cemmap

centre for microdata methods and practice

The cross-quantilogram: measuring quantile dependence and testing directional predictability between time series

Heejoon Han
Oliver Linton
Tatsushi Oka
Yoon-Jae Whang

The Institute for Fiscal Studies
Department of Economics, UCL

cemmap working paper CWP06/14

E·S·R·C ECONOMIC & SOCIAL RESEARCH COUNCIL

An ESRC Research Centre

Institute for Fiscal Studies, 7 Ridgmount Street, London WC1E 7AE **Tel:** 44 (0)20 7291 4800 **Email:** info@cemmap.ac.uk **www.cemmap.ac.uk**

The Cross-Quantilogram: Measuring Quantile Dependence and Testing Directional Predictability between Time Series

Heejoon Han Oliver Linton Tatsushi Oka Yoon-Jae Whang*

January 2014

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 the stationary bootstrap procedure; we show the consistency of this bootstrap. Also, we consider the self-normalized approach, which is shown to be asymptotically pivotal under

*Heejoon Han is Associate Professor, Department of Economics, Kyung Hee University, Seoul, Republic of Korea (E-mail: heejoon@khu.ac.kr). Oliver Linton is Fellow of Trinity College and Professor of Political Economy, Faculty of Economics, University of Cambridge, Cambridge, UK (E-mail: obl20@cam.ac.uk). Tatsushi Oka is Assistant Professor, Department of Economics, National University of Singapore, Singapore, Republic of Singapore (E-mail: oka@nus.edu.sg). Yoon-Jae Whang is Professor, Department of Economics, Seoul National University, Seoul, Republic of Korea (E-mail: whang@snu.ac.kr). Han's work was supported by the National Research Foundation of Korea (NRF-2013021502). Linton's work was supported by Cambridge INET and the ERC. Whang's work was supported by the National Research Foundation of Korea (NRF-2011-342-B00004).

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, Goldman Sachs and AIG. This article has supplementary materials online.

Keywords: Quantile, Correlogram, Dependence, Predictability, Systemic risk.

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 $\{x_t\}$ does not improve the prediction about whether x_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. Hagemann (2012) has introduced a Fourier domain version of the quantilogram, see also Dette et al. (2013) for an alternative approach. The quantilogram has recently been applied to stock returns and exchange rates, Laurini et al. (2008) and Chang and Shie (2011).

Our paper addresses two 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 tail 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. 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 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 or 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.

2 The Cross-Quantilogram

Let $\{x_t : t \in \mathbb{Z}\}$ be a two dimensional strictly stationary time series with $x_t \equiv (x_{1t}, x_{2t})^\intercal$ and let $F_i(\cdot)$ denote the distribution function of the series x_{it} with density function $f_i(\cdot)$ for $i = 1, 2$. The quantile function of the time series x_{it} is defined as $q_i(\alpha_i) = \inf\{v : F_i(v) \geq \alpha_i\}$ for $\alpha_i \in (0, 1)$, for $i = 1, 2$. Let $q_\alpha = (q_1(\alpha_1), q_2(\alpha_2))^\intercal$ for $\alpha \equiv (\alpha_1, \alpha_2)^\intercal$.

We consider a measure of serial dependence between two events $\{x_{1t} \leq q_1(\alpha_1)\}$ and $\{x_{2t-k} \leq q_2(\alpha_2)\}$ for arbitrary quantiles. In the literature, $\{1[x_{it} \leq q_i(\cdot)]\}$ is called the quantile-hit or quantile-exceedance process for $i = 1, 2$, where $1[\cdot]$ denotes the indicator function taking the value one when its argument is true, and zero otherwise. The cross-quantilogram is defined as the cross-correlation of the quantile-hit processes

$$\rho_\alpha(k) = \frac{E\left[\psi_{\alpha_1}(x_{1t} - q_1(\alpha_1))\psi_{\alpha_2}(x_{2t-k} - q_2(\alpha_2))\right]}{\sqrt{E\left[\psi^2_{\alpha_1}(x_{1t} - q_1(\alpha_1))\right]}\sqrt{E\left[\psi^2_{\alpha_2}(x_{2t} - q_2(\alpha_2))\right]}} \quad (1)$$

for $k = 0, \pm 1, \pm 2, \ldots$, where $\psi_a(u) \equiv 1[u < 0] - a$. The cross-quantilogram captures serial dependency between the two series at different 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 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.

To construct the sample analogue of the cross-quantilogram based on observations $\{x_1, \ldots, x_T\}$, we first estimate the unconditional quantile functions by solving the following minimization problems, separately:

$$\hat{q}_1(\alpha_1) = \arg\min_{v_1 \in \mathbb{R}} \sum_{t=1}^{T} \pi_{\alpha_1}(x_{1t} - v_1) \quad \text{and} \quad \hat{q}_2(\alpha_2) = \arg\min_{v_2 \in \mathbb{R}} \sum_{t=1}^{T} \pi_{\alpha_2}(x_{2t} - v_2),$$

where $\pi_a(u) \equiv u(a - 1[u < 0])$. Then, the sample cross-quantilogram is defined as

$$\hat{\rho}_\alpha(k) = \frac{\sum_{t=k+1}^{T} \psi_{\alpha_1}(x_{1t} - \hat{q}_1(\alpha_1))\psi_{\alpha_2}(x_{2t-k} - \hat{q}_2(\alpha_2))}{\sqrt{\sum_{t=k+1}^{T} \psi_{\alpha_1}^2(x_{1t} - \hat{q}_1(\alpha_1))}\sqrt{\sum_{t=k+1}^{T} \psi_{\alpha_2}^2(x_{2t-k} - \hat{q}_2(\alpha_2))}}, \tag{2}$$

for $k = 0, \pm 1, \pm 2, \ldots$. Given a set of quantiles, the cross-quantilogram considers dependency in terms of the direction of deviation from quantiles and thus measures the directional predictability from one series to another. This can be a useful descriptive device. By construction, $\hat{\rho}_\alpha(k) \in [-1, 1]$ with $\hat{\rho}_\alpha(k) = 0$ corresponding to the case of no directional predictability. The form of the statistic generalizes to the d dimensional multivariate case and the (i, j)th entry of the corresponding cross-correlation matrices $\Gamma_\alpha(k)$ is given by applying (2) for a pair of variables (x_{it}, x_{jt-k}) and a pair of quantiles $(\hat{q}_i(\alpha_i), \hat{q}_j(\alpha_j)))$ for $\alpha = (\alpha_1, \ldots, \alpha_d)^\intercal$. The cross-correlation matrices possess the usual symmetry property $\Gamma_\alpha(k) = \Gamma_\alpha(-k)^\intercal$ when $\alpha_1 = \cdots = \alpha_d$.

We may be interested in testing for the absence of directional predictability over a set of quantiles. Let $\mathcal{A} \equiv \mathcal{A}_1 \times \mathcal{A}_2$, where $\mathcal{A}_i \subset (0, 1)$ is a quantile range for each time series $\{x_{i1}, \ldots, x_{iT}\}$ $(i = 1, 2)$. We are interested in testing the hypothesis $H_0 : \rho_\alpha(1) = \cdots = \rho_\alpha(p) = 0, \quad \forall \alpha \in \mathcal{A}$, against the alternative hypothesis that $\rho_\alpha(k) \neq 0$ for some $k \in \{1, \ldots, p\}$ and some $\alpha \in \mathcal{A}$ with p fixed.[1] This is a test for the directional predictability of events up to p lags $\{x_{2t-k} \leq q_2(\alpha_2) : k = 1, \ldots, p\}$ for x_{1t}. To discriminate between these hypotheses we will use the test statistic

$$\sup_{\alpha \in \mathcal{A}} \hat{Q}_\alpha^{(p)} = \sup_{\alpha \in \mathcal{A}} T||\hat{\rho}_\alpha^{(p)}||^2 = \sup_{\alpha \in \mathcal{A}} T \sum_{k=1}^{p} \hat{\rho}_\alpha^2(k),$$

where $\hat{Q}_\alpha^{(p)}$ is the quantile specific Box-Pierce type statistic and $\hat{\rho}_\alpha^{(p)} \equiv [\hat{\rho}_\alpha(1), \ldots, \hat{\rho}_\alpha(p)]^\intercal$. To test the directional predictability in a specific quantile, or to provide confidence intervals for the population quantities, we use $\hat{\rho}_\alpha(k)$ or $\hat{Q}_\alpha^{(p)}$, which are special cases of the sup-type test statistic. In practice, we have found that the Box-Ljung version $\hat{Q}_\alpha^{(p)} \equiv T(T+2)\sum_{k=1}^{p} \hat{\rho}_\alpha^2(k)/(T-k)$ yields some small sample improvements.

3 Asymptotic Properties

Here we 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. We impose the following assumptions.

Assumption A1. $\{x_t : t \in \mathbb{Z}\}$ is a strictly stationary and strong mixing with coefficients $\{\alpha(n)\}$ that satisfy $\alpha(n) = O(n^{-a})$ for $a > 1$. **A2**. The distribution functions $F_i(\cdot)$ for $i = 1, 2$ have continuous densities $f_i(\cdot)$ uniformly bounded away from 0 and ∞ at $q_i(\alpha_i)$ uniformly over $\alpha_i \in \mathcal{A}_i$. **A3**. For any $\epsilon > 0$ there exists

[1] 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 is asymptotically normal. No doubt our results can be extended to accommodate this case, although in practice the desirability of such a test is questionable, and our limit theory may provide better critical values for even quite long lags.

a $\nu(\epsilon)$ such that $\sup_{\tau \in \mathcal{A}_i} \sup_{s:|s| \leq \nu(\epsilon)} |f_i(q_i(\tau)) - f_i(q_i(\tau) + s)| < \epsilon$ for $i = 1, 2$. **A4.** The joint distribution G_k of (x_{1t}, x_{2t-k}) has a bounded, continuous first derivative for each argument uniformly in the neighborhood of quantiles of interest for every $k \in \{1, \ldots, p\}$.

Assumption A1 imposes a mixing rate used in Rio (2000, Chapter 7). For a strong mixing process, $\rho_\alpha(k) \to 0$ as $k \to \infty$ for all $\alpha \in (0, 1)$. Assumption A2 ensures that the quantile functions are uniquely defined. 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.

To describe the asymptotic behavior of the cross-quantilogram, we define a set of 3-dimensional mean-zero Gaussian process $\{\mathbb{B}_k(\alpha) : \alpha \in [0, 1]^2\}_{k=1}^p$ with covariance-matrix function given by

$$\Gamma_{kk'}(\alpha, \alpha') \equiv E[\mathbb{B}_k(\alpha)\mathbb{B}_{k'}^\intercal(\alpha')] = \sum_{l=-\infty}^{\infty} \text{cov}\left(\xi_l(\alpha, k), \xi_0^\intercal(\alpha', k')\right),$$

for $k, k' \in \{1, \ldots, p\}$ and for $\alpha, \alpha' \in \mathcal{A}$, where

$$\xi_t(\alpha, k) = \left[1[\epsilon_{1t}(\alpha_1) \leq 0, \epsilon_{2t-k}(\alpha_2) \leq 0] - G_k(q_\alpha), \psi_{\alpha_1}(\epsilon_{1t}(\alpha_1)), \psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2))\right]^\intercal,$$

with $\epsilon_{it}(\cdot) = x_{it} - q_i(\cdot)$ for $i = 1, 2$ and for $t = 1, \ldots, T$. Define the $3p$-dimensional zero-mean Gaussian process $\mathbb{B}^{(p)}(\alpha) = [\mathbb{B}_1(\alpha)^\intercal, \ldots, \mathbb{B}_p(\alpha)^\intercal]^\intercal$ with the covariance-matrix function denoted by $\Gamma^{(p)}(\alpha, \beta)$. The next theorem establishes the asymptotic properties of the cross-quantilogram.

Theorem 1 *Suppose that Assumptions A1-A4 hold for some finite integer* $p > 0$*. Then, in the sense of weak convergence of the stochastic process we have:*

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

where $\Lambda_\alpha^{(p)} = \text{diag}(\lambda_{\alpha 1}^\intercal, \ldots, \lambda_{\alpha p}^\intercal)$ *and*

$$\lambda_{\alpha k} = \frac{1}{\sqrt{\alpha_1(1-\alpha_1)\alpha_2(1-\alpha_2)}} \text{diag}\left(1, \frac{1}{f_1(q_1(\alpha_1))}, \frac{1}{f_2(q_2(\alpha_2))}\right) \begin{bmatrix} 1 \\ \nabla G_k(q_\alpha) \end{bmatrix}, \quad (4)$$

with the gradient vector $\nabla G_k(\cdot)$ *of* $G_k(\cdot)$*.*

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

$$\sup_{\alpha \in \mathcal{A}} \hat{Q}_\alpha^{(p)} \Longrightarrow \sup_{\alpha \in \mathcal{A}} \|\Lambda_\alpha^{(p)} \mathbb{B}^{(p)}(\alpha)\|^2 \quad (5)$$

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. To estimate the critical values from the limiting distribution we could use nonparametric estimation, but that may suffer from a slow convergence rate. We address this issue by using the stationary bootstrap (SB) of Politis and Romano (1994). The SB is a block bootstrap method with blocks of random lengths. The SB resample is strictly stationary conditional on the original sample.

Let $\{L_i\}_{i \in \mathbb{N}}$ denote a sequence of block lengths, which are iid random variables having the geometric distribution with a scalar parameter $\gamma \equiv \gamma_T \in (0,1)$:

$$P^*(L_i = l) = \gamma(1-\gamma)^{l-1}, \quad \forall i \in \mathbb{N},$$

for each positive integer l, where P^* denotes the conditional expectation given the original sample. We assume that the parameter γ satisfies the following growth condition

Assumption A5. $\gamma + (\sqrt{T}\gamma)^{-1} \to 0$ as $T \to \infty$.

Let $\{K_i\}_{i\in\mathbb{N}}$ be iid random variables, which are independent of both the original data and $\{L_i\}_{i\in\mathbb{N}}$, and have the discrete uniform distribution on $\{k+1, \ldots, T\}$. We set $B_{K_i,L_i} = \{(x_{1t}, x_{2t-k})\}_{t=K_i}^{L_i-1}$ representing the block of length L_i starting with the K_i-th pair of observations. The SB procedure generate samples $\{(x_{1t}^*, x_{2t-k}^*)\}_{t=k+1}^{T}$ 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 $t > T$, (x_{1t}, x_{2t-k}) is set to be (x_{1j}, x_{2j-k}), where $j = k + (t \bmod (T-k))$ and $(x_{1k}, x_{20}) = (x_{1T}, x_{2T+k})$, where mod denotes the modulo operator.[2]

Using the SB resample, we obtain quantile estimates $\hat{q}_\alpha^* = (\hat{q}_1^*(\alpha_1), \hat{q}_2^*(\alpha_2))^\intercal$ for $q_\alpha = (q_1(\alpha_1), q_2(\alpha_2))^\intercal$ by solving the minimization problems:

$$\hat{q}_1^*(\alpha_1) = \arg\min_{v_1\in\mathbb{R}} \sum_{t=k+1}^{T} \pi_{\alpha_1}(x_{1t}^* - v_1) \quad \text{and} \quad \hat{q}_2^*(\alpha_2) = \arg\min_{v_2\in\mathbb{R}} \sum_{t=k+1}^{T} \pi_{\alpha_2}(x_{2t-k}^* - v_2).$$

We construct $\hat{q}_\alpha^*$ by using $(T-k)$ SB observations, while $\hat{q}_\alpha$ is based on T observations. The difference of sample sizes is asymptotically negligible given the finite lag order k.

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

$$\hat{\rho}_\alpha^*(k) = \frac{\sum_{t=k+1}^{T} \psi_{\alpha_1}(x_{1t}^* - \hat{q}_1^*(\alpha_1))\psi_{\alpha_2}(x_{2t-k}^* - \hat{q}_2^*(\alpha_2))}{\sqrt{\sum_{t=k+1}^{T} \psi_{\alpha_1}^2(x_{1t}^* - \hat{q}_1^*(\alpha_1))}\sqrt{\sum_{t=k+1}^{T} \psi_{\alpha_2}^2(x_{2t-k}^* - \hat{q}_2^*(\alpha_2))}}.$$

We consider the SB bootstrap to construct a confidence intervals for each statistic of p cross-quantilograms $\{\hat{\rho}_\alpha(1), \ldots, \hat{\rho}_\alpha(p)\}$ for a finite positive integer p and subsequently construct a confidence interval for the omnibus test based on the p statistics. To maintain possible dependence structures, we use $(T-p)$ pairs of observations $\{(x_{t-1}, \ldots, x_{t-p})\}_{t=p+1}^{T}$ to resample the blocks of random lengths.

Given a vector cross-quantilogram $\hat{\rho}_\alpha^{(p)*}$, we define the omnibus test based on the

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

SB resample as $\hat{Q}_{\alpha}^{(p)*} = T(\hat{\rho}_{\alpha}^{(p)*} - \hat{\rho}_{\alpha}^{(p)})^{\intercal}(\hat{\rho}_{\alpha}^{(p)*} - \hat{\rho}_{\alpha}^{(p)})$. The following lemma shows the validity of the SB procedure for the cross-quantilogram.

Theorem 2 *Suppose that Assumption A1-A5 hold. Then, in the sense of weak convergence conditional on the sample we have:*
(a) $\sqrt{T}\left(\hat{\rho}_{\alpha}^{(p)*} - \hat{\rho}_{\alpha}^{(p)}\right) \Longrightarrow^{*} \Lambda_{\alpha}^{(p)}\mathbb{B}^{(p)}(\alpha)$ *in probability;*
(b) Under the null hypothesis that $\rho_{\alpha}(1) = \cdots = \rho_{\alpha}(p) = 0$ *for every* $\alpha \in \mathcal{A}$*,*

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

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

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

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

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

$$H_1 : \rho_{\alpha}(k) \text{ is nonzero constant for some } \alpha \in \mathcal{A} \text{ and for some } k \in \{1,\ldots,p\}, \quad (6)$$

and the local alternative

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

where ζ is a finite non-zero scalar. Thus, under the local alternative, there exist a p-dimensional vector $\zeta_\alpha^{(p)}$ such that $\rho_\alpha^{(p)} = T^{-1/2}\zeta_\alpha^{(p)}$ with $\zeta_\alpha^{(p)}$ having at least one non-zero element for some $\alpha \in \mathcal{A}$.

While we consider the asymptotic power of test for the directional predictability over a range of quantiles with multiple lags in the following theorem, 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 $\sqrt{T}$-local alternatives.

Theorem 3 *Suppose that Assumptions A1-A5 hold. Then: (a) Under the fixed alternative in (6),*

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

(b) Under the local alternative in (7)

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

where $c_{Q,\tau} = \inf\{c : P(\sup_{\alpha\in\mathcal{A}} \|\Lambda_\alpha^{(p)}\mathbb{B}^{(p)}(\alpha)\|^2 \leq c)) \geq 1 - \tau\}$.

3.1.2 The Self-Normalized Cross-Quantilogram

The self-normalized approach was proposed in Lobato (2001) for testing the absence of autocorrelation of a time series that is not necessarily independent. The idea was recently extended by Shao (2010) to a class of asymptotically linear test statistics. 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 (2012) propose a procedure for improving the power of the M test, by dividing the original sample into subsamples before applying the self-normalization procedure. The self-normalized approach has a tight link with the fixed-b asymptotic framework, which was proposed by Kiefer et al. (2000) and has been 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. 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). The relation between size and power properties is consistent with simulation results reported in the cited papers above.

We use recursive estimates to construct a self-normalized cross-quantilogram. Given a subsample $\{x_t\}_{t=1}^{s}$, we can estimate sample quantile functions by solving minimization problems

$$\hat{q}_{1s}(\alpha_1) = \arg\min_{v_1 \in \mathbb{R}} \sum_{t=1}^{s} \pi_{\alpha_1}(x_{1t} - v_1) \text{ and } \hat{q}_{2s}(\alpha_2) = \arg\min_{v_2 \in \mathbb{R}} \sum_{t=1}^{s} \pi_{\alpha_2}(x_{2t} - v_2).$$

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 parameters.

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

$$\hat{\rho}_{\alpha,s}(k) = \frac{\sum_{t=k+1}^{s} \psi_{\alpha_1}(x_{1t} - \hat{q}_{1s}(\alpha_1))\psi_{\alpha_2}(x_{2t-k} - \hat{q}_{2s}(\alpha_2))}{\sqrt{\sum_{t=k+1}^{s} \psi_{\alpha_1}^2(x_{1t} - \hat{q}_{1s}(\alpha_1))}\sqrt{\sum_{t=k+1}^{s} \psi_{\alpha_2}^2(x_{2t-k} - \hat{q}_{2s}(\alpha_2))}},$$

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

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

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

$$\hat{S}_{\alpha}^{(p)} = T\hat{\rho}_{\alpha}^{(p)\intercal} \hat{A}_{\alpha p}^{-1} \hat{\rho}_{\alpha}^{(p)}.$$

The following theorem shows that $\hat{S}_{\alpha}^{(p)}$ is asymptotically pivotal. To distinguish the process used in the following theorem from the one used in the previous section, let $\{\bar{\mathbb{B}}^{(p)}(\cdot)\}$ denote a $3p$-dimensional, standard Brownian motion on $D^{3p}([0,1])$.

Theorem 4 *Suppose that Assumptions A1-A4 hold. Then, for each $\alpha \in \mathcal{A}$,*

$$\hat{S}_{\alpha}^{(p)} \Longrightarrow \bar{\mathbb{B}}^{(p)}(1)^\intercal \mathbb{A}_{(p)}^{-1} \bar{\mathbb{B}}^{(p)}(1),$$

where $\mathbb{A}_p = \int_\omega^1 \left\{\bar{\mathbb{B}}^{(p)}(r) - r\bar{\mathbb{B}}^{(p)}(1)\right\} \left\{\bar{\mathbb{B}}^{(p)}(r) - r\bar{\mathbb{B}}^{(p)}(1)\right\}^\intercal dr.$

The joint test based on finite multiple quantiles can be constructed in a similar manner, while the test based on a range of quantiles has a limiting distribution depending on the Kiefer process: this may be difficult to implement in practice.

The asymptotic null distribution in the above theorem can be simulated and a critical value, $c_{S,\tau}$, corresponding to a significance level τ is tabulated by using the

$(1-\tau)100\%$ percentile of the simulated distribution.[3] In the theorem below, we consider a power function of the self-normalized omnibus test statistic, $P(\hat{S}_{\alpha}^{(p)} > c_{S,\tau})$. For a fixed $\alpha \in \mathcal{A}$, we consider a fixed alternative

$$H_1 : \rho_{\alpha}(k) \text{ is nonzero constant for some } k \in \{1, \ldots, p\}, \tag{8}$$

and a local alternative

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

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

Theorem 5 *(a) Suppose that the fixed alternative in (8) and Assumptions A1-A4 hold. Then,*

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

(b) Suppose that the local alternative in (9) is true and Assumptions A1-A4 hold. Then,

$$\lim_{T\to\infty} P\left(\hat{S}_{\alpha}^{(p)} > c_{S,\tau}\right) = P\left(\left\{\bar{\mathbb{B}}^{(p)}(1) + (\Lambda_{\alpha}^{(p)}\Delta_{\alpha}^{(p)})^{-1}\zeta_{\alpha}^{(p)}\right\}^{\intercal} \mathbb{A}_p^{-1} \left\{\bar{\mathbb{B}}^{(p)}(1) + (\Lambda_{\alpha}^{(p)}\Delta_{\alpha}^{(p)})^{-1}\zeta_{\alpha}^{(p)}\right\} \geq c_{S,\tau}\right),$$

where $\Delta_{\alpha}^{(p)}$ is a $3p \times 3p$ matrix with $\Delta_{\alpha}^{(p)}(\Delta_{\alpha}^{(p)})^{\intercal} \equiv \Gamma^{p}(\alpha,\alpha)$.

4 The Partial Cross-Quantilogram

We define the partial cross-quantilogram, which measures the relationship between two events $\{x_{1t} \leq q_1(\alpha_1)\}$ and $\{x_{2t-k} \leq q_2(\alpha_2)\}$, while controlling for intermediate events between t and $t-k$ as well as whether some state variables exceed a given

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

quantile. Let $z_t \equiv (z_{1t}, \ldots, z_{lt})^\intercal$ be an l-dimensional vector for $l \geq 1$, which may include some of the lagged predicted variables $\{x_{1t-1}, \ldots, x_{1t-k}\}$, the intermediate predictors $\{x_{2t-1}, \ldots, x_{1t-k-1}\}$ and some state variables that may reflect some historical events up to t. We use $q_{z_i}(\beta_i)$ to denote the β_ith quantile of z_i given $\beta_i \in (0,1)$ for $i = 1, \ldots, l$ and define $q_{z,\beta} = (q_{z_1}(\beta_1), \ldots, q_{z_l}(\beta_l))^\intercal$, with $\beta = (\beta_1, \ldots, \beta_l)^\intercal$. To ease the notational burden in the rest of this section, we suppress the dependency of $q_{z_i}(\beta_i)$ on β_i for $i = 1, \ldots, l$ and use $q_z \equiv q_{z,\beta}$ and $q_{z_i} \equiv q_{z_i}(\beta_i)$. We present results for the single quantile α and a single lag k, although the results can be extended to the case of a range of quantiles and multiple lags.

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

$$R_{\alpha k} = E\left[h_{t,\alpha k} h_{t,\alpha k}^\intercal\right] \quad \text{and} \quad P_{\alpha k} = R_{\alpha k}^{-1},$$

where $h_{t,\alpha k} = [\psi_{\alpha_1}(\epsilon_{1t}(\alpha_1)), \psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2)), \Psi_\beta(z_t - q_z)^\intercal]^\intercal$ with $\Psi_\beta(u) = [\psi_{\beta_1}(u_1), \ldots, \psi_{\beta_l}(u_l)]^\intercal$ for $u = (u_1, \ldots, u_l)^\intercal \in \mathbb{R}^l$. For $i, j \in \{1, \ldots, l\}$, let $r_{\alpha k,ij}$ and $p_{\alpha k,ij}$ be the (i,j) element of $R_{\alpha k}$ and $P_{\alpha k}$, respectively. Notice that the cross-quantilogram is $\rho_\alpha(k) = r_{\alpha k,12}/\sqrt{r_{\alpha k,11} r_{\alpha k,22}}$, and the partial cross-quantilogram is defined as

$$\rho_{\alpha|z}(k) = -\frac{p_{\alpha k,12}}{\sqrt{p_{\alpha k,11} p_{\alpha k,22}}}.$$

To obtain the sample analogue of the partial cross-quantilogram, we first construct a vector of hit processes, $\hat{h}_{t,\alpha k}$, by replacing the population quantiles in $h_{t,\alpha k}$ by the sample analogues $(\hat{q}_1(\alpha_1), \hat{q}_2(\alpha_2), \hat{q}_{z_1}, \ldots, \hat{q}_{z_l})$. Then, we obtain the estimator for the correlation matrix and its inverse as

$$\hat{R}_{\alpha k} = \frac{1}{T} \sum_{t=k+1}^{T} \hat{h}_{t,\alpha k} \hat{h}_{t,\alpha k}^\intercal \quad \text{and} \quad \hat{P}_{\alpha k} = \hat{R}_{\alpha k}^{-1},$$

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

$$\hat{\rho}_{\alpha|z}(k) = -\frac{\hat{p}_{\alpha k,12}}{\sqrt{\hat{p}_{\alpha k,11}\hat{p}_{\alpha k,22}}}, \tag{10}$$

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

In Theorem 6 below, we show that $\hat{\rho}_{\alpha|z}(k)$ asymptotically follows the normal distribution, while the asymptotic variance depends on nuisance parameters as in the previous section. To address the issue of the nuisance parameters, we employ the stationary bootstrap or the self-normalization technique. For the bootstrap, we can use pairs of variables $\{(x_{1t}, x_{2t-k}, z_t)\}_{t=k+1}^{T}$ to generate the SB resample $\{(x_{1t}^*, x_{2t-k}^*, z_t^*)\}_{t=k+1}^{T}$ and we then obtain the SB version of the partial cross-quantilogram, denoted by $\hat{\rho}_{\alpha|z}^*(k)$, using the formula in (10). When we use the self-normalized test statistics, we estimate the partial cross-quantilogram $\rho_{\alpha,s|z}(k)$ based on the subsample up to s, recursively and then we use

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

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

To obtain the asymptotic results, we impose the following conditions on the distribution function $F_{z_i}(\cdot)$ and the density function $F_{z_i}(\cdot)$ of each controlling variable z_i for $i = 1,\ldots,l$.

Assumption A6. For every $i \in \{1,\ldots,l\}$: **(a)** $\{z_{it}\}_{t\in\mathbb{Z}}$ is strictly stationary and strong mixing as assumed in Assumption A1; **(b)** The conditions in Assumption A2 and A3 hold for the $F_{z_i}(\cdot)$ and $f_{z_i}(\cdot)$ at the relevant quantile; **(c)** $G_{2z_i}(v_1, v_2) \equiv P(x_{2t} \leq v_2, z_{it} \leq v_3)$ for $(v_1, v_2) \in \mathbb{R}^2$ is continuously differentiable.

Assumption A6(a) requires the control variables z_t to satisfy the same weak dependent property as x_{1t} and x_{2t}. Assumption A6(b)-(c) ensure the smoothness of the distribution, density function and the joint distribution of x_{2t} and z_{it}.

Define the covariance matrix

$$\Gamma_{k,z}(\alpha) = \sum_{l=-\infty}^{\infty} \operatorname{cov}\left(\tilde{\xi}_l(\alpha,k), \tilde{\xi}_0^{\intercal}(\alpha,k)\right),$$

where $\tilde{\xi}_t(\alpha,k) = [\tilde{\xi}_{1t}(\alpha,k)^{\intercal}, \tilde{\xi}_{2t}(\alpha,k)^{\intercal}]^{\intercal}$ with

$$\tilde{\xi}_{1t}(\alpha,k) = \begin{bmatrix} \psi_{\alpha_1}(\epsilon_{1t}(\alpha_1))\psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2)) \\ \psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2))\Psi_\beta(z_t - q_z) \end{bmatrix} \text{ and } \tilde{\xi}_{2t}(\alpha,k) = \begin{bmatrix} \psi_{\alpha_1}(\epsilon_{1t}(\alpha_1)) \\ \psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2)) \\ \Psi_\beta(z_t - q_z) \end{bmatrix}^{\intercal}.$$

Also, let $\lambda_{\alpha k,z} = (p_{\alpha k,11}p_{\alpha k,22})^{-1/2}(\lambda_{\alpha k,z}^{(1)\intercal}\lambda_{\alpha k,z}^{(2)\intercal})^{\intercal}$, where

$$\lambda_{\alpha k,z}^{(1)} = \begin{bmatrix} 1 \\ -\mu_{z1}^{\intercal}\Sigma_z^{-1} \end{bmatrix} \text{ and } \lambda_{\alpha k,z}^{(2)} = \Xi_f^{-1}\left(\begin{bmatrix} \nabla G_k(q_\alpha) \\ \mathbf{0}_l \end{bmatrix} - \begin{bmatrix} 0 \\ \mu_{1z}^{\intercal}\Sigma_z^{-1}\partial \mathbf{G}_{2z}(v,q_z)/\partial v|_{v=q_2(\alpha_2)} \\ \mu_{1z}^{\intercal}\Sigma_z^{-1}\partial \mathbf{G}_{2z}(q_2(\alpha_2),w)/\partial w'|_{w=q_z} \end{bmatrix}\right),$$

with $\Xi_f = \operatorname{diag}\{f_1(q_1(\alpha_1)), f_2(q_2(\alpha_2)), f_{z_1}(q_{z_1}), \ldots, f_{z_l}(q_{z_l})\}$, $\mathbf{0}_l = (0,\ldots,0)^{\intercal} \in \mathbb{R}^l$, $\mu_{1z} = E[\psi_{\alpha_1}(\epsilon_{1t}(\alpha_1))\Psi_\beta(z_t - q_z)]$, $\mu_{2z} = E[\psi_{\alpha_2}(\epsilon_{2t-k}(\alpha_2))\Psi_\beta(z_t - q_z)]$, $\Sigma_z = E[\Psi_\beta(z_t - q_z)\Psi_\beta(z_t - q_z)^{\intercal}]$ and $\mathbf{G}_{2z}(v,w) = [G_{2z_1}(v,w_1),\ldots,G_{2z_l}(v,w_l)]^{\intercal}$ for $(v,w) \in \mathbb{R}\times\mathbb{R}^l$ with $w = (w_1,\ldots,w_l)^{\intercal}$.

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

Theorem 6 *(a) Suppose that Assumption A1-A4 and A6 hold. Then,*

$$\sqrt{T}(\hat{\rho}_{\alpha|z}(k) - \rho_{\alpha|z}(k)) \to^d N(0, \sigma_{\alpha|z}^2),$$

for each $\alpha \in \mathcal{A}$ and for a finite positive integer k, where $\sigma_{\alpha|z}^2 = \lambda_{\alpha k,z}^{\intercal}\Gamma_{k,z}(\alpha)\lambda_{\alpha k,z}$.

(b) Suppose that Assumption A1-A6 hold. Then,

$$\sup_{s\in\mathbb{R}} \left| P^*\left(\hat{\rho}^*_{\alpha|z}(k) \leq s\right) - P\left(\hat{\rho}_{\alpha|z}(k) \leq s\right)\right| \to^p 0,$$

for each $\alpha \in \mathcal{A}$ and for a finite positive integer k.

(c) Suppose that Assumption A1-A4 and A6 hold. Then, under the null hypothesis that $\rho_{\alpha|z}(k) = 0$, we have

$$\frac{\sqrt{T}\hat{\rho}_{\alpha|z}(k)}{\hat{A}^{1/2}_{\alpha|z}} \to^d \frac{\mathbb{B}(1)}{\left\{\int_\omega^1 \{\mathbb{B}(1) - r\mathbb{B}(r)\}^2 dr\right\}^{1/2}},$$

for each $\alpha \in \mathcal{A}$ and for a finite positive integer k.

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

5 Monte Carlo Simulation

We investigate the finite sample performance of our test statistics in the following two data generating processes:

DGP1: $(x_{1t}, x_{2t})^\intercal \sim iid\ N(0, I_2)$ where I_2 is a 2×2 identity matrix.

DGP2:

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

where $(\varepsilon_{1t}, \varepsilon_{2t})^\intercal \sim iid\ N(0, I_2)$ and $\sigma^2_{1t} = 0.1 + 0.2x^2_{1t-1} + 0.2\sigma^2_{1t-1} + x^2_{2t-1}$.

Under DGP1, there is no predictability from x_{2t-k} to x_{1t} for all quantiles.[4] Under DGP2, (x_{1t}) is defined as the GARCH-X process, where its conditional variance is the

[4] Even when we allow for moderate dependence in the DGP1 such as AR processes, we obtain similar results and the results are available upon request.

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 (2013) and references therein). Under DGP2, there is no predictability for the median because x_{1t} is symmetrically distributed but, for other quantiles, there exists predictability from x_{2t-k} to x_{1t} for $k \geq 1$ through σ_{1t}^2.

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 $\alpha_1 = \alpha_2$ are reported here because the results for cases where $\alpha_1 \neq \alpha_2$ are similar. The Box-Ljung test statistics $\hat{Q}_{\alpha}^{(p)}$ are based on $\hat{\rho}_{\alpha}(k)$ for $\alpha_1 = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$,0.95 and $k = 1, 2, \ldots, 5$. Tables 1 and 2 report the empirical rejection frequency of the Box-Ljung test statistics based on bootstrap critical values at the 5% level. We chose sample sizes $T = 500$, $1,000$ and $2,000$. We computed $ns = 1,000$ replications. The bootstrap critical values are based on $1,000$ 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)). Since it is for univariate data, we apply it separately to each time series and define γ as the average value.

In general, our simulation results in Tables 1 and 2 show that the test has relatively good size and power properties. Table 1 reports the simulation results for the DGP1, which show the size performance. The rejection frequencies are mostly close to 0.05. Except for $\alpha_1 = 0.05$ and 0.95 cases, the rejection frequency is uniformly within simulation error of the nominal value. For all quantiles including $\alpha_1 = 0.05$ and 0.95 cases, the rejection frequency becomes closer to 0.05 as sample size increases.

Table 1. (size) Empirical rejection frequency of the Box-Ljung test statistic $\hat{Q}_{\alpha}^{(p)}$ (DGP1 and the nominal level: 5%)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.147	0.064	0.025	0.039	0.047	0.042	0.050	0.040	0.163
	2	0.020	0.020	0.026	0.035	0.026	0.034	0.039	0.024	0.039
	3	0.007	0.012	0.024	0.025	0.024	0.029	0.035	0.017	0.015
	4	0.006	0.008	0.019	0.026	0.017	0.029	0.028	0.010	0.004
	5	0.003	0.007	0.021	0.026	0.023	0.029	0.024	0.010	0.006
1000	1	0.071	0.050	0.039	0.044	0.042	0.040	0.044	0.034	0.078
	2	0.022	0.032	0.031	0.039	0.028	0.031	0.034	0.030	0.032
	3	0.018	0.028	0.027	0.039	0.030	0.023	0.035	0.017	0.022
	4	0.011	0.025	0.035	0.034	0.035	0.026	0.029	0.015	0.011
	5	0.014	0.020	0.018	0.031	0.031	0.026	0.024	0.016	0.007
2000	1	0.077	0.051	0.038	0.045	0.050	0.049	0.040	0.044	0.073
	2	0.036	0.026	0.036	0.036	0.043	0.037	0.036	0.024	0.034
	3	0.024	0.026	0.038	0.046	0.038	0.041	0.034	0.020	0.024
	4	0.015	0.035	0.042	0.038	0.041	0.042	0.030	0.015	0.020
	5	0.017	0.034	0.040	0.039	0.049	0.043	0.030	0.021	0.011

Notes: The first and second columns report the sample size T and the number of lags p for the Box-Ljung test statistics $\hat{Q}_{\alpha}^{(p)}$. The rest of columns show empirical rejection frequencies based on bootstrap critical values at the 5% significance level. The tuning parameter γ is chosen as explained above.

Table 2 reports the simulation results for the DGP2. As explained above, the null hypothesis of no predictability is satisfied at the median under the DGP2 and the rejection frequency is close to 0.05 at the median and moreover power is lower when quantile is closer to median in Table 2. However, for other quantiles, the rejection frequency approaches one in the largest sample size.

5.2 Results for the Self-Normalized Statistics

We also examine the performance of the self-normalized version of $\hat{Q}_{\alpha}^{(p)}$ under the same setup as above. We let the trimming value ω be 0.1. We also considered 0.03 and 0.05 for ω and the results are similar to those for $\omega = 0.1$. We provide the results

Table 2. (power) Empirical rejection frequency of the Box-Ljung test statistic $\hat{Q}_{\alpha}^{(p)}$
(DGP2: GARCH-X process)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.197	0.409	0.373	0.161	0.042	0.179	0.412	0.450	0.191
	2	0.223	0.403	0.377	0.134	0.037	0.149	0.389	0.446	0.213
	3	0.188	0.366	0.317	0.110	0.032	0.122	0.340	0.407	0.195
	4	0.166	0.323	0.276	0.098	0.030	0.100	0.290	0.370	0.168
	5	0.149	0.284	0.242	0.088	0.030	0.089	0.257	0.334	0.153
1000	1	0.555	0.813	0.739	0.347	0.040	0.337	0.765	0.817	0.590
	2	0.565	0.814	0.708	0.295	0.043	0.295	0.749	0.819	0.582
	3	0.531	0.779	0.665	0.254	0.035	0.251	0.693	0.780	0.553
	4	0.501	0.741	0.629	0.217	0.040	0.216	0.637	0.747	0.512
	5	0.463	0.702	0.579	0.194	0.034	0.181	0.594	0.725	0.485
2000	1	0.914	0.992	0.964	0.634	0.039	0.617	0.977	0.983	0.923
	2	0.932	0.989	0.968	0.591	0.038	0.583	0.978	0.987	0.923
	3	0.917	0.989	0.961	0.530	0.034	0.527	0.959	0.988	0.920
	4	0.909	0.983	0.952	0.492	0.040	0.472	0.946	0.987	0.909
	5	0.900	0.980	0.936	0.448	0.037	0.435	0.935	0.980	0.888

Notes: Same as Table 1.

for $\omega = 0.05$ in the supplementary material. The number of repetitions is $3,000$. The empirical sizes of the test are reported in Table 3, where the underlying process is DGP1 and the nominal size is 5%. The rejection frequencies for quantiles between 0.1 and 0.9 are close to 5% nominal level or a little conservative regardless of sample sizes. At 0.05 and 0.95 quantiles, the result shows size distortions for $T = 500$, but this becomes moderate for $T = 1000$ and empirical sizes become close to the nominal size for $T = 2000$. Using the GARCH-X process as in DGP2, we obtain empirical powers and present the result in Table 4. With a one-period lag ($p = 1$), the self-normalized quantilogram at $\alpha_1, \alpha_2 \in \{0.1, 0.2, 0.8, 0.9\}$ rejects the null by about 25.7-26.6%, 51.1-56.2% and 78.9-84.8% for sample sizes 500, 1,000 and 2,000, respectively. In general, the rejection frequencies increase as the sample size increases, decline as the maximum number of lags p increases, and is not sensitive to the choice of the trimming value.

Our results show that the self-normalized statistics have lower power in finite samples compared with the stationary bootstrap.

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

Table 3. (size) Empirical Rejection Frequencies of the Self-Normalized Statistics (DGP1 and the nominal level: 5%)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.176	0.038	0.032	0.030	0.027	0.028	0.026	0.034	0.196
	2	0.285	0.029	0.022	0.022	0.023	0.021	0.019	0.026	0.347
	3	0.391	0.031	0.020	0.017	0.019	0.015	0.017	0.031	0.459
	4	0.475	0.026	0.011	0.010	0.013	0.009	0.012	0.028	0.560
	5	0.539	0.032	0.008	0.007	0.007	0.008	0.008	0.035	0.643
1000	1	0.051	0.033	0.039	0.037	0.035	0.034	0.033	0.037	0.061
	2	0.095	0.027	0.029	0.024	0.026	0.026	0.028	0.027	0.102
	3	0.126	0.026	0.022	0.019	0.023	0.020	0.022	0.028	0.146
	4	0.162	0.021	0.015	0.012	0.015	0.017	0.016	0.020	0.182
	5	0.194	0.019	0.011	0.014	0.009	0.012	0.013	0.021	0.211
2000	1	0.042	0.037	0.040	0.040	0.041	0.045	0.036	0.035	0.053
	2	0.044	0.033	0.028	0.032	0.031	0.038	0.030	0.029	0.051
	3	0.048	0.026	0.020	0.030	0.030	0.028	0.025	0.025	0.052
	4	0.052	0.024	0.017	0.018	0.020	0.019	0.021	0.020	0.055
	5	0.055	0.022	0.018	0.019	0.019	0.016	0.016	0.018	0.054

Notes: The first and second columns report the sample size T and the number of lags p for the test statistics $\hat{Q}^p_\alpha$. 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 4. (power) Empirical Rejection Frequencies of the Self-Normalized Statistics (DGP2: GARCH-X process)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.134	0.259	0.266	0.110	0.031	0.121	0.263	0.257	0.105
	2	0.104	0.169	0.173	0.067	0.023	0.073	0.182	0.161	0.084
	3	0.149	0.109	0.111	0.042	0.015	0.049	0.118	0.095	0.159
	4	0.232	0.077	0.069	0.033	0.012	0.031	0.083	0.061	0.282
	5	0.328	0.059	0.051	0.020	0.008	0.020	0.060	0.050	0.402
1000	1	0.359	0.555	0.523	0.231	0.038	0.231	0.511	0.547	0.338
	2	0.232	0.449	0.417	0.164	0.025	0.176	0.420	0.449	0.214
	3	0.158	0.352	0.319	0.110	0.020	0.128	0.335	0.336	0.157
	4	0.136	0.269	0.241	0.086	0.018	0.093	0.253	0.256	0.141
	5	0.152	0.211	0.198	0.064	0.012	0.062	0.194	0.200	0.146
2000	1	0.677	0.842	0.789	0.443	0.044	0.445	0.804	0.842	0.660
	2	0.561	0.778	0.729	0.355	0.034	0.369	0.746	0.784	0.537
	3	0.451	0.700	0.661	0.279	0.025	0.292	0.660	0.700	0.438
	4	0.366	0.637	0.580	0.214	0.019	0.236	0.604	0.628	0.351
	5	0.305	0.580	0.525	0.183	0.014	0.191	0.547	0.559	0.282

Notes: Same as Table 3.

of an economic state variable is helpful in predicting the mean of stock return (mean-to-mean 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. They show that while the OLS estimates of regression coefficients for the mean return forecast are mostly insignificant even at the 10% level (see Table 2), the mean of some predictors is helpful in predicting certain quantiles of stock returns (see Table 3). Specifically, they considered the following linear quantile regression

$$q_\alpha(r_{t+1}|\mathcal{F}_t) = \beta_{0,\alpha} + \beta_{1,\alpha}x_t + \beta_{2,\alpha}q_\alpha(r_t|\mathcal{F}_{t-1}) + \beta_{3,\alpha}\left|r_t\right|,$$

where r_t and x_t are stock returns and an economic state variable, respectively, and $q_\alpha(r_{t+1}|\mathcal{F}_t)$ is the conditional quantile of stock return given the information $\mathcal{F}_t$ at time t. They include $q_\alpha(r_t|\mathcal{F}_{t-1})$ and $|r_t|$ in the model following Engle and Manganelli (2004). Maynard et al. (2011) investigate inferences in such linear quantile regressions when a predictive regressor has a near-unit root.

Our cross-quantilogram adds one more dimension to analyze predictability compared with the linear quantile regression, and so it shows a more complete relationship between a predictor and stock returns. Moreover, as pointed out by Linton and Whang (2007), we can consider very large lags in the framework of the quantilogram. For example, in our application, We consider lags $k = 60$, i.e. 5 years.

We use the data set of monthly stock returns and the predictor variables previously analyzed in Goyal and Welch (2008). Stock returns are measured by the S&P 500 index and include dividends. A Treasury-bill rate is subtracted from stock returns to give excess returns. There are a total of 16 variables considered as predictors in Goyal and Welch (2008). However, most time series used as predictors are highly persistent. For example, both the dividend-price ratio and the earnings-price ratio have autoregressive coefficients being estimated to be 0.99, and unit root tests show that there exists a near-unit root in each time series. This motivated the work by Maynard et al. (2011). Since our paper establishes asymptotic and bootstrap theories for the cross-quantilogram for strictly stationary time series, we do not consider these variables. Extension of the quantilogram for such persistent time series will be interesting and useful. We leave it as future work.

In our application, we consider stock variance as the predictor because the autoregressive coefficient for stock variance is estimated to be 0.59 and the unit root hypothesis is clearly rejected. The stock variance is a volatility estimate based on daily squared returns and is treated as an estimate of equity risk in the literature. The risk-return relationship is a very important issue in the literature. See Lettau and Ludvigson (2010) for an extensive review. The cross-quantilogram can provide quantile-to-quantile relationship from risk to return, which cannot be examined using existing methods. The sample period is from Feb. 1885 to Dec. 2005 with sample size 1,451. The sample mean and median of stock return are 0.0008 and 0.0032, respectively.

In Figures 1(a)-3(b), we provide the cross-quantilogram $\hat{\rho}_\alpha(k)$ and the portmanteau tests $\hat{Q}_\alpha^{(p)}$ (we use the Box-Ljung versions throughout) to detect directional predictability from stock variance, representing risk, to stock return. For the quantiles of stock return $q_1(\alpha_1)$, we consider a wide range for $\alpha_1 = 0.05, 0.1, 0.2, 0.3, 0.5, 0.7, 0.8, 0.9$ and 0.95. For the quantiles of stock variance $q_2(\alpha_2)$, we consider $\alpha_2 = 0.1$, 0.5 and 0.9. The tuning parameter γ is chosen as explained in Section 5.1. In each graph, we show the 95% bootstrap confidence intervals for no predictability based on $1,000$ bootstrapped replicates.

Figures 1(a) and 1(b) are for the case when the stock variance is in the lower quantile, i.e. $q_2(\alpha_2)$ for $\alpha_2 = 0.1$. The cross-quantilogram $\hat{\rho}_\alpha(k)$ for $\alpha_1 = 0.05$ is negative and significant for some lags. It means that when risk is very low, it is less likely to have a very large negative loss. The cross-quantilograms $\hat{\rho}_\alpha(k)$ for $\alpha_1 = 0.8$, 0.9 and 0.95 are positive and significant for some lags. For example, for $\alpha_1 = 0.95$, the cross-quantilogram is significantly positive for about a half year. This means that when risk is very low, it is less likely to have a very large positive gain for the next half year. Figure 1(b) shows that the Box-Ljung test statistics are significant for quantiles $\alpha_1 = 0.05$, 0.8, 0.9 and 0.95.

Figures 2(a) and 2(b) are for the case when the stock variance is in the median, i.e. $q_2(\alpha_2)$ for $\alpha_2 = 0.5$. If the distributions of stock returns and the predictor are symmetric, the median return forecast will be equal to the mean return forecast. In the case of $\alpha_1 = 0.5$ and $\alpha_2 = 0.5$, the cross-quantilograms are insignificant for almost all lags. This result corresponds to the mean return forecast result given in Table 2

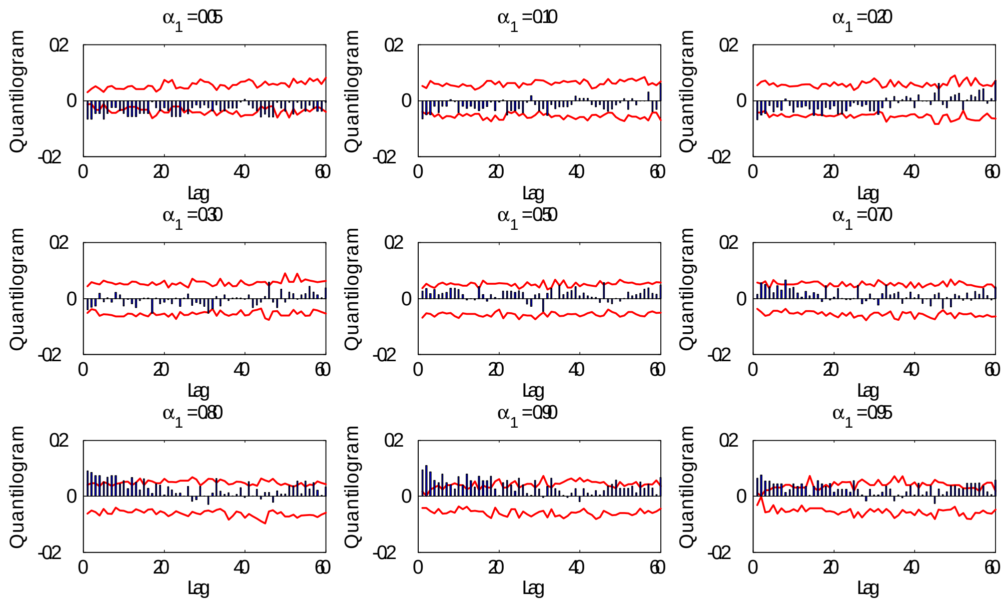

Figure 1(a). The sample cross quantilogram $\hat{\rho}_\alpha(k)$ for α_2= 0.1 to detect directional predictability from stock variance to stock return. Bar graphs describe sample cross-quantilograms and lines are the 95% bootstrap confidence intervals.

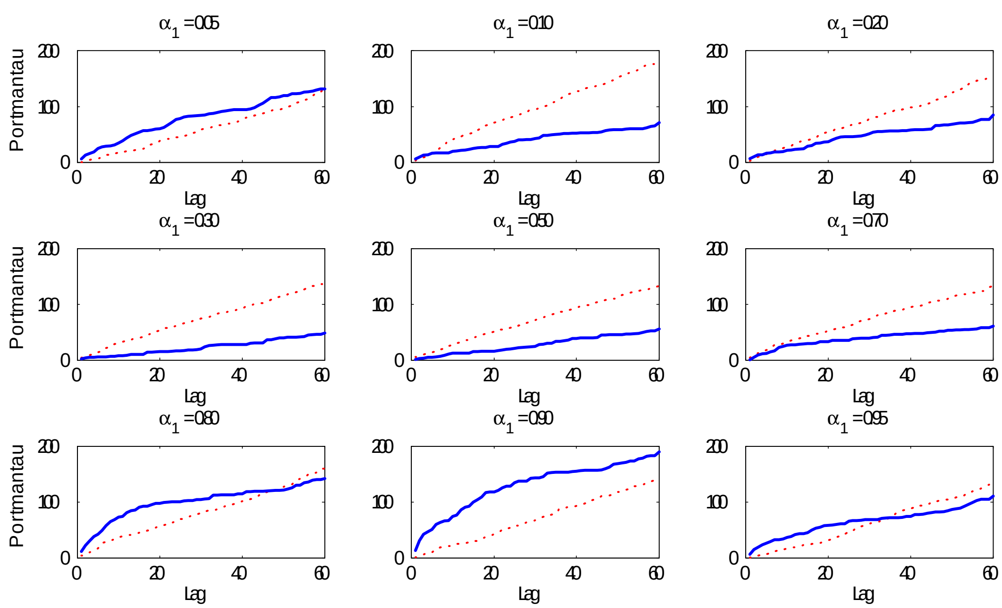

Figure 1(b). Box-Ljung test statistic $\hat{Q}_\alpha^{(p)}$ for each lag p and quantile α using $\hat{\rho}_\alpha(k)$ with α_2= 0.1. The dashed lines are the 95% bootstrap confidence intervals.

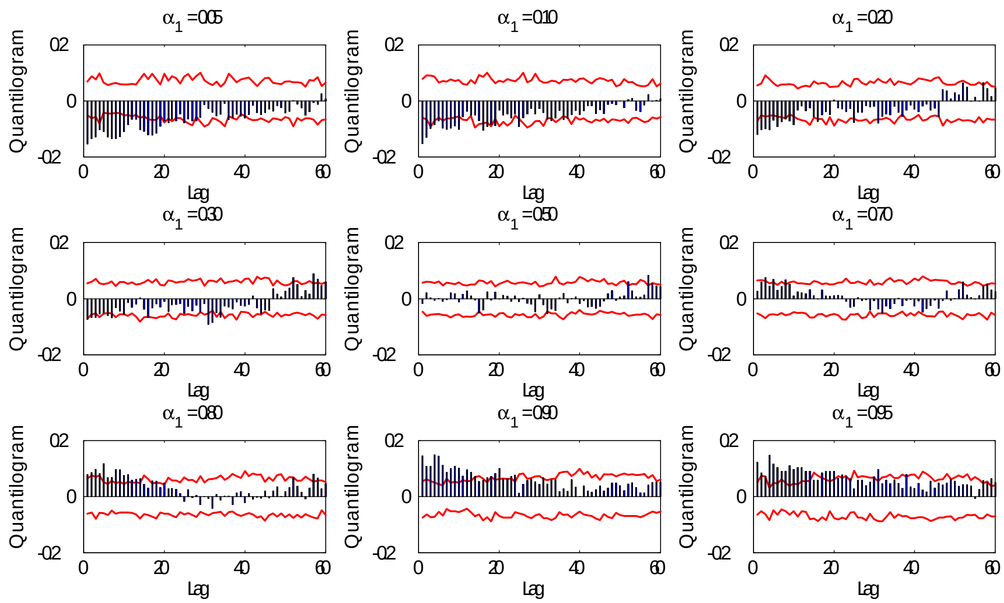

Figure 2(a). The sample cross quantilogram $\hat{\rho}_{\alpha}(k)$ for α_2= 0.5 to detect directional predictability from stock variance to stock return. Same as Figure 1(a)..

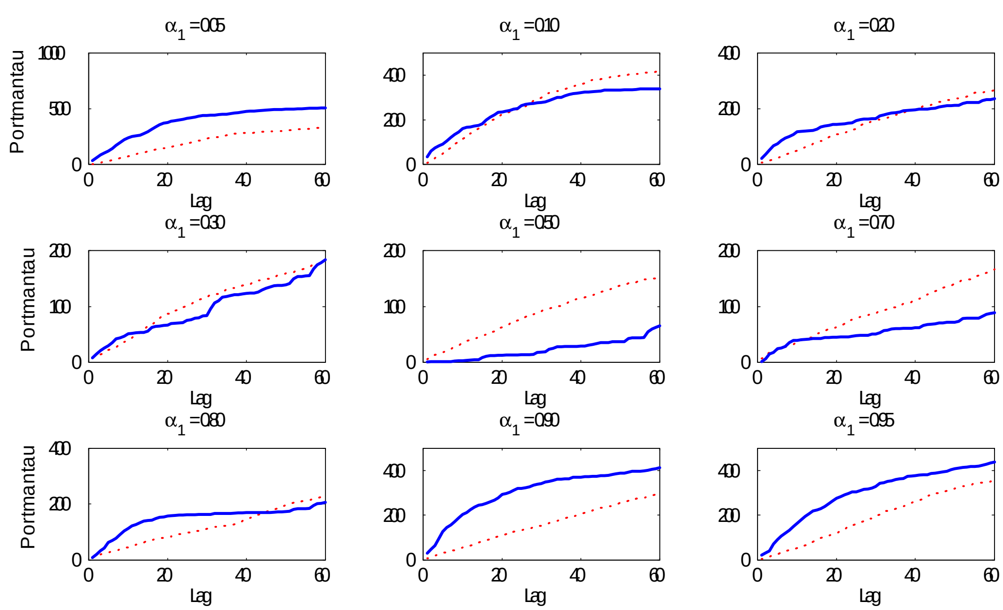

Figure 2(b). Box-Ljung test statistic $\hat{Q}_{\alpha}^{(p)}$ for each lag p and quantile α using $\hat{\rho}_{\alpha}(k)$ with α_2= 0.5. Same as Figure 1(b).

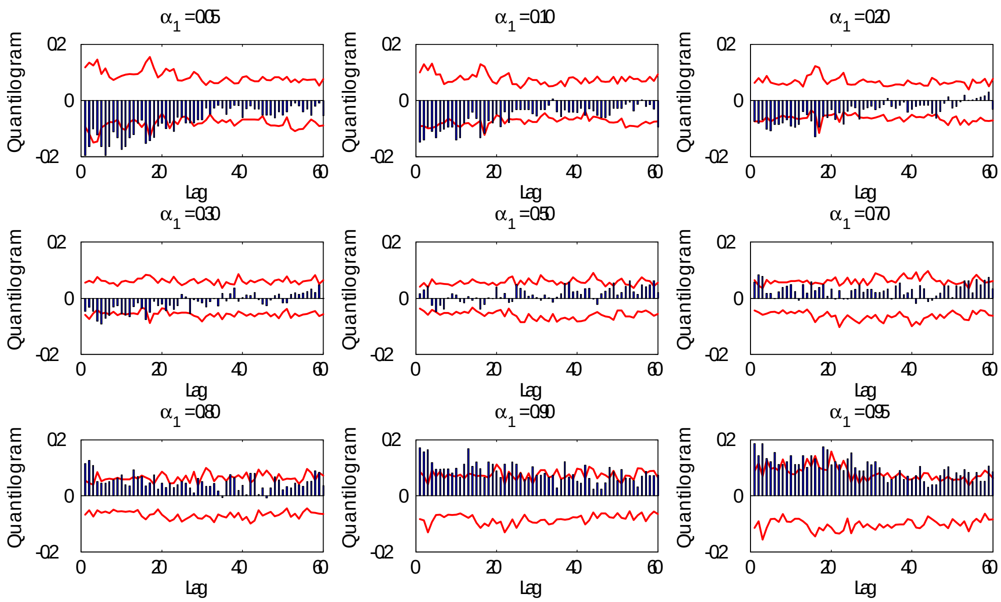

Figure 3(a). The sample cross quantilogram $\hat{\rho}_{\alpha}(k)$ for α_2= 0.9 to detect directional predictability from stock variance to stock return. Same as Figure 1(a).

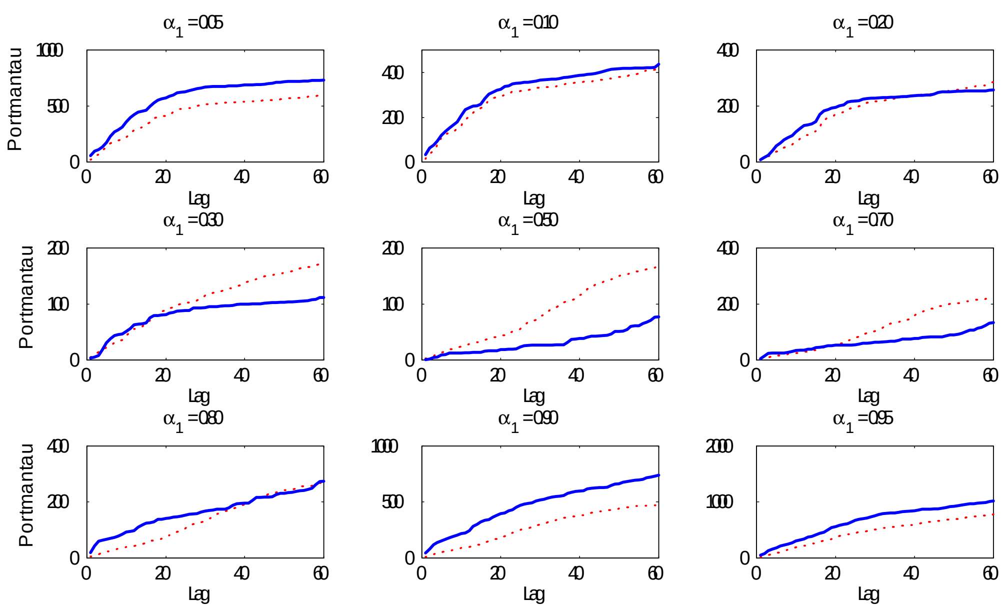

Figure 3(b). Box-Ljung test statistic $\hat{Q}^{(p)}_{\alpha}$ for each lag p and quantile α using $\hat{\rho}_{\alpha}(k)$ with α_2= 0.9. Same as Figure 1(b).

by Cenesizoglu and Timmermann (2008), where the OLS estimate for the mean return forecast is also insignificant. However, if we consider other quantiles of stock returns, the stock variance appears to have some predictability for stock returns. It should be noted that the results of the linear quantile regression given in Table 3 by Cenesizoglu and Timmermann (2008) correspond to our results for the first lag in Figure 2(a).

For $\alpha_1 = 0.05$, 0.1 and 0.2, the cross-quantilograms are negative and significant for some lags. This implies that when risk is lower than the median, it is less likely to have large negative losses. For $\alpha_1 = 0.8$, 0.9 and 0.95, the cross-quantilograms are positive and significant for some lags. For example, for $\alpha_1 = 0.95$, the cross-quantilogram is significantly positive for about one and a half years. This means that when risk is lower than median, it is less likely to have a very large positive gain for the next one and a half years. Figure 2(b) shows that the Box-Ljung test statistics are significant for both tails and both shoulders of the stock return distribution.

Figures 3(a) and 3(b) are for the case when stock variance is in the higher quantile, i.e. $q_2(\alpha_2)$ for $\alpha_2 = 0.9$. Compared to the previous case of $\alpha_2 = 0.5$, the cross quantilograms have similar trends but larger absolute values. For $\alpha_1 = 0.05$, the cross-quantilograms are negative and significant for about two years. This implies that when risk is higher than the 0.9 quantile, there is an increased likelihood of having very large negative losses up to two years. For $\alpha_1 = 0.95$, the cross quantilograms are positive and significant for about three years. This implies that when risk is very high (higher than the 0.9 quantile), there is an increased likelihood of having a very large positive gain for the next three years. Figure 3(b) shows that the Box-Ljung test statistics are significant for some lags in most quantiles of stock returns except for $\alpha_1 = 0.5$.

The results in Figures 1(a)-3(b) show that stock variance is helpful in predicting stock return and detailed features depend on each quantile of stock variance and stock return. Stock variance is helpful in predicting particularly both tails and both shoulders of stock return distribution. When stock variance is in higher quantile, in general 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 quantile-to-quantile 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. Bisias et al. (2012) summarized that the current approaches to measure systemic risk can be categorized as 1) tail measures, 2) contingent claims analysis, 3) network models and 4) dynamic stochastic macroeconomic models. Among these, the first approach is related to the cross-quantilogram and it measures co-dependence in the tails of equity returns of an individual financial institution and the financial system. 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), Goldman Sachs (GS) and AIG as individual financial institutions. JPM belongs to the Depositories group and Brownlees and Engle (2012) put GS with Broker-Dealers group. AIG belongs to the Insurance group. The sample period is from 2 Jan. 2001 to 30 Dec. 2011 with sample size 2,767. We investigate the cross-quantilogram $\hat{\rho}_\alpha(k)$ between an individual institution's stock return and the market index return for $k = 60$ and $\alpha_1 = \alpha_2 = 0.05$. In each graph, we show the 95% bootstrap confidence intervals for no quantile dependence based on 1,000 bootstrapped replicates.

Figure 4. (a1)-(a3) shows the cross-quantilogram from an individual institution to the market. The cross-quantilograms are positive and significant for large lags. The cross-quantilogram from JPM to the market reaches its peak (0.139) 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 GS to the market, it shows a similar pattern by reaching its peak (0.147) at $k = 10$. However, from AIG to the market, it exhibits a different trend by reaching its peak (0.136) at $k = 52$. When AIG is in distress, the systemic risk from AIG takes a longer time (more than two months) to reach its peak.

Figure 4. (b1)-(b3) 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 is similar to the stress tests performed by individual institutions. From the market to JPM, it reaches its peak (0.153) at $k = 2$. From the market to GS, it reaches its peak (0.162) at $k = 10$ and declines afterwards. Meanwhile, from the market to AIG, it reaches its peak (0.133) at $k = 27$. When the market is in distress, each institution is influenced by its impact in a different way.

As shown in Figure 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

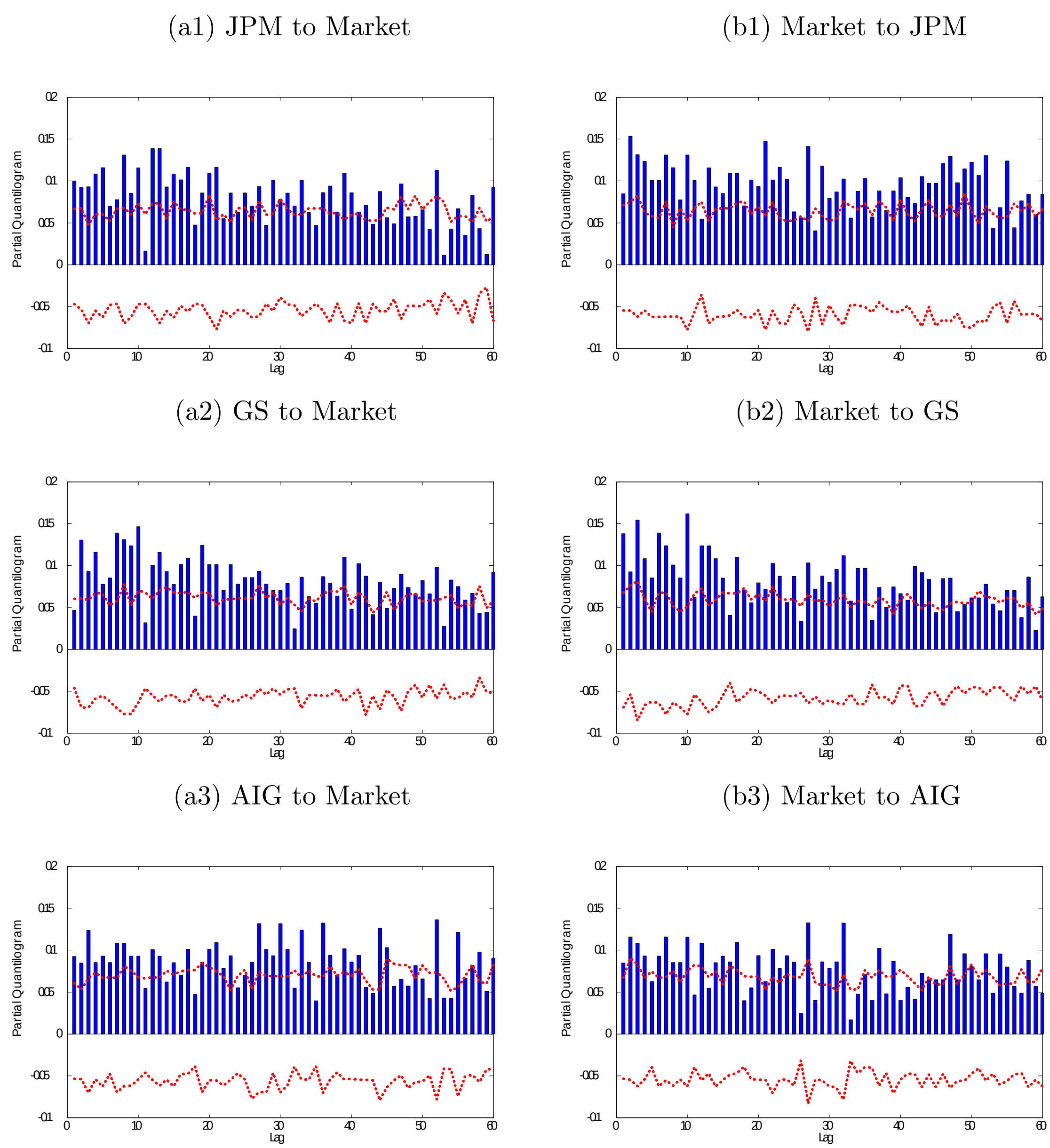

Figure 4. The sample cross quantilogram $\hat{\rho}_\alpha(k)$. Bar graphs describe sample cross-quantilograms and dotted lines are the 95% bootstrap confidence intervals.

example, White et al. (2012) adopts 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

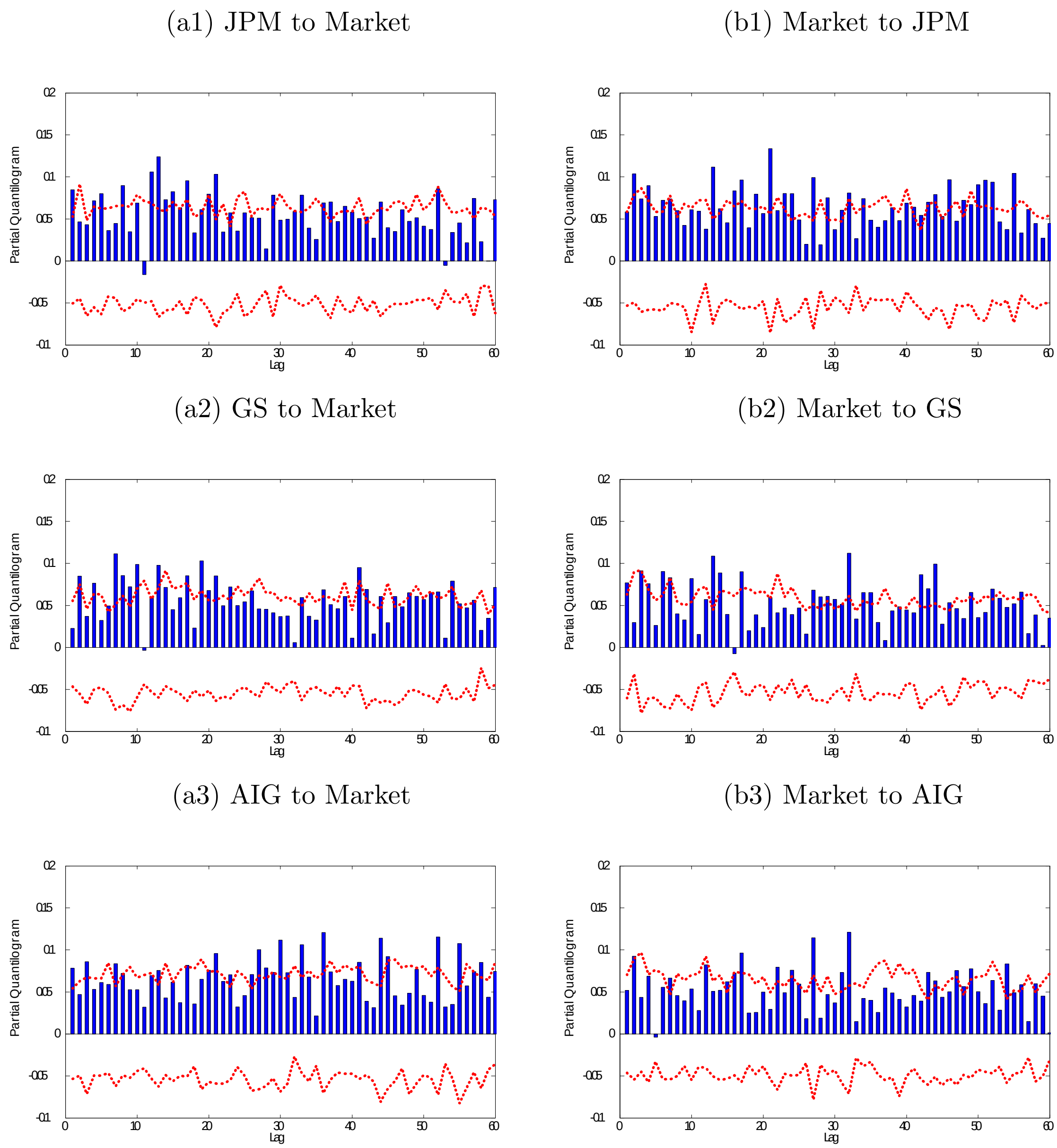

Figure 5. The sample partial cross-quantilogram $\hat{\rho}_{\alpha|z}(k)$. Bar graphs describe sample partial quantilograms and dotted lines are the 95% bootstrap confidence intervals.

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. β in (10), we let $\beta = 0.95$ because a low quantile of a stock return is generally associated with a rapid increase of the VIX index.

Figure 5 shows that the partial cross-quantilograms are still significant in some cases even if their values are generally smaller than the values of the cross-quantilograms in Figure 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.

Supplementary Materials

Section A: includes all proofs of theorems.

Section B: contains additional Figures.

Section C: provides computer codes and datasets.

References

Adrian, T. and M.K. Brunnermeier (2011) CoVaR, Tech. rep., Federal Reserve Bank of New York, Staff Reports

Bisias, D., M. Flood, A.W. Lo and S. Valavanis (2012) A survey of systemic risk analytics, working paper, Office of Financial Research.

Brownlees, C. and R.F. Engle (2012) Volatility, correlation and tails for systemic risk management, Tech. rep., SSRN.

Bedljkovic, M. (2010) Nonparametric test of conditional quantile independence with an applicationi to banks' systemic risk, working paper.

Bunzel, H., N.M. Kiefer and T.J. Vogelsang (2001) Simple robust testing of hypotheses in nonlinear models, *Journal of the American Statistical Association*, 96, 1088-1096.

Campbell, J.Y., A.W. Lo, and A.C. MacKinlay (1997): *The Econometrics of Financial Markets*, Princeton University Press, Princeton.

Cenesizoglu, T. and A. Timmermann (2008) Is the distribution of stock return predictable?, University of California at San Diego, working paper.

Chang, Chiao Yi and Fu Shuen Shie (2011). The Relation Between Relative Order Imbalance and Intraday Futures Returns: An Application of the Quantile Regression Model to Taiwan. Emerging Markets Finance and Trade, Volume 47, Number 3, pp 69 - 87.

Chen, Y.-T. and Z. Qu (2012) M tests with a new normalization matrix, *Econometric Reviews*, forthcoming.

Christoffersen, P.F., and F.X. Diebold (2002) Financial Asset Returns, market timing, and volatility dynamics, manuscript.

Cowles, A., and H. Jones (1937) Some A Posteriori Probabilities in Stock Market Action, *Econometrica* 5, 280-294.

Davis, R.A. and T. Mikosch (2009) The extremogram: a correlogram for extreme events, *Bernoulli*, 15, 977-1009.

Dette, H., M. Hallin, T. Kley, and Stanislav Volgushev (2013) Of Copulas, Quantiles, Ranks, and Spectra an L1 -approach to spectral analysis. Manuscript.

Dufour, J.M., M. Hallin, and I. Mizera (1998) Generalized Runs tests for heteroscedastic Time Series, *Nonparametric Statistics* 9, 39-86.

Embrechts, P., C. Kluppelberg and T. Mikosch (1997), Modelling extremal events for insurance and finance, Springer, New York.

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

Fama, E. (1965) The behavior of stock market prices, J. Business 38, 34 105.

Goyal, A. and I. Welch (2008) A comprehensive look at the empirical performance of equity premium prediction, *Review of Financial Studies*, 21, 1455-1508.

Hagemann, A (2012) Robust Spectral Analysis,
http://econ.la.psu.edu/papers/hagemann92413.pdf

Han, H. (2013) Asymptotic properties of GARCH-X processes, *Journal of Financial Econometrics*, forthcoming.

Hong, Y. (1996) Consistent Testing for Serial Correlation of Unknown Form. Econometrica, Vol. 64, No. 4, , pp. 837-864

Ibragimov, R. (2009), Heavy-tailed densities, in The New Palgrave Dictionary of Economics. Eds. Steven N. Durlauf and Lawrence E. Blume. Palgrave Macmillan.

Ibragimov, R., D. Jaffee and J. Walden (2009), Nondiversification traps in catastrophe insurance markets, Review of Financial Studies, 22, 959-993.

Kiefer, N.M. and T.J. Vogelsang (2002) Heteroskedasticity-autocorrelation robust standard errors using the Bartlett kernel without truncation, *Econometrica*, 70, 2093-2095.

Kiefer, N.M. and T.J. Vogelsang (2005) A new asymptotic theory for heteroskedasticity-autocorrelation robust tests, *Econometric Theory*, 21, 1130-1164.

Kiefer, N.M., T.J. Vogelsang and H. Bunzel (2000) Simple robust testing of regression hypotheses, *Econometrica*, 68, 695-714.

Kim, M.S. and Y. Sun (2011) Spatial heteroskedasticity and autocorrelation consistent estimation of covariance matrix, *Journal of Econometrics*, 160, 349-371.

Koenker, R. and W.G. Bassett (1978) Regression quantiles, *Econometrica*, 46, 33-50.

Kuan, C.-M. and W.-M. Lee (2006) Robust M tests without consistent estimation of the asymptotic covariance matrix, *Journal of the American Statistical Association*, 101, 1264-1275.

Laurini, Márcio Poletti, Luiz Gustavo Cassilatti Furlani, Marcelo Savino Portugal (2008), Empirical market microstructure: An analysis of the BRL/US$ exchange rate market, Emerging Markets Review, Volume 9, Issue 4, December 2008, Pages 247-265

Lettau, M. and S.C. Ludvigson (2010) Measuring and modeling variation in the risk-return trade-off, in *Handbook of Financial Econometrics*, ed. by Y. Ait-Sahalia and L.P. Hansen, vol. 1, 617-690, North-Holland.

Linton, O. and Y-J. Whang (2007) The quantilogram: With an application to evaluating directional predictability, *Journal of Econometrics*, 141, 250-282.

Lobato, I.N. (2001) Testing that a dependent process is uncorrelated, *Journal of the American Statistical Association*, 96, 1066-1076.

Mandelbrot, B.B., (1963) The variation of certain speculative prices. Journal of Business (Chicago) 36, 394–419. Reprinted in Cootner (1964), as Chapter E 14 of Mandelbrot (1997), in Telser (2000), and several other collections of papers on finance.

Maynard, A., K. Simotsu and Y. Wang (2011) Inference in predictive quantile regressions, University of Guelph, working paper.

Mikosch, T., and C. Starica (2000), Limit Theory for the sample autocorrelations and extremes of a GARCH(1,1) process, Annals of Statistics, 28, 1427-1451.

Patton, A., D.N. Politis and H. White (2009) Correction to "automatic block-length selection for dependent bootstrap" by D. Politis and H. White, *Econometric Reviews*, 28, 372-375.

Politis, D.N. and J.P. Romano (1994) The stationary bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.

Politis, D.N. and H. White (2004) Automatic block-length selection for dependent bootstrap, *Econometric Reviews*, 23, 53-70.

Pollard, D. (1991) Asymptotics for least absolute deviation regression estimators, *Econometric Theory*, 7, 186-199.

Rachev, S. and S. Mittnik (2000), *Stable Paretian Models in Finance*. Wiley, New York.

Rio, E. (2000) *Theorie Asymptoticque des Processus Aleatoires Faiblement Dependants*, no. 31 in Mathmaticues and Applications, Springer.

Shao, X. (2010) A self-normalized approach to confidence interval construction in time series, *Journal of the Royal Statistical Society Series B*, 72, 343-366.

Sun, Y. and M.S. Kim (2012) Simple and powerful GMM over-identification tests with accurate size, *Journal of Econometrics*, 166, 267-281.

Sun, Y. P.C.B. Phillips and S. Jin (2008) Optimal bandwidth Selection in Heteroskedasticity-autocorrelation robust testing, *Econometrica*, 76, 175-194.

White, H., T-H. Kim and S. Manganelli (2012) VAR for VaR: Measuring tail dependence using multivariate regression quantiles, working paper.

Supplementary Material for "The Cross-Quantilogram: Measuring Quantile Dependence and Testing Directional Predictability between Time Series"

Heejoon Han Oliver Linton Tatsushi Oka Yoon-Jae Whang

A Proofs of Main Results

We use $C, C_1, C_2, \ldots$ to denote generic positive constants without further clarification. In what follows, we assume that $\mathcal{A}$ is a compact subset of $(0,1)^2$, while the results can be easily extend to the case for which $\mathcal{A}$ is the union of a finite number of disjoint compact subsets of $(0,1)^2$.

We define an empirical processes indexed by $u = (u_1, u_2)^\intercal \in \mathbb{R}^2$ as

$$\mathbb{V}_{T,k}(u) = T^{-1/2} \sum_{t=k+1}^{T} \{1[x_{1t} \leq u_1, x_{2t-k} \leq u_2] - G_k(u)\},$$

for $k \in \{1, \ldots, p\}$. The empirical processes based on the stationary bootstrap is denoted by

$$\mathbb{V}^*_{T,k}(u) = T^{-1/2} \sum_{t=k+1}^{T} \left\{1[x^*_{1t} \leq u_1, x^*_{2t-k} \leq u_2] - 1[x_{1t} \leq u_1, x_{2t-k} \leq u_2]\right\}.$$

We also consider a sequential empirical process over $(s, u) \in [\omega, 1] \times \mathbb{R}^2$ for some $\omega \in (0,1)$, given by

$$\mathbb{V}_{T,k}(s, u) = T^{-1/2} \sum_{t=k+1}^{[Ts]} \{1[x_{1t} \leq u_1, x_{2t-k} \leq u_2] - G_k(u)\}.$$

The below lemma shows the weak convergence of the empirical process defined above. We define a set of mean-zero Gaussian process $\{\mathbb{V}_{\infty,k}(u) : u \in \mathbb{R}^2\}_{k=1}^{p}$ with covariance matrix function given by

$$\Xi_{kk'}(u,u') \equiv E\left[\mathbb{V}_{\infty,k}(u)\mathbb{V}_{\infty,k'}(u')\right] = \sum_{l=-\infty}^{\infty} \operatorname{cov}\left(\varphi_l(u,k), \varphi_0(u',k')\right),$$

for $k,k' \in \{1,\ldots,p\}$ and $u,u' \in \mathbb{R}^2$, where $\varphi_t(u,k) \equiv 1[x_{1t} \le u_1, x_{2t-k} \le u_2] - G_k(u)$. We also define a set of mean-zero Gaussian process $\{\mathbb{V}_{\infty,k}(s,u) : (s,u) \in [0,1]\times\mathbb{R}^2\}_{k=1}^{p}$ with covariance matrix function given by

$$E[\mathbb{V}_{\infty,k}(s,u)\mathbb{V}_{\infty,k'}(s',u')] = \min\{s,s'\}\Xi_{kk'}(u,u')$$

for $k,k' \in \{1,\ldots,p\}$ and $(s,u),(s',u') \in [0,1]\times\mathbb{R}^2$. Notice that $\Xi_{kk'}(u,u')$ is equivalent to the $(1,1)$th element of the matrix $\Gamma_{kk'}(\alpha,\alpha')$ defined in the main text, when $(u,u') = (q_\alpha, q_{\alpha'})$.

Lemma A.1 *Under Assumption A1-A4,*

(a) $[\mathbb{V}_{T,1}(\cdot),\ldots,\mathbb{V}_{T,p}(\cdot)]^\intercal \Longrightarrow [\mathbb{V}_{\infty,1}(\cdot),\ldots,\mathbb{V}_{\infty,p}(\cdot)]^\intercal$.

(b) $[\mathbb{V}^*_{T,1}(\cdot),\ldots,\mathbb{V}^*_{T,p}(\cdot)]^\intercal \Longrightarrow^* [\mathbb{V}_{\infty,1}(\cdot),\ldots,\mathbb{V}_{\infty,p}(\cdot)]^\intercal$ *in probability.*

(c) $[\mathbb{V}_{T,1}(\cdot,\cdot),\ldots,\mathbb{V}_{T,p}(\cdot,\cdot)]^\intercal \Longrightarrow [\mathbb{V}_{\infty,1}(\cdot,\cdot),\ldots,\mathbb{V}_{\infty,p}(\cdot,\cdot)]^\intercal$.

Proof. (a) is proved by the similar argument used in Theorem 7.3 of Rio (2000). (b) is shown by a slight modification of Theorem 3.1 of Politis and Romano (1994). (c) can be shown by using the argument in Bucher (2013) with (a). ■

Also, we define a vector of the one-parameter empirical process:

$$\mathbb{W}_{T,k}(u) = \left(\mathbb{W}^{(1)}_{T,k}(u_1), \mathbb{W}^{(2)}_{T,k}(u_2)\right)^\intercal,$$

where $\mathbb{W}^{(1)}_{T,k}(u_1) = \lim_{u_2\to\infty}\mathbb{V}_{T,k}(u)$ and $\mathbb{W}^{(2)}_{T,k}(u_2) = \lim_{u_1\to\infty}\mathbb{V}_{T,k}(u)$. Likewise, the bootstrap and sequential versions are denoted as follows:

$$\mathbb{W}^*_{T,k}(u) = \left(\mathbb{W}^{(1)*}_{T,k}(u_1), \mathbb{W}^{(2)*}_{T,k}(u_2)\right)^\intercal, \quad \text{and} \quad \mathbb{W}_{T,k}(u,s) = \left(\mathbb{W}^{(1)}_{T,k}(u_1,s), \mathbb{W}^{(2)}_{T,k}(u_2,s)\right)^\intercal,$$

respectively. In the following lemma, we shall obtain an asymptotic linear approximation for the cross-quantilogram.

Lemma A.2 *Suppose that Assumption A1-A4 hold. Then, for each* $k \in \{1, \ldots, p\}$,

$$\sqrt{T}\left(\hat{\rho}_{\alpha}(k) - \rho_{\alpha}(k)\right) = \frac{\mathbb{V}_{T,k}(q_{\alpha}) + \nabla G_k(q_{\alpha})^{\intercal} \sqrt{T}\left(\hat{q}_{\alpha} - q_{\alpha}\right)}{\sqrt{\alpha_1(1-\alpha_1)\alpha_2(1-\alpha_2)}} + o_p(1),$$

uniformly in $\alpha \in \mathcal{A}$.

Proof. Let $\hat{\gamma}_{\alpha k}(q) = T^{-1}\sum_{t=k+1}^{T} \psi_{\alpha_1}(x_{1t}-q_1)\psi_{\alpha_2}(x_{2t-k}-q_2)$ and $\gamma_{\alpha k}(q) = E[\psi_{\alpha_1}(x_{1t}-q_1)\psi_{\alpha_2}(x_{2t-k}-q_2)]$ for $q = (q_1, q_2)^{\intercal} \in \mathbb{R}^2$. Using a similar argument used to show Lemma 2.1 of Arcones (1998), we can show

$$\sup_{\alpha_i \in \mathcal{A}_i} \left| T^{-1/2} \sum_{t=1}^{T} \psi_{\alpha_i}\left(x_{it} - \hat{q}_i(\alpha_i)\right) \right| = o_p(1), \tag{A.1}$$

for $i = 1, 2$. It follows that, uniformly in $\alpha \in \mathcal{A}$,

$$\hat{\gamma}_{\alpha k}(\hat{q}_{\alpha}) = T^{-1} \sum_{t=k+1}^{T} 1[x_{1t} \leq \hat{q}_1(\alpha_1), x_{2t-k} \leq \hat{q}_2(\alpha_2)] - \alpha_1\alpha_2 + o_p(T^{-1/2}),$$

and $T^{-1}\sum_{t=k+1}^{T} \psi_{\alpha_i}^2(x_{it} - \hat{q}_i(\alpha_i)) = \alpha_i(1-\alpha_i) + o_p(1)$ for $i = 1, 2$. Thus, we have

$$\sqrt{T}\left\{\hat{\rho}_{\alpha}(k) - \rho_{\alpha}(k)\right\} = \frac{\sqrt{T}\left\{\hat{\gamma}_{\alpha k}(\hat{q}_{\alpha}) - \gamma_{\alpha k}(q_{\alpha})\right\}}{\sqrt{\alpha_1(1-\alpha_1)\alpha_2(1-\alpha_2)}} + o_p(1),$$

uniformly in $\alpha \in \mathcal{A}$.

The remaining task is to obtain a uniform asymptotic approximation of $\sqrt{T}\{\hat{\gamma}_{\alpha k}(\hat{q}_{\alpha}) - \gamma_{\alpha k}(q_{\alpha})\}$. Since $\gamma_{\alpha k}(q_{\alpha}) = G_k(q_{\alpha}) - \alpha_1\alpha_2$, we have

$$\sqrt{T}\left\{\hat{\gamma}_{\alpha k}(\hat{q}_{\alpha}) - \gamma_{\alpha k}(q_{\alpha})\right\} = \mathbb{V}_{T,k}(\hat{q}_{\alpha}) + \sqrt{T}\left\{G_k(\hat{q}_{\alpha}) - G_k(q_{\alpha})\right\}.$$

Let $\mathcal{E}_M$ denote the event that $\sup_{\alpha \in \mathcal{A}} \|\hat{q}_{\alpha} - q_{\alpha}\| \leq MT^{-1/2}$ for some $M > 0$ and $\mathcal{E}_M^c$

be the complement of $\mathcal{E}_M$. Given an arbitrary constants $\epsilon > 0$ and $\delta > 0$, we have

$$P\left(\sup_{\alpha\in\mathcal{A}} |\mathbb{V}_{T,k}(\hat{q}_\alpha) - \mathbb{V}_{T,k}(q_\alpha)| > \epsilon \text{ and } \mathcal{E}_M\right) \leq P\left(\sup_{\mathcal{U}(\delta)} \left|\mathbb{V}_{T,k}(u^{(2)}) - \mathbb{V}_{T,k}(u^{(1)})\right| > \epsilon\right), \tag{A.2}$$

for a sufficiently large T, where $\mathcal{U}(\delta) = \{(u^{(2)}, u^{(1)}) \in \mathbb{R}^2 : \|u^{(2)} - u^{(1)}\| \leq \delta\}$. From Lemma A.1, $\mathbb{V}_{T,k}(\cdot)$ is stochastically equicontinuous and thus the right-hand side of (A.2) converges to zero for some δ as $T \to \infty$. An application of Lemma A.1 and A.5 shows that $\lim_{T\to\infty} P(\mathcal{E}_M^c) = 0$ for some $M > 0$, which implies

$$\lim_{T\to\infty} P\left(\sup_{\alpha\in\mathcal{A}} |\mathbb{V}_{T,k}(\hat{q}_\alpha) - \mathbb{V}_{T,k}(q_\alpha)| > \epsilon\right) = 0.$$

Because $G_k(\cdot)$ is differentiable from Assumption A4, the mean value expansion yields

$$\sqrt{T}\left\{G_k(\hat{q}_\alpha) - G_k(q_\alpha)\right\} = \nabla G_k(\bar{q}_\alpha)^\intercal \sqrt{T}\left(\hat{q}_\alpha - q_\alpha\right),$$

uniformly in $\alpha \in \mathcal{A}$, where $\bar{q}_\alpha = (\bar{q}_1(\alpha_1), \bar{q}_2(\alpha_2))^\intercal$ with $\bar{q}_i(\alpha_i)$ being between $\hat{q}_i(\alpha_i)$ and $q_i(\alpha_i)$ for $i = 1, 2$. It follows that

$$\sqrt{T}\left\{\hat{\gamma}_{\alpha k}(\hat{q}_\alpha) - \gamma_{\alpha k}(q_\alpha)\right\} = \mathbb{V}_{T,k}(q_\alpha) + \nabla G_k(q_\alpha)^\intercal \sqrt{T}\left(\hat{q}_\alpha - q_\alpha\right) + o_p(1),$$

uniformly in $\alpha \in \mathcal{A}$. ■

Proof of Theorem 1. From the uniform Bahadur representation, we have

$$\sqrt{T}\left\{\hat{q}_i(\alpha_i) - q_i(\alpha_i)\right\} = \frac{1}{f_i(q_i(\alpha_i))}\mathbb{W}_{T,k}^{(i)}\left(q_i(\alpha_i)\right) + o_p(1), \tag{A.3}$$

uniformly in $\alpha_i \in \mathcal{A}_i$, for $i = 1, 2$. We define an empirical process indexed by $\alpha \in \mathcal{A}$ as

$$\nu_{T,k}(\alpha) = \left[\mathbb{V}_{T,k}(q_\alpha), \mathbb{W}_{T,k}(q_\alpha)^\intercal\right]^\intercal.$$

Then, Lemma A.2 together with (A.3) implies that, for every $k \in \{1, \ldots, p\}$,

$$\sqrt{T}\left(\hat{\rho}_\alpha(k) - \rho_\alpha(k)\right) = \lambda_{\alpha k}^\intercal \nu_{T,k}(\alpha) + o_p(1),$$

uniformly in $\alpha \in \mathcal{A}$, where $\lambda_{\alpha k}$ is defined in (4).

We first prove

$$\sup_{\alpha\in\mathcal{A}} \left\| \hat{\rho}_\alpha^{(p)} - \rho_\alpha^{(p)} \right\| = o_p(1).$$

Using the Cauchy-Schwarz inequality, we have $|\lambda_{\alpha k}^\intercal \nu_{T,k}(\alpha)| \leq \|\lambda_{\alpha k}\| \|\nu_{T,k}(\alpha)\|$. From Assumption A2 and A4, $\|\lambda_{\alpha k}\|$ is bounded uniformly in $\alpha \in \mathcal{A}$. Furthermore, Lemma A.1 implies that $\sup_{\alpha\in\mathcal{A}} \|\nu_{T,k}(\alpha)\| = O_p(1)$, which together with the c_r-inequality leads to that $\sup_{\alpha\in\mathcal{A}} \|\hat{\rho}_\alpha(k) - \rho_\alpha(k)\| \leq \sum_{k=1}^p \sup_{\alpha\in\mathcal{A}} |\hat{\rho}_\alpha^{(p)} - \rho_\alpha^{(p)}| = o_p(1)$.

To show (3), let $\Lambda_\alpha^{(p)} = \mathrm{diag}(\lambda_{\alpha 1}^\intercal, \ldots, \lambda_{\alpha p}^\intercal)$ and $\nu_T^{(p)}(\alpha) = [\nu_{T,1}(\alpha), \ldots, \nu_{T,p}(\alpha)]^\intercal$. Then, we have

$$\sqrt{T}\left(\hat{\rho}_\alpha^{(p)} - \rho_\alpha^{(p)}\right) = \Lambda_\alpha^{(p)} \nu_T^{(p)}(\alpha) + o_p(1),$$

uniformly in $\alpha \in \mathcal{A}$. From Lemma A.1, we can establish the the finite dimensional distributions convergence of $\Lambda_\alpha^{(p)} \nu_T^{(p)}(\alpha)$ over $\alpha \in \mathcal{A}$ and it suffices to show the stochastic equicontinuity of $\Lambda_\alpha^{(p)} \nu_T^{(p)}(\alpha)$. An application of the c_r-inequality with the uniform boundedness of $\lambda_{\alpha k}$ over $\alpha \in \mathcal{A}$ yields that, for any $\alpha, \beta \in \mathcal{A}$,

$$\left\| \Lambda_\alpha^{(p)} \nu_T^{(p)}(\alpha) - \Lambda_\beta^{(p)} \nu_T^{(p)}(\beta) \right\| \leq \sum_{k=1}^p \left\{ C \left\| \nu_{Tk}(\alpha) - \nu_{Tk}(\beta) \right\| + \|\lambda_{\alpha k} - \lambda_{\beta k}\| \|\nu_{Tk}(\alpha)\| \right\}.$$

Let $\alpha, \beta \in \mathcal{A}$ with $\|\alpha - \beta\| \leq \delta$ for some $\delta > 0$. From Assumption A2, $\inf_{\alpha_i \in \mathcal{A}_i} f_i(q(\alpha_i)) \geq C_1$ for some $C_1 > 0$. It follows that $\|q_\alpha - q_\beta\| \leq \tilde{\delta} \equiv C_1^{-1}\delta$, because $\alpha_i - \beta_i = \int_{q_i(\beta_i)}^{q_i(\alpha_i)} f_i(v)dv$ for each $i = 1, 2$, by definition. It follows that

$$\begin{aligned} \sup_{\alpha,\beta\in\mathcal{A}:\|\alpha-\beta\|\leq\delta} \|\nu_{T,k}(\alpha) - \nu_{T,k}(\beta)\| \quad \leq \quad & \sup_{\mathcal{U}(\tilde{\delta})} \left| \mathbb{V}_{T,k}(u^{(2)}) - \mathbb{V}_{T,k}(u^{(1)}) \right| \\ & + \sup_{\mathcal{U}(\tilde{\delta})} \left\| \mathbb{W}_{T,k}(u^{(2)}) - \mathbb{W}_{T,k}(u^{(1)}) \right\|, \end{aligned}$$

where $\mathcal{U}(\tilde{\delta}) = \{(u^{(2)}, u^{(1)}) \in \mathbb{R}^2 : \|u^{(2)} - u^{(1)}\| \leq \tilde{\delta}\}$. Lemma A.1 implies the stochastic equicontinuity of $\mathbb{V}_{T,k}(\cdot)$ and $\mathbb{W}_{T,k}(\cdot)$, which yields that, for arbitrarily small positive constants η and ξ, there exist an $\delta > 0$ such that

$$\lim_{T\to\infty} P\left(\sup_{\alpha,\beta\in\mathcal{A}:\|\alpha-\beta\|\leq\delta} \|\nu_{T,k}(\alpha) - \nu_{T,k}(\beta)\| > \eta\right) < \xi.$$

From Lemma A.1, $\sup_{\alpha\in\mathcal{A}} \|\nu_{Tk}(\alpha)\| = O_p(1)$ and the uniform continuity of $\lambda_{\alpha k}$ under Assumption A2 and A4 implies $\sup_{\alpha,\beta\in\mathcal{A}:\|\alpha-\beta\|\leq\delta} \|\lambda_{\alpha k} - \lambda_{\beta k}\| = o(1)$ for some $\delta > 0$. Collecting the results given the finite lag order, we establish the stochastic equicontinuity, from which (3) follows. ■

Lemma A.3 *Suppose that Assumption A1-A5 hold. Then,*

$$\sqrt{T}\left(\hat{\rho}^*_\alpha(k) - \hat{\rho}_\alpha(k)\right) = \frac{\mathbb{V}^*_{T,k}(q_\alpha) + \nabla G_k(q_\alpha)^\intercal \sqrt{T}\left(\hat{q}^*_\alpha - \hat{q}_\alpha\right)}{\sqrt{\alpha_1(1-\alpha_1)\alpha_2(1-\alpha_2)}} + r^*_{3,T}(\alpha)$$

where $P^*(\sup_{\alpha\in\mathcal{A}} |r^*_{3,T}(\alpha)| > \epsilon) \to^p 0$ *for every* $\epsilon > 0$*.*

Proof. This proof is similar to the proof of Lemma A.2 and thus we outline the key steps. First, we consider $\hat{\gamma}^*_{\alpha k}(\hat{q}^*_\alpha) - \hat{\gamma}_{\alpha k}(\hat{q}_\alpha)$, where $\hat{\gamma}^*_{\alpha k}(q) = \sum_{t=k+1}^T \psi_{\alpha_1}(x^*_{1t} - q_1)\psi_{\alpha_2}(x^*_{2t-k} - q_2)$ for $q = (q_1, q_2)^\intercal \in \mathbb{R}^2$. It can be shown that, for $i = 1, 2$,

$$P^*\left(\sup_{\alpha_i\in\mathcal{A}_i}\left|T^{-1/2}\sum_{t=1}^T \psi_{\alpha_i}\left(x^*_{it} - \hat{q}^*_i(\alpha_i)\right)\right| > \epsilon\right) = o_p(1). \qquad \text{(A.4)}$$

for any $\epsilon > 0$, which together with (A.1) implies

$$\sqrt{T}\left\{\hat{\gamma}^*_{\alpha k}(\hat{q}^*_\alpha) - \hat{\gamma}_{\alpha k}(\hat{q}_\alpha)\right\} = \mathbb{V}^*_{T,k}(\hat{q}^*_\alpha) + \mathbb{V}_{T,k}(\hat{q}^*_\alpha) - \mathbb{V}_{T,k}(\hat{q}_\alpha) + G_k(\hat{q}^*_\alpha) - G_k(\hat{q}_\alpha) + r^*_{1,T}(\alpha),$$

where $P^*(\sup_{\alpha\in\mathcal{A}} |r^*_T(\alpha)| > \epsilon) = o_p(1)$ for any $\epsilon > 0$.

From Theorem 2 of Doss and Gill (1992), $P^*(\sup_{\alpha\in\mathcal{A}} \|\hat{q}^*_\alpha - \hat{q}_\alpha\| > M_1 T^{-1/2}) = o_p(1)$ for some $M_1 > 0$. Since $\sup_{\alpha\in\mathcal{A}} \|\hat{q}_\alpha - q_\alpha\| = O_P(T^{-1/2})$, we have $P^*\left(\sup_{\alpha\in\mathcal{A}} \|\hat{q}^*_\alpha - q_\alpha\| \geq M_2 T^{-1/2}\right) =$

$o_p(1)$ for some $M_2 > 0$. Lemma A.1(b) implies the stochastic equicontinuity of $\mathbb{V}^*_{T,k}(\cdot)$ conditional on the original sample, thereby yielding that, for any $\epsilon > 0$,

$$P^*\left(\sup_{\alpha\in\mathcal{A}} |\mathbb{V}^*_{Tk}(\hat{q}^*_\alpha) - \mathbb{V}^*_{Tk}(q_\alpha)| > \epsilon\right) = o_p(1).$$

Similarly, Lemma A.1 implies that, for any $\epsilon > 0$,

$$P^*\left(\sup_{\alpha\in\mathcal{A}} |\mathbb{V}_{Tk}(\hat{q}^*_\alpha) - \mathbb{V}_{Tk}(\hat{q}_\alpha)| > \epsilon\right) = o_p(1).$$

Thus, it follows from the mean-value expansion that

$$\sqrt{T}\left\{\hat{\gamma}^*_{\alpha k}(\hat{q}^*_\alpha) - \hat{\gamma}_{\alpha k}(\hat{q}_\alpha)\right\} = \mathbb{V}^*_{\alpha k}(q_\alpha) + \nabla G_k(q_\alpha)^\intercal \sqrt{T}\left(\hat{q}^*_\alpha - \hat{q}_\alpha\right) + r^*_{2,T}(\alpha),$$

where $P^*(\sup_{\alpha\in\mathcal{A}} |r^*_{2,T}(\alpha)| > \epsilon) = o_p(1)$ for any $\epsilon > 0$.

Next, we consider the denominator. It can be shown that

$$\sup_{\alpha_1\in\mathcal{A}_1}\left|T^{-1}\sum_{t=k+1}^{T}\psi^2_{\alpha_1}\left(x^*_{1t} - \hat{q}^*_1(\alpha_1)\right) - \alpha_1(1-\alpha_1)\right| \leq \sup_{\alpha_1\in\mathcal{A}_1}\left|T^{-1}\sum_{t=1}^{T}\psi_{\alpha_i}\left(x^*_{it} - \hat{q}^*_i(\alpha_i)\right)\right|$$

where the right-hand side converges to zero in probability, conditional on the original sample from (A.4). A similar argument can apply for the other term of the denominator and thus we obtain desired result. ■

Proof of Theorem 2. As in the proof of Theorem 2 of Doss and Gill (1992), we have the uniform Bahadur representation

$$\sqrt{T}\left\{\hat{q}^*_i(\alpha_i) - \hat{q}_i(\alpha_i)\right\} = \frac{1}{f_i(\hat{q}_i(\alpha_i))}\mathbb{W}^{(i)*}_{T,k}\left(\hat{q}_i(\alpha_i)\right) + r^*_{T1}(\alpha),$$

for $i = 1, 2$, where $P^*(\sup_{\alpha_i\in\mathcal{A}_i} |r^*_{T1}(\alpha_i)| > \epsilon) = o_p(1)$ for any $\epsilon > 0$. Since $\sqrt{T}\{\hat{q}_\alpha - q_\alpha\} = O_p(1)$ uniformly in $\alpha \in \mathcal{A}$, Lemma A.1(b) and A.3 imply implies that, for every $k \in \{1, \ldots, p\}$,

$$\sqrt{T}\left(\hat{\rho}^*_\alpha(k) - \hat{\rho}_\alpha(k)\right) = \lambda^\intercal_{\alpha k}\nu^*_{T,k}(\alpha) + r^*_{T2}(\alpha),$$

where $P^*(\sup_{\alpha\in\mathcal{A}}|r^*_{T2}(\alpha)| > \epsilon) = o_p(1)$ for any $\epsilon > 0$, $\lambda_{\alpha k}$ is defined in (4), and $\nu^*_{T,k}(\alpha) = [\mathbb{V}^*_{T,k}(q_\alpha), \mathbb{W}^*_{T,k}(q_\alpha)^\intercal]^\intercal$. We define $\Lambda^{(p)}_\alpha = \text{diag}(\lambda^\intercal_{\alpha 1}, \ldots, \lambda^\intercal_{\alpha p})$ and $\nu^{(p)*}_T(\alpha) = [\nu^*_{T,1}(\alpha), \ldots, \nu^*_{T,p}(\alpha)]^\intercal$. Then, we have

$$\sqrt{T}\left(\hat{\rho}^{(p)}_\alpha - \rho^{(p)}_\alpha\right) = \Lambda^{(p)}_\alpha \nu^{(p)*}_T(\alpha) + r^*_{T3}(\alpha),$$

where $P^*(\sup_{\alpha\in\mathcal{A}}|r^*_{T3}(\alpha)| > \epsilon) = o_p(1)$ for any $\epsilon > 0$. The cross-quantilogram based on the SB resample is linearly depend on the empirical process based on the SB resample. Since Lemma A.1(b) shows the validity of SB method for the empirical distribution function and the similar argument used to show Theorem 1 can apply for the rest of the proof. ■

Proof of Theorem 3. Under both fixed and local alternatives, Lemma A.2 implies

$$\sqrt{T}\left(\hat{\rho}^{(p)}_\alpha - \rho^{(p)}_\alpha\right) = \Lambda^{(p)}_\alpha \nu^{(p)}_T(\alpha) + o_p(1).$$

uniformly in $\alpha \in \mathcal{A}$, and it follows from Theorem 1 that

$$\Lambda^{(p)}_\alpha \nu^{(p)}_T(\alpha) = O_P(1),$$

uniformly in $\alpha \in \mathcal{A}$.

(a) Under the fixed alternative, there is some $\alpha \in \mathcal{A}$ such that ρ_α is some non-zero constant and then

$$\sqrt{T}\hat{\rho}^{(p)}_\alpha = \Lambda^{(p)}_\alpha \nu^{(p)}_T(\alpha) + \sqrt{T}\rho^{(p)}_\alpha + o_p(1).$$

This implies that, under the fixed alternative,

$$\sup_{\alpha\in\mathcal{A}} \hat{Q}^{(p)}_\alpha = T \sup_{\alpha\in\mathcal{A}} \|\rho^{(p)}_\alpha\|^2 \left(1 + o_p(1)\right).$$

Thus, $\sup_{\alpha\in\mathcal{A}} \hat{Q}^{(p)}_\alpha \to^p \infty$ under the fixed alternative, whereas the critical value $c^*_{Q,\tau}$ is bounded in probability from Theorem 2. Therefore, $\lim_{T\to\infty} P(\sup_{\alpha\in\mathcal{A}} \hat{Q}^{(p)}_\alpha > c^*_{Q,\tau}) =$

1. Therefore, our test is shown to be consistent under the fixed alternative.

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

$$\hat{Q}_\alpha^{(p)} = \|\Lambda_\alpha^{(p)} \nu_T^{(p)}(\alpha) + \zeta_\alpha^{(p)}\|^2 + o_P(1),$$

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

$$\sup_{\alpha\in\mathcal{A}} \hat{Q}_\alpha^{(p)} \Longrightarrow \sup_{\alpha\in\mathcal{A}} \|\Lambda_\alpha^{(p)} \mathbb{B}^{(p)}(\alpha) + \zeta_\alpha^{(p)}\|^2.$$

Also, Theorem 2 implies

$$\sup_{\alpha\in\mathcal{A}} \hat{Q}_\alpha^{(p)*} \Longrightarrow^* \sup_{\alpha\in\mathcal{A}} \|\Lambda_\alpha^{(p)} \mathbb{B}^{(p)}(\alpha)\|^2,$$

in probability. Thus, the desired result follows. ■

Lemma A.4 *Suppose that Assumption A1-A4 hold. Then, for each $k \in \{1, \dots, p\}$ and for each $\alpha \in \mathcal{A}$,*

$$\sqrt{T}\left(\hat{\rho}_{\alpha,[Tr]}(k) - \rho_\alpha(k)\right) = \frac{r^{-1}\mathbb{V}_k(q_\alpha, r) + \nabla G_k(q_\alpha)^\intercal \sqrt{T}\left(\hat{q}_{\alpha,[Tr]} - q_\alpha\right)}{\sqrt{\alpha_1(1-\alpha_1)\alpha_2(1-\alpha_2)}} + o_p(1),$$

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

Proof. The proof follows the line of Lemma A.2, using Lemma A.1(c) and Lemma A.5. We omit the details. ■

Proof of Theorem 4. Lemma A.5 shows that, for $i = 1, 2$,

$$\sqrt{T}\left\{\hat{q}_{i[Tr]}(\alpha_i) - q_i(\alpha_i)\right\} = \frac{1}{r f_i(q_i(\alpha_i))} \mathbb{W}_{T,k}^{(i)}\left(q_i(\alpha_i), r\right) + o_p(1), \qquad \text{(A.5)}$$

uniformly in $r \in [\omega, 1]$. We define an empirical process indexed by $r \in [\omega, 1]$,

$$\bar{\nu}_{T,k,\alpha}(r) = \left[\mathbb{V}_{T,k}(q_\alpha, r), \mathbb{W}_{T,k}(q_\alpha, r)^\intercal\right]^\intercal,$$

and let $\bar{\nu}_{T,\alpha}^{(p)}(\cdot) = [\bar{\nu}_{T,1,\alpha}(\cdot), \ldots, \bar{\nu}_{T,p,\alpha}(\cdot)]^\intercal$. Then, Lemma A.4 together with (A.5) implies

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

uniformly in $r \in [\omega, 1]$. It follows that

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

uniformly in $r \in [\omega, 1]$. From Lemma A.1(c), $\{\bar{\nu}_{T,\alpha}^{(p)}(r) - r\bar{\nu}_{T,\alpha}^{(p)}(1) : r \in [\omega, 1]\}$ weakly converges to the Brownian bridge process $\{\Delta_\alpha^{(p)}(\bar{\mathbb{B}}^{(p)}(r) - r\bar{\mathbb{B}}^{(p)}(1)) : r \in [\omega, 1]\}$ with $\Delta_\alpha^{(p)}(\Delta_\alpha^{(p)})' \equiv \Gamma^{(p)}(\alpha, \alpha)$, and thus it follows from the continuous mapping theorem that

$$\hat{A}_{\alpha p} \Longrightarrow \left(\Lambda_\alpha^{(p)}\Delta_\alpha^{(p)}\right)\mathbb{A}_p\left(\Lambda_\alpha^{(p)}\Delta_\alpha^{(p)}\right)^\intercal.$$

Since $\sqrt{T}\hat{\rho}_{\alpha,T}^{(p)} \Longrightarrow (\Lambda_\alpha^{(p)}\Delta_\alpha^{(p)})\bar{\mathbb{B}}^{(p)}(1)$, we have

$$\hat{S}_\alpha^{(p)} \Longrightarrow \bar{\mathbb{B}}^{(p)}(1)^\intercal\mathbb{A}_p^{-1}\bar{\mathbb{B}}^{(p)}(1).$$

We complete the proof. ■

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

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

thereby yielding $\hat{A}_{\alpha p} \Longrightarrow (\Lambda_\alpha^{(p)}\Delta_\alpha^{(p)})\mathbb{A}_p(\Lambda_\alpha^{(p)}\Delta_\alpha^{(p)})^\intercal$.

(a) Under the fixed alternative, we have

$$\sqrt{T}\hat{\rho}^{(p)}_{\alpha,T} = \Lambda^{(p)}_{\alpha}\bar{\nu}^{(p)}_{T,\alpha}(1) + \sqrt{T}\rho^{(p)}_{\alpha} + o_p(1),$$

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,

$$\sqrt{T}\hat{\rho}^{(p)}_{\alpha,T} = \Lambda^{(p)}_{\alpha}\bar{\nu}^{(p)}_{T,\alpha}(1) + \xi^{(p)}_{\alpha} + o_p(1).$$

It follows that

$$\hat{S}^{(p)}_{\alpha} \to^{d} \left\{\bar{\mathbb{B}}^{(p)}(1) + (\Lambda^{(p)}_{\alpha}\Delta^{(p)}_{\alpha})^{-1}\xi^{(p)}_{\alpha}\right\}^{\intercal} \mathbb{A}_p^{-1} \left\{\bar{\mathbb{B}}^{(p)}(1) + (\Lambda^{(p)}_{\alpha}\Delta^{(p)}_{\alpha})^{-1}\xi^{(p)}_{\alpha}\right\}.$$

We complete the proof. ■

Proof of Theorem 6. Consider (a). Two l-dimensional vectors, μ_{z1} and μ_{z2}, and a $l \times l$ matrices Σ_z and their sample analogues, $\hat{\mu}_{z1}$ and $\hat{\mu}_{z2}$, satisfy the following partitioned matrices:

$$R_{\alpha k} = \begin{bmatrix} r_{\alpha k,11} & r_{\alpha k,12} & \mu_{z1}^{\intercal} \\ r_{\alpha k,21} & r_{\alpha k,22} & \mu_{z2}^{\intercal} \\ \mu_{z1} & \mu_{z1} & \Sigma_z \end{bmatrix} \quad \text{and} \quad \hat{R}_{\alpha k} = \begin{bmatrix} \hat{r}_{\alpha k,11} & \hat{r}_{\alpha k,12} & \hat{\mu}_{z1}^{\intercal} \\ \hat{r}_{\alpha k,21} & \hat{r}_{\alpha k,22} & \hat{\mu}_{z2}^{\intercal} \\ \hat{\mu}_{z1} & \hat{\mu}_{z1} & \hat{\Sigma}_z \end{bmatrix}.$$

Then, we can write

$$p_{\alpha k,ij} = r_{\alpha k,ij} - \mu_{zi}^{\intercal}\Sigma_z^{-1}\mu_{zj} \quad \text{and} \quad \hat{p}_{\alpha k,ij} = \hat{r}_{\alpha k,ij} - \hat{\mu}_{zi}^{\intercal}\hat{\Sigma}_z^{-1}\hat{\mu}_{zj},$$

for $i, j \in \{1, 2\}$. From Theorem 1(a), we have

$$\hat{R}_{\alpha k} = R_{\alpha k} + o_p(1),$$

which leads to that $\hat{p}_{\alpha k,12} = p_{\alpha k,12} + o_p(1)$ and that

$$\sqrt{T}\left(\hat{p}_{\alpha k,12} - p_{\alpha k,12}\right) = \sqrt{T}\left(\hat{r}_{\alpha k,12} - r_{\alpha k,12}\right) - \mu_{z1}^{\intercal}\Sigma_z^{-1}\sqrt{T}\left(\hat{\mu}_{z2} - \mu_{z2}\right) + o_p(1).$$

Applying arguments used to prove Lemma A.2, we can show

$$\sqrt{T}\left(\hat{p}_{\alpha k,12} - p_{\alpha k,12}\right) = \lambda_{\alpha k,z}^{(1)\intercal} T^{-1/2} \sum_{t=k+1}^{T} \tilde{\xi}_{1t}(\alpha, k) + \lambda_{\alpha k,z}^{(2)\intercal}\sqrt{T}\begin{bmatrix} \hat{q}_\alpha - q_\alpha \\ \hat{q}_z - q_z \end{bmatrix} + o_p(1).$$

This together with the Bahadur representation yields

$$\sqrt{T}\left\{\hat{\rho}_{\alpha|z}(k) - \rho_{\alpha|z}(k)\right\} = \lambda_{\alpha k,z}^{\intercal}\tilde{\nu}_{T,kz}(\alpha) + o_p(1)$$

where $\tilde{\nu}_{T,kz}(\alpha) = T^{-1/2}\sum_{t=k+1}^{T}\tilde{\xi}_t(\alpha, k)$.

The asymptotic normality of $\nu_{T,k,z}(\alpha)$ can be established by using the central limit theorem for mixing random vectors. The proof of (b) and (c) are similar to that of Theorem 2 and Theorem 5, respectively, and thus we omit the details. ■

Technical Lemma

The following lemma shows that the quantile estimator based on subsample has the Bahadur representation uniformly in both quantiles and subsample sizes. Let $\{y_t : t \in \mathbb{Z}\}$ be a strict stationary sequence of real-valued random variables with distribution function F_y and its density f_y. Assume that $\{y_t : t \in \mathbb{Z}\}$ satisfies Assumption A1-A4. We consider quantile estimators based on subsamples, $\{\hat{q}_{y,[Tr]}(\tau) : \tau \in \mathcal{T} \subset (0,1), \omega \leq r \leq 1\}$, for the true quantile function $q_y(\cdot)$. Define $\mathbb{W}_T(v,r) = T^{-1/2}\sum_{t=1}^{[Tr]}\{1[y_t \leq v] - F_y(v)\}$ for $(v,r) \in \mathbb{R} \times [\omega, 1]$

Lemma A.5 *Under Assumptions A1-A4,*

$$\sqrt{T}\left\{\hat{q}_{y,[Tr]}(\tau) - q_y(\tau)\right\} = \frac{1}{r f_y(q_y(\tau))}\mathbb{W}_T(q_y(\tau), r) + R_T(\tau, r),$$

where $\sup_{(\tau,r)\in\mathcal{T}\times[\omega,1]} |R_T(\tau,r)| = o_p(1)$.

Proof. We define a localized objective function based on the subsample:

$$M_{[Tr]}(v,\tau) = \sum_{t=1}^{[Tr]} \left\{\pi_\tau(y_t - q_y(\tau) - T^{-1/2}v) - \pi_\tau(y_t - q_y(\tau))\right\},$$

for every fixed $v \in \mathbb{R}$ with some constant $B > 0$. Then, $\sqrt{T}\{\hat{q}_{y,[Tr]}(\tau) - q_y(\tau)\} = \arg\min_{v:|v|\leq B} M_{[Tr]}(v,\tau)$ for every $(\tau,r) \in \mathcal{T}\times[\omega,1]$. Also, define a quadratic function:

$$Q_{[Tr]}(v,\tau) = -\mathbb{W}_T(q_y(\tau),r)v + rf_y(q_y(\tau))\frac{v^2}{2}.$$

Notice that $\{rf_y(q_y(\tau))\}^{-1}\mathbb{W}_T(q_y(\tau),r) = \arg\min Q_{[Tr]}(v,\tau)$. As in Theorem 2 of Kato (2009), the desired result follows from the convexity of $M_{[Tr]}(\cdot,\cdot)$ with respect to the first argument, if we show

$$\sup_{(\tau,r)\in\mathcal{T}\times[\omega,1]} \left|M_{[Tr]}(v,\tau) - Q_{[Tr]}(v,\tau)\right| = o_P(1), \tag{A.6}$$

for every fixed $v \in \mathbb{R}$. From Knight's identity (Knight (1998)), we have

$$\begin{aligned} & \pi_\tau(y_t - q_y(\tau) - T^{-1/2}v) - \pi_\tau(y_t - q_y(\tau)) + \psi_\tau(y_t - q_y(\tau))T^{-1/2}v \\ = \; & vT^{-1/2}\int_0^1 \psi_\tau(y_t - q_y(\tau) - T^{-1/2}vs) - \psi_\tau(y_t - q_y(\tau))ds, \end{aligned}$$

which leads to

$$M_{[Tr]}(v,\tau) - Q_{[Tr]}(v,\tau) = v\int_0^1 V_{[Tr]}\left(vs,\tau\right)ds,$$

where

$$\begin{aligned} V_{[Tr]}\left(vs,\tau\right) \; = \; & \mathbb{W}_T(q_y(\tau) + T^{-1/2}vs, r) - \mathbb{W}_T(q_y(\tau), r) \\ & + rvs\int_0^1 \left\{f_y(q_y(\tau) + T^{-1/2}vsl) - f_y(q_y(\tau))\right\}dl. \end{aligned}$$

Because $s \in (0,1)$ and $|v| \leq B$, we have

$$\begin{aligned}\left|M_{[Tr]}(v,\tau) - Q_{[Tr]}(v,\tau)\right| &\leq \sup_{b:|b|\leq B}\left|\mathbb{W}_T(q_y(\tau)+T^{-1/2}b,r) - \mathbb{W}_T(q_y(\tau),r)\right| \\ &\quad + B\sup_{b:|b|\leq B}\left|f_y(q_y(\tau)+T^{-1/2}b) - f_y(q_y(\tau))\right|,\end{aligned}$$

for any $(\tau,r) \in \mathcal{T} \times [\omega,1]$. The stochastic equicontinuity of $\mathbb{W}_T(\cdot,\cdot)$ from Lemma A.1(c) yields

$$\sup_{b:|b|\leq B}\ \sup_{(\tau,r)\in\mathcal{T}\times[\omega,1]}\left|\mathbb{W}_T(q_y(\tau)+T^{-1/2}b,r) - \mathbb{W}_T(q_y(\tau),r)\right| = o_p(1).$$

Also, it follows from Assumption A3 that

$$\sup_{b:|b|\leq B}\ \sup_{\tau\in\mathcal{T}}\left|f_y(q_y(\tau)+T^{-1/2}b) - f_y(q_y(\tau))\right| = o_p(1).$$

Thus, we establish (A.6) and complete the proof. ■

B Tables and Figures

Tables B1 and B2 provide simulation results for the self-normalized statistics using the trimming parameter $\omega = 0.05$. See Section 5.2 of our paper. In Figures B1(a)-B3(b), we provide the cross-quantilogram $\hat{\rho}_{\alpha}(k)$ and the portmanteau tests $\hat{Q}_{\alpha}^{(p)}$ with self-normalized confidence intervals for no predictability. See Section 6.1 of our paper.

Table B1. (size) Empirical Rejection Frequencies of the Self-Normalized Statistics for $\omega = 0.05$
(DGP1 and the nominal level: 5%)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.173	0.038	0.033	0.031	0.027	0.028	0.026	0.034	0.196
	2	0.277	0.031	0.022	0.021	0.023	0.021	0.018	0.030	0.342
	3	0.376	0.032	0.019	0.016	0.019	0.015	0.017	0.029	0.445
	4	0.456	0.025	0.011	0.010	0.015	0.010	0.012	0.030	0.537
	5	0.514	0.029	0.008	0.008	0.007	0.007	0.008	0.034	0.610
1000	1	0.051	0.034	0.037	0.037	0.034	0.034	0.032	0.036	0.061
	2	0.093	0.027	0.029	0.026	0.027	0.027	0.028	0.027	0.102
	3	0.124	0.026	0.021	0.018	0.023	0.020	0.022	0.029	0.144
	4	0.160	0.021	0.015	0.013	0.015	0.017	0.016	0.021	0.178
	5	0.189	0.019	0.011	0.013	0.010	0.012	0.011	0.023	0.208
2000	1	0.042	0.034	0.039	0.039	0.041	0.045	0.035	0.035	0.052
	2	0.046	0.032	0.029	0.031	0.032	0.037	0.030	0.028	0.050
	3	0.047	0.026	0.021	0.031	0.030	0.029	0.025	0.025	0.057
	4	0.051	0.024	0.018	0.018	0.021	0.019	0.022	0.020	0.057
	5	0.054	0.021	0.016	0.019	0.020	0.017	0.015	0.017	0.057

Notes: The first and second columns report the sample size T and the number of lags p for the test statistics $\hat{Q}_{\alpha}^{p}$. The rest of columns show empirical rejection frequencies given simulated critical values at 5% significance level. The trimming value ω is set to be 0.05.

Table B2. (power) Empirical Rejection Frequencies of the Self-Normalized Statistics for $\omega = 0.05$
(DGP2: GARCH-X process)

		Quantiles ($\alpha_1 = \alpha_2$)								
T	p	0.05	0.10	0.20	0.30	0.50	0.70	0.80	0.90	0.95
500	1	0.134	0.261	0.266	0.110	0.031	0.124	0.262	0.258	0.103
	2	0.102	0.170	0.177	0.068	0.024	0.075	0.185	0.160	0.082
	3	0.144	0.111	0.110	0.041	0.016	0.048	0.119	0.093	0.155
	4	0.221	0.076	0.074	0.033	0.012	0.033	0.086	0.061	0.268
	5	0.310	0.057	0.053	0.023	0.008	0.021	0.061	0.047	0.371
1000	1	0.364	0.562	0.527	0.235	0.038	0.232	0.517	0.555	0.339
	2	0.236	0.457	0.421	0.168	0.025	0.178	0.423	0.455	0.219
	3	0.164	0.356	0.326	0.114	0.021	0.131	0.342	0.345	0.160
	4	0.140	0.276	0.249	0.086	0.018	0.097	0.260	0.262	0.140
	5	0.147	0.219	0.201	0.067	0.011	0.062	0.201	0.207	0.144
2000	1	0.683	0.848	0.795	0.442	0.043	0.448	0.811	0.848	0.670
	2	0.567	0.787	0.735	0.359	0.035	0.370	0.752	0.790	0.549
	3	0.463	0.711	0.673	0.287	0.025	0.298	0.673	0.710	0.447
	4	0.379	0.647	0.593	0.220	0.017	0.242	0.614	0.639	0.368
	5	0.312	0.583	0.536	0.188	0.014	0.193	0.557	0.572	0.285

Notes: Same as Table 3.

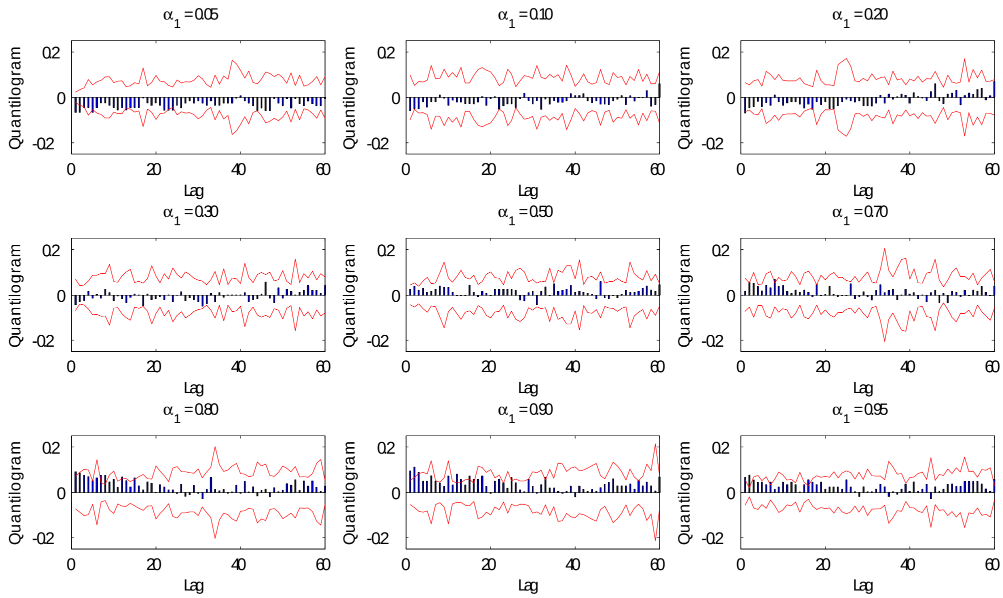

Figure B1(a). The sample cross quantilogram $\hat{\rho}_{\alpha}(k)$ for α_2= 0.1 to detect directional predictability from stock variance to stock return. Bar graphs describe sample cross-quantilograms and lines are the 95% self-normalized confidence intervals.

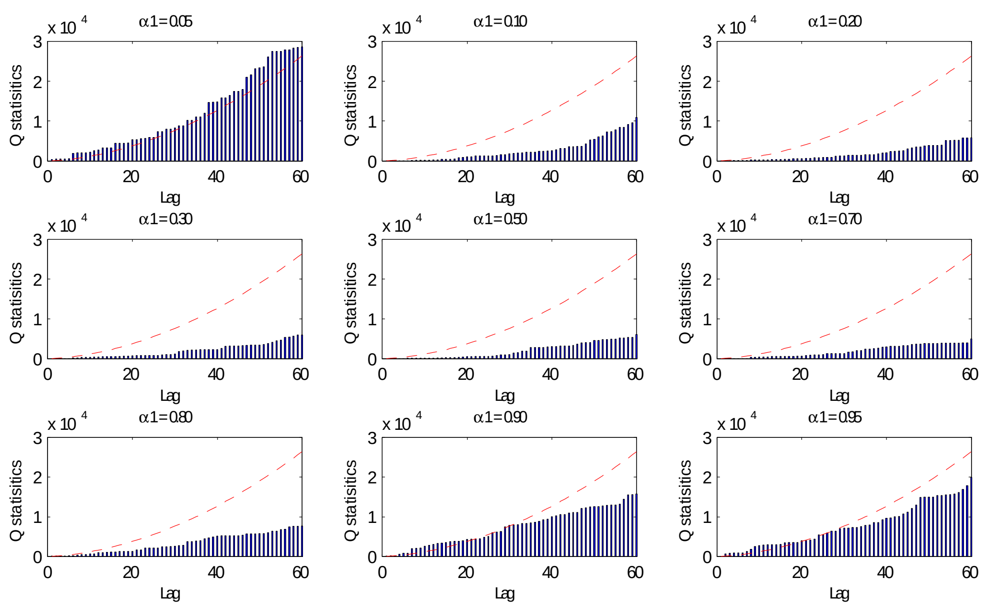

Figure B1(b). Box-Ljung test statistic $\hat{Q}_{\alpha}^{(p)}$ for each lag p and quantile α using $\hat{\rho}_{\alpha}(k)$ with α_2= 0.1. The dashed lines are the 95% self-normalized confidence intervals.

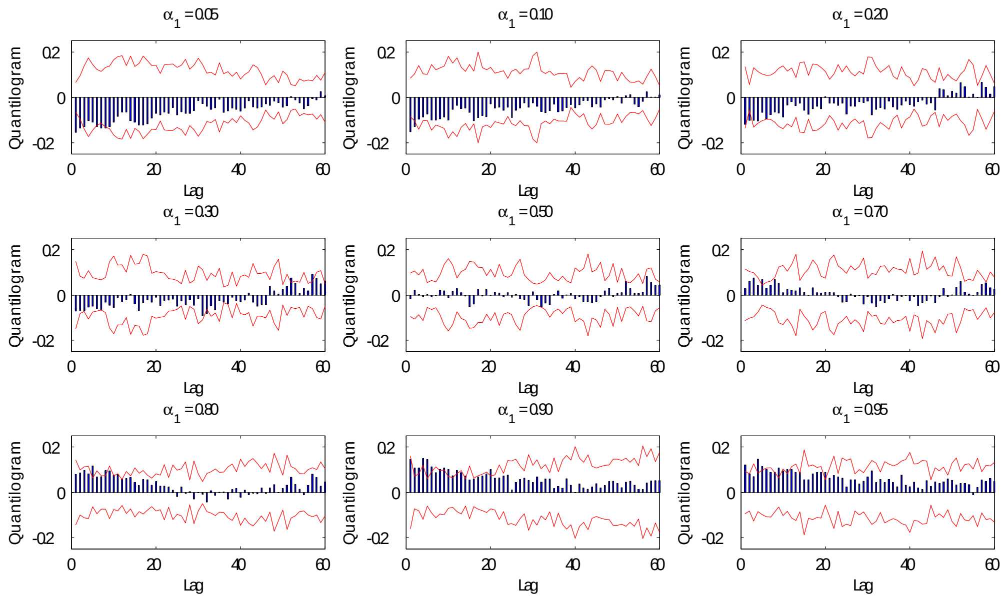

Figure 2(a). The sample cross quantilogram $\hat{\rho}_{\alpha}(k)$ for α_2= 0.5 to detect directional predictability from stock variance to stock return. Same as Figure B1(a)..

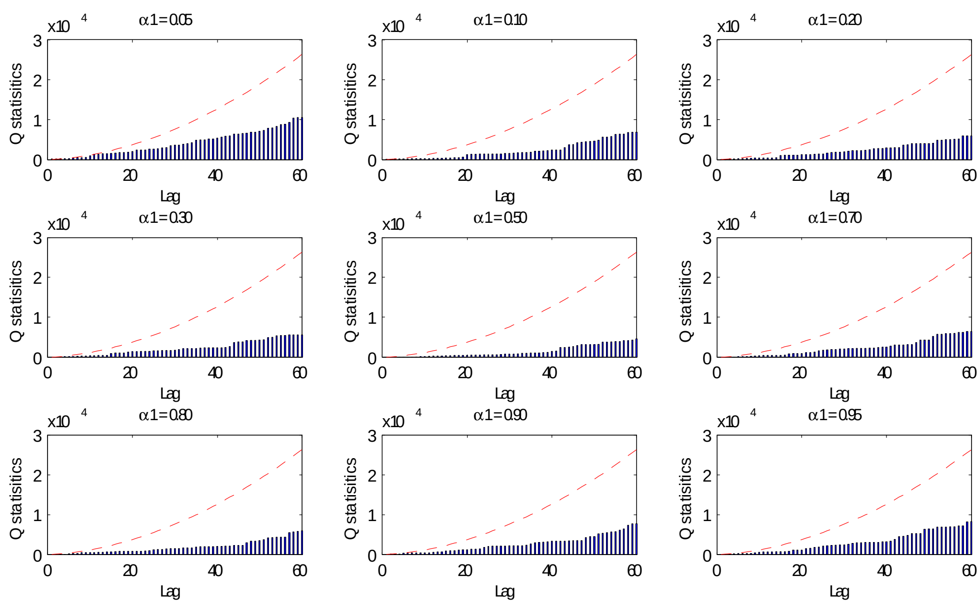

Figure 2(b). Box-Ljung test statistic $\hat{Q}_{\alpha}^{(p)}$ for each lag p and quantile α using $\hat{\rho}_{\alpha}(k)$ with α_2= 0.5. Same as Figure B1(b).

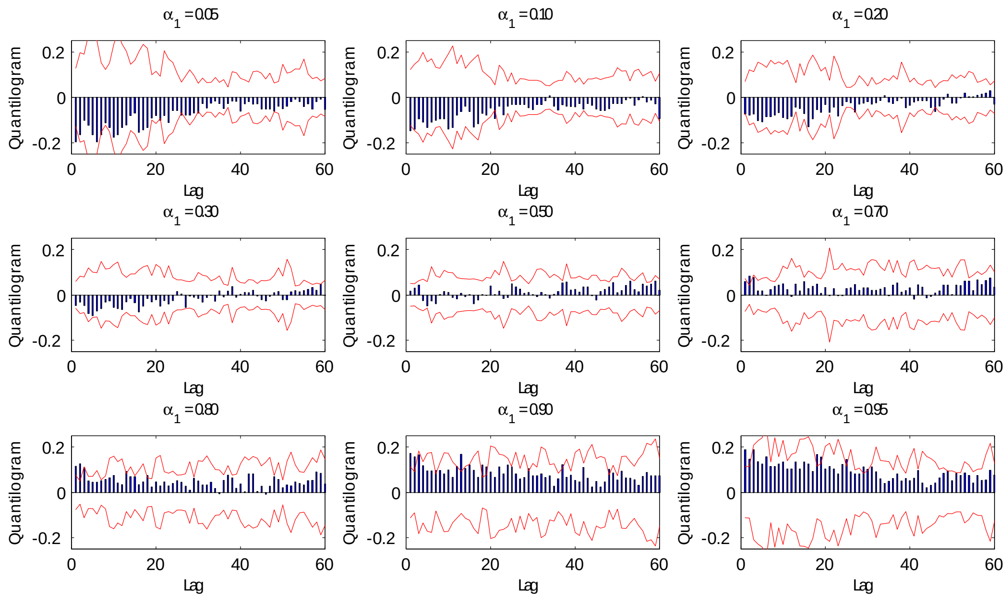

Figure 3(a). The sample cross quantilogram $\hat{\rho}_\alpha(k)$ for α_2= 0.9 to detect directional predictability from stock variance to stock return. Same as Figure B1(a).

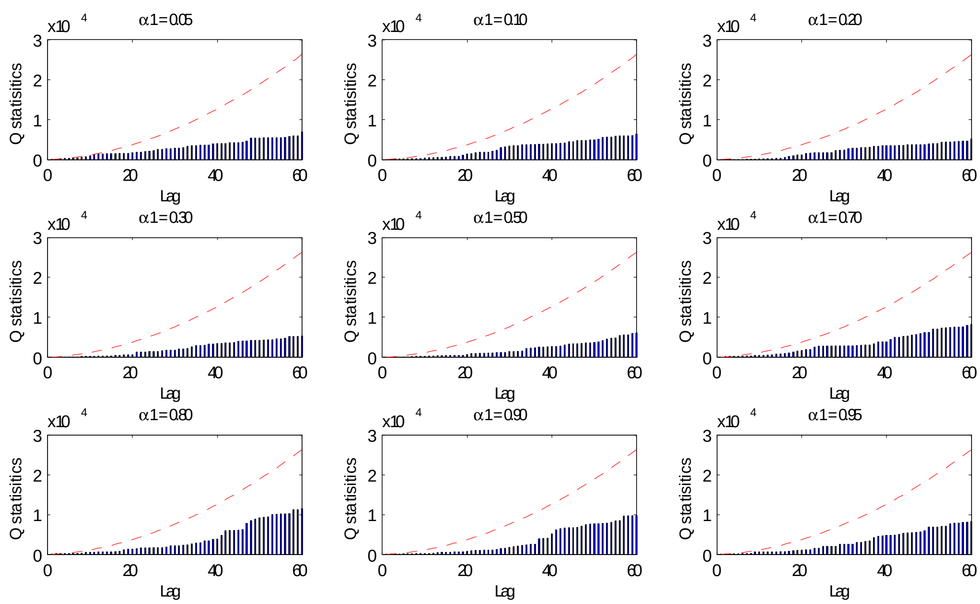

Figure 3(b). Box-Ljung test statistic $\hat{Q}_\alpha^{(p)}$ for each lag p and quantile α using $\hat{\rho}_\alpha(k)$ with α_2= 0.9. Same as Figure B1(b).

C Computer Codes

We provide the R-package, "quantilogram_0.1.tar.gz", and its manual. As explained in the manual, the package includes the datasets for two empirical applications presented in our paper and contains critical values for the self-normalized statistics. We also provide two R-codes, "Application_StockReturn.R" and Application_SystemicRisk.R for the applications, using the package.

References

Arcones, M. (1998) Second Order Representations of the Least Absolute Deviation Regression Estimator, *Annals of the Institute of Statistical Mathematics*, 50, 87-117.

Bucher, A. (2013) A note on weak convergence of the sequential multivariate empirical process under strong mixing, working paper.

Doss, H. and R.D. Gill (1992) An elementary approach to weak convergence for quantile processes, with applications to censored survival data, *Journal of the American Statistical Association*, 87, 869-877.

Kato, K. (2009) Asymptotics for argmin processes: Convexity arguments, *Journal of Multivariate Analysis*, 100, 1816-1829.

Knight, K. (1998) Limiting distributions for L_1 regression estimators under general conditions, *The Annals of Statistics*, 26, 755-770.

Politis, D.N. and J.P. Romano (1994) Limit theorems for weakly dependent Hilbert space valued random variables with application to the stationary bootstrap, *Statistica Sinica*, 4, 461-476.