}_{\tau} > c^{*}_{Q,\alpha}\right) \to 1.$$

(b) *Under the local alternative in* (7)

$$\lim_{T\to\infty} P\left(\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)}_{\tau} > c^{*}_{Q,\alpha}\right) = P\left(\sup_{\tau\in\mathcal{T}} \|\Lambda^{(p)}_{\tau}\mathbb{B}^{(p)}(\tau) + \zeta^{(p)}_{\tau}\|^{2} \ge c_{Q,\alpha}\right),$$

where $c_{Q,\alpha} = \inf\{c : P(\sup_{\tau\in\mathcal{T}} \|\Lambda^{(p)}_{\tau}\mathbb{B}^{(p)}(\tau)\|^{2} \le c) \ge 1-\alpha\}$.

3.1.2. The self-normalized cross-quantilogram

We use recursive estimates to construct a self-normalized cross-quantilogram. The self-normalized approach was proposed by Lobato (2001) and was recently extended by Shao (2010) to a class of asymptotically linear test statistics.[6] The self-normalized approach has a tight link with the fixed-b asymptotic framework proposed by Kiefer et al. (2000).[7] The self-normalized statistic has an asymptotically pivotal distribution whose critical values have been tabulated so that there is no need for long run variance estimation or even bootstrap. As discussed in section 2.1 of Shao (2010), the self-normalized and the fixed-b approach have better size properties, compared with the standard approach involving a consistent asymptotic variance estimator, while it may be asymptotically less powerful under local alternatives (see Lobato (2001) and Sun et al. (2008) for instance).

Given a subsample $\{(\mathbf{y}_t, \mathbf{x}_t)\}_{t=1}^{s}$, we can estimate sample quantile functions by solving minimization problems

$$\hat{\beta}_{i,s}(\tau_i) = \arg\min_{\beta_i\in\mathbb{R}^{d_i}} \sum_{t=1}^{s} \varrho_{\tau_i}\left(y_{it} - x_{it}^{\top}\beta_i\right),$$

for $i = 1, 2$. Let $\hat{q}_{i,t,s}(\tau_i) = x_{it}^{\top}\hat{\beta}_{i,s}(\tau_i)$. We consider the minimum subsample size s larger than $[T\omega]$, where $\omega \in (0, 1)$ is an arbitrary small positive constant. The trimming parameter, ω, is necessary to guarantee that the quantiles estimators based on subsamples have standard asymptotic properties and plays a different role to that of smoothing parameters in long-run variance estimators. Our simulation study suggests that the performance of the test is not sensitive to the trimming parameter.

A key ingredient of the self-normalized statistic is an estimate of cross-correlation based on subsamples:

$$\hat{\rho}_{\tau,s}(k) = \frac{\sum_{t=k+1}^{s} \psi_{\tau_1}(y_{1t} - \hat{q}_{1,t,s}(\tau_1))\psi_{\tau_2}(y_{2,t-k} - \hat{q}_{2,t-k,s}(\tau_2))}{\sqrt{\sum_{t=k+1}^{s} \psi^{2}_{\tau_1}(y_{1t} - \hat{q}_{1,t,s}(\tau_1))}\sqrt{\sum_{t=k+1}^{s} \psi^{2}_{\tau_2}(y_{2,t-k} - \hat{q}_{2,t-k,s}(\tau_2))}},$$

for $[T\omega] \le s \le T$. For a finite integer $p > 0$, let $\hat{\rho}^{(p)}_{\tau,s} = [\hat{\rho}_{\tau,s}(1), \ldots, \hat{\rho}_{\tau,s}(p)]^{\top}$. We construct an outer product of the cross-quantilogram using the subsample

$$\hat{V}_{\tau,p} = T^{-2}\sum_{s=[T\omega]}^{T} s^{2}\left(\hat{\rho}^{(p)}_{\tau,s} - \hat{\rho}^{(p)}_{\tau}\right)\left(\hat{\rho}^{(p)}_{\tau,s} - \hat{\rho}^{(p)}_{\tau}\right)^{\top}.$$

We can obtain the asymptotically pivotal distribution using $\hat{V}_{\tau,p}$ as the asymptotically random normalization. For testing the null of no directional predictability, we define the self-normalized omnibus test statistic

$$\hat{S}^{(p)}_{\tau} = T\hat{\rho}^{(p)\top}_{\tau}\hat{V}^{-1}_{\tau,p}\hat{\rho}^{(p)}_{\tau}.$$

The following theorem shows that $\hat{S}^{(p)}_{\tau}$ is asymptotically pivotal. To distinguish the process used in the following theorem from the one used in the previous section, let $\{\bar{\mathbf{B}}^{(p)}(\cdot)\}$ denote a p-dimensional, standard Brownian motion on $(\ell([0, 1]))^{p}$ equipped with the uniform topology.

Theorem 4. *Suppose that Assumptions* A1–A5 *hold. Then, for each* $\tau \in \mathcal{T}$,

$$\hat{S}^{(p)}_{\tau} \to^{d} \bar{\mathbf{B}}^{(p)}(1)^{\top}\left(\bar{\mathbf{V}}^{(p)}\right)^{-1}\bar{\mathbf{B}}^{(p)}(1),$$

where $\bar{\mathbf{V}}^{(p)} = \int_{\omega}^{1}\{\bar{\mathbf{B}}^{(p)}(r) - r\bar{\mathbf{B}}^{(p)}(1)\}\{\bar{\mathbf{B}}^{(p)}(r) - r\bar{\mathbf{B}}^{(p)}(1)\}^{\top}dr$.

6 Kuan and Lee (2006) apply the approach to a class of specification tests, the so-called M tests, which are based on the moment conditions involving unknown parameters. Chen and Qu (2015) propose a procedure for improving the power of the M test, by dividing the original sample into subsamples before applying the self-normalization procedure.

7 The fixed-b asymptotic has been further studied by Bunzel et al. (2001), Kiefer and Vogelsang (2002, 2005), Sun et al. (2008), Kim and Sun (2011) and Sun and Kim (2012) among others.

The joint test based on finite multiple quantiles can be constructed in a similar manner, while the extension of the self-normalized approach to a range of quantiles is not obvious. The asymptotic null distribution in the above theorem can be simulated and a critical value, $c_{S,\alpha}$, corresponding to a significance level α is tabulated by using the $(1-\alpha)100\%$ percentile of the simulated distribution.[8] In the theorem below, we consider a power function of the self-normalized omnibus test statistic, $P(\hat{S}_\tau^{(p)} > c_{S,\alpha})$. For a fixed $\tau \in \mathcal{T}$, we consider a fixed alternative

$$H_1 : \rho_\tau(k) \neq 0 \quad \text{for some } k \in \{1, \ldots, p\}, \tag{8}$$

and a local alternative

$$H_{1T} : \rho_\tau(k) = \zeta/\sqrt{T} \quad \text{for some } k \in \{1, \ldots, p\}, \tag{9}$$

where ζ is a finite non-zero scalar. This implies that there exists a p-dimensional vector $\zeta_\tau^{(p)}$ such that $\rho_\tau^{(p)} = T^{-1/2}\zeta_\tau^{(p)}$ with $\zeta_\tau^{(p)}$ having at least one non-zero element.

Theorem 5. (a) *Suppose that the fixed alternative in (8) and Assumption* A1–A5 *hold. Then,*

$$\lim_{T\to\infty} P\left(\hat{S}_\tau^{(p)} > c_{S,\alpha}\right) \to 1.$$

(b) *Suppose that the local alternative in (9) is true and Assumptions* A1–A5 *hold. Then,*

$$\lim_{T\to\infty} P\left(\hat{S}_\tau^{(p)} > c_{S,\tau}\right) = P\Bigg(\left\{\bar{\mathbf{B}}^{(p)}(1) + (\Lambda_\tau^{(p)}\Delta_\tau^{(p)})^{-1}\zeta_\tau^{(p)}\right\}^\top \times \left(\mathbf{V}^{(p)}\right)^{-1}\left\{\bar{\mathbf{B}}^{(p)}(1) + (\Lambda_\tau^{(p)}\Delta_\tau^{(p)})^{-1}\zeta_\tau^{(p)}\right\} \geq c_{S,\alpha}\Bigg),$$

where $\Delta_\tau^{(p)}$ is a $d_0p \times d_0p$ matrix with $\Delta_\tau^{(p)}(\Delta_\tau^{(p)})^\top \equiv \Xi^{(p)}(\tau,\tau)$.

4. The partial cross-quantilogram

We define the partial cross-quantilogram, which measures the relationship between two events $\{y_{1t} \leq q_{1,t}(\tau_1)\}$ and $\{y_{2,t-k} \leq q_{2,t-k}(\tau_2)\}$, while controlling for intermediate events between t and $t-k$ as well as whether some state variables exceed a given quantile. Let $\mathbf{z}_t \equiv [\psi_{\tau_3}(y_{3t} - q_{3,t}(\tau_3)), \ldots, \psi_{\tau_l}(y_{lt} - q_{l,t}(\tau_l))]^\top$ be an $(l-2) \times 1$ vector for $l \geq 3$, where $q_{i,t}(\tau_i) = x_{it}^\top\beta_i(\tau_i)$ for τ_i and a $d_i \times 1$ vector x_{it} $(i = 3, \ldots, l)$, and $\mathbf{z}_t$ may include the quantile-hit processes based on some of the lagged predicted variables $\{y_{1,t-1}, \ldots, y_{1,t-k}\}$, the intermediate predictors $\{y_{2,t-1}, \ldots, y_{1,t-k-1}\}$ and some state variables that may reflect some historical events up to t.[9]

For simplicity, we present the results for a single set of quantiles $\bar{\tau} = (\tau_1, \ldots, \tau_l)^\top$ and a single lag k, although the results can be extended to the case of a range of quantiles and multiple lags in an obvious way. To ease the notational burden in the rest of this section, we consider the case for which a lag $k = 0$ without loss of generality and suppress the dependence on k. Let $\bar{\mathbf{y}}_t = [y_{1t}, \ldots, y_{lt}]^\top$ and $\bar{\mathbf{x}}_t = [x_{1t}^\top, \ldots, x_{lt}^\top]^\top$.

We introduce the correlation matrix of the hit processes and its inverse matrix

$$R_{\bar{\tau}} = E\left[h_t(\bar{\tau})h_t(\bar{\tau})^\top\right] \quad \text{and} \quad P_{\bar{\tau}} = R_{\bar{\tau}}^{-1},$$

where an $l \times 1$ vector of the hit process is denoted by $h_t(\bar{\tau}) = [\psi_{\tau_1}(y_{1t} - q_{1,t}(\tau_1)), \ldots, \psi_{\tau_l}(y_{lt} - q_{l,t}(\tau_l))]^\top$. For $i, j \in \{1, \ldots, l\}$, let $r_{\bar{\tau},ij}$ and $p_{\bar{\tau},ij}$ be the (i,j) element of $R_{\bar{\tau}}$ and $P_{\bar{\tau}}$, respectively. Notice that the cross-quantilogram is $r_{\bar{\tau},12}/\sqrt{r_{\bar{\tau},11}r_{\bar{\tau},22}}$, and the partial cross-quantilogram is defined as

$$\rho_{\bar{\tau}|\mathbf{z}} = -\frac{p_{\bar{\tau},12}}{\sqrt{p_{\bar{\tau},11}p_{\bar{\tau},22}}}.$$

The partial cross-correlation also has a form

$$\rho_{\bar{\tau}|\mathbf{z}} = \delta\sqrt{\frac{\tau_1(1-\tau_1)}{\tau_2(1-\tau_2)}},$$

where δ is a scalar parameter defined in the following regression:

$$\psi_{\tau_1}(y_{1t} - q_{1,t}(\tau_1)) = \delta\psi_{\tau_2}(y_{2t} - q_{2,t}(\tau_2)) + \gamma^\top\mathbf{z}_t + u_t,$$

with a $(l-2) \times 1$ vector γ and an error term u_t. Thus, testing the null hypothesis of $\rho_{\bar{\tau}|\mathbf{z}} = 0$ can be viewed as testing predictability between two quantile hits with respect to information $\bar{z}$ as in Granger causality test based on the regression form (Granger, 1969). By choosing relevant variables $\bar{z}$, one can use $\rho_{\bar{\tau}|\mathbf{z}}$ for the purpose of testing Granger causality (Pierce and Haugh, 1977). See also Hong et al. (2009) for testing Granger causality in tail distribution.

To obtain the sample analogue of the partial cross-quantilogram, we first construct a vector of hit processes, $\hat{h}_t(\bar{\tau})$, by replacing the population conditional quantiles in $h_t(\bar{\tau})$ by the sample analogues $\{\hat{q}_{1,t}(\tau_1), \ldots, \hat{q}_{l,t}(\tau_l)\}$. Then, we obtain the estimator for the correlation matrix and its inverse as

$$\hat{R}_{\bar{\tau}} = \frac{1}{T}\sum_{t=1}^{T}\hat{h}_t(\bar{\tau})\hat{h}_t(\bar{\tau})^\top \quad \text{and} \quad \hat{P}_{\bar{\tau}} = \hat{R}_{\bar{\tau}}^{-1},$$

which leads to the sample analogue of the partial cross-quantilogram

$$\hat{\rho}_{\bar{\tau}|\mathbf{z}} = -\frac{\hat{p}_{\bar{\tau},12}}{\sqrt{\hat{p}_{\bar{\tau},11}\hat{p}_{\bar{\tau},22}}}, \tag{10}$$

where $\hat{p}_{\bar{\tau},ij}$ denotes the (i,j) element of $\hat{P}_{\bar{\tau}}$ for $i, j \in \{1, \ldots, l\}$.

In Theorem 6, we show that $\hat{\rho}_{\bar{\tau}|\mathbf{z}}$ asymptotically follows a normal distribution, while the asymptotic variance depends on nuisance parameters as in the previous section. To address the issue of the nuisance parameters, we may employ the stationary bootstrap or the self-normalization technique. For the bootstrap, we can use pairs of variables $\{(\bar{\mathbf{y}}_t, \bar{\mathbf{x}}_t)\}_{t=1}^T$ to generate the stationary bootstrap resample $\{(\bar{\mathbf{y}}_t^*, \bar{\mathbf{x}}_t^*)\}_{t=1}^T$ and then obtain the stationary bootstrap version of the partial cross-quantilogram, denoted by $\hat{\rho}_{\bar{\tau}|\mathbf{z}}^*$, using the formula in (10). When we use the self-normalized test statistics, we estimate the partial cross-quantilogram $\rho_{\bar{\tau},s|\mathbf{z}}$ based on the subsample up to s, recursively and then use

$$\hat{V}_{\bar{\tau}|\mathbf{z}} = T^{-2}\sum_{s=[T\omega]}^{T} s^2\left(\hat{\rho}_{\bar{\tau},s|\mathbf{z}} - \hat{\rho}_{\bar{\tau},T|\mathbf{z}}\right)^2,$$

to normalize the cross-quantilogram, thereby obtaining the asymptotically pivotal statistics.

To obtain the asymptotic results, we impose the following conditions on the conditional distribution function $F_{y_i|x_i}(\cdot|x_{it})$ and its density function $f_{y_i|x_i}(\cdot|x_{it})$ of each pair of additional variables (y_{it}, x_{it}) for $i = 1, \ldots, l$ and on the pairwise joint distribution $F_{ij}(v_1, v_2|x_{it}, x_{jt}) \equiv P(y_{it} \leq v_1, y_{jt} \leq v_2|x_{it}, x_{jt})$ for $(v_1, v_2) \in \mathbb{R}^2$.

Assumption A7. (a) $\{(\bar{\mathbf{y}}_t, \bar{\mathbf{x}}_t)\}_{t\in\mathbb{Z}}$ is a strictly stationary and strong mixing sequence satisfying the condition in Assumption A1; (b) The conditions in Assumptions A2 and A3 hold for the $F_{y_i|x_i}(\cdot|x_{it})$ and $f_{y_i|x_i}(\cdot|x_{it})$ at the relevant quantile for $t = 1, \ldots, T$, for $i = 1$,

[8] We provide the simulated critical values in our R package.

[9] In principle, the intermediate predictors and state variables do not need to be transformed into quantile hits. As emphasized earlier, however, one of the main advantages of considering quantile hits is its applicability to more general time series, being robust to the existence of moments. If needed, it is straightforward to extend the results here to the case of the original variables in $\mathbf{z}_t$ with additional moment conditions. We thank an anonymous referee for pointing this out.

$\dots, l$; (c) $F_{ij}(\cdot|x_{it}, x_{jt})$ satisfies the condition in Assumption A4 and there exists a vector $\nabla_r G_{ij} \equiv \partial/\partial b_r E[F_{ij}(x_{it}^\top b_1, x_{jt}^\top b_2|x_{it}, x_{jt})]$ evaluated at $(b_1, b_2) = (\beta_i(\tau_i), \beta_i(\tau_j))$ for $(r, i, j) \in \{1, 2\} \times \{1, \dots, l\}^2$; (d) There exist positive definite matrices M_i and $D_i(\tau_i)$ as in Assumption A5 for $i = 1, \dots, l$.

Assumption A7(a) requires the same weak dependence property as in Assumption A1. Assumption A7(b)–(c) ensure the smoothness of the marginal conditional distribution, marginal density function and the joint distribution of each pair (y_{it}, y_{jt}) given (x_{it}, x_{jt}) for $1 \le i, j \le l$. Assumption A7(d) is used to derive a Bahadur representation of $\hat{q}_{it}(\tau_i)$ for $i = 1, \dots, l$.

We now state the asymptotic properties of the partial cross-quantilogram and the related inference methods.

Theorem 6. (a) *Suppose that Assumption* A7 *holds. Then,*

$$\sqrt{T}(\hat{\rho}_{\bar{\tau}|\mathbf{z}} - \rho_{\bar{\tau}|\mathbf{z}}) \to^d N(0, \sigma^2_{\bar{\tau}|\mathbf{z}}),$$

for each $\bar{\tau} \in [0, 1]^l$, *where* $\sigma^2_{\bar{\tau}|\mathbf{z}} = \sum_{l=-\infty}^{\infty} \mathrm{cov}(\xi_{\bar{\tau}l}, \xi_{\bar{\tau}0})$ *with*

$$\xi_{\bar{\tau}t} = -\sum_{\substack{1\le i,j\le l\\ i\ne j}} p_{\bar{\tau},1i}p_{\bar{\tau},2j}\psi_{\tau_i}(y_{it} - q_{i,t}(\tau_i))\psi_{\tau_j}(y_{jt} - q_{j,t}(\tau_j))$$
$$+ \sum_{i=1}^{l} \lambda_{\bar{\tau}i}^\top D_i(\tau_i)^{-1} x_{it}\psi_{\tau_i}(y_{it} - q_{i,t}(\tau_i)),$$

and $\lambda_{\bar{\tau}i} = \sum_{\substack{1\le j\le l\\ j\ne i}} \left(p_{\bar{\tau},1i}p_{\bar{\tau},2j} + p_{\bar{\tau},2i}p_{\bar{\tau},1j}\right) \nabla_1 G_{ij}$.

(b) *Suppose that Assumptions* A6 *and* A7 *hold. Then,*

$$\sup_{s\in\mathbb{R}} \left| P^*\left(\hat{\rho}^*_{\bar{\tau}|\mathbf{z}} \le s\right) - P\left(\hat{\rho}_{\bar{\tau}|\mathbf{z}} \le s\right)\right| \to^p 0,$$

for each $\bar{\tau} \in [0, 1]^l$.

(c) *Suppose that Assumption* A7 *holds. Then, under the null hypothesis that* $\rho_{\bar{\tau}|\mathbf{z}} = 0$, *we have*

$$\frac{\sqrt{T}\hat{\rho}_{\bar{\tau}|\mathbf{z}}}{\hat{V}^{1/2}_{\bar{\tau}|\mathbf{z}}} \to^d \frac{\mathbf{B}(1)}{\left\{\int_\omega^1 \{\mathbf{B}(1) - r\mathbf{B}(r)\}^2 dr\right\}^{1/2}},$$

for each $\bar{\tau} \in [0, 1]^l$.

We can show that the partial cross-quantilogram has non-trivial local power against a sequence of $\sqrt{T}$-local alternatives, applying the similar arguments used in Theorems 3 and 5, and thus we omit the details.

5. Monte Carlo simulation

We investigate the finite sample performance of our test statistics. We adopt the following simple VAR model with covariates and consider two data generating processes for the error terms.

$$y_{1t} = 0.1 + 0.3y_{1,t-1} + 0.2y_{2,t-1} + 0.3z_{1t} + u_{1t}$$
$$y_{2t} = 0.1 + 0.2y_{2,t-1} + 0.3z_{2t} + u_{2t},$$

where $z_{it} \sim iid\chi^2(3)/3$ for $i = 1, 2$.

DGP1: $(u_{1t}, u_{2t})^\top \sim iidN(0, I_2)$ where I_2 is a 2×2 identity matrix. We let $(u_{1t}, u_{2t}, z_{1t}, z_{2t})$ be mutually independent.

DGP2:

$$\begin{pmatrix} u_{1t} \\ u_{2t} \end{pmatrix} = \begin{pmatrix} \sigma_{1t} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \end{pmatrix}$$

where $(\varepsilon_{1t}, \varepsilon_{2t})^\top \sim iidN(0, I_2)$ and $\sigma^2_{1t} = 0.1 + 0.2u^2_{1,t-1} + 0.2\sigma^2_{1,t-1} + u^2_{2,t-1}$. We let $(\varepsilon_{1t}, \varepsilon_{2t}, z_{1t}, z_{2t})$ be mutually independent.

The sample cross-quantilogram defined in (2) adopts conditional quantiles $\hat{q}_{it}(\tau_i) = x_{it}^\top \hat{\beta}_i(\tau_i)$. We first estimate $\beta(\tau) \equiv [\beta_1(\tau_1)^\top, \beta_2(\tau_2)^\top]^\top$ by quantile regression of the above VAR model, where $x_{1t} = \left(1, y_{1,t-1}, y_{2,t-1}, z_{1t}\right)^\top$ and $x_{2t} = \left(1, y_{2,t-1}, z_{2t}\right)^\top$ and then obtain the sample cross-quantilogram using $\hat{q}_{it}(\tau_i) = x_{it}^\top \hat{\beta}_i(\tau_i)$.

Under DGP1, there is no predictability from the event $\{y_{2,t-k} \le q_{2,t-k}(\tau_2)\}$ to the event $\{y_{1t} \le q_{1t}(\tau_1)\}$ for all quantiles τ_1 and τ_2, because $\Pr\left[y_{1t} \le q_{1t}(\tau_1) \mid y_{2,t-k}, x_{2,t-k}\right] = \Pr\left[u_{1t} \le \Phi^{-1}(\tau_1)\right] = \tau_1$ for all $t \ge 1$ and $\tau_1 \in (0, 1)$, where Φ denotes the standard normal cdf.

Under DGP2, (u_{1t}) is defined as the GARCH-X process, where its conditional variance is the GARCH(1,1) process with an exogenous covariate. The GARCH-X process is commonly used for modeling volatility of economic or financial time series in the literature, see Han (2015) and references therein. Under DGP2, there exists predictability from $\{y_{2,t-k} \le q_{2,t-k}(\tau_2)\}$ to $\{y_{1t} \le q_{1t}(\tau_1)\}$ through σ^2_{1t} for all quantiles $(\tau_1, \tau_2) \in (0, 1)^2$, except the case $\tau_1 = 0.5$ because the conditional distribution of u_{1t} given x_{1t} is symmetric around 0.[10]

5.1. Results based on the bootstrap procedure

We first examine the finite-sample performance of the Box–Ljung test statistics based on the stationary bootstrap procedure. To save space, only the results for the case where $\tau_1 = \tau_2$ are reported here because the results for the cases where $\tau_1 \ne \tau_2$ are similar. The Box–Ljung test statistics $\hat{Q}^{(p)}_\tau$ are based on $\hat{\rho}_\tau(k)$ for $\tau_i = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$ or 0.95 and $k = 1, 2, \dots, 5$. Tables 1 and 2 report empirical rejection frequencies of the Box–Ljung test statistics based on the bootstrap critical values at the 5% level. The sample sizes considered are $T = 500$, 1000 and 2000. The number of simulation repetitions is 1000. The bootstrap critical values are based on 1000 bootstrapped replicates. The tuning parameter γ is set to be 0.01.[11]

In general, our simulation results in Tables 1–3 show that the test has reasonably good size and power performance in finite samples. Table 1 reports the simulation results for the DGP1, which show the size performance. The rejection frequencies are close to 0.05 in mid quantiles, while the test tends to slightly under-reject in low and high quantiles.

Table 2 reports the simulation results for the DGP2, which show the power performance. Except for the median, the rejection frequencies approach one as the sample size increases, which shows that our test is consistent. As expected, the rejection frequencies are close to 0.05 at the median because there is no predictability at the median under the DGP2 (see Footnote 10 for an explanation).

Next, we examine the finite-sample performance of the sup-version of the Box–Ljung test statistic $\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)}_\tau$ over a range of quantiles.[12] The simulation results in Table 3 show that the sup-version test statistic $\sup_{\tau\in\mathcal{T}} \hat{Q}^{(p)}_\tau$ also has reasonably good finite

[10] To see this, note that the conditional distribution of u_{1t} given x_{1t} has median zero because $\Pr(u_{1t} \le 0 \mid x_{1t}) = \Pr(\sigma_{1t}\varepsilon_{1t} \le 0 \mid x_{1t}) = \Pr(\varepsilon_{1t} \le 0 \mid x_{1t}) = \Pr(\varepsilon_{1t} \le 0) = 0.5$ and likewise $\Pr(u_{1t} \ge 0 \mid x_{1t}) = 0.5$. Therefore, letting $\mathcal{F}_t = (y_{2,t-k}, x_{2,t-k})$, $\Pr\left(y_{1t} < q_{1,t}(0.5) \mid \mathcal{F}_t\right) = \Pr\left(u_{1t} < 0 \mid \mathcal{F}_t\right) = \Pr\left(\varepsilon_{1t} < 0 \mid \mathcal{F}_t\right) = 0.5$. This implies that there is no predictability from $\{y_{2,t-k} \le q_{2,t-k}(\tau_2)\}$ to $\{y_{1t} \le q_{1,t}(\tau_1)\}$ at $\tau_1 = 0.5$ under DGP2.

[11] Recall that $1/\gamma$ indicates the average block length. We tried different values for γ including one chosen by the data dependent rule suggested by Politis and White (2004) and the results are still similar particularly for a large sample. The details of the data dependent rule is explained in Section 6.

[12] Due to computational burden, we compute the Box–Ljung test statistic as a maximum over nine quantile levels $\tau_i = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$ and 0.95.

Table 1
(size) Empirical rejection frequency of the Box–Ljung test statistic $\hat{Q}_{\tau}^{(p)}$ based on the bootstrap procedure (VAR with DGP1 and the nominal level 5%).

T	p	Quantiles ($\tau_1 = \tau_2$)								
		0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.051	0.025	0.037	0.045	0.040	0.043	0.043	0.033	0.047
	2	0.017	0.032	0.043	0.072	0.068	0.060	0.057	0.036	0.012
	3	0.011	0.022	0.051	0.073	0.066	0.055	0.050	0.032	0.010
	4	0.007	0.022	0.047	0.062	0.059	0.057	0.046	0.026	0.008
	5	0.009	0.025	0.035	0.052	0.051	0.052	0.054	0.027	0.006
1000	1	0.033	0.030	0.037	0.048	0.047	0.039	0.037	0.052	0.042
	2	0.018	0.037	0.045	0.051	0.043	0.046	0.052	0.041	0.015
	3	0.011	0.031	0.049	0.056	0.044	0.054	0.045	0.028	0.006
	4	0.013	0.027	0.049	0.053	0.041	0.055	0.041	0.022	0.008
	5	0.007	0.022	0.044	0.040	0.044	0.040	0.036	0.021	0.006
2000	1	0.038	0.034	0.040	0.034	0.034	0.048	0.050	0.034	0.054
	2	0.028	0.025	0.043	0.035	0.045	0.051	0.050	0.035	0.024
	3	0.023	0.033	0.031	0.045	0.050	0.045	0.042	0.029	0.018
	4	0.017	0.023	0.042	0.052	0.038	0.036	0.038	0.025	0.016
	5	0.009	0.025	0.038	0.038	0.035	0.035	0.034	0.019	0.014

Notes: The first and second columns report the sample size T and the number of lags p for the Box–Ljung test statistics $\hat{Q}_{\tau}^{(p)}$, respectively. The rest of columns show empirical rejection frequencies based on bootstrap critical values at the 5% significance level. The tuning parameter γ is set to be 0.01.

Table 2
(power) Empirical rejection frequency of the Box–Ljung test statistic $\hat{Q}_{\tau}^{(p)}$ based on the bootstrap procedure (VAR with DGP2 (GARCH-X process)).

T	p	Quantiles ($\tau_1 = \tau_2$)								
		0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.361	0.701	0.722	0.383	0.042	0.383	0.713	0.684	0.362
	2	0.303	0.610	0.584	0.257	0.063	0.231	0.589	0.589	0.300
	3	0.270	0.541	0.491	0.202	0.053	0.174	0.467	0.515	0.246
	4	0.230	0.451	0.403	0.172	0.058	0.126	0.378	0.447	0.208
	5	0.203	0.393	0.344	0.134	0.060	0.115	0.314	0.386	0.177
1000	1	0.751	0.948	0.942	0.638	0.048	0.619	0.951	0.952	0.760
	2	0.708	0.916	0.912	0.425	0.046	0.431	0.908	0.932	0.712
	3	0.651	0.877	0.845	0.322	0.052	0.315	0.849	0.897	0.651
	4	0.589	0.838	0.784	0.255	0.048	0.250	0.778	0.854	0.596
	5	0.537	0.801	0.716	0.203	0.042	0.190	0.714	0.809	0.563
2000	1	0.969	0.999	0.999	0.905	0.044	0.923	0.999	0.998	0.974
	2	0.965	1.000	0.999	0.808	0.053	0.817	0.999	1.000	0.979
	3	0.959	1.000	0.997	0.688	0.053	0.673	0.998	1.000	0.967
	4	0.944	1.000	0.990	0.585	0.047	0.573	0.994	0.999	0.957
	5	0.930	1.000	0.982	0.510	0.037	0.485	0.987	0.997	0.938

Notes: Same as Table 1.

sample performance, though it tends to under-reject under DGP1. For DGP2, the rejection frequencies approach one as the sample size increases.

5.2. Results for the self-normalized statistics

We also examine the performance of the self-normalized version of $\hat{Q}_{\tau}^{(p)}$ under the same setup as above. We fix the trimming constant ω to be 0.1.[13] The number of repetitions is 3000. The empirical sizes of the test are reported in Table 4, where the underlying process is the VAR model with DGP1. The test generally under-rejects under the null hypothesis (DGP1), while at the extreme quantiles ($\tau = 0.05$ or 0.95) the test slightly over-rejects in the small sample ($T = 500$). This finding is not very surprising because the self-normalized statistic is based on subsamples and at the extreme quantiles there are effectively not enough observations to compute the test statistic accurately.

Using the GARCH-X process of DGP2, we obtain empirical powers and present the results in Table 5. With a one-period lag ($p = 1$), the self-normalized quantilogram at $\tau_1, \tau_2 \in \{0.1, 0.2, 0.8, 0.9\}$ rejects the null by about 23.0%–30.0%, 64.3%–68.3% and 91.7%–94.0% for sample sizes 500, 1000 and 2000, respectively. In general, the rejection frequencies increase as the sample size increases, decline as the maximum number of lags p increases, and are not sensitive to the choice of the trimming value. Our results suggest that the self-normalized statistics may have lower power in finite samples compared with the test statistics based on the stationary bootstrap procedure, see Lobato (2001) for a similar finding.

6. Empirical studies

6.1. Stock return predictability

We apply the cross-quantilogram to detect directional predictability from an economic state variable to stock returns. The issue of stock return predictability has been very important and extensively investigated in the literature; see Lettau and Ludvigson (2010) for an extensive review. A large literature has considered predictability of the mean of stock return. The typical mean return forecast examines whether the mean of an economic state variable is helpful in predicting the mean of stock return (mean-to-mean

13 We also considered 0.03 and 0.05 for ω and the results are similar to those for $\omega = 0.1$.

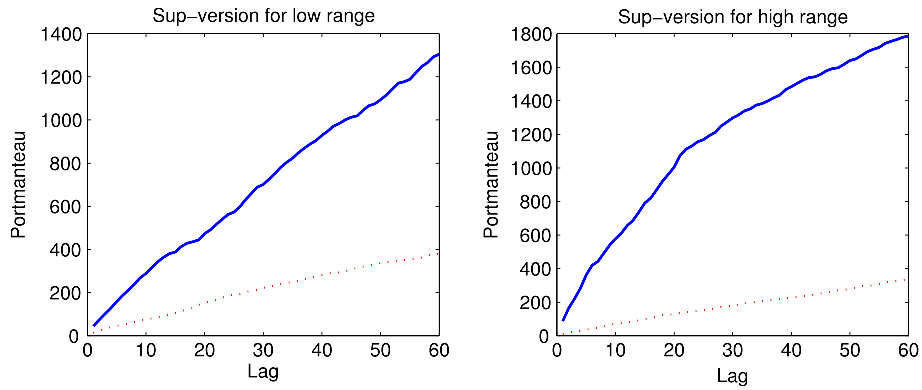

Fig. 1. Sup-version Box–Ljung test statistic $\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)}$ for each lag p to detect directional predictability from stock variance to stock return. For the low range, we set $\mathcal{T}=[0.1,0.3]$ and $\tau_i=0.1+0.02k$ for $k=0,1,\ldots,10$. We let $\tau_1=\tau_2$ for $\hat{\rho}_\tau(k)$. For the high range, we set $\mathcal{T}=[0.7,0.9]$ and $\tau_i=0.7+0.02k$ for $k=0,1,\ldots,10$. The dashed lines are the 95% bootstrap confidence intervals centered at the null hypothesis.

Table 3
Empirical rejection frequencies of the sup-version of the Box–Ljung test statistic $\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)}$ based on the bootstrap procedure (VAR with DGP1/DGP2 and the nominal level 5%).

T	p	DGP1 (size)	DGP2 (power)
500	1	0.004	0.624
	2	0.007	0.460
	3	0.008	0.356
	4	0.008	0.265
	5	0.009	0.221
1000	1	0.004	0.976
	2	0.011	0.946
	3	0.006	0.895
	4	0.003	0.825
	5	0.007	0.765
2000	1	0.012	1.000
	2	0.015	1.000
	3	0.020	1.000
	4	0.020	1.000
	5	0.017	0.999

Notes: The first and second columns report the sample size T and the number of lags p for the sup-version of the Box–Ljung test statistic $\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)}$, respectively. The sup-version test statistic is the Box–Ljung test statistic maximized over nine quantiles $\tau_i=0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$ and 0.95. The third and fourth columns show empirical rejection frequencies based on bootstrap critical values at the 5% significance level. The tuning parameter γ is set to be 0.01.

relationship). Recently, Cenesizoglu and Timmermann (2008) considered whether the mean of an economic state variable is helpful in predicting different quantiles of stock returns representing left tail, right tail or shoulders of the return distribution. The cross-quantilogram adds one more dimension to analyze predictability compared with the linear quantile regression, and so it provides a more complete picture on the relationship between a predictor and stock returns. Moreover, we can consider very large lags in the framework of the quantilogram.

We use daily data from 3 Jan. 1996 to 29 Dec. 2006 with sample size 2717.[14] Stock returns are measured by the log price difference of the S&P 500 index and we employ stock variance as the predictor. The stock variance is treated as an estimate of equity risk in the literature. The risk–return relationship is an important issue in the finance literature; see Lettau and Ludvigson (2010) for an extensive review. The cross-quantilogram can provide a more complete relationship from risk to return, which cannot be examined using existing methods. To measure stock variance, we use the realized variance given by the sum of squared 5-minute returns.[15] The autoregressive coefficient for stock variance is estimated to be 0.68 and the unit root hypothesis is clearly rejected. The sample mean and median of stock returns are 0.0003 and 0.0005, respectively.

In Figs. 1–3, we provide the sup-type test statistic $\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)}$, the cross-quantilogram $\hat{\rho}_\tau(k)$ and the portmanteau test $\hat{Q}_\tau^{(p)}$ (we use the Box–Ljung versions throughout) to detect directional predictability from stock variance, representing risk, to stock return. In each graph, we show the 95% bootstrap confidence intervals for no predictability based on 1000 bootstrapped replicates. The tuning parameter $1/\gamma$ is chosen by adapting the rule suggested by Politis and White (2004) (and later corrected in Patton et al. (2009)).[16] Since it is for univariate data, we apply it separately to each time series and define γ as the average value.

We first examine the sup-version Box–Ljung test statistic $\sup_{\tau\in\mathcal{T}}\hat{Q}_\tau^{(p)}$ and the results are provided in Fig. 1. We consider low and high ranges of quantiles. For the low range, we set $\mathcal{T}=[0.1,0.3]$ and $\tau_i=0.1+0.02k$ for $k=0,1,\ldots,10$. For the high range, we set $\mathcal{T}=[0.7,0.9]$ and $\tau_i=0.7+0.02k$ for $k=0,1,\ldots,10$. In each range, there are eleven different values of τ_i and we let $\tau_1=\tau_2$ in calculating $\hat{\rho}_\tau(k)$ for simplicity. Fig. 1 clearly shows that there exists predictability from stock variance to stock return in each range.

Next we investigate the cross-quantilogram $\hat{\rho}_\tau(k)$ and the portmanteau test $\hat{Q}_\tau^{(p)}$ for different quantile points in Figs. 2(a)–3(b). For the quantiles of stock return $q_1(\tau_1)$, we consider $\tau_1=0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$ and 0.95. For the quantiles of stock variance $q_2(\tau_2)$, we consider $\tau_2=0.1$ and 0.9. Figs. 2(a) and 2(b) are for the case when the stock variance is in the low quantile, i.e. $\tau_2=0.1$. The cross-quantilograms $\hat{\rho}_\tau(k)$ for $\tau_1=0.05, 0.1, 0.2$ and 0.3 are negative and significant for many lags. For example, in case of $\tau_1=0.05$, it means that when risk is very low, it is less likely to have a large negative loss. On the other hand, the cross-quantilograms for $\tau_1=0.7, 0.8, 0.9$ and 0.95 is positive and significant for many lags. For example, in case of $\tau_1=0.95$, it means that when risk is very low, it is less likely to have a large positive gain. However, the cross-quantilogram for $\tau_1=0.5$ is mostly insignificant, which means that risk is not helpful in predicting whether stock return is located below or above its median. Fig. 2(b) shows that the Box–Ljung test statistics are mostly significant except for $\tau_1=0.5$.

14 The working paper version of this paper provides the results using the monthly data previously analyzed in Goyal and Welch (2008).

15 The realized variance is obtained from 'Oxford-Man Institute's realized library'.

16 Specifically, $1/\hat{\gamma}=(2\hat{G}^2/\hat{D}_{SB})^{1/3}T^{1/3}$ where $\hat{D}_{SB}=2\hat{g}^2(0)$. The definitions of $\hat{g}$ and $\hat{G}$ are given on page 58 of Politis and White (2004).

Table 4
(size) Empirical rejection frequencies of the self-normalized statistics (VAR with DGP1 and the nominal level: 5%).

T	p	Quantiles ($\tau_1 = \tau_2$)								
		0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.043	0.000	0.000	0.007	0.003	0.013	0.007	0.000	0.047
	2	0.090	0.010	0.007	0.003	0.003	0.003	0.000	0.003	0.127
	3	0.130	0.007	0.000	0.007	0.003	0.000	0.003	0.000	0.143
	4	0.150	0.007	0.000	0.000	0.000	0.000	0.000	0.000	0.167
	5	0.187	0.003	0.000	0.000	0.000	0.000	0.000	0.000	0.177
1000	1	0.010	0.013	0.010	0.013	0.020	0.003	0.007	0.003	0.007
	2	0.023	0.007	0.000	0.007	0.000	0.003	0.003	0.007	0.037
	3	0.040	0.003	0.010	0.000	0.007	0.003	0.007	0.000	0.047
	4	0.043	0.000	0.007	0.000	0.007	0.003	0.003	0.000	0.047
	5	0.047	0.000	0.007	0.000	0.000	0.000	0.003	0.000	0.053
2000	1	0.013	0.030	0.017	0.017	0.033	0.013	0.020	0.017	0.027
	2	0.007	0.000	0.007	0.007	0.027	0.010	0.027	0.017	0.020
	3	0.017	0.000	0.003	0.003	0.013	0.010	0.003	0.003	0.013
	4	0.013	0.000	0.003	0.000	0.010	0.007	0.003	0.000	0.013
	5	0.010	0.003	0.003	0.000	0.007	0.003	0.000	0.000	0.017

Notes: The first and second columns report the sample size T and the number of lags p for the test statistics $\hat{Q}_{\tau}^{(p)}$, respectively. The rest of columns show empirical rejection frequencies given simulated critical values at 5% significance level. The trimming value ω is set to be 0.1.

Table 5
(power) Empirical rejection frequencies of the self-normalized statistics (VAR with DGP2: GARCH-X process).

T	p	Quantiles ($\tau_1 = \tau_2$)								
		0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.067	0.230	0.297	0.077	0.007	0.150	0.300	0.253	0.050
	2	0.030	0.070	0.113	0.033	0.000	0.037	0.113	0.077	0.010
	3	0.047	0.010	0.043	0.010	0.000	0.017	0.023	0.020	0.023
	4	0.063	0.007	0.023	0.000	0.000	0.010	0.013	0.003	0.050
	5	0.120	0.003	0.007	0.003	0.000	0.003	0.003	0.000	0.080
1000	1	0.347	0.643	0.683	0.313	0.010	0.323	0.673	0.663	0.317
	2	0.153	0.523	0.527	0.177	0.020	0.180	0.543	0.463	0.157
	3	0.063	0.300	0.347	0.090	0.010	0.097	0.377	0.283	0.063
	4	0.033	0.210	0.223	0.050	0.000	0.037	0.243	0.153	0.017
	5	0.047	0.097	0.133	0.030	0.000	0.023	0.127	0.097	0.020
2000	1	0.757	0.917	0.923	0.663	0.030	0.693	0.940	0.920	0.707
	2	0.577	0.873	0.917	0.513	0.013	0.540	0.883	0.863	0.577
	3	0.427	0.787	0.860	0.400	0.007	0.397	0.800	0.810	0.390
	4	0.270	0.680	0.807	0.323	0.017	0.297	0.740	0.680	0.250
	5	0.197	0.567	0.700	0.223	0.003	0.213	0.680	0.590	0.163

Notes: Same as Table 4.

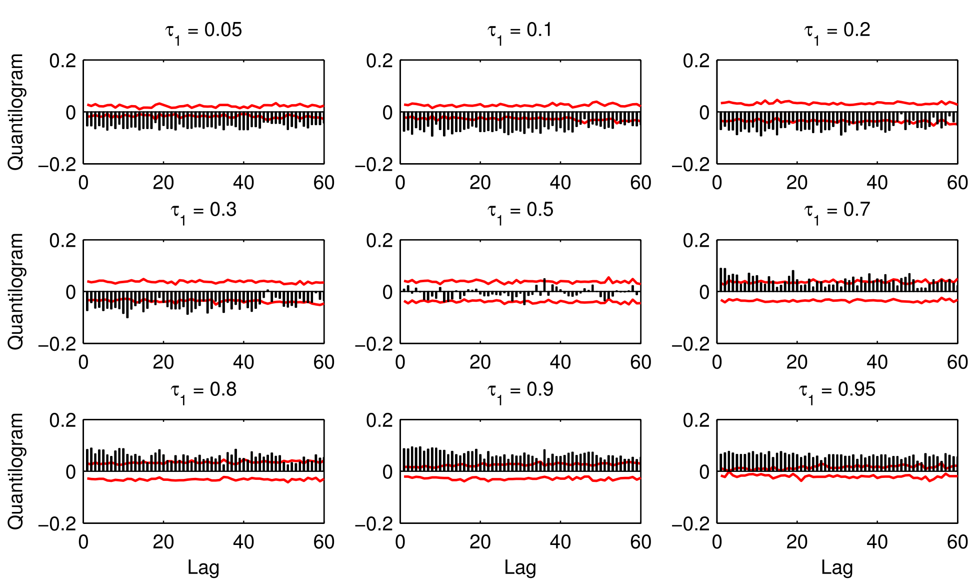

Fig. 2(a). The sample cross-quantilogram $\hat{\rho}_{\tau}(k)$ for $\tau_2 = 0.1$ to detect directional predictability from stock variance to stock return. Bar graphs describe sample cross-quantilograms and lines are the 95% bootstrap confidence intervals centered at zero.

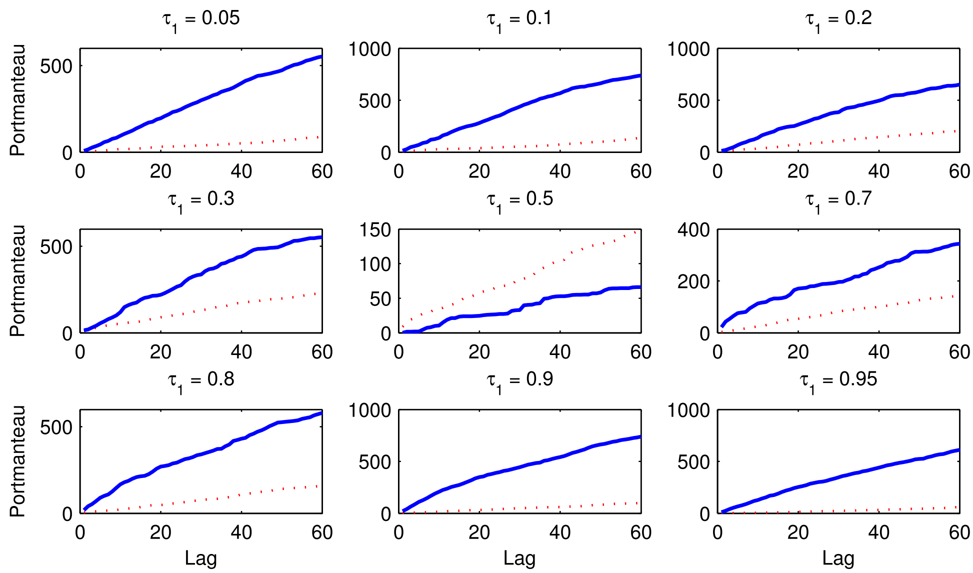

Fig. 2(b). Box–Ljung test statistic $\hat{Q}_{\tau}^{(p)}$ for each lag p and quantile τ using $\hat{\rho}_{\tau}(k)$ with $\tau_2 = 0.1$. The dashed lines are the 95% bootstrap confidence intervals centered at zero.

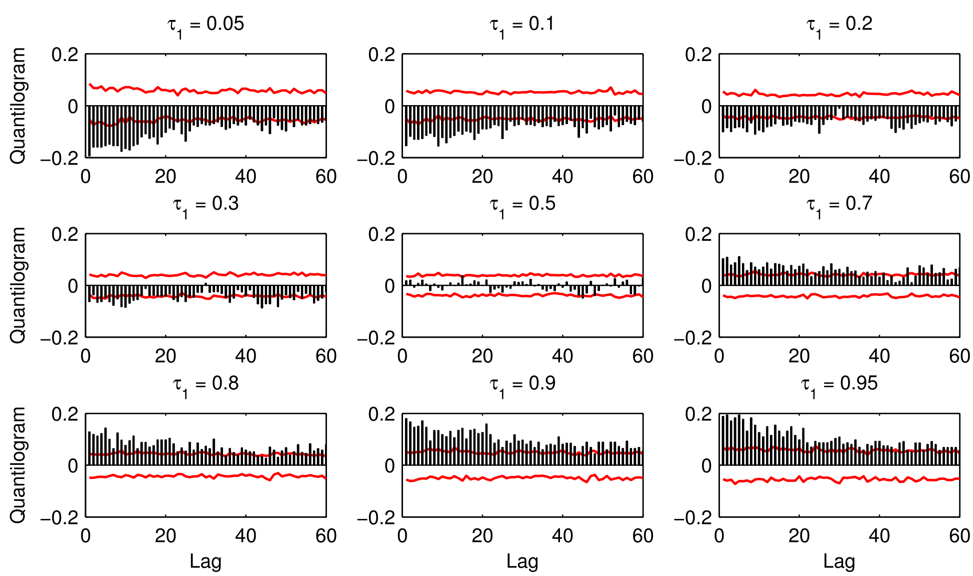

Fig. 3(a). The sample cross-quantilogram $\hat{\rho}_{\tau}(k)$ with $\tau_2 = 0.9$ to detect directional predictability from stock variance to stock return. Same as Fig. 1(a).

Figs. 3(a) and 3(b) are for the case when stock variance is in the high quantile, i.e. $\tau_2 = 0.9$. Compared to the previous case of $\tau_2 = 0.1$, the cross-quantilograms have similar trends but much larger absolute values. For $\tau_1 = 0.05$, the cross-quantilogram $\hat{\rho}_{\tau}(1)$ is -0.193, which implies that when risk is higher than its 0.9 quantile, there is an increased likelihood of having a very large negative loss in the next day. For $\tau_1 = 0.95$, the cross-quantilogram $\hat{\rho}_{\tau}(1)$ is 0.188, which implies that when risk is high (higher than its 0.9 quantile), there is an increased likelihood of having a very large positive gain in the next day. The cross-quantilogram for $\tau_1 = 0.5$ is mostly insignificant and the Box–Ljung test statistics in Fig. 3(b) are mostly significant except for $\tau_1 = 0.5$.

The results in Figs. 1–3 show that stock variance is helpful in predicting stock return and detailed features depend on each quantile of stock variance and stock return. When stock variance is in high quantile, the absolute value of the cross-quantilogram is higher and the cross-quantilogram is significantly different from zero for larger lags. Our results exhibit a more complete relationship between risk and return and additionally show how the relationship changes for different lags.

6.2. Systemic risk

The Great Recession of 2007–2009 has motivated researchers to better understand systemic risk—the risk that the intermediation capacity of the entire financial system can be impaired, with potentially adverse consequences for the supply of credit to the real economy. One approach to measure systemic risk is measuring co-dependence in the tails of equity returns of an individual financial institution and the financial system.[17] Prominent examples include the work of Adrian and Brunnermeier (2011), Brownlees and Engle (2012) and White et al. (2012). Since the cross-quantilogram measures quantile dependence between time series, we apply it to measure systemic risk.

We use the daily CRSP market value weighted index return as the market index return as in Brownlees and Engle (2012). We consider returns on JP Morgan Chase (JPM), Morgan Stanley (MS) and

[17] Bisias et al. (2012) categorize the current approaches to measuring systemic risk along the following lines: (1) tail measures, (2) contingent claims analysis, (3) network models, and (4) dynamic stochastic macroeconomic models.

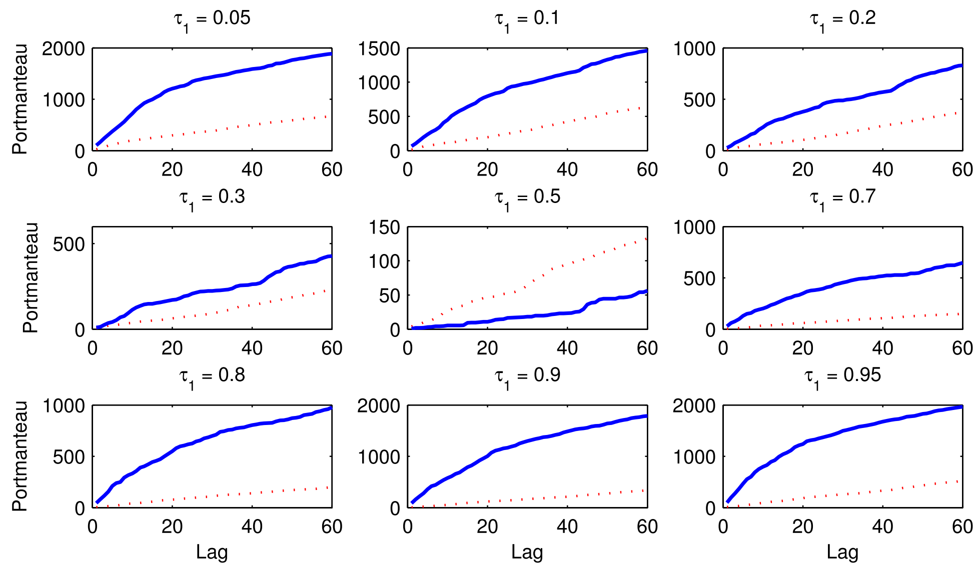

Fig. 3(b). Box–Ljung test statistic $\hat{Q}_\tau^{(p)}$ for each lag p and quantile τ using $\hat{\rho}_\tau(k)$ with $\tau_2 = 0.9$. Same as Fig. 1(b).

AIG as individual financial institutions. As in Brownlees and Engle (2012), JPM, MS and AIG belong to the Depositories group, the Broker–Dealers group and the Insurance group, respectively. We investigate the cross-quantilogram $\hat{\rho}_\tau(k)$ between an individual institution's stock return and the market index return for $k = 60$ and $\tau_1 = \tau_2 = 0.05$. In each graph, we show the 95% bootstrap confidence intervals for no quantile dependence based on 1000 bootstrapped replicates.

The sample period is from 24 Feb. 1993 to 31 Dec. 2014 with sample size 5505.[18] The data including the financial crisis from 2007 and 2009 might not be suitable to be viewed as a strictly stationary sequence and hence may not fit into our theoretical framework.[19] Nevertheless, we provide the empirical results because it would be practically interesting to consider a sample period that includes the recent crisis and post-crisis.[20]

In Fig. 4, each graph in the left column shows the cross-quantilogram from each individual institution to the market. The cross-quantilograms are positive and generally significant for large lags. The cross-quantilogram from JPM to the market reaches its peak (0.146) at $k = 12$ and declines steadily afterwards. This means that it takes about two weeks for the systemic risk from JPM to reach its peak once JPM is in distress. From MS to the market, the cross-quantilogram reaches its peak (0.127) at $k = 2$. From AIG to the market, the cross-quantilogram reaches its peak (0.127) at $k = 17$. When AIG is in distress, the systemic risk from AIG takes a longer time (about three weeks) to reach its peak. When an individual financial institution is in distress, each institution makes an influence on the market in a different way.

Each graph in the right column of Fig. 4 shows the cross-quantilograms from the market to an individual institution. The cross-quantilogram for this case is a measure of an individual institution's exposure to system wide distress and therefore it is similar to the stress tests performed by individual institutions. From the market to each institutions, the cross-quantilogram at $k = 1$ is relatively low for JPM (0.062) and MS (0.073) while it is higher for AIG (0.104). Overall, when the market is in distress, each institution is influenced by its impact in a different way. But the cross-quantilogram reaches its peak at $k = 2$ for all cases. The cross-quantilograms at $k = 2$ are 0.135, 0.131 and 0.139 for JPM, MS and AIG, respectively.

As shown in Fig. 4, the cross-quantilogram is a measure for either an institution's systemic risk or an institution's exposure to system wide distress. Compared to existing methods, one important feature of the cross-quantilogram is that it provides in a simple manner how such a measure changes as the lag k increases. For example, White et al. (2012) adopt an additional impulse response function within the multivariate and multi-quantile framework to consider tail dependence for a large k. Moreover, another feature of the cross-quantilogram is that it does not require any modeling. For example, the approach by Brownlees and Engle (2012) is based on the standard multivariate GARCH model and it requires the modeling of the entire multivariate distribution.

Next, we apply the partial cross-quantilogram to examine the systemic risk after controlling for an economic state variable. Following Adrian and Brunnermeier (2011) and Bedljkovic (2010), we adopt the VIX index as the economic state variable. Since the VIX index itself is highly persistent and can be modeled as an integrated process, we instead use the VIX index change, the first difference of the VIX index level, as the state variable. For the quantile of the state variable, i.e. τ_3 in (10), we let $\tau_3 = 0.95$ because a low quantile of a stock return is generally associated with a rapid increase of the VIX index.

Fig. 5 shows that the partial cross-quantilograms are still significant in some cases even if their values are generally lower than the values of the cross-quantilograms in Fig. 4. This indicates that there still remains systemic risk from an individual institution after controlling for an economic state variable. These significant partial cross-quantilograms will be of interest for the management of the systemic risk of an individual financial institution.

7. Conclusion

We have established the limiting properties of the cross-quantilogram in the case of a finite number of lags. Hong (1996) established the properties of the Box–Pierce statistic in the case that $p = p_n \to \infty$: after a location and scale adjustment the statistic

[18] The stock return series of Morgan Stanley are available from 24 Feb. 1993. The stock return series of individual financial institutions are obtained from Yahoo Finance.

[19] A rigorous treatment of nonstationary time series in our context is a challenging issue and will be reported in a future work.

[20] The results for the sample period from 24 Feb. 1993 to 29 Dec. 2006 are also available from the authors upon request.

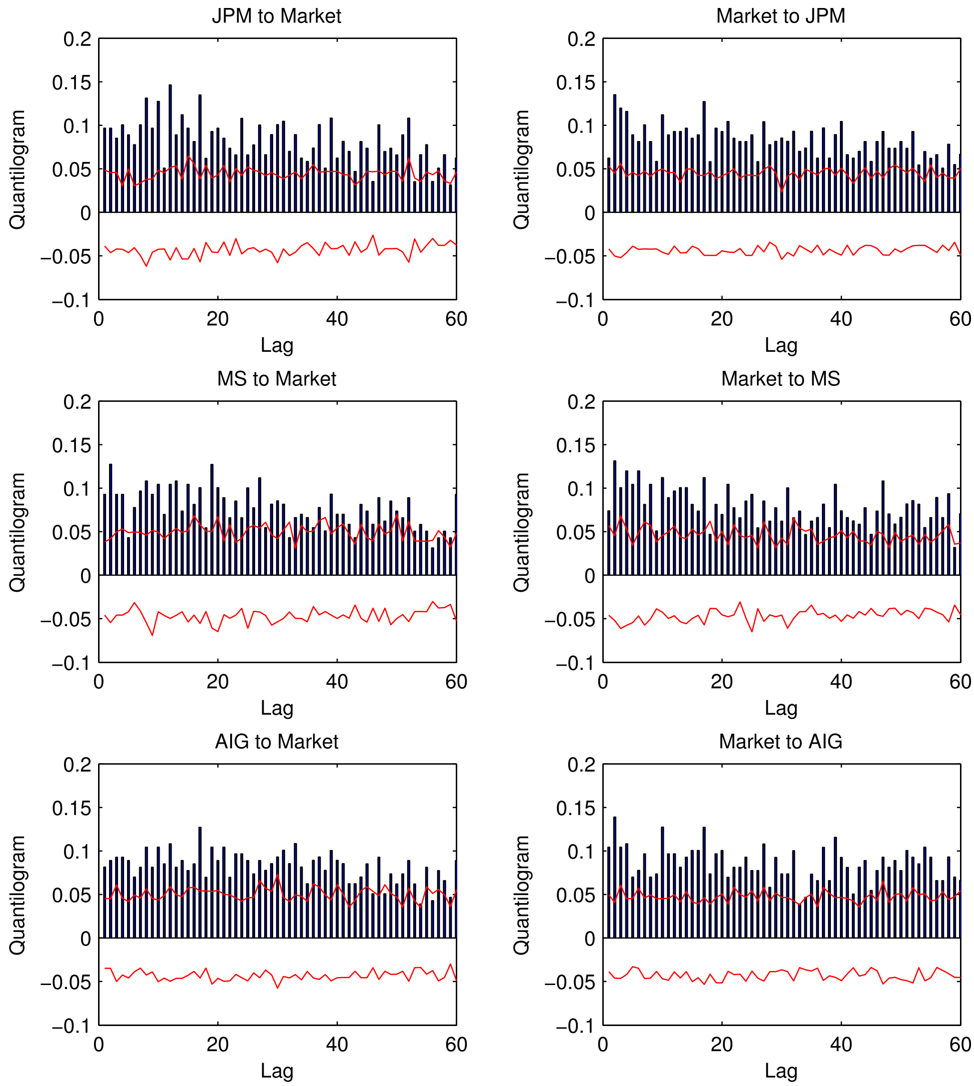

Fig. 4. The sample cross-quantilogram $\hat{\rho}_\tau(k)$. Bar graphs describe sample cross-quantilograms and lines are the 95% bootstrap confidence intervals centered at zero.

is asymptotically normal, see also Hong et al. (2009) for a related work. No doubt our results can be extended to accommodate this case, although in practice the desirability of such a test is questionable, and the chi-squared type limit in our theory may provide better critical values for even quite long lags. The cross-quantilogram is easy to compute and the bootstrap confidence intervals appear to represent modest enlargements of the Bartlett intervals in the series that we examined. The statistic shows the cross dependence structure of the time series in a granular fashion that is more informative than the usual methods.

Appendix

In appendix, we use $C, C_1, C_2, \ldots$ to denote generic positive constants without further clarification.

Appendix A. Asymptotic results of cross-quantilogram

Lemma A.1. *Let $\{z_t\}_{t\in\mathbb{Z}}$ be a strict stationary, strong mixing sequence of $\mathbb{R}^d$-valued random variables for some integer $d \geq 1$ with strong mixing coefficients $\{\alpha_j\}_{j\in\mathbb{Z}_+}$ satisfying $\sum_{j=0}^{\infty}(j+1)^{2s-2}\alpha_j^{\nu/(2s+\nu)}$ for some integer $s \geq 2$ and $\nu \in (0,1)$. Suppose that $E[z_1] = 0$ and $\|z_1\|_{2s+\nu} < \infty$. Then,*

$$E\left\|\sum_{t=1}^{T} z_t\right\|^{2s} \leq T^s C\left\{\|z_1\|_{2+\nu}^{2s} + T^{1-s}\|z_1\|_{2s+\nu}^{2s}\right\}.$$

Proof. See Supplemental material, Appendix E. ■

We define the process indexed by $\tau \in \mathcal{T}$:

$$\mathbb{V}_{t,k}(\tau) := \frac{1}{\sqrt{T}}\sum_{t=k+1}^{T}\left\{1[\mathbf{y}_{t,k} \leq \mathbf{q}_{t,k}(\tau)] - E[F_{\mathbf{y}|\mathbf{x}}^{(k)}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})]\right\}.$$

Also, define a $d_i \times 1$ vector of random variables indexed by $\tau_i \in \mathcal{T}_i$ for each $i = 1, 2$:

$$\mathbb{W}_{i,T}(\tau_i) := \frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{it}\psi_{\tau_i}\left(y_{it} - q_{i,t}(\tau_i)\right).$$

The below lemma shows the stochastic equicontinuity of the processes defined above, using a similar argument in Bai (1996).

Proposition A.1. *Suppose Assumptions* A1–A5 *hold. Let $k \in \{1,\ldots,p\}$ and define metrics $\rho_i(\tau_i,\tau_i') = |\tau_i' - \tau_i|$ for $\tau_i, \tau_i' \in \mathcal{T}_i$ $(i = 1, 2)$ and a metric $\rho(\tau,\tau') = \sum_{i=1}^{2}\rho_i(\tau_i,\tau_i')$ for $\tau, \tau' \in \mathcal{T}$. Then,*

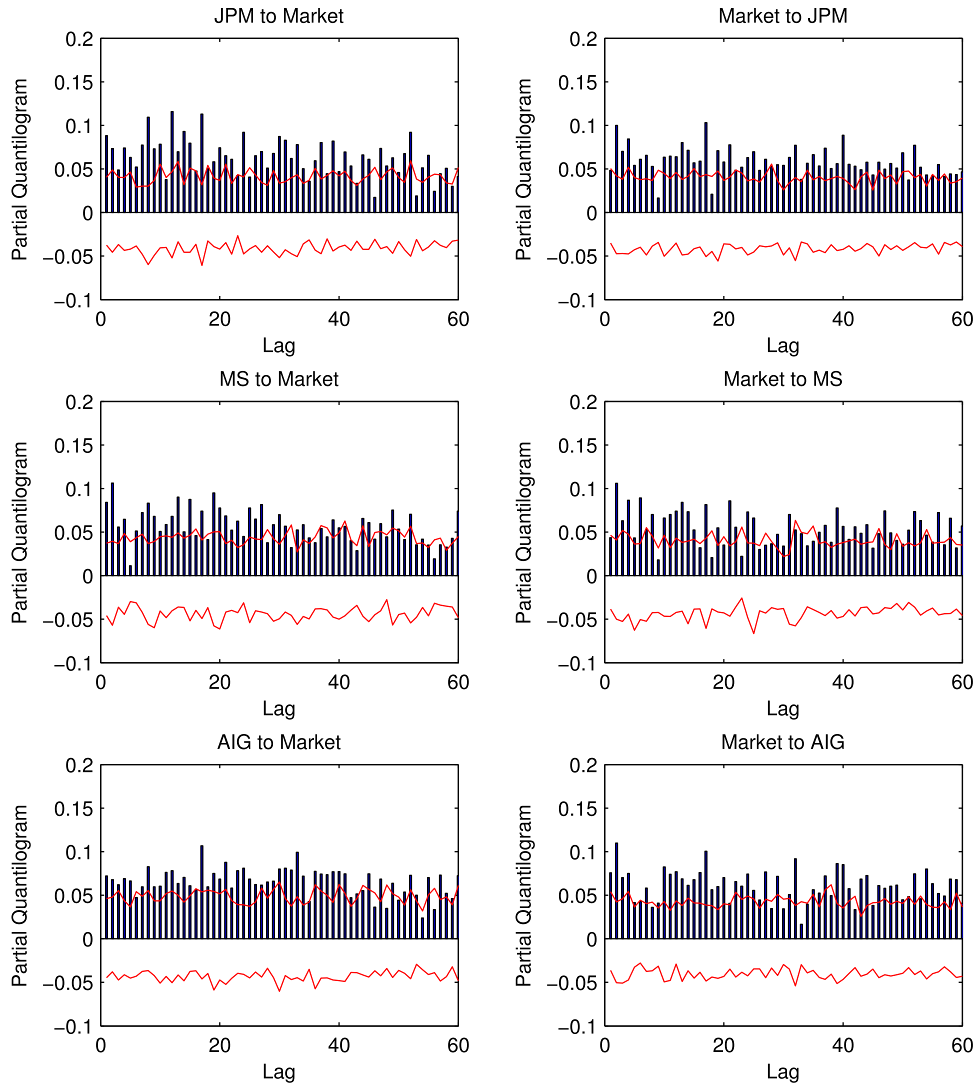

Fig. 5. The sample partial cross-quantilogram $\hat{\rho}_{\bar{\tau}|\mathbf{z}}(k)$. Bar graphs describe sample partial cross-quantilograms and lines are the 95% bootstrap confidence intervals centered at zero.

(a) $\mathbb{V}_{T,k}(\tau)$ *is stochastically equicontinuous on* $(\mathcal{T}, \rho)$;
(b) $\mathbb{W}_{i,T}(\tau_i)$ *is stochastically equicontinuous on* $(\mathcal{T}_i, \rho_i)$ *for each* $i = 1, 2$.

Proof. See Supplemental material, Appendix E. ■

Because of the importance of the result, we present the central limit theorem for strong mixing sequence in the lemma below. The proof can be found in Corollary 5.1 of Hall and Heyde (1980) or Rio (1997, 2013) among others.

Lemma A.2. *Suppose that the strict stationary sequence* $\{z_t\}_{t\in\mathbb{Z}}$ *satisfies the strong mixing condition with* $E[z_1] = 0$ *and* $E|z_1|^{2+\varsigma} < \infty$ *for some* $\varsigma \in (0, \infty)$, *while* $\sum_{j=1}^{\infty} \alpha_j^{\varsigma/(2+\varsigma)} < \infty$. *Then,* $\lim_{T\to\infty} E[(T^{-1/2}\sum_{t=1}^{T} z_t)^2] = \sigma^2$ *for some* $\sigma^2 \in [0, \infty)$. *If* $\sigma^2 > 0$, *then* $\sigma^{-1}T^{-1/2}\sum_{t=1}^{T} z_t \to^d N(0, 1)$.

Define a $d_0 \times 1$ vector $\mathbb{B}_{T,k}(\tau) = [\mathbb{V}_{T,k}(\tau), \mathbb{W}_{1,T}(\tau_1)^\top, \mathbb{W}_{2,T}(\tau_2)^\top]^\top$ for $\tau \in \mathcal{T}$ and $k = 1, \ldots, p$. The following proposition shows the weak convergence of the process $\{\mathbb{B}_{T,k}(\tau) : \tau \in \mathcal{T}\}_{k=1}^{p}$.

Proposition A.2. *Suppose Assumptions* A1–A5 *hold. Then,*

$$\left[\mathbb{B}_{T,1}(\cdot), \ldots, \mathbb{B}_{T,p}(\cdot)\right]^\top \Rightarrow \left[\mathbb{B}_1(\cdot), \ldots, \mathbb{B}_p(\cdot)\right]^\top .$$

Proof. Proposition A.1 shows that $[\mathbb{B}_{T,1}(\cdot), \ldots, \mathbb{B}_{T,p}(\cdot)]^\top$ is stochastic equicontinuous. Thus, it remains to establish convergence of the finite dimensional distributions. By the Cramer–Wold device, it suffices to show

$$\sum_{j=1}^{J} \theta_j \sum_{k=1}^{p} \kappa_k^\top \mathbb{B}_{T,k}\left(\tau^{(j)}\right) \to^d N\left(0, \sigma_{\theta,\kappa}^2\right),$$

for any $\{\theta_j \in \mathbb{R}\}_{j=1}^{J}$, $\{\kappa_k \in \mathbb{R}^d\}_{k=1}^{p}$, $\{\tau^{(j)} \in [0, 1]^2\}_{j=1}^{J}$, and $J \geq 1$, where

$$\sigma_{\theta,\kappa}^2 = \sum_{j=1}^{J}\sum_{j'=1}^{J} \theta_j\theta_{j'} \sum_{k=1}^{p}\sum_{k'=1}^{p} \kappa_k^\top \Xi_{k,k'}(\tau^{(j)}, \tau^{(j')})\kappa_{k'}. \tag{A.1}$$

The original time-series is a stationary sequence satisfying the strong mixing condition in Assumption A1 and a measurable transformation involving lagged variables satisfies the same mixing condition if the lag order is finite. Hence, the central limit theorem for strong-mixing sequences in Lemma A.2 shows that the convergence in distribution to the normal law with the finite variance. Therefore, we establish the weak convergence. ■

Let $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ and $\mathbf{v}_{t,k} = (v_{1,t}, v_{2,t-k})^\top \in \mathbb{R}^2$ with $v_{i,t} = x_{it}^\top v_i$ for $i = 1, 2$ and for $t = 1, \ldots, T$.

Define

$$\mathbb{V}_{T,k}(\tau,\mathbf{v}) := \frac{1}{\sqrt{T}}\sum_{t=k+1}^{T}\Big\{1[\mathbf{y}_{t,k}\le \mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}] - E[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}|\mathbf{x}_{t,k})]\Big\},$$

and

$$\mathbb{W}_{i,T}(\tau_i,v_i) := \frac{1}{\sqrt{T}}\sum_{t=1}^{T}x_{it}\left\{1[y_{it}\le q_{i,t}(\tau_i)+T^{-1/2}v_{i,t}] - F_{y_i|x_i}(q_{i,t}(\tau_i)+T^{-1/2}v_{i,t}|x_{it})\right\}.$$

Proposition A.3. *Suppose Assumptions* A1–A5 *hold. Then,*

(a) $\sup_{\tau\in\mathcal{T}}\sup_{\mathbf{v}\in\mathcal{V}_M}|\mathbb{V}_{T,k}(\tau,\mathbf{v})-\mathbb{V}_{T,k}(\tau)| = o_p(1)$ *for every* $M>0$*;*
(b) $\sup_{\tau_i\in\mathcal{T}_i}\sup_{v_i\in\mathcal{V}_{i,M}}\|\mathbb{W}_{i,T}(\tau_i,v_i)-\mathbb{W}_{i,T}(\tau_i)\| = o_p(1)$ *for every* $M>0$ *and* $i=1,2$*,*

where $\mathcal{V}_M=\mathcal{V}_{1,M}\times\mathcal{V}_{2,M}$ *with* $\mathcal{V}_{i,M}=\{v_i\in R^{d_i}:\|v_i\|\le M\}$ *for* $i=1,2$*.*

Proof. See Supplemental material, Appendix E. ■

Proposition A.4. *Suppose Assumptions* A1–A5 *hold. Then, for* $i=1,2$

$$\sqrt{T}\{\hat\beta_i(\tau_i)-\beta_i(\tau_i)\} = -D_i^{-1}(\tau_i)\mathbb{W}_{i,T}(\tau_i)+o_p(1),$$

uniformly in $\tau\in\mathcal{T}_i$*.*

Proof. See Supplemental material, Appendix E. ■

The below lemma shows that the limiting behavior of the cross-quantilogram process reflects the contributions of estimation errors due to the estimation of the conditional quantile function.

Proposition A.5. *Suppose that Assumptions* A1–A5 *hold. Then, for each* $k\in\{1,\dots,p\}$*,*

$$\sqrt{T}\left\{\hat\rho_\tau(k)-\rho_\tau(k)\right\} = \frac{\mathbb{V}_{T,k}(\tau)+\nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta(\tau)-\beta(\tau)\}}{\sqrt{\tau_1(1-\tau_1)\tau_2(1-\tau_2)}} + o_p(1),$$

uniformly in $\tau\in\mathcal{T}$*.*

Proof. Let $\hat\gamma_{\tau,k} = T^{-1}\sum_{t=k+1}^{T}\psi_{\tau_1}(y_{1t}-\hat q_{1,t}(\tau_1))\psi_{\tau_2}(y_{2,t-k}-\hat q_{2,t-k}(\tau_2))$ and $\gamma_{\tau,k}=E[\psi_{\tau_1}(y_{1t}-q_{1,t}(\tau_1))\psi_{\tau_2}(y_{2,t-k}-q_{2,t-k}(\tau_2))]$. Using a similar argument in Lemma 2.1 of Arcones (1988), we can show $\sup_{\tau_i\in\mathcal{T}_i}|T^{-1/2}\sum_{t=1}^{T}\psi_{\tau_i}(y_{it}-\hat q_{i,t}(\tau_i))| = o_p(1)$ for $i=1,2$, because x_{it} includes a constant term. It follows that, uniformly in $\tau\in\mathcal{T}$,

$$T^{-1}\sum_{t=1}^{T}\psi^2_{\tau_i}(y_{it}-\hat q_{i,t}(\tau_i)) = \tau_i(1-\tau_i)+o_p(1), \quad \text{for } i=1,2, \tag{A.2}$$

and

$$\sqrt{T}\left(\hat\gamma_{\tau,k}-\gamma_{\tau,k}\right) = T^{-1/2}\sum_{t=k+1}^{T}\Big\{1[\mathbf{y}_{t,k}\le\hat{\mathbf{q}}_{t,k}(\tau)] - E\Big[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})\Big]\Big\}+o_p(1).$$

Define $\mathcal{V}_M=\{\mathbf{v}\equiv(v_1,v_2)\in\mathbb{R}^{d_1}\times\mathbb{R}^{d_2}:\max_{i=1,2}\|v_i\|\le M\}$ for some $M>0$ and let $\mathbf{v}_{t,k}=(x_{1t}^\top v_1, x_{2,t-k}^\top v_2)^\top$. Then, Proposition A.3 implies

$$\begin{aligned}
&T^{-1/2}\sum_{t=k+1}^{T}\Big\{1[\mathbf{y}_{t,k}\le \mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}] - E\Big[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})\Big]\Big\}\\
&= \mathbb{V}_{T,k}(\tau)+\sqrt{T}E\Big[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}|\mathbf{x}_{t,k}) - F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})\Big]+o_p(1),
\end{aligned}$$

uniformly in $(\tau,\mathbf{v})\in\mathcal{T}\times\mathcal{V}_M$ for any $M>0$. Also, the mean-value theorem implies $\sqrt{T}E[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}|\mathbf{x}_{t,k}) - F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})] = \nabla G^{(k)}(\tau)^\top\mathbf{v}+o(1)$ uniformly in $(\tau,\mathbf{v})\in\mathcal{T}\times\mathcal{V}_M$. Thus, for any $M>0$,

$$\sup_{(\tau,\mathbf{v})\in\mathcal{T}\times\mathcal{V}_M}|R_T(\tau,\mathbf{v})| = o_p(1), \tag{A.3}$$

where

$$R_T(\tau,\mathbf{v}) := T^{-1/2}\sum_{t=k+1}^{T}\Big\{1[\mathbf{y}_{t,k}\le\mathbf{q}_{t,k}(\tau)+T^{-1/2}\mathbf{v}_{t,k}] - E\Big[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})\Big]\Big\} - \Big(\mathbb{V}_{T,k}(\tau)+\nabla G^{(k)}(\tau)^\top\mathbf{v}\Big).$$

Let ϵ be an arbitrary positive constant. Propositions A.2 and A.4 imply that there exists a constant $M>0$ such that $P(\sup_{\tau\in\mathcal{T}}\|\hat\beta(\tau)-\beta(\tau)\|>M/\sqrt{T})<\epsilon$ for a sufficiently large T. It follows that there exists an $M>0$ such that

$$\begin{aligned}
&P\left(\sup_{\tau\in\mathcal{T}}\left|R_T(\tau,\sqrt{T}\{\hat\beta(\tau)-\beta(\tau)\})\right|>\boldsymbol{\epsilon}\right)\\
&<\epsilon+P\Big(\sup_{(\tau,\mathbf{v})\in\mathcal{T}\times\mathcal{V}_M}|R_T(\tau,\mathbf{v})|>\boldsymbol{\epsilon}\Big),
\end{aligned}$$

for a sufficiently large T. Thus, (A.3) yields

$$\sqrt{T}\left(\hat\gamma_{\tau,k}-\gamma_{\tau,k}\right) = \mathbb{V}_{T,k}(\tau)+\nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta(\tau)-\beta(\tau)\}+o_p(1),$$

uniformly in $\tau\in\mathcal{T}$. This together with (A.2) yields the desired result. ■

Proof of Theorem 1. For each $i=1,2$, Proposition A.4 yields an asymptotic linear approximation, $\sqrt{T}\{\hat\beta_i(\tau_i)-\beta_i(\tau_i)\} = -D_i^{-1}(\tau_i)\mathbb{W}_{i,T}(\tau_i)+o_p(1)$ uniformly in $\tau_i\in\mathcal{T}_i$, which with Proposition A.5 shows that $\sqrt{T}\left\{\hat\rho_\tau(k)-\rho_\tau(k)\right\} = \lambda_{\tau,k}^\top\mathbb{B}_{T,k}(\tau)+o_p(1)$ uniformly in $\tau\in\mathcal{T}$. For a finite $p>0$, we have

$$\sqrt{T}\left(\hat\rho^{(p)}_\tau-\rho^{(p)}_\tau\right) = \Lambda^{(p)}_\tau\mathbb{B}^{(p)}_T(\tau)+o_p(1), \tag{A.4}$$

uniformly in $\tau\in\mathcal{T}$. The desired result is obtained from Proposition A.2 with the continuous mapping theorem. ■

Appendix B. Stationary bootstrap

A positive integer valued, possibly infinite random variable μ is said to be a *stopping time* with respect to a filtration $\{\mathcal{F}_n, n\ge1\}$ if $\{\mu=n\}\in\mathcal{F}_n$, $\forall n\in\mathbb{N}$. Given random block lengths $\{L_i\}_{i\in\mathbb{N}}$ under the stationary bootstrap, define $N=\inf\{i\in\mathbb{N}:L_1+\cdots+L_i\ge n\}$. Then, N is a stopping time with respect to $\{\sigma(L_1,\dots,L_i):1\le i\le n\}$. In the following lemma, we present a moment inequality using ideas found in the literature on the stopped random walk process. See Gut (2009) for a comprehensive treatment.

Lemma B.1. *Let* $\{z_t\}_{t\in\mathbb{Z}}$ *be a strict stationary, strong mixing sequence of* $\mathbb{R}^d$*-valued random variables for some integer* $d\ge1$ *with strong mixing coefficients* $\{\alpha_j\}_{j\in\mathbb{Z}_+}$ *satisfying* $\sum_{j=0}^{\infty}(j+1)^{2s-2}\alpha_j^{\nu/(2s+\nu)}$ *for some integer* $s\ge2$ *and* $\nu\in(0,1)$*. Suppose that* $\|z_1\|_{2s+\nu}<\infty$ *and a stationary bootstrap resample,* $\{z_t^*\}_{t=1}^T$*, from* $\{z_t\}_{t=1}^T$ *satisfies*

Assumption A6 with the sample size $T > 0$. Define $S_{k,l} = \sum_{t=k}^{k+l-1} z_t$ and $S^*_{k,l} = \sum_{t=k}^{k+l-1} z^*_t$. Then,

$$E\left\|S^*_{1,T} - E^*S^*_{1,T}\right\|^{2s} \le C\left\{(T\gamma)^s \sum_{l=1}^{\infty} \pi_l E\left\|\tilde{S}_{1,l}\right\|^{2s} + E\left\|\tilde{S}_{1,T}\right\|^{2s}\right\},$$

where $\tilde{S}_{k,l} = \sum_{t=k}^{k+l-1}(z_t - Ez_t)$ for $k, l \in \mathbb{N}$.

Proof. See Supplemental material, Appendix E. ■

Lemma B.2. *Suppose that the same conditions assumed in Lemma B.1 hold. Then,*

$$E\left\|S^*_{1,T} - E^*S^*_{1,T}\right\|^{2s} \le T^s C\left(\|z_1\|^{2s}_{2+\nu} + \gamma^{s-1}\|z_1\|^{2s}_{2s+\nu}\right),$$

for a sufficiently large T.

Proof. See Supplemental material, Appendix E. ■

We now turn to the asymptotic results of cross-quantilogram based on the stationary bootstrap. Define

$$\mathbb{V}^*_{T,k}(\tau) := \frac{1}{\sqrt{T}} \sum_{t=k+1}^{T} \left\{1[\mathbf{y}^*_{t,k} \le \mathbf{q}^*_{t,k}(\tau)] - 1[\mathbf{y}_{t,k} \le \mathbf{q}_{t,k}(\tau)]\right\}$$

and

$$\mathbb{W}^*_{i,T}(\tau_i) := \frac{1}{\sqrt{T}} \sum_{t=k+1}^{T} \left\{x^*_{it}\psi_{\tau_i}\left(y^*_{it} - q^*_{i,t}(\tau_i)\right) - x_{it}\psi_{\tau_i}\left(y_{it} - q_{i,t}(\tau_i)\right)\right\}$$

for each $i = 1, 2$. The lemma below shows the stochastic equicontinuity of the processes, $\mathbb{V}^*_{T,k}(\cdot)$ and $\mathbb{W}^*_{i,T}(\cdot)$, *unconditional* on the original sample.

Proposition B.1. *Suppose Assumptions A1–A6 hold. Let $k \in \{1, \ldots, p\}$ and define metrics $\rho_i(\cdot,\cdot)$ for $i = 1, 2$ and a metric $\rho(\cdot,\cdot)$ as in Proposition A.1. Then,*

(a) *$\mathbb{V}^*_{T,k}(\tau)$ is stochastically equicontinuous on $(\mathcal{T}, \rho)$;*
(b) *$\mathbb{W}^*_{i,T}(\tau_i)$ is stochastically equicontinuous on $(\mathcal{T}_i, \rho_i)$ for each $i = 1, 2$.*

Proof. See Supplemental material, Appendix E. ■

Let $\mathbb{B}^*_{T,k}(\tau) = [\mathbb{V}^*_{T,k}(\tau), \mathbb{W}^*_{1,T}(\tau_1)^\top, \mathbb{W}^*_{2,T}(\tau_2)^\top]^\top$ for $(k, \tau) \in \{1, \ldots, p\} \times \mathcal{T}$ and define $\mathbb{B}^{(p)*}_{T,k}(\tau) := [\mathbb{B}^*_{T,1}(\tau), \ldots, \mathbb{B}^*_{T,p}(\tau)]^\top$. As a norm that introduces the topology of $(\ell^\infty(\mathcal{T}))^{pd_0}$, we use $\sup_{\tau\in\mathcal{T}} \|\cdot\|$ defined on $(\ell^\infty(\mathcal{T}))^{pd_0}$, so that $\sup_{\tau\in\mathcal{T}} \|f(\tau)\|$ for any $f \in (\ell^\infty(\mathcal{T}))^{pd_0}$. Let BL_1 be the set of all Lipschitz continuous, real-valued functions on $(\ell^\infty(\mathcal{T}))^{pd_0}$ with a Lipschitz constant bounded by 1. We prove the following proposition by modifying the argument used in Theorem 2 of Galvao et al. (2014), where the approach of van der Vaart and Wellner (1996, Theorem 2.9.6) is extended for the dependent process but their setup differs from the one here.

Proposition B.2. *Suppose Assumptions A1–A6 hold. Then,*

$$\sup_{h\in BL_1} \left|E^*\left[h(\mathbb{B}^{(p)*}_T)\right] - E\left[h(\mathbb{B}^{(p)})\right]\right| \to^p 0.$$

Proof. Let $\delta > 0$. Given the compact set $\mathcal{T}$ in $[0, 1]^2$, there exists a finite partition $\{\mathcal{T}^{(j)}\}_{j=1}^J$ such that $\max_{1\le j\le J} \sup_{\tau',\tau''\in\mathcal{T}^{(j)}} \|\tau'' - \tau'\| \le \delta$. Pick up $\tau^{(j)} \equiv (\tau^{(j)}_1, \tau^{(j)}_2)^\top \in \mathcal{T}^{(j)}$ for $j = 1, \ldots, J$ and let Π_δ be a map from $\mathcal{T}$ to $\{\tau^{(j)}\}_{j=1}^J$ so that $\Pi_\delta(\tau) = \tau^{(j)}$ if $\tau \in \mathcal{T}^{(j)}$. Define $\mathbb{B}^{(p)*}_T \circ \Pi_\delta$ and $\mathbb{B}^{(p)} \circ \Pi_\delta$ as the stochastic processes on $\mathcal{T}$, given by $\mathbb{B}^{(p)*}_T \circ \Pi_\delta(\tau) = \mathbb{B}^{(p)*}_T(\Pi_\delta(\tau))$ and $\mathbb{B}^{(p)} \circ \Pi_\delta(\tau) = \mathbb{B}^{(p)}(\Pi_\delta(\tau))$ for $\tau \in \mathcal{T}$. It follows from the triangle inequality that, for any $h \in BL_1$,

$$\left|E^*\left[h(\mathbb{B}^{(p)*}_T)\right] - E\left[h(\mathbb{B}^{(p)})\right]\right|$$
$$\le \left|E^*\left[h(\mathbb{B}^{(p)*}_T)\right] - E^*\left[h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right]\right| \tag{B.5}$$
$$+ \left|E^*\left[h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right] - E\left[h(\mathbb{B}^{(p)} \circ \Pi_\delta)\right]\right| \tag{B.6}$$
$$+ \left|E\left[h(\mathbb{B}^{(p)} \circ \Pi_\delta)\right] - E\left[h(\mathbb{B}^{(p)})\right]\right|. \tag{B.7}$$

It suffices to show that (B.5)-(B.7) are $o_p(1)$ uniformly in $h \in BL_1$.

We first consider (B.5). We have

$$E\left[\sup_{h\in BL_1}\left|E^*\left[h(\mathbb{B}^{(p)*}_T)\right] - E^*\left[h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right]\right|\right] \le E\left[\sup_{h\in BL_1}\left|h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right|\right].$$

Let $I^*_{T,\delta,\epsilon} := 1[\sup_{\tau\in\mathcal{T}} \|\mathbb{B}^{(p)*}_T(\tau) - \mathbb{B}^{(p)*}_T \circ \Pi_\delta(\tau)\| > \epsilon]$ for $\epsilon > 0$. Proposition B.1 implies that $\lim_{\delta\downarrow 0}\lim_{T\to\infty} E[I^*_{T,\delta,\epsilon}] < \epsilon$ for every $\epsilon > 0$. Also $\sup_{h\in BL_1} |h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)| \le 2$ because the range of a function h is $[-1, 1]$. It follows that

$$\lim_{\delta\downarrow 0}\lim_{T\to\infty} E\left[\sup_{h\in BL_1}\left|h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right| \cdot I^*_{T,\delta,\epsilon}\right] \le 2\epsilon.$$

Since $\sup_{h\in BL_1} |h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)| \le \sup_{\tau\in\mathcal{T}} \|\mathbb{B}^{(p)*}_T(\tau) - \mathbb{B}^{(p)*}_T \circ \Pi_\delta(\tau)\|$, we have

$$E\left[\sup_{h\in BL_1}\left|h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)\right| \cdot (1 - I^*_{T,\delta,\epsilon})\right] \le \epsilon.$$

Thus, $\lim_{\delta\downarrow 0}\lim_{T\to\infty} E[\sup_{h\in BL_1} |h(\mathbb{B}^{(p)*}_T) - h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)|] \le 3\epsilon$. An application of the Markov inequality yields that (B.5) is $o_p(1)$ uniformly in $h \in BL_1$.

Next we shall show that $\sup_{h\in BL_1} |E^*[h(\mathbb{B}^{(p)*}_T \circ \Pi_\delta)] - E[h(\mathbb{B}^{(p)} \circ \Pi_\delta)]| \to^p 0$ for any $\delta > 0$. It suffices to show that $\{\mathbb{B}^{(p)*}_T(\tau^{(j)})\}_{j=1}^J \to^d \{\mathbb{B}^{(p)}(\tau^{(j)})\}_{j=1}^J$ conditional on the original sample, for almost every sequence. To this end, we use the Cramer–Wold device and consider $\sum_{j=1}^J \theta_j \sum_{k=1}^p \kappa_k^\top \mathbb{B}^*_{T,k}\left(\tau^{(j)}\right)$ for some $\{\theta_j \in \mathbb{R}\}_{j=1}^J$ and $\{\kappa_k \in \mathbb{R}^d\}_{k=1}^p$. Let $v^*_t = \sum_{j=1}^J \theta_j \sum_{k=1}^p \kappa_k^\top \xi^*_{t,k}(\tau^{(j)})$ and $v_t = \sum_{j=1}^J \theta_j \sum_{k=1}^p \kappa_k^\top \xi_{t,k}(\tau^{(j)})$, where $\xi_{t,k}(\cdot)$ is defined in Section 3 and $\xi^*_{t,k}(\cdot)$ is its bootstrap counterpart. Then, we can write

$$\sum_{j=1}^J \theta_j \sum_{k=1}^p \kappa_k^\top \mathbb{B}^*_{T,k}\left(\tau^{(j)}\right) = T^{-1/2} \sum_{t=k+1}^T (v^*_t - v_t).$$

As discussed in Proposition A.2, $\{v_t\}_{t\in\mathbb{N}}$ is a stationary time-series satisfying Assumption 1. As shown in p. 1237 of Kunsch (1989), the moment and strong-mixing assumption imposed on the original time series implies the condition imposed on the fourth joint cumulant in (8) of Politis and Romano (1994). Hence, Theorems 1 and 2 of Politis and Romano (1994) imply that the bootstrap estimate of the variance converges to $\sigma^2_{\theta,\kappa}$ in probability, where $\sigma^2_{\theta,\kappa}$ is defined in (A.1), and that we obtained the distribution convergence conditional on the original sample.

Finally, consider (B.7). The process $\mathbb{B}^{(p)}$ is uniformly continuous on $\mathcal{T}$, which with the dominated convergence theorem yields that $\lim_{\delta\downarrow 0}\sup_{h\in BL_1}\left|E\left[h(\mathbb{B}^{(p)} \circ \Pi_\delta)\right] - E\left[h(\mathbb{B}^{(p)})\right]\right| = 0$. Hence, we obtain the desired conclusion. ■

For $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, let $\mathbf{v}^*_{t,k} = (v^*_{1,t}, v^*_{2,t-k})^\top$ with $v^*_{i,t} = x^{*\top}_{it} v_i$ for $i = 1, 2$. Define

$$\mathbb{V}^*_{T,k}(\tau, \mathbf{v}) := T^{-1/2} \sum_{t=k+1}^{T} \left\{ 1[\mathbf{y}^*_{t,k} \le \mathbf{q}^*_{t,k}(\tau) + T^{-1/2}\mathbf{v}^*_{t,k}] - 1[\mathbf{y}_{t,k} \le \mathbf{q}_{t,k}(\tau) + T^{-1/2}\mathbf{v}_{t,k}] \right\},$$

and

$$\mathbb{W}^*_{i,T}(\tau_i, v_i) := T^{-1/2} \sum_{t=k+1}^{T} \left\{ x^*_{it}\psi_{\tau_i}\left(y^*_{it} - q^*_{i,t}(\tau_i) - T^{-1/2}v^*_{i,t}\right) - x_{it}\psi_{\tau_i}\left(y_{it} - q_{i,t}(\tau_i) - T^{-1/2}v_{i,t}\right) \right\}.$$

Proposition B.3. *Suppose Assumptions* A1–A6 *hold. Then,*

(a) $\sup_{\tau\in\mathcal{T}} \sup_{\mathbf{v}\in\mathcal{V}_M} |\mathbb{V}^*_{T,k}(\tau, \mathbf{v}) - \mathbb{V}^*_{T,k}(\tau)| = o_p(1)$ *for every* $M > 0$*;*
(b) $\sup_{\tau_i\in\mathcal{T}_i} \sup_{v_i\in\mathcal{V}_{i,M}} \|\mathbb{W}^*_{i,T}(\tau_i, v_i) - \mathbb{W}^*_{i,T}(\tau_i)\| = o_p(1)$ *for every* $M > 0$ *and* $i = 1, 2$,

where $\mathcal{V}_M = \mathcal{V}_{1,M} \times \mathcal{V}_{2,M}$ *with* $\mathcal{V}_{i,M} = \{v_i \in R^{d_i} : \|v_i\| \le M\}$ *for* $i = 1, 2$.

Proof. See Supplemental material, Appendix E. ■

Proposition B.4. *Suppose Assumptions* A1–A6 *hold. Then, for* $i = 1, 2$,

$$\sqrt{T}\{\hat\beta^*_i(\tau_i) - \beta_i(\tau_i)\} = -D_i^{-1}(\tau_i)\frac{1}{\sqrt{T}}\sum_{t=k+1}^{T} x^*_{it}\psi_{\tau_i}(y^*_{it} - q^*_{i,t}(\tau_i)) + o_p(1),$$

uniformly in $\tau_i \in \mathcal{T}_i$.

Proof. A similar argument used in Proposition A.4 completes the proof and thus the details are omitted. ■

Proposition B.5. *Suppose that Assumptions* A1–A6 *hold. Then, for each* $k \in \{1, \ldots, p\}$,

$$\sqrt{T}\left\{\hat\rho^*_\tau(k) - \hat\rho_\tau(k)\right\} = \frac{\mathbb{V}^*_{T,k}(\tau) + \nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta^*(\tau) - \hat\beta(\tau)\}}{\sqrt{\tau_1(1-\tau_1)\tau_2(1-\tau_2)}} + o_p(1),$$

uniformly in $\tau \in \mathcal{T}$.

Proof. Let $\hat\gamma^*_{\tau,k} = T^{-1}\sum_{t=k+1}^{T}\psi_{\tau_1}(y^*_{1t} - \hat q^*_{1,t}(\tau_1))\psi_{\tau_2}(y^*_{2,t-k} - \hat q^*_{2,t-k}(\tau_2))$. Using a similar argument used to show Lemma 2.1 of Arcones (1988), we can show $\sup_{\tau_i\in\mathcal{T}_i}|T^{-1/2}\sum_{t=1}^{T}\psi_{\tau_i}(y^*_{it} - \hat q^*_{i,t}(\tau_i))| = o_p(1)$ for $i = 1, 2$. It follows that

$$T^{-1}\sum_{t=k+1}^{T}\psi^2_{\tau_i}(y^*_{it} - \hat q^*_{it}(\tau_i)) = \tau_i(1-\tau_i) + o_p(1), \quad \text{for } i = 1, 2,$$

and

$$\sqrt{T}\left(\hat\gamma^*_{\tau,k} - \hat\gamma_{\tau,k}\right) = T^{-1/2}\sum_{t=k+1}^{T}\left\{1[\mathbf{y}^*_{t,k} \le \hat{\mathbf{q}}^*_{t,k}(\tau)] - 1[\mathbf{y}_{t,k} \le \hat{\mathbf{q}}_{t,k}(\tau)]\right\} + o_p(1),$$

uniformly in $\tau_i \in \mathcal{T}_i$ and $\tau \in \mathcal{T}$, respectively. As in Proposition A.5, we can show

$$\frac{1}{\sqrt{T}}\sum_{t=k+1}^{T}\left\{1[\mathbf{y}_{t,k} \le \hat{\mathbf{q}}_{t,k}(\tau)] - E\left[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)\mathbf{x}_{t,k})\right]\right\} = \mathbb{V}_{T,k}(\tau) + \nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta(\tau) - \beta(\tau)\} + o_P(1),$$

uniformly in $\tau \in \mathcal{T}$. A similar argument used in Proposition A.5 together with Propositions A.3 and B.3 yields that, uniformly in $\tau \in \mathcal{T}$,

$$\frac{1}{\sqrt{T}}\sum_{t=k+1}^{T}\left\{1[\mathbf{y}^*_{t,k} \le \hat{\mathbf{q}}^*_{t,k}(\tau)] - E\left[F^{(k)}_{\mathbf{y}|\mathbf{x}}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})\right]\right\} = \mathbb{V}_{T,k}(\tau)^* + \mathbb{V}_{T,k}(\tau) + \nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta^*(\tau) - \beta(\tau)\} + o_p(1).$$

It follows that $\sqrt{T}(\hat\gamma^*_{\tau,k} - \hat\gamma_{\tau,k}) = \mathbb{V}^*_{T,k}(\tau) + \nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat\beta^*(\tau) - \hat\beta(\tau)\} + o_p(1)$ uniformly in $\tau \in \mathcal{T}$. Thus, we obtained the desired result. ■

Proof of Theorem 2. (a) Define the processes $\hat{\mathbb{G}}^{(p)*}_T(\tau) := \sqrt{T}(\hat\rho^{*(p)}_\tau - \hat\rho^{(p)}_\tau)$ and $\mathbb{G}^{(p)}(\tau) := \Lambda^{(p)}_\tau\mathbb{B}^{(p)}(\tau)$ for $\tau \in \mathcal{T}$ and for an integer $p > 0$. Let $\widetilde{BL}_1$ denote the set of all Lipschitz continuous, real-valued functions on $(\ell^\infty(\mathcal{T}))^p$ with a Lipschitz constant bounded by 1. It suffices to show that

$$\sup_{h\in\widetilde{BL}_1}\left|E^*\left[h(\hat{\mathbb{G}}^{(p)*}_T)\right] - E\left[h(\mathbb{G}^{(p)})\right]\right| \to^p 0.$$

Let $\mathbb{G}^{(p)*}_T(\tau) := \Lambda^{(p)}_\tau\mathbb{B}^{(p)*}_T(\tau)$. We can write

$$\begin{aligned}
&\sup_{h\in\widetilde{BL}_1}\left|E^*\left[h(\hat{\mathbb{G}}^{(p)*}_T)\right] - E\left[h(\mathbb{G}^{(p)})\right]\right| \\
&\le \sup_{h\in\widetilde{BL}_1}\left|E^*\left[h(\hat{\mathbb{G}}^{(p)*}_T)\right] - E^*\left[h(\mathbb{G}^{(p)*}_T)\right]\right| \\
&\quad + \sup_{h\in\widetilde{BL}_1}\left|E^*\left[h(\mathbb{G}^{(p)*}_T)\right] - E\left[h(\mathbb{G}^{(p)})\right]\right|.
\end{aligned}$$

Propositions A.4 and B.4 imply that $\sqrt{T}\{\hat\beta^*_i(\tau_i) - \hat\beta_i(\tau_i)\} = -D_i^{-1}(\tau_i)\mathbb{W}^*_{1,T}(\tau_i) + o_p(1)$ uniformly in $\tau_i \in \mathcal{T}_i$ for each $i = 1, 2$. It follows from Proposition B.3 that

$$\sqrt{T}\left\{\hat\rho^*_\tau(k) - \hat\rho_\tau(k)\right\} = \lambda^\top_{\tau,k}\mathbb{B}^*_{T,k}(\tau) + o_p(1),$$

uniformly in $\tau \in \mathcal{T}$, where $\lambda_{\tau,k}$ is defined in (4). This leads to

$$\sup_{\tau\in\mathcal{T}}\left\|\hat{\mathbb{G}}^{(p)*}_T(\tau) - \mathbb{G}^{(p)*}_T(\tau)\right\| = o_p(1).$$

This implies that $\sup_{h\in\widetilde{BL}_1}|h(\hat{\mathbb{G}}^{(p)*}_T) - h(\mathbb{G}^{(p)*}_T)| = o_p(1)$, because $\sup_{h\in\widetilde{BL}_1}|h(\hat{\mathbb{G}}^{(p)*}_T) - h(\mathbb{G}^{(p)*}_T)| \le \sup_{\tau\in\mathcal{T}}\|\hat{\mathbb{G}}^{(p)*}_T(\tau) - \mathbb{G}^{(p)*}_T(\tau)\|$. It follows from the dominated convergence theorem that $\lim_{T\to\infty} E\sup_{h\in\widetilde{BL}_1}|E^*[h(\hat{\mathbb{G}}^{(p)*}_T)] - E^*[h(\mathbb{G}^{(p)*}_T)]| = 0$. An application of the Markov inequality shows that $\sup_{h\in\widetilde{BL}_1}|E^*[h(\hat{\mathbb{G}}^{(p)*}_T)] - E^*[h(\mathbb{G}^{(p)*}_T)]| = o_p(1)$.

Under Assumptions A4 and A5, $\Lambda^{(p)}_\tau$ is bounded uniformly in $\tau \in \mathcal{T}$, we have

$$\begin{aligned}
&\sup_{h\in\widetilde{BL}_1}\left|E^*\left[h(\mathbb{G}^{(p)*}_T)\right] - E\left[h(\mathbb{G}^{(p)})\right]\right| \\
&\le C_1\sup_{g\in BL_1}\left|E^*\left[g(\mathbb{B}^{(p)*}_T)\right] - E\left[g(\mathbb{B}^{(p)})\right]\right|,
\end{aligned}$$

where the right-hand side is negligible in probability from Proposition B.2. Hence, we obtain the desired result.

(b) From the continuous mapping theorem, the result in (a) of this theorem yields the desired result. See Theorem 10.8 of Kosorok (2007) for a general argument. ■

Proof of Theorem 3. As shown in (A.4), under both fixed and local alternatives,

$$\sqrt{T}\left(\hat\rho^{(p)}_\tau - \rho^{(p)}_\tau\right) = \Lambda^{(p)}_\tau\mathbb{B}^{(p)}_T(\tau) + o_p(1)$$

uniformly in $\tau \in \mathcal{T}$, and it follows from Theorem 1 that $\Lambda_\tau^{(p)}\mathbb{B}_T^{(p)}(\tau) = O_P(1)$ uniformly in $\tau \in \mathcal{T}$.

(a) Under the fixed alternative, there is some $\tau \in \mathcal{T}$ such that $\rho_\tau^{(p)}$ is some non-zero constant and then $\sqrt{T}\hat{\rho}_\tau^{(p)} = \Lambda_\tau^{(p)}\mathbb{B}_T^{(p)}(\tau) + \sqrt{T}\rho_\tau^{(p)} + o_p(1)$ uniformly in $\tau \in \mathcal{T}$. This implies that, under the fixed alternative, $\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)} = T \sup_{\tau\in\mathcal{T}} \|\rho_\tau^{(p)}\|^2(1+o_p(1))$. Thus, $\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)} \to^p \infty$ under the fixed alternative, whereas the critical value $c^*_{Q,\tau}$ is bounded in probability from Theorem 2. Therefore, $\lim_{T\to\infty} P(\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)} > c^*_{Q,\tau}) = 1$. Therefore, our test is shown to be consistent under the fixed alternative.

(b) Under the local alternative, we can write $\rho_\tau^{(p)} = \zeta_\tau^{(p)}/\sqrt{T}$, where $\zeta_\tau^{(p)}$ is a p-dimensional constant vector, at least one of elements is non-zero. Thus, we have

$$\hat{Q}_\tau^{(p)} = \|\Lambda_\tau^{(p)}\mathbb{B}_T^{(p)}(\tau) + \zeta_\tau^{(p)}\|^2 + o_P(1),$$

uniformly in $\tau \in \mathcal{T}$. From Theorem 1 and the continuous mapping theorem,

$$\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)} \Rightarrow \sup_{\tau\in\mathcal{T}} \|\Lambda_\tau^{(p)}\mathbb{B}^{(p)}(\tau) + \zeta_\tau^{(p)}\|^2.$$

Also, Theorem 2 implies $\sup_{\tau\in\mathcal{T}} \hat{Q}_\tau^{(p)*} \Rightarrow^* \sup_{\tau\in\mathcal{T}} \|\Lambda_\tau^{(p)}\mathbb{B}^{(p)}(\tau)\|^2$ in probability. Thus, the desired result follows. ■

Appendix C. Self-normalized cross-quantilogram

Lemma C.1. *Let $\{z_t\}_{t\in\mathbb{Z}}$ be a strict stationary, strong mixing sequence of $\mathbb{R}^d$-valued random variables for some integer $d \geq 1$ with strong mixing coefficients $\{\alpha_j\}_{j\in\mathbb{Z}_+}$ satisfying $\sum_{j=0}^\infty (j+1)^{2s-2}\alpha_j^{\nu/(2s+\nu)}$ for some integer $s \geq 2$ and $\nu \in (0, 1)$. Suppose that $E[z_1] = 0$ and $\|z_1\|_{2s+\nu} < \infty$. Then,*

$$E\left[\sup_{r\in[0,1]}\left\|\sum_{t=1}^{[Tr]} z_t\right\|^{2s}\right] \leq CT^s\left(\|z_1\|_{2+\nu}^{2s} + T^{1-s}\|z_1\|_{2s+\nu}\right).$$

Proof. The desired result follows from Theorem 6.3 and Annexes C of Rio (2013) as in Lemma A.1. ■

We define the process indexed by $r \in [0, 1]$

$$\bar{\mathbb{V}}_{T,k,\tau}(r) := \frac{1}{\sqrt{T}}\sum_{t=k+1}^{[Tr]}\left\{1[\mathbf{y}_{t,k} \leq \mathbf{q}_{t,k}(\tau)] - E[F_{\mathbf{y}|\mathbf{x}}^{(k)}(\mathbf{q}_{t,k}(\tau)|\mathbf{x}_{t,k})]\right\},$$

and

$$\bar{\mathbb{W}}_{i,T,\tau_i}(r) := \frac{1}{\sqrt{T}}\sum_{t=1}^{[Tr]} x_{it}\left\{1[y_{it} \leq q_{i,t}(\tau_i)] - \tau_i\right\},$$

for each $i = 1, 2$. The following proposition shows the stochastic equicontinuity of the processes defined above.

Proposition C.1. *Suppose Assumptions* A1–A5 *hold. Let $k \in \{1, \ldots, p\}$ and define metrics $\bar{\rho}(r, r') = |r' - r|$ for $r, r' \in [0, 1]$. Then,*

(a) *$\bar{\mathbb{V}}_{T,k,\tau}(r)$ is stochastically equicontinuous on $([0, 1], \bar{\rho})$.*
(b) *$\bar{\mathbb{W}}_{i,T,\tau_i}(r)$ is stochastically equicontinuous on $([0, 1], \bar{\rho})$ for each $i = 1, 2$.*

Proof. See Supplemental material, Appendix E. ■

Define a $d_0 \times 1$ vector $\bar{\mathbb{B}}_{T,k,\tau}(r) = [\bar{\mathbb{V}}_{T,k,\tau}(r), \bar{\mathbb{W}}_{1,T,\tau_1}(r)^\top, \bar{\mathbb{W}}_{2,T,\tau_2}(r)^\top]^\top$. for $r \in [0, 1]$ and $k = 1, \ldots, p$. The following proposition shows the weak convergence of the process $\{\bar{\mathbb{B}}_{T,k,\tau}(r) : r \in [0, 1]\}_{k=1}^p$.

Proposition C.2. *Suppose Assumptions* A1–A5 *hold. Then,*

$$\left[\bar{\mathbb{B}}_{T,1,\tau}(\cdot), \ldots, \bar{\mathbb{B}}_{T,p,\tau}(\cdot)\right]^\top \Rightarrow \left[\bar{\mathbb{B}}_{1,\tau}(\cdot), \ldots, \bar{\mathbb{B}}_{p,\tau}(\cdot)\right]^\top.$$

Proof. Proposition C.1 establishes the stochastic equicontinuity of $\left[\bar{\mathbb{B}}_{T,1,\tau}(\cdot), \ldots, \bar{\mathbb{B}}_{T,p,\tau}(\cdot)\right]^\top$ and it suffices to show convergence of the finite dimensional distributions. Since the finite dimensional convergences can be shown by a similar argument used in Proposition A.2, we omit the details. ■

For $\mathbf{v} = (v_1, v_2) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$, we define

$$\bar{\mathbb{V}}_{T,k,\tau}(r, \mathbf{v}) := \frac{1}{\sqrt{T}}\sum_{t=k+1}^{[Tr]}\Big\{1[\mathbf{y}_{t,k} \leq \mathbf{q}_{t,k}(\tau) + T^{-1/2}\mathbf{v}_{t,k}] - E\left[F_{\mathbf{y}|\mathbf{x}}^{(k)}\left(\mathbf{q}_{t,k}(\tau) + T^{-1/2}\mathbf{v}_{t,k}|\mathbf{x}_{t,k}\right)\right]\Big\},$$

and

$$\bar{\mathbb{W}}_{i,T,\tau_i}(r, v_i) := \frac{1}{\sqrt{T}}\sum_{t=k+1}^{[Tr]} x_{it}\Big\{1[y_{it} \leq q_{i,t}(\tau_i) + T^{-1/2}v_{i,t}] - F_{y_i|x_i}\left(q_{i,t}(\tau_i) + T^{-1/2}v_{i,t}|x_{it}\right)\Big\}.$$

Proposition C.3. *Suppose Assumptions* A1–A5 *hold. Then,*

(a) $\sup_{\omega\leq r\leq 1}\sup_{\mathbf{v}\in\mathcal{V}_M} |\bar{\mathbb{V}}_{T,k,\tau}(r, \mathbf{v}) - \bar{\mathbb{V}}_{T,k,\tau}(r)| = o_p(1)$ *for every $M > 0$;*
(b) $\sup_{\omega\leq r\leq 1}\sup_{v_i\in\mathcal{V}_{i,M}} \|\bar{\mathbb{W}}_{i,T,\tau_i}(r, v_i) - \bar{\mathbb{W}}_{i,T,\tau_i}(r)\| = o_p(1)$ *for every $M > 0$ and for $i = 1, 2$,*

where $\mathcal{V}_M = \mathcal{V}_{1,M} \times \mathcal{V}_{2,M}$ with $\mathcal{V}_{i,M} = \{v_i \in R^{d_i} : \|v_i\| \leq M\}$ for $i = 1, 2$.

Proof. See Supplemental material, Appendix E. ■

Proposition C.4. *Suppose Assumptions* A1–A5 *hold. Then, for $i = 1, 2$ and for each $\tau_i \in \mathcal{T}_i$,*

$$\sqrt{T}\{\hat{\beta}_{i,[Tr]}(\tau_i) - \beta_i(\tau_i)\} = -D_i^{-1}(\tau_i)r^{-1}\bar{\mathbb{W}}_{i,T,\tau_i}(r) + o_p(1),$$

uniformly in $r \in [\omega, 1]$.

Proof. The proof follows the line of Proposition A.4 with Proposition C.3(b). Hence, we omit the details. ■

Proposition C.5. *Suppose Assumptions* A1–A5 *hold. Then, for each $(k, \tau) \in \{1, \ldots, p\} \times \mathcal{T}$,*

$$\sqrt{T}\left\{\hat{\rho}_{\tau,[Tr]}(k) - \rho_\tau(k)\right\} = \frac{r^{-1}\bar{\mathbb{V}}_{T,k,\tau}(r) + \nabla G^{(k)}(\tau)^\top\sqrt{T}\{\hat{\beta}_{[Tr]}(\tau) - \beta(\tau)\}}{\sqrt{\tau_1(1-\tau_1)\tau_2(1-\tau_2)}} + o_p(1),$$

uniformly in $r \in [\omega, 1]$, where $\hat{\beta}_{[Tr]} = (\hat{\beta}_{1,[Tr]}^\top, \hat{\beta}_{2,[Tr]}^\top)^\top$.

Proof. A similar argument used in Proposition A.5 with Proposition C.3(a) yields the desired result and thus we omit the detail. ■

Proof of Theorem 4. Propositions C.4 and C.5 imply that, for each $(k, \tau) \in \{1, \ldots, p\} \times \mathcal{T}$,

$$\sqrt{T}\left\{\hat{\rho}_{\tau,[Tr]}(k) - \rho_\tau(k)\right\} = r^{-1}\lambda_{\tau,k}^\top\bar{\mathbb{B}}_{T,k,\tau}(r) + o_p(1),$$

uniformly in $r \in [\omega, 1]$. It follows that $\sqrt{T}(\hat{\rho}_{\tau,[Tr]}^{(p)} - \rho_\tau^{(p)}) = r^{-1}\Lambda_\tau^{(p)}\bar{\mathbb{B}}_{T,\tau}^{(p)}(r) + o_p(1)$ uniformly in $r \in [\omega, 1]$. This implies

$$\frac{[Tr]}{\sqrt{T}}\left(\hat{\rho}_{\tau,[Tr]}^{(p)} - \hat{\rho}_{\tau,T}^{(p)}\right) = \Lambda_\tau^{(p)}\left\{\bar{\mathbb{B}}_{T,\tau}^{(p)}(r) - r\bar{\mathbb{B}}_{T,\tau}^{(p)}(1)\right\} + o_p(1),$$

uniformly in $r \in [\omega, 1]$. From Proposition C.2, $\{\Lambda_\tau^{(p)}(\bar{\mathbb{B}}_{T,\tau}^{(p)}(r) - r\bar{\mathbb{B}}_{T,\tau}^{(p)}(1)) : r \in [\omega, 1]\}$ weakly converges to $\{\Lambda_\tau^{(p)}(\bar{\mathbb{B}}_{\tau}^{(p)}(r) - r\bar{\mathbb{B}}_{\tau}^{(p)}(1)) : r \in [\omega, 1]\}$, which is equivalent in distribution to a $p \times 1$ vector of the Brownian bridge process $\{\Delta_\tau^{(p)}(\bar{\mathbf{B}}^{(p)}(r) - r\bar{\mathbf{B}}^{(p)}(1)) : r \in [\omega, 1]\}$ with $\Delta_\tau^{(p)}(\Delta_\tau^{(p)})^\top \equiv \Xi^{(p)}(\tau, \tau)$, and thus it follows from the continuous mapping theorem that

$$\left(\sqrt{T}\hat{\rho}_{\tau,T}^{(p)}, \hat{V}_{\tau,p}\right) \to^d \left(\Delta_\tau^{(p)}\bar{\mathbf{B}}^{(p)}(1), \Delta_\tau^{(p)}\bar{\mathbf{V}}^{(p)}(\Delta_\tau^{(p)})^\top\right).$$

Thus, we obtain $\hat{S}_\tau^{(p)} \to^d \bar{\mathbf{B}}^{(p)}(1)^\top (\bar{\mathbf{V}}^{(p)})^{-1}\bar{\mathbf{B}}^{(p)}(1)$. This completes the proof. ∎

Proof of Theorem 5. Under both fixed and local alternative, the argument used in Theorem 4 gives

$$\sqrt{T}\left(\hat{\rho}_{\tau,[Tr]}^{(p)} - \rho_\tau^{(p)}\right) = r^{-1}\Lambda_\tau^{(p)}\bar{\mathbb{B}}_{T,\tau}^{(p)}(r) + o_p(1),$$

thereby yielding $\hat{V}_{\tau,p} \Rightarrow (\Lambda_\tau^{(p)}\Delta_\tau^{(p)})\bar{\mathbf{V}}^{(p)}(\Lambda_\tau^{(p)}\Delta_\tau^{(p)})^\top$.

(a) Under the fixed alternative, we have $\sqrt{T}\hat{\rho}_{\tau,T}^{(p)} = \Lambda_\tau^{(p)}\bar{\mathbb{B}}_{T,\tau}^{(p)}(1) + \sqrt{T}\rho_\tau^{(p)} + o_p(1)$, where the right-hand side diverges in probability as $T \to \infty$. Since the critical value we use is finite in probability from Theorem 4, we obtain the desired result.

(b) Under the local alternative, $\sqrt{T}\hat{\rho}_{\tau,T}^{(p)} = \Lambda_\tau^{(p)}\bar{\mathbb{B}}_{T,\tau}^{(p)}(1) + \xi_\tau^{(p)} + o_p(1)$. It follows that

$$\hat{S}_\tau^{(p)} \to^d \left\{\bar{\mathbf{B}}^{(p)}(1) + (\Lambda_\tau^{(p)}\Delta_\tau^{(p)})^{-1}\xi_\tau^{(p)}\right\}^\top \left(\bar{\mathbf{V}}^{(p)}\right)^{-1} \left\{\bar{\mathbf{B}}^{(p)}(1) + (\Lambda_\tau^{(p)}\Delta_\tau^{(p)})^{-1}\xi_\tau^{(p)}\right\}.$$

This completes the proof. ∎

Appendix D. Partial cross-quantilogram

For $1 \le i, j \le l$, let $\mathbf{1}_{ij} = 1[y_{it} \le q_{i,t}(\tau_i), y_{jt} \le q_{j,t}(\tau_j)]$ and define

$$\mathbb{V}_{T,ij} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\mathbf{1}_{ij} - E\left[\mathbf{1}_{ij}\right]\right) \quad \text{and}$$

$$\mathbb{W}_{i,T} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} x_{it}\psi_{\tau_i}\left(y_{it} - q_{i,t}(\tau_i)\right).$$

Proof of Theorem 6. We first consider (a). The correlation matrix $R_{\bar{\tau}}$ is symmetric and $\hat{R}_{\bar{\tau}}^{-1} - R_{\bar{\tau}}^{-1} = -\hat{R}_{\bar{\tau}}^{-1}(\hat{R}_{\bar{\tau}} - R_{\bar{\tau}})R_{\bar{\tau}}^{-1}$. It follows that $\mathrm{vec}(\hat{P}_{\bar{\tau}} - P_{\bar{\tau}}) = -P_{\bar{\tau}} \otimes \hat{P}_{\bar{\tau}}\mathrm{vec}(\hat{R}_{\bar{\tau}} - R_{\bar{\tau}})$, which implies

$$\sqrt{T}(\hat{p}_{\bar{\tau},12} - p_{\bar{\tau},12}) = -\sum_{i=1}^{l}\sum_{j=1}^{l} p_{\bar{\tau},1i}\hat{p}_{\bar{\tau},2j}\sqrt{T}(\hat{r}_{\bar{\tau},ij} - r_{\bar{\tau},ij}).$$

Following the line of proof of Theorem 1, we can show $\hat{P}_{\bar{\tau}} = P_{\bar{\tau}} + o_p(1)$ and also have $\sqrt{T}(\hat{r}_{\bar{\tau},ii} - r_{\bar{\tau},ii}) = o_p(1)$ for $i = 1, \ldots, l$, from argument in Lemma 2.1 of Arcones (1988). Thus, we have

$$\sqrt{T}(\hat{p}_{\bar{\tau},12} - p_{\bar{\tau},12}) = -\sum_{\substack{1\le i,j\le l\\ i\ne j}} p_{\bar{\tau},1i}p_{\bar{\tau},2j}\sqrt{T}(\hat{r}_{\bar{\tau},ij} - r_{\bar{\tau},ij}) + o_p(1).$$

Proposition A.5 implies

$$\sqrt{T}(\hat{r}_{\bar{\tau},ij} - r_{\bar{\tau},ij}) = \mathbb{V}_{T,ij} + \nabla_1 G_{ij}^\top\sqrt{T}\{\hat{\beta}_i(\tau_i) - \beta_i(\tau_i)\} + \nabla_2 G_{ij}^\top\sqrt{T}\{\hat{\beta}_j(\tau_j) - \beta_j(\tau_j)\} + o_p(1),$$

for $1 \le i, j \le l$ with $i \ne j$. Since $\mathbb{V}_{T,ij} = \mathbb{V}_{T,ji}$ and $\nabla_2 G_{ij} = \nabla_1 G_{ji}$ for $1 \le i, j \le l$,

$$\sqrt{T}(\hat{p}_{\bar{\tau},12} - p_{\bar{\tau},12}) = -\sum_{\substack{1\le i,j\le l\\ i\ne j}} p_{\bar{\tau},1i}p_{\bar{\tau},2j}\mathbb{V}_{T,ij} - \sum_{i=1}^{l}\lambda_{\bar{\tau}i}^\top\sqrt{T}\{\hat{\beta}_i(\tau_i) - \beta_i(\tau_i)\} + o_p(1),$$

where $\lambda_{\bar{\tau}i}$ is defined in Theorem 6. Proposition A.4 implies

$$\sqrt{T}(\hat{p}_{\bar{\tau},12} - p_{\bar{\tau},12}) = -\sum_{\substack{1\le i,j\le l\\ i\ne j}} p_{\bar{\tau},1i}p_{\bar{\tau},2j}\mathbb{V}_{T,ij} + \sum_{i=1}^{l}\lambda_{\bar{\tau}i}^\top D_i(\tau_i)^{-1}\mathbb{W}_{i,T} + o_p(1).$$

The asymptotic normality can be established by using the central limit theorem for mixing random vectors. The proofs of (b) and (c) are similar to those of Theorems 2 and 4, respectively, and thus we omit the details. ∎

Appendix E. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.jeconom.2016.03.001.

References

Adrian, T., Brunnermeier, M.K., 2011. CoVaR, Tech. Rep. Federal Reserve Bank of New York, Staff Reports.

Andrews, D.W.K., Pollard, D., 1994. An introduction to functional central limit theorems for dependent stochastic processes. Internat. Statist. Rev. 62, 119–132.

Arcones, M., 1988. Second order representations of the least absolute deviation regression estimator. Ann. Inst. Statist. Math. 50, 87–117.

Bai, J., 1996. Testing for parameter constancy in linear regressions: An empirical distribution function approach. Econometrica 64, 597–622.

Bedljkovic, M., 2010. Nonparametric test of conditional quantile independence with an applicationi to banks' systemic risk. Working Paper.

Bisias, D., Flood, M., Lo, A.W., Valavanis, S., 2012. A survey of systemic risk analytics. Working Paper, Office of Financial Research.

Brownlees, C., Engle, R.F., 2012. Volatility, Correlation and Tails for Systemic Risk Management, Tech. Rep. SSRN.

Bunzel, H., Kiefer, N.M., Vogelsang, T.J., 2001. Simple robust testing of hypotheses in nonlinear models. J. Amer. Statist. Assoc. 96, 1088–1096.

Campbell, J.Y., Lo, A.W., MacKinlay, A.C., 1997. The Econometrics of Financial Markets. Princeton University Press, Princeton.

Cenesizoglu, T., Timmermann, A., 2008. Is the distribution of stock return predictable? University of California at San Diego. Working Paper.

Chang, C.Y., Shie, F.S., 2011. The relation between relative order imbalance and intraday futures returns: An application of the quantile regression model to Taiwan. Emerg. Mark. Financ. Trade 47 (3), 69–87.

Chen, Y.-T., Qu, Z., 2015. M tests with a new normalization matrix. Econometric Rev. 34, 617–652.

Christoffersen, P.F., Diebold, F.X., 2002. Financial Asset Returns, market timing, and volatility dynamics. Working Paper.

Corradi, V., Swanson, N.R., 2006. Bootstrap conditional distribution tests in the presence of dynamic misspecification. J. Econometrics 133, 779–806.

Cowles, A., Jones, H., 1937. Some a posteriori probabilities in stock market action. Econometrica 5, 280–294.

Davis, R.A., Mikosch, T., 2009. The extremogram: a correlogram for extreme events. Bernoulli 15, 977–1009.

Davis, R.A., Mikosch, T., Cribben, I., 2012. Towards estimating extremal serial dependence via the bootstrapped extremogram. J. Econometrics 170 (1), 142–152.

Davis, R.A., Mikosch, T., Zhao, Y., 2013. Measures of serial extremal dependence and their estimation. Stochastic Process. Appl. 123 (7), 2575–2602.

Deheuvels, P., 1979. La fonction de dépendance empirique et ses propriétés. Un test non paramétrique d'indépendance. Acad. Roy. Belg. Bull. Cl. Sci. (5) 65, 274–292.

Dette, H., Hallin, M., Kley, T., Volgushev, S., 2015. Of copulas, quantiles, ranks, and spectra an L1 -approach to spectral analysis. Bernoulli 21, 781–831.

Dufour, J.M., Hallin, M., Mizera, I., 1998. Generalized runs tests for heteroscedastic time series. Nonparametr. Stat. 9, 39–86.

Embrechts, P., Kluppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer, New York.

Engle, R.F., Manganelli, S., 2004. CAViaR: conditional autoregressive value at risk by regression quantiles. J. Bus. Econom. Statist. 22, 367–381.

Fama, E., 1965. The behavior of stock market prices. J. Bus. 38, 34. 105.

Fermanian, J.-D., Radulovic, D., Wegkamp, M., 2004. Weak convergence of empirical copula processes. Bernoulli 10, 847–860.
Galvao, A.F., Kato, K., Montes-Rojas, G., Olmo, J., 2014. Testing linearity against threshold effects: uniform inference in quantile regression. Ann. Inst. Statist. Math. 66 (2), 413–439.
Goyal, A., Welch, I., 2008. A comprehensive look at the empirical performance of equity premium prediction. Rev. Financ. Stud. 21, 1455–1508.
Granger, C.W.J., 1969. Investigating causal relations by econometric models and cross-spectral methods. Econometrica 37, 424–438.
Gut, A., 2009. Stopped Random Walks: Limit Theorems and Applications. Springer, New York.
Hagemann, A., 2013. Robust Spectral Analysis, arXiv preprint arXiv:1111.1965.
Hall, P., Heyde, C.C., 1980. Martingale Limit Theory and its Application. Academic press, New York.
Han, H., 2015. Asymptotic properties of GARCH-X processes. J. Financ. Econom. 13, 188–221.
Hong, Y., 1996. Consistent testing for serial correlation of unknown form. Econometrica 64 (4), 837–864.
Hong, Y., 2000. Generalized spectral tests for serial dependence. J. R. Stat. Soc. Ser. B Stat. Methodol. 62 (3), 557–574.
Hong, Y., Liu, Y., Wang, S., 2009. Granger causality in risk and detection of extreme risk spillover between financial markets. J. Econometrics 150, 271–287.
Ibragimov, R., 2009. Heavy-tailed densities. In: Durlauf, Steven N., Blume., Lawrence E. (Eds.), The New Palgrave Dictionary of Economics. Palgrave Macmillan.
Ibragimov, R., Jaffee, D., Walden, J., 2009. Nondiversification traps in catastrophe insurance markets. Rev. Financ. Stud. 22, 959–993.
Kiefer, N.M., Vogelsang, T.J., 2002. Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation. Econometrica 70, 2093–2095.
Kiefer, N.M., Vogelsang, T.J., 2005. A new asymptotic theory for heteroskedasticity-autocorrelation robust tests. Econometric Theory 21, 1130–1164.
Kiefer, N.M., Vogelsang, T.J., Bunzel, H., 2000. Simple robust testing of regression hypotheses. Econometrica 68, 695–714.
Kim, M.S., Sun, Y., 2011. Spatial heteroskedasticity and autocorrelation consistent estimation of covariance matrix. J. Econometrics 160, 349–371.
Kley, T., Volgushev, S., Dette, H., Hallin, M., 2016. Quantile spectral processes: Asymptotic analysis and inference. Bernoulli 22, 1770–1807.
Koenker, R., Bassett, W.G., 1978. Regression quantiles. Econometrica 46, 33–50.
Kosorok, M.R., 2007. Introduction to Empirical Processes and Semiparametric Inference. Springer Science and Business Media.
Kuan, C.-M., Lee, W.-M., 2006. Robust M tests without consistent estimation of the asymptotic covariance matrix. J. Amer. Statist. Assoc. 101, 1264–1275.
Kunsch, H.R., 1989. The jackknife and the bootstrap for general stationary observations. Ann. Statist. 17, 1217–1241.
Laurini, M.P., Furlani, L.G.C., Portugal, M.S., 2008. Empirical market microstructure: An analysis of the BRL/US$ exchange rate market. Emerg. Mark. Rev. 9 (4), 247–265.
Lettau, M., Ludvigson, S.C., 2010. Measuring and modeling variation in the risk-return trade-off. In: Ait-Sahalia, Y., Hansen, L.P. (Eds.), Handbook of Financial Econometrics, vol. 1. North-Holland, pp. 617–690.
Li, T.H., 2008. Laplace periodogram for time series analysis. J. Amer. Statist. Assoc. 103 (482), 757–768.
Li, T.H., 2012. Quantile periodograms. J. Amer. Statist. Assoc. 107 (498), 765–776.
Li, T.H., 2014. Quantile periodogram and time-dependent variance. J. Time Series Anal. 35 (4), 322–340.
Linton, O., Whang, Y.-J., 2007. The quantilogram: With an application to evaluating directional predictability. J. Econometrics 141, 250–282.
Lobato, I.N., 2001. Testing that a dependent process is uncorrelated. J. Amer. Statist. Assoc. 96, 1066–1076.
Mandelbrot, B.B., 1963. The variation of certain speculative prices. J. Bus. 36, 394–419. (Chicago). Reprinted in Cootner (1964), as Chapter E 14 of Mandelbrot (1997), in Telser (2000), and several other collections of papers on finance.
Mikosch, T., Starica, C., 2000. Limit Theory for the sample autocorrelations and extremes of a GARCH(1,1) process. Ann. Statist. 28, 1427–1451.
Patton, A., Politis, D.N., White, H., 2009. Correction to "automatic block-length selection for dependent bootstrap" by D. Politis and H. White. Econometric Rev. 28, 372–375.
Pierce, D.A., Haugh, L.D., 1977. Causality in temporal systems: Characterization and a survey. J. Econometrics 5, 265–293.
Politis, D.N., Romano, J.P., 1994. The stationary bootstrap. J. Amer. Statist. Assoc. 89, 1303–1313.
Politis, D.N., White, H., 2004. Automatic block-length selection for dependent bootstrap. Econometric Rev. 23, 53–70.
Pollard, D., 1991. Asymptotics for least absolute deviation regression estimators. Econometric Theory 7, 186–199.
Rachev, S., Mittnik, S., 2000. Stable Paretian Models in Finance. Wiley, New York.
Rio, E., 1997. About the Lindeberg method for strongly mixing sequences. ESAIM Probab. Stat. 1, 35–61.
Rio, E., 2013. Inequalities and limit theorems for weakly dependent sequences. Working Paper.
Ruschendorf, L., 1976. Asymptotic distributions of multivariate rank order statistics. Ann. Statist. 4, 912–923.
Segers, J., 2012. Asymptotics of empirical copula processes under non-restrictive smoothness assumptions. Bernoulli 18, 764–782.
Shao, X., 2010. A self-normalized approach to confidence interval construction in time series. J. R. Stat. Soc. Ser. B 72, 343–366.
Stute, W., 1984. The oscillation behavior of empirical processes: The multivariate case. Ann. Probab. 12, 361–379.
Sun, Y., Kim, M.S., 2012. Simple and powerful GMM over-identification tests with accurate size. J. Econometrics 166, 267–281.
Sun, Y., Phillips, P.C.B., Jin, S., 2008. Optimal bandwidth selection in heteroskedasticity-autocorrelation robust testing. Econometrica 76, 175–194.
van der Vaart, A.W., Wellner, J.A., 1996. Weak Convergence and Empirical Processes. Springer, New York.
White, H., Kim, T.-H., Manganelli, S., 2012. VAR for VaR: Measuring tail dependence using multivariate regression quantiles. Working Paper.