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

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

November 4, 2005

Abstract

We develop asymptotically chi-squared tests of tail specific extremal serial dependence for possibly heavy-tailed time series, including infinite variance and infinite mean processes. Our test statistics have a chi-squared limit distribution under the null of "*extremal white-noise*" for processes near-epoch-dependent on a mixing process; and obtain a power of one for extremal dependent processes under general conditions. We restrict the NED property to hold only in the extreme support of the distribution, and characterize a broad array of linear and GARCH processes with geometric or hypoberbolic memory that are extremal NED. We apply one-tailed, two-tailed, and difference in tails tests to stock market and exchange rate returns data, and find low levels of significant, persistent, symmetric extremal dependence in the Yen and British Pound, and except for the Shanghai Stock Exchange we find no evidence of extremal dependence in any absolute returns series. A limited study of bivariate volatility spillover in exchange rates reveals extremes in the daily returns of the Yen symmetrically spillover briefly into the Euro after a four day delay, and positive extreme returns in the Euro immediately, and persistently, spillover into the Yen.

1. Introduction An "extreme" value in a stochastic process $\{X_t\}$ occurs when an observation X_t surpasses a threshold (positive or negative) and

*I would like to thank the participants of the European Meeting of Statisticians (Olso, 2005), John Galbraith and the participants of the Econometric Society World Congress (London, 2005), and Oliver Linton for helpful discussions. The comments and reference suggestions of two anonymous readers are gratefully acknowledged. All errors are 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; volatility spillover; near-epoch-dependence; regular variation; infinite variance; portmanteau test; exchange rates.

the threshold is allowed to expand to (positive or negative) infinity. See, e.g., Resnick (1987) and Rachev (2003). Extreme values occur in the macroeconomic and financial contexts of hyperinflation, asset market bubbles, collapses, and extreme value theory has gained rapid popularity in the economics (exchange rates), finance (value-at-risk), and telecommunications (network activity) literatures. In this paper we develop an extremal serial dependence measure and asymptotically chi-squared test statistic applicable to possibly heavy-tailed, dependent and heterogenous time series. 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 $h = 1, 2, ...$

When and why a time series obtains extreme values are of acute interest to empiricists and policy makers. Financial performance and risk management (e.g. value-at-risk), the fundamental ability to forecast a time series, formulate central bank policy and regulate financial markets, are all obviously intimately tied to volatility and serial dependence. An apparently noisy returns series like fluctuations in the daily Yen-Dollar rate may exhibit highly persistent serial extremes in either tail. The canonical efficient market hypothesis that today's rate best predicts tomorrow's, given *all available* information, may not be supported when today's rate represents an extreme value. All such extremal information should be used to help formalize a forecast model of extremes in exactly the same way the Box-Jenkins method is used for finite variance time series.

The leptokurtic properties of many time series have long been hypothesized to be driven by infinite variance/kurtosis data generating processes[1]. Moreover, the serial clustering of large movements in time series is predominantly couched in terms of a GARCH model, cf. Engle (1982) and Bollerslev (1986), including models which control for asymmetries and nonlinearities in volatility[2]. In most cases GARCH tests and models offer only parametric devices for detecting and modeling conditional volatility in time series, and by construction model volatility over the entire support of the distribution even when underlying innovations are assumed to have an infinite variance[3].

We are therefore interested in a non-parametric test of serial dependence for heterogeneous and dependent time series X_t based on extremal serial properties. The distribution tails of the process $\{X_t\}$ may exhibit any degree of tail thickness (or thinness). The theory developed here can be directly applied to broad classes of linear and power-GARCH processes with geometric ("short") or

[1] Examples of heavy tails abound in finance (stock returns; options), macroeconomics (exchange rates; prices), telecommunications (network traffic; cpu time to complete a job) and geophysics (soil diffusion of water), beginning with the seminal investigations of prices by Mandelbrot (1961,1963) and Fama (1965). See, also, Loretan (1991), Loretan and Phillips (1994), McCulloch (1997), and Rachev (2003).

For evidence of volatility clustering, see Mandelbrot (1963), Black (1976), Pagan and Schwert (1990), Cambell and Hentschel (1992), Ding *et al* (1993), and Engle and Ng (1993).

[2] In order to capture traits of serial assymetric leptokurtosis in finance and macroeconomics, smooth transition GARCH (González-Rivera, 1998), Markov switching GARCH (cf. Hamilton, 1989), and quadratic GARCH (Sentana, 1995) are just a few of the proposed models.

[3] See McCulloch (1985), Liu and Brorsen (1995), Mittnik *et al* (2002), and Hall and Yao (2003).

hyperbolic ("long") memory with possibly infinite variance errors, (eventually) augmented to bivariate processes in order to test for extremal volatility spillover (e.g. market contagion), applied to estimated residuals from parametric models, and employed for tests of nonlinearity in extremes. A complete theoretical treatment of bivariate dependence, however, will only deviate our attention from the pure ideas and is left for future endeavors. We do, however, provide a simple special case and a limited empirical study: see below[4].

Our test is based on a scaled difference in extreme tail unconditional marginal probabilities that reduces to a difference in extreme tail parameters, which is similar in spirit to the stable co-difference for stable random variables: see Samorodnitsky and Taqqu (1994). We scale the co-difference in order to generate left-, right-, two-tailed, and difference-in-tails serial "co-relation" coefficients in the spirit of the correlation, and create a portmanteau test statistic in the manner of Box and Pierce (1970) and Hong (2001). A process exhibits "*extremal white noise*" in either or both tails when tail-specific co-relations are identically zero for all displacements.

We use a generalized version of the co-difference because standard measures of dependence for finite variance processes do not exist in the infinite variance case (e.g. the correlation)[5], because measures of non-extremal and/or population dependence are inappropriate (e.g. fixed quantiles, multivariate regression, covariation), a generalized co-difference naturally articulates dependence between extremes, and the co-relation decays according to the memory properties of many linear and power-GARCH processes. Moreover, linear processes with symmetric shocks have inherently symmetric co-relations, hence the co-relation may be used to justify a test a nonlinearity. Nonetheless, like the correlation the co–relation does not have a multivariate counterpart: we only model pair-wise extremal dependence[6].

Use of the co-relation both compliments and augments standard models of memory[7]. A fractionally integrated power-ARCH(∞) process $|X_t|^p$ with series coefficients bounded by a regularly varying function, for example, will have hyperbolically decaying co-relations of the power process (see Section 3.5).

[4] Our dependence measure and the associated asymptotic theory in the bivariate case are fundamentally grounded on the results of the present work, but deviate enough to warrant a separate paper.

[5] Runde (1997) extends the Box Pierce (1970) test to infinite variance processes, but does not characterize the limiting properties of the test statistic under the alternative precisely because the alternative of "serial correlation" is nonsensical. See Davis and Resnick (1985). Even in the finite variance, infinite kurtosis case sample autocovariances of GARCH processes have infinite variances and non-standard limits. See Mikosch and Stărică (2000).

[6] Numerous measures of pair-wise dependence for infinite variance processes exist, including the covariation for stable processes, the association, and copula functionals based on bivariate extreme value theory. See Sibuya (1960), de Haan and Resnick (1977), Rachev and Xin (1993), Samorodnitsky and Taqqu (1994), Mittnik *et al* (1999), and Resnick (2002). Surprisingly few attempts to estimate the covariation and association measures of dependence exist, and when available the asymptotic theory is expidited by harsh restrictions on dependence. See, e.g. Gallagher (2002).

[7] See Kokoska and Taqqu (1996b), Leipus and Viano (2000), and Peng and Yao (2004) for model construction and estimation of long memory infinite variance processes.

This provides an extremal compliment to our understanding of the population memory properties of integrated process. But it does much more than that. A process need not have long or short memory population characteristics, per se: a hyperbolically decaying co-relation of a power process simply suggests the *extremes* of $|X_t|^p$ may be best modeled as a power FIGARCH, irrespective of the data generating process underlying the non-extremes.

This is particularly relevant when a time series like the daily Yen/Dollar exchange rate has noisy returns based on conventional dependence measures (i.e. $\Delta\text{Yen}_t \sim I(0)$), but highly persistent extremes (e.g. $\Delta\text{Yen}_t \sim I(d)$, $0 < d < 1$, when $|\Delta\text{Yen}_t| > \varepsilon$ as $\varepsilon \rightarrow \infty$): the co-relations *must* be zero if $\Delta\text{Yen}_t \sim I(0)$, which casts doubt on the standard $I(0)$ assumption. See Sections 3 and 9. In this sense standard non-extremal models of memory (e.g. population white noise, fractional integration) are not guaranteed to capture all of the forecastable traits of a time series: models of population memory may be misleading at best, or mis-specified at worst.

If we are to believe each set of evidence concerning daily returns of the Yen, we must conjecture a model that implies both the non-extremal returns are a noisy $I(0)$ and the extremal returns are a persistent $I(d)$. It may be useful, then, to develop an Extremal Threshold Autoregression, or Extremal Markov Switching power-GARCH process[8].

Although the co-difference does not have an intuitive regression counterpart[9], extremal dependence evidence may nevertheless be used to suggest an extremal model in the manner of the Box-Jenkins methodology. But this suggests we may study extremal dependence simply be estimating parameters of a given model. This requires knowledge of the specific parametric form, which in turn requires a complete theory of estimation and inference for infinite variance processes. Indeed, model construction and mis-specification analysis is one important use of the present extremal white noise test. Parametric and nonparametric estimators for infinite variance time series models typically have non-standard distribution limits, hence monte carlo techniques are required for critical value derivation for each model type and for each data set. See, e.g., Cline (1983), Knight (1993), Mikosch *et al* (1995), Kokoska and Taqqu (1996b), Davis and Wu (1997), and Hall and Yao (2003)[10].

---

[8] For example, an Extremal SETAR - Extremal Threshold power-GARCH would be $\Delta X_t - \phi\Delta X_{t-1} I(|\Delta X_{t-1}| > b_n) = \sigma_t \epsilon_t$, $|\phi| < 1$, where b_n is some sequence that diverges slowy: $b_n \rightarrow \infty$ as $n \rightarrow \infty$ (e.g. the m^{th} sample order statistic of $\{\Delta X_2, ..., \Delta X_n\}$), $E|\epsilon_t|^p = 1$ for some $0 < p < \alpha$, and σ_t^p is a power volatility process, $\sigma_t^p = \theta_0 + \sum_{i=1}^{\infty} \theta_i |\Delta X_t|^p I(|\Delta X_{t-1}| \leq b_n)$, $\sum_{i=1}^{\infty} \theta_i < 1$. When $|\Delta X_{t-1}| > b_n$ as $n \rightarrow \infty$, $\Delta X_t = \phi\Delta X_{t-1} + \theta_0 \epsilon_t$ is a stationary AR(1) process. When $|\Delta X_{t-1}| \leq b_n$ as $n \rightarrow \infty$, $\Delta X_t = \sigma_t \epsilon_t$, $\sigma_t^p = \theta_0 + \sum_{i=1}^{\infty} \theta_i |\Delta X_t|^p$.

[9] See Engle and Manganelli (2004) for a multivariate non-extremal quantile regression technique. The quantile regressor dimension must be bounded in order to ensure a non-singular asymptotic covariance matrix, and rather strict moment conditions are imposed.

[10] Ling (2005) establishes asymptotic normality of a weighted LAD estimator for stationary autoregressive processes. Using similar techniques, Peng and Yao (2004) find the LAD estimator of the ARFIMA parameters to have non-standard limits. Apparently there does not exist a result guaranteeing a standard limit for the LAD estimator (or any other estimator) of the parameters of a model of any fractionally integrated infinite variance process (e.g. ARFIMA, FIGARCH).

Under the null of extremal white noise our test statistics converge in distribution to a chi-squared random variable for the class of processes near-epoch-dependent on a mixing process. In this case the NED property is only assumed to hold in the extreme tails of the distribution ("extremal NED"), broad classes of "short" (geometric) or "long" (hyperbolic) memory linear or GARCH processes satisfy the extremal NED property, and the same classes of processes have similarly decaying co-relations. The test statistics obtain an asymptotic power of one under the alternative of extremal dependence (non-zero co-relation at some displacement). We use a non-parametric Newey-West type kernel estimator of the co-relation asymptotic variance allowing us to ignore essentially parametric issues like an underlying GARCH structure. To the best of our knowledge this is the first to attempt to provide an asymptotically Gaussian estimator of the co-relation[11].

When applied to stock market and exchange rate daily returns data we find particularly small levels of significant, symmetric, and persistent extremal dependence in the daily returns of the Yen and British Pound; shallow, asymmetric dependence in the Euro and NASDAQ daily returns; and extensive evidence that each absolute returns series, except the Shanghai Stock Exchange, is extremal white noise. Moreover, invoking somewhat strict assumptions under the null we perform a limited study of volatility spillover across exchange rates. We find extreme spikes in the Yen symmetrically spillover into the Euro after a delay of four days, where the spillover is short-lived. Conversely, predominantly positive extremes in the daily returns of the Euro immediately spillover into the Yen with some degree of persistence.

The concept of "extremal dependence" has (evidently) been exclusively applied to bivariate processes. Recent examples include contagion in international equity markets (e.g. Longin and Solnick, 2001; and Forbes and Rigobon, 2002); exchange rates and equity returns (Stărică, 1999; Poon *et al*, 2001; Patton, 2002; Chen *et al*, 2004); and emerging market bonds and equities (Bekeart and Harvey, 1997; Quintos, 2001, 2004)[12]. See, also, Ledford and Tawn (1996, 1997), Peng (1999) and Schmidt and Stadtmüller(2004). The standard procedure is to assume bivariate regular variation in the tails and construct a copula dependence function. The common themes are assumed marginal independence (Ledford and Tawn, 1996,1997; Schmidt and Stadtmüller (2004); assumed parametric conditional volatility structure (Stărică, 1999; Quintos, 2001,2004); covariance stationarity, mild leptokurtosis, and limited dependence under either null or alternative (Stărică, 1999; Quintos, 2001,2004); and/or a limiting distribution has not been established (Stărică, 1999)). In all cases dependence restrictions are enforced over the entire support of the distribution, and an explicit structure for the joint extreme distribution tail is enforced.

[11] Yang *et al* (2001) study population dependence and derive an estimator of a "generalized co-difference" based on characteristic functions, and only prove consistency for processes with hyperbolic memory. Our estimator is asymptotically Gaussian for a broad class of dependent processes that includes the same processes considered in Yang *et al* (2001).

[12] For background theory on extremal dependence, see Sibuya (1960), de Haan and Resnick (1977), Peng (1999), Resnick (2002), Quintos (2004) and Schmidt and Stadtmüller (2004).

Of course moment-free functional transformations may be utilized and standard L_2-dependence measures applied to such transforms. See de Lima (1996) for an extension of the correlation-integral BDS-test; Hong and Chung (2003) and Linton and Wang (2004) and for quantile and threshold indicator techniques; and Engle and Manganelli (2004) for quantile regression in value-at-risk models. Of particular note Linton and Wang (2004) analyze the correlations of quantile hits in order to adduce directional predictability. Their approach is essentially model and moment free, and extends to estimated residuals. Our "extremal white-noise" hypothesis simply fills in where the standard quantile method fails in the distant tails[13]. The decay properties of the quantilogram, however, do not easily relay basic information concerning memory or underlying functional form of the process[14]. See also Drees (2003) for extreme quantile and Chernozhukov (2005) for extreme quantile regression methods. The former is hampered with the same shortcomings as the quantilogram (memory and functional form are not characterized), in addition to the somewhat stronger memory assumption of absolute regularity which is enforced over the entire support of the distribution. The latter method forces one parametric specification (e.g. linear) for the process over all quantiles, and assumes the dependent and regressor variables are *iid* over the entire support of the distribution.

Finally, problems persist based on our use of the popular tail estimators by Hill (1975) and Hall (1982): there exists an order statistic nuisance index m well known to theorists and empiricists familiar with the Hill estimator, and the test statistics typically do not perform well for arbitrarily chosen m. We solve the nuisance index issue by ranking absolute sample co-relation coefficients and selecting for each lag an order index m based on a chosen rank. A simulation experiment demonstrates the superlative properties of the resulting test statistics.

The rest of the paper contains the following topics. In Sections 2 and 3 we detail parametric representations of regularly varying distribution tails, and develop the co-relation measure of dependence. Preliminary asymptotic theory is contained in Section 4, and Section 5 details the test statistic for tests of extremal white noise, and Section 6 define the new property of Extremal Near-Epoch-Dependence. In Section 7 we develop a strategy for handling the nuisance index m. Sections 8 and 9 contain a simulation experiment and an application to various exchange rate and asset returns series. Section 10 contains parting comments. Assumptions are presented in Appendix 1 and Appendix 2 contains proofs of main results. All table and figures are placed at the end of the paper.

We employ the following notation conventions. Denote by $\rightarrow$ convergence in probability, and by $\Longrightarrow$ convergence with respect to finite dimensional distributions. We write $X \equiv \{X_t\} \equiv \{X_t : -\infty \leq t \leq \infty\}$ to denote a stochastic

[13] Note that Linton and Wang's (2004) "quantilegale" definition states there is a θ probabililty that X_t will be below the θ-quantile given *all past information* on X_t, irrespective of whether any past X_{t-h} is actually below the same quantile. Our extremal white-noise hypothesis only relates extremes to extremes.

[14] The authors only discuss AR(1) processes (Linton and Wang, 2004, p. 6-7) and claim the quantilogram even in this case is "quite complicated".

process. $[x]$ denotes the integer part of x, where $[|x|] \leq |x|$ by convention. Let $z_+ \equiv \max\{z, 0\}$.

2. Regular Variation Denote by $X = \{X_t\}$ a random process defined on a proper probability measure space $(\Omega, \Im, P)$, $\Im = \cup_{t \in \mathbb{Z}} \Im_t$, with increasing σ-field $\Im_t = \sigma(X_s : s \leq t)$. Let $\bar{F}(x) \equiv P(X > x)$. We assume the process X has common marginal distribution tails that satisfy the following regular variation property:

$$\text{(1)} \quad \begin{aligned} F(x) &\equiv P(X_t \leq x) = c_1(x)|x|^{-\alpha}(1 + o(1)), \quad x < 0, \\ \bar{F}(x) &\equiv P(X_t > x) = c_2(x)x^{-\alpha}(1 + o(1)), \quad x > 0 \end{aligned}$$

as $|x| \to \infty$, where the "extreme scale" parameters $c_i(x)$ satisfy $c_1(x), c_2(x) \geq 0$, and $\alpha > 0$ denotes the index of regular variation. We say the process is maximally skewed if either $c_i(x) = 0$ (e.g. X_t^2 is maximally positively skewed). We assume $c_2(x) > 0$ as a convention.

When the tail index satisfies $\alpha < 2$ the population variance is infinite, and $E|X_t|^\rho < \infty$ for all $0 < \rho < \alpha$. Processes that abide (1) with $0 < \alpha < 2$ belong to the normal domain of attraction of the stable laws (see, e.g., Theorem 2.6.7 of Ibragimov and Linnik, 1971)[15]. Observe that we allow for $\alpha \geq 2$ such that variance is finite. Gaussian distribution tails decay exponentially hence not according to (1): see Feller (1971). Many stochastic recurrence equations (e.g. GARCH(p, q) processes) inherently satisfy (1) specifically, or the tails decay according to a more general form of regular variation. See Mikosch and Stărică (2000) and Basrak *et al* (2001).

3. Measures of Extremal Dependence Consider a stationary stochastic process X that satisfies (1), $\alpha > 0$. For convenience, denote by $Y_h = \{Y_{h,t}\}$ the convolution process $\{X_t + X_{t-h}\}$. The convolution Y_h satisfies (1) with the same tail index α: see Embrechts and Goldie (1980: Theorem 3; 1982); see also Cline (1983, 1986).

3.1 Co-Difference

Let $c_i(x) > 0$ for both $i = 1, 2$. We define the *generalized two-tailed co-difference* at lag $h > 0$, denoted $Co(h)^{(0)} \equiv Co(X_t, X_{t-h})^{(0)}$, as the asymptotic difference in the weighted two-tailed extremal marginal probabilities

$$\text{(2)} \quad Co(h)^{(0)} = \lim_{\epsilon \to \infty} \epsilon^\alpha \left[P\left(|X_t + X_{t-h}| > \epsilon\right) - P\left(|X_t| > \epsilon\right) - P\left(|X_{t-h}| > \epsilon\right) \right].$$

[15] The domain of attraction of the stables contains infinitely many distributions for which a central limit theorem applies: see Feller (1971) and Ibragimov and Linnik (1971). This empirically imperative distribution class has been utilized in the development of asymptotic theory for least squares estimators (Cline, 1983), t-ratios, the Durbin-Watson test (Loretan and Pihllips, 1991), the Box-Pierce test (Runde, 1997), tests of covariance stationarity (Loretan, 1991; Loretan and Pihllips, 1994), cointegration (Caner, 1998), unit roots (Chan and Tan, 1989; Phillips, 1990), extremal structural change tests (Quintos *et al*, 2001), etc.

By stationarity $P(|X_t| > \epsilon) = P(|X_{t-h}| > \epsilon)$ as $\epsilon \to \infty$, hence[16]:

$$(3) \quad Co(h)^{(0)} = \lim_{\epsilon\to\infty} |\epsilon|^{\alpha} \left[P\left(|X_t + X_{t-h}| > \epsilon\right) - 2P\left(|X_t| > \epsilon\right)\right] = \left[c_1(y_h) + c_2(y_h)\right] - 2\left[c_1(x) + c_2(x)\right],$$

where $c_i(y_h)$ and $c_i(x)$ denote respectively the extreme scales of the processes Y_h and X.

If large values of X_{t-h} are predominantly followed by large values of X_t of the same sign, for example, then $P(|X_t + X_{t-h}| > \epsilon)$ will be larger than the combined probabilities that $|X_t|$ or $|X_{t-h}|$ surpass ϵ as $\epsilon \to \infty$, hence $c_i(y_h)$ $> 2c_i(x)$ for both tails $i = 1, 2$. In this case $Co^{(0)}(h) > 0$ and we say $\{X_t\}$ is *two-tailed positively extremal dependent* at lag h. If $Co(h)^{(0)} = 0$ for all displacements $h \geq 1$, we say $\{X_t\}$ is a *two-tailed extremal white noise* process.

For serially independent processes that satisfy (1) it can be shown the convolution process $Y_h = \{X_t + X_{t-h}\}$ satisfies (1) with scales (see Feller, 1971: VIII.8):

$$(4) \quad c_1(y_h) = 2 \times c_1(x); \quad c_2(y_h) = 2 \times c_2(x)$$

For independent processes the co-difference identically equals zero:

$$(5) \quad Co(h)^{(0)} = \left[c_1(y_h) + c_2(y_h)\right] - 2\left[c_1(x) + c_2(x)\right] = 2\left[c_1(x) + c_2(x)\right] - 2\left[c_1(x) + c_2(x)\right] = 0.$$

Like the covariance, however, the relationship is not necessarily two-way: serially dependent processes may have a zero co-difference.

For non-maximally skewed processes the co-difference (2) as stated is insensitive to whether extremal dependence is more pronounced in one tail or the other. Moreover, for maximally skewed processes the co-difference requires a tail-specific form (e.g. a squared GARCH process X_t^2). In general, for *left- and right-tailed serial extremal dependence*, define respectively left- and right-co-differences

$$(6) \quad Co(h)^{(1)} = c_1(y_h) - 2c_1(x), \quad Co(h)^{(2)} = c_2(y_h) - 2c_2(x).$$

Trivially

$$(7) \quad Co(h)^{(0)} = Co(h)^{(1)} + Co(h)^{(2)}.$$

Serially independent processes satisfy (4) hence $Co(h)^{(1)} = Co(h)^{(2)} = 0$. If X displays left-tail extremal white-noise but right-tail positive extremal dependence at displacement $h > 0$, then $c_1(y_h) = 2c_1(x)$ and $c_2(y_h) > 2c_2(x)$ such

[16] The co-difference was originally conceptualized for symmetric stable random processes based on the convolution difference $X_t - X_{t-h}$ and based on differences in characteristic functions: see Samorodnitsky and Taqqu (1994), and Kokoszka and Taqqu (1994, 1996). The two-tailed co-difference is simply an extremal version of the covariance for processes with regularly varying tails: for finite variance processes, $cov(X_t, X_{t-h}) = (V[X_t + X_{t-h}] - V[X_t] - V[X_{t-h}])/2$.

that the left-tail $Co(h)^{(1)} = 0$, right-tail $Co(h)^{(2)} > 0$ and two-tailed $Co(h)^{(0)} > 0$. If the process X is maximally positively skewed, for example, then $c_1(x) = 0$ and only the right-tail co-difference is used.

3.2 Generalized Co-Relation

We define the left-, right-, and two-tailed *generalized co-relation coefficients* as scaled versions of the co-differences:

$$\text{(8)} \quad \rho_\alpha^{(0)}(h) = \frac{c_1(y_h) + c_2(y_h)}{2\left[c_1(x) + c_2(x)\right]} - 1, \;\; \rho_\alpha^{(i)}(h) = \frac{c_i(y_h)}{2c_i(x)} - 1 \qquad i = 1, 2.$$

The two-tailed co-relation coefficient is symmetric in the dual sense that $\rho_\alpha^{(0)}(h)$ $= \rho_\alpha^{(0)}(-h)$ and $\rho_\alpha^{(0)}(X_t, X_{t-h}) = \rho_\alpha^{(0)}(-X_t, -X_{t-h})$. The bounds can be shown to be $-1 \leq \rho_\alpha^{(0)}(h) \leq \max\{2^{\alpha-1}\text{-}1, 0\}$[17].

From (8), we deduce the coefficients satisfy

$$\text{(9)} \qquad \rho_\alpha^{(0)}(h) = \psi_1(x)\rho_\alpha^{(1)}(h) + \psi_2(x)\rho_\alpha^{(2)}(h), \;\; \psi_i(x) \equiv \frac{c_i(x)}{c_1(x) + c_2(x)}$$
$$\rho_\alpha^{(i)}(-X_t, -X_{t-h}) = \rho_\alpha^{(j)}(X_t, X_{t-h}),\; i \neq j$$

The bounds on the one-tailed co-relation can be deduced to be $-1 \leq \rho_\alpha^{(i)}(h) \leq \max\{2^{\alpha-1}\text{-}1, 0\}$ for each $i = 1, 2$. The two-tailed co-relation is a simple convex combination of the tail-specific co-relations, with more weight given to the fatter tail of X. Clearly $\rho_\alpha^{(1)}(h) = \rho_\alpha^{(2)}(h) = 0$ implies $\rho_\alpha^{(0)}(h) = 0$, but not visa-versa.

3.3 Further Extremal Dependence Measures

The traditional co-difference for symmetric stable laws uses the convolution difference $\tilde{Y}_{h,t} \equiv X_t - X_{t-h}$: see, e.g., Samorodnitsky and Taqqu (1994). For general processes with regularly varying tails (1) we may define an additional two-tailed co-difference and co-relation,

$$\text{(10)} \qquad \tilde{C}o(h) = 2\left[c_1(x) + c_2(x)\right] - \left[c_1(\tilde{y}_h) + c_2(\tilde{y}_h)\right]$$
$$\tilde{\rho}_\alpha(h) = 1 - \frac{c_1(\tilde{y}_h) + c_2(\tilde{y}_h)}{2\left[c_1(x) + c_2(x)\right]}.$$

For independent processes $c_1(\tilde{y}_h) = c_2(\tilde{y}_h) = c_1(x) + c_2(x)$, hence $\tilde{C}o(h) = \tilde{\rho}_\alpha(h)$ $= 0$. The bounds are now $\min\{2^{\alpha-1}\text{-}1, 0\} \leq \tilde{\rho}_\alpha(h) \leq 1$. Intuitively if the process X is *positively serially extremal dependent* at displacement $h > 0$, for example, then large positive (negative) values of X_{t-h} are in general followed by large positive (negative) values of X_t, hence $\tilde{Y}_{h,t}$ will tend to be small, the extremal dispersion of $\tilde{Y}_{h,t}$ will be small, and $\tilde{\rho}_\alpha(h) > 0$.

Tail-specific co-relations are not available when the difference $\tilde{Y}_{h,t} \equiv X_t - X_{t-h}$ is used: for example both $c_1(\tilde{y}_h)$ and $c_2(\tilde{y}_h)$ will be small for either left-

[17] When $\alpha \geq 1$, the maximum of $\rho_\alpha^{(0)}(h)$ occurs when $X_t = X_{t-h}$ with probability one: in this case $c_i(y_h) = 2^\alpha c_i(x)$ and the bound $2^{\alpha-1}$-1 follows from (8). The minimum follows by noting every $c_i \geq 0$ and specifically $c_i(y_h) = 0$ when $X_t = -X_{t-h}$ with probability one. When $\alpha < 1$, the bounds are $-1 \leq \rho_\alpha^{(0)}(h) \leq 0$: the upper bound follows by noting $2^{\alpha-1} - 1 < 0$ when $\alpha < 1$ (the greatest degree of "positive" co-relation is, in fact, negative), which does not account for the independence case, $\rho_\alpha^{(0)}(h) = 0$.

or right-tail positive extremal dependent processes. Simulation experiments, however, consistently demonstrate a sample version of this two-tailed measure dominates a sample version of $\rho_\alpha^{(0)}(h)$. See Section 7.

3.4 Example: Linear Processes and Memory

Many Linear and GARCH processes co-relations that decay according to the memory of the process. Consider a linear process $\{X_t\}$ defined by the moving average form

$$X_t = \sum\nolimits_{i=0}^{\infty} \psi_i \epsilon_{t-i-1},\ \psi_0 = 1,$$

where $\sum_{i=0}^{\infty} |\psi_i|^\alpha < \infty$, and $\{\epsilon_t\}$ is an *iid* stochastic process with marginal distributions that satisfy Assumption A.1 or A.2 with index $\alpha > 0$ and symmetric tails, $c_1(\epsilon) = c_2(\epsilon) = 1$. Then X_t satisfies Assumption A.1 or A.2 with index α: see Cline (1983, 1986). In particular, because X_t is a convolution of a unit-scaled, symmetric, *iid* process, it can be shown that both tails satisfy (e.g. Feller, 1971; and Cline, 1983)

$$c_i(x) = \sum\nolimits_{i=0}^{\infty} |\psi_i|^\alpha < \infty$$
$$c_i(X_t \pm X_{t-h}) = \sum\nolimits_{i=0}^{h-1} |\psi_i|^\alpha + \sum\nolimits_{i=0}^{\infty} |\psi_{i+h} \pm \psi_i|^\alpha < \infty$$

for each $i = 1$ and 2. The co-relations are

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

If $X_t = \phi X_{t-1} + \epsilon_t$, $|\phi| < 1$, such that $\psi_i = \phi^i$, then $\rho_\alpha^{(k)}(h) = .5(|1 + \phi^h|^\alpha - 1 - |\phi|^{h\alpha})$, which reduces to $\rho_2^{(k)}(h) = \phi^h$ when $\alpha = 2$. Similarly $\tilde{\rho}_\alpha(h) = .5(1 + |\phi|^{h\alpha} - |1 - \phi^h|^\alpha)$ likewise satisfies $\tilde{\rho}_2(h) = \phi^h$. If $\{\psi_i\}$ decays geometrically (e.g. ARMA) or hyperbolically (e.g. ARFIMA) the numerator summations in $\rho_\alpha^{(k)}(h)$ and $\tilde{\rho}_\alpha(h)$ decay geometrically or hyperbolically, hence the co-relations $\rho_\alpha^{(i)}(h)$ and $\tilde{\rho}_\alpha(h)$ decay likewise. See Kokoszka and Taqqu (1994, 1996a). Because the innovations have symmetric tails the two-tailed and one-tailed co-relations $\rho_\alpha^{(i)}(h)$ are all identical. Asymmetry in the tail-specific co-relations of inherently linear processes is therefore driven entirely by asymmetry in the underlying shocks. Table 1 of Appendix 2 contains values for $\rho_\alpha^{(i)}(h)$ and $\tilde{\rho}_\alpha(h)$, $h = 1...10$, for various ARMA(1,1) processes.

3.5 Example: Power-GARCH Processes and Memory

Consider an infinite order moving average of powers of realizations of some power time series $\{X_t\}$ (i.e. power-ARCH(∞))

$$X_t = \sigma_t Z_t$$
$$\sigma_t^p = \theta_0 + \sum\nolimits_{i=1}^{\infty} \theta_i |X_{t-i}|^p,\ \theta_i \geq 0\ \forall i,\ 0 < p < \alpha,$$

where $\{Z_t\}$ is an *iid* stochastic process that satisfies Assumption A.1 or A.2 with index $\alpha > 0$, $0 < p < \alpha$, $E|Z_t|^p = 1$, $c_1(z) = c_2(z)$, $\sum_{i=1}^{\infty} \theta_i < 1$ and $\sum_{i=1}^{\infty} \theta_i^{\alpha} < \infty$. It is straightforward to show the power process $\{|X_t|^p\}$ obtains an L_1-norm convergent volterra-type expansion of sums of independent products $|Z_{t-j_1}|^p \cdots |Z_{t-j_q}|^p$ $q = 1, 2, ...$ with coefficients $\theta_{j_1} \cdots \theta_{j_q}$, provided $\lim_{q \to \infty} \sum_{j_1,...,j_q=1}^{\infty} \theta_{j_1} \cdots \theta_{j_q} = 0$ (see Priestly, 1988; Giraitis *et al,* 2000; Davidson, 2004). Products of independent random variables with regularly varying tails, cf. Assumption A, also have regularly varying tails with the same index, and summations of independent processes the satisfy Assumption A also satisfy Assumption A with the same index. See Embrechts and Goldie (1980) and Cline (1983, 1986). Hence σ_t^p and $|X_t|^p$ satisfy Assumption A with tail index p/α. See Hill (2005: Lemma 9). The co-relations are still presented by $\rho_{\alpha}^{(k)}(h)$ and $\tilde{\rho}_{\alpha}(h)$, above, where

$$\psi_i = \theta_0 \sum\nolimits_{j_1,...,j_i=1}^{\infty} \theta_{j_1} \cdots \theta_{j_i}, \; i = 0, 1, 2, ...$$

If θ_i decays geometrically(e.g. IGARCH) or hyperbolically (FIGARCH) then $\psi_i = \theta_0 \sum_{j_1,...,j_i=1}^{\infty} \theta_{j_1} \cdots \theta_{j_i}$ will decay geometrically or hyperbolically, and the co-relations will have the same decay property. See Mikosch and Stărică (2000) and Basrak *et al* (2001) for general conditions in which $X_t = \sigma_t Z_t$ has regularly varying tails for any essentially any innovations process $\{Z_t\}$.

4. Co-Relation Sample Statistics and Asymptotics The distribution scales c_i and tail index α are the only parameters that require estimation. Simple estimators well known in the literature are due to Hill (1975) and Hall (1982). Denote by $X_{(i)}^{+} > 0$ the i^{th} right-tail order statistic of X_t, $X_{(1)}^{+} \geq X_{(2)}^{+} \geq ...$; by $X_{(i)}^{-} < 0$ the left tail order statistics; and by $X_{(i)}^{\pm}$ either order statistic. Similarly, we write X_t^{+} to denote $(X_t)_{+}$, the t^{th} observation or zero, whichever is larger. We have

$$(11) \quad \hat{c}_{1,m} \equiv \left(\frac{m}{n}\right) \left|X_{(m+1)}^{-}\right|^{\hat{\alpha}_m}, \quad \hat{c}_{2,m} \equiv \left(\frac{m}{n}\right) \left(X_{(m+1)}^{+}\right)^{\hat{\alpha}_m}$$

$$\hat{\alpha}_m \equiv \left(\frac{1}{m} \sum\nolimits_{j=1}^{m} \ln X_{(j)}^{+} - \ln X_{(m+1)}^{+}\right)^{-1},$$

for some integer $1 \leq m \leq n$ such that $m \to \infty$ as $n \to \infty$ and $m/n \to 0$. The inverted estimator $\hat{\alpha}_m^{-1}$ is widely known as the Hill estimator, cf. B. Hill (1975), the subject of theoretical and empirical studies too numerous to list[18].

[18] See, e.g., Resnick and Stărică (1998) and Hill (2005) and the numerous citations therein. See Hall (1982) for an intuitive account of $\hat{c}_{i,m}$, and see Hsing (1991) for an interpretation of the Hill estimator as a method of moments estimator. A multitude of alternative estimators of α exist. See, e.g., Pickands (1975), Dekkers et al (1989), de Haan and Peng (1998). We opt for the Hill estimator because it is simple to compute, and because a broad asymptotic theory exists for a large class of dependent and heterogenous processes, cf. Hill (2005). See de Haan and Peng (1998) for derivations of the mean-squared-error of four tail estimators for *iid* data,

Because we assume the single tail index α describes both extreme right and left distribution tail shapes, and the process X_t is not maximally negatively skewed, we lose nothing asymptotically by using only the right-tail information in the computation of $\hat{\alpha}_m$.

4.1 Sample Co-Relation

We deduce estimators of the co-relation coefficients

$$(12)\quad \hat{\rho}^{(0)}_{\alpha,m}(h) = \frac{\hat{c}_{1,m}(y_h)+\hat{c}_{2,m}(y_h)}{2\left[\hat{c}_{1,m}(x)+\hat{c}_{2,m}(x)\right]}-1,\quad \hat{\rho}^{(i)}_{\alpha,m}(h)=\frac{\hat{c}_{i,m}(y_h))}{2\hat{c}_{i,m}(x)}-1$$

$$\widehat{\widetilde{\rho}}_{\alpha,m}(h) = 1-\frac{\hat{c}_{1,m}(\tilde{y}_h)+\hat{c}_{2,m}(\tilde{y}_h)}{2\left[\hat{c}_{1,m}(x)+\hat{c}_{2,m}(x)\right]}.$$

For the two-tailed estimators is it understood that $c_i(x) > 0$ for each $i = 1, 2$; for left or right tailed co-relations, it is understood that $c_1(x) > 0$ or $c_2(x) > 0$. By assumption the extreme tails of X_t are for each t identically represented by (1) with common exponent α. Similarly, the convolutions $Y_{h,t} = X_t + X_{t-h}$ and $\tilde{Y}_{h,t} = X_t - X_{t-h}$ satisfy (1) with exponent α. It suffices, therefore, to use a tail index estimator $\hat{\alpha}_m$ based solely on X_t in all components of the co-relation estimators.

4.2 Assumptions

We employ five sets of assumptions regarding tail convergence, dependence and degree of heterogeneity: consult Appendix 1. Assumptions A.1 and A.2 restrict tail convergence; Assumption B restricts the rate $m \to \infty$; and Assumptions C and D define regulatory conditions. Assumption E.1 defines functionals of X_t, denoted $U_{n,t}$ and $U^*_{n,t}$, as L_1-mixingales; and Assumption E.2 restricts the functionals to be L_2-NED of size $-1/2$ on a mixing process. It is straightforward to prove $U_{n,t}$ and $U^*_{n,t}$ are bounded in the L_r-norm for any $r \geq 1$, hence the L_2 assumption is non-binding: see Lemma 1 of Hill (2005). See Section 5 for a discussion on restricting the dependence properties of X_t itself[19].

4.3 Preliminary Theory

The first result establishes a weak limit for $\hat{\alpha}_m$ for mixingale $U_{n,t}$ and $U^*_{n,t}$ with regularly varying tails. For a proof, consult Hill (2005: Theorem 2)[20]. The probability limit is of paramount importance for our proposed test: test power relies entirely on consistency for dependent processes.

Theorem 1 *Let X satisfy Assumptions A.1, B.1 and E.1 hold. Then $\hat{\alpha}_m \to \alpha$.*

Remark 1: Because strong and uniform mixing processes and processes NED on a mixing processes are special cases of mixingales, consistency holds in

including the Hill estimator. Their conclusion is that no one estimator dominates the others. None of the estimators they consider, however, with the exception of the Hill estimator, has been studied in the context of dependent and heterogenous data.

[19] Consult Hall and Hyde (1985), Gallant and White (1988), Pötscher and Prucha (1991), and Davidson (1994) for details on the dependence properties used in this paper.

[20] For brevity, proofs of Theorems 1-3 have been allotted to a separate paper, "On Tail Index Estimation Using Dependent, Heterogenous Data" (Hill, 2005). In particular, Theorems 1-3 of the present paper are Theorems 2, 11, and 4, respectively, of that work. The paper is available at http://econwpa.wustl.edu:80/eps/em/papers/0505/0505005.pdf.

these cases as well for the functionals $U_{n,t}$ and $U^*_{n,t}$. Moreover, under suitable conditions, Hill (2005: Theorem 3) proves the estimator $\hat{\alpha}_m$ is consistent for the classes of L_0- and L_1-approximable processes $\{U_{n,t}\}$ and $\{U^*_{n,t}\}$ (cf. Pötscher and Prucha, 1991; see also Davidson, 1994). See also Resnick and Stărică (1995, 1997, 1998)[21].

Theorem 2 *Let X satisfy Assumptions A.2, B.2, C, D, and E.2, for any $h > 0$. Then*

$$(13)\quad \hat{c}_{i,m} \rightarrow c_i, \;\; \hat{\rho}^{(\cdot)}_{\alpha,m}(h) \rightarrow \rho^{(\cdot)}_{\alpha}(h), \;\; \text{and } \widehat{\tilde{\rho}}_{\alpha,m}(h) \rightarrow \tilde{\rho}_{\alpha}(h).$$

Remark 1: Theorem 11 of Hill (2005) proves $\hat{c}_{i,m} \rightarrow c_i$: the remaining limits follow from the functional invariance property of probability limits and (12).

The limit distributions of the proposed co-relation based test statistics hinge entirely on the asymptotic properties of $\hat{\alpha}_m$. We next provide a Gaussian limit for $\hat{\alpha}_m$ for processes X_t that belong to the domain of attraction of the stable laws when $\alpha < 2$, where the functionals $\{U_{n,t}, U^*_{n,t}\}$ are L_2-NED on a uniform mixing process. For a proof, see Hill (2005: Theorem 4).

Define the asymptotic variances

$$(14)\quad \begin{aligned} \sigma^2_m &= m^{-1} E\left(\hat{\alpha}^{-1}_m - \alpha^{-1}\right)^2 \\ \sigma^2_m(\omega) &= m^{-1} \sum\nolimits_{s,t=1}^{n} E\left(\omega_1 U_{n,t} + \omega_2 \alpha^{-1} U^*_{n,t}\right)^2, \end{aligned}$$

where $\omega \in \mathbb{R}^2$ is arbitrary. Define $\tilde{\sigma}^2_m = \alpha^2 \sigma^2_m$

Assumption F 1. For some $\gamma \geq 0$ let $\inf_{\omega \in \mathbb{R}^2} \sigma^2_m(\omega) = O(n^{\gamma})$; 2. Let $\inf_{\omega \in \mathbb{R}^2} \sigma^2_m(\omega) = O(1)$ (i.e. $\gamma = 0$)[22].

Theorem 3 *Let X satisfy Assumptions A.2, B.2, E.2, and F.1. Then*

$$(15)\quad \sqrt{m}(\hat{\alpha}_m - \alpha)/\tilde{\sigma}_m \Rightarrow N(0, \alpha^2),$$

if $r > 2$, $\delta \geq 1 - r\gamma/(r-2)$ and $\gamma > 0$; or $r = 2$ and $\gamma \geq 0$. Moreover $\sigma^2_m = O(n^{\gamma})$. If additionally Assumption E.2 holds with $r = 2$ (i.e. the NED-mixing base size is -1) and Assumption F.2 holds (i.e. $\gamma = 0$), then $\sqrt{m}(\hat{\alpha}_m - \alpha) \Rightarrow N(0, \alpha^2 \tilde{\sigma}^2)$, where $\tilde{\sigma}^2 = \alpha^2 \sigma^2 = \alpha^2 \lim_{n\rightarrow\infty} \sigma^2_m$. If X_t is iid, then $\sigma^2 = \alpha^{-2}$.

Remark 1: Assumption F ensures a non-degenerate $\sigma^2_m(\omega)$ for any $\omega \in \mathbb{R}^2$, and Hill (2005: Theorem 4) proves $\sigma^2_m / \sigma^2_m(1,-1) \rightarrow 1$, hence $\sigma^2_m = O(n^{\gamma})$ must hold.

[21] L_0-approximability includes processes $\{U_{n,t}, U^*_{n,t}\}$ that are infinite order distributed lags of *iid* innovations that satisfy (1); simple bilinear processes; and stochastic difference equations, including GARCH-type processes. See Resnick and Stărică (1995, 1997, 1998). Moreover, Davidson (2004) proves many IGARCH and FIGARCH processes are L_0-approximable, hence consistency holds for these processes $\{U_{n,t}, U^*_{n,t}\}$ as well.

[22] It is understood that $\inf_{\omega \in \mathbb{R}^2} \sigma^2_m(\omega) = O(1)$ implies $\lim_{n\rightarrow\infty} \sigma^2_m(\omega)$ is finite and positive uniformly in $\mathbb{R}^2$.

Remark 2: Hill (2005: Lemma 5) proves a non-parametric Newey–West-type kernel estimator $\hat{\sigma}_m^2$ satisfies $\hat{\sigma}_m^2/\sigma_m^2 \to 1$ under the conditions of Theorem 3. See Section 8, below. The estimator $\hat{\sigma}_m^2$ does not require knowledge of an underlying parametric structure (e.g. GARCH) and therefore different variance estimators for different parametric classes are not required. The only primitive assumptions required are tail decay (Assumption A.2) and memory (Assumption E.2, or E.3 below). See Section 5 for a discussion of processes which satisfy the tail decay and memory assumptions (ARFIMA, IGARCH, etc.).

4.4 Co-Relation Estimator

The basis of a portmanteau test statistic is grounded on the following limit for the co-relation estimator for extremal NED processes. Let $\hat{\rho}_{\alpha,m}^{(i)} = [\hat{\rho}_{\alpha,m}^{(i)}(1)$,..., $\hat{\rho}_{\alpha,m}^{(i)}(h)]'$.

Theorem 4 *Let X satisfy Assumptions A.2, B.2, C, D, E.2 and F.1. Define $\tilde{\sigma}_m^2 = \alpha^2\sigma_m^2$. Under the hypotheses $H_0 : \rho_\alpha^{(i)}(h) = 0$, $i = 0, 1, 2, ...$,*

$$(16) \quad \sqrt{m}\hat{\rho}_{\alpha,m}^{(i)}/\tilde{\sigma}_m \Rightarrow N(0, \Omega_h^{(i)}), \quad \Omega_h^{(i)} = [\varpi_j^{(i)}\varpi_k^{(i)}]_{j,k=1}^h,$$

if $r > 2$, $\delta \geq 1 - r\gamma/(r-2)$ and $\gamma > 0$, or $r = 2$ and $\gamma \geq 0$, where

$$(17) \quad \varpi_h^{(0)} \equiv \ln\left[\frac{c_1(y_h)^{\psi_1(y_h)}c_2(y_h)^{\psi_2(y_h)}}{c_1(x)^{\psi_1(x)}c_2(x)^{\psi_2(x)}}\right], \quad \psi_j \equiv \frac{c_j}{c_1+c_2}$$
$$\varpi_h^{(i)} \equiv \ln[c_i(y_h)/c_i(x)] = \ln[2],\ i = 1, 2.$$

Moreover, under $H_1 : \rho_\alpha^{(i)}(h) \neq 0$, $m\hat{\rho}_{\alpha,m}^{(i)}(h)^2 \to \infty$ with probability one.

Remark 1: The line of proof does not make use of the property of the convolution summation, per se: use of the summation $X_t + X_{t-h}$ or difference $X_t - X_{t-h}$ is irrelevant with respect to the asymptotic behavior of $\hat{\rho}_{\alpha,m}^{(i)}(h)$ and $\hat{\tilde{\rho}}_{\alpha,m}(h)$. A limit identical to (16) holds for $\hat{\tilde{\rho}}_{\alpha,m}(h)$:

$$(16') \quad \sqrt{m}\ \hat{\tilde{\rho}}_{\alpha,m}/\tilde{\sigma}_m \Rightarrow N(0, \tilde{\Omega}_h), \quad \tilde{\Omega}_h = [\tilde{\varpi}_j, \tilde{\varpi}_k]_{j,k=1}^h,$$

where $\tilde{\varpi}_h$ is simply $\varpi_h^{(0)}$ defined in (17) with $c_j(y_h)$ and $\psi_j(y_h)$ replaced by $c_j(\tilde{y}_h)$ and $\psi_j(\tilde{y}_h)$.

Remark 2: Because we consider the non-maximal skewness case such that each $c_j(x) > 0$, the parameters $\varpi_{i,h}$ are well-defined, finite, and may be negative. We discuss the possibility that $\varpi_h^{(i)} = 0$ in Section 6.2.

Remark 3: The general asymptotic variance term $\tilde{\sigma}_m^2$ is discussed in Theorem 3. From that result we do not require knowledge of an underlying parametric structure (e.g. GARCH) as long as the minimal tail decay and memory properties hold.

Remark 4: Each $\hat{\rho}_{\alpha,m}^{(i)}(j)$ is fundamentally grounded on the same random variable $\hat{\alpha}_m$, hence the limiting covariance matrix $\Omega^{(i)}$ is singular. In particular, asymptotically any pair $\hat{\rho}_{\alpha,m}^{(i)}(j)$ and $\hat{\rho}_{\alpha,m}^{(i)}(k)$ is perfectly correlated. This will make the derivation of a test statistic quite simple: see Section 5.

4.5 Tail Difference Estimator

In order to test tail difference in extremal dependence for non-maximally skewed processes we have the following result.

Theorem 5 *Let X satisfy Assumptions A.2, B.2, C, D, E.2 and F.1. Then*

$$(18)\quad \sqrt{m}\left(\left[\hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h)\right] - \left[\rho^{(1)}_{\alpha}(h) - \rho^{(2)}_{\alpha}(h)\right]\right)/\tilde{\sigma}_m \Rightarrow N\left(0, v(h)\right),$$

if $r > 2$, $\delta \geq 1 - r\gamma/(r-2)$ and $\gamma > 0$, or $r = 2$ and $\gamma \geq 0$, where $v(h) = [\varpi^{(1)}_h(1 + \rho^{(1)}_{\alpha}(h)) - \varpi^{(2)}_h(1 + \rho^{(2)}_{\alpha}(h))]^2$. If $\rho^{(1)}_{\alpha}(h) = \rho^{(2)}_{\alpha}(h)$ then

$$(19)\quad \sqrt{m}\left(\hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h)\right)/\tilde{\sigma}_m \Rightarrow 0,$$

and if $\rho^{(1)}_{\alpha}(h) \neq \rho^{(2)}_{\alpha}(h)$, then $m(\hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h))^2 \rightarrow \infty$ with probability one.

Remark 1: A test of identical left- and right-tailed extremal dependence involves a degenerate distribution due to the asymptotic variance term $v(h)$: if $\rho^{(1)}_{\alpha}(h) = \rho^{(2)}_{\alpha}(h)$, then from (8) and (18) by construction $\varpi^{(1)}_h = \varpi^{(2)}_h$, hence $v(h) = 0$ and $\sqrt{m}(\hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h)) \Rightarrow 0$.

4.6 A Simple Bivariate Case

Suppose $X_t = (X_{1,t}, X_{2,t}) \in \mathbb{R}^2$ is a bivariate process on the measure space $(\Omega, \Im_t, P)$, $\Im_t = \sigma(\{X_{1,s}, X_{2,s}\} : s \leq t)$. Assume the marginal distributions satisfy (1) with indices $\alpha(x_1) > 0$ and $\alpha(x_2) > 0$. Note that the convolutions $X_{1,t-h} \pm X_{2,t}$ satisfy (1) with index $\min\{\alpha(x_1), \alpha(x_2)\}$. An analysis of the asymptotic properties of $\hat{\rho}_{\alpha,m}(X_{1,t-h}, X_{2,t})$ requires a joint limit law for $\hat{\alpha}_m(x_1)$ and $\hat{\alpha}_m(x_2)$ under general dependence assumptions. This is well beyond the scope of the present paper. However it is not difficult to derive a Gaussian limit under the hypothesis that $X_{1,t}$ and X_{2t} are mutually independent (with serial near-epoch-dependence in the individual processes). A proof is omitted for brevity, but mimics the line of proof of Theorem 4[23]. Let $\sigma^2_{i,m} = m^{-1}E(\hat{\alpha}^{-1}_{i,m} - \alpha^{-1}_i)^2$, $\sigma^2_i = \lim_{n\to\infty} \sigma^2_{i,m} \in (0,\infty)$, $\hat{\alpha}_{i,m} = \hat{\alpha}_m(x_i)$, $i = 1, 2$.

Theorem 6 *Let $\{X_{t,1}, X_{t,2}\}$ satisfy Assumptions A.2, B.2, C, D and E.2. Assume $X_{t,1}$ and $X_{t,2}$ are mutually independent, and assume Assumption F.2 holds for each $\sigma^2_{1,m}(\omega)$ and $\sigma^2_{2,m}(\omega)$. Then each $\sqrt{m}/\ln(n/m)]\hat{\rho}^{(\cdot)}_{\alpha,m}(h) \Rightarrow N(0, \varpi^2)$ if $r > 2$, $\delta \geq 1 - r\gamma/(r-2)$ and $\gamma > 0$, or $r = 2$ and $\gamma \geq 0$, where $\varpi^2 = \psi^2_1\alpha^4_1\sigma^2_1 + \psi^2_2\alpha^4_2\sigma^2_2$, and $\psi^2_i = [c_1(x_i) + c_2(x_i)]/[c_1(x_1) + c_2(x_1) + c_1(x_2) + c_2(x_2)]$.*

Remark 1: The slower rate of convergence $\sqrt{m}/\ln(n/m)$ is due to the added error introduced by having multiple estimators $\hat{\alpha}_m(x_1)$ and $\hat{\alpha}_m(x_2)$.

Remark 2: Notice the limiting variance ϖ^2 does not depend on the displacement h: the variance of $\hat{\rho}^{(\cdot)}_{\alpha,m}(h)$ depends entirely on the variances of $\hat{\alpha}_m(x_1)$ and $\hat{\alpha}_m(x_2)$, which are not functions of h.

[23] A proof is available upon request.

4.7 Estimated Residuals

In practice we will want to analyze the statistical fit of model in part based on how well the model characterizes the extremes of the time series. Similarly, we may want to analyze the (extremal) dependence properties of estimated GARCH innovations as in Hong (2001) and Linton and Wang (2004). As long as the estimated residuals, say $\hat{\epsilon}_t$, satisfy $\hat{\epsilon}_t = \epsilon_t + o_p(1)$ for some underlying process $\{\epsilon_t\}$ then $\hat{\epsilon}_t \to \epsilon_t$ in distribution. Hence $\hat{\epsilon}_t$ will have all the tail and memory properties of ϵ_t as $n \to \infty$. If $\{\epsilon_t\}$ satisfies Assumptions A.2, B.2, C, D and E.2 then asymptotically $\{\hat{\epsilon}_t\}$ will as well. 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., Cline (1983), Knight (1993), Mikosch *et al* (1995), Kokoska and Taqqu (1996b), Davis and Wu (1997), Hall and Yao (2003), and Ling (2005).

5. Extremal Near-Epoch-Dependence The L_2-NED assumption is enforced on the functionals $U_{n,t}$ and $U^*_{n,t}$: see Assumption E.2. In Hill (2005) we bridge a dependence link between the process X_t and functionals $U_{n,t}$ and $U^*_{n,t}$ by defining the sufficient property of "*extremal near-epoch-dependence*" (E-NED). Indeed, because we are only concerned with dependence between extremes we are interested in a dependence property that does not involve the non-extremal support of the distribution.

5.1 E-NED Definition

Let $\digamma_t$ denote an increasing σ-algebra, and let $b_{2,n}(m)$ be a sequence of increasing positive numbers satisfying $\lim_{n\to\infty}(n/m)P(X_t > b_{2,n}(m)) = 1$, where $b_{2,n}(m) \to \infty$ as $n \to \infty$.

Assumption E.3 (L_2-E-NED) Let $\{X_t\}$ satisfy (1) with tail index $\alpha > 0$. Let $x_n : \mathbb{R} \to \mathbb{R}_+$ be a sequence of functions satisfying $x_n(u) \to \infty$ for arbitrary $u \in \mathbb{R}$. For each t there exists a Lebesgue measurable function $e^*_t : \mathbb{R} \to \mathbb{R}_+$, integrable on $\mathbb{R}_+$, and a sequence of constants $\{\upsilon^*_q\}_{q=0}^{\infty}$, with $\upsilon^*_q \to 0$ as $q \to \infty$, such that

$$\limsup_{n\to\infty} \sqrt{\frac{n}{m}} \left\| P\left(X_t > x_n(u) | \Im_{t-q}^{t+q}\right) - P\left(X_t > x_n(u) | \digamma_{t-q}^{t+q}\right) \right\|_2 \leq e^*_t(u)\upsilon^*_q.$$

Remark 1: As with all of the proceeding theory the E-NED definition holds for processes with regularly varying tails with any degree of tail thickness (i.e. any $\alpha < 2$ or $\alpha \geq 2$ is allowed).

Remark 2: The above condition is defined only in the extreme right-tail. Assumption E.3 does not make any reference to the serial dependence properties of X_t over the remaining non-extremal support, and is therefore quite general.

Remark 3: If Assumption E.3 holds with $x_n(u) = b_{2,n}(m)e^u$, for $\digamma_t = \sigma(\varepsilon_s : s \leq t)$, where $\{\varepsilon_t\}$ is a uniform mixing process of size -1 and the constants ν^*_q have size $-1/2$, then each processes $\{U_{n,t}\}$ and $\{U^*_{n,t}\}$ is L_2-NED on $\{\varepsilon_t\}$ with coefficient size $-1/2$ such that Assumption E.2 holds as $n \to \infty$. See Hill

(2005: Theorem 7). In Theorems 2-6, and all subsequent results, we may simply replace Assumption E.2 with Assumption E.3. See Hill (2005: Theorem 7) for a related "*extremal mixingale*" property, and an extension of Theorem 1 to this case.

Remark 4: Linton and Wang (2004) define the conditional θ-quantile $E[\psi_\theta(x_t - \mu_\theta)|\Im_{t-1}]$, where μ_θ denotes the θ-quantile and $\psi_\theta(z) = \theta - I(z < 0)$. If we define $\tilde{\psi}_\theta(z) = \theta - I(z > 0)$, then the E-NED assumption implies the $1 - m/n^{th}$ conditional quantile $E[\tilde{\psi}_{1-m/n}(x_t - \mu_{1-m/n})|F_{t-q}^{t+q}]$ is adequately approximated by $E[\tilde{\psi}_{1-m/n}(x_t - \mu_{1-m/n})|\Im_{t-q}^{t+q}]$ as $q \to \infty$, as $n \to \infty$. Notice that $\mu_{1-m/n}$ is simply $b_{2,n}(m)$ as $n \to \infty$, and $X_{(m+1)}$ consistently estimates $\mu_{1-m/n}$ for any $m < n$.

5.2 "Long" and "Short" Memory Linear and GARCH Processes

In Sections 3.4 and 3.5 we discussed the similitude between population memory properties and the co-relation decay rate. We now study the relationship between population memory and extremal dependence, a la the E-NED property.

Consider the form

$$(20)\quad X_t = \sum\nolimits_{i=0}^{\infty} \psi_i \epsilon_{t-i}, \quad \sum\nolimits_{i=0}^{\infty} |\psi_i|^\alpha < \infty,$$

where $\{\epsilon_t\}$ are *iid* and satisfy (1) with index $\alpha > 0$. Cline (1983) proves $\{X_t\}$ satisfies (1). In particular if $\{\epsilon_t\}$ satisfies Assumption A.2 then so does $\{X_t\}$. The process $\{X_t\}$ is L_2-E-NED on $\{\epsilon_t\}$, cf. Lemma 8 of See Hill (2005). If the coefficients $\{\psi_i\}$ decay geometrically, $\psi_i = O(\rho^{-i})$, $\rho > 1$, then $\tilde{v}_q^* = O(\rho^{-(q+1)\alpha})$. If $\{\psi_i\}$ decay hyperbolically, $\psi_i = O(i^{-\mu})$, $\mu > 1/\alpha$, then the E-NED size is $-(\alpha\mu - 1)/2$. If $\mu \geq 2/\alpha$ then the size satisfies $(\alpha\mu - 1)/2 \geq 1/2$. The geometric case covers causal-invertible ARMA processes; the hyperbolic case covers ARFIMA processes.

Moreover, Hill (2005: Lemmas 9 and 10) analyzes the E-NED property of power-GARCH process

$$(21)\quad X_t = \sigma_t Z_t, \;\; \sigma_t^p = \theta_0 + \sum\nolimits_{i=0}^{\infty} \theta_i |X_{t-i}|^p, \; \theta_i \geq 0 \; \forall i, \; 0 < p < \alpha$$

where $\sum_{i=0}^{\infty} |\theta_i|^\alpha < \infty$, $\sum_{i=0}^{\infty} \theta_i < 1$, $\{Z_t\}$ are *iid* and satisfy Assumption A.2 with index $\alpha > 0$, and $E|Z_t^p| = 1$. The processes $\{|X_t|^p\}$ and $\{\sigma_t^p\}$ satisfy Assumption A.2 with index p/α. Furthermore $\{X_t\}$ is L_2-E-NED on $\{Z_t\}$ with constants $\tilde{e}_t^*(u) : \mathbb{R} \to \mathbb{R}_+$ integrable on $\mathbb{R}_+$. If $\theta_i = O(i^{-\mu})$, $\mu > 1/\alpha$, then E-NED size is $(1 - \alpha\mu)/2$. If $\theta_i = O(\rho^{-i})$, $\rho > 1$, then $\tilde{v}_q^* = O(\rho^{-(q+1)\alpha})$.

Thus, the class of processes covering (20) and (21) satisfy the tail and memory requirements of Assumptions A.2 and E.3. We conjecture without proof that any stochastic process that satisfies the NED property will satisfy the E-NED property.

6. Tests of Extremal White Noise

In this section we develop the test of two-tailed extremal white noise using the two-tailed sample estimator $\widehat{\tilde{\rho}}_{\alpha,m}(h)$

for the cases of $\tilde{\varpi}_h \neq 0$ and $\tilde{\varpi}_h = 0$ for all h. The associated asymptotic theory for one- or two-tailed tests involving $\hat{\rho}^{(i)}_{\alpha,m}(h)$ follows identically.

6.1 $\quad \tilde{\varpi}_h \neq 0$

The "extremal white noise" hypothesis states X has zero serial co-relation for each displacement. In the two-tailed case $H_0 : \tilde{\rho}_\alpha(i) = 0,\ i = 1, 2, ...$ For compactness, define the standardized co-relation coefficient,

$$(22)\quad \tilde{\tau}_{\alpha,m}(i) = \sqrt{m}\,\widehat{\tilde{\rho}}_{\alpha,m}(i)/\sqrt{\widehat{\tilde{\varpi}}^2_{h,m}\widehat{\tilde{\sigma}}^2_m},$$

where $\widehat{\tilde{\sigma}}^2_m$ is any consistent estimator of $\tilde{\sigma}^2_m = \alpha^2\sigma^2_m$, where $\sigma^2_m = E[\sqrt{m}(\hat{\alpha}^{-1}_m - \alpha^{-1})]^2$. See Lemma 5 of Hill (2005) for a proof that a conventional Newey-West type estimator of σ^2_m with bandwidth $l_n = O(n^\varsigma)$, $\varsigma \in [0, 1]$, is consistent under the assumptions of Theorem 3.

We estimate $\tilde{\varpi}_h$ by

$$(23)\quad \widehat{\tilde{\varpi}}_{h,m} = \ln\left[\frac{\hat{c}_{1,m}(\tilde{y}_h)^{\hat{\psi}_1(\tilde{y}_h)}\hat{c}_{2,m}(\tilde{y}_h)^{\hat{\psi}_2(\tilde{y}_h)}}{\hat{c}_{1,m}(x)^{\hat{\psi}_1(x)}\hat{c}_{2,m}(x)^{\hat{\psi}_2(x)}}\right], \quad \hat{\psi}_i = \frac{\hat{c}_{i,m}}{\hat{c}_{1,m} + \hat{c}_{2,m}}.$$

From Theorem 2 and functional invariance of probability limits, for extremal NED processes $\widehat{\tilde{\varpi}}^2_{h,m} \to \tilde{\varpi}^2_h$ where $0 < \tilde{\varpi}^2_h < \infty$ by construction. By Theorem 4, Cramér's theorem and the continuous mapping theorem, for any $i > 0$ under the null of extremal white-noise we infer for extremal NED processes $(\tilde{\tau}_{\alpha,m}(i))^2 \Rightarrow \chi^2(1)$.

The limiting null distribution of each $\widehat{\tilde{\rho}}_{\alpha,m}(i)$ is grounded on the same random variable, $\hat{\alpha}_m$, thus we consider the average portmanteau statistic $h^{-1}\sum_{i=1}^h (\tilde{\tau}_{\alpha,m}(i))^2$.

Theorem 7 *Assume the conditions of Theorem 4 hold. If $\tilde{\rho}_\alpha(i) = 0$, $i = 1...h$, for any $h \geq 1$, then*

$$(24)\quad \tilde{Q}_{\alpha,m}(h) = \frac{1}{h}\sum_{i=1}^h (\tilde{\tau}_{\alpha,m}(i))^2 \Rightarrow \chi^2(1).$$

Moreover, provided $\tilde{\rho}_\alpha(i) \neq 0$ for at least one $i = 1...h$, then $\tilde{Q}_{\alpha,m}(h) \to \infty$ with probability one.

6.2 $\quad \tilde{\varpi}_h = 0$

If the asymptotic variance component $\tilde{\varpi}_h$ is identically zero for all h, then the sample co-relation limiting distribution is degenerate: under the null of $\tilde{\rho}_\alpha(h) = 0$, $\sqrt{m}\,\widehat{\tilde{\rho}}_{\alpha,m}(h) \Rightarrow 0$ in distribution, cf. Theorem 4. In this case the statistic $\tilde{\tau}_{\alpha,m}(h)$ is not well defined asymptotically. If we know $\tilde{\varpi}_{0,h} = 0$, however, then we simply use $\tilde{\tau}_{\alpha,m}(h) = \sqrt{m}\,\widehat{\tilde{\rho}}_{\alpha,m}(h)$: under the null hypothesis of extremal white noise, $\tilde{Q}_{\alpha,m}(h) \Rightarrow 0$, and under the alternative that at least one $\tilde{\rho}_\alpha(i) \neq 0$, $1 \leq i \leq h$, $\tilde{Q}_{\alpha,m}(h) \to \infty$ with probability one. A test of whether $\tilde{\varpi}_h = 0$ can be constructed from the estimator $\widehat{\tilde{\varpi}}_{h,m}$ using the asymptotic properties of $\hat{\alpha}_m$. If $\tilde{\varpi}_i = 0$ for some, but not all, $i = 1...h$, then the analyst can perform a standard Q-test for those displacements such that $\tilde{\varpi}_i \neq 0$ and the above alternative Q-test over those displacements satisfying $\tilde{\varpi}_i = 0$. For brevity we consider only the case $\tilde{\varpi}_i \neq 0$ for all $i = 1...h$.

7. Order Statistic Index m Selection We focus on the two-tailed statistic $\tilde{Q}_{\alpha,m}(h)$ using $\widehat{\tilde{\rho}}_{\alpha,m}(i)$: the subsequent method and theory carry over to statistics based on the one- or -two-tailed $\hat{\rho}^{(i)}_{\alpha,m}(h)$.

In order to construct the test statistic of Section 6 a rather arbitrary choice of the order statistic index m is required. Most practitioners pursue Dumouchel's (1983) suggestion of using observations X_t in the lower and/or upper 10^{th}-percentile, $m = [.1 \times n]$. Other methods are considered in Hall (1982), Hall and Welsh (1985), Resnick (1996), Draisma *et al* (1997), Resnick and Stărică (1997), and Danielson *et al* (1998), including bootstrap, minimum mean-squared-error and so-called Hill-plot methods. Each method performs reasonably well in certain environments, however each renders an essentially miserable performance for the Q-test.

Asymptotic theory, however, only requires $m \to \infty$ and $m/n^{2\theta/(2\theta+\alpha)} \to 0$, as $n \to \infty$, for unknown $\theta > 0$ and $\alpha > 0$. In order to remedy the small sample performance and arbitrariness of the index m we consider constructing statistic functionals over all feasible values of m by ranking the sample co-relation coefficients[24].

Let m_i denote any index, $1 \leq m_i \leq n$, and define the set

$$(25) \quad S_n = \left\{ 1 \leq m_i \leq n : \forall i, j,\ m_i \ \sim\ [n^{\delta}],\ \delta < \frac{2\theta}{2\theta+\alpha},\ \frac{m_i}{m_j} \to 1 \right\}.$$

We therefore do not consider the possibility of two different sequences (m_i, m_j) $= O(n^{\delta})$ with $m_i/m_j \to a \neq 1$. For example, $m_i = [n^{\delta}] \pm a_{n,i}$ is appropriate for any integer sequence $a_{n,i} = o(n^{\delta})$ provided $1 \leq m_i \leq n$.

For each lag i, define the following order statistic indices $m_i^{(j)} \in S_n$ by the j^{th}-rank of $\widehat{\tilde{\rho}}_{\alpha,m}(i)$:

$$(26) \quad \left| \widehat{\tilde{\rho}}_{\alpha,m_i^{(1)}}(i) \right| \leq \left| \widehat{\tilde{\rho}}_{\alpha,m_i^{(2)}}(i) \right| \leq ....$$

The Hill estimator nuisance index $m_i^{(j)}$ is then selected according to our chosen co-relation rank. The ranking scheme is performed for each displacement $i = 1...h$, and each $\tilde{\tau}_{\alpha,m_i^{(j)}}(i)$ is then constructed.

We then construct the test statistic functional $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$, where $\underline{m}^{(j)} \equiv (m_1^{(j)}, ..., m_h^{(j)})$ for each rank $j = 1, 2, ...$,

$$(27) \quad \tilde{Q}_{\alpha,\underline{m}^{(j)}}(h) = \frac{1}{h} \sum\nolimits_{i=1}^{h} \left(\tilde{\tau}_{\alpha,m_i^{(j)}}(i) \right)^2 .$$

In general, the rank-specific indices need not equate across displacements ($m_{i_1}^{(j)}$ $\gtreqless m_{i_2}^{(j)}$ for $i_1 \neq i_2$), and the test statistic $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$ may not have the same

[24] Of course, which values are feasible is unknown and depends on n, θ and α, and therefore also subject to arbitrariness in practice. However, simulation experiments in Section 8 in which essentially all possible m are considered provides very encouraging evidence that knowledge of the exact set of "feasible" m is not important.

rank order as $\widehat{\tilde{\rho}}_{\alpha,m_i^{(j)}}(i)$: for example, $\tilde{Q}_{\alpha,\underline{m}^{(1)}}(h)$ may not be the minimum test statistic because the scale $\widehat{\tilde{\varpi}}_{h,m_i^{(1)}}\widehat{\tilde{\sigma}}_{m_i^{(1)}}$ used in (22) need not, for example, be the maximum of all $\widehat{\tilde{\varpi}}_{h,m_i^{(j)}}\widehat{\tilde{\sigma}}_{m_i^{(i)}}$. In simulation experiments not presented in this paper, ranking the Q-statistic itself (i.e. replacing $\widehat{\tilde{\rho}}_{\alpha,m_i^{(j)}}(i)$ in (26) with $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$) generates an essentially degenerate test statistic, where empirical size *and* power are similar and often below 5% for tests at the 5%-level.

As $n \rightarrow \infty$, as long as the sequences of each chosen m_i satisfies Assumption B.2 and $m_i/m_j \rightarrow 1$ the previous asymptotic theory goes through.

Lemma 8 *Let $\sigma_m^2 = O(n^{\gamma})$, $0 \leq \gamma < \delta$. Let $\tilde{m} = \min\{m : m \in S_n\}$. Then for each $i = 1...h$ and any $m_i \in S_n$, $\widehat{\tilde{\rho}}_{\alpha,m_i}(i) = \widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(i) + o_p(1)$ and $\tilde{\tau}_{\alpha,m_i}(i) = \tilde{\tau}_{\alpha,\tilde{m}}(i) + o_p(1)$.*

Remark 1: A simple Cramer's theorem and continuous mapping theorem argument suffices to show Theorem 7 carries over to $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$. Notice $\sigma_m^2 = O(n^{\gamma})$, $0 \leq \gamma < \delta$, imitates a standard assumption effectively restricting the rate at which the asymptotic variance is allowed to diverge. Because $m = O(n^{\delta})$, the restriction implies $\sigma_m^2/m \rightarrow 0$.

The above result implies that we may use an average co-relation over a window of indices $m \in S_n$. Let $\widehat{\tilde{\rho}}_{\alpha,m} = [\widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(1), ..., \widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(h)]'$.

Corollary 9 *Let the subset $M \subseteq S_n$ have n_M elements. Then $(1/n_M)\sum_{m\in M}\widehat{\tilde{\rho}}_{\alpha,m} = \widehat{\tilde{\rho}}_{\alpha,\tilde{m}} + o_p(1)$. Specifically, $\sqrt{m}(1/n_M)\sum_{m\in M}\widehat{\tilde{\rho}}_{\alpha,m}/\tilde{\sigma}_m \Rightarrow N(0, \tilde{\Omega}_h)$.*

Remark 1: We may simply substitute $(1/n_M)\sum_{m\in M}\widehat{\tilde{\rho}}_{\alpha,m}(i)/\tilde{\sigma}_m$ for $\widehat{\tilde{\rho}}_{\alpha,m}(i)/\tilde{\sigma}_m$ in any test given above.

8. Small Sample Performance

Based on simulations not reported here the two-tailed co-relation based on the convolution summation $X_t + X_{t+h}$, and the associated Q-statistic are dominated in performance by the two-tailed $\widehat{\tilde{\rho}}_{\alpha,m_h^{(j)}}(h)$ and $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$, based on the convolution difference $X_t - X_{t+h}$. In this section, we perform a monte carlo study in order to analyze the small sample behavior of the two-tailed statistics $\widehat{\tilde{\rho}}_{\alpha,m_h^{(j)}}(h)$ and $\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)$, and for brevity we then only consider the difference of one-tailed co-relations, $\hat{\rho}_{\alpha,m}^{(1)}(h) - \hat{\rho}_{\alpha,m}^{(2)}(h)$. We comment on simulations of the one-tailed co-relations not reported here. We explicitly student Paretian random variables and extremal dependence in linear and nonlinear symmetric and asymmetric processes.

8.1 Q-Test Simulation Study

For sample sizes $n \in \{100, ..., 500\}$ we draw random samples of *iid* mean-zero time series ϵ_t from a family of Pareto distributions satisfying

(28) $F(\epsilon) = c_1|\epsilon|^{-\alpha} \quad \epsilon < 0, \quad \bar{F}(\epsilon) = c_2\epsilon^{-\alpha} \quad \epsilon > 0,$

for $\alpha \in \{1.3, 1.7\}$. We set $(c_1, c_2) \in \{(1,1), (1,2), (2,1)\}$ such that the innovation ϵ_t is, respectively, symmetric, asymmetric right and asymmetric left[25].

For empirical size we simply set $X_t = \epsilon_t$. For linear alternative models with dependent processes, we construct Self-Exciting ARMA(1,1) processes of the form

$$(29) \qquad X_t = (\phi_1 X_{t-1} + \eta_1 \epsilon_{t-1})\, I(X_{t-1} > 0) \\ + (\phi_2 X_{t-1} + \eta_2 \epsilon_{t-1})\, I(X_{t-1} \leq 0) + \epsilon_t.$$

The respective cases are *i*. AR(1): $(\phi_i, \eta_i) = (.9, 0)$, $i = 1, 2$; *ii*. MA(1): $(\phi_i, \eta_i) = (0, .8)$, $i = 1, 2$; and SETAR: $\phi_1 = \eta_1 = 0$, and $(\phi_2, \eta_2) = (.9, 0)$. In the SETAR case the process is serially independent when $X_{t-1} > 0$, and AR(1) when $X_{t-1} \leq 0$, hence asymmetrically serially extremal dependent, $\rho_\alpha^{(2)} = 0 < \rho_\alpha^{(1)}$.

We also simulate power Hyperbolic-ARCH(∞) processes of the form (cf. Davidson, 2004)

$$(30) \quad X_t = \sigma_t \epsilon_t, \ \sigma_t^p = \theta_0 + \sum\nolimits_{i=1}^{L_n} \theta_i |X_{t-i}|^p.$$

We fix $\alpha = 1.5$, randomly select $\theta_0 \in [.01, .5]$, fix $\theta_i = i^{-\mu}$, $p = 1.2$,, and use $\mu = 2$ or 4, and let $L_n = [.25n]$.

We use AR(1) and MA(1) processes of symmetric and asymmetric shocks to analyze the properties of the two-tailed co-relation; we use the SETAR process with symmetric shocks to analyze the difference in one-tailed co-relations. We use power-ARCH processes to analyze extremal dependence in X_t (independent) and $|X_t|^p$ (hyperbolic memory). We simulate $3 \times n$ observations and retain the last n. We simulate 100 series of each process, and tests are performed at the 5%-level for each $h = 1...5$. For compactness, we report only the maximum rejection frequency over all $h = 1...5$.

For the asymptotic variances $\varpi_{i,h}^2 \tilde{\sigma}_m^2 = \varpi_{i,h}^2 \alpha^2 \sigma_m^2$ and $\tilde{\varpi}_{i,h}^2 \tilde{\sigma}_m^2 = \tilde{\varpi}_{i,h}^2 \alpha^2 \sigma_m^2$ we estimate σ_m^2 using a Newey-West kernel estimator,

$$(31) \quad \hat{\sigma}_m^2 = \frac{1}{m} \sum\nolimits_{s=1}^{n} \sum\nolimits_{t=1}^{n} w((s-t)/l_n) \hat{Z}_s \hat{Z}_t$$

where $\hat{Z}_t = \left(\ln X_t / X_{(m+1)}\right)_+ - (m/n)\hat{\alpha}_m^{-1}$, and $w((s-t)/l_n)$ denotes a Bartlett kernel function with bandwidth $l_n = [\sqrt{m}]$. See Hill (2005: Section 3.3). We then construct

$$(32) \quad \tilde{\tau}_{\alpha,m}(i) = \sqrt{m}\, \widehat{\rho}_{\alpha,m}(i) / |\widehat{\tilde{\varpi}}_{h,m} \widehat{\tilde{\sigma}}_m|$$

where the estimator $\widehat{\tilde{\varpi}}_{h,m}$ is presented in (23).

In order to handle the nuisance index m problem we pursue the strategy outlined in Section 7. Test statistics $\tilde{Q}_{\alpha,\text{m}^{(j)}}(h)$ are generated for co-relation

[25] For simulations not presented here, we also drew *iid* mean-zero stable random variables: the results under null and alternative hypotheses are qualitatively similar to the Paretian case, and can be obtained from the author upon request.

ranks $j \in \{1, 5, 10, 15, 20, 25\}$: for each displacement i, each $\widehat{\widehat{\rho}}_{\alpha,m_i^{(j)}}(i)$ is derived based on the same rank j. For $\tilde{\tau}_{\alpha,m_i^{(j)}}(i)$, we use the same index $m_i^{(j)}$ for each m-dependent component: $\tilde{\tau}_{\alpha,m_i^{(j)}}(i) = (m_i^{(j)})^{1/2}\widehat{\widehat{\rho}}_{\alpha,m_i^{(j)}}(i)/|\widehat{\widehat{\varpi}}_{0,h,m_i^{(j)}}\widehat{\widehat{\sigma}}_{m_i^{(j)}}|$.

Results are compiled in Tables 2-6. Because test results for left and right skewed processes ϵ_t are qualitatively similar, we report only the symmetric and right-skewed cases. Table 2 contains comprehensive results under the null hypothesis relative to co-relation rank choice j. We highlight in bold rejection frequencies up to the 5% target level. Using 100 repetitions, the sample rejection frequencies at the 5%-level have asymptotic 99% interval lengths of $\pm 2.577 \times \sqrt{.01 \times .99/100} = \pm.0.0256$ (given the samples are independently drawn). Thus rejection frequencies near 7%-8% are expected, and are therefore highlighted as well. Let j^* denote the maximum rank such that the rejection frequency is at or below .07.

8.1.1 Null Hypothesis

Overall, empirical rejection rates under the benchmark null of independence are near the nominal 5%-level, with a uniform pattern of increasing j^* as n increases for either $\alpha \in \{1.3, 1.7\}$.

8.1.2 Alternative Hypotheses

ARMA Under either AR(1) or MA(1) hypothesis, the Q-test performs extremely well: see Table 3. For AR(1) processes with $\phi = .9$, rejection rates reach 100%, predominantly for $\alpha = 1.7$. The Q-test generates empirical powers above 90% in many cases for MA(1) processes, where empirical power rises monotonically as the sample size increases.

P-ARCH The process $\{X_t\}$ is independent noise and therefore has zero auto-co-relations. Conversely, the power process $\{|X_t|^p\}$ exhibits hyperbolic memory with co-relations bounded by a slowly varying function. Table 4 demonstrates the accuracy of the Q-test in these two cases. Interestingly, for stationary ARMA processes (which have geometric memory) the Q-test displays slowly monotonically decreasing power as the rank increases, whereas empirical power declines sharply as the rank increases for a process with hyperbolic memory. This suggests sample co-relations of a very small rank work best for hyperbolic processes.

8.1.3 Co-Relation Rank Rule: $\mathbf{j} \in [\mathbf{.01 \times n}]...[\mathbf{.03 \times n}]$

For two-tailed estimators and the two-tailed Q-test, we consistently find as a rule of thumb that ranks $j \in [.01 \times n,\ .03 \times n]$ work best depending on α and skewness. Indeed, inspecting Tables 2-4, if we simply set $j = [.01 \times n]$ *in all cases*, empirical sizes range uniformly between .00 and .03, while empirical powers are still reasonably large (above .80), in particular for $\alpha = 1.7$ and $n \geq 200$. Averaging the co-relations over ranks $j = [.01 \times n]...[.03 \times n]$ also works quite well: see below.

8.2 Co-Relation Simulation Study

We inspect the sample two-tailed co-relation coefficient $\widehat{\widehat{\rho}}_{\alpha,m_h^{(j)}}(h)$ as a measure of dependence. We restrict attention to symmetric Paretian random variables ($c_1 = c_2 = 1$), with $n = 500$. Using *iid*, AR(1) and MA(1) processes, we derive $\widehat{\rho}_{\alpha,m_h^{(j)}}(h)$ for each $h = 1...5$ and report the 95% interval length. All reported values are simulation averages over 100 repetitions. The co-relation rank j is either set to $j = [.01 \times n] = 5$ (Table 5), or the ranked co-relations are averaged over $j = [.01 \times n]...[.03\times n] = 5...15$ (Table 6).

When the rank is fixed at $[.01 \times n]$ there is a slight bias toward zero in the sample co-relation coefficients, as expected. However, zero never occurs in the AR(1) intervals and occurs in each MA(1) interval for displacements greater than 1. The rank averaged co-relations perform better under each *iid* noise and ARMA hypothesis: the true co-relation occurs in every estimated interval for the noise and AR(1) simulations, and the downward bias in the co-relations of an MA(1) process is less pronounced than when the rank is fixed.

8.3 Difference in One-Tailed Co-Relations

We use independent Paretian innovations, AR(1), and asymmetric SETAR processes. Independent and AR(1) processes are inherently symmetric, hence $\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h) = 0$ for all $h \geq 1$; for the simulated SETAR process, X_t is independent noise in the right-tail, hence $\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h) = \rho_\alpha^{(1)}(h) - 0 > 0$.

In simulations not reported here we consistently find a co-relation rank in the interval $[.03 \times n...06 \times n]$ works best for one-tailed co-relations, either alone or as a difference in co-relations. For brevity we set $\alpha = 1.5$, $n = 500$, and put $j = [.03 \times n] = 15$ (averaging over the rank interval produces qualitatively similar results). We perform one-sided tests of the hypothesis $\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h) = 0$ against $\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h) > 0$ at the 5%-level for each $h = 1...5$. Because the statistic $\tau_{\alpha,m_h^{(j)}}(h) = (m_h^{(j)})^{1/2}(\hat{\rho}^{(1)}_{\alpha,m_h^{(j)}}(h) - \hat{\rho}^{(2)}_{\alpha,m_h^{(j)}}(h))$ is degenerate under the null, for speed we simply use critical-values from the standard normal distribution, and report the maximum rejection frequency over all $h = 1...5$. Moreover, we report average estimated magnitudes of the difference, $\hat{\rho}^{(1)}_{\alpha,m_h^{(j)}}(h) - \hat{\rho}^{(2)}_{\alpha,m_h^{(j)}}(h)$, as well as approximate 95% asymptotic bounds based on the approximate asymptotic variance[26] $1/(m_h^{(j)})^{1/2}$. See Table 7.

The one-tailed co-relation difference works extremely well for inherently symmetric processes. The value 0 lies within the 95% interval of the sample co-relation difference for each h. The test rejection rate for AR(1) processes of 5.2% is relatively "large" due simply to the sample size: given the degenerate asymptotic distribution of the statistic $\tau_{\alpha,m_h^{(j)}}(h)$ we expect an under-rejection of the null for large n. For example, in a simulation with $n = 1000$ (not shown) the 5%-level empirical rejection rate was .004 for independent processes, and .008 for AR(1) processes.

For asymmetric SETAR processes the estimated tail-differences and bounds are extremely encouraging: all interval bounds averages are positive (0 does not occur in within the intervals at each displacement), and the maximum rejection frequency over all h is above 94%. When $n = 1000$ (not shown) the rejection

[26] The tail difference estimator $\hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h)$ is asymptotically degenerate under the null of no tail difference, $\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h) = 0$. Hence, the true asymptotic variance and 95% interval length are identically zero.

frequency in favor of a positive co-relation difference increases to 100%.

8.4 One-Tailed Co-Relations

For brevity we omit all simulations of one-tailed co-relations. Either right- or left-tailed co-relations, and associated Q-tests, work as well as their two-tailed counterparts. We find that a rank of $j = [.03 \times n]$, or averaging the co-relations over ranks $[.03 \times n]...[.06 \times n]$, works best as a quick rule of thumb.

8.5 Co-Relation Weakness

The above study is arguably slanted toward revealing when the sample co-relation is sharp. Consider a simple bilinear process $X_t = \sum_{i=1}^{p} \phi_i X_{t-i} u_{t-1} + u_t$ where p and ϕ are randomly selected as above. Using either a fixed co-relation rank of $[.01 \times n]$ or averaging over ranks $[.01 \times n]...[.03 \times n]$, sample co-relations and two-sided Q-tests are rather challenged to detect any degree of dependence. When $n = 500$ the maximum Q-test rejection rate over $h = 1...5$ is only .19 in a simulation of 100 series. Moreover, zero occurs in the averaged co-relation 95% intervals for each displacement $h = 1...5$.

9. Application We now analyze the serial extremal dependence properties of daily exchange rates and asset returns series for the period 1/3/00 to 8/31/05, and we perform a limited bivariate volatility spillover study on exchange rates. We consider daily logged Yen/Dollar [YEN], Euro/Dollar [EURO], and British Pound/Dollar [BP] spot rates; and the NASDAQ, S&P500 [SP500], and Shanghai Stock Exchange [SSE] composite daily open-close averages[27]. We analyze the daily difference, $\Delta X_t = X_t - X_{t-1}$ and absolute difference, $|\Delta X_t| = |X_t - X_{t-1}|$. 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.

For the daily differences ΔX_t we estimate the two-tailed co-relation and difference in one-tailed co-relations, and test the extremal white noise hypothesis for displacements $h = 1...4$. The absolute difference $|\Delta X_t|$ is maximally positively skewed, hence we measure and test only the right-tail. We perform both two-tailed and difference in tails Q-tests, and used the chi-squared distribution as a small sample approximation for the tests of difference in tails.

Results are compiled in Table 8 and Figures 1 and 2. We average the co-relations over ranks $[.01 \times n]...[.03 \times n]$ for all two-tailed applications, and $[.03 \times n]... [.06 \times n]$ for all one-tailed applications. We report the average sample tail index $\hat{\alpha}_m$ for each series ΔX_t, where the average is computed over those indices m which are derived from ranking the first auto-co-relation $\widehat{\rho}_{\alpha,m}(1)$. We report 95% bounds of $\hat{\alpha}_m$ based on a Newey-West variance estimator with Bartlett kernel.

[Insert Table 8 Here]

9.1 Tail Thickness

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

For each series, except EURO and SP500, we reject the hypothesis that the kurtosis is finite ($\alpha \geq 4$) at the 5%-level. Only YEN and SSE have unambiguously heavy tails: we reject the finite kurtosis hypothesis at the 1%-level for both rates. Notice that BP and NASDAQ have $\alpha = 4$ in the 99% interval. On the surface this suggests the tails of most of the time series analyzed here have thin enough tails to warrant classic dependence analysis. However, the fact that values $\alpha < 4$ lie in the 95% intervals for every series, and values $\alpha < 2$ lie in the 95% intervals for YEN, NASDAQ, SP500 and SSE, all suggest a method robust to thick tails is warranted. Nonetheless notice that for SP500 both $\alpha = 2$ and $\alpha = 4$ lie within the 95% interval. This is due to the large dispersion of the Hill estimator for serially dependent processes[28].

9.2 Exchange Rate Returns

Each FX returns series demonstrates persistent, low level extremal dependence, in particular YEN and BP. See Figure 1 for two-tailed and tail-difference co-relations of returns, and right-tailed co-relations of absolute returns, plotted out to 50 daily lags.

[Insert Figure 1 Here]

We strongly reject the two-tailed extremal white-noise hypothesis up to 4 lags for YEN and BP, and do not find any evidence that the nature of dependence is asymmetric in YEN. There is weak evidence that BP displays asymmetric extremal dependence favoring the right-tail: sharp devaluations in the Pound (positive spikes) are more serially dependent than large increases in value at a displacement of one day. EURO has active co-relation differences with evidence that extremal serial dependence is more acute after large devaluations in the Euro against the Dollar.

After 50 daily lags the auto-co-relations continue to be positive and significant. Notice, however, that a conjecture for any form of "decay" would be tenuous at best: both YEN and BP sample co-relations fluctuate substantially. Nonetheless the evidence strongly suggests the "efficient market hypothesis" is not supported in the extreme tails of major exchange rates. The unambiguous evidence in favor of symmetric extremal dependence suggests the nature of extremal dependence is linear.

Furthermore, all evidence supports the extremal white-noise hypothesis for the absolute return series $|\Delta X_t|$. Compare this to the stockpile of evidence suggesting absolute returns in asset markets exhibit highly persistent (possibly hyperbolic) serial dependence, cf. Ding *et al* (1993) and Granger and Ding (1996a,b).

[28] The lion's share of empirical research on tail thickness in the economics and finance literatures assumes the data are *iid* when the Hill estimator is used. Not only is this an erroneous assumption with respect to the time series typically analyzed (e.g. stock returns), but this is doubly erroneous given the extant theory of the Hill estimator under data dependence. See Hsing (1990) and Hill (2005a). The standard error of $\hat{\alpha}_m$ in the *iid* case is simply $\alpha m^{-1/2}$, but in the encompassing E-NED case it is $\alpha^2 \sigma_m m^{-1/2}$, effectively scaling the *iid* standard error by $\alpha\sigma_m$. The value of $\alpha\sigma_m$ in the *iid* case is identically 1, but can be quite large for serially dependent processes, as shown in Hill (2005a).

9.3 Asset Market Returns

The NASDAQ and SP500 returns series uniformly demonstrate a substantially smaller and more shallow degree of extremal dependence than the exchange rate series. Only the SSE series provides strong evidence in favor of persistent, symmetric extremal dependence. For NASDAQ we find asymmetric tail effects at a displacement of two days: after a one-day lag extreme values are more serially dependent when negative spikes occur than when positive spikes occur. Only SSE demonstrates significant, low level extremal dependence in absolute returns, with rhythmic fluctuations from positive to negative.

[Insert Figure 2 Here]

Recall from Section 3 that linear processes with symmetric iid shocks have inherently symmetric co-relations. Thus, we have evidence that NASDAQ, BP and EURO daily returns are either governed by a linear data generating process with asymmetric shocks, or is inherently nonlinear.

9.4 Bivariate Volatility Spillover

Finally, we perform a limited study of bivariate volatility spillover in exchange rates. An extended study will require a general asymptotic theory for bivariate processes that may be serially and mutually E-NED, but simply not extremally dependent under the null.

The volatility dynamics in international currency markets, and spillover effects across exchange rates, have been extensively studied. See Baillie and Bollerslev (1990), Engle *et al* (1990), and Hong (2001), and see Cheung and Ng (1996) for a study of causality-in-variance of financial returns. We perform tests of whether extremes in the daily returns of the Yen spillover into the Euro, and visa-versa. For a bivariate process $X_t = (X_{1,t}, X_{2,t}) \in \mathbb{R}^2$ with marginal distributions that satisfy (1), Theorem 4 implies Q-tests based on bivariate co-relations will have a limiting chi-squared distribution under the restricted null hypothesis that the two series are mutually independent. An interpretation of the co-relation for bivariate processes is straightforward. If the two-tailed $\rho_\alpha^{(0)}(X_{1,t-h}, X_{2,t}) > 0$ then spikes in $X_{1,t}$ "cause", or spillover into, $X_{2,t}$. If $\rho_\alpha^{(1)}(X_{1,t-h}, X_{2,t}) > \rho_\alpha^{(2)}(X_{1,t-h}, X_{2,t})$ then the spillover occurs predominantly in the left tail. The fundamental null hypothesis is no extremal volatility spillover, $\rho_\alpha^{(0)}(X_{1,t-h}, X_{2,t}) = 0$, $h \geq 1$.

[Insert Table 9 Here]

Results are compiled in Table 9. Two-tailed extremes in YEN significantly and symmetrically spillover into EURO after a four-day lag (i.e. ΔYEN $\rightarrow$ ΔEURO at five-days ahead), although weak evidence suggests spillover occurs at four days ahead. This suggests traders in the Euro require several trading days in order to asses the information content of large negative or positive spikes in the daily returns of the Yen.

While extremal volatility spillover from the Yen to the Euro occurs after a lengthy delay and is short lived, extreme fluctuations in the Euro spillover into the Yen the next day, and have a lasting, damping effect. We find significant and positive co-relations for displacements up to 10 trading days (i.e. two weeks). Moreover, there is weak evidence that the spillover asymmetrically occurs in the right tail at displacements of one and three days: large positive spikes in the Euro (devaluations against the Dollar) are predominantly followed by large positive spikes in the Yen.

10. Conclusion The information generated above can and should be incorporated into the forecaster's information set. While well beyond the scope of the present paper the next logical step is an attempt to incorporate serial extremal co-relation information into the specification of forecast models of extreme events. Over the period 1/3/00 - 8/31/05 the Yen-Dollar exchange rate exhibits significant, low level, symmetric and persistent serial extremal dependence in daily returns, but not daily absolute returns. This deflates classical evidence in favor of the efficient market hypothesis, suggests the nature of extremal dependence is linear rather than nonlinear, and provides rather sharp evidence the absolute returns are not fractionally integrated. Any absolute returns series with hyperbolic memory (e.g. the FIGARCH model of Granger and Ding, 1996a,b) *must have serial co-relations that exhibit hyperbolic decay.* Evidence in favor of extremal white noise, however, convincingly suggests the absolute returns series have some other form of possibly short memory that the co-relation cannot detect. All of these implications are left for future efforts.

Appendix 1: Assumptions

Define the quantile functions $b_{1,n}(m) < 0$ and $b_{2,n}(m) > 0$ by the inverse probabilities $(n/m)P(X_t < b_{1,n}(m)) \to 1, \quad (n/m)P(X_t > b_{2,n}(m)) \to 1$. See, for example, Hsing (1991). For any $\epsilon \in \mathbb{R}$ and ρ in a neighborhood of 1, define the functions $\{U_{n,t}, U^*_{n,t}(\epsilon, \rho)\}$:

$$U_{n,t} = (\ln X_t - \ln b_{2,n}(m))_+ - E(\ln X_t - \ln b_{2,n}(m))_+$$
$$U^*_{n,t}(\epsilon, \rho) = I(\ln X_t - \ln b_{2,n}(\rho m) > \epsilon) - E[I(\ln X_t - \ln b_{2,n}(\rho m) > \epsilon)].$$

Assumption A

1. For some slowly varying functions L_i, and for some $\alpha > 0$, as $|x| \to \infty$

$$F(x) = |x|^{-\alpha} L_1(x), \quad x < 0; \quad \bar{F}(x) = x^{-\alpha} L_2(x), \quad x > 0. \tag{33}$$

2. The distribution tails satisfy for some $\alpha > 0$ and $\theta > 0$, as $|x| \to \infty$

$$\begin{aligned} F(x) &= c_1(x)|x|^{-\alpha}(1 + o(|x|^{-\theta})), \quad x < 0 \\ \bar{F}(x) &= c_2(x)x^{-\alpha}(1 + o(x^{-\theta})), \quad x > 0. \end{aligned} \tag{34}$$

Assumption B 1. $m = o(n)$; 2. $m = [n^\delta]$, $0 < \delta < 2\theta/(2\theta + \alpha)$.

Assumption C Let $X^{\pm}_{(m+1)}/b_{i,n}(m) = 1 + o_p(n^{-\xi})$, $\xi \geq 0$.

Assumption D Let $\xi > (1 - \delta)\theta/\alpha$.

Assumption E Denote by $\mathcal{F}_t$ an increasing sigma field such that $\mathcal{F}_t = \sigma(\varepsilon_s < s \leq t)$ where $\{\varepsilon_t\}_{-\infty}^{\infty}$ is a stochastic process.

1. For each $U \in \{U_{n,t}, U^*_{n,t}(\epsilon, \rho)\}$, the sequence $\{U_t, \mathcal{F}_t\}_{-\infty}^{\infty}$ is an L_1-mixingale with mixingale coefficients of size $-\lambda(1 - 1/r)$ for some $\lambda > 0$, $r > 1$;

2. For each $U \in \{U_{n,t}, U^*_{n,t}(\epsilon/\sqrt{m}, 1)\}$, the sequence $\{U_t\}$ is L_r-bounded, $r \geq 1$, L_2-NED of size $-1/2$ on $\{\varepsilon_t\}_{-\infty}^{\infty}$, where ε_t is a uniform mixing process of size $-r/[2(r - 1)]$, $r \geq 2$; or strong mixing of size $-r/(r - 2)$, $r > 2$.

Appendix 2: Proofs

Proofs of the main theorems require the following lemmas. Recall the sequences $b_{1,n} = b_{1,n}(m) < 0$ and $b_{2,n} = b_{2,n}(m) > 0$ to satisfy $(m/n)P(X < b_{1,n}) \to 1$ and $(n/m)P(X > b_{2,n}) \to 1$ as $n \to \infty$. For brevity, and without loss of generality, we simply assume $P(X < b_{1,n}) = P(X > b_{2,n}) = m/n$.

Throughout we write $X^{+}_{(m+1)}$ to denote the right-tail (positive) order statistic, $X^{-}_{(m+1)}$ to denote the left-tail (negative) order statistic, and $X^{\pm}_{(m+1)}$ denotes either order statistic. For compactness, we write $Z_{(m+1)}$ to denote any of $X^{\pm}_{(m+1)}$ or $Y^{\pm}_{h,(m+1)}$, and write $b_{i,n}$ to denote any of $b_{i,n}(x)$ or $b_{i,n}(y_n)$. We will employ an arbitrary variate $\alpha_* \in (\hat{\alpha}_m, \alpha)$ throughout. Although α_* and $b_{i,n}$ depend on m, we suppress such notation. Assumption E.2 may be replaced with Assumption E.3: see Section 5.

Lemma A.1 *Under Assumption A.2*

(35) $(m/n)|b_{i,n}|^{\alpha} - c_i = o(n^{-(1-\delta)\theta/\alpha})$,

and under Assumptions A.2 and B.2, $(m/n)|b_{i,n}|^{\alpha} - c_i = o(1/\sqrt{m})$. *Moreover, under Assumptions A.2, B.2, C and D,*

(36) $(m/n)|Z_{(m+1)}|^{\alpha} = c_i + o(1/\sqrt{m})$.

Finally, under Assumptions A.2, B.2, C, D and E.2, for any $\alpha_* \in (\hat{\alpha}_m, \alpha)$

(37) $(m/n)|Z_{(m+1)}|^{\alpha_*} = c_i + o_p(1)$.

Remark 1: The subsequence proof relies on the implied $\sqrt{m}$-consistency results of Theorem 2 for $\hat{\alpha}_m$ which does not rely on the present result.

Proof of Lemma A.1. Consider the right-tail claims: the left-tail proof follows identically. Let $Z_{(i)}$ denote any right tail order statistic.

Step 1 ($(m/n)b^{\alpha}_{2,n}$): From (34) of Assumption A.2 and the definition of $b_{2,n}$, we deduce

(38) $(m/n) = c_2 b_{2,n}^{-\alpha}(1 + o(b_{2,n}^{-\theta}))$

hence

$$(39) \qquad \begin{aligned} (m/n)b^{\alpha}_{2,n} &= c_2(1 + o(b_{2,n}^{-\theta})) = c_2 + o(b_{2,n}^{-\theta})) = c_2 + o(1) \\ (m/n)b^{\alpha}_{2,n} - c_2 &= o(b_{2,n}^{-\theta}). \end{aligned}$$

Define $\tilde{c}_2 \equiv (m/n)b^{\alpha}_{2,n}$ and recall $m \approx n^{\delta}$. Then from (39) $b^{\theta}_{2,n}$ can be written as

$$
\begin{aligned}
(40) \qquad (m/n)b_{2,n}^{\alpha} &= c(1+o(b_{2,n}^{-\theta})) \\
(m/n)^{\theta/\alpha}b_{2,n}^{\theta} &= c^{\theta/\alpha}(1+o(b_{2,n}^{-\theta}))^{\theta/\alpha} \\
b_{2,n}^{\theta} &= (n/m)^{\theta/\alpha}\, c^{\theta/\alpha}(1+o(b_{2,n}^{-\theta}))^{\theta/\alpha} \\
&= (n/m)^{\theta/\alpha}\, \tilde{c}^{\theta/\alpha} \\
&= \left(n/n^{\delta}\right)^{\theta/\alpha}\, \tilde{c}^{\theta/\alpha} \\
&= n^{(1-\delta)\theta/\alpha)}\tilde{c}^{\theta/\alpha} = n^{(1-\delta)\theta/\alpha)}\left(c+o(1)\right)^{\theta/\alpha},
\end{aligned}
$$

which gives

(41) $o(b_{2,n}^{-\theta}) = o(n^{-(1-\delta)\theta/\alpha)})$.

Together, (39)-(41) give the convergence rate of the estimator $(m/n)b_{2,n}^{\alpha}$:

(42) $(m/n)b_{2,n}^{\alpha} - c_2 = o(n^{-(1-\delta)\theta/\alpha)})$.

This implies

(43) $(m/n)b_{2,n}^{\alpha} - c_2 = o(1/\sqrt{m})$,

if and only if

(44) $n^{\delta/2} \times o(n^{-(1-\delta)\theta/\alpha)}) = o(1)$,

if

(45) $\delta/2 - (1-\delta)\theta/\alpha < 0$,

which follows from simple manipulation and Assumption B.2:

$$
\begin{aligned}
(46) \quad \delta/2 - (1-\delta)\theta/\alpha &< 0 \\
\frac{\delta}{2} - \frac{\theta}{\alpha} + \frac{\delta\theta}{\alpha} &< 0 \implies \delta < \frac{2\theta}{2\theta+\alpha}.
\end{aligned}
$$

Step 2 ($(m/n)Z_{(m+1)}^{\alpha}$): Under Assumptions A.2, B.2 and C, and (34), we can write

$$
\begin{aligned}
(47) \quad (m/n)Z_{(m+1)}^{\alpha} &= \left(Z_{(m+1)}/b_{2,n}\right)^{\alpha}(m/n)b_{2,n}^{\alpha} \\
&= (1+o_p(n^{-\xi}))^{\alpha}\frac{m}{n}b_{2,n}^{\alpha} \\
&= (1+o_p(n^{-\xi}))^{\alpha}c_2(1+o(1/\sqrt{m})).
\end{aligned}
$$

The term $(1+o_p(n^{-\xi}))^{\alpha}$ is bounded by $(1+o(1/\sqrt{m}))$. In order to see this, for any $\alpha > 0$ let $d(\alpha) = [\alpha+1]$, the next integer great than α. Then for any

$\alpha > 0$

$$(48)\quad \begin{aligned} \left(1+o(n^{-\xi})\right)^{\alpha} &\leq \left(1+|o(n^{-\xi})|\right)^{\alpha} \\ &\leq \left(1+|o(n^{-\xi})|\right)^{d(\alpha)} \\ &= \sum\nolimits_{i=0}^{d(\alpha)} 1^{i}|o(n^{-\xi})|^{d(\alpha)-i}\begin{pmatrix} d(\alpha) \\ i \end{pmatrix} \\ &= 1+o(n^{-\xi}) = 1+o(1/\sqrt{m}). \end{aligned}$$

The last line follows from Assumptions B.2, C and D: $o(n^{-\xi})$ is $o(1/\sqrt{m})$ if $\sqrt{m}n^{-\xi} \approx n^{\delta/2-\xi} \to 0$, *if and only if* $\xi > \delta/2$. The equality $\xi > \delta/2$ holds sufficiently if $\xi > (1-\delta)\theta/\alpha > \delta/2$, cf. Assumption D, which is true by Assumption B.2: see (45)-(46).

Together, (47) and (48) imply

$$(49)\quad \begin{aligned} (m/n)Z_{(m+1)}^{\alpha} &= (1+o_p(n^{-\xi}))^{\alpha} c_2(1+o(1/\sqrt{m})) \\ &= c_2(1+o(1/\sqrt{m})), \end{aligned}$$

which proves (36).

Step 3 ($(m/n)Z_{(m+1)}^{\alpha_*}$): Define

$$(50)\quad \check{c}_{2,m} \equiv (m/n)Z_{(m+1)}^{\alpha}, \quad \check{c}_{2,*} \equiv (m/n)Z_{(m+1)}^{\alpha_*},$$

and observe that

$$(51)\quad \begin{aligned} \ln\ \check{c}_{2,*} &= \ln(m/n) + \alpha_* \ln Z_{(m+1)} \\ &= \ln(m/n) + (\alpha_* - \alpha)\ln Z_{(m+1)} + \alpha \ln Z_{(m+1)} \\ &= \ln(m/n)Z_{(m+1)}^{\alpha} + (\alpha_* - \alpha)\,\alpha^{-1}\ln Z_{(m+1)}^{\alpha} \\ &= \ln \check{c}_{2,m} + (\alpha_* - \alpha)\,\alpha^{-1}\ln(m/n)Z_{(m+1)}^{\alpha} \\ &\quad + \ln(n/m)\alpha^{-1}\,(\alpha_* - \alpha) \\ &= \ln \check{c}_{2,m} + (\alpha_* - \alpha)\,\alpha^{-1}\ln \check{c}_{2,m} \\ &\quad + \ln(n/m)\alpha^{-1}\,(\alpha_* - \alpha)\,. \end{aligned}$$

From (36), under Assumptions A.2, B.2, C, and D, $\check{c}_{2,m} = c_2(1 + o(1/\sqrt{m}))$, hence $\ln \check{c}_{2,m} = \ln c_2 + o(1/\sqrt{m})$, giving

$$(52)\quad \begin{aligned} \ln\ \check{c}_{2,*}/c_2 &= o(1/\sqrt{m}) + (\alpha_* - \alpha)\,\alpha^{-1}\ln c_2(1+o(1/\sqrt{m})), \\ &\quad + \ln(n/m)\alpha^{-1}\,(\alpha_* - \alpha) \end{aligned}$$

This implies $\check{c}_{2,*} = c_2 + o_p(1)$ *if and only if* $\alpha_* - \alpha$ converges at rate $o_p(1/\ln(n/m))$. Under Assumptions A.2, B.2, C, D and E.2 we deduce from Theorem 3, $\hat{\alpha}_m - \alpha = o_p(1/\sqrt{m})$, therefore, because $\alpha_* \in (\hat{\alpha}_m, \alpha)$ we obtain $\alpha_* - \alpha = o_p(1/\sqrt{m})$.

From (52) it suffices to show $\ln(n/m)/\sqrt{m} \to 0$. By Assumption B.2, $m \approx n^{\delta}$, $\delta \in (0,1)$, hence

$$(53)\quad \frac{\ln(n/m)}{\sqrt{m}} \approx \frac{(1-\delta)\ln n}{n^{\delta/2}},$$

and from l'Hôspital's rule we obtain for any $\delta \in (0,1)$

$$(54)\quad \lim_{n\to\infty} \frac{\ln(n/m)}{\sqrt{m}} \approx \lim_{n\to\infty} \frac{(\partial/\partial n)(1-\delta)\ln n}{(\partial/\partial n)n^{\delta/2}} = \lim_{n\to\infty} \frac{(1-\delta)n^{-1}}{(\delta/2)n^{\delta/2-1}}$$
$$= \lim_{n\to\infty} \frac{(1-\delta)}{(\delta/2)n^{\delta/2}} = 0.$$

Therefore, $\ln(n/m)(\alpha^* - \alpha)$ is $o_p(1)$, proving $\check{c}_{2,*} = c_2 + o_p(1)$. ■

Lemma A.2 *For any $\alpha_* \in (\hat{\alpha}_m, \alpha)$, the following point partial derivatives hold:*

$$(55)\quad \frac{\partial \hat{c}_{i,m}}{\partial \hat{\alpha}_m}|_{\alpha_*} = \check{c}_{i,*} \ln|Z_{(m+1)}|, \quad \frac{\partial \hat{\rho}^{(0)}_{\alpha,m}(h)}{\partial \hat{\alpha}_m}|_{\alpha_*} = [\check{\rho}_*(h)+1]\,\frac{1}{\alpha}\check{\varpi}_{*,m},$$

where we define

$$(56)\quad \check{\rho}_*(h) \equiv \frac{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}{2\left(\check{c}_{1,*}(x)+\check{c}_{2,*}(x)\right)} - 1$$
$$\check{\varpi}_{*,m} \equiv \ln\left[\frac{\check{c}_{1,m}(y_h)^{\psi_{1,*}(y_h)}\check{c}_{2,m}(y_h)^{\psi_{2,*}(y_h)}}{\check{c}_{1,m}(x)^{\psi_{1,*}}\check{c}_{2,m}(x)^{\psi_{2,*}}}\right]$$
$$\psi_{i,*} \equiv \frac{\check{c}_{i,*}}{\check{c}_{1,*}+\check{c}_{2,*}}, \quad \check{c}_i \equiv (m/n)|Z_{(m+1)}|^{\alpha}, \quad \check{c}_{i,*} \equiv (m/n)|Z_{(m+1)}|^{\alpha_*}.$$

Remark 1: Similar derivatives can be derived for the one-tailed $\hat{\varphi}^{(1)}_{\alpha,m}(h)$ and $\hat{\varphi}^{(2)}_{\alpha,m}(h)$.

Proof of Lemma A.2. Recalling $\hat{c}_{i,m} = (m/n)\left|Z_{(m+1)}\right|^{\hat{\alpha}_m}$ we derive

$$(57)\quad \frac{\partial \hat{c}_{i,m}}{\partial \hat{\alpha}_m}|_{\alpha_*} = \frac{\partial}{\partial \hat{\alpha}_m}\left(\frac{m}{n}|Z_{(m+1)}|^{\hat{\alpha}_m}\right)|_{\alpha_*}$$
$$= \frac{m}{n}|Z_{(m+1)}|^{\hat{\alpha}_m}\ln|Z_{(m+1)}|_{\alpha_*}$$
$$= \frac{m}{n}|Z_{(m+1)}|^{\alpha_*}\ln|Z_{(m+1)}| = \check{c}_{i,*}\ln|Z_{(m+1)}|.$$

Next, differentiating $\hat{\rho}^{(0)}_{\alpha,m}(h)$ with respect to $\hat{\alpha}_m$ and evaluating at α_*, we obtain

$$
\begin{aligned}
(58)\qquad & \frac{\partial \hat{\rho}^{(0)}_{\alpha,m}(h)}{\partial \hat{\alpha}_m}|_{\alpha_*,b_*} \\
&= \frac{\partial}{\partial \hat{\alpha}_m}\left(\frac{\hat{c}_{1,m}(y_h)+\hat{c}_{2,m}(y_h)}{2\left(\hat{c}_{1,m}(x)+\hat{c}_{2,m}(x)\right)}-1\right)|_{\alpha_*} \\
&= \frac{1}{2\left(\check{c}_{1,*}(x)+\check{c}_{2,*}(x)\right)}\left(\frac{\partial}{\partial \hat{\alpha}_m}\hat{c}_{1,m}(y_h)+\frac{\partial}{\partial \hat{\alpha}_m}\hat{c}_{2,m}(y_h)\right)|_{\alpha_*} \\
&\quad - \frac{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}{2\left(\check{c}_{1,*}(x)+\check{c}_{2,*}(x)\right)^2}\left(\frac{\partial}{\partial \hat{\alpha}_m}\hat{c}_{1,m}(x)+\frac{\partial}{\partial \hat{\alpha}_m}\hat{c}_{2,m}(x)\right)|_{\alpha_*} \\
&= \frac{1}{2\left(\check{c}_{1,*}(x)+\check{c}_{2,*}(x)\right)}\left[\check{c}_{1,*}(y_h)\ln|Y^-_{h,(m+1)}|+\check{c}_{2,*}(y_h)\ln Y^+_{h,(m+1)}\right] \\
&\quad - \frac{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}{2\left(\check{c}_{1,*}(x)+\check{c}_{2,*}(x)\right)^2}\left[\check{c}_{1,*}(x)\ln|X^-_{(m+1)}|+\check{c}_{2,*}(x)\ln X^+_{(m+1)}\right] \\
&= \left[\check{\rho}_*(h)+1\right]\left(\frac{\check{c}_{1,*}(y_h)\ln|Y^-_{h,(m+1)}|+\check{c}_{2,*}(y_h)\ln Y^+_{h,(m+1)}}{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}\right) \\
&\quad - \left[\check{\rho}_*(h)+1\right]\left(\frac{\check{c}_{1,*}(x)\ln|X^-_{(m+1)}|+\check{c}_{2,*}(x)\ln X^+_{(m+1)}}{\check{c}_{1,*}(x)+\check{c}_{2,*}(x)}\right)
\end{aligned}
$$

$$
\begin{aligned}
= \quad & [\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n}|Y^-_{h,(m+1)}|^\alpha + \check{c}_{2,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n}\left(Y^+_{h,(m+1)}\right)^\alpha}{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}\right) \\
& -[\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n} + \check{c}_{2,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n}}{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}\right) \\
& -[\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(x)\frac{1}{\alpha}\ln\frac{m}{n}|X^-_{(m+1)}|^\alpha + \check{c}_{2,*}(x)\frac{1}{\alpha}\ln\frac{m}{n}(X^+_{(m+1)})^\alpha}{\check{c}_{1,*}(x)+\check{c}_{2,*}(x)}\right) \\
& +[\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(x)\frac{1}{\alpha}\ln\frac{m}{n} + \check{c}_{2,*}(x)\frac{1}{\alpha}\ln\frac{m}{n}}{\check{c}_{1,*}(x)+\check{c}_{2,*}(x)}\right) \\
= \quad & [\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n}|Y^-_{h,(m+1)}|^\alpha + \check{c}_{2,*}(y_h)\frac{1}{\alpha}\ln\frac{m}{n}\left(Y^+_{h,(m+1)}\right)^\alpha}{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}\right) \\
& -[\check{\rho}_*(h)+1]\left(\frac{\check{c}_{1,*}(x)\frac{1}{\alpha}\ln\frac{m}{n}|X^-_{(m+1)}|^\alpha + \check{c}_{2,*}(x)\frac{1}{\alpha}\ln\frac{m}{n}\left(X^+_{(m+1)}\right)^\alpha}{\check{c}_{1,*}(x)+\check{c}_{2,*}(x)}\right) \\
= \quad & [\check{\rho}_*(h)+1]\frac{1}{\alpha}\left[\check{\psi}_{1,*}(y_h)\ln\check{c}_1(y_h)+\check{\psi}_{2,*}(y_h)\ln\check{c}_2(y_h)\right] \\
& -[\check{\rho}_*(h)+1]\frac{1}{\alpha}\left[\check{\psi}_{1,*}(x)\ln\check{c}_1(x)+\check{\psi}_{2,*}(x)\ln\check{c}_2(x)\right] \\
= \quad & [\check{\rho}_*(h)+1]\frac{1}{\alpha}\ln\left[\frac{\check{c}_1(y_h)^{\check{\psi}_{1,*}(y_h)}\check{c}_2(y_h)^{\check{\psi}_{2,*}(y_h)}}{\check{c}_1(x)^{\check{\psi}_{1,*}(x)}\check{c}_2(x)^{\check{\psi}_{2,*}(x)}}\right] = \check{\varphi}_*(h)\frac{1}{\alpha}\varpi_{*,m}.
\end{aligned}
$$

■

Lemma A.3 *Define*

$$
(59) \quad \check{\rho}_m(h) \equiv \frac{\check{c}_1(y_h)+\check{c}_2(y_h)}{2\left[\check{c}_1(x)+\check{c}_2(x)\right]} - 1, \quad \check{c}_i \equiv \frac{m}{n}|Z_{(m+1)}|^\alpha.
$$

Under Assumption A.2, B.2, C, and D, $\check{\rho}_m(h) - \rho_\alpha(h) = o(1/\sqrt{m})$.

Proof of Lemma A.3. Write

(60)

$$
\begin{aligned}
& \check{\rho}_m(h) - \rho_\alpha(h) \\
& \quad = \left(\frac{\check{c}_1(y_h)+\check{c}_2(y_h)}{2\left(\check{c}_1(x)+\check{c}_2(x)\right)} - \frac{c_1(y_h)+c_2(y_h)}{2\left(c_1(x)+c_2(x)\right)}\right) \\
& \quad = \frac{1}{2}\left(\frac{\left[\check{c}_1(y_h)+\check{c}_2(y_h)\right]\left[c_1(x)+c_2(x)\right] - \left[c_1(y_h)+c_2(y_h)\right]\left[\check{c}_1(x)+\check{c}_2(x)\right]}{\left[\check{c}_1(x)+\check{c}_2(x)\right]\left[c_1(x)+c_2(x)\right]}\right)
\end{aligned}
$$

$$
\begin{aligned}
= \quad & \frac{1}{2}\left(\frac{[\check{c}_1(y_h)-c_1(y_h)]\,c_1(x)+[\check{c}_1(x)-c_1(x)]\,c_1(y_h)}{[\check{c}_1(x)+\check{c}_2(x)]\,[c_1(x)+c_2(x)]}\right)\\
& +\frac{1}{2}\left(\frac{[\check{c}_1(y_h)-c_1(y_h)]\,c_2(x)+[\check{c}_2(x)-c_2(x)]\,c_1(y_h)}{[\check{c}_1(x)+\check{c}_2(x)]\,[c_1(x)+c_2(x)]}\right)\\
& +\frac{1}{2}\left(\frac{[\check{c}_2(y_h)-c_2(y_h)]\,c_1(x)+[\check{c}_1(x)-c_1(x)]\,c_2(y_h)}{[\check{c}_1(x)+\check{c}_2(x)]\,[c_1(x)+c_2(x)]}\right)\\
& +\frac{1}{2}\left(\frac{[\check{c}_2(y_h)-c_2(y_h)]\,c_2(x)+[\check{c}_2(x)-c_2(x)]\,c_2(y_h)}{[\check{c}_1(x)+\check{c}_2(x)]\,[c_1(x)+c_2(x)]}\right).
\end{aligned}
$$

From Lemma A.1, cf. (36), under the maintained assumptions each term $\check{c}_i - c_i$ is $o(1/\sqrt{m})$. Therefore each denominator can be written as

$$
\begin{aligned}
(61) \qquad & (\check{c}_1(x)+\check{c}_2(x))\,(c_1(x)+c_2(x))\\
= \quad & [\check{c}_1(x)-c_1(x)+\check{c}_2(x)-c_2(x)]\,[c_1(x)+c_2(x)]+[c_1(x)+c_2(x)]^2\\
= \quad & o(n^{-(1-\delta)\theta/\alpha})\times(c_1(x)+c_2(x))+(c_1(x)+c_2(x))^2\\
= \quad & (c_1(x)+c_2(x))^2+o(n^{-(1-\delta)\theta/\alpha}).
\end{aligned}
$$

Similarly, each numerator is $o(1/\sqrt{m})$. We conclude $\check{\rho}_m(h) \;-\; \rho_\alpha(h) = o(1/\sqrt{m})$.
■

Proof of Theorem 4. We prove the limit for the two-tailed coefficient $\hat{\rho}^{(0)}_{\alpha,m}(h)$. Proofs of the one-tailed $\hat{\rho}^{(i)}_{\alpha,m}(h)$, $i = 1, 2$, and for the two-tailed $\widehat{\bar{\rho}}_{\alpha,m}(h)$ follow using identical methods (which slight adjustments to the line of proof of Lemma A.2).

Expanding $\hat{\varphi}^{(0)}_{\alpha,m}(h)$ around α, by the mean-value theorem for each n there exists some $\alpha_* \in (\hat{\alpha}_m, \alpha)$ such that

$$
(62) \quad \hat{\rho}^{(0)}_{\alpha,m}(h)=\check{\rho}^{(0)}_{m}(h)+\frac{\partial\hat{\rho}^{(0)}_{\alpha,m}(h)}{\partial\hat{\alpha}_m}|_{\alpha_*}(\hat{\alpha}_m-\alpha),
$$

where we define

$$
(63) \quad \check{\rho}^{(0)}_{m}(h)\equiv\frac{\check{c}_1(y_h)+\check{c}_2(y_h)}{2\,(\check{c}_1(x)+\check{c}_2(x))}-1, \quad \check{c}_i\equiv(m/n)|Z_{(m+1)}|^{\alpha}.
$$

Using the partial derivatives $\partial\hat{\rho}^{(0)}_{\alpha,m}(h)/\partial\hat{\alpha}_m$ from Lemma A.2, (62) is written as

$$
(64) \quad \hat{\rho}^{(0)}_{\alpha,m}(h)=\check{\rho}^{(0)}_{m}(h)+\left[\check{\rho}^{(0)}_{*}(h)+1\right]\frac{1}{\alpha}\check{\varpi}_{*,m}(h)(\hat{\alpha}_m-\alpha)
$$

where

$$
\begin{aligned}
(65) \qquad & \check{\rho}^{(0)}_{*}(h)\equiv\frac{\check{c}_{1,*}(y_h)+\check{c}_{2,*}(y_h)}{2\,(\check{c}_{1,*}(x)+\check{c}_{2,*}(x))}+1\\
& \check{\varpi}_{*,m}(h)\equiv\ln\left[\frac{\check{c}_1(y_h)^{\check{\psi}_{1,*}(y_h)}\check{c}_2(y_h)^{\check{\psi}_{2,*}(y_h)}}{\check{c}_1(x)^{\check{\psi}_{1,*}(x)}\check{c}_2(x)^{\check{\psi}_{2,*}(x)}}\right]\\
& \check{\psi}_{i,*}\equiv\frac{\check{c}_{i,*}}{\check{c}_{1,*}+\check{c}_{2,*}}, \quad \check{c}_{i,*}\equiv\frac{m}{n}\left|Z_{(m+1)}\right|^{\alpha_*},
\end{aligned}
$$

hence (64) simplifies to

$$\sqrt{m}\left(\frac{\hat{\rho}^{(0)}_{\alpha,m}(h)-\rho^{(0)}_{\alpha}(h)}{\check{\rho}^{(0)}_{*}(h)+1}\right) \tag{66}$$
$$= \sqrt{m}\left(\frac{\check{\rho}^{(0)}_{m}(h)-\rho^{(0)}_{\alpha}(h)}{\check{\rho}^{(0)}_{*}(h)+1}\right)+\check{\varpi}_{*,m}(h)\sqrt{m}\left(\frac{\hat{\alpha}_m-\alpha}{\alpha}\right)$$

Under the maintained assumptions, we deduce by Lemma A.1 the weak limits $\check{c}_{i,*} \to c_i$ and $\check{c}_i \to c_i$, hence by the Slutsky theorem $\psi_{i,*} \to \psi_i$, $\check{\rho}^{(0)}_{*}(h) \to \rho^{(0)}_{\alpha}(h)$ and $\check{\varpi}_{*,m}(h) \to \varpi^{(0)}_h$, where $0 \leq |\varpi_h| < \infty$ by construction, and

$$\varpi^{(0)}_h = \ln\left[\frac{c_1(y_h)^{\psi_1(y_h)}c_2(y_h)^{\psi_2(y_h)}}{c_1(x)^{\psi_1(x)}c_2(x)^{\psi_2(x)}}\right]. \tag{67}$$

Moreover, by Lemma A.3 we have $\check{\rho}^{(0)}_{m}(h) - \rho^{(0)}_{\alpha}(h) = o_p(1/\sqrt{m})$ under the maintained assumptions. Together, we deduce the first term on the right-hand-side of (66) is $o_p(1)$. By Cramér's theorem, $\tilde{\sigma}^2_m = O(n^\gamma)$, $\gamma \geq 0$, and Theorem 3 we deduce from (66)

$$\sqrt{m}\left(\frac{\hat{\rho}^{(0)}_{\alpha,m}(h)-\rho^{(0)}_{\alpha}(h)}{\left[\check{\rho}^{(0)}_{*}(h)+1\right]\tilde{\sigma}_m}\right) \tag{68}$$
$$= o_p(1)+\sqrt{m}\left(\frac{\hat{\alpha}_m-\alpha}{\alpha\tilde{\sigma}_m}\right)\check{\varpi}_{*,m}(h) \Rightarrow N\left(0,[\varpi^{(0)}_h]^2\right),$$

Because $\check{\rho}^{(0)}_{*}(h) \to \rho^{(0)}_{\alpha}(h)$, by Cramér's theorem

$$\sqrt{m}\left(\frac{\hat{\rho}^{(0)}_{\alpha,m}(h)-\rho^{(0)}_{\alpha}(h)}{\left[\rho^{(0)}_{\alpha}(h)+1\right]\tilde{\sigma}_m}\right) \Rightarrow N(0,[\varpi^{(0)}_h]^2). \tag{69}$$

When $\rho^{(0)}_{\alpha}(h) = 0$, (68) and (69) imply for any h

$$\sqrt{m}\hat{\rho}^{(0)}_{\alpha,m}(h)/\tilde{\sigma}_m = o_p(1)+\sqrt{m}\left(\frac{\hat{\alpha}_m-\alpha}{\alpha\tilde{\sigma}_m}\right)\varpi^{(0)}_h \Rightarrow N(0,[\varpi^{(0)}_h]^2). \tag{70}$$

Because each $\hat{\rho}^{(0)}_{\alpha,m}(h)$ is stochastically grounded on the same random variable $\hat{\alpha}_m$, cf. (70), the joint distribution is simply

$$\sqrt{m}\hat{\rho}^{(0)}_{\alpha,m}/\tilde{\sigma}_m = \sqrt{m}\left[\hat{\rho}^{(0)}_{\alpha,m}(1)/\tilde{\sigma}_m,...,\hat{\rho}^{(0)}_{\alpha,m}(h)/\tilde{\sigma}_m\right]' \tag{71}$$
$$= \sqrt{m}\left(\frac{\hat{\alpha}_m-\alpha}{\alpha}\right)[\varpi_{0,1},...,\varpi_{0,h}]' \Rightarrow N\left(0,\Omega^{(0)}_h\right)$$

where $\Omega^{(0)}_h = [\varpi^{(0)}_j\varpi^{(0)}_k]^h_{j,k=1}$.

Finally, if $\rho_\alpha^{(0)}(h) \neq 0$ then $\hat{\rho}_{\alpha,m}^{(0)}(h) \to \rho_\alpha^{(0)}(h) \neq 0$, cf. Theorem 2, hence

$$(72)\quad m \times \hat{\rho}_{\alpha,m}^{(0)}(h)^2 \to \infty$$

with probability one. ■

Proof of Theorem 5. Using the logic of the lines of proof of Theorem 4 and Lemma A.2, it is straightforward to prove for each $i = 1, 2$ the one-tailed $\hat{\rho}_{\alpha,m}^{(i)}(h)$ satisfy

$$(73)\quad \sqrt{m}\left(\frac{\hat{\rho}_{\alpha,m}^{(i)}(h) - \rho_{\alpha,m}^{(i)}(h)}{\left[\rho_{\alpha,m}^{(i)}(h) + 1\right]\tilde{\sigma}_m}\right) = \varpi_h^{(i)}\sqrt{m}\left(\frac{\hat{\alpha}_m - \alpha}{\alpha\tilde{\sigma}_m}\right) + o_p(1) \Rightarrow N(0, [\varpi_h^{(i)}]^2),$$

where $\varpi_h^{(i)} = \ln[c_i(y_h)/c_i(x)]$. By Cramér's theorem

$$\begin{aligned}
(74)\quad & \sqrt{m}\left(\left[\hat{\rho}_{\alpha,m}^{(1)}(h) - \hat{\rho}_{\alpha,m}^{(2)}(h)\right] - \left[\rho_\alpha^{(1)}(h) - \rho_\alpha^{(2)}(h)\right]\right)/\tilde{\sigma}_m \\
&= \sqrt{m}\left(\left[\hat{\rho}_{\alpha,m}^{(1)}(h) - \rho_\alpha^{(1)}(h)\right] - \left[\hat{\rho}_{\alpha,m}^{(2)}(h) - \rho_\alpha^{(2)}(h)\right]\right)/\tilde{\sigma}_m \\
&= \sqrt{m}\left((1 + \rho_\alpha^{(1)}(h))\varpi_h^{(1)} - (1 + \rho_\alpha^{(2)}(h))\varpi_h^{(2)}\right)\left(\frac{\hat{\alpha}_m - \alpha}{\alpha\tilde{\sigma}_m}\right) + o_p(1) \\
&\Rightarrow N\left(0, \left(\varpi_h^{(1)}(1 + \rho_\alpha^{(1)}(h)) - \varpi_h^{(2)}(1 + \rho_\alpha^{(2)}(h))\right)^2\right).
\end{aligned}$$

If $\rho_\alpha^{(1)}(h) = \rho_\alpha^{(2)}(h)$ then by definition $\rho_\alpha^{(1)}(h) = c_1(y_h)/2c_1(x) - 1 = \rho_\alpha^{(2)}(h) = c_2(y_h)/2c_2(x) - 1$, hence $c_1(y_h)/c_1(x) = c_2(y_h)/c_2(x)$, giving $\varpi_h^{(1)} = \varpi_h^{(2)}$ by definition. From (74) we conclude

$$(75)\quad \sqrt{m}\left(\hat{\rho}_{\alpha,m}^{(1)}(h) - \hat{\rho}_{\alpha,m}^{(2)}(h)\right)/\tilde{\sigma}_m \Rightarrow 0.$$

■

Proof of Theorem 7. Recall $H_0 : \tilde{\rho}_\alpha(i) = 0$, $i = 1...h$. By Theorem 2 we have the weak limit $\hat{c}_{i,m} \to c_i$, hence by the Slutsky theorem $\hat{\tilde{\varpi}}_{0,j,m} \to \tilde{\varpi}_{0,j}$ for any $j = 1, 2, ...$ From Theorem 4, cf. (70) in the line of proof, and by Theorem 2 and Cramér's theorem, under the null we deduce for each i,

$$\begin{aligned}
(76)\quad \tilde{\tau}_{\alpha,m}(i) &= \sqrt{m}\frac{\hat{\tilde{\rho}}_{\alpha,m}(i)}{\sqrt{\hat{\tilde{\varpi}}_{i,m}^2\hat{\tilde{\sigma}}_m^2}} \\
&= \sqrt{m}\frac{(\hat{\alpha}_m - \alpha)}{\alpha\tilde{\sigma}_m}\left(\frac{\tilde{\sigma}_m\breve{\varpi}_{*,m}(i)}{\hat{\tilde{\sigma}}_m|\hat{\tilde{\varpi}}_{i,m}|}\right) + o_p(1) \Rightarrow N(0,1),
\end{aligned}$$

where $\breve{\varpi}_{*,m}(i)$ is defined in (63); $\breve{\varpi}_{*,m}(i)/|\hat{\tilde{\varpi}}_{i,m}| \to 1$ or -1 by Theorem 2 and the construction of $\breve{\varpi}_{*,m}(i)$; and $\tilde{\sigma}_m/\hat{\tilde{\sigma}}_m \to 1$ follows from the definitions $\tilde{\sigma}_m^2 = \alpha^2\sigma_m^2$ and $\hat{\tilde{\sigma}}_m^2 = \hat{\alpha}_m^2\hat{\sigma}_m^2$, Theorem 1, and $\text{plim}_{n\to\infty}\hat{\sigma}_m^2/\sigma_m^2 = 1$, cf. Hill

(2005: Lemma 5). Therefore, by the continuous mapping theorem and Cramér's theorem, the portmanteau-statistic $\tilde{Q}_{\alpha,m}(h)$ satisfies for each $h \geq 1$

$$
\begin{aligned}
(77)\quad \tilde{Q}_{\alpha,m}(h) &= \frac{1}{h}\sum\nolimits_{i=1}^{h} \tilde{\tau}_{\alpha,m}(i)^2 \\
&= \frac{1}{h}\sum\nolimits_{i=1}^{h} \left(m\left(\frac{\hat{\alpha}_m - \alpha}{\alpha\tilde{\sigma}_m}\right)^2 \frac{\tilde{\sigma}_m^2 \breve{\varpi}_{*,m}^2(i)}{\widehat{\tilde{\sigma}}_m^2 |\widehat{\tilde{\varpi}}_{i,m}^2|} + o_p(1)\right) \\
&= \left[\sqrt{m}\left(\frac{\hat{\alpha}_m - \alpha}{\alpha\tilde{\sigma}_m}\right)\right]^2 \times \frac{1}{h}\sum\nolimits_{i=1}^{h} \left(\frac{\tilde{\sigma}_m^2 \breve{\varpi}_{*,m}^2(i)}{\widehat{\tilde{\sigma}}_m^2 |\widehat{\tilde{\varpi}}_{i,m}^2|}\right) + o_p(1) \\
&\Rightarrow \chi^2(1).
\end{aligned}
$$

Under the alternative $\tilde{\rho}_\alpha(i) \neq 0$ for at least one $i = 1...h$. By Theorem 2 we have under the maintained assumptions the weak limits $\hat{c}_{i,m} \rightarrow c_i$ and $\widehat{\tilde{\rho}}_{\alpha,m}(i) \rightarrow \tilde{\rho}_\alpha(i) \neq 0$, which implies $\tilde{Q}_{\alpha,m}(h) \rightarrow \infty$ with probability one. ■

Proof of Lemma 8. We will prove $|\hat{\alpha}_{m_i} - \hat{\alpha}_{\tilde{m}}| = o_p(1/\sqrt{m_i})$ and $\tilde{\sigma}_{m_i}^2/\tilde{\sigma}_{\tilde{m}}^2 = 1 + o_p(1)$ in two steps. We then prove $\widehat{\tilde{\rho}}_{\alpha,m_i}(i) = \widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(i) + o_p(1)$ and $\tilde{\tau}_{\alpha,m_i}(j) = \tilde{\tau}_{\alpha,\tilde{m}}(j) + o_p(1)$.

Step 1: Consider any $m_i \sim [n^\delta]$ such that $m_i/m_j \rightarrow 1$, $i \neq j$. First observe that

$$
(78)\quad \alpha^{-1}(\hat{\alpha}_{m_i} - \alpha) = -\hat{\alpha}_{m_i}\left(\hat{\alpha}_{m_i}^{-1} - \alpha^{-1}\right),
$$

where

$$
\begin{aligned}
(79)\qquad \hat{\alpha}_{m_i}^{-1} &= \frac{1}{m_i}\sum\nolimits_{j=1}^{m_i} \ln X_{(j)}^{+} - \ln X_{(m_i+1)}^{+} \\
&= \frac{\tilde{m}}{m_i}\left(\frac{1}{\tilde{m}}\sum\nolimits_{j=1}^{\tilde{m}} \ln X_{(j)}^{+} - \ln X_{(\tilde{m}+1)}^{+}\right) \\
&\quad + \frac{1}{m_i}\sum\nolimits_{j=\tilde{m}+1}^{m_i} \ln X_{(j)}^{+} + \frac{\tilde{m}}{m_i}\ln X_{(\tilde{m}+1)}^{+} - \ln X_{(m_i+1)}^{+} \\
&= \frac{\tilde{m}}{m_i}\hat{\alpha}_{\tilde{m}}^{-1} + \frac{1}{m_i}\sum\nolimits_{j=\tilde{m}+1}^{m_i} \ln X_{(j)}^{+} + \frac{\tilde{m}}{m_i}\ln X_{(\tilde{m}+1)}^{+} - \ln X_{(m_i+1)}^{+}
\end{aligned}
$$

hence

$$
\begin{aligned}
(80)\qquad \left|\hat{\alpha}_{m_i}^{-1} - \hat{\alpha}_{\tilde{m}}^{-1}\right| &\leq o(1/m_i) + \left|\frac{1}{m_i}\sum\nolimits_{j=\tilde{m}+1}^{m_i} \ln X_{(j)}^{+}\right| \\
&\quad + \left|\frac{\tilde{m}}{m_i}\ln X_{(\tilde{m}+1)}^{+} - \ln X_{(m_i+1)}^{+}\right|,
\end{aligned}
$$

where the $o(1/m_i)$ term is identically $\hat{\alpha}_{\tilde{m}}^{-1}|m_i - \tilde{m}|/m_i$, due to $|m_i - \tilde{m}| \rightarrow 0$.

We will show the 2nd and 3rd terms on the right-hand-side of (80) are $o_p(1/\sqrt{m_i})$ implying $|\hat{\alpha}_{m_i}^{-1} - \hat{\alpha}_{\tilde{m}}^{-1}| = o_p(1/\sqrt{m_i})$, hence $|\hat{\alpha}_{m_i} - \hat{\alpha}_{\tilde{m}}| = o_p(1/\sqrt{m_i})$, cf. (78) and Theorem 1.

For the 3rd term, by Assumptions A.2, B.2, C, and D, Lemma A.1 and the fact that $\tilde{m}/m_i \to 1$,

$$
\begin{aligned}
(81)\quad & \frac{\tilde{m}}{m_i} \ln X^+_{(\tilde{m}+1)} - \ln X^+_{(m_i+1)} \\
&= (1 + o(1/m_i)) \ln X^+_{(\tilde{m}+1)} - \ln X^+_{(m_i+1)} \\
&= \ln X^+_{(\tilde{m}+1)}/X^+_{(m_i+1)} + o(1/m_i) \ln X^+_{(\tilde{m}+1)} \\
&= \ln(1 + o_p(n^{-\xi}))/(1 + o_p(n^{-\xi})) + o(1/m_i) \ln \left[b_{2,n}(\tilde{m})(1 + o_p(n^{-\xi}))\right] \\
&= o_p(n^{-\xi}) + o(1/m_i) \ln b_{2,n}(\tilde{m}) \\
&= o_p(n^{-\xi}) + o(1/m_i)\alpha^{-1} \ln(\tilde{m}/n) b^{a}_{2,n}(\tilde{m}) + o(1/m_i) \ln(n/\tilde{m}) \\
&= o_p(n^{-\xi}) + o(1/m_i)\alpha^{-1} \ln c_2(1 + o_p(1/\sqrt{\tilde{m}})) + o(1/m_i) \ln(n/\tilde{m}).
\end{aligned}
$$

Using l'Hôspital's rule, notice for all $\delta \in (0, 1)$

$$
(82)\quad \lim_{n\to\infty} \frac{\ln(n/\tilde{m})}{\sqrt{m_i}} = \lim_{n\to\infty} \frac{(1-\delta)\ln n}{n^{\delta/2}} = \lim_{n\to\infty} \frac{(1-\delta)n^{-1}}{(\delta/2)n^{\delta/2-1}} = 0,
$$

hence $o(1/m_i)\ln(n/\tilde{m}) = o(1/\sqrt{m_i})$. Moreover the $o_p(n^{-\xi})$ term is $o_p(1/\sqrt{m_i})$ due to $m_i \approx n^{\delta}$, and $\xi > \delta/2$ under Assumption B.2 (see the line of proof of Lemma A.1). Thus,

$$
\begin{aligned}
(83)\quad & \frac{\tilde{m}}{m_i} \ln X^+_{(\tilde{m}+1)} - \ln X^+_{(m_i+1)} \\
&= o_p(n^{-\xi}) + o(1/m_i)\alpha^{-1} \ln c_2(1 + o_p(1/\sqrt{\tilde{m}})) + o(1/m_i) \ln(n/\tilde{m}) \\
&= o_p(1/\sqrt{m_i}).
\end{aligned}
$$

For the 2nd term on the right-hand-side of (80), observe that

$$
(84)\quad (n/j)P\left(X_i \geq b_{2,n}(j)\right) \to 1,
$$

by construction of the sequence $b_{2,n}(\cdot)$, hence for any $j = \tilde{m} + 1 ... m_i$, as $n \to \infty$

$$
(85)\quad \frac{(n/j)P\left(X_i \geq b_{2,n}(j)\right)}{(n/\tilde{m})P\left(X_i \geq b_{2,n}(\tilde{m})\right)} = \frac{\tilde{m}}{j} \times \frac{P\left(X_i \geq b_{2,n}(j)\right)}{P\left(X_i \geq b_{2,n}(\tilde{m})\right)} \to 1.
$$

This implies $P(X_i \geq b_{2,n}(j)) \geq P(X_i \geq b_{2,n}(\tilde{m}))$ as $n \to \infty$, hence $b_{2,n}(j) \leq b_{2,n}(\tilde{m})$ as $n \to \infty$. By Assumption C, Lemma A.1, $m_i \sim [n^{\delta}]$ and $\tilde{m}/m_i \to 1$, we have as $n \to \infty$

$$(86)\quad \frac{1}{m_i}\sum\nolimits_{j=\tilde{m}+1}^{m_i} \ln X_{(j)}^{+}$$
$$= \frac{1}{m_i}\sum\nolimits_{j=\tilde{m}+1}^{m_i} \left[\ln b_{2,n}(j) + o_p(n^{-\xi})\right]$$
$$\le \left(\frac{m_i-\tilde{m}}{m_i}\right)\ln b_{2,n}(\tilde{m}) + \left(\frac{m_i-\tilde{m}}{m_i}\right) o_p(n^{-\xi})$$
$$= o(1/m_i)\times\alpha^{-1}\ln(\tilde{m}/n)b_{2,n}^{\alpha}(\tilde{m})$$
$$+\left(\frac{m_i-\tilde{m}}{m_i}\right)\ln(n/m_i) + o(1/\tilde{m})$$
$$= o(1/m_i)\times\alpha^{-1}\ln c_2(1+o_p(1/\sqrt{m_i})) + \left(\frac{m_i-\tilde{m}}{m_i}\right)\ln\frac{n}{m_i} + o(1/m_i)$$
$$\approx o_p(1/m_i) + o_p(1/m_i)\ln\frac{n}{m_i} = o_p(1/\sqrt{m_i}),$$

due to $o(1/\tilde{m}) = o(1/m_i)$ from the assumption $\tilde{m}/m_i \to 1$, and (82). Therefore, for any $m_i \in S_n$

$$(87)\quad \hat{\alpha}_{m_i}^{-1} = \hat{\alpha}_{\tilde{m}}^{-1} + o_p(1/\sqrt{m_i}).$$

Step 2: The proof that $\tilde{\sigma}_{m_i}^2/\tilde{\sigma}_{\tilde{m}}^2 = 1 + o_p(1)$ runs more or less parallel to the Step 1. Briefly,

$$(88)\quad \tilde{\sigma}_{m_i}^2 = \frac{1}{m_i}\sum\nolimits_{s,t=1}^{n} w_{s,t}[\left(\ln X_s/X_{(m_i+1)}\right)_+ - (m_i/n)\hat{\alpha}_{m_i}^{-1}]$$
$$\times[\left(\ln X_t/X_{(m_i+1)}\right)_+ - (m_i/n)\hat{\alpha}_{m_i}^{-1}]$$
$$= (1+o(1/m_i))\frac{1}{\tilde{m}}\sum\nolimits_{s,t=1}^{n} w_{s,t}[\left(\ln X_s/X_{(\tilde{m}+1)}\right)_+ - (\tilde{m}/n)\hat{\alpha}_{\tilde{m}}^{-1}]$$
$$\times[\left(\ln X_t/X_{(\tilde{m}+1)}\right)_+ - (\tilde{m}/n)\hat{\alpha}_{m_i}^{-1}] + R_n$$

where the remaining term R_n contains summations involving $w_{s,t}$-weighted products and cross-products of $[(\tilde{m}/n)\hat{\alpha}_{\tilde{m}}^{-1} - (m_i/n)\hat{\alpha}_{m_i}^{-1}]$, $[\left(\ln X_t/X_{(m_i+1)}\right)_+ - (m_i/n)\hat{\alpha}_{m_i}^{-1}]$ and $[\left(\ln X_t/X_{(\tilde{m}+1)}\right)_+ - \left(\ln X_t/X_{(m_i+1)}\right)_+]$. Step 1 implies

$$(89)\quad (\tilde{m}/n)\hat{\alpha}_{\tilde{m}}^{-1} - (m_i/n)\hat{\alpha}_{m_i}^{-1} = o_p(1/\sqrt{m_i})$$
$$[\left(\ln X_t/X_{(\tilde{m}+1)}\right)_+ - \left(\ln X_t/X_{(m_i+1)}\right)_+] = o_p(1/\sqrt{m_i}),$$

and using Theorem 2 it is not difficult to prove

$$(90)\quad \frac{1}{m_i}\sum\nolimits_{s,t=1}^{n} w_{s,t}[\left(\ln X_t/X_{(m_i+1)}\right)_+ - (m_i/n)\hat{\alpha}_{m_i}^{-1}]$$
$$= \frac{1}{m_i}\sum\nolimits_{s,t=1}^{n} w_{s,t}\left(\ln X_t/X_{(m_i+1)}\right)_+ - \hat{\alpha}_{m_i}^{-1}\frac{1}{n}\sum\nolimits_{s,t=1}^{n} w_{s,t} \to 0.$$

Repeated application of the Cauchy-Schwartz inequality then suffices to show $R_n = o_p(1)$.

Step 3: Imitating the expansion in (70), we deduce for any m

$$\sqrt{m}\,\widehat{\tilde{\rho}}_{\alpha,m}(h)/\tilde{\sigma}_m = \sqrt{m}\,(\hat{\alpha}_m - \alpha)\,\alpha^{-1}\tilde{\sigma}_m^{-1}\varpi_h^{(0)} + o_p(1) \tag{91}$$

hence

$$\begin{aligned}
&\sqrt{m_i}\,\widehat{\tilde{\rho}}_{\alpha,m_i}(h)/\tilde{\sigma}_{m_i} - \sqrt{\tilde{m}}\,\widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(h)/\tilde{\sigma}_{\tilde{m}} \\
&\quad = \left(\sqrt{m_i}\hat{\alpha}_{m_i}\alpha^{-1}\tilde{\sigma}_{m_i}^{-1} - \sqrt{\tilde{m}}\hat{\alpha}_{\tilde{m}}\alpha^{-1}\tilde{\sigma}_{\tilde{m}}^{-1}\right)\varpi_h^{(0)} \\
&\quad\quad - \left(\sqrt{m_i}\tilde{\sigma}_{m_i}^{-1} + \sqrt{\tilde{m}}\tilde{\sigma}_{\tilde{m}}^{-1}\right)\varpi_h^{(0)} + o_p(1)
\end{aligned} \tag{92}$$

which reduces to

$$\begin{aligned}
&\hat{\rho}_{\alpha,m_i}^{(0)}(h) - \hat{\rho}_{\alpha,\tilde{m}}^{(0)}(h) \\
&\quad = -\hat{\rho}_{\alpha,\tilde{m}}^{(0)}(h)\left(1 - \frac{\sqrt{\tilde{m}}}{\sqrt{m_i}}\frac{\tilde{\sigma}_{m_i}}{\tilde{\sigma}_{m_i}}\right) + \hat{\alpha}_{m_i}\left(1 - \frac{\sqrt{\tilde{m}}}{\sqrt{m_i}}\frac{\hat{\alpha}_{\tilde{m}}\tilde{\sigma}_{m_i}}{\hat{\alpha}_{m_i}\tilde{\sigma}_{\tilde{m}}}\right)\alpha^{-1}\varpi_h^{(0)} \\
&\quad\quad - \left(1 - \frac{\sqrt{\tilde{m}}}{\sqrt{m_i}}\frac{\tilde{\sigma}_{m_i}}{\tilde{\sigma}_{\tilde{m}}}\right)\varpi_h^{(0)} + \tilde{\sigma}_{m_i} \times o_p(1/\sqrt{m_i}).
\end{aligned} \tag{93}$$

From Steps 1 and 2, and the assumption $\tilde{m}/m_i \to 1$, the first three terms on the right-hand-side are $o_p(1)$. Moreover $\tilde{\sigma}_{m_i}^2 \times o_p(1/m_i) = o_p(1)$ by the assumption $\tilde{\sigma}_{m_i}^2 = O(n^{\gamma})$, $\gamma < \delta$, given $m_i = O(n^{\delta})$. Thus

$$\hat{\rho}_{\alpha,m_i}^{(0)}(h) - \hat{\rho}_{\alpha,\tilde{m}}^{(0)}(h) = o_p(1) \tag{94}$$

as claimed.

Combined with Step 2 we conclude

$$\begin{aligned}
\tilde{\tau}_{\alpha,m_i}(i) &= \sqrt{m_i}\frac{\widehat{\tilde{\rho}}_{\alpha,m_i}(i)}{\sqrt{\widehat{\tilde{\varpi}}_{i,m_i}^2\widehat{\tilde{\sigma}}_{m_i}^2}} \\
&= \sqrt{\tilde{m}}\frac{\widehat{\tilde{\rho}}_{\alpha,\tilde{m}}(i)}{\sqrt{\widehat{\tilde{\varpi}}_{i,\tilde{m}}^2\widehat{\tilde{\sigma}}_{\tilde{m}}^2}}\left(\frac{\sqrt{m_i}}{\sqrt{\tilde{m}}}\right)\left(\frac{\sqrt{\widehat{\tilde{\varpi}}_{i,\tilde{m}}^2\widehat{\tilde{\sigma}}_{\tilde{m}}^2}}{\sqrt{\widehat{\tilde{\varpi}}_{i,m_i}^2\widehat{\tilde{\sigma}}_{m_i}^2}}\right) + o_p(1) \\
&= \tilde{\tau}_{\alpha,\tilde{m}}(i) \times [1 + o_p(1)] + o_p(1) = \tilde{\tau}_{\alpha,\tilde{m}}(i) + o_p(1).
\end{aligned} \tag{95}$$

■

References

[1] Baillie, R. and T. Bollerslev, 1990, Intra-Day and Inter-Market Volatility in Foreign Exchange Rates, Review of Economic Studies 58, 565-585.

[2] Basrak, B., R. Davis, and T. Mikosch, 2001, Regular Variation and GARCH Processes, mimeo, Laboratory of Actuarial Mathematics, University of Copenhagen.

[3] Bekeart, G. and C. R. Harvey, 1997, Emerging Equity Market Volatility, Journal of Financial Economics 43, 29-77.

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

[5] Black, F., 1976, Studies of Stock Market Volatility, in American Statistical Association (eds.), Proceedings of the American Statistical Association: Business and Economic Statistics Section (Chicago).

[6] Bollerslev, T., 1986, Generalized Autoregressive Conditional Heteroscedasticity, Journal of Econometrics 31, 307-327.

[7] Box G. and D. Pierce, 1970, Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, Journal of the American Statistical Association 65, 1509-1526.

[8] Cambell, J. Y. and L. Hentschel, 1992, No News is Good News - An Asymmetric Model of Changing Volatility in Stock Returns, Journal of Financial Economics 31, 281-318.

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

[10] Chan, N.H., and L.T. Tran, 1989, On the First Order Autoregressive Process with Infinite Variance, Econometric Theory 5, 354-362.

[11] Chen, X., Y. Fan and A. Patton, 2004, Simple Tests for Models of Dependence Between Multiple Financial Time Series, with Applications to U.S. Equity Returns and Exchange Rates, London Economics Financial Markets Group Working Paper No. 483.

[12] Chernozhukov, V., 2005, Extremal Quantile Regression, Annals of Statistics 33, 806–839.

[13] Cheung, Y. and L. Ng, 1996, A Causality-in-Variance Test and Its Applications to Financial Markets, Journal of Econometrics 72, 33-48.

[14] Cline, D. B. H., 1983, Estimation and Linear Prediction for Regression, Autoregression and ARMA with Infinite Variance Data, Unpublished Ph.D. Dissertation, Department of Statistics, Colorado State University.

[15] Cline, D. B. H., 1986, Convolution Tails, Product Tails and Domains of Attraction, Probability Theory and Related Fields 72, 525-557.

[16] Danielsson, J., L. de Haan, L. Peng and C. G. de Vries, 1998, Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation, mimeo, Eramus University Rotterdam.

[17] Davidson, J., 1994, Stochastic Limit Theory (Oxford Univ. Press: Oxford).

[18] Davidson, J., 2004, Moment and Memory Properties of Linear Conditional Heteroscedasticity Models, and a New Model, Journal of Business and Economics Statistics 22, 16-29.

[19] Davis, R.A. and S.I. Resnick 1985, More limit theory for the sample correlation function of moving averages, Stochastic Processes and Their Applications 20, 257-279.

[20] Davis, R.„ and W. Wu, 1997, Bootstrapping M-Estimates in Regression and Autoregression with Infinite Variance, Staistica Sinica 7, 1135-1154.

[21] de Lima, Pedro J. F., 1996, Nuisance Parameter Free Properties of Correlation Integral Based Statistics, Econometric Reviews 15, 237-259.

[22] Dekkers, A.L.M., J.H.J. Einmahl, and L. de Haan, 1989, A Moment Estimator for the Index of an Extreme-Value Distribution, Annals of Statistics 17, 1833-1855.

[23] Ding, Z., R. Engle and C. Granger, 1993, A Long Memory Property of Stock Market Returns and a New Model, Journal of Empirical Finance 1, 83-106.

[24] Draisma, G., L. de Haan, L. Peng, and T.T. Periera, 1997, A Bootstrap-based Method to Achieve Optimality in Estimating the Extreme-value Index, Eramus University, Econometric Institute Report EI 2000-18/A.

[25] Drees, H., 2003, Extreme Quantile Estimation for Dependent Data, with Applications to Finance, Bernoulli 9, 617-657.

[26] Dumouchel, W. H., 1983, Estimating the Stable Index α in order to Measure Tail Thickness, Annals of Statistics 11, 1019-1036.

[27] Embrechts, P. and C. M. Goldie, 1980, On Closure and Factorization Properties of Subexponential Distributions, Journal of Aust. Math. Soc. (Ser. A) 29, 243-256.

[28] Embrechts, P. and C. M. Goldie, 1982, On Convolution Tails, Stochastic Processes and their Applications 13, 263-278.

[29] Engle, R., 1982, Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of U.K. Inflation, Econometrica 50, 987-1008.

[30] Engle, R. and V. K. Ng, 1993, Measuring and Testing the Impact of News of Volatility, Journal of Finance 48, 1749-1778.

[31] Engle, R., T. Ito, and W. Lin, 1990, Meteor Showers or Heat Waves? Heteroscedastic Intra-Day Volatility in the Foreign Exchange Rate Market, Econometrica 58, 525-542.

[32] Fama, E., 1965, Portfolio Analysis in a Stable Paretian Market, Management Science 11, 404-419.

[33] Feller, William, 1971, An Introduction to Probability Theory and its Applications, 2 ed., vol. 2, (Wiley: New York).

[34] Forbes, K.J. and R. Rigobon, 2002, No Contagion, Only Interdependence: Measuring Stock Market Comovements, Journal of Finance 57, 2223-2261.

[35] Gallagher, C., 2002, Testing for Linear Dependence in Heavy-Tailed Data, Communications in Statatistics: Theory and Methods 31, 611-623.

[36] Gallant, A. R. and H. White, 1988, A Unified Theory of Estimation and Inference for Nonlinear Dynamic Models (Basil Blackwell: Oxford).

[37] González-Rivera, G., 1998, Smooth-Transition GARCH Models, Nonlinear Dynamics and Econometrcs 3, No. 2, Article 1.

[38] Granger, C.W.J., and Z. Ding, 1996a, Modeling Volatility Persistence of Speculative Returns: A New Approach, Journal of Econometrics 73, 185-215.

[39] Granger, C.W.J., and Z. Ding, 1996b, Some Properties of Absolute Returns: An Alternative Measure of Risk, Annales d'Economie et de Statistique 40, 67-90.

[40] Haan, L. de, and L. Peng, 1998, Comparison of Tail Index Estimators, Statsitica Neerlandica 52, 60-70.

[41] Haan, L. de, and S. Resnick, 1977, Limit Theory for Multivariate Sample Extremes, Z. Wahr. verw. Geb. 40, 317-37.

[42] Hall, P., 1982, On Some Estimates of an Exponent of Regular Variation, Journal of the Royal Statistical Society 44, 37-42.

[43] Hall, P., and C.C. Hyde, 1980, Martingale Limit Theory and Its Applications. New York: Academic Press.

[44] Hall, P. and A.H. Welsh, 1985, Adaptive Estimates of Parameters of Regular Variation, Annals of Statistics 3, 1163-1174.

[45] Hall, P. and Q. Yao, 2003, Inference in Arch and Garch Models with HeavyTailed Errors , Econometrica 71, 285-318.

[46] Hamilton, J., 1989, A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, Econometrica 57, 357-384.

[47] Hill, B.M., 1975, A Simple General Approach to Inference about the Tail of a Distribution, Annals of Mathematical Statistics 3, 1163-1174.

[48] Hill, J.B., 2005, On Tail Index Estimation Using Dependent, Heterogenous Data, Dept. of Economics, Florida International; available at http://econwpa.wustl.edu:80/eps/em/papers/0505/0505005.pdf.

[49] Hong, Y., 2001, A Test for Volatility Spillover with Application to Exchange Rates, Journal of Econometrics 103, 183-224.

[50] Hsing, T., 1991, On Tail Index Estimation Using Dependent Data, Annals of Statistics 19, 1547-1569.

[51] Ibragimov, I. A., 1962, Some Limit Theorems for Stationary Processes, Theory of Probability and its Applications 7, 349-382.

[52] Ibragimov, I. A. and Y. V. Linnik, 1971, Independent and Stationary Sequences of Random Variables (Wolters-Noordhof).

[53] Knight, K., 1993, Estimation in Dynamic Linear Regression Models with Infinite Variance Errors, Econometric Theory 9, 570-588.

[54] Kokoszka, P. S. and M. S. Taqqu, 1994, Infinite Variance Stable ARMA Processes, Journal of Time Series Analysis 15, 203-220.

[55] Kokoszka, P. S. and M. S. Taqqu, 1996a, Infinite Variance Stable Moving Averages with Long Memory, Journal of Econometrics 73, 79-99.

[56] Kokoszka, P. S. and M. S. Taqqu, 1996b, Parameter Estimation for Infinite Variance Fractional ARIMA, Annals of Statitics 24, 1880-1913.

[57] Ledford, A. W. and J. A. Tawn, 1996, Statistics for Near Independence in Multivariate Extreme Values, Biometrika 83, 169-187.

[58] Ledford, A. W. and J. A. Tawn, 1997, Modeling Dependence within Joint Tail Regions, Journal of the Royal Statistical Society B 59, 475-499.

[59] 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.

[60] Ling, S., 2005, Self-Weighted Least Absolute Deviation Estimation for Infinite Variance Autoregressive Models, Journal of the Royal Statistical Society, Series B 67, 381-393.

[61] Linton, O. and Y-J Wang, 2004, A Quantilogram Approach to Evaluating Directional Probability, manuscript, London School of Economics.

[62] Liu, S. and B. W. Brorsen, 1995, Maximum Likelihood Estimation of a Garch-Stable Model, Journal of Applied Econometrics 10, 273-285.

[63] Longin, F. and B. Solnik, 2001, Extreme Correlation of International Equity Markets, Journal of Finance 56, 649-676.

[64] Loretan, M., 1991, Testing Covariance Stationarity of Heavy-Tailed Economic Time Series, Unpublished Ph.D. Dissertation, Dept. of Economics, Yale University.

[65] Loretan, M. and P.C.B. Phillips, 1991, The Durbin-Watson Ratio under Infinite-Variance Errors, Journal of Econometrics 47, 85-114 .

[66] Loretan, M. and P.C.B. Phillips, 1994, Testing the Covariance Stationarity of Heavy-Tailed Economic Time Series: An Overview of the Theory with Applications to Financial Data Sets, Journal of Empirical Finance 1, 211-48.

[67] Mandelbrot, B., 1961, Stable Paretian Random Fluctuations and the Multiplicative Variation of Income, Econometrica, 29, 517-543.

[68] Mandelbrot, B., 1963, The Variation of Certain Speculative Prices, Journal of Business 36, 394-419.

[69] McLeish, D. L., 1975, A Maximal Inequality and Dependent Strong Law, Annals of Probability 3, 829-839.

[70] McCulloch, J. Huston, 1985, Interest-Risk Sensitive Deposit Insurance Premia: Stable ACH Estimates, Journal of Banking and Finance 9, 137-56.

[71] McCulloch, J. H., 1997, Measuring Tail Thickness in Order to Estimate the Stable Index α: A Critique, Journal of Business and Economic Statistics 15, 74-81.

[72] 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.

[73] Mikosch, T., and C. Stărică, 2000, Limit Theory for the Sample Autocorrelations and Extremes of the GARCH(1,1) Process, Annals of Statistics 28, 1427-1451.

[74] Mittnik, S., M.S. Paolella, and S.T. Rachev, 2002, Stationarity of Stable Power-GARCH Processes, Journal of Econometrics 106, 97-107,

[75] Mittnik, S., S.T. Rachev, and L. Rüschendorf, 1999, Test of Association Between Multivariate Stable Vectors, Mathematical and Computer Modelling 29, 181-195.

[76] Pagan, A., and G. Schwert, 1990, Alternative Models of Conditional Volatility, Journal of Econometrics 45, 267-290.

[77] Patton, A., 2002, Modeling Time-Varying Exchange Rate Dependence Using the Conditional Copula, Dept. of Economics, University of California-San Diego.

[78] Peng, L., 1999, Estimation of the Coefficient of Tail Dependence in Bivariate Extremes, Statistical Probability Letters 43, 399-409.

[79] Peng, L., and Q. Yao, 2004, Nonparametric Regression under Dependent Errors with Infinite Variance, Annals of the Institute of Statistical Mathematics 56, 73-86.

[80] Pickands, J., 1975, Statistical Inference usng Extreme Order Statistics, Annals of Statistics 3, 119-131.

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

[82] Pötscher, B. M., and I. R. Prucha, 1991, Basic Structure of the Asymptotic Theory in Dynamic Nonlinear Econometric Models, Part I: Consistency and Approximation Concepts, Econometric Reviews 10, 125-216.

[83] Quintos, C., Z. Fan and P. C. B. Phillips, 2001, Structural Change Tests in Tail Behavior and the Asian Crisis, Review of Economics Studies 68, 633-663.

[84] Quintos, C., 2004, Extremal Correlation for GARCH Data, manuscript, Dept. of Applied Statistics, University of Rochester.

[85] Rachev, S. T., 2003, Handbook of Heavy Tailed Distributions in Finance (Elsevier Science, New York).

[86] Rachev, S.T., and H. Xin, 1993, Test on Association of Random Variables in the Domain of Attraction of a Multivariate Stable Law,

[87] Resnick, S., 1987, Extreme Values, Regular Variation and Point Processes (Springer-Verlag: New York).

[88] Resnick, S., 1996, Why Non-Linearities Can Ruin the Heavy Tailed Modeler's Day, mimeo, Dept. Statistics, Cornel University.

[89] Resnick, S., 2002, The Extremal Dependence Measure and Asymptotic Independence, mimeo, Dept. Statistics, Cornel University.

[90] Resnick, S. and C. Stărică, 1995, Consistency of Hill's Estimator for Dependent Data, Journal of Applied Probability 32, 139-167.

[91] Resnick, S. and C. Stărică, 1997, Smoothing the Hill Estimator, Advances in Applied Probability 29, 271-293.

[92] Resnick, S. and C. Stărică, 1998, Tail Index Estimation for Dependent Data, The Annals of Applied Probability 8, 1156-1183.

[93] Runde, Ralf, 1997, The Asymptotic Null Distribution of the Box-Pierce Q-Statistic for Random Variables with Infinite Variance with An Application to German Stock Returns, Journal of Econometrics 78, 205-216.

[94] Samorodnitsky, G. and M. S. Taqqu, 1994, Stable Non-Gaussian Random Processes (Chapman and Hall: New York).

[95] Schmidt, R. and U. Stadtmüller, 2004, Nonparametric Estimation of Tail Dependence, manuscript, London School of Economics.

[96] Sentana, E., 1995, Quadratic ARCH Models, Review of Economic Studies 62, 639-661.

[97] Sibuya, M., 1960, Bivariate Extreme Statistics, Annals of the Institute of Statistical Mathematics 11, 195-210.

[98] Stărică, C., 1999, Multivariate Extremes for Models with Constant Conditional Correlations, Journal of Empirical Finance 6, 515-553.

[99] Teräsvirta, T., 1994, Specification, Estimation, and Evaluation of Smooth Transition Autoregressive Models, Journal of the American Statistical Association 89, 208-218.

[100] Yang, A., A.P. Petropulu, and J-C. Pesquet, 2001, Estimating Long-Range Dependence in Impulsive Traffic Flows, Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing 6, 3413-3416.

Table 1

$\tilde{\rho}_\alpha(h)$ and $\rho_\alpha^{(i)}(h)$, $i = 0, 1, 2$, for ARMA(1,1)[a]

	$\tilde{\rho}_\alpha(h)^b$				$\rho_\alpha^{(i)}(h)^c$			
	AR(1)		MA(1)		AR(1)		MA(1)	
	$\phi = .9$, $\eta = 0$		$\phi = .0$, $\eta = .8$		$\phi = .9$, $\eta = 0$		$\phi = .0$, $\eta = .8$	
$h\backslash\alpha$	1.3	1.7	1.3	1.7	1.3	1.7	1.3	1.7
1	.8892	.8746	.4647	.4808	.2157	.5708	.1141	.3063
2	.7900	.7652	.0000	.0000	.2011	.5215	.0000	.0000
3	.7016	.6700	.0000	.0000	.1873	.4761	.0000	.0000
4	.6232	.5871	.0000	.0000	.1743	.4345	.0000	.0000
5	.5535	.5150	.0000	.0000	.1619	.3963	.0000	.0000
6	.4916	.4522	.0000	.0000	.1503	.3612	.0000	.0000
7	.4367	.3975	.0000	.0000	.1394	.3291	.0000	.0000
8	.3880	.3498	.0000	.0000	.1292	.2996	.0000	.0000
9	.3448	.3080	.0000	.0000	.1200	.2727	.0000	.0000
10	.3065	.2716	.0000	.0000	.1106	.2480	.0000	.0000

Notes: a. The process follows $X_t = \phi X_{t-1} + \eta\epsilon_{t-1} + \epsilon_t$, $\epsilon_t \sim$Pareto, $c_1 = c_2 = 1$.
b. $\tilde{\rho}_\alpha(h)$ is based on the convolution difference $X_t - X_{t-h}$.
c. $\rho_\alpha^{(i)}(h)$ is based on the convolution summation $X_t + X_{t-h}$.

Table 2

Two-Tailed $\max_{1 \le h \le 5}\{\tilde{Q}_{\alpha,\mathrm{m}^{(j)}}(h)\}^a$

H_0 : *extremal white noise*

		$c_2/c_1 = 2$						$c_2/c_1 = 1$					
α	$n \setminus j$	1	5	10	15	20	25	1	5	10	15	20	25
1.30	100	**.07**	.24	.70	.94	1.0	1.0	**.02**	.16	.55	.87	1.0	1.0
	200	**.02**	.08	.18	.37	.60	.74	**.00**	**.03**	**.05**	.14	.29	.47
	300	**.01**	**.05**	.08	.16	.29	.37	**.00**	**.01**	**.02**	**.05**	.09	.17
	400	**.00**	**.01**	**.04**	**.07**	.15	.20	**.00**	**.01**	**.01**	**.01**	**.06**	.10
	500	**.00**	**.01**	**.02**	**.03**	**.06**	.09	**.00**	**.00**	**.00**	**.02**	**.02**	**.03**
1.70	100	**.05**	.37	.74	.97	1.0	1.0	**.02**	.31	.61	.88	.97	.96
	200	**.00**	.21	.32	.57	.71	.84	**.00**	**.06**	.35	.49	.540	.65
	300	**.00**	**.06**	.27	.42	.59	.70	**.00**	**.04**	.19	.33	.51	.56
	400	**.00**	**.03**	.15	.23	.42	.54	**.00**	**.02**	.12	.26	.39	.47
	500	**.00**	**.03**	.14	.21	.37	.42	**.00**	**.01**	**.06**	.20	.32	.41

Notes: a. Tabulated frequencies are for the maximum Q-statistic over displacements h = 1...5.

Table 3

Two-Tailed $\max_{1\leq h\leq 5}\{\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)\}$

ARMA(1,1)

			$c_2/c_1 = 2$						$c_2/c_1 = 1$					
	α	$n \setminus j$	1	5	10	15	20	25	1	5	10	15	20	25
MA(1)[a]	1.3	100	**.35**	.85	.99	.95	.75	.42	**.37**	.76	.94	.99	.97	.90
		200	**.43**	.81	.93	.96	.99	1.0	**.38**	**.69**	**.83**	.93	.93	1.0
		300	**.32**	**.69**	.90	.96	.99	1.0	**.31**	**.74**	**.92**	**.98**	1.0	1.0
		400	**.37**	**.73**	**.92**	**.98**	.99	1.0	**.32**	**.72**	**.89**	**.98**	**.98**	.99
		500	**.47**	**.79**	**.91**	**.96**	**.98**	1.0	**.34**	**.73**	**.89**	**.95**	**.98**	**.99**
	1.7	100	**.37**	.91	.96	.94	.81	.44	**.42**	.77	.97	.99	.97	.93
		200	**.37**	.77	.97	1.0	1.0	1.0	**.33**	**.73**	.89	.95	.99	1.0
		300	**.38**	**.73**	.90	.95	.99	1.0	**.33**	**.65**	.79	.91	.94	.98
		400	**.40**	**.74**	.86	.97	.98	1.0	**.30**	**.65**	.87	.92	.96	.99
		500	**.58**	**.81**	.85	.91	.97	1.0	**.31**	**.70**	**.88**	.97	.96	.97
AR(1)[b]	1.3	100	**.61**	.39	.31	.24	.17	.12	**.77**	.79	.72	.66	.57	.38
		200	**.68**	.67	.67	.57	.50	.43	**.95**	**1.0**	**1.0**	.98	.93	.93
		300	**.74**	**.82**	.74	.70	.69	.60	**.96**	**1.0**	**1.0**	**1.0**	1.0	1.0
		400	**.87**	**.91**	**.91**	**.91**	.85	.85	**.97**	**1.0**	**1.0**	**1.0**	**1.0**	1.0
		500	**.85**	**.98**	**.95**	**.94**	**.94**	.93	**.98**	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**
	1.7	100	**.63**	.42	.38	.29	.29	.19	**.85**	.82	.70	.59	.52	.40
		200	**.70**	**.70**	.58	.50	.48	.36	**.96**	**1.0**	.95	.91	.89	.89
		300	**.85**	**.89**	.78	.76	.73	.73	**.98**	**1.0**	1.0	1.0	1.0	1.0
		400	**.86**	**.91**	.84	.81	.80	.78	**.98**	**1.0**	1.0	1.0	1.0	1.0
		500	**.91**	**.95**	.94	.93	.93	.91	**.98**	**1.0**	**1.0**	1.0	1.0	1.0

Notes: a. The process is $X_t = .8\epsilon_{t-1} + \epsilon_t$; b. The process is $X_t = .9X_{t-1} + \epsilon_t$.

Table 4

Two-Tailed $\max_{1\leq h\leq 5}\{\tilde{Q}_{\alpha,\underline{m}^{(j)}}(h)\}$

Power Hyberbolic ARCH$(\infty)^a$, $p = 1.2$, $\alpha = 1.5$

		$X_t = \sigma_t\epsilon_t$						$\lvert X_t\rvert^p = \sigma_t^p\lvert\epsilon_t\rvert^p$					
μ	$n \setminus j$	1	5	10	15	20	25	1	5	10	15	20	25
2	100	**.04**	.32	.59	.72	.84	.94	**.92**	.41	.13	.04	.00	.00
	200	**.02**	**.07**	.16	.30	.44	.51	**.97**	**.93**	.75	.15	.08	.03
	300	**.00**	**.04**	**.06**	.10	.16	.22	**1.0**	**.95**	**.90**	**.85**	.64	.45
	400	**.00**	**.03**	**.06**	.08	.13	.15	**1.0**	**1.0**	**.95**	**.86**	.79	.63
	500	**.00**	**.02**	**.05**	**.07**	.12	.13	**1.0**	**1.0**	**.99**	**.93**	**.87**	.79
4	100	.13	.37	.65	.79	.83	.94	**.97**	.49	.01	.00	.00	.00
	200	**.01**	**.06**	.19	.33	.50	.63	**1.0**	**.99**	.81	.22	.02	.01
	300	**.01**	**.04**	.09	.14	.18	.31	**1.0**	**1.0**	**.94**	**.68**	.30	.22
	400	**.00**	**.01**	**.03**	.13	.17	.25	**1.0**	**1.0**	**1.0**	**.92**	.83	.67
	500	**.00**	**.00**	**.02**	**.07**	.12	.18	**1.0**	**1.0**	**1.0**	**1.0**	**1.0**	.92

Notes: a. The process is $X_t = \sigma_t\epsilon_t$, $\sigma_t^p = \theta_0 + \sum_{i=1}^{L_n}\theta_i\lvert X_{t-i}\rvert^p$, $\theta_i = i^{-\mu}$, $\mathrm{L_n} = [.25\mathrm{n}]$

Table 5

Two-Tailed Sample Co-Relation[a]

ARMA(1,1), $c_1 = c_2 = 1$, $n = 500$

	$\alpha = 1.3$		$\alpha = 1.7$	
	AR(1)	MA(1)	AR(1)	MA(1)
h	$\hat{\tilde{\rho}}_{\alpha,m_h^{(j)}} \pm k^b$	$\hat{\tilde{\rho}}_{\alpha,m_h^{(j)}} \pm k$	$\hat{\tilde{\rho}}_{\alpha,m_h^{(j)}} \pm k$	$\hat{\tilde{\rho}}_{\alpha,m_h^{(j)}} \pm k$
1	$.675 \pm .12^c$	$.239 \pm .14$	$.660 \pm .02$	$.252 \pm .12$
2	$.540 \pm .04$	$.012 \pm .16$	$.591 \pm .07$	$.012 \pm .13$
3	$.436 \pm .20$	$.006 \pm .17$	$.482 \pm .01$	$.001 \pm .14$
4	$.364 \pm .57$	$.015 \pm .15$	$.386 \pm .02$	$.009 \pm .14$
5	$.231 \pm .11$	$.014 \pm .15$	$.281 \pm .06$	$.013 \pm .13$

Notes: a. Co-relation rank is j = [.01n].

b. 95% interval width $1.96 \times (\hat{\tilde{\varpi}}^2_{h,m_h^{(j)}}\hat{\tilde{\sigma}}^2_{m_h^{(j)}}/m_h^{(j)})^{1/2}$.

c. True co-relation values are in Table 1.

Table 6
Two-Tailed Sample Co-Relation Averages[a]
$\widehat{\widehat{\rho}}_{\alpha,*}(h) = 1/n_M \sum_{m \in M} \widehat{\widehat{\rho}}_{\alpha,m}(h)$
ARMA(1,1), $c_1 = c_2 = 1$, $n = 500$

	$\alpha = 1.3$			$\alpha = 1.7$		
	iid	AR(1)	MA(1)	iid	AR(1)	MA(1)
h	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K^b$	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K$	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K$	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K$	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K$	$\widehat{\widehat{\rho}}_{\alpha,*}\pm K$
1	-.013±.17	.746±.15	.308±.09	.019±.14	.887±.15	.337±.06
2	-.012±.17	.663±.10	.028±.15	.012±.15	.842±.13	.026±.13
3	-.003±.17	.545±.03	.029±.14	.011±.15	.731±.09	.025±.13
4	-.009±.17	.420±.04	.036±.15	.018±.15	.604±.05	.025±.13
5	-.008±.17	.342±.07	.031±.14	.017±.15	.487±.02	.028±.13

Notes: a. Ranked co-relations are averaged over ranks j = [.01n]...[.03n].
b. K denotes the 95% interval width averaged over co-relation ranks j = [.01n]...[.03n].

Table 7
Co-Relation Difference
$\Delta\hat{\rho}^{(i)}_{\alpha,m} = \hat{\rho}^{(1)}_{\alpha,m}(h) - \hat{\rho}^{(2)}_{\alpha,m}(h)$
SETAR[a]

	iid	AR(1)	SETAR
h	$\Delta\hat{\rho}^{(i)}_{\alpha,m}\pm K$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}\pm K$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}\pm K$
1	.002±.15	.006±.17	.327±.32
2	.001±.17	-.006±.17	.405±.30
3	.002±.15	.008±.15	.612±.32
4	.001±.15	-.010±.19	.532±.32
5	.001±.16	-.004±.18	.813±.21
	(.013)[b]	(.052)	(.936)

Notes: a. The processes is $X_t = \phi X_{t-1} \times I(X_{t-1} \leq 0) + \epsilon_t$: iid noise when $X_t > 0$, and AR(1) when $X_t \leq 0$. For iid processes, $\phi = 0$; for AR and SETAR, $\phi = .9$. c_1=c_2=1, α=1.5, n=500, and the co-relation rank is fixed at j=[.03×n]
b. Parenthetical values contain maximum rejection frequencies over h = 1...5 of the Z-test of the hypothesis that the difference in one-tailed co-relations is zero against a one-sided alternative, performed at the 5%-level using the standard normal distribution.

Table 8

Q-Tests, Co-relations and Differences in Co-relations[a]

	ΔYEN[b]		ΔEURO		ΔBP	
h	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$
1	.244* [.00][c]	-.001 [.82][d]	.255* [.00]	-.203* [.00]	.139* [.01]	.004 [.06]
2	.162* [.00]	-.001 [.66]	.050 [.00]	.016 [.00]	.159* [.00]	.030 [.11]
3	.172* [.00]	.008 [.68]	.058 [.00]	.012 [.01]	.276* [.00]	-.003 [.17]
4	.127* [.00]	-.001 [.70]	.051 [.00]	-.194* [.01]	.227* [.00]	.002 [.19]
$\hat{\alpha}_m$	2.55±.70[e]		3.37±1.20		2.96±1.03	
	ΔNASDAQ		ΔSP500		ΔSSE	
h	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$	$\widehat{\tilde{\rho}}_{\alpha,m}$	$\Delta\hat{\rho}^{(i)}_{\alpha,m}$
1	.167* [.01]	.011 [.20]	-.054 [.45]	-.016 [.82]	.253* [.01]	.050 [.27]
2	.009 [.01]	.176* [.00]	.012 [.36]	.018 [.72]	.129* [.00]	.030 [.20]
3	.029 [.01]	.029 [.00]	-.061 [.30]	-.011 [.74]	.278* [.00]	-.013 [.26]
4	.002 [.02]	.075 [.00]	-.107* [.04]	.007 [.76]	-.022 [.00]	.055 [.11]
$\hat{\alpha}_m$	2.44±1.55		2.61±1.52		2.50±1.28	
	|ΔYEN|	|ΔEURO|	|ΔBP|	|ΔNAS|	|ΔSP500|	|ΔSSE|
h	$\hat{\rho}^{(2)}_{\alpha,m}$	$\hat{\rho}^{(2)}_{\alpha,m}$	$\hat{\rho}^{(2)}_{\alpha,m}$	$\hat{\rho}^{(2)}_{\alpha,m}$	$\hat{\rho}^{(2)}_{\alpha,m}$	$\hat{\rho}^{(2)}_{\alpha,m}$
1	-.003 [.53][f]	-.002 [.25]	-.010 [.28]	-.024 [.77]	.010 [.12]	.052* [.05]
2	-.003 [.55]	-.003 [.29]	-.006 [.28]	-.012 [.66]	.009 [.09]	-.060* [.02]
3	-.000 [.53]	-.004 [.31]	-.002 [.28]	.044 [.68]	.008 [.08]	-.060* [.01]
4	.003 [.53]	-.008 [.32]	-.009 [.27]	.032 [.68]	.005 [.06]	-.069* [.01]

Notes: a. Ranked two-tailed co-relations are averaged over ranks [.01n]...[.03n]; one-tailed and difference-in-tails co-relations are averaged over ranks [.03n]...[.06n].

b. BP = British pound; SSE = Shanghai Stock Exchange.

c. Two-tailed co-relation, with Q-statistic p-value in [...]. An asterisk * denotes significance of the co-relation at the 5%-level.

d. Difference in one-tailed co-relations with two-sided test p-value in [...] based on a $\chi^2(1)$ distribution. A positive value implies more extremal dependence in the left tail.

e. Tail index and 95% interval length based on a Newey-West asymptotic variance.

f. Right-tailed co-relation, with Q-statistic p-value in [...].

Table 9

Bivariate Volatility Spillover[a]

	ΔYEN→ ΔEURO		ΔEURO→ ΔYEN	
h	$\widehat{\tilde{\rho}}_{\alpha,m} \pm K$	$\Delta\hat{\rho}^{(i)}_{\alpha,m} \pm K$	$\widehat{\tilde{\rho}}_{\alpha,m} \pm K$	$\Delta\hat{\rho}^{(i)}_{\alpha,m} \pm K$
1	.052±.253 (.348)[b]	-.007±.120 (.415)	.295±.002* (.000)	-.147±.191 (.130)
2	-.001±.121 (.338)	-.004±.263 (.503)	.132±.017* (.000)	.032±.516 (.283)
3	.041±.142 (.285)	.001±.341 (.530)	.109±.040* (.000)	-.104±.090* (.116)
4	.163±.177 (.166)	.003±.340 (.564)	.128±.033* (.000)	.049±.188 (.166)
5	.134±.097* (.028)	.008±.541 (.572)	.020±.011* (.000)	.022±.624 (.215)

Notes: a. * denotes co-relation significance at the 5%-level; b. Two-tailed Q-test p-values.

Figure 1
Daily Exchange Rate Returns

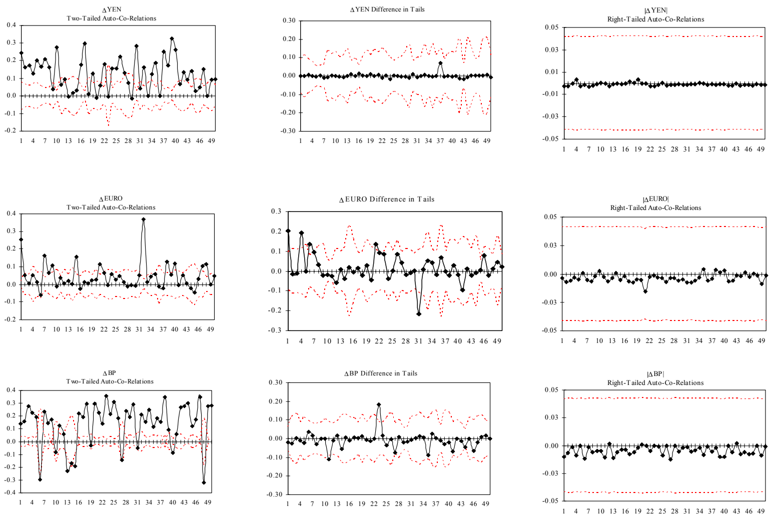

Figure 2
Daily Stock Market Index Returns