

## Quantile Dependence between Stock Markets and Its Application in Volatility Forecasting\*

Heejoon Han<sup>†</sup>

**Abstract** This paper examines quantile dependence between international stock markets and evaluates its use for improving volatility forecasting. First, we adopt the cross-quantilogram, a correlation statistic of quantile hit processes, and analyze quantile dependence and directional predictability between the US stock market and stock markets in the UK, Germany, France and Japan. Second, we consider a simple quantile-augmented volatility model that accommodates the quantile dependence and directional predictability from the US market to these other markets. The quantile-augmented volatility model provides superior in-sample and out-of-sample volatility forecasts. Finally, we set up a generalized quantile-based approach to improve volatility forecasting for a wide class of asset portfolios.

**Keywords** Quantile dependence, Cross-quantilogram, Spillover, Volatility Forecasting

**JEL Classification** C14, C22, G12

---

\*I would like to thank Robert F. Engle, Soosung Hwang, Chang Sik Kim, Simone Menganeli and the seminar participants at the 9th SoFiE (Society of Financial Econometrics) annual conference (Hong Kong), Time Series Workshop on Macro and Financial Economics (Seoul), the 10th Cross-Strait Conference on Statistics and Probability (Chengdu), Korean Econometric Society Summer Meeting (Jeju), Hanyang University, University of Missouri and Sungkyunkwan University for their valuable comments and suggestions. This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2016S1A5A8019134).

<sup>†</sup>Department of Economics, Sungkyunkwan University (E-mail: heejoonhan@skku.edu).

## 1. INTRODUCTION

In many circumstances, investors are interested in dependence between financial markets such as dependence between international stock markets, dependence between currency markets, dependence between stock markets and bond markets or dependence between stock markets and commodity markets. It is essential for investors to have an understanding of the dependence between financial markets because this can be used to improve asset allocation and risk management. Therefore, volatility spillover, co-movement and contagion of financial markets have been extensively investigated in the literature. Researchers typically adopted a vector autoregressive model, a multivariate generalized autoregressive conditional heteroskedasticity (GARCH) model or a combination of both models to analyze volatility spillover, co-movement and contagion of financial markets (Baele (2005), Dungey et al. (2005), Forbes and Rigobon (2002), Karolyi (1995), King et al. (1994) and the references therein). Additionally, a copula model or a combination of a copula and an existing multivariate model has been used to investigate dependence between financial markets (Garcia and Tsafack (2011), Lee and Long (2009), and Rodriguez (2007), among others).

While these existing methods generally depend on parametric modeling of conditional variance, conditional correlation or copula of multivariate financial time series, researchers recently introduced some methods that do not require any modeling and focus directly on the quantile dependence of financial time series (Baruník and Kley (2015), Cappiello et al. (2014), Han et al. (2016), Li et al. (2015), Schmitt et al. (2015), and Sim and Zhou (2015)). These works provide various new methods to measure quantile dependence that is not captured by classical measures based on linear correlation. Some methods such as that used in Cappiello et al. (2014) test contagion or constant correlation between financial time series, which can provide useful implications for asset allocation. However, little research has explored beyond basic measurement of quantile dependence between financial time series to investigate how to directly make use of measured quantile dependence in volatility forecasting, asset allocation or risk management.

The main motivation of this paper is to address this gap. We first measure detailed quantile dependence between stock markets and examine quantile-based directional predictability between stock markets. Using the quantile-based dependence and directional predictability, we introduce and evaluate a method to improve volatility forecasting in each stock market. Finally, we generalize these results and set up a quantile-based approach to improve volatility forecasting for a wide class of asset portfolios.

We consider the daily S&P 500 index, FTSE 100 index, DAX index, CAC 40 index and Nikkei 250 index and examine quantile dependence between the US stock return and stock return series for the UK, Germany, France and Japan, i.e. quantile dependence between US-UK, US-Germany, US-France or US-Japan bivariate stock market returns. To examine detailed quantile-based relationships between stock markets, we adopt the cross-quantilogram recently proposed by Han et al. (2016). The cross-quantilogram is a correlation statistic of quantile hit processes and measures dependence between the quantile range of one time series and the quantile range of the other time series. Therefore, it can provide quantile-based dependence between two financial markets. One can set up a cross-quantilogram for specific quantile ranges of interest or for an arbitrary large lag, and it is simple to interpret these results. The cross-quantilogram was recently adopted in Baumöhl and Lyócsa (2017), Jiang et al. (2016), Shahzad et al. (2017), Todorova (2017) and Yifan (2017) to measure quantile dependence between financial time series.

The results based on the cross-quantilogram show the following. First, negative spillover (lower-quantile dependence between stock markets) is stronger than positive spillover (upper-quantile dependence between stock markets). The cross-quantilogram has higher values and remains significant for larger lags when we consider lower-quantile dependence between stock markets. Second, there exists stronger quantile dependence or directional predictability from the US stock market to the UK, Germany, France and Japan markets than the other way around. Third, when stock returns are devolatilized and standardized residuals are used, directional predictability remains significant only at the first lag in the lower or upper quantile from the US market to other markets (UK, Germany, France or Japan), but it disappears from other markets (UK, Germany, France or Japan) to the US market.

Using these findings, we consider a simple way to improve volatility forecasting. In particular, we use the cross-quantilogram results of standardized residuals and modify a volatility model to exploit the quantile-based directional predictability from the US market to markets in the UK, Germany, France and Japan. In a volatility model for stock markets in the UK, Germany, France and Japan, we introduce an additional multiplicative component that can be predicted from a lower-quantile or upper-quantile event in the US stock market. We show that the quantile-augmented volatility model provides superior in-sample and out-of-sample volatility forecasts regardless of the choice of a base volatility model or a forecasting horizon. We also find that our multiplicative model provides better volatility forecasts than the usual additive GARCH-X model even

if both models use the same information. Finally, based on these results, we set up and discuss a generalized quantile-based approach to improve volatility forecasting for a wide class of asset portfolios.

The rest of the paper is organized as follows. Section 2 explains the cross-quantilogram and related Box-Ljung type test statistic, and provides a simulation study showing how cross-quantilogram measures quantile dependence of data generated by various copulas. Section 3 provides the data description and results on quantile dependence between stock markets. It presents results of auto-quantilogram and cross-quantilogram for stock return series and standardized residuals. Section 4 presents the application of quantile dependence to volatility forecasting and Section 5 discusses a generalized quantile-based approach to improve volatility forecasting for a wide class of asset portfolios. Section 6 concludes the paper.

## 2. MEASURE OF QUANTILE DEPENDENCE

### 2.1. CROSS-QUANTILOGRAM AND ITS ADVANTAGES

Linton and Whang (2007) introduced the (auto-) quantilogram to measure dependence in different parts of the distribution of a stationary time series based on the correlogram of quantile hits. Han et al. (2016) developed a multivariate version called the cross-quantilogram. The cross-quantilogram can be used 1) to measure quantile dependence between two series, 2) to test directional predictability between two series, and 3) to test model specification. They proposed and investigated the stationary bootstrap procedure and a self-normalized approach to construct the confidence intervals of the cross-quantilogram.

As explained in Linton and Whang (2007) and Han et al. (2016), the advantages of the cross-quantilogram are as follows: 1) it is simple to interpret, 2) no moment condition is required for time series, 3) it captures the properties of a joint distribution, 4) it can consider arbitrary lags. The second advantage is particularly important when we use the cross-quantilogram to analyze financial time series. While commonly used models such as multivariate GARCH models in general assume the existence of finite fourth moments of time series, it is well known that finite fourth moments do not exist for most stock return or exchange rate return series due to heavy tails. The appeal of cross-quantilogram is its simplicity compared to the existing methods. If one uses a method based on copula as in Reboredo and Ugolini (2016), one should specify and estimate a copula for each lag  $k$  and there may be a misspecification error. For example, if one is interested in quantile dependence between  $y_{1t}$  and  $y_{2,t-k}$  for  $k = 1, 2, \dots, 22$

(one month interval for daily data), one should specify and estimate a copula for each  $k$  (total 22 times). On the contrary, cross-quantilogram can be very easily calculated for all lag  $k$  at one time. Moreover, since cross-quantilogram is a nonparametric statistic based on quantile hits, there is no misspecification error.

We let  $q_{i,t}(\tau_i)$  be either  $\tau_i$  conditional or unconditional quantile of  $y_{i,t}$ . The cross-quantilogram measures dependence between two events  $\{y_{1,t} < q_{1,t}(\tau_1)\}$  and  $\{y_{2,t-k} < q_{2,t-k}(\tau_2)\}$  for an arbitrary pair of  $\tau = (\tau_1, \tau_2)'$  and a positive integer  $k$ . In the literature,  $\{1[y_{i,t} < q_{i,t}(\cdot)]\}$  is called the quantile-hit or quantile-exceedance process for  $i = 1, 2$ , where  $1[\cdot]$  denotes the indicator function.

The cross-quantilogram is the cross-correlation of the quantile-hit processes and is defined as

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

for  $k = 0, \pm 1, \pm 2, \dots$ , where

$$\psi_{\tau_i}(y_{i,t} - q_{i,t}(\tau_i)) = 1[y_{i,t} < q_{i,t}(\tau_i)] - \tau_i.$$

Its sample counterpart is

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

where  $\hat{q}_{i,t}(\tau_i)$  is the estimate of either  $\tau_i$  conditional or unconditional quantile of  $y_{i,t}$ . As an example, Figure 1 provides a pair of events:  $\{y_{1,t} < q_{1,t}(\tau_1)\}$  for  $\tau_1 = 0.05$  and  $\{y_{2,t-k} < q_{2,t-k}(\tau_2)\}$  for  $\tau_2 = 0.5$ . Given  $y_{2,t-k}$  is located below its median, the cross-quantilogram  $\rho_\tau(k)$  is zero if the probability of  $y_{1,t}$  being located below its 0.05 quantile is the same as 0.05.

Instead of two events  $\{y_{1,t} < q_{1,t}(\tau_1)\}$  and  $\{y_{2,t-k} < q_{2,t-k}(\tau_2)\}$ , one may be interested in measuring the dependence between two events  $\{q_{1,t}(\tau_1^l) < y_{1,t} < q_{1,t}(\tau_1^h)\}$  and  $\{q_{2,t-k}(\tau_2^l) < y_{2,t-k} < q_{2,t-k}(\tau_2^h)\}$  for arbitrary quantile ranges  $[\tau_1^l, \tau_1^h]$  and  $[\tau_2^l, \tau_2^h]$ . Figure 2 provides various events  $\{q_{i,t}(\tau_i^l) < y_{i,t} < q_{i,t}(\tau_i^h)\}$  for different quantiles for  $\tau_i^l$  and  $\tau_i^h$ . To obtain the dependence of such events, one can use an alternative version of the cross-quantilogram that is defined by replacing  $\psi_{\tau_i}(y_{it} - q_{i,t}(\tau_i))$  in (1) with

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

See footnote 4 in Han et al. (2016). This alternative version could be easier to interpret and therefore we will adopt this alternative version of the cross-quantilogram in this paper.

If  $\rho_\tau(k) = 0$ , there is no dependence or directional predictability from an event  $\{q_{2,t-k}(\tau_2^l) \leq y_{2,t-k} \leq q_{2,t-k}(\tau_2^h)\}$  to an event  $\{q_{1,t}(\tau_1^l) \leq y_{1,t} \leq q_{1,t}(\tau_1^h)\}$ . If  $\rho_\tau(k) \neq 0$ , there exists quantile dependence or directional predictability between two events. If  $\rho_\tau(k) > 0$ , it is more likely for  $y_{1,t}$  to be located in the range  $[q_{1,t}(\tau_1^l), q_{1,t}(\tau_1^h)]$  when  $y_{2,t-k}$  is located in the range  $[q_{2,t-k}(\tau_2^l), q_{2,t-k}(\tau_2^h)]$ . If  $\rho_\tau(k) < 0$ , it is less likely for  $y_{1,t}$  to be located in the range  $[q_{1,t}(\tau_1^l), q_{1,t}(\tau_1^h)]$  when  $y_{2,t-k}$  is located in the range  $[q_{2,t-k}(\tau_2^l), q_{2,t-k}(\tau_2^h)]$ . The stationary bootstrap inference procedure is still valid for this alternative version, as mentioned in Han et al. (2016) and, therefore, we will use it to construct confidence bands.<sup>1</sup>

Using the cross-quantilogram, we can conduct related Portmanteau tests. Suppose that  $\tau \in \mathcal{T}$  and  $p$  are given. One may be interested in testing

$$\begin{aligned} H_0 &: \rho_\tau(1) = \cdots = \rho_\tau(p) = 0, \\ H_1 &: \rho_\tau(k) \neq 0 \text{ for some } k \in \{1, \dots, p\}. \end{aligned}$$

For this test, the Box-Pierce type test statistic  $\hat{Q}_\tau^{(p)} = T \sum_{k=1}^p \hat{\rho}_\tau^2(k)$  can be used. We will use the Box-Ljung version  $\check{Q}_\tau^{(p)} = T(T+2) \sum_{k=1}^p \hat{\rho}_\tau^2(k) / (T-k)$  in this paper because it has better finite sample performance for a large  $p$  and a small sample size. Han et al. (2016) also analyze the sup-version test statistic over a set of quantiles and the partial cross-quantilogram.

## 2.2. COPULA AND CROSS-QUANTILOGRAM

We conduct a simulation study to show how cross-quantilogram measures quantile dependence of data generated by various copulas. We consider Gaussian, Clayton, Frank, Gumbel and Student's  $t$  copulas that are commonly used in the literature. The Gaussian and Student's  $t$  copulas belong to elliptical copulas and the rest copulas (Clayton, Frank and Gumbel) are classified as Archimedian copulas. It is well known in the literature that dependence between two series can vary across a quantile even if linear correlation between two time series is constant. Figure 3 provides six scatter plots of data generated by various copulas

<sup>1</sup>We conducted a Monte Carlo simulation study for this alternative version of cross-quantilogram. As in section 5 in Han et al. (2016), we examined the finite-sample performance of the Box-Ljung test statistics based on the stationary bootstrap procedure, which showed that the test had reasonably good size and power performance in finite samples. The simulation results are available upon request.

and shows that, except for Figure 3(a), dependence between two series depends on a specific quantile range.

For all  $v_1, v_2$  in  $[0,1]$ , the Gaussian copula is defined by

$$C(v_1, v_2) = \int_{-\infty}^{\Phi^{-1}(v_1)} \int_{-\infty}^{\Phi^{-1}(v_2)} \frac{1}{2\pi\sqrt{1-\kappa^2}} \exp\left(-\frac{s^2 - 2\kappa st + t^2}{2(1-\kappa^2)}\right) ds dt$$

where  $\Phi$  is the univariate standard normal distribution function and  $\kappa$  is the linear correlation coefficient.  $\kappa = 0$  implies independence between  $v_1$  and  $v_2$ . Figure 3(a) provides a scatter plot of  $v_1$  and  $v_2$  generated by the Gaussian copula with  $\kappa = 0$ . Figure 3(b) is that with  $\kappa = 0.7$ .

The Clayton copula is defined as

$$C(v_1, v_2) = (v_1^{-\kappa} + v_2^{-\kappa} - 1)^{-1/\kappa}$$

for  $\kappa \in (0, \infty)$  and  $\kappa \rightarrow 0$  leads to independence between  $v_1$  and  $v_2$ . It is an asymmetric copula with higher probability concentrated in the lower quantile and therefore exhibits stronger dependence in the lower quantile than in the upper quantile. Figure 3(c) provides a scatter plot  $v_1$  and  $v_2$  generated by the Clayton copula with  $\kappa = 2$ .

The Frank copula is defined as

$$C(v_1, v_2) = -\frac{1}{\kappa} \ln \left( 1 + \frac{(\exp(-\kappa v_1) - 1)(\exp(-\kappa v_2) - 1)}{\exp(-\kappa) - 1} \right)$$

for  $\kappa \in (-\infty, +\infty)$  and  $\kappa \rightarrow 0$  leads to independence. Unlike the Clayton copula, it is symmetric. Figure 3(d) provides a scatter plot  $v_1$  and  $v_2$  generated by the Frank copula with  $\kappa = 7$ .

The Gumbel copula is defined by

$$C(v_1, v_2) = \exp \left[ - \left\{ (-\ln v_1)^\kappa + (-\ln v_2)^\kappa \right\}^{1/\kappa} \right]$$

for  $\kappa \in (1, \infty)$  and  $\kappa = 1$  implies independence. In contrast to the Clayton copula, it is an asymmetric copula with higher probability concentrated in the upper quantile and exhibits stronger dependence in the upper quantile than in the lower quantile. Figure 3(e) provides a scatter plot  $v_1$  and  $v_2$  generated by the Gumbel copula with  $\kappa = 2$ .

The Student's  $t$  copula is defined by

$$C(v_1, v_2) = \int_{-\infty}^{t_{\kappa_2}^{-1}(v_1)} \int_{-\infty}^{t_{\kappa_2}^{-1}(v_2)} \frac{1}{2\pi\sqrt{1-\kappa_1^2}} \left( 1 + \frac{s^2 - 2\kappa_1 st + t^2}{\kappa_2(1-\kappa_1^2)} \right)^{-\frac{\kappa_2+2}{2}} ds dt$$

where  $t_{\kappa_2}^{-1}$  is the inverse of the CDF of the standard univariate Student's  $t$  distribution with degree of freedom  $\kappa_2$  and  $\kappa_1$  is the linear correlation coefficient. Figure 3(f) provides a scatter plot  $v_1$  and  $v_2$  generated by the Student's  $t$  copula with  $\kappa_1 = 0.7$  and  $\kappa_2 = 3$ . For more details on these copulas, see Joe (1997) and Nelsen (2006).

For each copula, we generate data  $\{v_{1,t}, v_{2,t}\}$  with sample size 2,000 and calculate sample cross-quantilogram for various quantile ranges;

$$\hat{\rho}_\tau(0) = \frac{\sum_{t=k+1}^T \Psi_{[\tau_1^l, \tau_1^h]}(v_{1,t} - \hat{q}_{1,t}([\tau_1^l, \tau_1^h])) \Psi_{[\tau_2^l, \tau_2^h]}(v_{2,t} - \hat{q}_{2,t}([\tau_2^l, \tau_2^h]))}{\sqrt{\sum_{t=1}^T \Psi_{[\tau_1^l, \tau_1^h]}^2(v_{1,t} - \hat{q}_{1,t}([\tau_1^l, \tau_1^h]))} \sqrt{\sum_{t=1}^T \Psi_{[\tau_2^l, \tau_2^h]}^2(v_{2,t} - \hat{q}_{2,t}([\tau_2^l, \tau_2^h]))}}$$

where

$$\Psi_{[\tau_i^l, \tau_i^h]}(v_{i,t} - \hat{q}_{i,t}([\tau_i^l, \tau_i^h])) = 1[\hat{q}_{i,t}(\tau_i^l) < v_{i,t} < \hat{q}_{i,t}(\tau_i^h)] - (\tau_i^h - \tau_i^l).$$

We let  $[\tau_1^l, \tau_1^h] = [\tau_2^l, \tau_2^h]$  and set  $[\tau_i^l, \tau_i^h]$  to be  $[0,0.1]$ ,  $[0.1,0.2]$ ,  $\dots$ , or  $[0.9,1.0]$ . After repeating this procedure 10,000 times, we obtained average of sample cross-quantilograms for each quantile range, which are provided in Table 1.<sup>2</sup>

For Gaussian copula with  $\kappa = 0$ , two series  $v_{1,t}$  and  $v_{2,t}$  are independent. As shown in Figure 3(a), there exists no dependence for all quantile ranges. Therefore, cross-quantilogram is supposed to be zero for each quantile range, which we can confirm in the first row of Table 1. For example, when  $v_{2,t}$  is located in  $[0,0.1]$  quantile range, the percentage of  $v_{1,t}$  that is also located in the same quantile range is 10% (same as the percentage of  $[0,0.1]$  quantile range). This makes the value of cross-quantilogram be zero.

On the other hand, for the rest copulas in Figure 3, there exists dependence for certain quantile ranges and, more importantly, the degree of quantile de-

<sup>2</sup>It should be noted that the cross-quantilogram  $\rho_\tau(0)$  for  $\tau = [\tau_i^l, \tau_i^h] = [0,0.1]$  or  $[0.9, 1]$  in Table 1 does not measure the tail dependence defined in the extreme value theory. Tail dependence is a measure of the dependence between extreme events, and lower tail dependence  $\lambda^L$  and upper tail dependence  $\lambda^U$  are defined as the following limits:

$$\lambda^L = \lim_{\tau \rightarrow 0^+} P \left[ v_1 \leq F_1^{-1}(\tau) | v_2 \leq F_2^{-1}(\tau) \right]$$

$$\lambda^U = \lim_{\tau \rightarrow 1^-} P \left[ v_1 > F_1^{-1}(\tau) | v_2 > F_2^{-1}(\tau) \right]$$

where  $F_i$  is the distribution function of  $v_i$ . For dependence of such extreme quantiles, Davis and Mikosch (2009) introduced the extremogram. The Gaussian and Frank copulas impose zero tail dependence. For the Clayton copula,  $\lambda^L = 2^{-1/\kappa}$  and  $\lambda^U = 0$ . For the Gumbel copula,  $\lambda^L = 0$  and  $\lambda^U = 2 - 2^{1/\kappa}$ . For the Student's  $t$  copula,  $\lambda^L = \lambda^U = 2t_{\kappa_1+1} \left( -\sqrt{(\kappa_1+1) \frac{1-\kappa_2}{1+\kappa_2}} \right)$  as shown in Demarta and McNeil (2005, Proposition 1).

pendence is not constant across a quantile range. Table 1 shows that cross-quantilogram properly measures such a varying quantile dependence. For example, Figure 3(b) shows that the area for  $v_1 \in [0, 0.1]$  and  $v_2 \in [0, 0.1]$  is more dense than that for  $v_1 \in [0.4, 0.5]$  and  $v_2 \in [0.4, 0.5]$ . This indicates that dependence between  $v_1$  and  $v_2$  in the former area is stronger than that in the latter area. When  $v_2 \in [0, 0.1]$ , there are more  $v_1$  located in the range of  $[0, 0.1]$  and this leads to a higher value of cross-quantilogram. On the other hand, when  $v_2 \in [0.4, 0.5]$ ,  $v_1$  is more widely located, which leads to a lower value of cross-quantilogram.

Figure 3(c) exhibits the case of Clayton copula with  $\kappa = 2$ , which shows that dependence of two series in the lower-quantile is much stronger than that in mid-quantile or upper-quantile. As shown in Figure 3(c), when  $v_{2,t}$  is located in  $[0, 0.1]$  quantile range  $v_{1,t}$  is also likely to be located in the same quantile range. But such a strong dependence does not appear in mid-quantile ranges such as  $[0.4, 0.5]$  or  $[0.5, 0.6]$  quantile range. In Table 1, the cross-quantilogram in  $[0, 0.1]$  quantile range is 0.68, which is much higher than 0.17 in  $[0.9, 1]$  quantile range. Meanwhile, the cross-quantilogram is close to zero in mid-quantile ranges such as  $[0.4, 0.5]$  or  $[0.5, 0.6]$  quantile range because dependence is very weak in mid-quantile ranges.

Similarly in the rest copulas, dependence between two series varies across a quantile range and the results in Table 1 show that cross-quantilogram properly measures quantile-dependence of two series. For the case of Gumbel copula with  $\kappa = 2$  in Figure 3(e), dependence in the upper-quantile is much stronger than that in mid-quantile or lower-quantile and the values of cross-quantilogram correspond to such a varying quantile dependence. For the case of Student's  $t$  copula in Figure 3(f), dependence in both tails is much stronger and, correspondingly, the values of cross-quantilogram are high in both  $[0, 0.1]$  and  $[0.9, 1]$  quantile ranges.

### 3. QUANTILOGRAM ANALYSIS

#### 3.1. DATA AND SETUP

We investigate quantile dependence and directional predictability between the US stock market and stock markets in the UK, Germany, France and Japan, i.e. quantile dependence and directional predictability between US-UK, US-Germany, US-France and US-Japan bivariate stock market returns. We consider the daily S&P 500 index, FTSE 100 index, DAX index, CAC 40 index and Nikkei 250 index. To calculate the cross-quantilogram between the US stock return and the stock return series for the UK, Germany, France and Japan, we only

consider days  $t$  for which we have observations from both indices for each pair. The sample period and sample size for each pair of indices is given in Table 2.<sup>3</sup> We consider samples until the end of 2007 so that strict stationarity holds for the data.<sup>4</sup> We demean each stock return series by subtracting its sample mean.

We let  $\tau_i$  denote a quantile range  $[\tau_i^l, \tau_i^h]$  in this section. The quantile range of stock return  $\tau_i$  is set to be  $[0,0.05]$ ,  $[0.05,0.1]$ ,  $[0.1,0.2]$ ,  $[0.2,0.4]$ ,  $[0.4,0.6]$ ,  $[0.6,0.8]$ ,  $[0.8,0.9]$ ,  $[0.9,0.95]$  or  $[0.95,1]$ . We first let  $\tau_1 = \tau_2$  for the next two subsections and consider the case with  $\tau_1 \neq \tau_2$  later. We let lag  $k = 1, \dots, 20$ . We use the stationary bootstrapping procedure by Politis and Romano (1994) to obtain confidence intervals based on 1,000 bootstrap replicates. The tuning parameter is chosen by adapting the rule suggested by Politis and White (2004) (and later corrected in Patton et al. (2009)).

### 3.2. AUTO-QUANTILOGRAM AND CROSS-QUANTILOGRAM

We first examine the auto-quantilogram in the US stock market and the UK stock market. The results for the German, French or Japanese stock market are in general similar to those for the UK stock market and, therefore, we do not include them in the paper. Figures 4(a) and 4(b) show the auto-quantilogram and the Box-Ljung test statistic for the S&P 500 index return series. The auto-quantilogram is significantly positive at some lags for  $\tau_1=[0,0.05]$ ,  $[0.4,0.6]$  or  $[0.95,1.0]$ , which makes the Box-Ljung test statistic in Figure 4(b) significant for the same quantile range  $\tau_1$ .

Figures 5(a) and 5(b) present the auto-quantilogram and the Box-Ljung test statistic for the FTSE 100 index return series. The results of the UK stock market are in general similar to those of the US stock market. For the lower-quantile or upper-quantile ( $\tau_1=[0,0.05]$  or  $[0.95,1.0]$ ) and the mid-range ( $\tau_1=[0.4,0.6]$ ), the auto-quantilogram is significantly positive for some lags.

Next, we investigate the cross-quantilogram between the US stock market and the UK stock market. Figures 6(a) and 6(b) provide the cross-quantilogram and the Box-Ljung test statistic from the US stock market to the UK stock market, i.e.,  $y_{1,t}$  is the FTSE 100 index return and  $y_{2,t-k}$  is the S&P 500 index return. This shows that there exists directional predictability from the US market to the UK market for various quantile ranges. When we consider only the first lag,  $k = 1$ , the cross-quantilogram is significantly positive for  $\tau_1=[0,0.05]$ ,  $[0.05,0.1]$ ,  $[0.1,0.2]$ ,  $[0.9,0.95]$  or  $[0.95,1.0]$ .

<sup>3</sup>The data set is from realized library 0.1 by the Oxford-Man Institute.

<sup>4</sup>To the best of our knowledge, all existing methods to measure quantile dependence of time series require strict stationarity for valid inference.

It is not surprising to note that the quantile dependence is asymmetric. For the lower-quantile ( $\tau_1=[0,0.05]$ ), the cross-quantilogram exhibits much higher values and it is significant for larger lags. This implies that when there is a very large negative loss in the US stock market, it is more likely that there is also a very large loss in the UK stock market for quite a long time. Table 4 provides the value of  $\hat{\rho}_{\tau_1}(1)$ , the cross-quantilogram at the first lag, for the lower-quantile and the upper-quantile; it is 0.25 for the lower-quantile ( $\tau_1=[0,0.05]$ ) and 0.13 for the upper-quantile ( $\tau_1=[0.95,1.0]$ ). This implies that the negative spillover (risk spillover) is stronger than the positive spillover. Such an asymmetric dependence is in accordance with what we commonly observe in international stock markets and has been reported in the financial literature. See Ang and Chen (2002), Das and Uppal (2004), Garcia and Tsafack (2011), Longin and Solnik (2001), Poon et al. (2004) and references therein.

Figures 7(a) and 7(b) present the cross-quantilogram and the Box-Ljung test statistic from the UK stock market to the US stock market, i.e.,  $y_{1,t}$  is the S&P 500 index return and  $y_{2,t-k}$  is the FTSE 100 index return. Compared to the results in Figures 6(a) and 6(b), the dependence is much weaker. The cross-quantilogram in general has a lower value and is significant at some lags only for  $\tau_1=[0,0.05]$ ,  $[0.4,0.6]$  or  $[0.95,1.0]$ . The cross-quantilogram from the UK market to the US market exhibits similar patterns to the auto-quantilogram for the US market in Figure 4(a).

### 3.3. RESULTS OF DEVOLATIZED RETURN SERIES

The results in the previous subsection show that dependence or predictability still exists from the UK stock market to the US stock market despite it being much weaker than the case from the US market to the UK market. However, the auto-quantilogram in the US market exhibits similar patterns to the cross-quantilogram from the UK market to the US market, while it is obviously different from the cross-quantilogram from the US market to the UK market. Therefore, the quantile dependence from the UK stock return to the US stock return could be an artifact due to persistence and synchronicity in the marginal volatilities of the two stock return series. As discussed in Section 3 in Davis et al. (2013), this phenomenon is similar to the well-known issue with the cross-correlation function of linear bivariate time series. Unless one or all time series are whitened, the cross-correlation may appear to be spuriously significant (see Chapter 11 in Brockwell and Davis (1991)).

Hence, in this subsection, we devolatilize each stock return series and examine the cross-quantilogram using standardized residuals. For each return series, we

estimate the GJR-GARCH(1,1) model:

$$\begin{aligned} y_{i,t} &= \sigma_{i,t} \varepsilon_{i,t}, \\ \sigma_{i,t}^2 &= \omega + \alpha y_{i,t-1}^2 + \gamma y_{i,t-1}^2 I(y_{i,t-1} < 0) + \beta \sigma_{i,t-1}^2. \end{aligned}$$

We adopt the GJR-GARCH model to accommodate the asymmetric relationship between stock return and volatility. The innovation  $\varepsilon_{i,t}$  is assumed to be iid (0,1) and therefore the standardized residual  $\hat{\varepsilon}_{i,t} = y_{i,t}/\hat{\sigma}_{i,t}$  is presumed to be serially uncorrelated. Testing serial correlation in the standardized residual is one of the most common ways to check model specification in the literature. Table 3 reports the ‘usual’ Ljung-Box Q-statistic based on auto-correlations of  $\hat{\varepsilon}_{i,t}$  or  $\hat{\varepsilon}_{i,t}^2$ . For all stock return series, the p-values of the Ljung-Box Q-statistic for lag 10 or 20 are larger than 0.05. This shows that  $\hat{\varepsilon}_{i,t}$  or  $\hat{\varepsilon}_{i,t}^2$  are serially uncorrelated and suggests that the GJR-GARCH model is an appropriate volatility model for this return series.

Now we use the standardized residual instead of the stock return series and conduct quantilogram analysis. Figures 8(a)-9(b) provide the auto-quantilogram and the Box-Ljung test statistic using the standardized residual  $\hat{\varepsilon}_{i,t}$  for the US market or the UK market. The auto-quantilogram is insignificant in most cases for both stock markets, which is in accordance with the results of the ‘usual’ Ljung-Box Q-statistic on  $\hat{\varepsilon}_{i,t}$  or  $\hat{\varepsilon}_{i,t}^2$  in Table 3 and suggests the GJR-GARCH model is appropriate for modeling each stock return series. Poon et al. (2004) also used an asymmetric version of the GARCH(1,1) model to filter return series and showed that tail indices reduced for filtered return series.

Figures 10(a) and 10(b) present the cross-quantilogram and the Box-Ljung test statistic from the US market to the UK market using the standardized residual, i.e.  $y_{1,t}$  is the standardized residual for the FTSE 100 index return and  $y_{2,t-k}$  is that for the S&P 500 index return. The cross-quantilogram has a large positive value at the first lag for the lower-quantile ( $\tau_1=[0,0.05]$ ), while it is mostly insignificant in the rest of the cases. Even after devolatilizing the returns series, there still exists directional predictability from the US market to the UK market in the lower-quantile. Figures 11(a) and 11(b) provide the cross-quantilogram and the Box-Ljung test statistic from the UK market to the US market using the standardized residual, i.e.  $y_{1,t}$  is the standardized residual for the S&P 500 index return and  $y_{2,t-k}$  is that for the FTSE 100 index return. The cross-quantilogram is insignificant in almost all cases and consequently the Box-Ljung test statistic is insignificant in all cases.

When we devolatilize only one stock return series, the results are in general similar. For example, when  $y_{1,t}$  is the standardized residual for the FTSE 100

index return and  $y_{2,t-k}$  is the S&P 500 index return itself, predictability still exists at the first lag for the lower-quantile from the US market to the UK market. However, when  $y_{1,t}$  is the standardized residual for the S&P 500 index return and  $y_{2,t-k}$  is the FTSE 100 index return itself, no predictability exists from the UK market to the US market in all quantile ranges. It should be emphasized that the directional predictability from the US market depends on quantile ranges (i.e., exists only in the lower-quantile). In the next section, we will exploit such a quantile dependence from the US market to the UK market to improve volatility forecasting in the UK stock market.

When one or both stock return series is devolatilized, directional predictability still appears from the US market to the UK market in the lower-quantile, but disappears from the UK market to the US market in all quantile ranges. This could be due to the dominance of the US stock market. Another possibility is the difference in stock market opening times. The stock market opening times are Japan (00:00-06:00), UK/Germany/France (08:00-16:30) and US (14:30-21:00) in GMT. There are two hours of overlap between the European and US stock market opening times. One may surmise that a shock in the UK market on day  $t$  will be transmitted to the US market on the same day and, consequently, directional predictability will disappear from the UK market to the US market at the first lag. However, this does not make sense considering that the US-Japan case presented in Tables 4 and 5 shows similar results as the US-UK case despite no overlap between the US and Japan stock market opening times. We conjecture that the market dominance of the US causes large significant values of the cross-quantilogram at the first lag for tails from the US market to each stock market in Europe and Japan.

When we replace the UK stock market with the German or French stock market, the cross-quantilogram exhibits similar patterns as the US-UK case. Table 4 provides the cross-quantilograms at the first lag from the US stock market to each stock market and Table 5 presents those from each stock market to the US stock market. For example, when we consider the US-Germany case, we observe the following: 1) dependence is stronger for the case from the US market to the German market than the other way around, 2) the negative spillover is stronger than the positive spillover, 3) when the standardized residual is used, directional predictability still exists in both tails from the US market to the German market, but disappears from the German market to the US market.

There is an interesting difference in the US-Japan case. The positive spillover from the US market to the Japanese market is similar to the negative spillover, i.e.,  $\hat{\rho}_\tau(1) = 0.14$  for  $\tau_1=[0,0.05]$  and  $\hat{\rho}_\tau(1) = 0.15$  for  $\tau_1=[0.95,1.0]$  when stan-

standardized residuals are used, whereas the negative spillover is stronger from the US market to three European markets. Figures 12(a) and 12(b) present the cross-quantilogram from the US market to the Japanese market. When we use the standardized residuals from the GJR-GARCH model, Figure 12(b) shows that the cross-quantilogram is significantly positive for both tails at the first lag.

### 3.4. RESULTS OF CROSS-QUANTILE RANGES

Instead of letting  $\tau_1 = \tau_2$ , we now consider the case with  $\tau_1 \neq \tau_2$ . We let the quantile range of the US stock market  $\tau_2$  be either  $[0,0.05]$  or  $[0.95,1.0]$ . We set the quantile range of the UK stock market  $\tau_1$  to be  $[0,0.05]$ ,  $[0.05,0.1]$ ,  $[0.1,0.2]$ ,  $[0.2,0.4]$ ,  $[0.4,0.6]$ ,  $[0.6,0.8]$ ,  $[0.8,0.9]$ ,  $[0.9,0.95]$  or  $[0.95,1]$  as in previous subsections.

First, we examine dependence and directional predictability from the lower-quantile event in the US market to various quantile ranges of the UK stock market. Figure 13(a) presents the cross-quantilogram from the US market to the UK market, i.e.,  $y_{1,t}$  is the FTSE 100 index return,  $y_{2,t-k}$  is the S&P 500 index return and  $\tau_2=[0,0.05]$ . The first plot in the first row in Figure 13(a) is identical to that in Figure 6(a) where  $\tau_1 = \tau_2 = [0,0.05]$ . For mid-quantile ranges of the UK market ( $\tau_1=[0.2,0.4]$ ,  $[0.4,0.6]$  or  $[0.6,0.8]$ ), the cross-quantilogram is significantly negative for some lags. This means that it is less likely for the UK stock return to be located in mid-quantile ranges when there is a large loss in the US market at day  $t - k$ . For the upper-quantile of the UK stock market ( $\tau_1=[0.95,1]$ ), the cross-quantilogram is close to zero and insignificant at the first lag but it is mostly significantly positive from the second lag to the last lag. This could be due to the bouncing effect after a large negative shock. It is interesting to note that values of the cross-quantilogram are higher in the upper-quantile than in the lower-quantile from the second lag, while the value is very high in the lower-quantile only at the first lag.

Second, we consider the case from the upper-quantile event in the US market. Figure 13(b) presents the cross-quantilogram from the US market to the UK market, i.e.,  $y_{1,t}$  is the FTSE 100 index return,  $y_{2,t-k}$  is the S&P 500 index return and  $\tau_2=[0.95,1]$ . The last plot in the third row in Figure 13(b) is identical to that in Figure 6(a) where  $\tau_1 = \tau_2 = [0.95,1]$ . In general, the dependence is weaker than the case in Figure 13(a). On various quantile ranges of the UK stock market return, a large negative shock in the US stock market has a stronger influence than a large positive shock. For  $\tau_1=[0.9,0.95]$ , the cross-quantilogram is significantly positive at the first lag. The figure shows that, when there is a large gain in the US stock market, it is more likely for the UK stock market to have a large

or a relatively large gain on the next day.

Next, we use the standardized residuals from the GJR-GARCH model and examine the same cross-quantile range aspects. When the standardized residuals are used, the cross-quantilogram is mostly insignificant except for some quantile ranges at the first lag. Figure 14(a) considers the lower-quantile case corresponding to Figure 13(a). At the first lag, the cross-quantilogram is significantly positive for  $\tau_1=[0,0.05]$ ,  $[0.05,0.1]$  or  $[0.1,0.2]$ . Figure 14(b) presents the upper-quantile case corresponding to Figure 13(b). The cross-quantilogram is mostly close to zero and insignificant.

## 4. APPLICATION IN VOLATILITY FORECASTING

### 4.1. QUANTILE-AUGMENTED VOLATILITY MODEL

In this section, we consider a method that uses the findings in the previous section to improve volatility forecasting. The cross-quantilogram analysis in the previous section shows that there exists directional predictability from a lower or upper quantile event in  $y_{2,t-1}$ , i.e., US stock return at day  $t-1$  for  $\tau_2 = [0, 0.05]$  or  $[0.95, 1]$ , to the standardized residual  $\hat{\epsilon}_{1,t}$  in each market in the UK, Germany, France and Japan. This result suggests that we can decompose  $\hat{\epsilon}_{1,t}$  into two parts; one is a predictable component from a lower- or upper-quantile event in the US market and the other is an unpredictable component. It will be more desirable to accommodate such a predictable component from a lower or upper quantile event in the US market in modeling volatility in each stock market in the UK, Germany, France and Japan.

We decompose the standardized residual in a multiplicative way such that  $\hat{\epsilon}_{1,t} = \sqrt{\hat{f}_{1,t}} \hat{\eta}_{1,t}$  where  $f_{1,t}$  is the predictable component from a lower or upper quantile event in  $y_{2,t-1}$  and  $\eta_{1,t}$  is an unpredictable component. Using this, we consider the following volatility model for stock return series  $y_{1,t}$  of each market in the UK, Germany, France and Japan;

$$y_{1,t} = \sqrt{h_{1,t} f_{1,t}} \eta_{1,t}$$

where  $h_{1,t}$  is a base volatility model such as the GJR-GARCH model,  $f_{1,t}$  is a function of a lower or upper quantile event in  $y_{2,t-1}$  and  $\eta_{1,t}$  is *iid* (0,1). The return series in each market has three multiplicative components. The first component  $h_{1,t}$  is a function of past values of  $y_{1,t}$  and it is possible to specify it as another GARCH-type model. While we can model the second component  $f_{1,t}$  in various ways including nonparametric methods, it should be noted that the

specification of  $f_{1,t}$  must properly accommodate the cross-quantilogram results of standardized residuals in Section 3.3. In this paper, we consider the following simple specification:

$$f_{1,t}(\delta) = \delta_0 + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \leq q_2(0.05)) + \delta_2 y_{2,t-1}^2 I(y_{2,t-1} \geq q_2(0.95)) \quad (2)$$

where  $y_{2,t}$  is the return series in the US stock market and  $q_2(0.05)$  or  $q_2(0.95)$  is 0.05 or 0.95 quantile of  $y_{2,t}$ , respectively.<sup>5</sup> In this manner, the conditional variance of  $y_{1,t}$  is augmented as

$$\sigma_{1,t}^2 = h_{1,t} \times f_{1,t}.$$

If the stock index return in US is below the 5% quantile or above the 95% quantile, the volatility in each stock market will be influenced on the next day. We call this model the *quantile-augmented* volatility model (QA model). Compared to a base volatility model, the QA model accommodates additional information from the US market in the previous day (particularly a lower or upper quantile event). We expect that this augmented model will provide better volatility forecasts than a base volatility model.

Another way to accommodate the directional predictability from the US market to each stock market in volatility modeling is to adopt the following additive GARCH-X model;

$$y_{1,t} = \sqrt{h_{1,t}} \eta_{1,t} \quad (3)$$

where

$$h_{1,t} = \omega + \alpha y_{1,t-1}^2 + \gamma y_{1,t-1}^2 I(y_{t-1} < 0) + \beta h_{1,t-1} + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \leq q_2(0.05)) + \delta_2 y_{2,t-1}^2 I(y_{2,t-1} \geq q_2(0.95))$$

and  $\eta_{1,t}$  is *iid* (0,1). The additive GARCH-X model is a typical way to accommodate exogenous covariates in volatility modeling (see Han and Kristensen (2015) and references therein).

---

<sup>5</sup>Instead of (2), one can also adopt

$$f_{1,t}(\delta) = \delta_0 + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \leq q_2(0.05)),$$

which accounts for only left-tail events in the US stock market. The quantile-augmented model using this specification provides similar in-sample and out-of-sample forecasting performance as that adopting (2) for our data set described in Tables 2 and 7.

Now we discuss the estimation method of the QA model. We can rearrange the model

$$y_{1,t} = \sqrt{h_{1,t}(\theta) f_{1,t}(\delta)} \eta_{1,t} \text{ for } \eta_{1,t} \sim iid(0, 1),$$

into

$$y_{1,t}^2/h_{1,t}(\theta) = f_{1,t}(\delta) + u_{1,t}$$

where  $u_{1,t} = f_{1,t}(\delta) (\eta_{1,t}^2 - 1)$ . Here  $u_{1,t}$  is a Martingale difference sequence. The estimation procedure is as follows:

1. Estimate  $\theta$  in the base model  $h_{1,t}(\theta)$ , for example the GJR-GARCH model, using the quasi-maximum likelihood estimation (QMLE) method from

$$y_{1,t} = \sqrt{h_{1,t}(\theta)} \varepsilon_{1,t} \text{ for } \varepsilon_{1,t} \sim iid(0, 1).$$

2. Rescale the squared return and estimate  $\delta$  in the following model using the least squares method

$$y_{1,t}^2/h_{1,t}(\hat{\theta}) = f_{1,t}(\delta) + u_{1,t}.$$

3. Use the estimates from the previous steps and obtain

$$\hat{\sigma}_{1,t}^2 = h_{1,t}(\hat{\theta}) \times f_{1,t}(\hat{\delta}).$$

#### 4.2. FORECAST EVALUATION METHOD

We evaluate the within-sample and out-of-sample predictive power of the QA model. We will compare the within-sample and out-of-sample forecasts of the base model (GJR-GARCH,  $\hat{\sigma}_{t,base}^2 = \hat{h}_t$ ) and the QA model ( $\hat{\sigma}_{t,QA}^2 = \hat{h}_t \times \hat{f}_t$ ). To evaluate the volatility forecast, we adopt the following standard procedure. First, we use the realized kernel as a proxy for actual volatility. Barndorff-Nielsen et al. (2008) introduced the realized kernel and it has some robustness to market microstructure noise. The realized kernels of the return series are available in the ‘Oxford-Man Institute’s realised library’ database produced by Heber et al. (2009).

Second, we use the QLIKE loss function defined as

$$L(\hat{\sigma}_t^2, \sigma_t^2) = \frac{\sigma_t^2}{\hat{\sigma}_t^2} - \log \frac{\sigma_t^2}{\hat{\sigma}_t^2} - 1$$

where  $(\sigma_t^2)$  is the proxy for actual volatility and  $(\hat{\sigma}_t^2)$  is the within-sample or out-of-sample volatility forecast. Even if realized measures are known to be better measures, they are imperfect and noisy proxies for actual volatility. There has been research on loss functions that are robust to the use of a noisy volatility proxy (see Hansen and Lunde (2006), Patton (2010) and Patton and Sheppard (2009)). Patton (2010) shows that the QLIKE loss function is robust and, in particular, Patton and Sheppard (2009) show in their simulation study that the QLIKE loss function has the highest power.

Third, the significance of any difference in the QLIKE loss is tested via a Diebold-Marinao and West (DMW) test. See Diebold-Marinao (1995) and West (1996). A DMW statistic is computed using the difference in the losses of two models

$$\begin{aligned} d_t &= L(\hat{\sigma}_{t,base}^2, \sigma_t^2) - L(\hat{\sigma}_{t,QA}^2, \sigma_t^2) \\ DMW_T &= \frac{\sqrt{T}\bar{d}_T}{\sqrt{\widehat{avar}(\sqrt{T}\bar{d}_T)}} \end{aligned} \quad (4)$$

where  $\bar{d}_T$  is the sample mean of  $d_t$  and  $T$  is the number of forecasts. The asymptotic variance of the average is computed using a Newey-West variance estimator with the number of lags set to  $[T^{1/3}]$ . If  $DMW_T$  is positive, it means that the QA model has a smaller loss than the base model. The DMW test for equal predictability is for

$$H_0 : \mathbb{E}[d_t] = 0$$

and the asymptotic distribution of the test statistic is standard normal under the null hypothesis.

#### 4.3. FORECAST EVALUATION RESULTS

We first compare fitted values of volatility for the entire sample period. Table 6 shows the DMW test results for each series. In all cases, the DMW test statistics are positive and statistically significant at the 1% level. This shows that the QA model significantly outperforms the GJR-GARCH model.

Next we compare one-step ahead out-of-sample forecasts. We adopt the rolling window procedure with a moving window of eight years (2016 days) and produce one-step ahead out-of-sample forecasts. The forecast period and number of forecasts for each series are given in Table 7.

The second row in Table 6 shows the DMW test results for the out-of-sample forecasts. The results are similar to those for the in-sample comparison. The QA

model significantly outperforms the GJR-GARCH model. Both in-sample and out-of-sample comparison results show that a simple augmented model using quantile dependence and directional predictability from the US market can significantly improve volatility forecasting.

**Remark 1:** Instead of a lower or upper quantile event in  $y_{2,t-1}$ , one may use a lower or upper quantile event in  $\hat{\varepsilon}_{2,t-1}$  that is the standardized residual of the GJR-GARCH model for  $y_{2,t-1}$ . Accordingly, we can adjust  $f_{1,t}$  as follows:

$$f_{1,t}(\delta) = \delta_0 + \delta_1 \hat{\varepsilon}_{2,t-1}^2 I(\hat{\varepsilon}_{2,t-1} \leq q_2(0.05)) + \delta_2 \hat{\varepsilon}_{2,t-1}^2 I(\hat{\varepsilon}_{2,t-1} \geq q_2(0.95))$$

where  $q_2(0.05)$  or  $q_2(0.95)$  are 0.05 or 0.95 quantile of  $\hat{\varepsilon}_{2,t}$ , respectively. We still obtain similar results. For all cases, the QA model significantly outperforms the base model in both in-sample and out-of-sample forecasts.

**Remark 2:** We consider two different base models instead of the GJR-GARCH model and conduct the same in-sample and out-of-sample forecast evaluations. One is the GJR-GARCH model with  $t$ -distribution, in which the innovation  $\varepsilon_{1,t}$  follows the  $t$ -distribution. The other is the HEAVY model by Shephard and Sheppard (2010). Specifically, we use their HEAVY-r model:

$$\begin{aligned} y_{1,t} &= \sigma_{1,t} \varepsilon_{1,t} \\ \sigma_{1,t}^2 &= \omega + \beta \sigma_{1,t-1}^2 + \pi RM_{1,t-1} \end{aligned}$$

where  $RM_{1,t}$  is the realized volatility measure of  $y_{1,t}$  at time  $t$ . Shephard and Sheppard (2010) and Hansen et al. (2012) show that this GARCH-X type model using a realized volatility measure as the covariate performs better than the standard GARCH model. Following Shephard and Sheppard (2010), we use the realized kernel as  $RM_{1,t}$ . Tables 8 and 9 show the results of in-sample and out-of-sample forecast comparisons using the alternative base models. They show that the quantile-augmented approach still significantly improves volatility forecasting.

**Remark 3:** We investigate whether the additive GARCH-X model given in (3) provides better in-sample and out-of-sample forecasts than the GJR-GARCH model. It should be noted that the QA model uses the information of lower or upper quantile events in the US stock market, and so does the additive GARCH-X model. If one wants to exploit directional predictability from lower or upper quantile events in the US market, the additive GARCH-X model would be a typical approach to be adopted. The QA model augments the conditional variance of  $y_{1,t}$  in a multiplicative way as  $\sigma_{1,t}^2 = h_{1,t} \times f_{1,t}$ , instead of the usual additive

way as in the additive GARCH-X model, because of the cross-quantilogram results of devolatilized return series in Section 3.3. While both models use the same information, it is shown that the additive GARCH-X model is not as effective as the multiplicative approach in the QA model. Table 10 shows that the DMW test statistics between the additive GARCH-X model and the GJR-GARCH model are insignificant in both in-sample and out-of-sample cases, except for the Nikkei index.

**Remark 4:** In addition to one-step ahead ( $N = 1$ ) out-of-sample forecasts, we consider multi-step ahead ( $N = 5$  and  $N = 10$ ) forecasts. We let  $var(y_{1,T+h}|F_T)$  denote the  $h$ -step ahead (pointwise) volatility forecast, and the  $N$ -step ahead cumulative forecast is defined as

$$var(y_{1,T+1} + y_{1,T+2} + \cdots + y_{1,T+N}|F_T) = \sum_{h=1}^N var(y_{1,T+h}|F_T), \quad N = 5, 10$$

where  $T$  is the last day of the moving window of 2016 days and  $F_T$  contains all available information at time  $T$ . Table 11 shows that the QA model provides superior multi-step ahead forecasts than the base model. For all cases, the DMW test statistics are positive and significant.

**Remark 5:** In addition to the estimation window of eight years (2016 days), we try various estimation windows in the rolling window forecasting procedure. Since there are approximately 252 days in a year, we consider 4 years, 5 years and 6 years for window sizes. For each window size adopted, we produce one-step ahead out-of-sample forecasts and calculate the DMW test statistic. The results are reported in Table 12. When the estimation window size is large (1512 days or 1260 days), the QA model still significantly outperforms the GJR-GARCH model. However, for a smaller estimation window size (1008 days), the DMW test cannot reject the null hypothesis of equal predictability. When the estimation window size is 1008, there are only approximately 50 observations for quantile range  $\tau_1=[0,0.05]$  or  $[0.95,1.0]$  and this may indicate that there must be sufficient number of observations so that the QA model significantly outperforms a base model.

**Remark 6:** We apply the same quantile-augmented approach in volatility modeling of the US stock return. For the US stock return  $y_{2,t}$ , we consider

$$y_{2,t} = \sqrt{h_{2,t}f_{2,t}}\eta_{2,t}$$

where  $h_{2,t}$  is the GJR-GARCH model,  $\eta_{2,t}$  is *iid* (0,1) and

$$f_{2,t} = \delta_0 + \delta_1 y_{1,t-1}^2 I(y_{1,t-1} \leq q_1(0.05)) + \delta_2 y_{1,t-1}^2 I(y_{1,t-1} \geq q_1(0.95)).$$

$y_{1,t}$  is the stock return of one of the markets in the UK, Germany, France and Japan, and  $q_1(0.05)$  and  $q_1(0.95)$  are the 0.05 and 0.95 quantile of  $y_{1,t}$ , respectively. Since the cross-quantilogram analysis in Section 3 shows that there is no quantile dependence or directional predictability from each market (UK, Germany, France or Japan) to the US market after devolatilizing, there is no reason to expect that the QA model outperforms the base model in this case. When we compare in-sample forecasts, the QA model does not provide any significant improvement: DMW test statistics are either insignificantly positive or significantly negative. This confirms that the quantile-augmented approach should be based on the quantile dependence or directional predictability revealed in cross-quantilogram analysis.

**Remark 7:** We produce one-step ahead out-of-sample forecasts for the period from January 2008 to December 2009 and conduct forecast evaluation for the crisis period. The QA model with

$$f_{1,t}(\delta) = \delta_0 + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \leq q_2(0.05))^6$$

still provides smaller forecast losses than the GJR-GARCH model for FTSE, DAX and CAC but the null hypothesis of equal predictability is not rejected at the 5% significance level for each case. The cross-quantilogram analysis for the period from January 2007 to December 2010<sup>7</sup> shows that there may not exist directional predictability in the lower-quantile from the US stock market to the European stock markets for the period. This could explain why the DMW test statistics are insignificant for the crisis period.

## 5. GENERALIZED APPROACH FOR ASSET PORTFOLIOS

The results in Sections 3 and 4 can lead us to set up a generalized quantile-based approach to improve volatility forecasting for a wide class of asset portfolios. Suppose that  $y_{1,t}$  is the log return series of a given asset portfolio and can be modeled as

$$y_{1,t} = \mu_{1,t}(\mathbf{z}_{t-1}) + \sqrt{h_{1,t}(\mathbf{z}_{t-1})} \varepsilon_{1,t} \text{ for } \mathbf{z}_{t-1} \in \mathcal{F}_{t-1},$$

where  $\mathcal{F}_t$  is the filtration containing all available information available up to time  $t$  and  $\mu_{1,t}(\mathbf{z}_{t-1})$  and  $h_{1,t}(\mathbf{z}_{t-1})$  are the conditional mean and conditional variance of  $y_{1,t}$ , respectively. One can adopt the following steps to improve volatility forecasting of  $y_{1,t}$ .

<sup>7</sup>It should be noted that inference of the cross-quantilogram could be invalid for this period because strict stationarity may not hold. The results are available upon request.

1-step Model the conditional mean and conditional variance of  $y_{1,t}$  and obtain the standardized residual  $\hat{\varepsilon}_{1,t}$ ;

$$\hat{\varepsilon}_{1,t} = \frac{y_{1,t} - \mu_{1,t}(\mathbf{z}_{t-1}; \hat{\theta})}{\sqrt{h_{1,t}(\mathbf{z}_{t-1}; \hat{\theta})}}$$

where  $\hat{\theta}$  is the vector of estimated parameters for the models for the conditional mean and conditional variance of  $y_{1,t}$ .

2-step: Using the cross-quantilogram, find an economic/financial variable  $y_{2,t}$  such that there exists directional predictability from a certain quantile range  $[\tau_2^l, \tau_2^h]$  of  $y_{2,t-k}$  to the standardized residual  $\hat{\varepsilon}_{1,t}$  for some  $k > 0$ .

3-step: Based on the result in the previous step, specify  $f_{1,t}$  and modify the volatility of  $y_{1,t}$  as  $h_{1,t}(\mathbf{z}_{t-1}; \hat{\theta}) \times f_{1,t}$ . It should be noted that the specification of  $f_{1,t}$  must properly accommodate the result of the cross-quantilogram analysis in the previous step, particularly specific quantile range  $[\tau_2^l, \tau_2^h]$  and lag  $k$  of  $y_{2,t-k}$ .

In the first step, the conditional mean  $\mu_{1,t}(\mathbf{z}_{t-1})$  can be modeled as an AR model and the conditional variance  $h_{1,t}(\mathbf{z}_{t-1})$  can be specified as a GARCH-type model. One can use standard specification test procedures to specify them. See Patton (2013, Section 1.1) for an empirical illustration. In the second step, it should be noted that the cross-quantilogram analysis is conducted for a pair of  $\{\hat{\varepsilon}_{1,t}, y_{2,t}\}$  not for a pair of  $\{y_{1,t}, y_{2,t}\}$ . As we discussed in Sections 3.3, the cross-quantilogram between  $y_{1,t}$  and  $y_{2,t}$  could be spuriously significant due to persistence and synchronicity in the marginal volatilities of two return series. Therefore, one or all time series should be devolatilized, which means that either a pair of  $\{\hat{\varepsilon}_{1,t}, y_{2,t}\}$  or  $\{\hat{\varepsilon}_{1,t}, \hat{\varepsilon}_{2,t}\}$  should be considered for cross-quantilogram. Here  $\hat{\varepsilon}_{2,t}$  is the standardized residual of  $y_{2,t}$ . It is shown that the cross-quantilogram between  $\hat{\varepsilon}_{1,t}$  and  $y_{2,t}$  is generally similar to that between  $\hat{\varepsilon}_{1,t}$  and  $\hat{\varepsilon}_{2,t}$ .

In the third step, the specific form of  $f_{1,t}$  should depend on the result of the cross-quantilogram analysis in the previous step. If the cross-quantilogram analysis shows that there exists directional predictability only from  $[0.9, 1]$  quantile range of  $y_{2,t-1}$  to both tails of  $\hat{\varepsilon}_{1,t}$ , the specific definition of  $f_{1,t}$  should be able to accommodate such a result and we may model  $f_{1,t}$  as

$$f_{1,t}(\delta) = \delta_0 + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \geq q_2(0.9)).$$

For example, we may consider a case where  $y_{2,t}$  is an index presenting uncertainty in the economy or financial market and  $y_{1,t}$  is a portfolio of various stocks.

Suppose that only a rapid increase of the uncertainty index affects volatility of the portfolio in the next day while its rapid decrease does not. In this case, there exists directional predictability only from a high quantile range of  $y_{2,t-1}$  such as  $[0.95, 1]$  or  $[0.9, 1]$  and a lower-quantile event of  $y_{2,t-1}$  should not be included in the specification of  $f_{1,t}$ .

It is also possible that directional predictability exists for more than one day. If cross-quantilogram results indicate that there exists directional predictability not only from  $y_{2,t-1}$  but also from  $y_{2,t-2}$  for  $[0, 0.5]$  quantile range, it is desirable to accommodate the directional predictability from  $y_{2,t-2}$  as well and we may model  $f_{1,t}$  as

$$f_{1,t}(\delta) = \delta_0 + \delta_1 y_{2,t-1}^2 I(y_{2,t-1} \leq q_2(0.05)) + \delta_2 y_{2,t-2}^2 I(y_{2,t-2} \leq q_2(0.05)).$$

While these examples adopt quite simple specifications of  $f_{1,t}$ , one can construct a more sophisticated specification of  $f_{1,t}$  as long as it properly accommodate a specific quantile range  $[\tau_2^l, \tau_2^h]$  and lag  $k$  of  $y_{2,t-k}$  that are revealed in the cross-quantilogram analysis.

To obtain information for constructing  $f_{1,t}$ , the cross-quantilogram is not the only available method and one may choose another method to measure dependence between  $y_{2,t-k}$  and  $\hat{\epsilon}_{1,t}$ . However, as mentioned in Section 2.1, the cross-quantilogram is much simpler and more convenient for practitioners to use than other existing methods. One may model the dependence between  $y_{2,t-k}$  and  $\hat{\epsilon}_{1,t}$  through a copula. However, using a copula method is more cumbersome in the second step. One should conduct estimation of various copulas between  $\hat{\epsilon}_{2,t-k}$  (or  $y_{2,t-k}$ ) and  $\hat{\epsilon}_{1,t}$  for each  $k$  for  $k = 1, 2, \dots, 20$ . There may be inevitably a misspecification error. Moreover, an estimated and selected copula does not provide specific quantile range  $[\tau_2^l, \tau_2^h]$  of  $y_{2,t-k}$ . On the contrary, the cross-quantilogram can be easily calculated for all lag  $k$  at one time and provides a specific quantile range of  $y_{2,t-k}$  with directional predictability.

## 6. CONCLUSION

The paper examines quantile dependence and directional predictability between international stock markets and investigates how to apply these measures in volatility forecasting. We consider dependence between the US stock return and stock return series in the UK, Germany, France and Japan, i.e., quantile dependence between US-UK, US-Germany, US-France and US-Japan bivariate stock market returns. The results based on the cross-quantilogram show that the negative spillover is in general much stronger than the positive spillover.

We apply the cross-quantilogram on standardized residuals as well as stock return series. There exists directional predictability from the US stock market to markets in the UK, Germany, France and Japan. In particular, lower or upper quantile events in the US stock market influence these stock markets. However, when standardized residuals are used, there is no directional predictability from markets in the UK, Germany, France and Japan to the US market. Using these results on quantile dependence and directional predictability, we consider a simple method to improve volatility forecasting in stock markets in the UK, Germany, France and Japan. The quantile-augmented volatility model significantly improves both in-sample and out-of-sample volatility forecasting, which is robust to the choice of a base volatility model. These results lead us to set up a generalized quantile-based approach to improve volatility forecasting for a wide class of asset portfolios. Recently, researchers have developed various methods to measure quantile dependence between financial time series. While little research has explored beyond basic measurement of quantile dependence between financial time series, this paper considers a simple method to make use of quantile dependence in order to improve volatility forecasting. The quantile-based approach discussed in this paper could be applied to a wide class of asset portfolios to improve volatility forecasting. The information provided on detailed quantile dependence can be used for various purposes, such as modeling univariate or multivariate volatility and estimating value at risk. It will be also possible to directly improve value at risk prediction by using results of cross-quantilogram analysis. We leave it as future work to develop more sophisticated methods that apply quantile dependence in asset allocation and risk management.

## A. TABLES AND FIGURES

Table 1. Average of sample cross-quantilogram  $\hat{\rho}_\tau(0)$ 

| Copula      | Parameter value(s) | Quantile ranges |           |           |           |           |           |           |           |           | Linear correlation |         |      |
|-------------|--------------------|-----------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|--------------------|---------|------|
|             |                    | [0,0.1]         | [0.1,0.2] | [0.2,0.3] | [0.3,0.4] | [0.4,0.5] | [0.5,0.6] | [0.6,0.7] | [0.7,0.8] | [0.8,0.9] |                    | [0.9,1] |      |
| Gaussian    | 0                  | 0.00            | 0.00      | 0.00      | 0.00      | 0.00      | 0.00      | 0.00      | 0.00      | 0.00      | 0.00               | 0.00    | 0.00 |
| Gaussian    | 0.7                | 0.41            | 0.13      | 0.08      | 0.05      | 0.04      | 0.04      | 0.05      | 0.08      | 0.13      | 0.41               | 0.68    | 0.68 |
| Clayton     | 2                  | 0.68            | 0.27      | 0.14      | 0.08      | 0.06      | 0.05      | 0.05      | 0.07      | 0.10      | 0.17               | 0.68    | 0.68 |
| Frank       | 7                  | 0.35            | 0.17      | 0.12      | 0.10      | 0.09      | 0.09      | 0.10      | 0.12      | 0.17      | 0.35               | 0.76    | 0.76 |
| Gumbel      | 2                  | 0.32            | 0.12      | 0.08      | 0.06      | 0.06      | 0.06      | 0.08      | 0.11      | 0.20      | 0.57               | 0.68    | 0.68 |
| Student's t | 0.7, 3             | 0.48            | 0.18      | 0.11      | 0.08      | 0.07      | 0.07      | 0.08      | 0.11      | 0.18      | 0.48               | 0.66    | 0.66 |

Note: For each copula, we generate data  $\{v_{1,t}, v_{2,t}\}$  with sample size 2,000 and calculate sample cross-quantilogram  $\hat{\rho}_\tau(0)$  for various quantile ranges from [0,0.1] to [0.9,1]. After repeating this procedure 10,000 times for a given copula, we obtain average of sample cross-quantilograms for each quantile range. Linear correlation is average of sample correlation for each case.

Table 2. Sample period and sample size for each pair of stock return series

| Pair of indices  | Sample period (sample size)        |
|------------------|------------------------------------|
| FTSE - S&P 500   | 21 Oct. 1997 - 31 Dec. 2007 (2470) |
| DAX - S&P 500    | 3 Jan. 1996 - 28 Dec. 2007 (2907)  |
| CAC - S&P 500    | 3 Jan. 1996 - 31 Dec. 2007 (2908)  |
| Nikkei - S&P 500 | 8 Jan. 1996 - 27 Dec. 2007 (2763)  |

Table 3. Results of the 'usual' Ljung-Box Q-statistic

|                         |                  | S&P  | FTSE | DAX  | CAC  | Nikkei |
|-------------------------|------------------|------|------|------|------|--------|
| $\hat{\varepsilon}_t$   | p-value of Q(10) | 0.19 | 0.67 | 0.99 | 0.11 | 0.63   |
|                         | p-value of Q(20) | 0.17 | 0.55 | 0.86 | 0.39 | 0.80   |
| $\hat{\varepsilon}_t^2$ | p-value of Q(10) | 0.81 | 0.59 | 0.08 | 0.31 | 0.39   |
|                         | p-value of Q(20) | 0.70 | 0.28 | 0.19 | 0.25 | 0.35   |

Note: The table reports the Ljung-Box Q-statistic on  $\hat{\varepsilon}_t$  or  $\hat{\varepsilon}_t^2$ , where  $\hat{\varepsilon}_t$  is the standardized residual from the GJR-GARCH model.

Table 4. Cross-quantilograms at the first lag from the US market to other markets

|               | $\tau_1 (= \tau_2)$ | FTSE  | DAX   | CAC   | Nikkei |
|---------------|---------------------|-------|-------|-------|--------|
| Return        | [0, 0.05]           | 0.25* | 0.17* | 0.20* | 0.18*  |
|               | [0.95, 1]           | 0.13* | 0.12* | 0.11* | 0.21*  |
| Std. residual | [0, 0.05]           | 0.16* | 0.14* | 0.14* | 0.14*  |
|               | [0.95, 1]           | 0.04  | 0.06* | 0.03  | 0.15*  |

Note: The table reports  $\hat{\rho}_\tau(1)$ , a sample cross-quantilogram at the first lag, from the US stock market to other stock markets, i.e.,  $y_{1,t}$  is the return series of FTSE, DAX, CAC or Nikkei and  $y_{2,t-1}$  is the S&P 500 index return. The second and third rows are the cases where stock return series are used. The fourth and fifth rows are the cases where standardized residuals from the GJR-GARCH model are used. \* indicates significance at the 5% level.

Table 5. Cross-quantilograms at the first lag from other markets to the US market

|               | $\tau_1 (= \tau_2)$ | FTSE  | DAX   | CAC   | Nikkei |
|---------------|---------------------|-------|-------|-------|--------|
| Return        | [0, 0.05]           | 0.06* | 0.08* | 0.06* | 0.05*  |
|               | [0.95, 1]           | 0.03  | 0.06* | 0.06  | -0.01  |
| Std. residual | [0, 0.05]           | 0.02  | 0.01  | -0.00 | 0.02   |
|               | [0.95, 1]           | -0.02 | -0.01 | 0.01  | -0.04* |

Note: The table reports  $\hat{\rho}_\tau(1)$ , a sample cross-quantilogram at the first lag, from each stock market to the US stock market, i.e.,  $y_{1,t}$  is the S&P 500 index return and  $y_{2,t-1}$  is the return series of FTSE, DAX, CAC or Nikkei. Same as Table 4.

Table 6. DMW test results against GJR-GARCH model

|               | FTSE    | DAX      | CAC     | Nikkei   |
|---------------|---------|----------|---------|----------|
| In-sample     | 3.29*** | 3.13***  | 4.03*** | 3.59***  |
| Out-of-sample | 7.67*** | 10.88*** | 12.90** | 10.01*** |

Note: The table reports the DMW statistics given in (4). The base model is the GJR-GARCH model. A positive DMW statistic means that the quantile-augmented model has a smaller forecast loss. \*, \*\* and \*\*\* indicate that the null hypothesis of equal predictability between the base model and the quantile-augmented model is rejected at the 10%, 5% and 1% significance level, respectively.

Table 7. Out-of-sample forecast period and number of forecasts

| Index  | Forecast period (number of forecasts)       |
|--------|---|
| FTSE   | 2 Mar. 2006 - 31 Dec. 2007 (454 forecasts)  |
| DAX    | 1 June 2004 - 31 Dec. 2007 (891 forecasts)  |
| CAC    | 2 June 2004 - 31 Dec. 2007 (892 forecasts)  |
| Nikkei | 29 Oct. 2004 - 31 Dec. 2007 (746 forecasts) |

Note: The table reports the out-of-sample forecast period and number of forecasts for each return series. For each return series, one-step ahead out-of-sample forecasts are produced via the rolling window procedure with a moving window of eight years (2016 days).

Table 8. DMW test results against GJR-GARCH model with  $t$ -distribution

|               | FTSE    | DAX     | CAC      | Nikkei  |
|---------------|---------|---------|----------|---------|
| In-sample     | 2.96*** | 2.85*** | 3.96***  | 2.91*** |
| Out-of-sample | 7.28*** | 8.71*** | 12.56*** | 8.98*** |

Note: The base model is the GJR-GARCH model with  $t$ -distribution. Same as Table 6.

Table 9. DMW test results against HEAVY-r model

|               | FTSE    | DAX     | CAC      | Nikkei  |
|---------------|---------|---------|----------|---------|
| In-sample     | 3.78*** | 3.20*** | 4.86***  | 2.89*** |
| Out-of-sample | 7.22*** | 8.19*** | 10.84*** | 8.43*** |

Note: The base model is the HEAVY-r model by Shephard and Sheppard (2010). Same as Table 6.

Table 10. DMW test results between additive GARCH-X model and GJR-GARCH model

|               | FTSE | DAX   | CAC  | Nikkei  |
|---------------|------|-------|------|---------|
| In-sample     | 1.52 | 0.62  | 1.28 | 2.30**  |
| Out-of-sample | 1.17 | -0.79 | 0.63 | 6.60*** |

Note: The table reports the DMW statistics given in (4) where  $d_t = L(\hat{\sigma}_{t,base}^2, \sigma_t^2) - L(\hat{\sigma}_{t,Additive}^2, \sigma_t^2)$ . The base model is the GJR-GARCH model and  $\hat{\sigma}_{t,Additive}^2$  is in-sample or out-of-sample forecast of the additive GARCH-X model described in (3). Same as Table 6.

Table 11. DMW test results for multi-step ahead out-of-sample forecasting

|                          | FTSE    | DAX     | CAC      | Nikkei   |
|--------------------------|---------|---------|----------|----------|
| 5-step ahead pointwise   | 2.04**  | 7.02*** | 8.35***  | 10.82*** |
| 5-step ahead cumulative  | 3.96*** | 8.80*** | 10.73*** | 9.04***  |
| 10-step ahead pointwise  | 1.67*   | 7.23*** | 8.35***  | 8.58***  |
| 10-step ahead cumulative | 2.74*** | 8.03*** | 9.65***  | 9.64***  |

Note: The base model is the GJR-GARCH model. Same as Table 6.

Table 12. DMW test results for various estimation windows

| Window size    | FTSE     | DAX     | CAC     | Nikkei  |
|----------------|----------|---------|---------|---------|
| 1008 (4 years) | 0.25     | -1.40   | 0.76    | 0.97    |
| 1260 (5 years) | 9.71***  | 1.40    | 2.83*** | 4.50*** |
| 1512 (6 years) | 11.20*** | 3.17*** | 3.25*** | 9.19*** |

Note: The base model is the GJR-GARCH model. For each window size adopted, we produce one-step ahead out-of-sample forecasts and calculate the DMW test statistic. Same as Table 6.

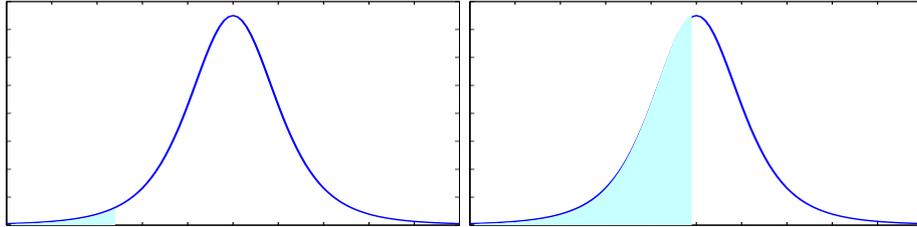


Figure 1. Event  $\{y_{i,t} < q_{i,t}(\tau_i)\}$ . The left figure describes an event  $\{y_{1,t} < q_{1,t}(\tau_1)\}$  for  $\tau_1 = 0.05$  and the right figure provides an event  $\{y_{2,t-k} < q_{2,t-k}(\tau_2)\}$  for  $\tau_2 = 0.5$ .

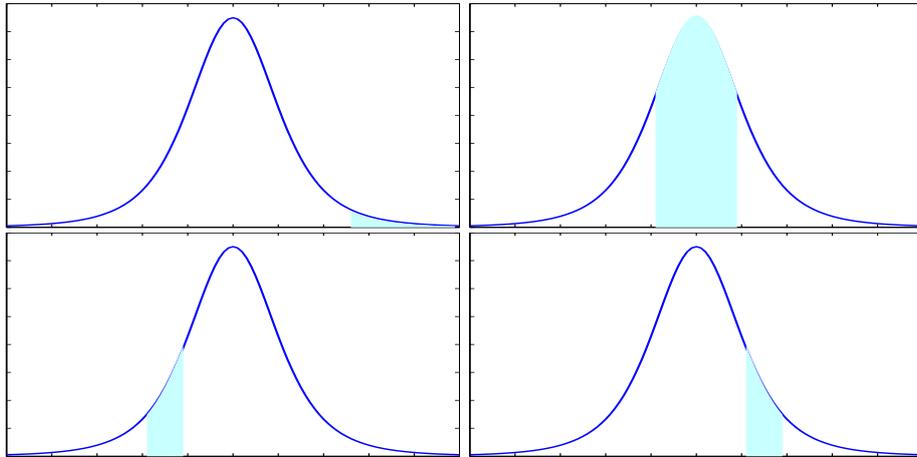
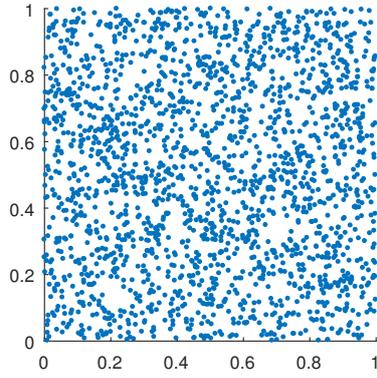
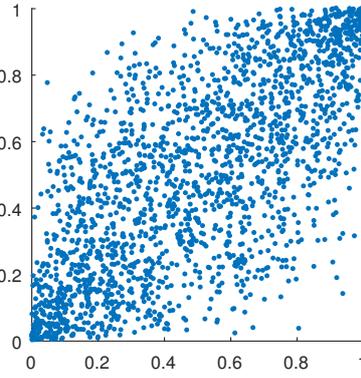


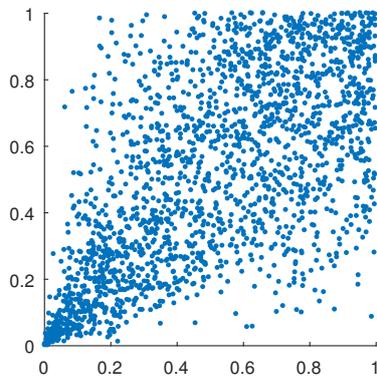
Figure 2. Event  $\{q_{i,t}(\tau_i^l) < y_{i,t} < q_{i,t}(\tau_i^h)\}$ . The figures describe various events  $\{q_{i,t}(\tau_i^l) < y_{i,t} < q_{i,t}(\tau_i^h)\}$  for different quantiles for  $\tau_i^l$  and  $\tau_i^h$ . The top left figure provides an upper-quantile event and the top right figure gives a mid-range event. The bottom figures present events for the left and right shoulders of the distribution.



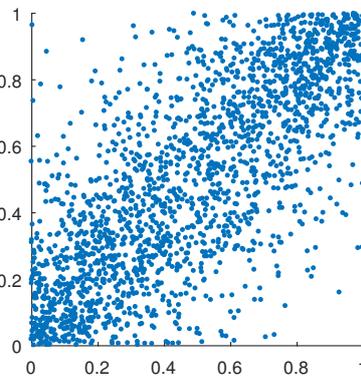
(a) Gaussian ( $\kappa = 0$ )



(b) Gaussian ( $\kappa = 0.7$ )



(c) Clayton ( $\kappa = 2$ )



(d) Frank ( $\kappa = 7$ )

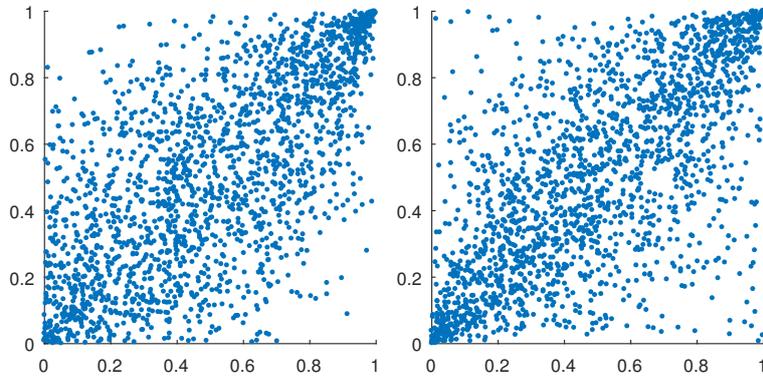
(e) Gumbel ( $\kappa = 2$ )(f) Student's t ( $\kappa_1=0.7, \kappa_2=3$ )

Figure 3. Scatter plots of various colulas

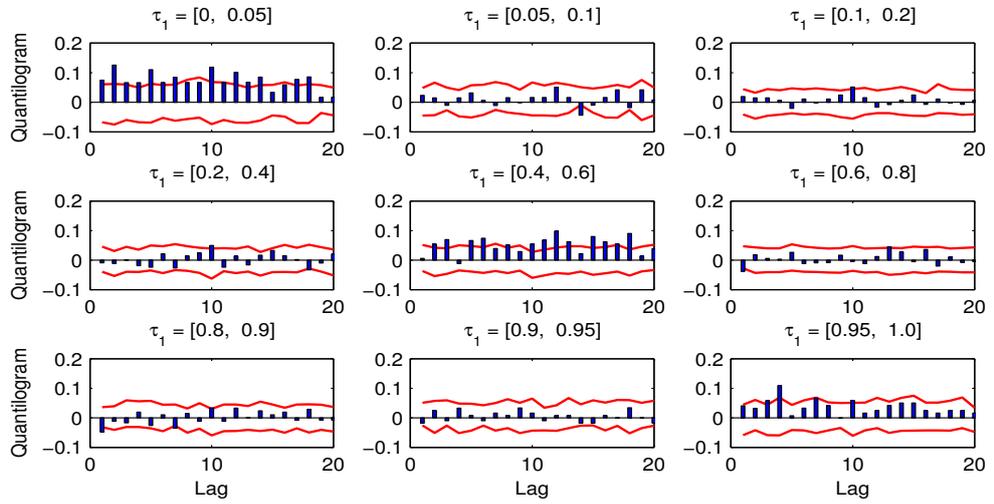


Figure 4(a). [US] Auto-quantilogram  $\hat{\rho}_\tau(k)$  of the S&P 500 index return series.  $\tau_1$  is the quantile range. Bar graphs describe sample quantilograms and lines are the 95% bootstrap confidence intervals centered at zero.

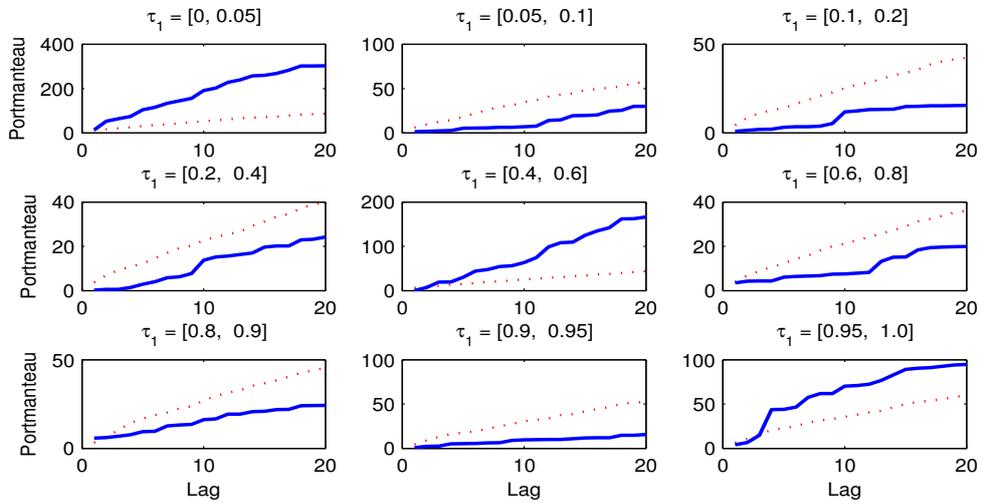


Figure 4(b). [US] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(a). The dashed lines are the 95% bootstrap confidence intervals centered at zero.

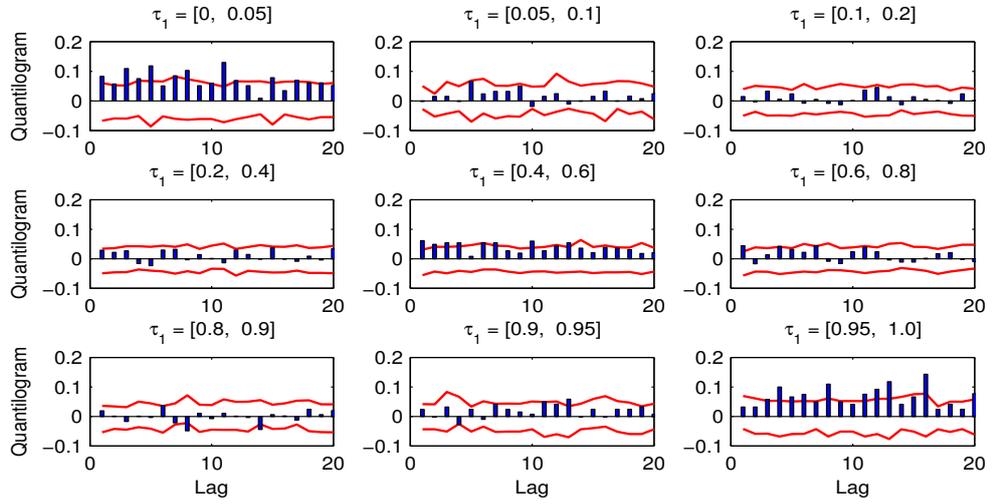


Figure 5(a). [UK] Auto-quantilogram  $\hat{\rho}_\tau(k)$  of the FTSE index return series. Same as Figure 1(a).

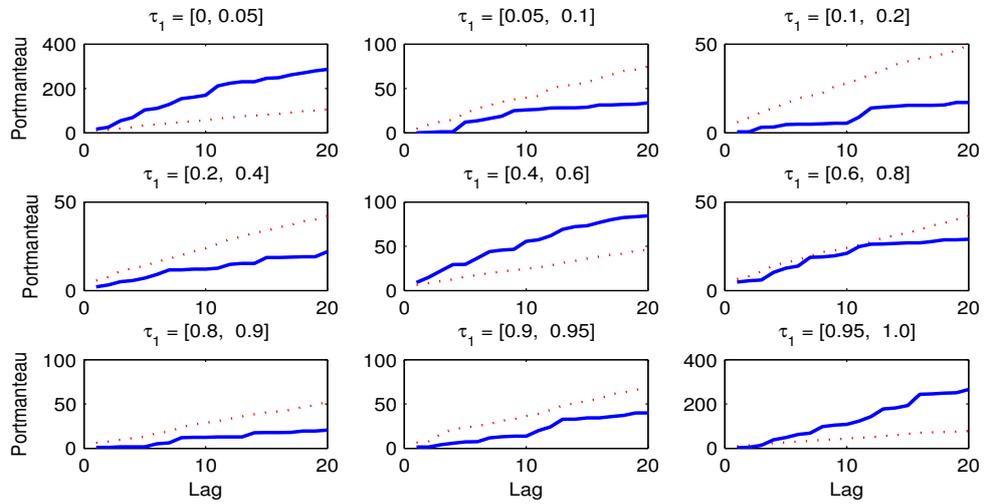


Figure 5(b). [UK] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

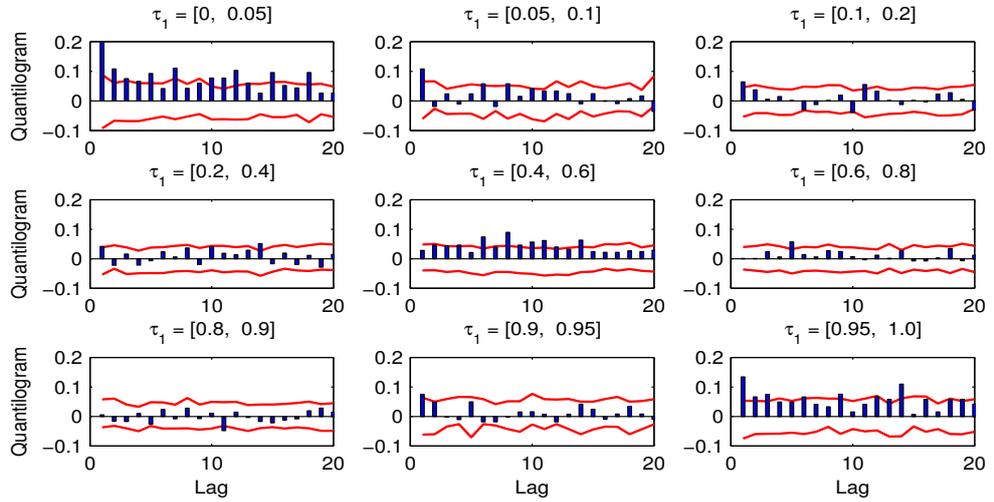


Figure 6(a). [US to UK] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to UK.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

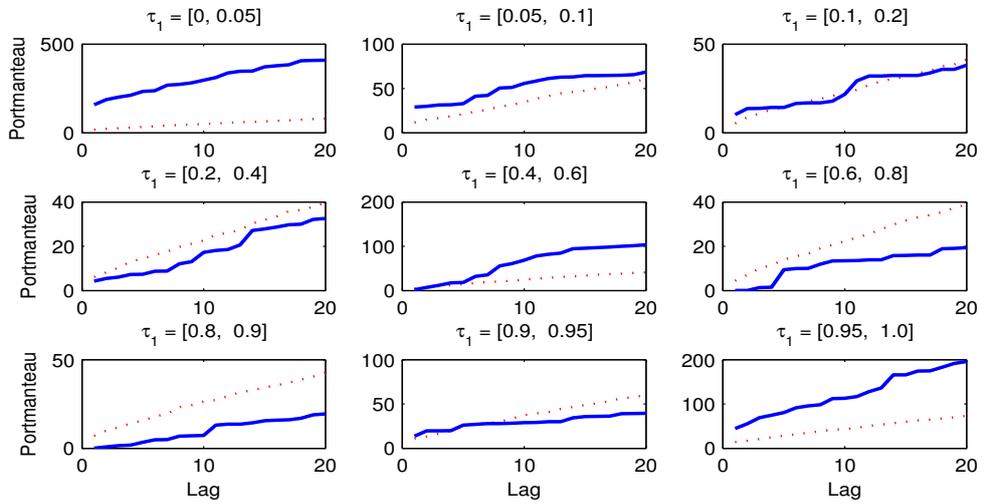


Figure 6(b). [US to UK] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

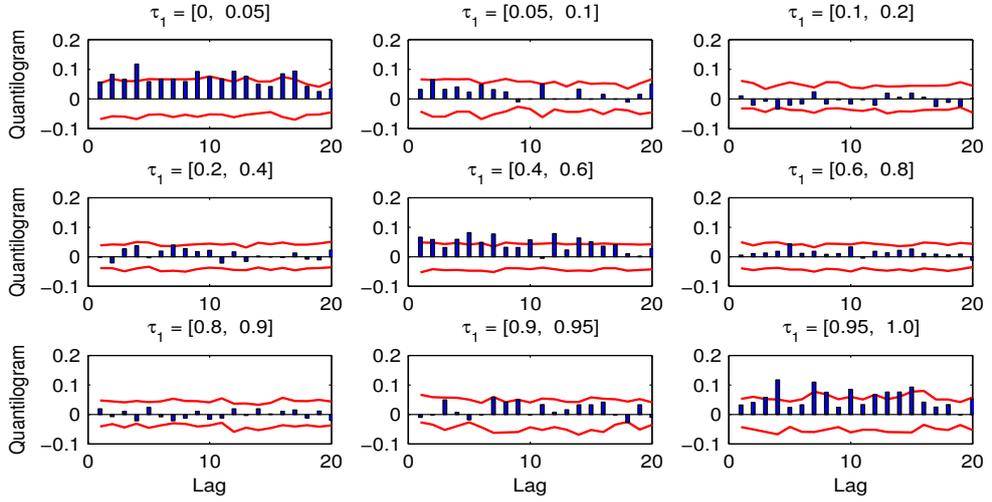


Figure 7(a). [UK to US] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from UK to US  $\tau_1 = \tau_2$ . Same as Figure 1(a).

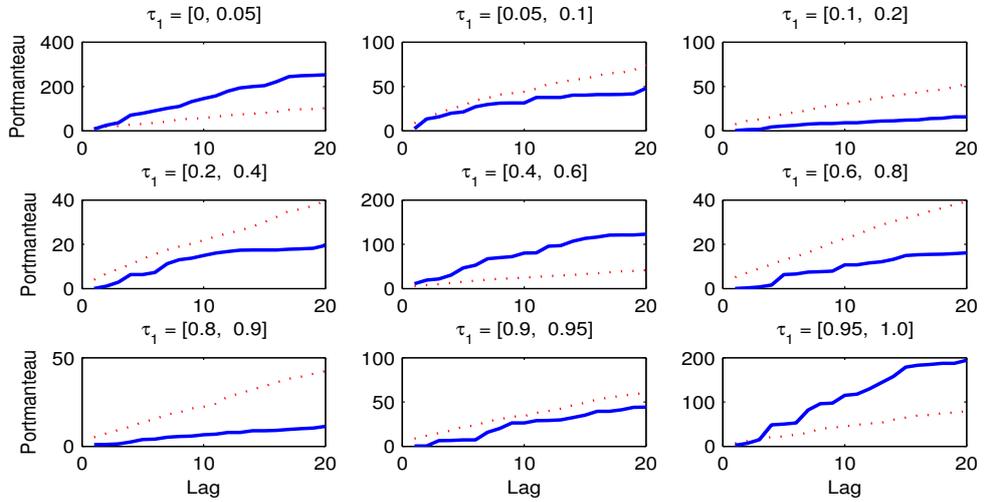


Figure 7(b). [UK to US] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 4(b).

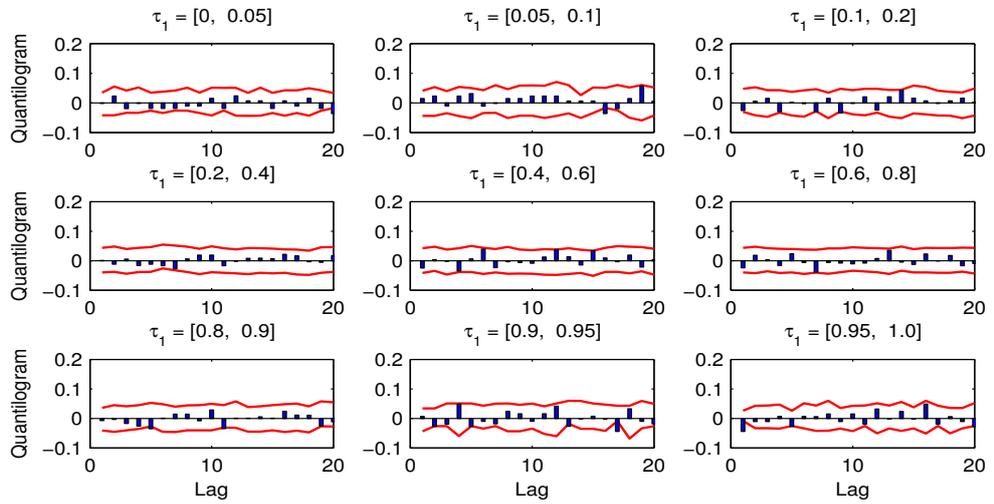


Figure 8(a). [US, std. residual] Auto-quantilogram  $\hat{\rho}_\tau(k)$  of the S&P 500 index return series using the standardized residual from the GJR-GARCH model.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

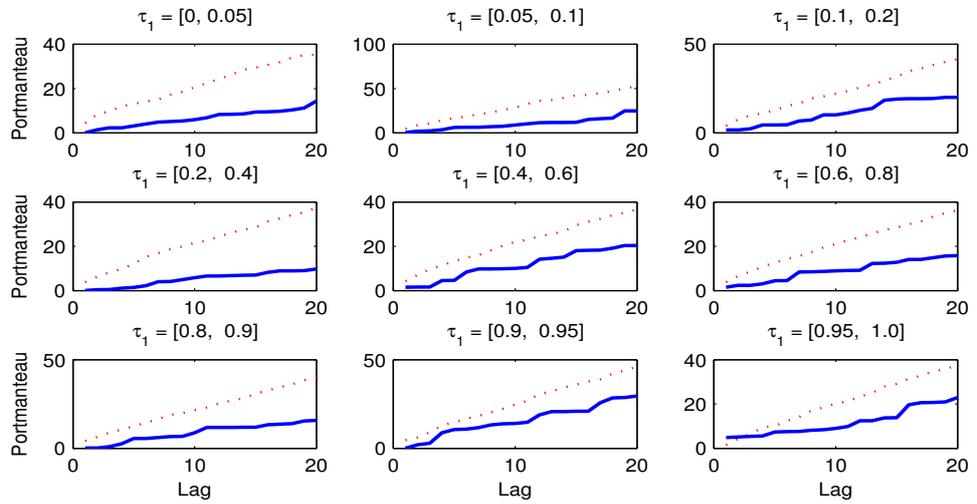


Figure 8(b). [US, std. residual] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

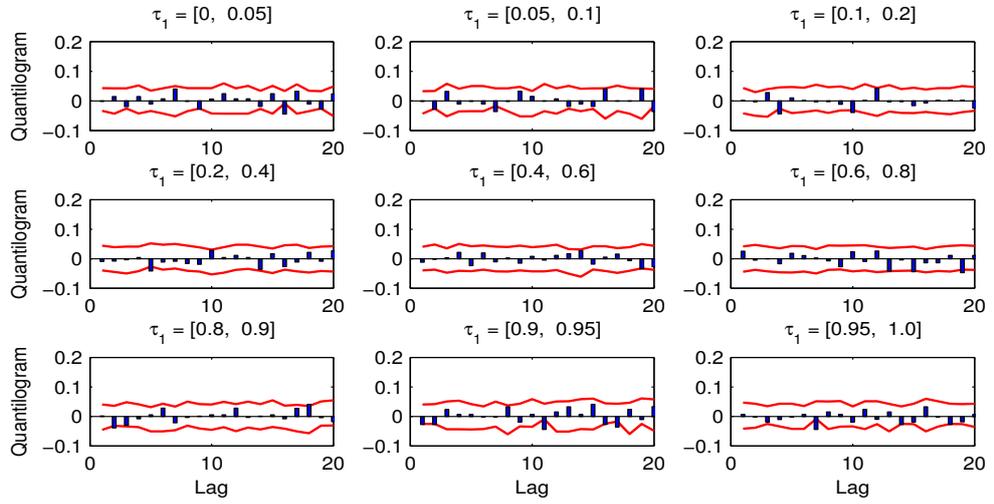


Figure 9(a). [UK, std. residual] Auto-quantilogram  $\hat{\rho}_\tau(k)$  of the FTSE index return series using the standardized residual from the GJR-GARCH model.  $\tau_1=\tau_2$ . Same as Figure 1(a).

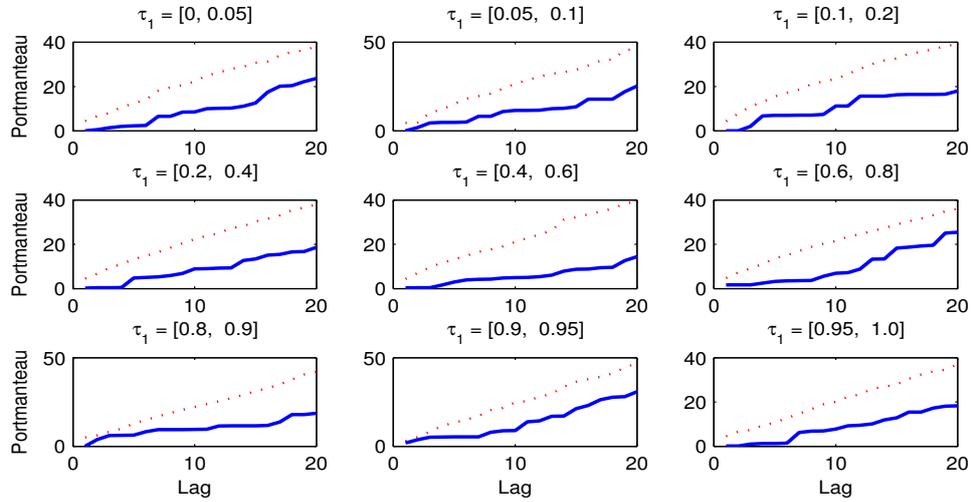


Figure 9(b). [UK, std. residual] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

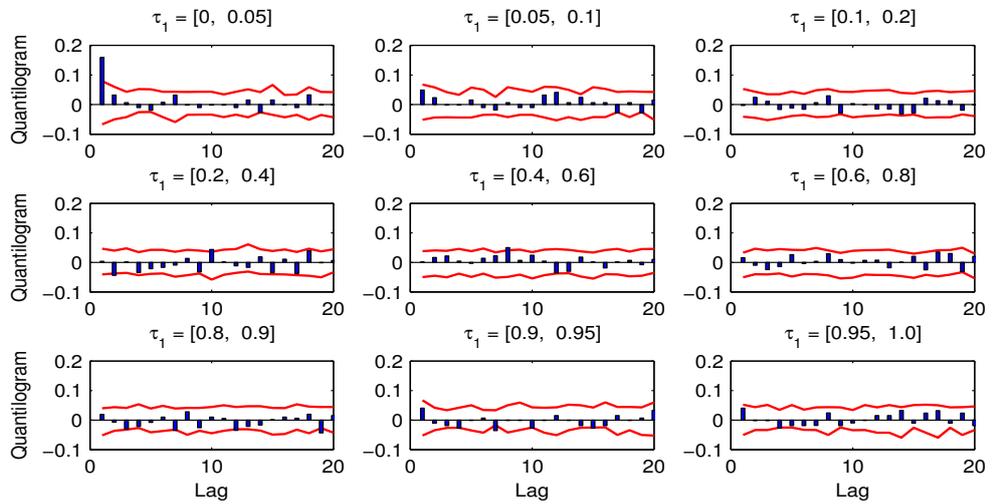


Figure 10(a). [US to UK, std. residual] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to UK using the standardized residual from the GJR-GARCH model.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

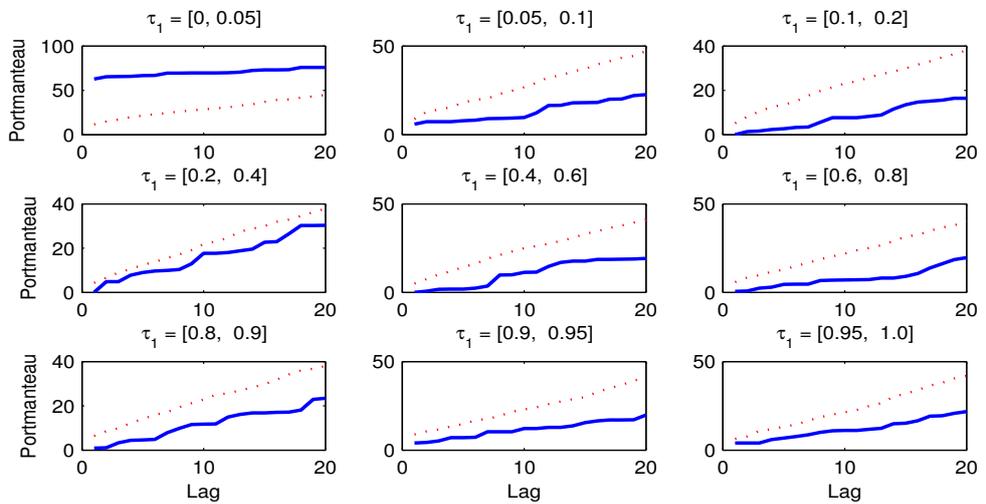


Figure 10(b). [US to UK, std. residual] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

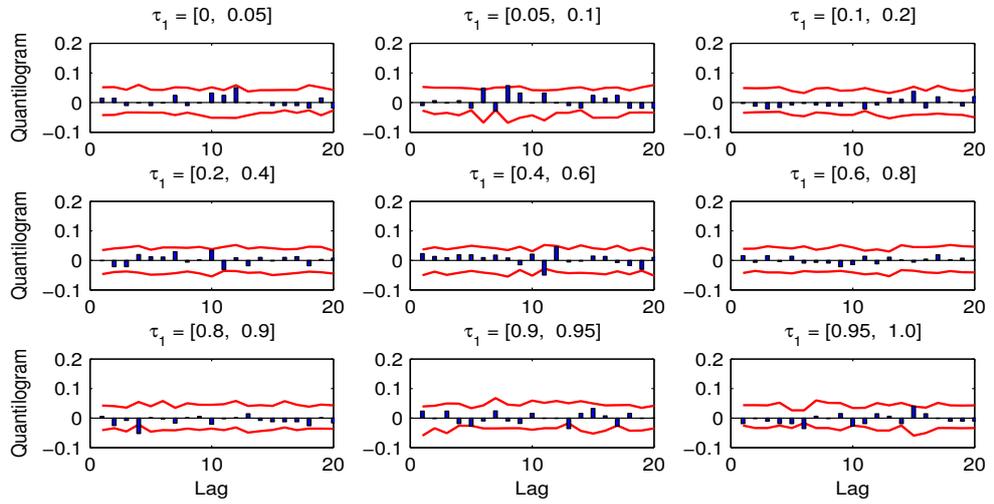


Figure 11(a). [UK to US, std. residual] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from UK to US using the standardized residual from the GJR-GARCH model.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

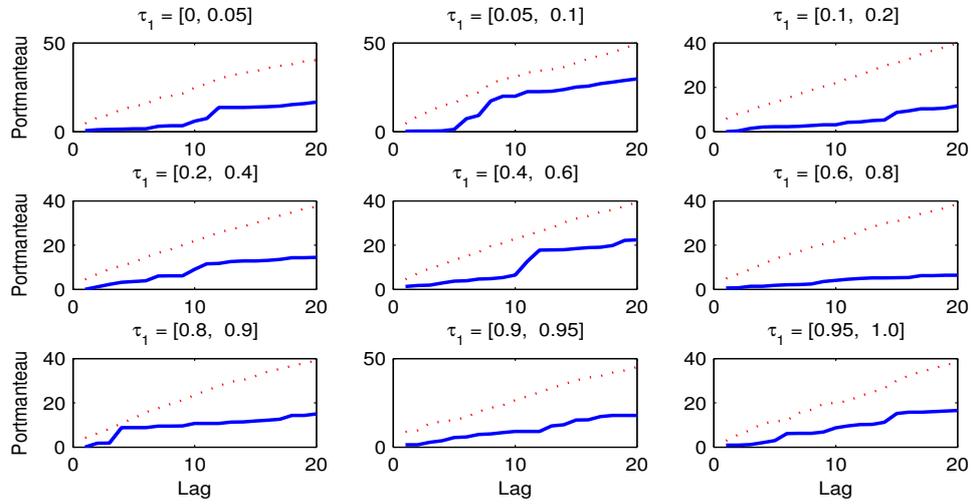


Figure 11(b). [UK to US, std. residual] Box-Ljung test statistic  $\hat{Q}_\tau^{(p)}$  for each lag  $p$  using  $\hat{\rho}_\tau(k)$ . Same as Figure 1(b).

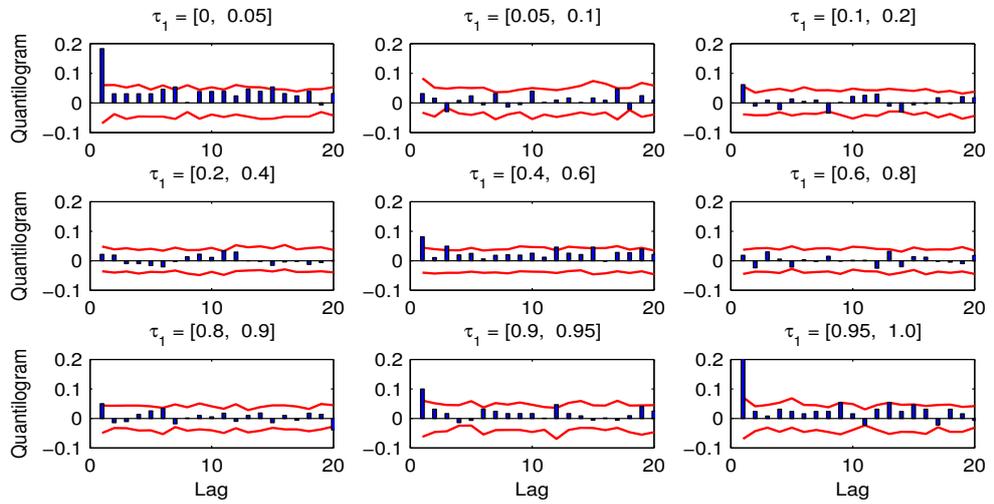


Figure 12(a). [US to Japan] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to Japan.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

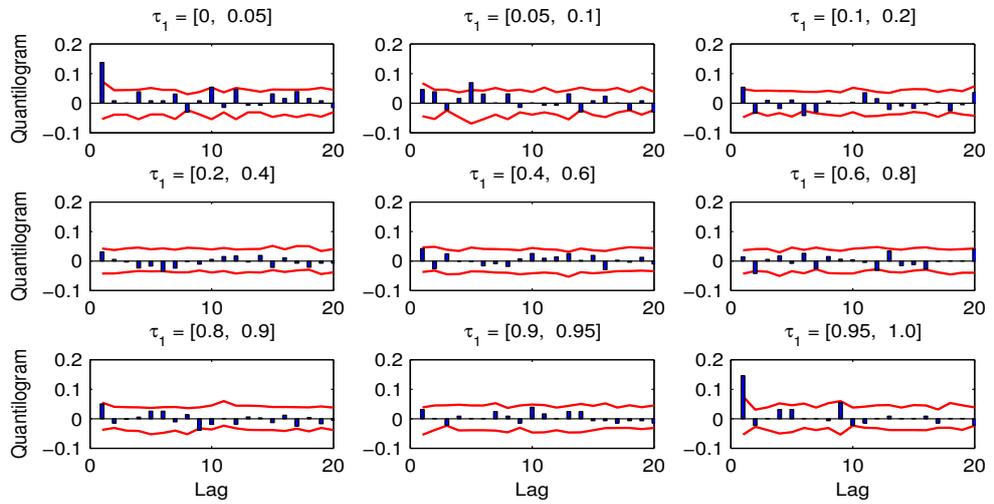


Figure 12(b). [US to Japan, std. residual] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to Japan using the standardized residual from the GJR-GARCH model.  $\tau_1 = \tau_2$ . Same as Figure 1(a).

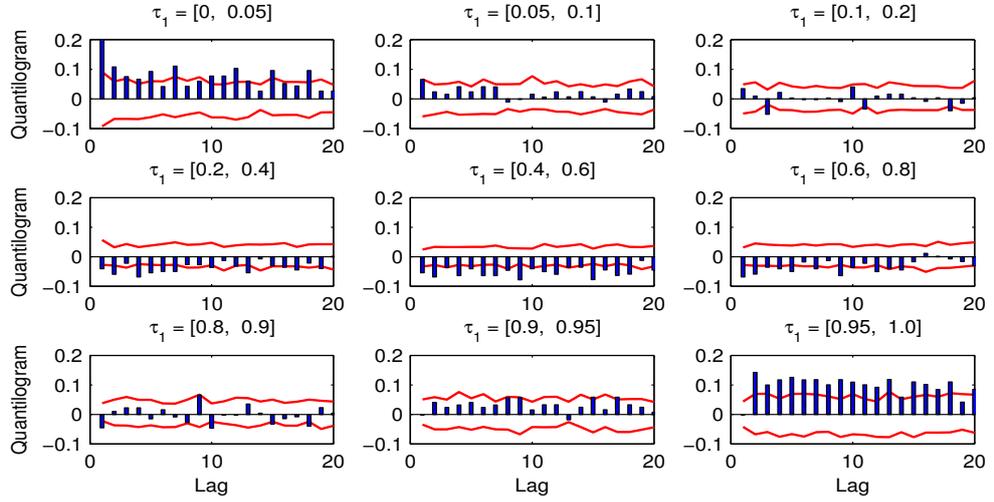


Figure 13(a). [US to UK, from the lower-quantile] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to UK.  $\tau_1 \neq \tau_2$  and  $\tau_2=[0,0.05]$  where  $\tau_2$  is the quantile range of US the stock return. Same as Figure 1(a).

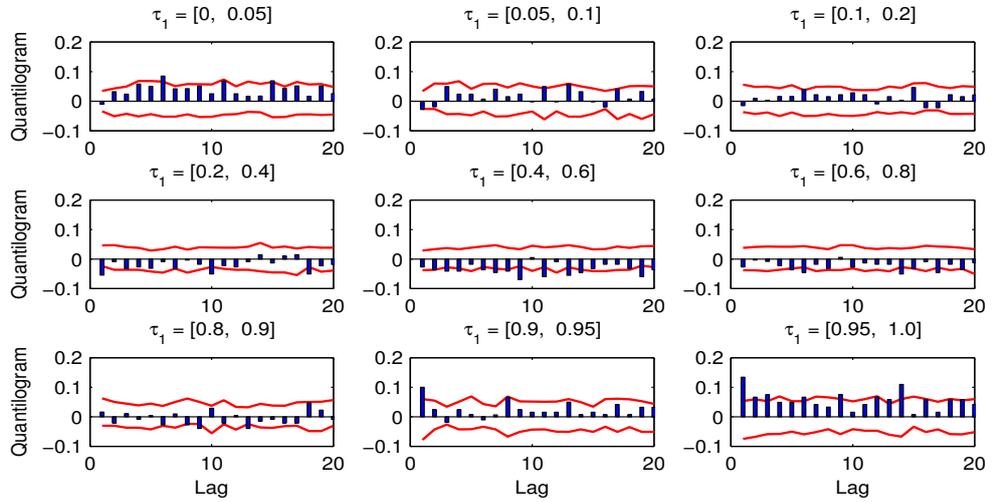


Figure 13(b). [US to UK, from the upper-quantile] Cross-quantilogram  $\hat{\rho}_\tau(k)$  to detect directional predictability from US to UK.  $\tau_1 \neq \tau_2$  and  $\tau_2=[0.95, 1]$  where  $\tau_2$  is the quantile range of the US stock return. Same as Figure 1(a).

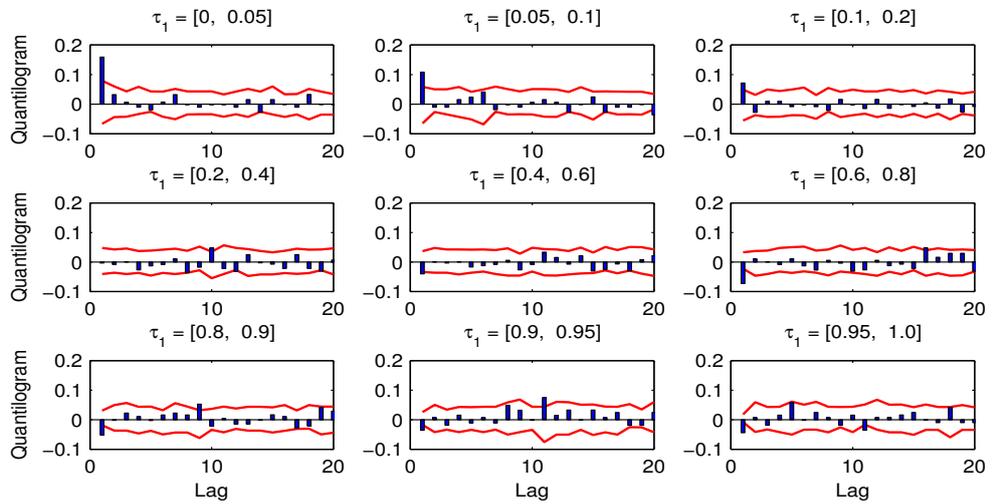


Figure 14(a). [US to UK, std. residual, from the lower-quantile] Cross-quantilogram  $\hat{\rho}_\tau(k)$  from US to UK using the standardized residual from the GJR-GARCH model.  $\tau_1 \neq \tau_2$  and  $\tau_2 = [0, 0.05]$  where  $\tau_2$  is for the US stock return. Same as Figure 1(a).

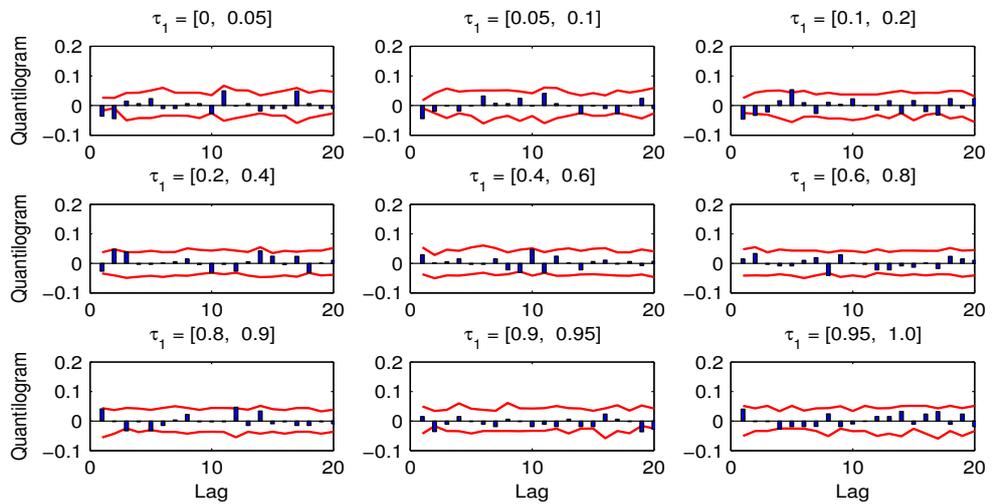


Figure 14(b). [US to UK, std. residual, from the upper-quantile] Cross-quantilogram  $\hat{\rho}_\tau(k)$  from US to UK for  $\tau_1 \neq \tau_2$  and  $\tau_2 = [0.95, 1]$ . Same as Figure 13(a).

## REFERENCES

- Ang, A. and J. Chen (2002) Asymmetric correlations of equity portfolios, *Journal of Financial Economics*, 63, 443-494.
- Barndorff-Nielsen, O. E., P. R. Hansen, A. Lunde and N. Shephard (2009). Realized kernels in practice: trades and quotes, *Econometrics Journal* 12, C1-C32.
- Baele, L. (2005) Volatility spillover effects in European equity markets, *Journal of Financial and Quantitative Analysis*, 40, 373-401.
- Baumöhl, E. and Š. Lyócsa (2017) Directional predictability from stock market sector indices to gold: A cross-quantilogram analysis, *Finance Research Letters*, forthcoming.
- Brockwell, P. and R.A. Davis (1991) *Time Series: Theory and Methods*, 2nd ed., Springer, New York.
- Baruník, J. and T. Kley (2015) Quantile cross-spectral measures of dependence between economic variables (arXiv:1510.06946v1), *ArXiv e-prints*.
- Cappiello, L., B. Gérard, A. Kadareja and S. Manganelli (2014) Measuring comovements by regression quantiles, *Journal of Financial Econometrics*, 12, 645-678.
- Das, S. and R. Uppal (2004) Systemic risk and international portfolio choice, *Journal of Finance*, 59, 2809-2834.
- Davis, R.A. and T. Mikosch (2009) The extremogram: a correlogram for extreme events, *Bernoulli*, 15, 977-1009.
- Davis, R.A., T. Mikosch, and Y. Zhao (2013) Measures of serial extremal dependence and their estimation, *Stochastic Processes and their Applications*, 123, 2575-2602.
- Demarta, S. and A.J. McNeil (2005) The  $t$  copula and related copulas, *International Statistical Review*, 73, 111-129.
- Diebold, F. X. and R. S. Mariano (1995). Comparing predictive accuracy, *Journal of Business & Economic Statistics* 13, 253-263.
- Dungey, M., R. Fry, B. González-Hermosillo and V.L. Martin (2005) Empirical modelling of contagion: A review of methodologies, *Quantitative Finance*, 5, 9-24.

- Forbes, K.J. and R. Rigobon (2002) No contagion, only interdependence: Measuring stock market comovement, *Journal of Finance*, 57, 2223-2261.
- Garcia, R. and G. Tsafack (2011) Dependence structure and extreme comovements in international equity and bond markets, *Journal of Banking and Finance*, 35, 1954-1970.
- Han, H. and D. Kristensen (2014) Asymptotic theory for the QMLE in GARCH-X models with stationary and non-stationary covariates, *Journal of Business & Economic Statistics*, 32, 416-429.
- Han, H., O. Linton, T. Oka and Y.-J. Whang (2016) The cross-quantilogram: measuring quantile dependence and testing directional predictability between time series, *Journal of Econometrics*, 193, 251-270.
- Hansen, P. R. and A. Lunde (2006) Consistent ranking of volatility models, *Journal of Econometrics* 131, 97-121.
- Hansen, P.R., Z. Huang and H.H. Shek (2012) Realized GARCH: a joint model for returns and realized measures of volatility, *Journal of Applied Econometrics*, 27, 877-906.
- Heber, G., A. Lunde, N. Shephard and K. Sheppard (2009). *OMI's realised library*, Version 0.1. Oxford-Man Institute, University of Oxford.
- Jiang, H., J.-J. Su, N. Todorova and E. Roca (2016) Spillovers and directional predictability with a cross-quantilogram analysis: The case of U.S. and Chinese agricultural futures, *Journal of Futures Markets*, 36, 1231-1255.
- Joe, H. (1997) *Multivariate Models and Dependence Concepts*, Monographs in Statistics and Probability 73, Chapman and Hall, London.
- Karolyi, G.A. (1995) A multivariate GARCH models of international transmission of stock returns and volatility: the case of United States and Canada, *Journal of Business & Economic Statistics*, 13, 11-25.
- King, M., E. Sentana and S. Wadhvani (1994) Volatility and links between national stock markets, *Econometrica*, 62, 901-933.
- Lee, T.-H. and X. Long (2009) Copula-based multivariate GARCH model with uncorrelated dependent errors, *Journal of Econometrics*, 150, 207-218.

- Li, G., Y. Li and C.-L. Tsai (2015) Quantile correlations and quantile autoregressive modeling, *Journal of the American Statistical Association*, 110(509), 246-261.
- Linton, O. and Y.-J. Whang (2007) The quantilegram: With an application to evaluating directional predictability, *Journal of Econometrics*, 141, 250-282.
- Login, F. and B. Solnik (2001) Extreme correlations in international equity markets, *Journal of Finance*, 56, 649-676.
- Nelsen, R.B. (2006) *An Introduction to Copulas*, Second Edition, Springer, U.S.A.
- Patton, A.J. (2010) Volatility forecast comparison using imperfect volatility proxies, *Journal of Econometrics*, 160, 246-256.
- Patton, A.J. (2013) Copula methods for forecasting multivariate time series, in G. Elliott and A. Timmermann (eds.) *Handbook of Economic Forecasting*, Volume 2, Springer, Verlag.
- Patton, A.J., 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.
- Patton, A.J. and K. Sheppard (2009) Evaluating volatility and correlation forecasts. in Andersen, T.G., Davis, R.A., Kreiss, J.P. and Mikosch, T. (Eds.), *Handbook of Financial Time Series*, 801-838, Springer, New York.
- 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.
- Poon, S., M. Rockinger and J. Tawn (2004) Extreme value dependence in financial markets: diagnostics, models, and financial implications, *Review of Financial Studies*, 17, 581-610.
- Reboredo, J. and A. Ugolini (2016) Quantile dependence of oil price movements and stock returns, *Energy Economics*, 54, 33-49.
- Rodriguez, J.C. (2007) Measuring financial contagion: A Copula approach, *Journal of Empirical Finance*, 14, 401-423.

- Schmitt, T.A., R. Schäfer, H. Dette and T. Guhr (2015) Quantile correlation: Uncovering temporal dependencies in financial time series, *Journal of Theoretical and Applied Finance*, forthcoming.
- Shahzad, S.J.H., N. Naifar, S. Hammoudeh and D. Roubaud (2017) Directional predictability from oil market uncertainty to sovereign credit spreads of oil-exporting countries: Evidence from rolling windows and cross-quantilegram analysis, *Energy Economics*, forthcoming.
- Sim, N. and H. Zhou (2015) Oil prices, US stock return, and the dependence between their quantiles, *Journal of Banking & Finance*, 55, 1-8.
- Shephard, N. and K. Sheppard (2010), Realising the future: Forecasting with high frequency based volatility (HEAVY) models, *Journal of Applied Econometrics*, 25, 197-231.
- Todorova, N. (2017) The intraday directional predictability of large Australian stocks: A cross-quantilegram analysis, *Economic Modelling*, 64, 221-230.
- West, K.D. (1996) Asymptotic inference about predictive ability, *Econometrica*, 64, 1067-1084.
- Yifan, S. (2017) International risk transmission of stock market movements, *Economic Modelling*, forthcoming.