Gaussian Tests of "Extremal White Noise" for Dependent, Heterogeneous, Heavy Tailed Stochastic Processes with an Application*

Jonathan B. Hill†
Dept. of Economics
Florida International University

Original Version: Oct. 2005
This Version:

November 5, 2006

Abstract

We develop a non-parametric test of tail-specific extremal serial dependence for possibly heavy-tailed time series. The test statistic is asymptotically chi-squared under a null of "extremal white noise", as long as extremes of the time series are Near-Epoch-Dependent on the extremes of some mixing process. The theory covers ARFIMA, FIGARCH, bilinear, and Extremal Threshold processes, and a wide range of nonlinear distributed lags. In this setting the test statistic obtains an asymptotic power of one under the alternative.

Of separate interest, we deliver a joint distribution limit for an arbitrary vector of tail index estimators under extraordinarily general conditions, complete with a consistent kernel estimator of the covariance matrix.

We apply tail specific tests to equity market and exchange rate returns data.

*I would like to thank the participants of the European Meeting of Statisticians in Olso, 2005, participants of the Econometric Society World Congress in London, 2005, and Oliver Linton for helpful discussions. Finally, I would like to thank two anonymous referees, an Associate Editor, and Peter Robinson for detailed suggestions that lead to substantial improvements. All errors remain mine alone.

†Dept. of Economics, Florida International University, Miami, FL; www.fiu.edu/~hilljona; jonathan.hill@fiu.edu.

JEL classifications: C12, C16, C52.

Keywords: extremal dependence; white-noise; near-epoch-dependence; regular variation; infinite variance; portmanteau test.

1. Introduction We develop a test of extremal serial dependence applicable to possibly heavy-tailed, dependent and heterogenous stochastic processes. The null hypothesis of interest is "*extremal white noise*": extreme values of X_{t-h} are not predominantly followed by extremes in X_t for all $h = 1, 2, ...$ An extreme value occurs when an observation X_t surpasses a threshold (positive or negative) as the threshold diverges to infinity with the sample size. Extreme-value-theory has been applied to the analysis of hyperinflation, asset market bubbles, and exchange rate contagion; and has gained enormous popularity in the telecommunications (network activity), meteorological, geological, electrical and insurance sciences. See Resnick (1987), Embrechts, Klueppelberg, and Mikosch (2003), Beirlant, Goegebeur, Segers, Teugels, and de Waal (2004) and the citations therein.

The concept of "extremal dependence" has been exclusively applied to bivariate processes. See Tawn (1990), Ledford and Tawn (1996, 1997), Stărică (1999), Hefferman and Tawn (2004), and Schmidt and Stadtmüller (2006), to name a few. The predominant approaches involve the assumption of bivariate regular variation, the specification of a multivariate extreme value distribution tail, the use of a copula dependence function, and a transformation of marginal distributions into unit Fréchets. See Section 4 for comparisons with our method. In all cases either the processes are assumed to be marginally serially independent; a GARCH structure is arbitrarily imposed; and/or the limiting distribution of a proposed sample tail dependence estimator is not established. Moreover, serial dependence is ignored entirely in this literature, and memory and heterogeneity restrictions are enforced over the entire distribution support.

Quantile-based methods have bridged the gap between population and extreme-tail dependence analysis (e.g. Linton and Wang 2004). The so-called "extreme quantile regression" method of Drees (2003) and Chernozhukov (2005) has only been developed for *iid* and mixing processes, and only under harsh parametric restrictions (e.g. linearity).

In the present paper we develop the *co-relation* measure of linear serial extremal dependence for possibly heavy-tailed processes. We develop tests of left-, right-, and two-tailed extremal serial dependence for a wide range of heterogenous, dependent, and possibly heavy-tailed processes without imposing parametric structure. We only require the marginal distributions to have regularly varying tails, a less restrictive assumption than the typical joint bivariate regular variation assumption. Cf. Basrak *et al* (2002).

Although the co-relation is only one of many measures of extremal dependence, it serves multiple purposes here. The co-relation does not require a joint tail specification as long as convolutions $X_t \pm X_{t-h}$ satisfy a tail equivalence property that holds for a broad class of linear, nonlinear, conditional volatility and extremal threshold processes. The co-relation naturally measures linear "extreme tail" dependence; the decay properties of the co-relation coincide with population or tail memory properties; and we are able to provide an estimator of the co-relation that is consistent and asymptotically normal under substantially general conditions.

The test statistic itself is a weighted-average portmanteau statistic in the tradition of Ljung and Box (1979). It is essentially irrelevant whether the source of leptokurtosis is unknown, or a parametric model is assumed, a la GARCH, IGARCH, or stable-GARCH. In order to ensure a standard distribution limit, we only require mild restrictions on tail decay and memory which cover at least Extremal Threshold, ARFIMA, FIGARCH, and bilinear processes, and many nonlinear distributed lags. In particular, using new extremal dependence properties developed in Hill (2005) we only require the extremes of the process $\{X_t\}$ to be Near-Epoch-Dependent on the extremes of some mixing process.

Of separate interest we deliver under the above general conditions a joint distribution limit for an arbitrary vector of tail index estimators, complete with a consistent kernel estimator of the covariance matrix. See Theorem 2. To the best of our knowledge, this is the first instance of such theory and the development and use of such a covariance kernel estimator.

We apply two-tailed and difference in tails tests to exchange rate and equity market daily returns, including one emerging market (Shanghai Stock Exchange). We find small levels of significant, symmetric, positive and persistent extremal dependence in the daily returns of the Yen and Euro against the U.S. Dollar, and asymmetric extremal dependence in the British Pound. Symmetric extremal dependence suggests some returns cannot be governed by an asymmetric regime switching process, contrary to popular applications: see Section 9. We find a wider range of extremal dependence characteristics in daily asset market returns, including weak, positive and persistent dependence; and extremely weak, asymmetric dependence.

In Sections 2-4 we characterize distribution tails, develop the co-relation measure of dependence, and compare the co-relation to alternative extremal dependence measures. Preliminary asymptotic theory is contained in Section 5, Section 6 develops the test statistic, and in Section 7 we develop a strategy for selecting the sample tail fractile. Sections 8 and 9 contain a simulation study and the empirical application. Appendix 1 contains proofs of the main results and Appendix 2 contains preliminary results. All table and figures are placed at the end of the paper.

Denote by $\rightarrow$ convergence in probability, and by $\Rightarrow$ convergence with respect to finite dimensional distributions. $[x]$ denotes the integer part of x, $[|x|] \leq |x|$. $z_+ \equiv \max\{z, 0\}$. $1_h \equiv [1, ..., 1] \in \mathbb{N}^h$. $|\cdot|$ denotes the l_1-matrix norm.

2. Regular Variation Let $\{X_t\}$ denote a stochastic process defined on a probability measure space, $(\quad, \Im, P)$, $\Im = \sigma(\cup_t \Im_t)$, $\Im_t = \sigma(X_\tau : \tau \leq t) \subset \Im_{t+1}$. Assume X_t has for each t common regularly varying marginal distribution tails: as $z \rightarrow \infty$

$$(1) \quad P(X_t \leq -\varepsilon) = c_1(x)\varepsilon^{-\alpha_1}(1 + o(1)), \quad P(X_t > \varepsilon) = c_2(x)\varepsilon^{-\alpha_2}(1 + o(1)),$$

where $0 \leq c_i(x) < \infty$ and $c_2(x) > 0$ by convention, and $\alpha_1, \alpha_2 > 0$ denote the tail-specific indices of regular variation. When $\min\{\alpha_1, \alpha_2\} < 2$, the population variance is infinite, however we allow for any $\alpha_i > 0$. The above structure allows for asymmetry in the tails both through extremal dispersion c_i and tail thickness α_i. See Jansen and de Vries (1991) and Jondeau and Rockinger (2003) for evidence of asymmetric tail thickness in equity markets. Notice that we do not make any reference to the finite distributions of $\{X_t, X_{t-1}, ..., X_{t-h}\}$, $h > 1$, nor the non-extremal support of X_t.

Processes that abide (1) belong to the normal domain of attraction of the stable laws, include the maximum domain of attraction of extreme value distributions, and include many stochastic recurrence equations. See Bingham *et al* (1987), Resnick (1987), and Basrak *et al* (2002). Gaussian distributions do not satisfy (1): see Resnick (1987). The tail form (1) has been used widely in the statistics and econometrics literatures. See, for example, Hall (1982) and Caner (1998) to name a very few.

We require two convolution processes

$$Y_{h,t} \equiv X_t + X_{t-h} \quad \text{and} \quad \tilde{Y}_{h,t} \equiv X_t - X_{t-h}.$$

We assume $Y_{h,t} \sim (1)$ with the same indices $\{\alpha_1, \alpha_2\}$ and scales $0 \leq \{c_1(y_h), c_2(y_h)\} < \infty$, and $\tilde{Y}_{h,t} \sim (1)$ with common left/right index $\alpha_0 \equiv \min\{\alpha_1, \alpha_2\}$ and scales $0 \leq \{c_1(\tilde{y}_h), c_2(\tilde{y}_h)\} < \infty$. Any α-stable process satisfies such convolution tail equivalence, cf. Ibragimov and Linnik (1971). In general it suffices for X_t to have an infinite order distributed lag representation with independent innovations that satisfy (1), including at least ARFIMA, FIGARCH, bilinear and extremal threshold processes. Consult the technical appendix to this paper in Hill (2006a)[1].

3. Measures of Extremal Dependence We now develop and characterize the co-relation coefficient.

3.1 Two-Tailed Co-Relation

We define the *two-tailed co-relation* coefficient at displacement $h > 0$, $\rho^{(0)}(h) = \rho^{(0)}(X_t, X_{t-h})$, as a scaled ratio of extreme tail probabilities

$$\rho^{(0)}(h) \equiv 1 - \lim_{\varepsilon \to \infty} \frac{P(|\tilde{Y}_{h,t}| > \varepsilon)}{[P(|X_t| > \varepsilon) + P(|X_{t-h}| > \varepsilon)]}.$$

For any $Z_t \sim (1)$ define

$$c_0(z) \equiv c_1(z) I(\alpha_1 \leq \alpha_2) + c_2(z) I(\alpha_2 \leq \alpha_1).$$

From (1), the assumption that X_t has the same tail structure for all t, and properties of regularly varying functions (e.g. Feller 1971, Resnick 1987) we obtain

$$\begin{aligned} \lim_{\varepsilon \to \infty} \varepsilon^{\alpha_0} P(|Z_t| > \varepsilon) &= \lim_{\varepsilon \to \infty} \varepsilon^{\alpha_0 - \alpha_1} \varepsilon^{\alpha_1} P(Z_t < -\epsilon) + \varepsilon^{\alpha_0 - \alpha_2} \varepsilon^{\alpha_2} P(Z_t > \varepsilon) \\ &= \lim_{\varepsilon \to \infty} \varepsilon^{\alpha_0 - \alpha_1} c_1(z) + \lim_{\varepsilon \to \infty} \varepsilon^{\alpha_0 - \alpha_2} c_2(z) + o(1) = c_0(z). \end{aligned}$$

[1] The only extant literature on the topic that we are aware of concerns convolutions of stable random variables, or *iid* variates with regularly varying tails. See, for example, Embrechts and Goldie (1982), Cline (1986), Datta and McCormick (1998), and Geluk, Liang and de Vries (2000).

Therefore

$$(2)\qquad \rho^{(0)}(h) = 1 - \lim_{\varepsilon\to\infty} \frac{\varepsilon^{\alpha_0} P(|\tilde{Y}_{h,t}| > \varepsilon)}{2\varepsilon^{\alpha_0} P(|X_t| > \varepsilon)} = 1 - \frac{c_0(\tilde{y}_h)}{2c_0(x)},$$

If X_t has symmetric tails a la $\alpha_1 = \alpha_2$ then $c_0(x) = c_1(x) + c_2(x)$, hence by construction $c_0(\tilde{y}_h) = c_1(\tilde{y}_h) + c_2(\tilde{y}_h)$. For serially independent processes it easy to show $c_i(\tilde{y}_h) = c_0(x)$ for each $i = 1, 2$, hence $\rho^{(0)}(h) = 0\ \forall h \neq 0$. See Breiman (1965), Feller (1971) and Cline (1986).

Suppose $\min\{\alpha_1, \alpha_2\} > 1$ (see Section 3.3.2 for the case $\alpha_i \leq 1$). If extremes of X_{t-h} are predominantly followed by extremes of X_t with the same sign, then $|X_t - X_{t-h}|$ will have a comparatively low tail dispersion $c_i(\tilde{y}_h) < c_0(x)$ for each $i = 1, 2$, hence $\rho^{(0)}(h) > 0$. Similarly, if extremes of X_{t-h} are predominantly followed by extremes of X_t with the opposite sign, then $|X_t - X_{t-h}|$ will have comparatively large tail dispersion $c_i(\tilde{y}_h) > c_0(x)$, hence $\rho^{(0)}(h) < 0$. In this sense the co-relation measures *linear* extremal dependence. See Section 3.3.4 for a comparison with the correlation.

In general the property $\rho^{(0)}(h) = 0$ is interesting in its own right.

Extremal White Noise *If $\rho^{(0)}(h) = 0$ for all displacements $h \geq 1$, we say $\{X_t\}$ is two-tailed* extremal white noise.

An *iid* process is trivially extremal white noise. However, processes $\{X_t\}$ with a switching or threshold mechanism in which non-extremes are highly persistent and extremes are independent are also extremal white noise. See Section 3.4 for examples.

3.2 Co-Relation Tails

The two-tailed co-relation cannot reveal whether extremal serial dependence is more pronounced when X_{t-h} obtains a large negative or positive value. For tail-specific measures we use the convolution summation $Y_{h,t} \equiv X_t + X_{t-h}$. Convolution tail equivalence implies

$$\begin{aligned}
\rho^{(1)}(h) &= 1 - \lim_{\varepsilon\to\infty} \frac{P(Y_{h,t} < -\varepsilon)}{[P(X_t < -\varepsilon) + P(X_{t-h} < -\varepsilon)]} = c_1(y_h)/2c_1(x) - 1 \\
\rho^{(2)}(h) &= 1 - \lim_{\varepsilon\to\infty} \frac{P(Y_{h,t} > \varepsilon)}{[P(X_t > \varepsilon) + P(X_{t-h} > \varepsilon)]} = c_2(y_h)/2c_2(x) - 1.
\end{aligned}$$

If positive (negative) extremes of X_{t-h} are predominantly followed by positive (negative) extremes of X_t, then $X_t + X_{t-h}$ will have a comparatively high tail dispersion $c_2(y_h) > 2c_2(x)$ ($c_1(y_h) > 2c_1(x)$), hence $\rho^{(2)}(h) > 0$ ($\rho^{(1)}(h) > 0$), and so on. If X_t is serially independent then $\rho^{(i)}(h) = 0\ \forall h \geq 1$, $i = 0, 1, 2$.

Remark 1: The convolution difference $\tilde{Y}_{h,t} = X_t - X_{t-h}$ does not lend itself to tail-specific co-relations because the scales $c_i(\tilde{y}_h)$ incorporate *both* right *and* left tail information from $\{X_t\}$.

Remark 2: A two-tailed co-relation based on the convolution summation $Y_{h,t} \equiv X_t + X_{t-h}$ exists. We do not consider it here because it is vastly dominated by $\rho^{(0)}(h)$ in simulation experiments.

3.3 Co-Relation Properties

Consult Hill (2006a: Lemmas B.1-B.8) for complete details on the following properties.

3.3.1 Symmetry: The co-relation coefficient is symmetric in the senses $\rho^{(\cdot)}(h) = \rho^{(\cdot)}(-h)$, $\rho^{(0)}(-X_t, -X_{t-h}) = \rho^{(0)}(X_t, X_{t-h})$ and $\rho^{(1)}(-X_t, -X_{t-h}) = \rho^{(2)}(X_t, X_{t-h})$.

3.3.2 Bounds and α_0< 1: In general

$$\min\{1-2^{\alpha_0-1}, 0\} \leq \rho^{(0)}(h) \leq 1$$
$$-1 \leq \rho^{(i)}(h) \leq \max\{2^{\alpha_i-1}-1, 0\}\}, \quad i = 1, 2.$$

The co-relation obtains a boundary value when the extremes of X_t are a linear function of the extremes of X_{t-h}: $X_t = g(X_{t-h}) + o_p(X_{t-h})$ for some $g : \mathbb{R} \rightarrow \mathbb{R}$ where $g(z)/z \rightarrow -1$ or 1 as $z \rightarrow \pm\infty$. For example $X_t = a + X_{t-h} + u_t \exp\{-|X_{t-h}|\}$, for any $a \in \mathbb{R}$ and $u_t \overset{iid}{\sim} (1)$, is maximally positively co-related[2]. If $\alpha_1 = \alpha_2 = 2$ then each $\rho^{(i)}(h) \in [-1, 1]$.

If $\alpha_0 = \min\{\alpha_1, \alpha_2\} \leq 1$ then the strongest degree of "negative" two-tailed dependence is greater than or equal to zero, and at least one one-tailed upper bound is less then or equal to zero. This suggests interpreting small co-relation values is problematic when the data are highly volatile, but it does not preclude a test of $\rho^{(\cdot)}(h) = 0$. Nevertheless most time series of practical interest in macroeconomics and finance suggest $\min\{\alpha_1, \alpha_2\} > 1$.

3.3.3 Damping: If $\rho^{(0)}(h_1) \gtreqless \rho^{(0)}(h_2)$ for any $h_2 > h_1$ then

$$\lim_{\varepsilon\rightarrow\infty} P\left(|X_t - X_{t-h_2}| > \varepsilon\right) / P\left(|X_t - X_{t-h_1}| > \varepsilon\right) \gtreqless 1.$$

One-sided damping $\rho^{(0)}(h) \searrow 0$ implies $X_t - X_{t-h}$ is increasingly likely to obtain an extreme value as $h \rightarrow \infty$. Synonymously, X_t is increasingly likely to surpass the stochastic thresholds $X_{t-h} \pm \varepsilon$ as $\varepsilon \rightarrow \infty$ and $h \rightarrow \infty$. In general the co-relation decays according to the memory properties of $\{X_t\}$: see Section 3.4. Damping properties apply to each $\rho^{(\cdot)}(h)$.

3.3.4 Linear Extremal Dependence: For stationary, mean-zero finite variance processes the serial correlation is

$$corr(X_t, X_{t-h}) = 1 - \frac{\int_{-\infty}^{0} \tilde{y}^2 dF_{\tilde{y}_h}(\tilde{y}) d\tilde{y} + \int_{0}^{\infty} \tilde{y}^2 dF_{\tilde{y}_h}(\tilde{y}) d\tilde{y}}{2\left(\int_{-\infty}^{0} x^2 dF(dx) dx + \int_{0}^{\infty} x^2 dF(d) dx\right)} = 1 - \frac{\sigma_0^2(\tilde{y}_h)}{2\sigma_0^2(x)},$$

[2] Notice $X_t = a + bX_{t-h}$ for any $b \neq 1$ fails to satisfy the assumption that X_t has the identical tail shape (1) *for all* t: if $b > 0$ and X_{t-h} has right scale c_2 then X_t has right scale $|b|^{\alpha_2} c_2$.

say, where F and $F_{\tilde{y}_h}$ denote the continuously differentiable marginal distribution functions of X_t and $X_t - X_{t-h}$. The two-tailed $\rho^{(0)}(h)$ co-relation is therefore a generalized tail-version of the correlation, measuring linear extremal dependence:

$$\rho^{(0)}(h) = 1 - \lim_{\varepsilon \to \infty} \frac{\int_{-\infty}^{-\varepsilon} \varepsilon^{\alpha_0} dF_{\tilde{y}_h,t}(\tilde{y})d\tilde{y} + \int_{\varepsilon}^{\infty} \varepsilon^{\alpha_0} dF_{\tilde{y}_h,t}(\tilde{y})d\tilde{y}}{2\left(\int_{-\infty}^{-\varepsilon} \varepsilon^{\alpha_0} dF_t(x)dx + \int_{\varepsilon}^{\infty} \varepsilon^{\alpha_0} dF_t(x)dx\right)} = 1 - \frac{c_0(\tilde{y}_h)}{2c_0(x)},$$

where F_t and $F_{\tilde{y}_h,t}$ capture non-extremal non-stationarity. Notice the co-relation $\rho^{(\cdot)}(h)$ does not require X_t to be centered, nor in general do we require F_t to be differentiable, cf. (1) and (2). An identical comparison holds for either tail by noting, for example, $corr(X_t I(X_t > 0), X_{t-h}(X_{t-h} > 0)) = 1 - \sigma_2^2(\tilde{y}_h)/2\sigma_2^2(x)$, where $\sigma_2^2(x) = \int_0^{\infty} x^2 dF(d)dx$, etc.

3.3.5 Linearity and Symmetry: If $\{X_t\}$ is linear (e.g. ARFIMA) with *iid* innovations governed by symmetric tails (1) then the co-relation is symmetric: $\rho^{(0)}(h) = \rho^{(1)}(h) = \rho^{(2)}(h)$. Thus, a process with asymmetric co-relations is either nonlinear, or linear with asymmetric innovations.

3.4 Co-Relation Examples

In the following we characterize convolution tail equivalence and the co-relation for linear, nonlinear and conditional volatility processes. Throughout $u_t \sim (1)$ is symmetric and independent: $c_1(u) = c_2(u) = 1$ and $\alpha_1 = \alpha_2 = \alpha$ for each t. Consult Lemmas B.4-B.8 of Hill (2006a) for complete details. All results can be extended to the case $c_1(u) \lesseqgtr c_2(u)$ and $\alpha_1 \lesseqgtr \alpha_2$.

Example 1 (Distributed Lag): Suppose $X_t = \sum_{i=0}^{\infty} \psi_i u_{t-i}$, $\psi_0 = 1$, and $\sum_{i=0}^{\infty} |\psi_i|^{\alpha_0} < \infty$. In general $\{u_t\}$ need not be identically distributed, nor have a zero mean. Each $\{X_t\}$ and $\{X_t \pm X_{t-h}\}$ satisfies (1) with index α, and

$$\rho^{(0)}(h) = \frac{\sum_{i=0}^{\infty} |\psi_i|^{\alpha} + \sum_{i=h}^{\infty} |\psi_i|^{\alpha} - \sum_{i=0}^{\infty} |\psi_{i+h} - \psi_i|^{\alpha}}{2\sum_{i=0}^{\infty} |\psi_i|^{\alpha}}$$

$$\rho^{(j)}(h) = \frac{\sum_{i=0}^{\infty} |\psi_{i+h} + \psi_i|^{\alpha} - \sum_{i=h}^{\infty} |\psi_i|^{\alpha} - \sum_{i=0}^{\infty} |\psi_i|^{\alpha}}{2\sum_{i=0}^{\infty} |\psi_i|^{\alpha}}, \quad j = 1, 2.$$

Co-relation decay rates (e.g. geometric or hyperbolic) follow from decay rates of ψ_i. For an AR(1) process $X_t = \phi X_{t-1} + u_t$, $|\phi| < 1$, $\rho^{(0)}(h) = (1 + |\phi^h|^{\alpha} - |\phi^h - 1|^{\alpha})/2$ and $\rho^{(j)}(h) = (|\phi^h + 1|^{\alpha} - |\phi^h|^{\alpha} - 1)/2$, $j = 1,2$

Example 2 (Bilinear): Suppose $X_t = \beta X_{t-1} u_{t-1} + u_t$, $u_t \overset{iid}{\sim} (1)$, $P(u_t \geq 0) = 1$, $\beta > 0$, and $\beta^{\alpha/2} E[u_t^{\alpha/2}] < 1$. Each $\{X_t\}$ and $\{X_t \pm X_{t-h}\}$ satisfies (1) with index $\alpha/2$, and $\rho^{(\cdot)}(h)$ are represented above with $\psi_i = \beta^i$.

Example 3 (Extremal Threshold): Construct *Extremal Threshold* double arrays $X_{n,t} = \phi X_{n,t-1} I(|u_{t-1}| \leq v_n) + u_t$ and $W_{n,t} = \phi W_{n,t-1} I(|u_{t-1}| > v_n) + u_t$, $u_t \overset{iid}{\sim} (1)$, where $\{v_n\}$ is a non-stochastic sequence, $v_n \to \infty$ as $n \to \infty$, and $|\phi| < 1$. Clearly if $|u_{t-1}| > v_n$ then $X_{n,t}$ is independent noise and $W_{n,t}$ is AR(1). Each $\{X_{n,t}, X_{n,t} \pm X_{n,t-h}\}$ and $\{W_{n,t}, W_t \pm W_{t-h}\}$ satisfies (1).

$\{X_{n,t}\}$ is left-, right-, and two-tailed extremal white noise, and $\rho^{(0)}(W_t, W_{t-h})$ $= (1 + |\phi^h|^\alpha - |\phi^h - 1|^\alpha)/2 = 0$.

Example 4 (Power-ARCH(∞)): Let $X_t = \sigma_t u_t$, $u_t \overset{iid}{\sim} (1)$, $P(\epsilon_t > 0) = 1$, $E|u_t|^p = 1$, $0 < p < \alpha$, and $\sigma_t^p = \theta_0 + \sum_{i=1}^\infty \theta_i |X_{t-i}|^p$, $\theta_i \geq 0$, $\sum_{i=1}^\infty \theta_i < 1$ and $\sum_{i=1}^\infty \theta_i^\alpha < \infty$. Each $\{X_t, X_t \pm X_{t-h}\}$, $\{|X_t|^p, |X_t|^p \pm |X_{t-h}|^p\}$ and $\{\sigma_t^p \pm \sigma_{t-h}^p\}$ satisfies (1) with indices α, α/p and α/p, respectively. $\{X_t\}$ and $\{|X_t|^p\}$ are extremal white noise. The co-relations between σ_t^p and σ_{t-h}^p are represented above where $\{\psi_i\}$ satisfies $\sum_{i=0}^\infty \psi_i = \theta_0 + \sum_{r=1}^\infty \sum_{j_1,\cdots,j_r=1}^\infty \theta_0 \theta_{j_1} \cdots \theta_{j_r}$.

4. Comparison with Extant Extremal Dependence Measure The following is necessarily brief, and the reader may consult Hill (2006a) for additional details, as well as Tawn (1990), Coles *et al* (1999), Ledford and Tawn (1996, 1997) and the numerous citations therein. Consider only right-tail dependence.

We adopt bivariate dependence notation to the serial dependence case. The fundamental objective concerns

$$P\left(X_t > \varepsilon | X_{t-h} > \varepsilon\right), \quad \varepsilon > 0.$$

If $\lim_{\varepsilon \to \infty} P(X_t > \varepsilon | X_{t-h} > \varepsilon) = 0$ for all $h \geq 1$ then $\{X_t\}$ is serially asymptotically independent. Cf. Sibuya (1960).

The bulk of the effort concerns specifying a joint survival $P(X_t > \varepsilon, X_{t-h} > \varepsilon)$ with unit Frechét marginals. Ledford and Tawn (1996, 1997) investigate the joint regular variation form $P(W_t > \varepsilon, W_{t-h} > \varepsilon) = \varepsilon^{-1/\eta_h^{(2)}} L(\varepsilon)$ where $L(\varepsilon)$ is slowly varying, $\eta_h^{(2)}$ is the so-called "coefficient of tail dependence", and W_t is the unit Frechét transform of X_t with distribution function F_w. At displacement h negative dependence, independence, and positive dependence respectively imply $0 < \eta_h^{(2)} < 1/2$, $\eta_h^{(2)} = 1/2$ and $1/2 < \eta_h^{(2)} \leq 1$. We might, therefore, say $\{X_t\}$ is right-tailed "extremal serially independent" if $\eta_h^{(2)} = 1/2$ $\forall h \geq 1$. The remaining arguments in this literature concern specifications for $L(\varepsilon)$ and estimation of $\eta_h^{(2)}$.

A parametric strategy for abstracting from the marginal distributions is the use of a copula dependence function $C(W_t, W_{t-h})$. In general $C(W_t, W_{t-h}) = C(g(W_t), g(W_{t-h}))$ for $g : \mathbb{R} \to \mathbb{R}$ positive monotonic. Coles *et al* (1999), for example, consider the joint extreme value distribution: $C(w_t, w_{t-h}) = \exp\{-[(-\ln w_t)^{-1/\gamma} + (-\ln w_{t-h})^{-1/\gamma})^\gamma]\}$, where $w_t \equiv F_w(W_t)$.

Remark 1: In all cases that we are aware of tail dependence is defined between otherwise *iid* random variables *without* displacement. The above extension to the serial dependence case has never been entertained[3].

Remark 2: The index $\eta_h^{(2)}$ is bounded between 0 and 1 where values of $\eta_h^{(2)}$ have clear dependence interpretations. The co-relation $\rho^{(2)}(h)$ has (typically)

[3] Starica (1999) considers tail dependence for non-identically distributed Constant Conditional Correlation GARCH processes, but does not derive the distribution limit of the proposed tail dependence estimator.

asymmetric bounds depending on tail thickness, with a clear interpretation when $\alpha_2 > 1$.

Remark 3: The index $\eta_h^{(2)}$ captures all forms of right tail dependence, while the co-relation measures linear extremal dependence. Nevertheless, the co-relation provides an extremal conjugate to covariance orthogonality that can be applied to extremal model selection (we leave this idea for future endeavors).

Remark 4: Unlike copula-based measures the co-relation $\rho^{(0)}(h)$ is not invariant to *any* monotonic transformation because such transformations need not satisfy (1). In fact, if $g : \mathbb{R} \to \mathbb{R}$ is asymptotically monotonic (e.g. $g(z_2) \geq g(z_1) \to \infty$ as $z_2 \geq z_1 \to \infty$) and $X_t \sim (1)$ then $g(X_t) \sim (1)$ with indices $\alpha_i(g)$ and scales $c_i(g)$ *if and only if* g is asymptotically proportional to a power function. The co-relation is invariant to asymptotically positive affine transformations $g(X_t)$: $g)z)/z \to b > 0$ as $z \to \pm\infty$.

Remark 5: The co-relation decays according to the memory properties of $\{X_t\}$. We are not aware of a related property for the tail dependence coefficient.

5. Co-Relation Sample Statistics and Asymptotics It is convenient to work with tail-preserving transformations:

$$X_t^{(0)} = |X_t|, \;\; X_t^{(1)} = |X_t I(X_t \leq 0)|, \;\; X_t^{(2)} = X_t I(X_t > 0).$$

Each $X_t^{(i)}$ has a marginal distribution function with support $[0, \infty)$ and tail

$$P(X_t^{(i)} > \varepsilon) = c_i(x)\varepsilon^{-\alpha_i}(1 + o(1)).$$

Tail-preserving transformations for the convolutions are denoted $Y_{h,t}^{(\cdot)}$ and $\tilde{Y}_{h,t}^{(\cdot)}$.

Simple estimators for α_i and c_i well known in the literature are due to B. Hill (1975) and Hall (1982): for each $i = 0, 1, 2$, and $Z_t \in \{X_t, Y_{1,t}, \tilde{Y}_{1,t}\}$

$$\hat{c}_{i,m}(z) \equiv (m/n)\,(Z_{(m+1)}^{(i)})^{\hat{\alpha}_{i,m}(z)} \;\; \text{and} \;\; \hat{\alpha}_{i,m}^{-1}(z) \equiv \; 1/m \sum\nolimits_{j=1}^{m} \ln Z_{(j)}^{(i)}/Z_{(m+1)}^{(i)},$$

for some integer $1 \leq m \leq n$ such that $m \to \infty$ as $n \to \infty$ and $m/n \to 0$. The estimator $\hat{\alpha}_{i,m}^{-1}$ is widely known as the Hill-estimator, cf. B. Hill (1975), the subject of theoretical and empirical studies too numerous to list. See Hsing (1991) and Hill (2005) and the citations therein.

We deduce estimators of the co-relation coefficients

$$\hat{\rho}_m^{(0)}(h) = 1 - \hat{c}_{0,m}(\tilde{y}_h)/2\hat{c}_{0,m}(x), \;\; \hat{\rho}_m^{(i)}(h) = \hat{c}_{i,m}(y_h)/2\hat{c}_{i,m}(x) - 1, \; i = 1, 2,$$

where $\hat{c}_{i,m}(\tilde{y}_h) = (m/n)(\tilde{Y}_{h,(m+1)}^{(i)})^{\hat{\alpha}_{0,m}(\tilde{y}_1)}$ and $\{\tilde{Y}_{h,(\cdot)}^{(i)}\}$ denote the corresponding order statistics for $\{\tilde{Y}_{h,t}^{(i)}\}$. Similarly $\hat{c}_{i,m}(y_h) = (m/n)(Y_{h,(m+1)}^{(i)})^{\hat{\alpha}_{i,m}(y_1)}$.

We need only estimate the tail index of $Y_{1,t}$ and $\tilde{Y}_{1,t}$ (i.e. only for $h = 1$) due to convolution tail equivalence. Of course, by tail equivalence we could simply use $\hat{\alpha}_{i,m}(x)$ in each component of $\hat{\rho}_m^{(i)}(h)$. Somewhat surprisingly, asymptotic theory is greatly expedited[4] by using process specific tail index estimators in the numerical and denominator of $\hat{\rho}_m^{(i)}(h)$.

5.1 Assumptions

Define sequences of thresholds $\{b_m^{(i)}\}_{i=0}^2$ by

$$(3) \quad (n/m)P(X_t^{(i)} > b_m^{(i)}) \rightarrow 1,$$

where $b_m^{(\cdot)} \rightarrow \infty$ as $n \rightarrow \infty$. See Leadbetter *et al* (1983: Theorem 1.7.13).

We require extremal versions of mixing and Near-Epoch-Dependence properties developed in Hill (2005). Consult Gallant and White (1988) and Davidson (1994) for details on conventional properties. Denote by $\{\pi_{n,t}\}_{t=1}^n$ a sequence of thresholds $\pi_{n,t} \rightarrow \infty$ as $n \rightarrow \infty$ for each t, and define the σ-sub-algebra associated with extreme events of ϵ_t:

$$\mathcal{F}_{n,s}^t \equiv \sigma(I(|\epsilon_\tau| > \pi_{n,t}) : 1 \leq s \leq \tau \leq t \leq n).$$

Define the following coefficients, where $1 \leq q_n < n$, $q_n \rightarrow \infty$ as $n \rightarrow \infty$:

$$\varepsilon_{n,q_n} \equiv \sup_{A_{n,t} \in \mathcal{F}_{n,1}^t, B_{n,t+q} \in \mathcal{F}_{n,t+q_n}^n : 1 \leq t \leq n-q_n} |P(A_{n,t} \cap B_{n,t+q_n}) - P(A_{n,t})P(B_{n,t+q_n})|$$

$$\varpi_{n,q_n} \equiv \sup_{A_{n,t} \in \mathcal{F}_{n,1}^t, B_{n,t+q} \in \mathcal{F}_{n,t+q_n}^n : 1 \leq t \leq n-q_n} |P(B_{n,t+q_n}|A_{n,t}) - P(B_{n,t+q_n})|.$$

E-Mixing *If* $(n/m)q_n^\lambda \varepsilon_{n,q_n} \rightarrow 0$ *as* $n \rightarrow \infty$ *we say* $\{\epsilon_t\}$ *is Extremal-Strong Mixing with size* $\lambda > 0$. *If* $(n/m)q_n^\lambda \varpi_{n,q_n} \rightarrow 0$ *we say* $\{\epsilon_t\}$ *is Extremal-uniform mixing with size* $\lambda > 0$.

L_2**-E-NED** $\{X_t\}$ *is* L_2*-Extremal-NED on some array of* σ*-fields* $\{\mathcal{F}_{n,t}\}$ *with size* $\lambda > 0$ *if*

$$\left\| P(|X_t| > b_m^{(0)} e^v | \Im_{t-q}^{t+q}) - P(|X_t| > b_m^{(0)} e^v | \mathcal{F}_{n,t-q}^{t+q}) \right\|_2 \leq d_{n,t}(v)\psi_{n,q_n},$$

where $d_{n,t} : \mathbb{R} \rightarrow \mathbb{R}_+$ *is Lebesgue measurable,* $\sup_t d_{n,t}(v) = O((m/n)^{1/r})$ *for each* $v \in \mathbb{R}$, *and* $(n/m)^{1/2-1/r} q_n^\lambda \psi_{n,q_n} \rightarrow 0$ *as* $n \rightarrow \infty$ *for some* $r \geq 2$.

Remark 1: E-mixing properties are simply mixing properties for the extremal event process $\{I(|\epsilon_t| > \pi_{n,t})\}$. The E-NED property is simply an NED property applied to the two-tailed extremal event $I(|X_t| > b_m^{(0)} e^v)$, and implies the infinite history of the extremal event $\{I(|\epsilon_\tau| > \pi_{n,t}) : \tau \leq t\}$ can be used to predict almost surely whether the extreme $|X_t| > b_m^{(0)} e^v$ occurs or not.

[4] If we use $\hat{\alpha}_{i,m}(x)$ in both the numerator and demoninator of $\hat{\rho}_m^{(i)}(h)$, asymptotic theory then necessitates a characterization of the joint distribution, from which we want to abstract.

Remark 2: Hill (2005: Lemma 7) shows any L_p-NED process, $p > 0$, with tails (1) has the L_2-E-NED property. This covers linear, bilinear, certain nonlinear distributed lags, and conditional volatility processes. In general, ARFIMA, FIGARCH, bilinear and Extremal Threshold processes are L_2-E-NED as long as the underlying innovations are *iid* and satisfy a general regularly varying tail property. Furthermore, the Extremal Threshold process $\{W_{n,t}\}$ in Example 3 of Section 3.4 is L_2-E-NED if the threshold $v_n \to \infty$ slowly enough (e.g. $v_n < b_m^{(0)}$), but is *not* L_2-NED for any threshold sequence $v_n \to \infty$ when $\alpha_0 > 2$. This implies E-NED has a clear advantage for extreme value applications. See Lemma 9 of Hill (2005).

Remark 3: Memory restrictions are not imposed on the non-extremal support of $X_t \in [-b_m^{(0)}, b_m^{(0)}] \to (-\infty, \infty)$.

Let $Z_t^{(\cdot)}$ denote any of $X_t^{(\cdot)}$, $Y_{h,t}^{(\cdot)}$, or $\tilde{Y}_{h,t}^{(\cdot)}$, and denote by $b_m^{(\cdot)}$ the associated threshold sequences, a la (3).

Assumption A Each $Z_t \in \{X_t, Y_{t,h}, \tilde{Y}_{t,h}\}$ satisfies for some $\alpha_i(z), \theta_i(z) > 0$

$$P(Z_t < -\varepsilon) = c_1(z)\varepsilon^{-\alpha_1(z)}(1 + O(\varepsilon^{-\theta_1(z)}))$$
$$P(Z_t > \varepsilon) = c_2(z)\varepsilon^{-\alpha_2(z)}(1 + O(\varepsilon^{-\theta_2(z)})).$$

In particular, $\alpha_i(x) = \alpha_i(y_h) = \alpha_i$, $\theta_i(x) = \theta_i(y_h) = \theta_i$, $\alpha_i(\tilde{y}_h) = \alpha_0 = \min\{\alpha_1, \alpha_2\}$, and $c_i(z) > 0$ for at least one $i \in \{1, 2\}$

Assumption B $m \sim n^\delta$, $\delta \in (0, 1)$. One of the following holds:

1. $0 < \delta < \min_{0 \leq i \leq 2}\{2\theta_i/(2\theta_i + \alpha_i)\}$.
2. $1/2 < \delta < \min_{0 \leq i \leq 2}\{2\theta_i/(2\theta_i + \alpha_i)\}$.

Assumption C Let $\{\epsilon_t\}$ be E-uniform mixing with size $r/[2(r-1)]$, $r \geq 2$; or E-strong mixing with size $r/(r-2)$, $r > 2$.

1. $\{X_t\}$ is L_2-E-NED on $\{\mathcal{F}_{n,t}\}$ of size $-1/2$ with constants $d_{n,t}^{(x)}(v)$, where $\sup_t d_{n,t}^{(x)}(v) = O((m/n)^{1/r})$ and $(\int_0^\infty d_{n,t}^{(x)}(v)^2 dv) = O((m/n)^{1/r})$ for some $r \geq 2$.
2. Each $\{Z_t\} \in \{Y_{h,t}, \tilde{Y}_{h,t}\}$ is L_2-E-NED on $\{\mathcal{F}_{n,t}\}$ of size $-1/2$ with constants $d_{n,t}^{(z)}(v)$, where $\sup_t d_{n,t}^{(z)}(v) = O((m/n)^{1/r})$, $(\int_0^\infty d_{n,t}^{(z)}(v)^2 dv) = O((m/n)^{1/r})$ for some $r \geq 2$.

Remark 1: Assumption B.1 is a direct consequence of Assumption A: see Haeusler and Teugels (1985) and Hill (2005). Assumption B.2 expedites consistency of a kernel estimator of the asymptotic covariance matrix of $\hat{\alpha}_{i,m}^{-1}(z)'s$.

Remark 2: Hill (2005b) proves Assumption C.1 holds for each linear, nonlinear and conditional volatility process detailed in Section 3.4. It is straightforward to verify that Assumption C.2 also holds given these processes have distributed lag representations. See Lemma 8 of Hill (2005).

5.2 Preliminary Theory

Consistency and asymptotic normality of the Hill-estimator under general conditions is of paramount importance for the limiting properties of our test statistic under both hypotheses.

THEOREM 1 *Under Assumptions A, B.1 and C.1,* $\hat{\alpha}_{i,m}(z) = \alpha_i + O_p(1/\sqrt{m})$, $Z^{(\cdot)}_{(m+1)}/b^{(\cdot)}_m = 1 + O_p(1/\sqrt{m})$, $\hat{c}_{i,m}(z) = c_i(z) + O_p(\ln(n/m)/\sqrt{m})$, *and* $\hat{\rho}^{(j)}_m(h) \rightarrow \rho^{(j)}(h)$ *for each* $j = 0, 1, 2$.

Tests of extremal white noise and co-relation tail equivalence are grounded on $\hat{\rho}^{(j)}_m(h)$ and $\hat{\rho}^{(2)}_m(h) - \hat{\rho}^{(1)}_m(h)$, functions of two and four tail index estimators respectively. We therefore require a joint limit theory for a vector of Hill-estimators and a consistent covariance matrix estimator.

Let $\{W_t\} = \{W_{i,t} : i = 1...k\}$ be a k-vector stochastic process on $\mathbb{R}^k_+$ with marginal distribution tails (1) and indices $\alpha = [\alpha_1, ..., \alpha_k]' > 0$, $k \geq 1$. In the sequel we will use $W_t = [X^{(i)}_t, \tilde{Y}^{(i)}_{1,t}]'$ and $W_t = [X^{(1)}_t, Y^{(1)}_{1,t}, X^{(2)}_t, Y^{(2)}_{1,t}]'$.

Write $\hat{\alpha}^{-1}_m = [\hat{\alpha}^{-1}_{1,m}, ..., \hat{\alpha}^{-1}_{k,m}]'$ and define the covariance matrix

$$\Sigma_m \equiv mE(\hat{\alpha}^{-1}_m - \alpha^{-1})(\hat{\alpha}^{-1}_m - \alpha^{-1})'.$$

Define a kernel estimator

$$(4) \quad \hat{\Sigma}_m = m^{-1} \sum\nolimits_{s=1}^{n} \sum\nolimits_{t=1}^{n} w((s-t)/\gamma_n) \hat{Z}_s \hat{Z}'_t, \quad \hat{Z}_t \equiv [\hat{Z}_{1,t}, ..., \hat{Z}_{k,t}]'$$

where $\hat{Z}_{i,t} \equiv [(\ln W_{i,t}/W_{i,(m+1)})_+ - (m/n)\hat{\alpha}^{-1}_{i,m}]$, and $w((s-t)/\gamma_n)$ denotes a standard kernel function with bandwidth $\gamma_n \rightarrow \infty$ as $n \rightarrow \infty$, $w(0) = 1$ and $w(z) = w(-z)$. Assume $w(\cdot)$ satisfies Assumption 1 of de Jong and Davidson (2000) such that $\hat{\Sigma}_m$ is positive definite. The following result holds for Barlett, Parzen, Quadratic Spectral and Tukey-Hanning kernels.

THEOREM 2

i. If each $\{W_{i,t}\}$ *satisfies Assumptions A, B.1 and C.1, then*

$$\sqrt{m}(\hat{\alpha}^{-1}_m - \alpha^{-1}) \Rightarrow N(0, \Sigma), \quad \Sigma = \lim_{n \rightarrow \infty} \Sigma_m, \quad ||\Sigma_m||_1 = O(1).$$

ii. Let each $\{W_{i,t}\}$ *satisfy Assumptions A, B.2 and C.1, and let* $\gamma_n \rightarrow \infty$, $\gamma_n/m = o(n^{-1/2})$ *and* $1/m \sum_{s,t=1}^{n} |w((s-t)/\gamma_n)| = o(n^{1/2})$. *Then* $|\hat{\Sigma}_m - \Sigma_m| \rightarrow 0$.

Remark 1: Standard arguments easily verify

$$\tilde{\Sigma}^{-1/2}_m \sqrt{m}(\hat{\alpha}_m - \alpha) \Rightarrow N(0, I_k), \quad \text{where } \tilde{\Sigma}_m = [\alpha^2_i \alpha^2_j \Sigma_{m,i,j}]^k_{i,j=1}.$$

Remark 2: Write $\sigma^2_{i,m} = E(\sqrt{m}(\hat{\alpha}^{-1}_{i,m} - \alpha^{-1}_i))^2$, and define

$$\hat{\sigma}^2_{i,m} = m^{-1} \sum\nolimits_{s=1}^{n} \sum\nolimits_{t=1}^{n} w((s-t)/\gamma_n) \hat{Z}_{i,s} \hat{Z}_{i,t}.$$

The scalar case simply represents Theorems 5 and 6 of Hill (2005): $\sqrt{m}(\hat{\alpha}_{i,m} - \alpha_i)/\tilde{\sigma}_{i,m} \Rightarrow N(0,1)$, $\tilde{\sigma}^2_{i,m} = \alpha^4_i \sigma^2_{i,m}$, where $\sigma^2_{i,m} = E(\sqrt{m}(\hat{\alpha}^{-1}_{i,m} - \alpha^{-1}_i))^2$; and $|\hat{\sigma}^2_{i,m} - \sigma^2_{i,m}| \rightarrow 0$.

5.3 Co-Relation Estimator

Define $\alpha_{i,i} = [\alpha_i, \alpha_i]'$, $\hat{\alpha}_{0,0,m} \equiv [\hat{\alpha}_{i,m}(x), \hat{\alpha}_{i,m}(\tilde{y}_1)]'$, $\hat{\alpha}_{i,i,m} \equiv [\hat{\alpha}_{i,m}(x), \hat{\alpha}_{i,m}(y_1)]'$ for $i = 1, 2$, and

$$\Sigma^{(i,i)}_m \equiv mE\left(\hat{\alpha}^{-1}_{i,i,m} - \alpha^{-1}_{i,i}\right)\left(\hat{\alpha}^{-1}_{i,i,m} - \alpha^{-1}_{i,i}\right)' \in \mathbb{R}^2, \quad i = 0, 1, 2.$$

Write $\hat{\rho}^{(i)}_{m,h} = [\hat{\rho}^{(i)}_m(1), ..., \hat{\rho}^{(i)}_m(h)]'$ and $\rho^{(i)}_h = [\rho^{(i)}(1), ..., \rho^{(i)}(h)]'$.

THEOREM 3 *If Assumptions A, B.1, and C hold, and* $\Sigma^{(i,i)} = \lim_{n\rightarrow\infty} \Sigma^{(i,i)}_m$ *is positive definite, then*

$$\left(\sqrt{m}/\ln(n/m)\right)\left(\hat{\rho}^{(i)}_{m,h} - \rho^{(i)}_h\right) \Rightarrow N\left(0, v^2_i \times [1_h \times 1'_h]\right)$$

where

$$v^2_i = \alpha^2_i A' \Sigma^{(i,i)} A > 0, \;\; A = [1, -1]'.$$

If $\rho^{(i)}(j) \neq 0$ *then* $m \times \hat{\rho}^{(i)}_m(j)^2 \rightarrow \infty$ *with probability one.*

Remark: The covariance matrices $v^2_i[1_h \times 1'_h]$ are valid for any $\rho^{(i)}(j) \lesseqgtr 0$, and are singular because $\hat{\rho}^{(i)}_m(j)$ is grounded on the same random variables for each displacement j. Notice the variances v^2_i depend only on the tail $i \in \{0, 1, 2\}$, and not displacement because we use $\hat{\alpha}_{i,m}(\tilde{y}_1)$ and $\hat{\alpha}_{i,m}(y_1)$ for each displacement.

5.4 Tail Difference Estimator

In order to test the difference in co-relation tails we use

$$\Delta\hat{\rho}_{m,h} \equiv [\hat{\rho}^{(2)}_m(1) - \hat{\rho}^{(1)}_m(1), ..., \hat{\rho}^{(2)}_m(h) - \hat{\rho}^{(1)}_m(h)].$$

Define $\alpha_{2,1} \equiv [\alpha_2, \alpha_2, \alpha_1, \alpha_1]'$, $\hat{\alpha}_{2,1,m} \equiv [\hat{\alpha}_{2,m}(x), \hat{\alpha}_{2,m}(y_1), \hat{\alpha}_{1,m}(x), \hat{\alpha}_{1,m}(y_1)]'$ and

$$\Sigma^{(2,1)}_m \equiv mE\left(\hat{\alpha}^{-1}_{2,1,m} - \alpha^{-1}_{2,1}\right)\left(\hat{\alpha}^{-1}_{2,1,m} - \alpha^{-1}_{2,1}\right)'.$$

THEOREM 4 *If Assumptions A, B.1, C, $\rho_m^{(i)}(h) > -1$ for each $i = 1, 2$, and $\Sigma^{(2,1)} = \lim_{n\to\infty} \Sigma_m^{(2,1)}$ is positive definite, then*

$$\left(\sqrt{m}/\ln(n/m)\right)\left(\Delta\hat{\rho}_{m,h} - \Delta\rho_h\right) \Rightarrow N\left(0, [v(i) \times v(j)]_{i,j=1}^h\right),$$

where $v^2(j) = A(j)'\Sigma^{(2,1)}A(j) > 0$, and

$$A(j) \equiv \left[(\rho^{(2)}(j)+1)\alpha_2, -(\rho^{(2)}(j)+1)\alpha_2, -(\rho^{(1)}(j)+1)\alpha_1, (\rho^{(1)}(j)+1)\alpha_1\right]'.$$

If $\rho^{(1)}(j) \neq \rho^{(2)}(j)$ for some $j \in \{1, ..., h\}$ then $m \times (\hat{\rho}_m^{(1)}(j) - \hat{\rho}_m^{(2)}(j))^2 \to \infty$ with probability one.

Remark: Most time series encountered in practice will not be maximally negatively extremal dependent (i.e. $X_t = -X_{t-h}$ *a.s.*). Assuming $\rho_m^{(i)}(h) > -1$ does not reduce the generality of the result by much and, together with positive definiteness of $\Sigma^{(2,1)}$, ensures $v^2(j) > 0$. Notice the variances $v^2(j)$ depend on the displacement $j \in \{1, ..., h\}$ through $\rho^{(i)}(j)$.

5.5 Estimated Residuals

As long as estimated residuals, say $\hat{u}_t$, satisfy $\hat{u}_t = u_t + o_p(1)$ for some underlying process $\{u_t\}$ then $\hat{u}_t \to u_t$ in distribution. If $\{u_t\}$ satisfies Assumptions A-D then asymptotically $\{\hat{u}_t\}$ will. Such a condition is satisfied for linear least squares residuals, residuals derived from least absolute deviation estimation of ARIMA time series, Whittle estimated residuals from ARIMA and ARFIMA models, estimated residuals in GARCH models with infinite variance errors, etc. See, e.g., Knight (1993), Mikosch *et al* (1995), Kokoszka and Taqqu (1996), Davis and Wu (1997), and Hall and Yao (2003).

6. Tests of Extremal White Noise

We develop a test of extremal white noise using the sample estimators $\hat{\rho}_m^{(i)}(h)$. All ideas extend to a test of difference in co-relation tails.

Recall the limiting null distribution of $\hat{\rho}_m^{(i)}(j)$ is grounded on the same $\hat{\alpha}_{i,m}$'s for each j. Thus, a standard portmanteau statistic is not available, but a weighted-average portmanteau statistic is. For some increasing sequence of positive integers $h_n \to \infty$ as $n \to \infty$, and some set of weights $\{w(j) > 0\}_{j=1}^{\infty}$, $\sum_{j=1}^{h_n} w(j) \to 1$, define

$$Q_m^{(i)}(h_n) \equiv \left(m/\left(\ln n/m\right)^2\right)\hat{v}_{i,m}^{-2}\sum_{i=1}^{h_n} w(i) \times \hat{\rho}_m^{(i)}(i)^2$$

where

$$\hat{v}_{i,m}^2 = \hat{\alpha}_{i,m}^2(x) \times [1, -1] \times \hat{\Sigma}_m^{(i,i)} \times [1, -1]',$$

and $\hat{\Sigma}_m^{(i,i)}(x, \tilde{y}_1)$ is based on (4) with $W_t = [X_t^{(i)}, \tilde{Y}_{1,t}^{(i)}]'$.

For a straight average fix $w(i) = 1/h_n$ for each i. Weights that augment co-relations at distant displacements, a la Ljung and Box (1978), include $w(i) = (m+2)/h_n(m-i)$ provided $m > h_n = o(m)$.

THEOREM 5 *Let Assumptions A, B.2, and C hold and let $h_n \to \infty$ as $n \to \infty$, $h_n = o(m)$. If $\rho^{(i)}(j) = 0\ \forall j \geq 1$ then $Q_m^{(i)}(h_n) \Rightarrow \chi^2(1)$. If $\rho^{(i)}(j) \neq 0$ for at least one $j = 1 \geq 1$ then $Q_m^{(i)}(h_n) \to \infty$ with probability one.*

7. Order Statistic Index Selection In order to construct the test statistic a rather arbitrary choice of the sample tail fractile m is required. Using all "feasible" fractile sequences $\{m\}$ we consider a ranked co-relation strategy based on the simple logic that under the null any sequence $\{m\}$ should render $\hat{\rho}_m^{(0)}(j)$ close to zero: $\hat{\rho}_m^{(0)}(j) \in [-k_n^{\pi}, k_n^{\pi}]$ where, say, $k_n^{.05} = 1.96 \times \hat{v}_n/(\sqrt{m}/\ln(n/m))$, and $(\sqrt{m}/\ln(n/m)\hat{\rho}_m^{(0)}(j)/\hat{v}_n \Rightarrow N(0,1)$. The subsequent method and theory carry over to $\hat{\rho}_m^{(i)}(j)$, $i = 1, 2$, and $\Delta\hat{\rho}_m(j)$.

Define the set of proportional sequences $\{m\}$ under Assumption B.2:

$$\begin{aligned} S_m &= \{m : 1 \leq m \leq n,\ m \sim n^{\delta},\ 1/2 < \delta < \min\nolimits_{0 \leq 1 \leq 2}\{2\theta_i/(2\theta_i + \alpha_i)\}, \\ &\qquad \text{and } m/\tilde{m} = 1 + o(1)\}. \end{aligned}$$

For example $m = [n^{\delta}] \pm [a_n]$ is appropriate for any $a_n = o(n^{\delta})$.

For each lag i define $m_i^{(r)} \in S_m$ by the r^{th}-rank of $\hat{\rho}_m^{(0)}(i)$:

$$\left|\hat{\rho}^{(0)}_{m_i^{(1)}}(i)\right| \leq \left|\hat{\rho}^{(0)}_{m_i^{(2)}}(i)\right| \leq \ldots.$$

Write $m^{(r)} \equiv [m_1^{(r)}, ..., m_h^{(r)}]$ and construct the test statistic functional

$$Q^{(0)}_{m^{(r)}}(h_n) = \sum\nolimits_{i=1}^{h_n} \left(m_i^{(r)}/\left(\ln(n/m_i^{(r)})\right)^2\right) \hat{v}^{-2}_{0,m_i^{(r)}} \times w(i) \left(\hat{\rho}^{(0)}_{m_i^{(r)}}(i)\right)^2.$$

LEMMA 6 *Let Assumptions A, B.2, and C hold. For each $j = 1...h$ and any $(m, \tilde{m}) \in S_m$: $\hat{\alpha}_{i,m} = \hat{\alpha}_{i,\tilde{m}} + o_p(1)$, $\hat{v}_{i,m} = \hat{v}_{i,\tilde{m}} + o_p(1)$, and $\hat{v}_m(j) = \hat{v}_{\tilde{m}}(j) + o_p(1)$. Moreover, if $\rho^{(i)}(j) = 0$ for some $j \geq 1$ then*

$$(\sqrt{m}/\ln(n/m))\hat{\rho}_m^{(i)}(j/\hat{v}_{i,m} = (\sqrt{\tilde{m}}/\ln(n/\tilde{m}))\hat{\rho}_{\tilde{m}}^{(i)}(j)/\hat{v}_{i,\tilde{m}} + o_p(1).$$

Now let M denote an arbitrary subset of S_m with n_m sequences. It is irrelevant whether n_m is constant or $n_m \to \infty$ as $n \to \infty$. Lemma 6 implies the following result.

COROLLARY 7 *Let Assumptions A, B.2, and C hold, let $\rho^{(0)}(i) = 0$, $\forall i \geq 1$, and consider any $(m, \tilde{m}) \in S_m$.*

i. $n_m^{-1}\sum_{m \in M}(\sqrt{m}/\ln(n/m))\hat{\rho}_m^{(0)}(i) = (\sqrt{\tilde{m}}/\ln(n/\tilde{m}))\hat{\rho}_{\tilde{m}}^{(0)}(i) + o_p(1)$;
ii. $n_m^{-1}\sum_{m \in M} Q_m^{(0)}(h_n) = Q_{\tilde{m}}^{(0)}(h_n) + o_p(1)$;
iii. $n_m^{-1}\sum_{m \in M}\sum_{i=1}^{h_n} Q_m^{(0)}(i) = Q_{\tilde{m}}^{(0)}(h_n) + o_p(1)$;

If $\rho^{(0)}(i) \neq 0$ for some $i \geq 1$ then $n_m^{-1}\sum_{m \in M} Q_m^{(0)}(h_n) \to \infty$ and $n_m^{-1}\sum_{m \in M}\sum_{i=1}^{h_n} Q_m^{(0)}(i) \to \infty$.

Remark: Claims $(i)-(iii)$ imply an average Q-statistic over an arbitrary window of proportional fractiles m, or over proportional $m's$ and displacements $i = 1...h_n$, are asymptotically equivalent to any one $Q_m^{(0)}(h_n)$.

8. Small Sample Performance

We draw random samples of *iid* time series u_t from the Pareto distribution

$$F(\varepsilon) = |\varepsilon|^{-\alpha} \quad \varepsilon < 0, \;\; \bar{F}(\varepsilon) = \varepsilon^{-\alpha} \quad \varepsilon > 0, \quad \alpha = 1.5.$$

Simulation results based on asymmetric Paretian shocks are qualitatively similar. For empirical size we consider both $X_t = \epsilon_t$ and an Extremal Threshold process

$$X_{n,t} = \phi_1 X_{n,t-1} I(|u_{t-1}| \leq v_n) + u_t, \quad v_n = n^{1/5}.$$

By Example 3 of Section 3.4, $\{X_{n,t}\}$ is left-, right- and two-tailed extremal white noise.

Under the alternative we construct AR(1), MA(1), and Self-Exciting AR processes from

$$X_t = (\phi_1 X_{t-1} + \eta_1 u_{t-1})\, I(X_{t-1} > 0) + (\phi_2 X_{t-1} + \eta_2 \epsilon_{t-1})\, I(X_{t-1} \leq 0) + \epsilon_t.$$

The respective cases are AR(1): $(\phi_j, \eta_j) = (.9, 0)$, $j = 1, 2$; MA(1): $(\phi_j, \eta_j) = (0, .9)$, $j = 1, 2$; and SETAR: $(\phi_1, \eta_1) = (.9, 0)$ and $\phi_2 = \eta_2 = 0$. In the SETAR case the process is *iid* noise when $X_{t-1} \leq 0$, and AR(1) when $X_{t-1} > 0$, hence $\rho^{(1)} = 0 < \rho^{(2)} = (1 + |\phi_1^h|^\alpha - |1 - \phi_1^h|^\alpha)/2$.

Finally, we simulate a power Hyperbolic-ARCH(∞) process of the form

$$X_t = \sigma_t u_t, \;\; \sigma_t^p = \theta_0 + \sum_{i=1}^{L_n} \theta_i |X_{t-i}|^p.$$

We randomly select $\theta_0 \in [.01, .5]$, fix $\theta_i = i^{-\mu}$, $p = 1.2$, $\mu = 2$, and $L_n = [.25n]$. The process $\{X_t\}$ is non-identically distributed extremal white noise.

We simulate 100 series of each process, and tests are performed at the 5%-level for each $h = 1...4$. We compute two-tailed $Q_{m^{(r)}}^{(0)}$ and difference in tails $\Delta Q_{m^{(r)}}$ based on co-relation ranks $r \in \{1, 5, 10, 15, 20, 25, 30, 60\}$. For the sake of brevity we simply use uniform weights $w(j) = 1/h_n$.

8.1 Two-Tailed Tests

Table 1 contains comprehensive two-tailed Q-test results for sample sizes $n \in \{400, 600\}$. Rejection rates for simple extremal white noise and HYARCH random variables are exceptional for all ranks $r \leq [.04 \times n]$ presented. Moreover, empirical power is essentially 100% for both MA(1) and AR(1) processes for even the smallest rank $r = 1$.

8.2 Tail Index Estimates

We report tail index estimates and 95% confidence bands $\hat{\alpha}_{0,m} \pm k_{0,m}$ for $|X_t|$ by averaging $\hat{\alpha}_{0,m_i^{(r)}}$ and $k_{0,m_i^{(r)}} \equiv 1.96 \times \hat{\bar{\sigma}}_{0,m_i^{(r)}}/[(m_i^{(r)})^{1/2}/\ln(n/m_i^{(r)})]$

over ranks $r \in \{1,...,[.05 \times n]\}$ and displacements $i = 1...h$. As long as the fractiles $m_i^{(r)}$ are chosen from a set of proportional sequences S_m, Lemma C.3 of Hill (2006a) guarantees $\hat{\alpha}_{0,m_i^{(r)}}$ and $\widehat{\overline{\sigma}}_{0,m_i^{(r)}}$ have the same properties as any one $\hat{\alpha}_{0,m}$ and $\widehat{\overline{\sigma}}_{0,m}$. The resulting estimator is interesting in and of itself, and is clearly accurate for all random variables simulated. In all cases $\alpha = 1.5$ lies in the confidence band, where 2 does not for all but AR(1) processes. The persistence of the AR(1) process flattens small sample tails, hence the estimates are positively biased with decreasing bias as n increases (see, also, Table 2).

8.3 Difference in Tails

In simulation experiments not reported here we found a comparatively larger sample size is required for the co-relation difference-based test statistic, $\Delta Q_m(h_n)$, to obtain a reasonable empirical size for dependent, symmetric data (e.g. AR(1)). The challenge lies in the fact that $\Delta\hat{\rho}_m(j)$ requires tail-specific Hill estimators $\hat{\alpha}_{i,m}$. For any n-sample there may be very few observations from one tail even if the process is symmetric, hence the corresponding $\hat{\alpha}_{i,m}$ may be heavily biased with a large dispersion. This occurs because one large spike in ϵ_t coupled with strong serial dependence in X_t may produce a sample path asymmetrically dominated by positive of negative values. A large sample size ensures adequate left-tail and right- sample sizes. We consider sizes $n \in \{1000, 1200\}$, comparable to our empirical study. See Table 2.

We test inherently symmetric extremal white noise and AR(1) processes, producing exceptional empirical size for low ranks $r \leq [.01n]$. Empirical power for an asymmetric SETAR process is as large as 99% for the same low ranks ($n = 1200$). Again, two-tailed tail index estimates are exceptional.

8.4 Sample Co-Relations

Table 3 presents sample two-tailed and difference-in-tails co-relations and confidence bands. We also present true co-relation values based on the formula in Example 1 of Section 3.4. For all processes we average $\hat{\rho}^{(0)}_{m_h^{(r)}}(h)$ and $\Delta\rho_{m_h^{(r)}}(h)$ over ranks $r = [.005n]...[.05n]$, the asymptotic validity of which is guaranteed by Lemma 6. The sample two-tailed co-relation is extremely sharp in all cases where the true value, and not zero (when $\rho^{(0)}(h) \neq 0$), appear in each 95% band, *except* for the MA(1) at displacement $h = 1$. Because the co-relation ranks are relatively small there exists a slight negative bias in nearly all cases, and a larger bias for moving averages due to the inherently weak nature of dependence.

When the process is SETAR the difference-in-tails estimator requires a relatively large sample size for the obvious positive bias to vanish. Bias arises due to the highly asymmetric and persistent nature of the data favoring the right-tail. Nevertheless the associated Q-test is exceptional.

Figure 1 plots sample co-relations of Extremal AR(1) $X_{n,t} = .9X_{n,t-1}I(|\epsilon_{t-1}| > n^{1/5}) + \epsilon_t$ and Extremal MA(1) $X_{n,t} = .9\epsilon_{t-1}I(|\epsilon_{t-1}| > n^{1/5}) + \epsilon_t$ processes out to 50 displacements based on a sample of size $n = 1000$. The plots are averages over ranks $r = [.005n]...[.05n]$, and clearly demonstrate the sharpness of the sample co-relations.

9. Application We now analyze the serial extremal dependence properties of exchange rate and equity market daily log returns for the period 1/3/00 to 8/31/05. We consider Yen, Euro, and British Pound [BP] daily 10am spot rates against the U.S. Dollar; and the NASDAQ, S&P500, and Shanghai Stock Exchange [SSE] composite daily open-close averages[5]. Exchange rate trading does not occur on the weekends, and all weekends, holidays and unscheduled closures are treated as missing values. We linearly filter all variables to remove day effects using a standard daily dummy regression.

We estimate two-tailed and difference-in-tails co-relations, test extremal white noise hypotheses for daily displacements $h = 1...4$, and report the minimum p-value over h. Results are compiled in Table 4. Figures 2 and 3 contains plots of sample co-relation 95% intervals for 50 daily displacements. We average all co-relations over ranks $r = [.005n]...[.05n]$. For each series we also report the sample two-tailed tail index band $\hat{\alpha}_{0,m} \pm \widehat{\widetilde{\sigma}}_{0,m}/\sqrt{m}$ averaged over $m_i^{(r)}$ for each $i = 1...h$ and $r = [.005n]...[.05n]$.

9.1 Exchange Rate Fluctuations

The daily return to the Yen and Euro exhibit weak positive serial extremal dependence up to (and beyond) 50 daily displacements. See Figure 2. The large confidence bands are due, in part, to the slow rate of convergence $\sqrt{m}/\ln(n/m)$. Co-relation estimates at several displacements are significant at the 10% level for each series (not shown). The British Pound exhibits extremely weak, positive extremal dependence, where significance is at the 5% level only at a displacement of 13 trading days.

There is strong evidence in favor of symmetric co-relations in the Yen and Euro. Symmetric movements near a fixed (non-extremal) threshold is typically modeled as an Exponential Smooth Transition Autoregression. Cf. Teräsvirta (1994). A time series governed by a regime switching mechanism with respect to the non-extremal support (e.g. SETAR; Markov Switching; ESTAR) or merely with respect to extremes (e.g. Extremal TAR), or linear processes with asymmetric innovations, will have asymmetric co-relations. Thus, the strong evidence in favor of symmetric co-relation tails suggests extreme exchange rate returns may be governed by a linear data generating process with symmetric shocks, and implies non-extremes *and* extremes cannot be governed by a switching mechanism that is sensitive to the sign of X_t or the underlying shock. This evidence runs sharply contrary to the now *de riguer* wisdom that many exchange rates are $I(1)$ process with non-*iid* shocks governed by a regime switching data generating process. See Kräger and Kugler (1993) and Clements and Smith (2001).

Although the British Pound is *nearly* two-tailed extremal white noise, it displays significant asymmetric serial dependence favoring the left-tail beginning at a two day displacement. This implies sharp devaluations of the Pound are relatively noise compared to sharp devaluations of the Dollar. This result is likely due to the continuing decline of the U.S. Dollar againt major currencies

[5] Exchange rates are daily 10am spot rates taken from the New York Federal Reserve Bank statistical releases. Stock indices were obtained from Finance.yahoo.com.

over the past decade.

9.2 Asset Market Returns

The nature of extremal dependence in asset returns is both more diverse, and much stronger. See Figure 3. The NASDAQ is symmetrically two-tailed co-related well beyond 50 daily displacements. The strongest degree of serial extremal dependence exists at one daily displacement, and the tail of the auto-co-relogram suggests long ("hyperbolic") memory.

The S&P500 exhibits highly significant and weak symmetric first order serial co-relation, and extremely weak, persistent, serial extremal dependence at subsequence displacements (significance is at the 10% level). Like the NASDAQ, extreme spikes in the S&P500 appear to be governed by a long memory data generating process. We leave for future research any attempt to test for long memory in extremes.

Conversely, the SSE exhibits low levels of positive and negative serial co-relation up to 22 daily displacements. Evidence suggests asymmetric co-relation tails favoring the right-tail: positive spikes are more persistent than negative spikes. This may be due to the emerging market status of the SSE (start: Dec. 1990) and the evolving financial sector in China. Over the sample period traders may view a sharp increase as indicative of a lasting trend.

Appendix 1: Proofs

Proof of Theorem 1. See Lemma 4 and Theorem 5 of Hill (2005) for $Z^{(i)}_{(m+1)}/b^{(i)}_m = 1 + O_p(1/\sqrt{m})$ and $\hat{\alpha}_{i,m} = \alpha_i + O_p(1/\sqrt{m})$. By invoking functional invariance of probability limits, we need only show $\hat{c}_{i,m}(z) = c_i(z) + O_p(\ln(n/m)/\sqrt{m}) = o_p(1)$ to complete the proof. Write

$$\begin{aligned}\ln \hat{c}_{i,m}(z) &= \hat{\alpha}_{i,m} \ln Z^{(i)}_{(m+1)} + \ln\left(\frac{m}{n}\right) \\ &= \frac{\hat{\alpha}_{i,m}}{\alpha_i} \ln\left(\frac{m}{n}\left(b^{(i)}_m\right)^{\alpha_i}\right) + \frac{\hat{\alpha}_{i,m}}{\alpha_i} \ln\left(Z^{(i)}_{(m+1)}/b^{(i)}_m\right) + \left(\frac{\hat{\alpha}_{i,m} - \alpha_i}{\alpha_i}\right) \ln\left(\frac{m}{n}\right).\end{aligned}$$

By Lemma C.1 of Hill (2006a) $(m/n)(b^{(i)}_m)^{\alpha_i} = c_i(z) + o(1/\sqrt{m})$. Therefore, using $\hat{\alpha}_{i,m} = \alpha_i + O_p(1/\sqrt{m})$

$$\begin{aligned}\ln\left(\hat{c}_{i,m}(z)/c_i(z)\right) &= \left(\frac{\hat{\alpha}_{i,m} - \alpha_i}{\alpha_i}\right) \ln c_i(z) + \left(\frac{\hat{\alpha}_{i,m} - \alpha_i}{\alpha_i}\right) \ln\left(\frac{m}{n}\right) + o_p(1/\sqrt{m}) \\ &= O_p(\ln(m/n)/\sqrt{m}).\end{aligned}$$

■

Proof of Theorem 2. For claim (i), let $\pi \in \mathbb{R}^k$, $\pi'\pi = 1$, and define the sequence $\{b_{i,m}\}$ by $(n/m)P(W_{i,t} > b_{i,m}) \to 1$. Define for any $v \in \mathbb{R}$

$$\begin{aligned}&U_{i,m,t} = (\ln W_{i,t}/b_{i,m})_+ - E(\ln W_{i,t}/b_{i,m})_+ \\ &U^*_{i,m,t}(v/\sqrt{m}) = I(W_{i,t} > b_{i,m}e^{v/\sqrt{m}}) - E[I(W_{i,t} > b_{i,m}e^{v/\sqrt{m}})],\end{aligned}$$

and $\sigma_n^2(\pi) = E(m^{-1/2}\sum_{i=1}^k \pi_i \sum_{t=1}^n [U_{i,m,t} - \alpha^{-1}U^*_{i,m,t}(v/\sqrt{m})])^2$. From Lemmas A.1 and A.2

$$\begin{aligned}&\sqrt{m}\sum_{i=1}^k \pi_i\left(\hat{\alpha}_{i,m}^{-1} - \alpha_i^{-1}\right)\\ &= m^{-1/2}\left(\sum_{i=1}^k \pi_i \sum_{t=1}^n \left(U_{i,m,t} - \alpha^{-1}U^*_{i,m,t}(v/\sqrt{m})\right)\right) + o_p(1)\\ &\Rightarrow N(0, \lim_{n\to\infty}\sigma_n^2(\pi)).\end{aligned}$$

Lemma A.1 implies $|\pi'\Sigma_m\pi - \sigma_n^2(\pi)| \to 0$ and $\sigma_n^2(\pi) = O(1)$ by Lemma A.2, hence $|\Sigma_m| = O(1)$. A Cramér-Wold device delivers the desired joint limit.

Claim (ii) follows easily from the scalar case deliverd in Theorem 6 of Hill (2005), and standard arguments, cf. Newey and West (1987). ■

LEMMA A.1 (Hill 2006a) *Under the conditions of Theorem 2, for each* $i = 1...k$

$$\sqrt{m}\left(\hat{\alpha}_{i,m}^{-1} - \alpha_i^{-1}\right) = m^{-1/2}\sum_{t=1}^n \left[U_{i,m,t} - \alpha_i^{-1}U^*_{i,m,t}(v/\sqrt{m})\right] + o_p(1).$$

LEMMA A.2 (Hill 2006a) *For any* $\pi \in \mathbb{R}^k$, $\pi'\pi = 1$, *define*

$$\sigma_n^2(\pi) = E\left(\sum_{i=1}^k \pi_i\sqrt{m}\left(1/m\sum_{i=1}^m \left[U_{i,m,t} - \alpha_i^{-1}U^*_{i,m,t}(v/\sqrt{m})\right]\right)\right)^2.$$

Under the conditions of Theorem 2, $\sigma_n^2(\pi) = O(1)$ *and*

$$\sum_{i=1}^k \pi_i\sqrt{m}\left(1/m\sum_{i=1}^m \left[U_{n,t} - \alpha_i^{-1}U^*_{n,t}(v/\sqrt{m})\right]\right) \Rightarrow N(0, \lim_{n\to\infty}\sigma_n^2(\pi)).$$

Proof of Theorem 3. We prove the limit for the two-tailed coefficient $\hat{\rho}_m^{(0)}(h)$. Proofs of the one-tailed $\hat{\rho}_m^{(i)}(h)$, $i = 1, 2$, follow similarly.

Expand $\hat{\rho}_m^{(0)}(h)$ around α_0: by the mean-value theorem for each n there exist stochastic sequences $\{\alpha_{m,*}(x), \alpha_{m,*}(\tilde{y}_1)\}$ satisfying $\alpha_{m,*}(x) \in (\hat{\alpha}_{0,m}(x), \alpha_0)$, $\alpha_{m,*}(\tilde{y}_1) \in (\hat{\alpha}_{0,m}(\tilde{y}_1), \alpha_0)$, and

$$\begin{aligned}\hat{\rho}_m^{(0)}(h) &= 1 - \frac{(m/n)\left(\tilde{Y}^{(0)}_{h,(m+1)}\right)^{\hat{\alpha}_0(\hat{y}_t)}}{2(m/n)\left(X^{(0)}_{(m+1)}\right)^{\hat{\alpha}_0(x)}}\\ &= 1 - \frac{(m/n)\left(\tilde{Y}^{(0)}_{h,(m+1)}\right)^{\alpha_0}}{2(m/n)\left(X^{(0)}_{(m+1)}\right)^{\alpha_0}}\\ &\quad - \frac{(m/n)\left(\tilde{Y}^{(0)}_{h,(m+1)}\right)^{\alpha_{m,*}(\tilde{y}_1)}\left(\ln Y^{(0)}_{h,(m+1)}\right)}{2(m/n)\left(X^{(0)}_{(m+1)}\right)^{\alpha_{m,*}(x)}} \times \left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right)\\ &\quad + \frac{(m/n)\left(\tilde{Y}^{(0)}_{h,(m+1)}\right)^{\alpha_{m,*}(\tilde{y}_1)}\left(\ln X^{(0)}_{(m+1)}\right)}{2(m/n)\left(X^{(0)}_{(m+1)}\right)^{\alpha_{m,*}(x)}} \times \left(\hat{\alpha}_{0,m}(x) - \alpha_0\right).\end{aligned}$$

For each $Z_t \in \{X_t, \tilde{Y}_{h,t}\}$ define

$$\check{c}_0(z) \equiv (m/n)\left(Z^{(0)}_{(m+1)}\right)^{\alpha_0}, \quad \check{c}_{0,*}(z) \equiv (m/n)\left(Z^{(0)}_{(m+1)}\right)^{\alpha_{m,*}(z)}$$

$$\check{\rho}^{(0)}(h) \equiv 1 - \check{c}_0(\tilde{y}_h)/2\check{c}_0(x), \quad \rho^{(0)}_*(h) \equiv 1 - \check{c}_{0,*}(\tilde{y}_h)/2\check{c}_{0,*}(x).$$

Rearranging terms, and noting

$$\ln Z^{(0)}_{h,(m+1)} = \alpha_0^{-1} \ln \check{c}_0(z) + \alpha_0^{-1} \ln(n/m),$$

we obtain

$$\begin{aligned}
&\frac{\sqrt{m}}{(\ln n/m)}\left(\hat{\rho}^{(0)}_m(h) - \rho^{(0)}(h)\right) \\
&\quad = \frac{\sqrt{m}}{(\ln n/m)}\left(\check{\rho}^{(0)}(h) - \rho^{(0)}(h)\right) \\
&\quad - \left(1 - \rho^{(0)}_*(h)\right)\alpha_0^{-1}\left(\ln \check{c}_0(\tilde{y}_h)\right)\frac{\sqrt{m}}{(\ln n/m)}\left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right) \\
&\quad - \left(1 - \rho^{(0)}_*(h)\right)\alpha_0^{-1}\sqrt{m}\left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right) \\
&\quad + \left(1 - \rho^{(0)}_*(h)\right)\alpha_0^{-1}\left(\ln \check{c}_0(x)\right)\frac{\sqrt{m}}{(\ln n/m)}\left(\hat{\alpha}_{0,m}(x) - \alpha_0\right) \\
&\quad + \left(1 - \rho^{(0)}_*(h)\right)\alpha_0^{-1}\sqrt{m}\left(\hat{\alpha}_{0,m}(x) - \alpha_0\right).
\end{aligned}$$

Lemma C.2 of Hill (2006a) states $\check{\rho}^{(0)}(h) = \rho^{(0)}(h) + O_p(1/\sqrt{m})$, by Theorem 1 $\hat{\alpha}_{0,m}(z) = \alpha_0 + O_p(1/\sqrt{m})$, and $\check{c}_{i,*}(z) = c_i(z) + o_p(1)$ by Lemma C.1 of Hill (2006a). Therefore, Theorem 2 and the continuous mapping theorem give

$$\begin{aligned}
(5) \qquad &\frac{\sqrt{m}}{(\ln n/m)}\left(\frac{\hat{\rho}^{(0)}_m(h) - \rho^{(0)}(h)}{1 - \rho^{(0)}(h)}\right) \\
&= \alpha_0^{-1}\sqrt{m}\left(\hat{\alpha}_{0,m}(x) - \alpha_0\right) - \alpha_0^{-1}\sqrt{m}\left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right) + o_p(1) \\
&\Rightarrow N\left(0, v_0^2\right),
\end{aligned}$$

where $v_0^2 = \alpha_0^{-2} A' \tilde{\Sigma}^{(0,0)} A$, $A = [1, -1]'$. Notice $\tilde{\Sigma}^{(0,0)} = \alpha_0^4 \Sigma^{(0,0)}$ follows from Remark 2 of Theorem 2 and convolution tail equivalence. Hence $v_0^2 = \alpha_0^2 A' \Sigma^{(0,0)} A > 0$ if $\Sigma^{(0,0)}$ is positive definite.

Because each $\hat{\rho}^{(0)}_m(h)$ is stochastically grounded on the same random variables $\hat{\alpha}_{0,m}(x)$ and $\hat{\alpha}_{0,m}(\tilde{y}_1)$, the joint distribution under the null $\rho^{(0)}(j) = 0$, $j = 1...h$, is simply

$$\frac{\sqrt{m}}{(\ln n/m)}\left(\hat{\rho}^{(0)}_{m,h} - \rho^{(0)}_h\right)' \Rightarrow N\left(0, v_0^2 \times [1_h \times 1_h']\right).$$

Under the maintained assumptions $\hat{\rho}^{(0)}_m(j) \to \rho^{(0)}(j)$ by Theorem 1. Hence, $m \times \hat{\rho}^{(0)}_m(j)^2 \to \infty$ with probability one if $\rho^{(0)}(j) \neq 0$. ■

Proof of Theorem 4. Let $\rho^{(2)}(j) - \rho^{(1)}(j) = 0$, and define

$$\tilde{A}(j) \equiv \left[(\rho^{(2)}(j)+1)\alpha_2^{-1}, -(\rho^{(2)}(j)+1)\alpha_2^{-1}, -(\rho^{(1)}(j)+1)\alpha_1^{-1}, (\rho^{(1)}(j)+1)\alpha_1^{-1}\right]'.$$

By imitating the logic of the line of proof of Theorem 3, from the mean-value-theorem, Theorem 2 and the continuous mapping theorem

$$\begin{aligned}
&\frac{\sqrt{m}}{(\ln n/m)}\left[\left(\hat{\rho}_m^{(2)}(j) - \rho^{(2)}(j)\right) - \left(\hat{\rho}_m^{(1)}(j) - \rho^{(1)}(j)\right)\right] \\
&\quad = \left(\rho^{(2)}(j)+1\right)\alpha_2^{-1}\sqrt{m}\left(\hat{\alpha}_{2,m}(y_1) - \alpha_2\right) \\
&\qquad - \left(\rho^{(2)}(j)+1\right)\alpha_2^{-1}\sqrt{m}\left(\hat{\alpha}_{2,m}(x) - \alpha_2\right) \\
&\qquad - \left(\rho^{(1)}(j)+1\right)\alpha_1^{-1}\sqrt{m}\left(\hat{\alpha}_{1,m}(y_1) - \alpha_1\right) \\
&\qquad + \left(\rho^{(1)}(j)+1\right)\alpha_1^{-1}\sqrt{m}\left(\hat{\alpha}_{1,m}(x) - \alpha_1\right) + o_p(1) \Rightarrow N(0, v_j^2),
\end{aligned}$$

where

$$v_j^2 = \tilde{A}(j)'\tilde{\Sigma}^{(1,2)}\tilde{A}(j) > 0, \quad \tilde{\Sigma}^{(1,2)} = \lim_{n\to\infty}\tilde{\Sigma}_m^{(1,2)}.$$

Remark 2 of Theorem 2 and convolution tail equivalence imply

$$\tilde{\Sigma}^{(1,2)} = \operatorname{diag}\{\alpha_2^2, \alpha_2^2, \alpha_1^2, \alpha_2^2\} \times \Sigma^{(1,2)} \times \operatorname{diag}\{\alpha_2^2, \alpha_2^2, \alpha_1^2, \alpha_2^2\},$$

hence $v_j^2 = A(j)'\Sigma^{(1,2)}A(j)$, where

$$A(j) \equiv \left[(\rho^{(2)}(j)+1)\alpha_2, -(\rho^{(2)}(j)+1)\alpha_2, -(\rho^{(1)}(j)+1)\alpha_1, (\rho^{(1)}(j)+1)\alpha_1\right]'.$$

Notice $v_j^2 > 0$ by positive definiteness of $\Sigma^{(1,2)}$, and $A(j) \neq 0$ given $\rho^{(i)}(j) > -1$. The remaining joint limit proof mimicks the proof of Theorem 3. ■

Proof of Theorem 5. Theorems 1 and 2, the continuous mapping theorem, Cramér's Theorem, the assumptions $\rho^{(0)}(j) = 0\ \forall j \geq 1$ and $\sum_{j=1}^{h_n} w(j) \to 1$, and (5) imply

$$\begin{aligned}
Q_m^{(0)}(h_n) &= \left(m/(\ln n/m)^2\right)\hat{v}_{0,m}^{-2}\sum_{j=1}^{h_n} w(j) \times \hat{\rho}_m^{(0)}(j)^2 \\
&= \hat{v}_{0,m}^{-2}\sum_{j=1}^{h_n} w(j) \times \alpha_0^{-1}\left(\sqrt{m}\left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right) - \sqrt{m}\left(\hat{\alpha}_{0,m}(x) - \alpha_0\right) + o_p(1)\right)^2 \\
&= \left(\sqrt{m}\left(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0\right) - \sqrt{m}\left(\hat{\alpha}_{0,m}(x) - \alpha_0\right)\right)^2 \times v_0^{-2} \times \sum_{j=1}^{h_n} w(j) + o_p(1) \\
&\Rightarrow \chi^2(1).
\end{aligned}$$

If $\rho^{(0)}(j) \neq 0$ for some $j \geq 1$ and $w(j) > 0\ \forall j \geq 1$, then Theorem 1 implies

$$\begin{aligned}
Q_m^{(0)}(h_n)/\left(m/(\ln n/m)^2\right) &= \hat{v}_{0,m}^{-2}\sum_{j=1}^{h_n} w(j) \times \rho^{(0)}(j)^2 + o_p(1) \\
&\to v_0^{-2}\sum_{j=1}^{\infty} w(j) \times \rho^{(0)}(j)^2 > 0.
\end{aligned}$$

An identical argument applies to each $Q_m^{(i)}(h_n)$. ■

Proof of Lemma 6.

Step 1 ($\hat{\alpha}_{i,m}$): From Theorem 1 $\hat{\alpha}_{i,\tilde{m}}(z) - \alpha_i = O_p(1/\sqrt{\tilde{m}})$, and Lemma C.3 of Hill (2006a) states $\hat{\alpha}_{i,m}(z) = \hat{\alpha}_{i,\tilde{m}}(z) + o_p(1/\sqrt{\tilde{m}})$ for any sequences $(m, \tilde{m}) \in S_m$.

Step 2 ($\hat{\rho}_m^{(i)}$): Under the maintained assumptions, (5) implies $\forall m \in S_m$

$$\left(\sqrt{m}/(\ln n/m)\right)\left(\hat{\rho}_m^{(0)}(h) - \rho^{(0)}(j)\right)$$
$$= \alpha_0^{-1}\sqrt{m}\,(\hat{\alpha}_{0,m}(\tilde{y}_1) - \alpha_0) - \alpha_0^{-1}\sqrt{m}\,(\hat{\alpha}_{0,m}(x) - \alpha_0) + o_p(1)$$

Therefore, for each $\forall(m, \tilde{m}) \in S_m$

$$(\sqrt{m}/(\ln n/m))\hat{\rho}_m^{(0)}(j) - (\sqrt{\tilde{m}}/(\ln n/\tilde{m}))\ \hat{\rho}_{\tilde{m}}^{(0)}(j)$$
$$= \alpha_0^{-1}\left[\sqrt{m}\,(\hat{\alpha}_{0,m}(x) - \hat{\alpha}_{0,\tilde{m}}(x)) + \sqrt{\tilde{m}}\,(\hat{\alpha}_{0,\tilde{m}}(x) - \alpha_0)\,(\sqrt{m}/\sqrt{\tilde{m}} - 1)\right]$$
$$- \alpha_0^{-1}\left[\sqrt{m}\,(\hat{\alpha}_{0,m}(\tilde{y}_1) - \hat{\alpha}_{0,\tilde{m}}(\tilde{y}_1)) + \sqrt{\tilde{m}}\,(\hat{\alpha}_{0,\tilde{m}}(\tilde{y}_1) - \alpha_0)\,(\sqrt{m}/\sqrt{\tilde{m}} - 1)\right] + o_p(1).$$

The claim $(\sqrt{m}/(\ln n/m))\hat{\rho}_m^{(0)}(j) - (\sqrt{\tilde{m}}/(\ln n/\tilde{m}))\ \hat{\rho}_{\tilde{m}}^{(0)}(j) = o_p(1)$ now follows from Step 1 and $m/\tilde{m} - 1 = o(1)$. An identical argument holds for each $\hat{\rho}_m^{(i)}$.

Step 3 ($\hat{v}_{i,m}$ and $\hat{v}_m(j)$): Consider $\hat{v}_{i,m}$: the same logic, along with Step 2, applies to $\hat{v}_m(j)$. Using Step 1 we need only show $\hat{\Sigma}_m^{(i,i)} = \hat{\Sigma}_{\tilde{m}}^{(i,i)} + o_p(1)$. Theorem 3 implies $|\hat{\Sigma}_m^{(i,i)} - \Sigma_m^{(i,i)}| = o_p(1)$ for all $m \in S_m$, and $|\Sigma_m^{(i,i)} - \Sigma_{\tilde{m}}^{(i,i)}| = o_p(1)$ from Lemma C.4 of Hill (2006a). Now apply the triangular inequality. ■

REFERENCES

[1] Basrak, B., R.A. Davis, and T, Mikosch (2002). A Characterization of Multivariate Regular Variation, Annals of Applied Probability 12, 908-920.

[2] Beirlant, J. Y. Goegebeur, J. Segers, J. Teugels, D. De Waal (2004). Statistics of Extremes: Theory and Applications (John Wiley & Sons: New York).

[3] Bingham, N. H., C. M. Goldie and J. L. Teugels (1987). *Regular Variation* (Cambridge Univ. Press: Great Britain).

[4] Breiman, L. (1965). On Some Limit Theorems Similar to the Arc-Sine Law, *Theory of Probability and its Applications* 10, 351–360.

[5] Caner, M. (1998). Tests for Cointegration with Infinite Variance Errors, *Journal of Econometrics* 86,155-175.

[6] Chernozhukov, V. (2005). Extremal Quantile Regression, *Annals of Statistics* 33, 806–839.

[7] Clements, M. and J. Smith (2001). Evaluating Forecasts from SETAR Models of Exchange Rates, *Journal of International Money and Finance* 20, 133-148.

[8] Cline, D.B.H. (1986). Convolution Tails, Product Tails and Domains of Attraction,

Probability Theory and Related Fields 72, 529–557.
[9] Coles, S., J. Hefferman, and J.A. Tawn (1999). Dependence Measures for Extreme Value Analyses, Extremes 2, 339-365.
[10] Datta, s. and W.P. McCormick (1998). Inference for the Tail Parameters of a Linear Process with Heavy Tail Innovations. Annals of the Institute of Stati*stical Mathematics* 50, 337-359.
[11] Davidson, J. (1994). Stochastic Limit Theory (Oxford Univ. Press: Oxford).
[12] Davis, R., and S. Resnick (1996). Limit Theory for Bilinear Processes with Heavy-Tailed Noise, *Annals of Applied Probability* 6, 1191–1210.
[13] Davis, R., and W. Wu (1997). Bootstrapping M-Estimates in Regression and Autoregression with Infinite Variance, *Staistica Sinica* 7, 1135-1154.
[14] de Jong, R.M., and J. Davidson (2000). Consistency of Kernel Estimators of Heteroscedastic and Autocorrelated Covariance Matrices, *Econometrica* 68, 407-423.
[15] Drees, H. (2003). Extreme Quantile Estimation for Dependent Data, with Applications to Finance, *Bernoulli* 9, 617-657.
[16] Embrechts, P. and C. M. Goldie (1982). On Convolution Tails, *Stochastic Processes and their Applications* 13, 263-278.
[17] Embrechts, P., Klueppelberg, C. and Mikosch, T. (2003). Modelling Extremal Events for Insurance and Finance (Springer-Verlag).
[18] Fama, E. (1965). Portfolio Analysis in a Stable Paretian Market, *Management Science* 11, 404-419.
[19] Feller, William (1971). *An Introduction to Probability Theory and its Applications*, 2nd ed., vol. 2, (Wiley: New York).
[20] Gallant, A. R. and H. White (1988). *A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models* (Basil Blackwell: Oxford).
[21] Geluk, J., P. Liang and C. G. de Vries (2000). Convolutions of Heavy Tailed Random Variables and Applications to Portfolio Diversification and MA(1) Time Series. *Advances in Applied Probability* 32, 1011 - 1026.
[22] Haeusler, E., and J. L. Teugels (1985). On Asymptotic Normality of Hill's Estimator for the Exponent of Regular Variation, *Annals of Statistics* 13, 743-756.
[23] Hall, P. (1982). On Some Estimates of an Exponent of Regular Variation, *Journal of the Royal Statistical Society* 44, 37-42.
[24] Hall, P. and Q. Yao (2003). Inference in Arch and Garch Models with HeavyTailed Errors , *Econometrica* 71, 285-318.
[25] Hefferman, J.E. and J.A. Tawn (2004). A Conditional Approach for Multivariate Extreme Values, *Journal of the Royal Statistical Society Ser. B* 66, 497-546.
[26] Hill, B.M. (1975). A Simple General Approach to Inference about the Tail of a Distribution, *Annals of Mathematical Statistics* 3, 1163-1174.
[27] Hill, J.B. (2005). On Tail Index Estimation for Dependent, Heterogenous Data, Florida International; available at http:// econwpa.wustl.edu:80/eps/em /papers/ 0505/0505005.pdf.
[28] Hill, J.B. (2006). Technical Appendix for "Gaussian Tests of 'Extremal White Noise' for Dependent, Heterogeneous, Heavy Tailed Stochastic Processes with an Application", Florida International University; available at http://www.fiu.edu/~hilljona/ tech_append_co_rel.pdf.
[29] Hsing, T. (1991). On Tail Index Estimation Using Dependent Data, *Annals of*

Statistics 19, 1547-1569.
[30] Ibragimov, I. A. and Y. V. Linnik (1971). *Independent and Stationary Sequences of Random Variables* (Wolters-Noordhof).
[31] Jansen, D. and C. de Vries (1991). On the Frequency of Large Stock Returns: Putting Booms and Busts into Perspective, *The Review of Economics and Statistics* 73, 18-24.
[32] Jondeau, E. and M. Rockinger (2003). Testing the Difference in the Tails of Stock Market Returns, *Journal of Empirical Finance* 9, 1-23.
[33] Knight, K. (1993) Estimation in Dynamic Linear Regression Models with Infinite Variance Errors, *Econometric Theory* 9, 570-588.
[34] Kräger, H. and p. Kugler (1993). Non-Linearities in Foreign Exchange Markets: A Different Perspective, *Journal of International Money and Finance* 12, 195-208.
[35] Kokoszka, P. S. and M. S. Taqqu (1996). Parameter Estimation for Infinite Variance Fractional ARIMA, *Annals of Statistics* 24, 1880-1913.
[36] Leadbetter, M.R., G. Lindgren and H. Rootzén (1983). *Extremes and Related Properties of Random Sequences and Processes* (Springer-Verlag: New York).
[37] Ledford, A. W. and J. A. Tawn (1996). Statistics for Near Independence in Multivariate Extreme Values, *Biometrika* 83, 169-187.
[38] Ledford, A. W. and J. A. Tawn (1997). Modeling Dependence within Joint Tail Regions, *Journal of the Royal Statistical Society Ser B* 59, 475-499.
[39] Leipus R; and M. Viano (2000). Modelling Long-memory Time Series with Finite or Infinite Variance: a General Approach, *Journal of Time Series Analysis* 21 61-74.
[40] Ljung, G.M. and G.E.P. Box (1978). On a Measure of Lack of Fit in Time Series Models, *Biometrika* 65, 297-303.
[41] Linton, O. and Y-J Wang (2004). A Quantilogram Approach to Evaluating Directional Probability, London School of Economics.
[42] Mikosch, T., T. Gadrich, C. Klüppelberg, and R.J. Adler (1995). Parameter Estimation for ARMA Models with Infinite Variance, *Annals of Statistics* 23, 305-326.
[43] Newey, W.K., and K.D. West (1987) A Simple, Positive Semi-Definite, Heteroscedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* 55, 703-708.
[44] Poon, S., M. Rockinger and J. Tawn (2001). New Extreme Value Dependence Measures and Finance Applications, Les Cahiers de Recherche, Groupe HEC.
[45] Resnick, S. (1987). *Extreme Values, Regular Variation and Point Processes* (Springer-Verlag: New York).
[46] Schmidt R. and U. Stadtmuller (2006). Non-Parametric Estimation of Tail Dependence, *Scandinavian Journal of Statistics* 33, 307-335.
[47] Sibuya, M. (1960). Bivariate Extremal Statistics. *Ann. Ins. Statist. Math. XI*, 195-210.
[48] Stărică, C. (1999). Multivariate Extremes for Models with Constant Conditional Correlations, *Journal of Empirical Finance* 6, 515-553.
[49] Tawn, J. (1990). Modeling Multivariate Extreme Value Distributions, *Biometrika* 77, 245-253.
[50] Teräsvirta, T. (1994). Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, *Journal of the American Statistical Association* 89, 208-

Table 1. Two-Tailed Tests: $Q^{(0)}_{m^{(r)}}(h)$, $\alpha = 1.5$

Process	$n \setminus r^a$	1	5	10	15	20	25	30	60	Av Q[b]	$\hat{\alpha}_{0,m} \pm k_{0,m}$
EWN[d]	400	.00[e]	.00	.05	.07	.11	.15	.28	.81	.372	1.57±0.25[c]
	600	.00	.00	.00	.00	.00	.06	.13	.70	.407	1.54±0.19
MA(1)[f]	400	.97	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.000	1.59±0.41
	600	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.000	1.57±0.32
AR(1)[g]	400	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.000	2.02±1.24
	600	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.000	1.75±0.61
HYARCH[h]	400	.01	.02	.05	.14	.22	.34	.42	.78	.592	1.54±0.26
	600	.00	.01	.04	.07	.18	.24	.35	.57	.416	1.55±0.22

Notes: a. Co-relation rank.

b. The p-value of a rank- and h-averaged $Q^{(0)}_{m^{(r)}}(h)$, r = [.005n]...[.05], h = 1...4.

c. Average of $\hat{\alpha}_{0,m_i^{(r)}} \pm 1.96\hat{\hat{\sigma}}_{0,m_i^{(r)}}/(m_i^{(r)})^{1/2}$ over r = [.005n]...[.05n], i = 1...4.

d. The "extremal white noise" process $X_{n,t} = \phi_1 X_{n,t-1} I(|u_{t-1}| \leq n^{1/5}) + u_t$.

e. Frequencies are for the *maximum* Q-statistic over h = 1...4.

f. $X_t = .9u_{t-1} + u_t$. g. $X_t = .9X_{t-1} + u_t$.

h. $X_t = \sigma_t u_t$, $\sigma_t^p = \theta_0 + \sum_{i=1}^{L_n} \theta_i |X_{t-i}|^{1.2}, \theta_i = i^{-2}$, $\alpha = 1.5$, and $L_n = [.25n]$.

Table 2. Difference in Tails Tests: $\Delta Q_{m^{(r)}}(h)$, $\alpha = 1.5$

Process	$n \setminus r$	1	5	10	15	20	25	30	60	100	150	200	Av Q[a]	$\hat{\alpha}_{0,m} \pm k_{0,m}$
EWN	1000	.02	.03	.04	.05	.05	.08	.09	.12	.18	.31	.46	.365	1.57±.21
	1200	.02	.03	.03	.04	.04	.05	.06	.08	.12	.18	.23	.345	1.56±.16
AR(1)	1000	.03	.06	.08	.13	.18	.24	.24	.24	.24	.29	.41	.167[b]	1.49±.09
	1200	.02	.04	.05	.08	.10	.17	.17	.21	.21	.24	.37	.161	1.56±.11
SETAR	1000	.78	.81	.84	.85	.86	.91	.91	.92	.93	.94	.98	.000	1.59±.18
	1200	.92	.94	.99	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	.000	1.66±.16

Notes: a. p-value of rank- and h -averaged $\Delta Q_{m^{(r)}}(h)$, r = [.005]...[.05n], h = 1...4.

b. The p-value it rather small because we are averaging over ranks up to [.05n]. When we average over r = [.005]...[.01n], the p-values are .311 and .365 respectively.

Table 3. Sample Co-Relations ($\alpha = 1.5$)

	Two-Tailed[a]			Difference in Tails		
h	EWN	MA(1)	AR(1)	EWN	AR(1)	SETAR
1	.002±.24[b]	.455±.04	.855±.09	.005±.198	.076±.21	1.13±.28
	(.000)[c]	(.491)	(.911)	(.000)	(.000)	(.911)
2	-.001±.23	-.000±.05	.801±.07	.003±.160	.083±.23	1.08±.32
	(.000)	(.000)	(.823)	(.000)	(.000)	(.823)
3	.000±.21	.001±.05	.711±.06	.006±.162	.055±.21	1.09±.44
	(.000)	(.000)	(.741)	(.000)	(.000)	(.741)
4	-.005±.24	.000±.05	.624±.06	.001±.167	.049±.21	0.96±.51
	(.000)	(.000)	(.665)	(.000)	(.000)	(.665)
5	.003±.25	.001±.05	.535±.05	-.001±.162	.054±.22	0.92±.48
	(.000)	(.000)	(.596)	(.000)	(.000)	(.596)

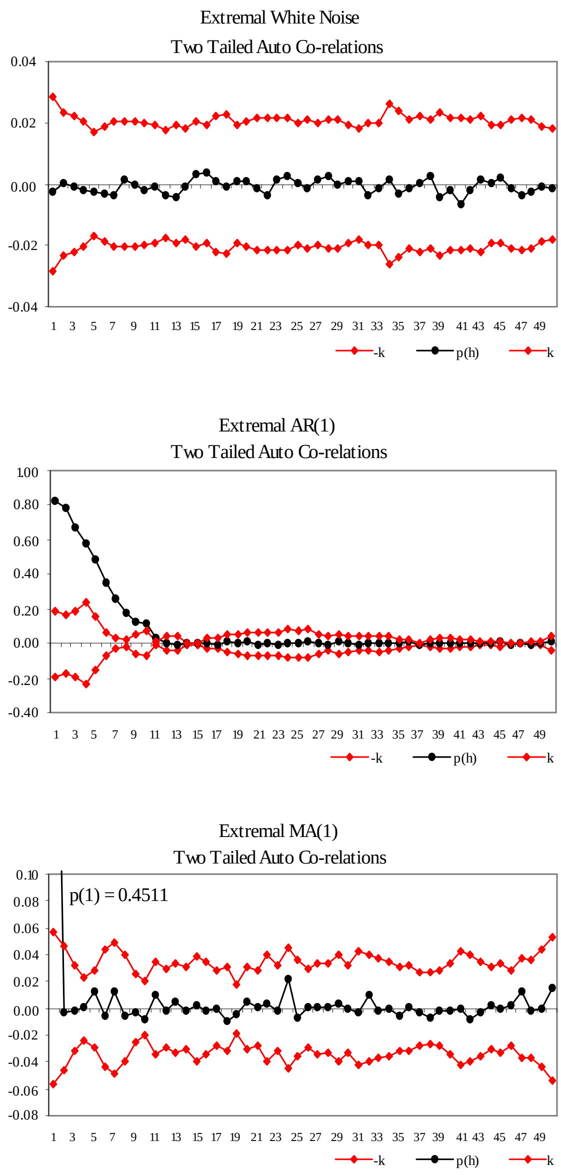
Notes: a. For two-tailed co-relations, n = 500; for difference-in-tails, n = 1000.
b. Co-relation bands are averages over ranks r = [.005n],...,[.05n].
c. True co-relation: see Example 1 of Section 3.4.

Table 4. Daily USD Exchange Rates and Asset Markets

	ΔYEN		ΔEURO		ΔBP[a]	
h	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$
1	-.004±.02	-.001±.06	.002±.02	-.005±.07	-.002±.02	-.058±.08
2	-.000±.03	-.004±.07	-.010±.02	.001±.04	.011±.02	-.055±.04*[b]
3	.010±.02	-.000±.05	.003±.01	-.002±.07	-.018±.03	-.007±.04
4	-.003±.02	-.002±.05	.008±.02	.000±.06	-.004±.02	.000±.07
Av Q	.044	.875	.009	.823	.065	.154
$\hat{\alpha}_{0,m} \pm k_{0,m}$	1.90±.196		1.97±.221		1.89±.220	
	ΔNASDAQ		ΔSP500		ΔSSE[c]	
h	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$	$\hat{\rho}_m^{(0)}$	$\Delta\hat{\rho}_m$
1	.062±.01*	.001±.08	.426±.01*	.000±.14	.012±.02	.002±.08
2	.003±.01	.000±.07	.003±.01	.000±.05	-.001±.02	.031±.06
3	.005±.01	.005±.08	-.005±.01	.001±.06	-.012±.02	.000±.07
4	.024±.02*	.001±.06	-.007±.02	-.001±.04	.012±.02	-.004±.11
Av Q	.000	.762	.000	.848	.060	.591
$\hat{\alpha}_{0,m} \pm k_{0,m}$	2.00±.694		1.99±.307		1.72±.531	

Notes: a. BP = British pound; b. * = significant at the 5% level (see also Figures 2-3).
c. SSE = Shanghai Stock Exchange.

Figure 1
Extremal White Noise,
Extremal-AR(1) and Extremal-MA(1)



Notes: EWN: $X_{n,t} = .9X_{n,t-1}I(|\epsilon_{t-1}| \leq n^{1/5}) + \epsilon_t$
E-AR(1): $X_{n,t} = .9X_{n,t-1}I(|\epsilon_{t-1}| > n^{1/5}) + \epsilon_t$.
E-MA(1): $X_{n,t} = .9\epsilon_{t-1}I(|\epsilon_{t-1}| > n^{1/5}) + \epsilon_t$.

Figure 2
Daily US Dollar Exchange Rates
Auto-Co-Relogram with 95% Confidence Bands

Figure 3
Daily Stock Market Returns
Auto-Co-Relogram with 95% Confidence Bands