Extremal Correlation for GARCH Data

Carmela Quintos*

University of Rochester

Abstract

This paper derives the distribution of the extremal correlation estimator allowing for GARCH dependence in the data. In contrast to existing tests for tail dependence that impose the iid assumption, our test is applicable to financial time series. Our asymptotic theory is based on extreme value theory methods. So long as stationarity is satisfied, the difference between the distribution under the iid and non-iid case is a scaling variance. The variance corrections needed for the extremal correlation estimator under GARCH dependence are derived.

JEL Classification: C22 and G15.

Keywords and Phrases: Extremal dependence, Extreme Value Theory, Hill's estimator, Tail dependence.

*My thanks go to seminar participants at the University of Rochester for comments. Address for Correspondence: Simon Graduate School of Business, University of Rochester, Box 270100, Rochester, New York 14627-0100. Email: quintos@simon.rochester.edu.

1 Introduction

Relationships between extreme events are of interest for various measures of financial performance. The dependence between assets, particularly dependence of extreme movements, is an issue in risk management. Since the objective is to minimize the simultaneous occurrence of large losses across assets in a portfolio, particular attention is paid to the correlation between assets during periods of crisis when constructing value-at-risk (VaR) models. Extremal dependence is also of interest in the contagion literature. Whether a shock to one economic variable is contagious is often investigated by testing whether markets move closely together during turbulent periods. See for example Longin-Solnick (2001) and Forbes-Rigobon (2001) who test for contagion in extremal events in international equity markets and Patton (1999) in exchange rates.

There are several methods available to measure tail dependence. Tail dependence is defined as the probability that a variable exceeds a threshold value, given that the realization of another variable exceeds a threshold value, and the thresholds are defined far out in the tail of its distribution. Parametric estimation requires specification of tail dependence by a copula (dependence function). Longin and Solnick (2001) uses this approach on equity markets and Patton (1999) does so for exchange rates. The copula approach is a parametric approach that requires knowledge of the full distribution, and closed form solutions for the copula do not exist for most joint densities. Bae, Karolyi and Stulz (2000) avoid the copula approach by estimating the probability using a multinomial logistic regression on risk factors. However, a key feature of these papers is that their extreme value methods are restricted to iid data. This may be viewed as a strong weakness of these studies as the iid assumption is usually violated and a GARCH type process is often used to successfully model the data.

This paper uses the extremal correlation estimator as a measure of tail dependence and derives its statistical properties under the GARCH framework. Our estimator differs from those in the literature in that we work with expectations (rather than probabilities) of tail events. The advantage is that our tail dependence measure can be written explicitly in terms of the *tail index* which determines the shape of the tail. The statistical properties of the tail index have been studied extensively, and the extension of the asymptotics to the GARCH framework is contained in Quintos-Fan-Phillips (2001) and Starica (2000). Because these asymptotics have been developed, it is feasible for us to derive the statistical properties of the extremal correlation estimator. In particular, we show that it is consistent and, when appropriately normalized, has an asymptotic Normal distribution. This greatly simplifies the construction of any test statistic based on our estimator (say, a test of constant extremal correlation over time).

Our estimator is derived using an approach that assumes an asymptotic form for the tails and then nonparametrically computes the corresponding probability of interest. The tail specification in this case requires knowledge of the asymptotic form of the tails rather than the full data generating distribution. Typically, a Pareto tail is assumed since it allows for tail fatness which is a condition known to occur in financial data. Poon-Rockinger-Twan (2001), Starica (1999), Hartmann, et.al. (2001) are studies that have used this approach. In all these papers, tail dependence is measured from the conditional probability of extremes with varying suggestions of how to appropriately calculate the tail probability. Only Starica (1999) takes into account GARCH dependence in the data.

The remainder of the paper is organized as follows. Section 2 discuses the notion of tail dependence and gives background material on the tail index as a measure of tail shape. We use Hill's maximum likelihood estimator as the tail index estimator although the asymptotic results (consistency and Normality) using this estimator will hold as well if we use alternative estimators (say Pickands estimator). Hill's estimator is the mean estimator for exceedances. This section contains the extension of Hill's asymptotics to the kth moment of exceedances estimator. Section 3 derives the properties of the test statistic for extremal correlation, proving its consistency and asymptotic Normality. Section 4 provides the size and power simulations for the test of upper tail correlation. Conclusions are given in Section 5 and Section 6 contains all proofs.

2 Preliminaries

In this section, we provide the intuition behind our measure of tail correlation and analyze its statistical properties.

2.1 Measures of Tail Dependence

Measures of tail dependence capture the comovement of the extremes of random variables. If two random variables X_1 and X_2 have marginal distributions F_{X_1} and F_{X_2} respectively, then upper tail dependence is usually defined as

$$\lambda_{u_1,u_2} = \lim_{u_1,u_2 \to 1} P\left(X_1 > F_{X_1}^{-1}(u_1) \,|X_2 > F_{X_2}^{-1}(u_2)\right), \tag{1}$$

where $\lambda_{u_1,u_2} \in [0,1]$. The variables are asymptotically extremal dependent if $\lambda_{u_1,u_2} \in (0,1]$ and extremal independent if $\lambda_{u_1,u_2} = 0$. A measure of the degree that extremal independence is violated is the ratio

$$\chi = \lim_{u_1,u_2 \to 1} \frac{P\left(X_1 > F_{X_1}^{-1}(u_1) \,|X_2 > F_{X_2}^{-1}(u_2)\right)}{P\left(X_1 > F_{X_1}^{-1}(u_1)\right)} - 1. \tag{2}$$

When $\chi = 0$, the variables are independent in the tails and when $\chi < 0$ ($\chi > 0$), the variables are negatively (positively) extremal dependent. Coles, et al. (1999), for example, modify (2) and define their measure of extremal dependence as

$$\ddot{\chi} = \lim_{\upsilon_1,u_2 \to 1} \frac{2\log P\left(X_1 > F_{X_1}^{-1}(u_1)\right)}{\log P\left(X_1 > F_{X_1}^{-1}(u_1) \,|X_2 > F_{X_2}^{-1}(u_2)\right)} - 1.$$

Hwang and Salmon (2002) apply $\ddot{\chi}$ to the analysis of the properties of traditional portfolio performance measures.

Tail dependence can be estimated either parametrically or nonparametrically. The parametric route requires the specification of a copula or a dependence function and is a straightforward approach if the function has a closed form. Typically, a Gumbel copula is used which, for iid data, takes the form

$$C_\delta(x_1, x_2) = \exp\left[-\left\{(-\log x_1)^{1/\delta} + (-\log x_2)^{1/\delta}\right\}^{\delta}\right],$$

where $0 < \delta \leq 1$ controls the amount of dependence between X_1 and X_2. The special cases $\delta = 1$ and $\delta \to 0$ correspond respectively to independence and perfect dependence. The relation of δ to the correlation coefficient is $\rho = 1 - \delta^2$. Kendall's tau is another commonly used extremal measure that can be calculated directly from the copula.

A nonparametric approach estimates the probabilites of exceedance with weaker assumptions on the generating distributions than the copula specification. In particular, our approach assumes only an asymptotic form for the tail. When X is known to have fat tails as in most financial time series then its tails behave asymptotically like a power function in x,

$$\bar{F}(x) = P(X > x) = x^{-\alpha}\left(1 + dx^{-\beta} + o\left(x^{-\beta}\right)\right), \qquad x > 0, \tag{3}$$

where $\beta > 0$ and $d \in \Re$ control the approximation of the tail function. The term α is the *tail index* or *tail slope* which controls the rate of tail decay. The smaller the α, the slower the rate of decay, and the thicker the tails.

Under the heavy tail assumption (3),

$$
\begin{aligned}
E\left(X > x\right) &= \int_0^\infty P\left(X > x\right) dx \\
&= \int_1^\infty P\left(\log X > \log x\right) \tfrac{dx}{x} \\
&\sim 1/\alpha,
\end{aligned}
\tag{4}
$$

which states that α is the mean of log exceedances, $\log\left(X/x\right)$. A measure of tail dependence can then be written in terms of α,

$$
\psi = \lim_{x_1,x_2\to\infty} \frac{E\left(X_1 > x_1, X_2 > x_2\right)}{E\left(X_1 > x_1\right) E\left(X_2 > x_2\right)} - 1 = \frac{\alpha_{12}^{-1}}{\alpha_1^{-1}\alpha_2^{-1}} - 1 \tag{5}
$$

using (4).

More generally, ψ can be interpreted as the extremal correlation measure. For a series X define its exceedances relative to the threshold b as $\xi = \frac{X}{b}$. Define the notation for any variable z_t, $z_{t+} = z_t 1_{(A)}$ where $1_{(A)}$ is an indicator function which takes the value 1 when A is satisfied. Also let $z_t^* = \log z_t$. Note that if we let $\xi_+^* = \log\left(X/b\right) 1_{(\xi>1)}$ then the *kth* moment is

$$
E\left(\xi_+^*\right)^k \sim \frac{k!}{\alpha^k}, \tag{6}
$$

from (see Hsing, 1991, equation 1.5). The *extremal variance* for any univariate series X_i can be computed as

$$
\sigma^2_{\xi^*_{i+}} = E\left(\log\left(X_i/b_i\right) 1_{(\xi_i>1)}\right)^2 - \left[E\left(\log\left(X_i/b_i\right) 1_{(\xi_i>1)}\right)\right]^2 \sim \frac{2!}{\alpha_i^2} - \frac{1}{\alpha_i^2} = \alpha_i^{-2} > 0 \tag{7}
$$

using (6) on the first term with $k = 2$. Then in terms of our notation the upper tail correlation is

$$
\begin{aligned}
\psi &= \lim_{b_1,b_2\to\infty} \frac{Cov(X_1>b_1, X_2>b_2)}{\sqrt{\sigma^2_{\xi^*_{1+}}}\sqrt{\sigma^2_{\xi^*_{2+}}}} \\
&= \lim_{b_1,b_2\to\infty} \frac{E\left(\log(X_1/b_1)\log(X_2/b_2) 1_{(\xi_{1t}>1,\xi_{2t}>1)}\right) - E\left(\xi^*_{1+}\right)E\left(\xi^*_{2+}\right)}{\sqrt{\sigma^2_{\xi^*_{1+}}}\sqrt{\sigma^2_{\xi^*_{2+}}}} \\
&= \frac{\alpha_{12}^{-1} - \alpha_1^{-1}\alpha_2^{-1}}{\sqrt{\alpha_1^{-2}}\sqrt{\alpha_2^{-2}}} = \frac{\alpha_{12}^{-1}}{\alpha_1^{-1}\alpha_2^{-1}} - 1
\end{aligned}
\tag{8}
$$

as in (5).

Our concern is in deriving the distribution of the correlation of upper tail exceedances for both iid and non-iid data, unlike the copula based measures that are defined only for the iid case. The upper tail correlation ψ is of interest in many financial applications where the concern is whether or not portfolios be uncorrelated during extreme market movements. There has been little statistical theory, particularly in the non-iid case, for inference using this statistic. Results on the lower tail correlation are obviously obtained by changing the sign of the series.

2.2 EVT Asymptotics

Extreme value theory (EVT) is used to find the distribution of $\hat{\psi}$. Since $\hat{\psi}$ depends on $\hat{\sigma}^2_{\xi^*_+}$, the distribution theory of $\hat{\psi}$ is driven by the asymptotics of the tail index estimator $\hat{\alpha}^{-k}$.

We define $\hat{\alpha}^{-k}$ using Hill's nonparametric method (Hill, 1975). It is based on EVT in that it makes use of only the m largest exceedances. For a time series $\{X_t\}_{t=1}^T$ ordered as $X_{(1)} \leq ... \leq X_{(m)} ... \leq X_{(T)}$, Hill's estimator takes the mean over the threshold $\log X_{(T-m)}$ (note that log is a monotonic transformation so order is preserved),

$$\hat{\alpha}^{-1} = \frac{1}{m}\sum_{t=1}^{m} \log\left(\frac{X_{(T-t+1)}}{X_{(T-m)}}\right) = \frac{1}{m}\sum_{t=1}^{T}\left[\log\left(\frac{X_t}{X_{(T-m)}}\right)\right] 1_{(\xi>1)} = \frac{1}{m}\sum_{t=1}^{T}\xi^*_{t+}. \tag{9}$$

The number of the largest order statistics $m = \sum_{t=1}^T 1_{(\xi_t>1)}$ can be chosen using the Hill plot of m on $\hat{\alpha}^{-1} = \hat{\alpha}^{-1}(m)$, and to select m in the region where the estimator is stable (see, for example, Resnick (1998) for the use of the Hill plot).

The asymptotics of $\hat{\alpha}^{-1}$ have been derived for the iid case by Hall (1982) and for GARCH data by Quintos-Fan-Phillips (2001). The major result is that, for iid data,

$$m^{1/2}\alpha^{-1}(\hat{\alpha}-\alpha) \xrightarrow{d} N(0,1), \tag{10}$$

i.e., Hill's estimator is asymptotically Normally distributed provided $m = o(T)$, and for GARCH data a scaling parameter is required,

$$m^{1/2}\zeta^{-1/2}\alpha^{-1}(\hat{\alpha}-\alpha) \xrightarrow{d} N(0,1), \tag{11}$$

where the scaling parameter ζ depends on the parameters of the GARCH process.

To develop the asymptotics of $\hat{\psi}$, we extend (10) and (11) in our-set up here. Consider the k-th moment of the exceedances estimator as an extension of (9),

$$\hat{\alpha}^{-k}(p) = \frac{1}{k!m}\sum_{t=1}^{m}\left[\log\left(\frac{X_{(T-t+1)}}{X_{(T-m)}}\right)^p\right]^k = \frac{1}{k!m}\sum_{t=1}^{T}\left[\log\left(\frac{X_t}{X_{(T-m)}}\right)^p 1_{(\xi>1)}\right]^k = \frac{1}{k!m}\sum_{t=1}^{T}\xi^{k*}_{t+}(p).$$

For example if $k=1$ and $p=2$,

$$\hat{\alpha}^{-1}(2) = \frac{1}{m}\sum_{t=1}^{m}\log\left(\frac{X^2_{(T-t+1)}}{X^2_{(T-m)}}\right)$$

is just Hill's (mean) estimator for the upper tail exceedances of the squared sequence $\{X_t^2\}$. Similarly, if $k=2$ and $p=1$ then

$$\hat{\alpha}^{-2}(1) = \frac{1}{2m}\sum_{t=1}^{m}\left[\log\left(\frac{X_{(T-t+1)}}{X_{(T-m)}}\right)\right]^2$$

is interpreted as the second moment of the upper tail exceedances of the sequence $\{X_t\}$. The generalization of Hill's asymptotics for the iid case is given in the following Lemma.

Lemma 1. Define the k-th moment of exceedances for the sequence $\{X_t^p\}$ as

$$\hat{\alpha}^{-k}(p) = \frac{1}{k!m} \sum_{t=1}^{m} \left(\log X_{(T-t+1)}^p - \log X_{(T-m)}^p \right)^k . \tag{12}$$

Then with $m = o(T)$,

$$\hat{\alpha}^{-k}(p) \xrightarrow{p} p^k / \alpha^k, \tag{13}$$

and

$$m^{1/2} \hat{\alpha}^k(p) \left(\hat{\alpha}^{-k}(p) - p^k \alpha^{-k} \right) \xrightarrow{d} N(0,1). \tag{14}$$

Equation (13) states that powers of X enter only as scale coefficients and in fact cancel out in the asymptotics of $\hat{\psi}$, while equation (14) provides the asymptotic distribution of $\hat{\alpha}^{-k}$.

An example where tail estimation is applied to the power of the series is GARCH data. Assume the standard GARCH(1, 1) process for the data,

$$\begin{aligned} X_t^2 &= \sigma_t^2 Z_t^2, \quad Z_t \sim iid\ N(0,1), \\ \sigma_t^2 &= \beta_0 + \beta_1 \sigma_{t-1}^2 + \lambda X_{t-1}^2. \end{aligned} \tag{15}$$

The distribution Z is assumed Normal however the results can accomodate the case that Z also has fat tails. The important condition is that stationarity holds, i.e. $0 < \beta_1 + \lambda < 1$.

We can write (15) as

$$\sigma_t^2 = \left(\beta_1 + \lambda Z_{t-1}^2 \right) \sigma_{t-1}^2 + \beta_0 = A_t \sigma_{t-1}^2 + B_t. \tag{16}$$

Set $Y_t = \hat{\sigma}_t$ and assume that the process starts at Z_0 (i.e. $(Z_{-1}, ..., Z_{-\infty}) = 0$). Note that Y_t can be written in finite moving average form (with random coefficients),

$$Y_t^2 = \prod_{i=1}^{t} A_i Y_0^2 + \sum_{j=1}^{t} \left(\prod_{i=j+1}^{t} A_i \right) B_j = G_1^t Y_0^2 + Y_t^2(A,B), \tag{17}$$

where we have set $\prod_{i=j+1}^{t} A_i = 1$ for $j \geq t$. Stationarity in this case requires $E\left(G_1^t\right) < 1$. So long as the condition of stationarity is satisfied the tails of σ_t^2 (and hence X_t^2) behaves like (3). Then α can be estimated from

$$\hat{\alpha}^{-1}(2) = \frac{1}{m} \sum_{t=1}^{m} \left(\log \hat{\sigma}_{(T-t+1)}^2 - \log \hat{\sigma}_{(T-m)}^2 \right). \tag{18}$$

The tails of X behaves like the tails of σ, i.e. $P\left(X_t^2 > b\right) \sim E\left(Z_t^2\right)^{\alpha} P\left(\sigma_t^2 > b\right)$. Extending Lemma 1 to the non-iid case requires a scale parameter in the asymptotics as in (11).

Lemma 2. If $\{X_t\}$ follows a stationary GARCH process, we calculate the tail slope as

$$\hat{\alpha}^{-k}(p) = \frac{1}{k!m}\sum_{t=1}^{T}\left[\log\left(\frac{Y_t}{Y_{(T-m)}}\right)^p 1_{(\xi>1)}\right]^k = \frac{1}{k!m}\sum_{t=1}^{T}\xi_{t+}^{k*}(p). \tag{19}$$

Then with $m = o(T)$,

$$\hat{\alpha}^{-k}(p) \xrightarrow{p} p^k/\alpha^k \tag{20}$$

and

$$m^{1/2}\hat{\zeta}^{-1/2}\left(\hat{\alpha}^{-k}(p) - p^k\alpha^{-k}\right) \xrightarrow{d} N(0,1). \tag{21}$$

The scale parameter is calculated by defining

$$C_t = \frac{\xi_{t+}^{k*}(p)}{k!} \text{ and } D_t = 1\left(\log Y_t > \log Y_{(T-m)}\right),$$

then $\hat{\zeta} = \frac{1}{(k!)^2}\hat{\alpha}^{-2k}(p)\left[(2k)! + \hat{\chi} + 1 + \hat{\omega} - 2k! - 2\hat{\phi}\right]$ so that from (19),

$$\hat{\alpha}^{-2k}(p) = \frac{1}{(2k)!m}\sum_{t=1}^{T}\left[\log\left(\frac{Y_t}{Y_{(T-m)}}\right)^p 1_{(\xi>1)}\right]^{2k}$$

and

$$\begin{aligned}
\hat{\chi} &= 2\hat{\alpha}^{2k}(p)\tfrac{1}{m}\textstyle\sum_{j=1}^{T-1}\sum_{i=j+1}^{T} C_j C_i \to \chi, \\
\hat{\phi} &= \hat{\alpha}^{k}(p)\tfrac{1}{m}\textstyle\sum_{j=1}^{T-1}\sum_{i=j+1}^{T}\left(C_j D_i + C_j D_i\right) \to \phi, \\
\hat{\omega} &= \ 2\tfrac{1}{m}\textstyle\sum_{j=1}^{T-1}\sum_{i=j+1}^{T} D_j D_i \to \omega
\end{aligned}$$

for $\chi, \phi, \omega < \infty$ and $\zeta = \frac{1}{(k!)^2}\left(\frac{p}{\alpha}\right)^{2k}\left[(2k)! + \chi + 1 + \omega - 2k! - 2\phi\right]$.
If $p = 2$ and $k = 1$, then the results of Lemma 2 coincide with those in Quintos, et al. (2001, Theorem 8), i.e., the tail index of a GARCH process is distributed as

$$m^{1/2}\hat{\zeta}^{-1/2}\left(\hat{\alpha}^{-1}(2) - 2/\alpha\right) \xrightarrow{d} N(0,1),$$

where $\hat{\zeta} = \hat{\alpha}^{-2}(2)\left[1 + \hat{\chi} + \hat{\omega} - 2\hat{\phi}\right]$. The only difference between the iid and non-iid case is the scale parameter ζ. So long as the process is stationary, Hill's estimator is still consistent and Normally distributed. This standard asymptotic result simplifies inference.

3 Inference

Consider the null hypothesis of no correlation in the upper tail,

$$H_0 : \qquad \psi = 0. \tag{22}$$

The estimator for ψ is the correlation of upper exceedances which we can write in tail index notation

$$\hat{\psi} = \frac{\hat{\alpha}_{12}^{-1} - \hat{\alpha}_1^{-1}\hat{\alpha}_2^{-1}}{\hat{\sigma}^*_{\xi_{1+}}\hat{\sigma}^*_{\xi_{2+}}}, \tag{23}$$

where, recalling that $\xi^*_{it+} = \log\left(\frac{X_{it}}{X_{i(T-m_i)}}\right) 1_{(\xi_{it}>1)}$,

$$\begin{aligned} \hat{\alpha}_{12}^{-1} &= \tfrac{1}{L}\textstyle\sum_{t=1}^T \xi^*_{1t+}\xi^*_{2t+}, \\ \hat{\alpha}_i^{-1} &= \tfrac{1}{L}\textstyle\sum_{t=1}^T \xi^*_{it+}, \\ \hat{\sigma}^{*2}_{\xi_{i+}} &= \tfrac{1}{L}\textstyle\sum_{t=1}^T \xi^{*2}_{it+} - \left[\tfrac{1}{L}\textstyle\sum_{t=1}^T \xi^*_{it+}\right]^2 \end{aligned}$$

for $i = 1, 2$ and $L = \sum_{t=1}^T 1_{(\xi_{1t}>1,\xi_{2t}>1)}$ is the number of joint exceedance. We note that since

$$\hat{\sigma}^{*2}_{\xi_{i+}} \to \alpha_i^{-2}$$

from (7) we can construct the estimator from

$$\hat{\psi}' = \frac{\hat{\alpha}_{12}^{-1}}{\hat{\alpha}_1^{-1}\hat{\alpha}_2^{-1}} - 1.$$

The estimators $\hat{\psi}$ and $\hat{\psi}'$ will behave differently in finite sample but will not change our asymptotic results.

Theorem 3. To derive the distribution of $\hat{\psi}$ note that $\hat{\alpha}_{12}^{-1}$ converges to

$$\hat{\alpha}_{12}^{-1} \xrightarrow{p} \alpha_1^{-1}\alpha_2^{-1} E\left(X_1 - X_{1(T-m_1)}\right)\left(X_2 - X_{2(T-m_2)}\right). \tag{24}$$

If X_1 and X_2 are tail independent (or tail dependent) we have consistency

$$\begin{aligned} \hat{\alpha}_{12}^{-1} &\xrightarrow{p} \alpha_1^{-1}\alpha_2^{-1} \left(\text{or } \hat{\alpha}_{12}^{-1} \xrightarrow{p} 2\alpha_1^{-1}\alpha_2^{-1}\right) \\ \hat{\sigma}^*_{\xi_{i+}} &\xrightarrow{p} \alpha_i^{-1} \end{aligned} \tag{25}$$

and (25) implies that

$$\hat{\psi} \xrightarrow{p} \frac{\alpha_1^{-1}\alpha_2^{-1} - \alpha_1^{-1}\alpha_2^{-1}}{\alpha_1^{-1}\alpha_2^{-1}} = 0 \ \left(\text{or } \hat{\psi} \xrightarrow{p} 1\right). \tag{26}$$

Asymptotic Normality also holds,

$$L^{1/2}\left(\hat{\psi}-\psi\right)\xrightarrow{d}N\left(0,1\right) \tag{27}$$

so long as $L=o\left(T\right)$.

Remarks. The results of Theorem 3 are valid when the generating distribution has fat tails, say a t-distribution. It is not valid however when the generating distribution is Normal.

To extend Theorem 3 to $\hat{\psi}\left(2\right)$, we need the asymptotics for the comovement of extremes $\hat{\alpha}_{12}^{-1}\left(2\right)$ in the GARCH(1,1) case. Once again the only difference is that the scaling variance needs to be defined and its asymptotics derived.

For the GARCH case, we let

$$\xi_{it+}^{*}\left(2\right)=\log\left(\frac{Y_{it}^{2}}{Y_{i(T-m_i)}^{2}}\right)1_{(\xi_{it}>1)}, \tag{28}$$

where $\xi_{it}=\frac{Y_{it}}{Y_{i(T-m_i)}}$ and $Y_{it}=\hat{\sigma}_{it}$. Moreover, define

$$\begin{aligned}
\hat{\alpha}_{12}^{-1}\left(2\right) &= \tfrac{1}{L}\textstyle\sum_{t=1}^{T}\xi_{1t+}^{*}\left(2\right)\xi_{2t+}^{*}\left(2\right),\\
\hat{\alpha}_{i}^{-1}\left(2\right) &= \tfrac{1}{L}\textstyle\sum_{t=1}^{T}\xi_{it+}^{*}\left(2\right),\\
\hat{\sigma}_{\xi_{i+}}^{2*}\left(2\right) &= \tfrac{1}{L}\textstyle\sum_{t=1}^{T}\xi_{it+}^{2*}\left(2\right)-\left[\tfrac{1}{L}\textstyle\sum_{t=1}^{T}\xi_{it+}^{*}\left(2\right)\right]^{2}
\end{aligned}$$

for $i=1,2$ and $L=\sum_{t=1}^{T}1_{(\xi_{1t}>1,\xi_{2t}>1)}$ is the number of joint exceedance. Then the test statistic for the GARCH process case can be written as

$$\hat{\psi}\left(2\right)=\frac{\hat{\alpha}_{12}^{-1}\left(2\right)-\hat{\alpha}_{1}^{-1}\left(2\right)\hat{\alpha}_{2}^{-1}\left(2\right)}{\hat{\sigma}_{\xi_{1+}}^{*}\left(2\right)\hat{\sigma}_{\xi_{2+}}^{*}\left(2\right)}. \tag{29}$$

Theorem 4. Denote $Y_{it}=\hat{\sigma}_{it}$ and let Hill's estimator be calculated as

$$\hat{\alpha}_{12}^{-1}\left(2\right)=\frac{1}{L}\sum_{t=1}^{T}\xi_{1t+}^{*}\left(2\right)\xi_{2t+}^{*}\left(2\right)$$

and define

$$C_t=\xi_{1t+}^{*}\left(2\right)\xi_{2t+}^{*}\left(2\right)\text{ and }D_t=1\left(\log Y_{1t}>\log Y_{1(T-m_1)},\log Y_{2t}>\log Y_{2(T-m_2)}\right).$$

Then if $\{X_{it}\}$ follows a GARCH(1,1) process (a)-(c) hold,

(a) $2\hat{\alpha}_{12}^{2}\frac{1}{L}\sum_{j=1}^{T-1}\sum_{i=j+1}^{T}C_jC_i\rightarrow\chi_0,$

(b) $\hat{\alpha}_{12}\frac{1}{L}\sum_{j=1}^{T-1}\sum_{i=j+1}^{T}\left(C_jD_i+C_jD_i\right)\rightarrow\phi_0,$

(c) $2\frac{1}{L}\sum_{j=1}^{T-1}\sum_{i=j+1}^{T}D_jD_i\rightarrow\omega_0.$

where $\chi_0, \phi_0, \omega_0 < \infty$. Set

$$\hat{\zeta}_0 = \left(1 + \hat{\chi}_0 + \hat{\omega}_0 - 2\hat{\psi}_0\right)$$

then Hill's estimator is distributed as

$$L^{1/2}\hat{\zeta}_0^{-1/2}\hat{\alpha}_{12}(2)\left(\hat{\alpha}_{12}^{-1}(2) - \alpha_{12}^{-1}(2)\right) \xrightarrow{d} N(0,1). \tag{30}$$

Tests based on $\hat{\psi}(2)$ are constructed from

$$L^{1/2}\hat{\zeta}_0^{-1/2}\left(\hat{\psi}(2) - \psi(2)\right) \xrightarrow{d} N(0,1). \tag{31}$$

Theorem 4 gives the variance calculation for the test of extremal correlation under the GARCH assumption.

Remarks. Other tests can be constructed based on the asymptotics of Theorems 3 and 4. For example, tests for contagion can be constructed as a structural change test on ψ. The null hypothesis for contagion is whether there is a significant break in correlation, $H_0 : \psi_t = \psi \ \forall \ t \in (0, T)$. Following Quintos-Fan-Phillips (2001), we can use the rolling test for structural change (the recursive and sequential tests have consistency problems, see Theorem 4). We let w_t denote the size of our rolled sample, $m_{w_t} = [\kappa w_t]$ for $0 < \kappa < 1$ denotes the number of order statistics used to calculate Hill's estimator, and $\hat{\psi}_t$ is the extremal correlation estimate until time t. The rolling estimator fixes the subsample size w_t and estimates α using w_t rolled through time. Let $\gamma_0 \in (0,1)$ denote the fraction of the fixed sample length and restrict $r \in (\gamma_0, 1)$. The calculation of the rolling correlation starts from $t_0 = [T(r - \gamma_0)] + 1$ so each subsample is of length $w_t^* = t - t_0 + 1 = [T\gamma_0]$. The extremal correlation estimator using sample w_t^* is denoted as $\hat{\psi}_t^*$. The test statistic is

$$V_T(t) = \max_t \left(\frac{w_t^* m_{w_t^*}}{T}\right)\left(\frac{\hat{\psi}_t^*}{\hat{\psi}_T} - 1\right) \xrightarrow{d} \left(W(r,\gamma_0) - (r-s)\,W(1,1)\right)^2$$

and the critical values are tabulated in Quintos, et al. (2001). The notation $W(r)$ denotes a Wiener process with $t = [Tr]$ and $W(r,\gamma_0) = W(r) - W(s)$ where $w_t/T \to r - s = \gamma_{0.}$

4 Simulations

This section explores the finite sample properties of the extremal correlation test using 5000 Monte Carlo simulations. Section 4.1 considers only the iid case to highlight the main properties of the test. Section 4.2 extends the simulation to the ARCH case. All simulations use a sample size of 1000.

4.1 Size and Power Properties

We generate Pareto tails from the t-distribution with v degrees of freedom. The tail index α in this case corresponds to v. The variance is infinite when $v = 2$ and thinner tails occur as $v \to \infty$. The choice of m in the calculation of α is based on the Hill plot generated by the EVIM software.

Table 1 gives the size and power of the test under the null of no extremal correlation. The size and power are analyzed by generating correlated t distributions with correlation ρ. We give performance statistics for three size levels: 1%, 5% and 10% levels. With no correlation of $\rho = 0$, the test has correct size when the generating process is thick tailed (i.e. $v = 2$). In this best case scenario, the test rejects 5.4% for a correct size of 5%. For whatever correlation level the power falls as v increases since the Pareto assumption is less relevant. Also, given the degrees of freedom, the power increases as dependence increases since the distance from the null is increased. Indeed with $v = 6$ and a correlation of $\rho = .38$, the test correctly rejects the null only 34% of the time at the 5% level as compared to 88% rejection with $\rho = .7$.

4.2 Effects of Serial Dependence

Serial dependence in the data creates size distortion unless the test statistic is properly adjusted for it. Table 2 shows the size distortion in the test when serial dependence is introduced using the ARCH process. The data generating process is

$$\begin{aligned} X_t^2 &= \sigma_t^2 Z_t^2, \\ \sigma_t^2 &= \beta_0 + \lambda X_{t-1}^2, \end{aligned}$$

and Z_t are iid Normals. The parameter β_0 is set to .1 and the parameter that carries the clustering information λ is varied.

We work with the ARCH process since the direct correspondence between the tail slope of X_t and λ has been tabulated in Embrechts et al. (1997) and is repeated here,

λ	.3125	.4	.5	.5773	.6	.7	.8	.9	1	1.57
α	8.00	6.09	4.74	4.00	3.82	3.17	2.68	2.30	2.00	1.00

.

Note that we apply our test to X_t^2 so the true tail index is $2/\alpha$. It is evident that as persistency increases in the data (i.e. as λ increases), the size distortion of the test without the size correction increases while the test with the correction remains close to the true size of the test.

5 Conclusion

Correlation during periods of extreme market movements is of interest in portfolio selection and risk management. This paper provides the distribution of the extremal correlation estimator that is applicable to GARCH data. The asymptotics used to derive the distribution is based on extreme value theory. The tests based on our correlation estimator are widely applicable since the theory requires little knowledge of the full data generating distribution. It only assumes that the limiting tail distribution is heavy tailed which is a condition satisfied by most financial time series.

6 Proofs

Proof of Lemma 1:

To prove (13), from Hall-Welsh (1985, equations 3.9 and 3.10) we make use of the following representation,

$$\log X_{(.)} = -\alpha^{-1} Y_{(.)} - \alpha^{-1} d \exp\left(-\rho Y_{(.)}\right) + o_p(1),$$

where $\rho = \beta/\alpha$, $Y_{(n)} = \sum_{j=1}^{T-n+1} Z_j / T - j + 1$ and Z are independent exponential random variables with mean 1. Assuming $m = o(T)$ the second term in the right hand side is $o_p(1)$ (see Quintos-Fan-Phillips, 2001, proof of Lemma 10, equation 32) so

$$\begin{aligned} \log X_{(m+1)} - \log X_{(t)} &\sim -\alpha^{-1} Y_{(m+1)} + \alpha^{-1} Y_{(t)} \\ &= \alpha^{-1} \textstyle\sum_{j=T-m+1}^{T-t+1} Z_j / T - j + 1. \end{aligned} \tag{32}$$

Using the transformation $x \to x^{-1}$, we have

$$F(x) = c_\alpha x^\alpha \left(1 + d x^\beta + o\left(x^\beta\right)\right) \text{ as } x \downarrow 0,$$

and so (12) becomes

$$\hat{\alpha}^{-k}(p) = \frac{1}{k! m} \sum_{t=1}^{m} \left(\log\left(X_{(m+1)}^p / X_{(t)}^p\right)\right)^k.$$

From (32),

$$X_{(m+1)} / X_{(t)} = \exp\left(\alpha^{-1} \sum_{j=T-m+1}^{T-t+1} Z_j / T - j + 1\right), \tag{33}$$

and hence we can write

$$\begin{aligned} \hat{\alpha}^{-k}(p) &= \tfrac{1}{k! m} \textstyle\sum_{t=1}^{m} \left(\log\left(X_{(m+1)}^p / X_{(t)}^p\right)\right)^k \\ &= \tfrac{1}{k! m} \textstyle\sum_{t=1}^{m} \left(p\alpha^{-1} \sum_{j=T-m+1}^{T-t+1} Z_j / T - j + 1\right)^k + o_p(1) \text{ using (33)} \\ &= p^k \alpha^{-k} \left[\tfrac{1}{k! m} \textstyle\sum_{t=1}^{m} \left(\sum_{j=T-m+1}^{T-t+1} Z_j / T - j + 1\right)^k\right] + o_p(1). \end{aligned} \tag{34}$$

Now for each $t = 1, ..., m$ the term $\left(\sum_{j=T-m+1}^{T-t+1} Z_j/T-j+1\right)^k$ is

$$\begin{array}{ll} t=1: & \sum_{k_1,...,k_m} \frac{k!}{k_1!...k_m!} \left(\frac{Z_{T-m+1}}{m}\right)^{k_1} ... \left(\frac{Z_T}{1}\right)^{k_m}, \\ t=2: & \sum_{k_1,...,k_{m-1}} \frac{k!}{k_1!...k_{m-1}!} \left(\frac{Z_{T-m+1}}{m}\right)^{k_1} ... \left(\frac{Z_{T-1}}{2}\right)^{k_{m-1}}, \\ \vdots & \vdots \\ t=m: & \left(\frac{Z_{T-m+1}}{m}\right)^k, \end{array}$$

so there are m terms for the case $k_i = k$, $i = 1, \ldots, m$. Therefore, we have

$$\sum_{t=1}^{m} \left(\sum_{j=T-m+1}^{T-t+1} Z_j/T-j+1 \right)^k = \sum_{t=1}^{m} \left(\sum_{j=T-m+1}^{T-t+1} \frac{Z_j^k}{(T-j+1)^k} \right) + o_p\left(m^{-1}\right),$$

and from (34),

$$\begin{aligned} \hat{\alpha}^{-k}(p) &= \frac{p^k}{k!\alpha^k} \frac{1}{m} \sum_{t=1}^{m} \left(\sum_{j=T-m+1}^{T-t+1} \frac{Z_j^k}{(T-j+1)^k} \right) + o_p(1) \\ &= \frac{p^k}{k!\alpha^k} \frac{1}{m} \sum_{j=T-m+1}^{T} \frac{Z_j^k}{(T-j+1)^{(k-1)}} + o_p(1), \; k>1 \\ &\xrightarrow{p} \frac{p^k k!}{k!\alpha^k} = p^k \alpha^{-k}, \end{aligned} \tag{35}$$

where the last line follows from Hall (1978, equation 14) and by noting that $E\left(Z^k\right) = k!$.

To prove (14), we write from (35),

$$\begin{aligned} m^{1/2}\left(\hat{\alpha}^{-k}(p) - p^k\alpha^{-k}\right) &= \frac{p^k}{k!\alpha^k} \frac{1}{\sqrt{m}} \sum_{j=T-m+1}^{T} \frac{Z_j^k}{(T-j+1)^{(k-1)}} - m^{1/2} p^k \alpha^{-k} + o_p(1) \\ &\quad \text{from line 2 of (35)} \\ &\sim p^k \alpha^{-k} m^{-1/2} \sum_{j=T-m+1}^{T} \frac{Z_j^k - k!}{k!(T-j+1)^{(k-1)}} + o_p(1) \\ &\xrightarrow{d} p^k \alpha^{-k} N(0,1) \end{aligned}$$

as required since $m^{-1/2} \sum_{j=T-m+1}^{T} \frac{Z_j^k - k!}{k!(T-j+1)^{(k-1)}} \xrightarrow{d} N(0,1)$ from Hall (1978, equation 14). □

Proof of Lemma 2:

Consistency and asymptotic Normality holds so long as the process is stationary. We need only to derive the variance components. In particular, we work with extending equation (3.2) of Hsing (1991) to the case of $k, p > 1$,

$$\sqrt{m}\left[\left(\hat{\alpha}^{-k}(p) - \alpha^{-k}(p)\right) + \frac{p^k}{k!\alpha^k}\left(S_T - E\left(S_T\right)\right)\right], \tag{36}$$

where $S_T = \frac{1}{m}\sum_{t=1}^{T} 1\left(\log Y_t > \log Y_{(T-m)}\right)$. The variance components are then (see page 1558 of

Hsing, 1991),

$$\begin{aligned}
&\tfrac{1}{m}Var\left(\textstyle\sum_{j=1}^{T}C_j\right)\\
&=\tfrac{1}{(k!)^2 m}\left(Var\left(C_1\right)+2\sum_{1\le i\le j\le T}\sum Cov\left(C_i,C_j\right)\right)\\
&\to\tfrac{1}{(k!)^2}\left(\left(\tfrac{p}{\alpha}\right)^{2k}(2k)!+2\sum_{1\le i\le j\le T}\sum Cov\left(C_i,C_j\right)\right)\\
&=\tfrac{1}{(k!)^2}\left(\tfrac{p}{\alpha}\right)^{2k}\left((2k)!+\chi\right),
\end{aligned}\tag{37}$$

where $\chi=2\left(\frac{\alpha}{p}\right)^{2k}\sum_{1\le i\le j\le T}\sum Cov\left(C_i,C_j\right)<\infty$. Also for the second term,

$$\begin{aligned}
&\left(\tfrac{p^k}{k!\alpha^k}\right)^2\tfrac{1}{m}Var\left(\textstyle\sum_{t=1}^{T}1\left(\log Y_t>\log Y_{(T-m)}\right)\right)\\
&=\left(\tfrac{p^k}{k!\alpha^k}\right)^2\left[Var\left(D_1\right)+2\sum_{1\le i\le j\le T}\sum Cov\left(D_i,D_j\right)\right]\\
&\to\left(\tfrac{p^k}{k!\alpha^k}\right)^2(1+\omega)<\infty,
\end{aligned}\tag{38}$$

and the covariance term,

$$\begin{aligned}
&\left(\tfrac{p}{\alpha}\right)^k\tfrac{2}{(k!)^2 m}Cov\left(\textstyle\sum_{i=1}^{T}C_i,\sum_{j=1}^{T}Dj\right)\\
&=\left(\tfrac{p}{\alpha}\right)^k\tfrac{2}{(k!)^2 m}\left[Cov\left(C_1D_1\right)+\sum_{1\le i\le j\le T}\sum Cov\left(C_iD_j\right)+Cov\left(D_iC_j\right)\right]\\
&\to\left(\tfrac{p}{\alpha}\right)^k\tfrac{2}{(k!)^2}\left[\tfrac{k!p^k}{\alpha^k}+\sum_{1\le i\le j\le T}\sum Cov\left(C_iD_j\right)+Cov\left(D_iC_j\right)\right]\\
&=\left(\tfrac{p}{\alpha}\right)^{2k}\tfrac{2}{(k!)^2}\left[k!+\phi\right],
\end{aligned}\tag{39}$$

where $\phi=\left(\frac{\alpha}{p}\right)^k\sum_{1\le i\le j\le T}\sum Cov\left(C_iD_j\right)+Cov\left(D_iC_j\right)<\infty$.

The result follows since the variance of (36) is a combination of (37), (38), and (39),

$$\begin{aligned}
&\tfrac{1}{(k!)^2}\left(\tfrac{p}{\alpha}\right)^{2k}\left((2k)!+\chi\right)+\left(\tfrac{p^k}{k!\alpha^k}\right)^2(1+\omega)-\left(\tfrac{p}{\alpha}\right)^{2k}\tfrac{2}{(k!)^2}\left(k!+\phi\right)\\
&=\tfrac{1}{(k!)^2}\left(\tfrac{p}{\alpha}\right)^{2k}\left[(2k)!+\chi+1+\omega-2k!-2\phi\right].
\end{aligned}$$

Finiteness of χ,ω and ϕ follows the arguments contained in Quintos, et al. (2001, Theorem 8). □

Proof of Theorem 3:

Let t_1,t_2 and l denote the position of X_1,X_2 and Y in $(0,m_1)$, $(0,m_2)$, $(0,L)$ respectively so that $t_1=[\kappa_1 m_1],t_2=[\kappa_2 m_2]$ and $l=[\kappa L]$. The notation $[x]$ denotes the integer part of x. Since $L=\min(m_1,m_2)$, we let $L=[\lambda_1 m_1]$ or $L=[\lambda_2 m_2]$ with $\lambda_1,\lambda_2\le 1$. The relation of t_1 and t_2 to l is then $t_1=\left[\frac{\kappa_1}{\kappa\lambda_1}l\right],t_2=\left[\frac{\kappa_2}{\kappa\lambda_2}l\right]$.

We write the summation indexed by l,

$$\begin{aligned}\hat{\alpha}_{12}^{-1} &= \tfrac{1}{L}\sum_{l=1}^{L} Y_{(l)} \\ &= \tfrac{1}{L}\sum_{l=1}^{L}\left(\log X_{1(t_1)} - \log X_{1(T-m_1+1)}\right)_+ \left(\log X_{2(t_2)} - \log X_{2(T-m_2+1)}\right)_+ .\end{aligned} \tag{40}$$

Now make the transformation $x \to x^{-1}$and (40) becomes

$$\hat{\alpha}_{12}^{-1} = \frac{1}{L}\sum_{l=1}^{L}\left(\log X_{1(m_1+1)} - \log X_{1(t_1)}\right)_+ \left(\log X_{2(m_2+1)} - \log X_{2(t_2)}\right)_+ . \tag{41}$$

To prove (24) we write $\hat{\alpha}_{12}^{-1}$ in terms of standard exponential variables Z,

$$\begin{aligned}\hat{\alpha}_{12}^{-1} &= \left(\tfrac{1}{\alpha_1\alpha_2}\right)\tfrac{1}{L}\sum_{l=0}^{L-1}\left[\left(\sum_{j=1}^{T-t_1}\tfrac{Z_{1j}}{T-j+1} - \sum_{j=1}^{T-m_1}\tfrac{Z_{1j}}{T-j+1}\right)\left(\sum_{j=1}^{T-t_2}\tfrac{Z_{2j}}{T-j+1} - \sum_{j=1}^{T-m_2}\tfrac{Z_{2j}}{T-j+1}\right)\right] \\ &= \left(\tfrac{1}{\alpha_1\alpha_2}\right)\tfrac{1}{L}\sum_{l=0}^{L-1}\left[\left(A_1 - A_2\right)\left(B_1 - B_2\right)\right].\end{aligned} \tag{42}$$

Consider first expanding $\sum_{l=0}^{L-1}(A_1 - A_2)$ assuming for now $B_1 = B_2 = 0$ and $m_1 \leq m_2$. In expanding the summation over l, notice that for each l, the term A_1 is

$$\begin{aligned} l = 0: &\quad \tfrac{Z_{11}}{T} + \tfrac{Z_{12}}{T-1} + ... + \tfrac{Z_{1T-m_1}}{m_1+1} + ... + \frac{Z_{1T-\left[\frac{\kappa_1}{\kappa\lambda_1}\right]}}{\left[\frac{\kappa_1}{\kappa\lambda_1}\right]+1} + ... + \tfrac{Z_{1T}}{1}, \\ l = 1: &\quad \tfrac{Z_{11}}{T} + \tfrac{Z_{12}}{T-1} + ... + \tfrac{Z_{1T-m_1}}{m_1+1} + ... + \frac{Z_{1T-\left[\frac{\kappa_1}{\kappa\lambda_1}\right]}}{\left[\frac{\kappa_1}{\kappa\lambda_1}\right]+1}, \\ \vdots &\quad \vdots \\ l = L-1: &\quad \tfrac{Z_{11}}{T} + \tfrac{Z_{12}}{T-1} + ... + \tfrac{Z_{1T-m_1}}{m_1+1}.\end{aligned}$$

Thus we can write

$$\sum_{l=0}^{L-1} A_1 = L\left(\sum_{j=1}^{T-m_1}\frac{Z_{1j}}{T-j+1}\right) + \sum_{j=T-m_1+1}^{T} Z_{1j},$$

and so since $L = m_1$

$$\begin{aligned}\tfrac{1}{L}\sum_{l=0}^{L-1} A_1 - A_2 \quad & \sum_{j=1}^{T-m_1}\tfrac{Z_{1j}}{T-j+1} - A_2 + \tfrac{1}{L}\sum_{j=T-L+1}^{T} Z_{1j} \\ &= o_p(1) + \tfrac{1}{L}\sum_{j=T-L+1}^{T} Z_{1j} \\ &\xrightarrow{p} E(Z_1).\end{aligned}$$

Returning to (42), we are concerned with the nonzero terms of A_1 and B_1,

$$\begin{aligned}\hat{\alpha}_{12}^{-1} &= \tfrac{1}{\alpha_1\alpha_2}\tfrac{1}{L}\sum_{l=0}^{L-1}\left[\left(A_1 - A_2\right)\left(B_1 - B_2\right)\right] \\ &= \tfrac{1}{\alpha_1\alpha_2}\tfrac{1}{L}\sum_{j=T-L+1}^{T} Z_{1j}Z_{2j} + o_p(1) \\ &\xrightarrow{p} \alpha_1^{-1}\alpha_2^{-1}E\left(Z_1 Z_2\right)\end{aligned} \tag{43}$$

as required.

To prove (25), we have the following three steps:

1. We show that $\hat{\alpha}_{12}^{-1} \xrightarrow{p} \alpha_1^{-1}\alpha_2^{-1}$ under H_0. Under H_0, $E(Z_1Z_2) = E(Z_1)E(Z_2) = 1$ given our assumption on Z_i as standard expoonential variables with mean 1. To show $\hat{\alpha}_{12}^{-1} \xrightarrow{p} 2\alpha_1^{-1}\alpha_2^{-1}$ we can use Lemma 1 or (43). Perfect correlation is given by $\left(X_2 - X_{2(T-m_2)}\right) = a\left(X_1 - X_{1(T-m_1)}\right)$, say. Then $E(Z_1Z_2) = aE\left(Z_1^2\right) = a2!$ for any $a \neq 0$ using (43). Since the coefficient a cancels out in the construction of the correlation coefficient we take $a = 1$ without loss of generality. The same result can be derived using Lemma 1 by setting $p = 2$ since $\left(X_2 - X_{2(T-m_2)}\right) = \left(X_1 - X_{1(T-m_1)}\right)$ under perfect correlation.

2. The result for $\hat{\sigma}_{\xi_{i+}}^{*} \xrightarrow{p} \alpha_i^{-1}$ follows from Lemma 1,

$$\hat{\sigma}_{\xi_{i+}}^{2*} = \frac{1}{L}\sum_{t=1}^{T}\xi_{it+}^{2*} - \left[\frac{1}{L}\sum_{t=1}^{T}\xi_{it+}^{*}\right]^2 \xrightarrow{p} \left(\frac{1}{\alpha_i}\right)^2 2! - \left(\left(\frac{1}{\alpha_i}\right)^1 1!\right)^2 = \alpha_i^{-2}, \tag{44}$$

   with $p = 1$, $k = 2$ for the first term, and $k = 1$ for the second term.

3. To prove $L^{1/2}\left(\hat{\psi} - \psi\right) \xrightarrow{d} N(0,1)$, first note that the numerator of $\hat{\psi}$,

$$\hat{\psi}_{num} = \hat{\alpha}_{12}^{-1} - \hat{\alpha}_1^{-1}\hat{\alpha}_2^{-1} \xrightarrow{p} \alpha_1^{-1}\alpha_2^{-1}\left(E(Z_1Z_2) - 1\right) = 0$$

   under H_0. Also,

$$\begin{aligned}\hat{\psi}_{num} &= \alpha_1^{-1}\alpha_2^{-1}\tfrac{1}{L}\textstyle\sum_{j=T-L+1}^{T} Z_{1j}Z_{2j} - \left(\tfrac{1}{\alpha_1 L}\sum_{j=T-L+1}^{T} Z_{1j}\right)\left(\tfrac{1}{\alpha_2 L}\sum_{j=T-L+1}^{T} Z_{2j}\right) + o_p(1) \\ &= \alpha_1^{-1}\alpha_2^{-1}\tfrac{1}{L}\left[\textstyle\sum_{j=T-L+1}^{T}\left(Z_{1j}^2 - \bar{Z}_1^2\right)\right] + o_p(1).\end{aligned} \tag{45}$$

   By the CLT,

$$L^{1/2}\alpha_1\alpha_2\left(\hat{\psi}_{num}\right) \xrightarrow{d} N(0,1). \tag{46}$$

   The result follows by noting that $\hat{\sigma}_{\xi_{i+}}^{*} \xrightarrow{p} \alpha_i^{-1}$ so

$$L^{1/2}\left(\hat{\psi}\right) \overset{d}{=} L^{1/2}\hat{\sigma}_{\xi_{1+}}^{*-1}\hat{\sigma}_{\xi_{2+}}^{*-1}\left(\hat{\psi}_{num}\right) \xrightarrow{d} N(0,1)$$

   from (44) and (46). □

Proof of Theorem 4:

To prove (30) we follow the proof of Lemma 2. As in (36) we are concerned with the variance of

$$\sqrt{L}\left[\left(\hat{\alpha}_{12}^{-1}(2) - \alpha_{12}^{-1}(2)\right) + \frac{1}{\alpha_1\alpha_2}\left(S_T' - E\left(S_T'\right)\right)\right], \tag{47}$$

where $S'_T = \frac{1}{L}\sum_{t=1}^{T} 1\left(\log Y_{1t} > \log Y_{1(T-m_1)}, \log Y_{2t} > \log Y_{2(T-m_2)}\right)$. The variance terms are calculated from

$$C_t = \log\left(\frac{Y_{1t}^2}{Y_{1(T-m_i)}^2}\right) 1_{(\xi_{1t}>1)} \log\left(\frac{Y_{2t}^2}{Y_{2(T-m_i)}^2}\right) 1_{(\xi_{2t}>1)},$$
$$D_t = 1\left(\log Y_{1t} > \log Y_{1(T-m_1)}, \log Y_{2t} > \log Y_{2(T-m_2)}\right),$$

where the first term is

$$\begin{aligned} & \tfrac{1}{L} Var\left(\textstyle\sum_{j=1}^{T} C_j\right) \\ &= \tfrac{1}{L}\left(Var\left(C_1\right) + 2\sum_{1\le i\le j\le T}\sum Cov\left(C_i, C_j\right)\right) \\ &\to \left(\tfrac{2E(Z_1Z_2)^2}{(\alpha_1\alpha_2)^2} + 2\sum_{1\le i\le j\le T}\sum Cov\left(C_i, C_j\right)\right) \\ &= \tfrac{1}{(\alpha_1\alpha_2)^2}\left(2 + 2\left(\alpha_1\alpha_2\right)^2\sum_{1\le i\le j\le T}\sum Cov\left(C_i, C_j\right)\right) = \tfrac{1}{(\alpha_1\alpha_2)^2}\left(2+\chi_0\right) \end{aligned} \tag{48}$$

under H_0. The second term is

$$\begin{aligned} & \tfrac{1}{(\alpha_1\alpha_2)^2}\tfrac{1}{L} Var\left(\textstyle\sum_{t=1}^{T} D_t\right) \\ &= \tfrac{1}{(\alpha_1\alpha_2)^2}\left[Var\left(D_1\right) + 2\sum_{1\le i\le j\le T}\sum Cov\left(D_i, D_j\right)\right] \\ &\to \tfrac{1}{(\alpha_1\alpha_2)^2}\left(1+\omega_0\right), \end{aligned} \tag{49}$$

and the final term can be written as

$$\begin{aligned} & \tfrac{2}{(\alpha_1\alpha_2)^2}\tfrac{1}{L} Cov\left(\textstyle\sum_{i=1}^{T} C_i, \sum_{j=1}^{T} Dj\right) \\ &= \tfrac{2}{(\alpha_1\alpha_2)^2}\tfrac{1}{L}\left[Cov\left(C_1D_1\right) + \sum_{1\le i\le j\le T}\sum Cov\left(C_iD_j\right) + Cov\left(D_iC_j\right)\right] \\ &\to \tfrac{2}{(\alpha_1\alpha_2)^2}\left[1 + \left(\alpha_1\alpha_2\right)^2\sum_{1\le i\le j\le T}\sum Cov\left(C_iD_j\right) + Cov\left(D_iC_j\right)\right] \\ &= \tfrac{2}{(\alpha_1\alpha_2)^2}\left[1+\phi_0\right]. \end{aligned} \tag{50}$$

The result for ζ_0 follows by combining (48), (49) and (50). □

7 References

Bae, K.H., Karolyi, G.A., Stulz, R.M., 2001, A new approach to measuring financial contagion, *Review of Financial Studies*, forthcoming.

Coles, S.G., J. Hefferman and J.A. Tawn, 1999, Dependence Measures for Extreme Value Analysis, *Extremes* 3, 5–38.

Danielsson, J. and C.G. de Vries, 1997a, Beyond the Sample: Extreme Quantile and Probability Estimation, Working paper, Tinbergen Institute, Rotterdam.

Danielsson, J. and C.G. de Vries, 1997b, Tail Index and Quantile Estimation with very high Frequency Data, *Journal of Empirical Finance* 4, 241–257.

Davis, R.D. and T. Mikosch, 1998, The Sample ACF of Heavy-Tailed Stationary Processes with Applications to ARCH, *Annals of Statistics* 26, 2049–2080.

DuMouchel, W.H., 1983, Estimating the Stable Index α in Order to Measure Tail Thickness: A Critique, *Annals of Statistics* 11, 1019–1031.

Embrechts, P., Kluppelberg, C. and T. Mikosch, 1997, Modelling Extremal Events. Springer Verlag.

Forbes, K., Rigobon, R., 2002, No contagion, only interdependence: measuring stock market co-movements, *Journal of Finance* 57, 2223–2261.

Hall, P., 1978, Representations and Limit Theorems for Extreme Value Distributions, *Journal of Applied Probability* 15, 639–644.

Hall, P., 1982, On Some Simple Estimates of an Exponent of Regular Variation," *Journal of the Royal Statistical Society* B 44, 37–42.

Hall P. and A.H. Welsh, 1985, Adaptive Estimates of Parameters of Regular Variation, *Annals of Statistics* 13, 331–341.

Hartmann, P., S. Straetmans and C.G. de Vries, 2001, Asset Market Linkages in Crisis Periods, Working paper, Center for Economic Policy Research.

Hwang, S. and M. Salmon, 2002, An Analysis of Performance Measures Using Copulae, Performance Measurement in Finance: Firms, Funds and Managers. Butterworth-Heineman, London.

Hill, B.M., 1975, A Simple General Approach to Inference about the Tail of a Distribution, *Annals of Statistics* 3, 1163–1174.

Hsing, T., 1991, On Tail Index Estimation Using Dependent Data, *Annals of Statistics* 19, 1547–1569.

Hsing, T., 1993, Extremal Index Estimation for a Weakly Dependent Stationary Sequence, *Annals of Statistics* 21, 2043–2071.

Longin, F., Solnik, B., 2001, Extreme correlation of international equity markets, *Journal of Finance* 56, 649–676.

Mikosch, T. and C. Starica, 2000, Limit Theory for the Sample Autocorrelations and Extremes of a GARCH(1,1) Process, *Annals of Statistics* 28, 1427–1451.

Patton, A., 1999, Modeling Time-Varying Exchange Rate Dependence Using the Conditional Copula, Working paper, University of California, San Diego.

Poon, S., M. Rockinger and J. Tawn, 2001, New Extreme Value Dependence Measures and Finance Applications, working paper, Les Cahiers de Recherche, Groupe HEC.

Quintos, C., Z. Fan and P.C.B. Phillips, 2001, Structural Change Tests in Tail Behavior and the Asian Crisis, *Review of Economic Studies* 68, 633–663.

Resnick, S., 1998, Why Nonlinearities can Ruin the Heavy-Tailed Modeler's Day, A Practical Guide to Heavy Tails. Boston: Birkhauser.

Resnick, S. and C. Starica, 1995, Consistency of Hill's Estimator for Dependent Data, *Journal of Applied Probability* 32, 129–167.

Resnick, S. and C. Starica, 1998, Tail Index Estimation for Dependent Data, *Annals of Applied Probability* 8, 1156–1183.

Samorodnitsky, G. and M. Taqqu, 1994, Stable Non-Gaussian Random Processes. New York: Chapman & Hall.

Starica, C., 1999, Multivariate Extremes for Models with Constant Conditional Correlations, *Journal of Empirical Finance* 6, 515–553.

Starica, C., 2000, On the Tail Empirical Process of Solutions of Stochastic Difference Equations, Working paper, Chalmers University.

Table 1
Size and Power of Tail Dependence Test

The data generating process used for the table below are correlated t distributions each with υ degrees of freedom. Size and Power are given for three size levels – 1%, 5% and 10% levels. The variable ρ denotes the correlation of the t random variables.

	$\rho = 0$	$\rho = .38$	$\rho = .7$
	$\hat{\psi}$	$\hat{\psi}$	$\hat{\psi}$
$\upsilon = 2$			
.01	.006	.2258	.8742
.05	.054	.3642	.8904
.10	.106	.5783	.9331
$\upsilon = 4$			
.01	.008	.1961	.8691
.05	.063	.3544	.9003
.10	.112	.5390	.9291
$\upsilon = 6$			
.01	.014	.1914	.8548
.05	.071	.3432	.8804
.10	.118	.5368	.9152

Table 2
Effects of Serial Dependence on Size of the Test

The data generating process used for the table below are uncorrelated ARCH processes. The ARCH process is generated as

$$X_t^{\ 2} = \sigma_t^{\ 2} Z_t^{\ 2}$$
$$\sigma^2{}_t = \beta_0 + \lambda X_{t-1}^{\ 2}\,.$$

The parameter β_0 is set to .1 and λ is varied. The results for $\hat{\psi}$ with and without the serial dependence correction are given.

	Without correction	With correction
$\lambda = .4$		
.01	.0187	.0064
.05	.0856	.067
.10	.1252	.1087
$\lambda = .6$		
.01	.0231	.0065
.05	.0889	.0621
.10	.1268	.1050
$\lambda = .8$		
.01	.0296	.0057
.05	.0974	.0583
.10	.1337	.0965