EXTREMES OF STOCHASTIC VOLATILITY MODELS

By F. Jay Breidt[1] and Richard A. Davis[2]

Iowa State University and Colorado State University

Extreme value theory for a class of stochastic volatility models, in which the logarithm of the conditional variance follows a Gaussian linear process, is developed. A result for the asymptotic tail behavior of the transformed stochastic volatility process is established and used to prove that the suitably normalized extremes converge in distribution to the double exponential (Gumbel) distribution. Explicit normalizing constants are obtained, and point process convergence is discussed.

1. Introduction. Financial variables such as stock returns are often modeled using martingale difference (MD) sequences, since such models are theoretically justified by efficient economic markets and are empirically useful. If a sequence of random variables is MD with constant unconditional variance, then it is also serially uncorrelated, and so it is white noise in the sense of uncorrelated $(0, \sigma^2)$ random variables.

An independent and identically distributed (iid) sequence with finite variance is both a MD sequence and a white noise, but series arising in finance and econometrics often cannot be assumed iid even when the MD property appears plausible [e.g., Clark (1973); Tauchen and Pitts (1983); Melino and Turnbull (1990)]. Instead, the variance in a given realization seems to change smoothly over time, in the sense that large observations tend to be followed by large observations and small observations by small observations, as Mandelbrot (1963) noted.

Two approaches have been proposed to model time-dependent variances, while maintaining the MD property. The first approach, proposed by Engle (1982) and later generalized by Bollerslev (1986) and by others, uses an autoregressive conditionally heteroscedastic (ARCH, or its generalized version, GARCH) process to model the serial autocorrelation in the variances. In this approach, the variance of the series at time t is assumed to be a deterministic function of lagged values of the squared observations and of past variances. For a review of this approach, see Bollerslev, Chou and Kroner (1992).

The second approach, pioneered in its earliest version in the work of Clark (1973), uses models known as stochastic volatility (SV) models. In this context, it is assumed that smooth functions of the time-dependent variances are random variables generated by an underlying stochastic process, for example an autoregressive process. Stochastic volatility models also result from dis-

cretizing continuous-time diffusion processes such as those proposed for asset pricing [Hull and White (1987)].

Given the potential impact on investors, maxima and minima of such processes are naturally of interest. Extremal behavior for a first-order ARCH process is described in de Haan, Resnick, Rootzén and de Vries (1989). Geweke (1994) has compared the observed maximum absolute returns for five stocks and two stock portfolios to maxima of simulated SV processes as an informal goodness-of-fit diagnostic. The purpose of this paper is to show that, for a broad class of Gaussian SV models, the suitably normalized extremes converge in distribution to the double exponential distribution.

2. Model and transformation. Consider the simple stochastic volatility model

$$(2.1) \qquad Y_t^* = \zeta \exp(\alpha_t/2)\xi_t, \qquad \alpha_t = \sum_{j=0}^{\infty} \theta_j Z_{t-j},$$

where $\{\xi_t\}$ is iid $N(0, 1)$, $\{Z_t\}$ is iid $N(0, \sigma_Z^2)$ independent of $\{\xi_t\}$, ζ is a positive constant, and $\sum_{j=0}^{\infty} \theta_j^2 < \infty$. The mean of α_t is taken to be zero and the variance of ξ_t is taken to be 1 without loss of generality, since any other values could be absorbed in the parameter ζ. Define

$$\sigma_\alpha^2 = \mathrm{Var}(\alpha_t) = \sigma_Z^2 \sum_{j=0}^{\infty} \theta_j^2$$

and let

$$Y_t = Y_t^*/\zeta.$$

In what follows, we work with the rescaled process $\{Y_t\}$ for simplicity.

The SV model in which $\{\alpha_t\}$ is a first-order autoregressive process, suggested by Taylor (1986), has been considered by many authors. See, for example, Jacquier, Polson and Rossi (1994) and the references therein. Other linear processes are also of interest, such as the long memory stochastic volatility models in Breidt, Crato and de Lima (1995).

Note that $\{Y_t\}$ is an identically distributed martingale difference sequence with finite variance, hence a white noise sequence in the sense of uncorrelated $(0, \sigma^2)$ random variables. The dynamics of the series (2.1) are in the conditional second moments, which can be seen after transformation to the stationary process

$$(2.2) \qquad X_t = \ln Y_t^2 = \alpha_t + \ln \xi_t^2.$$

The process $\{X_t\}$ is thus a Gaussian linear process plus an iid log-χ_1^2 noise [Wishart (1947)], with

$$\mathrm{E}[X_t] = \mathrm{E}[\ln \xi_t^2] = -(\text{Euler's constant}) - \ln 2$$

and

$$\mathrm{Cov}(X_t, X_{t+h}) = \sigma_\alpha^2 \rho(h) + \frac{\pi^2}{2} I_{\{h=0\}},$$

where $\rho(\cdot)$ denotes the autocorrelation function of $\{\alpha_t\}$ and $I_{\{h=0\}}$ is 1 if $h = 0$ and zero otherwise.

Since $\ln \xi_t^2$ is distributed as the log of a χ_1^2 random variable, the log of its moment-generating function is

$$\ln \mathrm{E}\left[\exp\left\{\lambda \ln \chi_1^2\right\}\right] = \lambda \ln 2 + \ln \Gamma(1/2 + \lambda) - \ln \Gamma(1/2)$$
$$= \lambda \ln \lambda + \lambda(\ln 2 - 1) + \ln 2/2 + O(1/\lambda).$$

The log of the moment-generating function of X_t is

$$\text{(2.3)} \quad \begin{aligned} \ln C(\lambda) &= \lambda^2 \sigma_\alpha^2/2 + \lambda \ln 2 + \ln \Gamma(1/2 + \lambda) - \ln \Gamma(1/2) \\ &= \lambda^2 \sigma_\alpha^2/2 + \lambda \ln \lambda + \lambda(\ln 2 - 1) + \ln 2/2 + O(1/\lambda), \end{aligned}$$

from which we can obtain the first two derivatives,

$$\text{(2.4)} \quad \begin{aligned} m(\lambda) &= \frac{d}{d\lambda} \ln C(\lambda) = \lambda \sigma_\alpha^2 + \ln 2 + \psi(1/2 + \lambda) \\ &= \lambda \sigma_\alpha^2 + \ln \lambda + \ln 2 + O(1/\lambda^2) \end{aligned}$$

and

$$\text{(2.5)} \quad \begin{aligned} S^2(\lambda) &= \frac{d^2}{d\lambda^2} \ln C(\lambda) = \sigma_\alpha^2 + \psi'(1/2 + \lambda) \\ &= \sigma_\alpha^2 + O(1/\lambda), \end{aligned}$$

where $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions, respectively [e.g., Abramowitz and Stegun (1965)].

Let

$$\text{(2.6)} \quad m^{-1}(x) = \frac{x}{\sigma_\alpha^2} - \frac{\ln x}{\sigma_\alpha^2} - \frac{k}{\sigma_\alpha^2} + \frac{\ln x}{\sigma_\alpha^2 x} + \frac{k}{\sigma_\alpha^2 x},$$

where

$$k = \ln(2/\sigma_\alpha^2),$$

and note that

$$\text{(2.7)} \quad \ln\{m^{-1}(x)\} = \ln\left(\frac{x}{\sigma_\alpha^2}\right) - \frac{\ln x}{x} - \frac{k}{x} + O\left(\frac{\ln x}{x^2}\right).$$

Substituting (2.6) into (2.4) and using (2.7), we obtain $m\{m^{-1}(x)\} = x + O(\ln x/x^2)$, justifying the inverse notation.

From (2.4), (2.5) and (2.6), we also have the following expansions:

(2.8)
$$\frac{\sigma_\alpha^2}{2}\{m^{-1}(x)\}^2 = \frac{x^2}{2\sigma_\alpha^2} - \frac{x\ln x}{\sigma_\alpha^2} - \frac{kx}{\sigma_\alpha^2} + \frac{k+1}{\sigma_\alpha^2}\ln x + \frac{\ln^2 x}{2\sigma_\alpha^2}$$
$$+ \frac{2k+k^2}{2\sigma_\alpha^2} + O\left(\frac{\ln^2 x}{x}\right),$$

(2.9)
$$m^{-1}(x)\ln\{m^{-1}(x)\} = \frac{x\ln x}{\sigma_\alpha^2} - \frac{x\ln\sigma_\alpha^2}{\sigma_\alpha^2} + \frac{\ln\sigma_\alpha^2 - 1 - k}{\sigma_\alpha^2}\ln x - \frac{\ln^2 x}{\sigma_\alpha^2}$$
$$+ \frac{k\ln\sigma_\alpha^2 - k}{\sigma_\alpha^2} + O\left(\frac{\ln^2 x}{x}\right),$$

(2.10)
$$\frac{1}{2}\ln S^2\{m^{-1}(x)\} = \frac{1}{2}\ln\sigma_\alpha^2 + O\left(\frac{1}{x}\right).$$

3. Extreme value results.

Tail behavior. Using a Tauberian argument as in Feigin and Yashchin (1983) [see also Davis and Resnick (1991)], the asymptotic approximation to the tail distribution of X_t can be expressed in terms of $m(\lambda)$ and $S(\lambda)$. This relationship is described in the following proposition.

PROPOSITION. *If F is the distribution of the random variable X_t defined in (2.2), then*

$$\bar{F}(x) = P[X_t > x]$$
$$\sim \frac{\exp\{-xm^{-1}(x)\}C\{m^{-1}(x)\}}{m^{-1}(x)S\{m^{-1}(x)\}(2\pi)^{1/2}}$$
$$= \frac{\sigma_\alpha}{\sqrt{\pi}}\exp\left\{-\frac{x^2}{2\sigma_\alpha^2} + \frac{x\ln x}{\sigma_\alpha^2} + \frac{(k-1)x}{\sigma_\alpha^2} - \frac{(k+\sigma_\alpha^2)\ln x}{\sigma_\alpha^2} - \frac{\ln^2 x}{2\sigma_\alpha^2}\right.$$
$$\left. -\frac{k^2}{2\sigma_\alpha^2} + O\left(\frac{\ln^2 x}{x}\right)\right\},$$

as $x \to \infty$, where C, S and m^{-1} are defined in (2.3), (2.5), and (2.6), respectively, and $k = \ln(2/\sigma_\alpha^2)$.

While the proof of this proposition is similar to the argument given on pages 38–39 of Feigin and Yashchin (1983), we supply the details of the argument in the present situation. The proof relies on the establishment of asymptotic normality for the normalized Esscher transform of F.

PROOF. Consider the family of probability density functions defined by

$$f_\lambda(x) = \lambda e^{\lambda x} \bar{F}(x)/C(\lambda), \qquad \lambda > 0.$$

It is a simple matter to check that the mean and variance associated with f_λ are $m(\lambda) - \lambda^{-1}$ and $S^2(\lambda) + \lambda^{-2}$, respectively. Now let g_λ be equal to f_λ after renormalizing the density to have approximate mean zero and variance 1; that is, define

$$
\begin{aligned}
g_\lambda(x) &= Sf_\lambda(Sx + m) \\
&= \lambda S e^{\lambda(Sx+m)} \bar{F}(Sx + m)/C(\lambda),
\end{aligned}
$$
(3.11)

where, for notational simplicity, the dependence of S and m on λ has been suppressed. The moment-generating function of g_λ is given by

$$
\begin{aligned}
\varphi_\lambda(t) &= \int_{-\infty}^{\infty} e^{tx} g_\lambda(x)\, dx \\
&= \int_{-\infty}^{\infty} \lambda S e^{(t+\lambda S)x + \lambda m} \bar{F}(Sx + m)\, dx/C(\lambda) \\
&= \lambda/(\lambda + t/S) e^{-tm/S} C(\lambda + t/S)/C(\lambda).
\end{aligned}
$$

Using properties (2.3)–(2.5), we have

$$
\begin{aligned}
\varphi_\lambda(t) ={}& \exp\Big\{ \ln \lambda - \ln\Big(\lambda + \frac{t}{S}\Big) - tm/S + \ln C\Big(\lambda + \frac{t}{S}\Big) - \ln C(\lambda) \Big\} \\
\sim{}& \exp\Big\{ -\frac{t}{S}(\lambda \sigma_\alpha^2 + \ln \lambda + \ln 2) + (\lambda + t/S)^2 \frac{\sigma_\alpha^2}{2} \\
& + (\lambda + t/S)\Big(\ln \lambda + \frac{t}{S\lambda}\Big) + (\lambda + t/S)(\ln 2 - 1) + \ln 2/2 \\
& - \lambda^2 \sigma_\alpha^2/2 - \lambda \ln \lambda - \lambda(\ln 2 - 1) - \ln 2/2 + O(1/\lambda) \Big\} \\
\sim{}& \exp\Big\{ \frac{t^2}{2} + O(1/\lambda) \Big\} \\
\to{}& \exp\Big\{ \frac{t^2}{2} \Big\}
\end{aligned}
$$

as $\lambda \to \infty$. It follows that

$$G_\lambda(x) \to_w \Phi(x),$$

where $G_\lambda(x) = \int_{-\infty}^x g_\lambda(y)\, dy$ and Φ is the standard normal cdf. It is a routine matter to check that $\varphi_\lambda(it)$ can be dominated by an integrable function (dominated by $K \exp\{-t^2/2\}$ for some sufficiently large constant K) and hence by

the dominated convergence theorem and the inversion formula for characteristic functions,

$$g_\lambda(x) = (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itx} \varphi_\lambda(it)\, dt$$

$$\to (2\pi)^{-1} \int_{-\infty}^{\infty} e^{-itx} \exp(-t^2/2)\, dt = (2\pi)^{-1/2} \exp(-x^2/2).$$

Choosing $x = 0$ through this limit, we have from (3.11) that

$$\bar{F}(m(\lambda)) \sim \frac{\exp\{-\lambda m(\lambda)\} C(\lambda)}{(2\pi)^{1/2} \lambda S(\lambda)},$$

which, upon making the substitution $\lambda \mapsto m^{-1}(\lambda)$ and invoking (2.6)–(2.10), establishes the proposition. $\square$

Convergence in distribution. We now have the following result.

THEOREM. *Let $\{X_t\}$ be given by (2.2), where $\{Y_t\}$ is a (standardized) stationary stochastic volatility process, with $\rho(h) \ln h \to 0$ as $h \to \infty$. Then*

$$M_n = \max\{X_1, \dots, X_n\}$$

has the Type I limiting distribution,

$$P[a_n(M_n - b_n) \le x] \to \exp(-e^{-x})$$

as $n \to \infty$, where

$$(3.12) \qquad a_n = \sigma_\alpha^{-1}(2\ln n)^{1/2} = (2/\sigma_\alpha^2)^{1/2}\, d_n$$

[where $d_n = (\ln n)^{1/2}$] and b_n, a solution of

$$\ln n + \ln \bar{F}(b_n) = o(1),$$

is given by

$$(3.13) \qquad b_n = c_1 d_n + c_2 \ln d_n + c_3 + c_4 \frac{\ln d_n}{d_n} + c_5 \frac{1}{d_n},$$

where

$$c_1 = (2\sigma_\alpha^2)^{1/2}, \qquad c_2 = 1, \qquad c_3 = \tfrac{3}{2}\ln 2 - \tfrac{1}{2}\ln \sigma_\alpha^2 - 1, \qquad c_4 = \frac{-\sigma_\alpha}{\sqrt{2}},$$

and

$$c_5 = \frac{-1}{2(2\sigma_\alpha^2)^{1/2}}\{1 + \sigma_\alpha^2 \ln(2\pi)\}.$$

PROOF. For x fixed, let $u_n = x/a_n + b_n$, so that from the proposition it can be shown that

$$(3.14) \qquad \bar{F}(u_n) = n^{-1}e^{-x} + o(n^{-1}),$$

and hence $F^n(u_n) \to \exp(-e^{-x})$. So it suffices to show that $|P[M_n \le u_n] - F^n(u_n)| \to 0$.

Conditioning on $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_n)'$, where $\eta_i = \ln \xi_i^2$, and using the Normal Comparison Lemma [see Leadbetter, Lindgren and Rootzén (1983), page 81], we have

$$|P[M_n \le u_n] - F^n(u_n)|$$

$$= \left| \mathrm{E}\left[P\left[\bigcap_{i=1}^n \{\alpha_i \le u_n - \eta_i\} \,|\, \boldsymbol{\eta} \right] \right] - \prod_{i=1}^n \mathrm{E}P[\alpha_i \le u_n - \eta_i \,|\, \eta_i] \right|$$

$$(3.15) \qquad \le K\mathrm{E} \sum_{1 \le i < j \le n} |\rho(i-j)| \exp\left\{ -\frac{(u_n - \eta_i)^2 + (u_n - \eta_j)^2}{2\sigma_\alpha^2(1 + |\rho(i-j)|)} \right\}$$

$$\le nK \sum_{i=1}^n |\rho(i)| \left(\mathrm{E}\exp\left\{ -\frac{(u_n - \eta_1)^2}{2\sigma_\alpha^2(1 + |\rho(i)|)} \right\} \right)^2.$$

(Throughout this argument, K will be used to represent a generic constant whose value may change from line to line.) Now let $w_n = 8\ln\ln n$ and note that $u_n - w_n \to \infty$ and

$$P[\eta_1 > w_n] = P[\xi_1^2 > (\ln n)^8] = P[|\xi_1| > (\ln n)^4] = o(n^{-1}).$$

The last equality follows from the asymptotic relation for the tail probability of the standard normal, $\bar{\Phi}(x) \sim \exp\{-x^2/2\}/(x\sqrt{2\pi})$ as $x \to \infty$, where Φ is the standard normal cdf. Using this relation, we have, for $\eta_1 \le w_n$ and n large,

$$\exp\left\{ -\frac{(u_n - \eta_1)^2}{2\sigma_\alpha^2(1 + |\rho(i)|)} \right\}$$

$$= (u_n - \eta_1)^{1/(1+|\rho(i)|)}(u_n - \eta_1)^{(-1)/(1+|\rho(i)|)} \exp\left\{ -\frac{(u_n - \eta_1)^2}{2\sigma_\alpha^2(1 + |\rho(i)|)} \right\}$$

$$\le K u_n^{1/(1+|\rho(i)|)} \bar{\Phi}^{1/(1+|\rho(i)|)}(\sigma_\alpha^{-1}(u_n - \eta_1)).$$

It follows that

$$\mathrm{E}\left[\exp\left\{ -\frac{(u_n - \eta_1)^2}{2\sigma_\alpha^2(1 + |\rho(i)|)} \right\} \right]$$

$$\le K u_n^{1/(1+|\rho(i)|)} \mathrm{E}\left[\bar{\Phi}^{1/(1+|\rho(i)|)}(\sigma_\alpha^{-1}(u_n - \eta_1)) I_{\{\eta_1 \le w_n\}} \right] + P[\eta_1 > w_n]$$

$$\le K(\ln n)^{1/[2(1+|\rho(i)|)]} \mathrm{E}\left[\bar{\Phi}^{1/(1+|\rho(i)|)}(\sigma_\alpha^{-1}(u_n - \eta_1)) \right] + o(n^{-1}),$$

which, by Jensen's inequality, is less than or equal to

$$K(\ln n)^{1/[2(1+|\rho(i)|)]}\big[\mathrm{E}\bar{\Phi}(\sigma_\alpha^{-1}(u_n-\eta_1))\big]^{1/(1+|\rho(i)|)} + o(n^{-1})$$

$$= K(\ln n)^{1/[2(1+|\rho(i)|)]}P^{1/(1+|\rho(i)|)}[\alpha_1+\eta_1>u_n] + o(n^{-1})$$

$$= K(\ln n)^{1/[2(1+|\rho(i)|)]}n^{(-1)/(1+|\rho(i)|)} + o(n^{-1}),$$

where the last equality is from (3.14). We conclude that

$$(3.16) \qquad \left[\mathrm{E}\exp\left\{-\frac{(u_n-\eta_1)^2}{2\sigma_\alpha^2(1+|\rho(i)|)}\right\}\right]^2 \le K(\ln n)^{1/(1+|\rho(i)|)}n^{(-2)/(1+|\rho(i)|)}.$$

The remainder of the proof follows the argument given on pages 86–87 of Leadbetter, Lindgren and Rootzén (1983). Set $\delta = \sup_{h\ge 1}|\rho(h)|$ and let β be a constant satisfying $0 < \beta < (1-\delta)/(1+\delta)$. Split the sum of the bound in (3.15) into the two parts, $i \le m$ and $i > m$, where $m = m_n = [n^\beta]$. Using (3.16), the sum over $i \le m$ is bounded by

$$(3.17) \qquad Knn^\beta(\ln n)n^{(-2)/(1+\delta)} = K(\ln n)n^{1+\beta-2/(1+\delta)} \to 0,$$

since $1+\beta-2/(1+\delta) < 0$ by the choice of β. Turning to the case $i > m$, let $\delta_n = \sup_{h\ge n}|\rho(h)|$ and note that $\delta_m(\ln n) \sim \beta^{-1}\delta_m \ln m \le \beta^{-1}\sup_{h\ge m}|\rho(h)|\ln h \to 0$. We then have, using (3.16) once again,

$$nK\sum_{i=m+1}^{n}|\rho(i)|\left(\mathrm{E}\exp\left\{-\frac{(u_n-\eta_1)^2}{2\sigma_\alpha^2(1+|\rho(i)|)}\right\}\right)^2 \le K\delta_m n^{2-2/(1+\delta_m)}\ln n$$

$$= K\exp\left\{\frac{2\delta_m}{1+\delta_m}\ln n\right\}\delta_m\ln n$$

$$\le K\exp\{2\delta_m\ln n\}\delta_m\ln n$$

$$\to 0.$$

This, combined with (3.17), completes the proof of the theorem. □

REMARKS.

1. The scale-normalizing sequence $\{a_n\}$ is the same as in the iid Gaussian case [see Leadbetter, Lindgren and Rootzén (1983), page 14]. The location normalizing sequence $\{b_n\}$ differs from the location-normalizing sequence for the iid Gaussian case, say $\{b_n^*\}$, by

$$b_n - b_n^* = c_2\ln d_n + c_3 + \left(\frac{\sigma_\alpha\ln 2}{2\sqrt{2}} - \frac{1}{2\sqrt{2}\sigma_\alpha}\right)\frac{1}{d_n},$$

so that $\{b_n\}$ is "slightly larger" than $\{b_n^*\}$.
2. The condition $\rho(h)\ln h \to 0$ as $h \to \infty$ is satisfied by both the short-memory ARMA stochastic volatility models considered by many authors and the long memory stochastic volatility models in Breidt, Crato and de Lima (1995).

3. The theorem shows that, for a wide class of stochastic volatility models, the extremes can be normalized independently of the covariance structure in $\{\alpha_t\}$, and the same limiting distribution is obtained in all cases. In finite samples, however, the degree of dependence in $\{\alpha_t\}$ does affect the goodness-of-fit of the limiting distribution, as can be seen by looking across the columns of Figure 1. The three columns correspond to the iid stochastic volatility model with

$$\alpha_t = Z_t, \qquad \{Z_t\} \text{ iid } N(0, \sigma_\alpha^2),$$

the autoregressive stochastic volatility model with

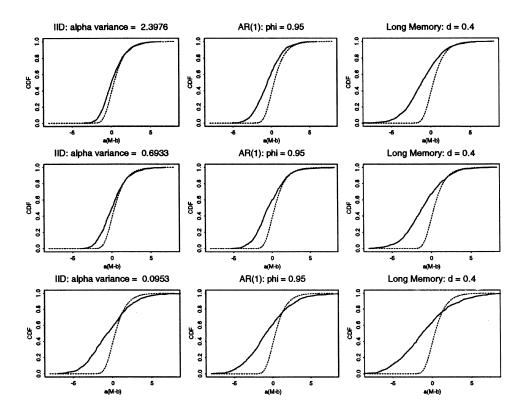$$\alpha_t = \phi\alpha_{t-1} + Z_t, \qquad \{Z_t\} \text{ iid } N(0, (1 - \phi^2)\sigma_\alpha^2),$$

FIG. 1.   *Empirical distribution functions (solid lines) for* 1000 *normalized extremes* $[a_n(M_n - b_n)$ *with* $n = 1000]$, *and limiting double exponential distribution function (dotted lines). In each row,* σ_α^2 *is constant. First column is the iid case. Second column is for the autoregressive stochastic volatility model, with parameter values taken from column two, Table* 4, *of Jacquier, Polson and Rossi* (1994). *Third column is for the long-memory stochastic volatility model in which* $\{\alpha_t\}$ *is fractionally integrated Gaussian noise.*

and the long-memory stochastic volatility model with

$$(1 - B)^d \alpha_t = Z_t, \qquad \{Z_t\} \text{ iid } N\left(0, \frac{\sigma_\alpha^2 \Gamma^2(1-d)}{\Gamma(1-2d)}\right)$$

[see Brockwell and Davis (1991), Section 13.2]. The parameters in the second column of Figure 1 are taken from column two, Table 4 of Jacquier, Polson and Rossi (1994) and have been chosen to represent a wide range of cases of practical interest in modeling returns on financial assets. The sample size $n = 1000$ is reasonable in this context, as many years of daily or weekly data are available in a typical application.

4. For very large n, the extremes of the process are effectively drawn from the upper tail of the Gaussian component of X_t, since the Gaussian tail dominates the $\ln \chi_1^2$ tail. In finite samples, however, the extent to which the Gaussian tail dominates the $\ln \chi_1^2$ tail is indexed by the parameter σ_α^2: as σ_α^2 decreases, the extremes increasingly come from the $\ln \chi_1^2$, the mass of the distribution shifts to the left, and the limiting distribution is an increasingly poor approximation to the finite sample distribution. This can be seen by looking down the rows of Figure 1.

Point process convergence. The proof of the theorem above can also be easily adapted to show that Leadbetter's conditions D and D', as modified by Adler (1978), hold for the stochastic volatility process $\{X_t\}$ defined in (2.2). These conditions are sufficient for weak convergence of the associated point processes

$$N_n(B) = \#\{j: (j/n, a_n(X_j - b_n)) \in B, j = 1, 2, \ldots\},$$

where B is any Borel subset of $(0, \infty) \times (-\infty, \infty)$, to a Poisson limit as $n \to \infty$. The intensity measure in the limit, ν, is defined on rectangles of the form $B = (a, b] \times (x, y]$ as $\nu(B) = (b - a)(e^{-x} - e^{-y})$. In other words, the limiting behavior of $\{N_n\}$ is the same for $\{X_t\}$ and for the independent sequence associated with $\{X_t\}$ (i.e., the sequence $\{\hat{X}_t\}$ which is defined to be iid with $\hat{X}_1 =_d X_1$).

A host of ancillary results ensue from the point process convergence. For example, the joint limiting distribution of any collection of extreme order statistics can be determined and the convergence of the sample extremal process can be established [see Adler (1978) and Resnick (1975)]. These results have a number of standard applications in extreme value theory which are described in detail in Resnick [(1987), Chapter 4].

One other application is for the limiting distribution of the normalized extremes of the untransformed stochastic volatility process, $\{Y_t^*\}$. The argument follows that of de Haan et al. (1989).

For $x \in \mathbb{R}$ fixed, let $u_n = a_n^{-1}x + b_n$, $v_n^* = \exp(u_n/2)\zeta$ and $v_n = v_n^*/\zeta$. Note that $Y_t^* = |Y_t^*|B_t$, where $\{B_t\} = \{\text{sign}(Y_t^*)\}$ is an iid sequence uniformly distributed on $\{-1, 1\}$ and independent of $\{|Y_t^*|\}$. Let $N_n = N_n\{(0, 1] \times (x, \infty)\}$

and let $1 \le \tau_1 < \tau_2 < \cdots$ be the times at which $\{Y_t^2\}$ exceeds v_n^2. Then

$$\{N_n = k\} = \{\ln Y_{\tau_1}^2 > u_n, \ldots, \ln Y_{\tau_k}^2 > u_n, \tau_k \le n, \tau_{k+1} > n\}$$

$$= \{Y_{\tau_1}^2 > v_n^2, \ldots, Y_{\tau_k}^2 > v_n^2, \tau_k \le n, \tau_{k+1} > n\}$$

$$\subset \{Y_{\tau_1}^2 > v_n^2, \ldots, Y_{\tau_k}^2 > v_n^2\},$$

so

(3.18) $$\{N_n = k\} = \{N_n = k, Y_{\tau_1}^2 > v_n^2, \ldots, Y_{\tau_k}^2 > v_n^2\}.$$

Define $M_n = \max\{Y_1, \ldots, Y_n\}$. Then

$$P[\max\{Y_1^*, \ldots, Y_n^*\} \le v_n^*]$$

$$= P[M_n \le v_n]$$

(3.19)

$$= \sum_{k=0}^{\infty} P[N_n = k, Y_{\tau_1}^2 > v_n^2, \ldots, Y_{\tau_k}^2 > v_n^2, M_n \le v_n]$$

$$= \sum_{k=0}^{\infty} P[N_n = k, Y_{\tau_1} \le 0, \ldots, Y_{\tau_k} \le 0,$$

$$Y_{\tau_1}^2 > v_n^2, \ldots, Y_{\tau_k}^2 > v_n^2, M_n \le v_n],$$

because $Y_{\tau_i}^2 > v_n^2$ implies $Y_{\tau_i} > v_n$ if $Y_{\tau_i} > 0$.

Using (3.18), (3.19) becomes

$$\sum_{k=0}^{\infty} P[N_n = k, Y_{\tau_1} \le 0, \ldots, Y_{\tau_k} \le 0]$$

(3.20)

$$= \sum_{k=0}^{\infty} P[N_n = k, B_{\tau_1} = -1, \ldots, B_{\tau_k} = -1]$$

$$= \sum_{k=0}^{\infty} P[N_n = k] 2^{-k},$$

since N_n is independent of the signs of the Y_t. Also, the random times τ_i depend only on $|Y_{\tau_i}|$ and are independent of $\{B_{\tau_i}\}$. Hence, by dominated convergence as $n \to \infty$, (3.20) becomes

$$\sum_{k=0}^{\infty} P[N\{(0, 1] \times (x, \infty)\} = k] 2^{-k} = \sum_{k=0}^{\infty} \frac{(e^{-x}/2)^k \exp\{-e^{-x}\}}{k!}$$

$$= \exp\{-e^{-x}/2\}.$$

REFERENCES

ABRAMOWITZ, M. and STEGUN, I. A. (1965). *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. National Bureau of Standards, Washington, D.C.

ADLER, R. J. (1978). Weak convergence results for extremal processes generated by dependent random variables. *Ann. Probab.* **6** 660–667.

BOLLERSLEV, T. (1986). Generalized autoregressive conditional heteroskedasticity. *J. Econometrics* **31** 307–327.

BOLLERSLEV, T., CHOU, R. Y. and KRONER, K. F. (1992). ARCH modeling in finance. *J. Econometrics* **52** 5–59.

BREIDT, F. J., CRATO, N. and DE LIMA, P. (1995). On the detection and estimation of long memory in stochastic volatility. *J. Econometrics*. To appear.

BROCKWELL, P. J. and DAVIS, R. A. (1991). *Time Series: Theory and Methods,* 2nd ed. Springer, New York.

CLARK, P. K. (1973). A subordinated stochastic process model with finite variances for speculative prices. *Econometrica* **41** 135–156.

DAVIS, R. A. and RESNICK, S. I. (1991). Extremes of moving averages of random variables with finite endpoint. *Ann. Probab.* **19** 312–328.

DE HAAN, L., RESNICK, S. I., ROOTZÉN, H. and DE VRIES, C. G. (1989). Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes. *Stochastic Process. Appl.* **32** 213–224.

ENGLE, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* **50** 987–1007.

FEIGIN, P. D. and YASHCHIN, E. (1983). On a strong Tauberian result. *Z. Wahrsch. Verw. Gebiete* **65** 35–48.

GEWEKE, J. (1994). Comment on "Bayesian analysis of stochastic volatility models" by E. Jacquier, N. G. Polson and P. E. Rossi. *J. Bus. Econom. Statist.* **12** 397–399.

HULL, J. and WHITE, A. (1987). The pricing of options on assets with stochastic volatilities. *J. Finance* **42** 281–300.

JACQUIER, E., POLSON, N. G. and ROSSI, P. E. (1994). Bayesian analysis of stochastic volatility models (with discussion). *J. Bus. Econom. Statist.* **12** 371–417.

LEADBETTER, M. R., LINDGREN, G. and ROOTZÉN, H. (1983). *Extremes and Related Properties of Random Sequences and Processes.* Springer, New York.

MANDELBROT, B. (1963). The variation of certain speculative prices. *J. Business* **36** 394–419.

MELINO, A. and TURNBULL, S. M. (1990). Pricing foreign currency options with stochastic volatility. *J. Econometrics* **45** 239–265.

RESNICK, S. (1975). Weak convergence to extremal processes. *Ann. Probab.* **3** 951–960.

RESNICK, S. (1987). *Extreme Values, Regular Variation and Point Processes*. Springer, New York.

TAUCHEN, G. and PITTS, M. (1983). The price variability–volume relationship on speculative markets. *Econometrica* **51** 485–505.

TAYLOR, S. (1986). *Modelling Financial Time Series*. Wiley, New York.

WISHART, J. (1947). The cumulants of the z and of the logarithmic χ^2 and t distributions. *Biometrika* **34** 170–178.

DEPARTMENT OF STATISTICS
IOWA STATE UNIVERSITY
AMES, IOWA 50011–1210
E-MAIL: jbreidt@iastate.edu

DEPARTMENT OF STATISTICS
COLORADO STATE UNIVERSITY
FT. COLLINS, COLORADO 80523–1877
E-MAIL: rdavis@stat.colostate.edu