

# A note on the relationship between GARCH and symmetric stable processes

Patrick A. Groenendijk, André Lucas, Casper G. de Vries <sup>\*,1</sup>

*Tinbergen Institute and Erasmus University Rotterdam, Rotterdam, The Netherlands*

---

## Abstract

This note provides some explanations and extensions for the interesting results in Ghose and Kroner (1995). Specifically, we address the following points: (1) It is shown that the stable distribution and the stationary ARCH distributions are partially nested with respect to their tail shapes; (2) A novel interpretation of the McCulloch estimator is developed from the vantage point of extreme value theory; (3) This interpretation not only explains the apparent bias in some of the reported estimates, but it also helps in remedying the problem. Taken together, all three points reinforce the main conclusion of Ghose and Kroner.

*JEL classification:* C14; C22

*Keywords:* GARCH; Stable distributions; Tail index estimators

---

## 1. Introduction

The intriguing contribution by Ghose and Kroner (1995) in this issue compares the non-normal stable distributions to the class of GARCH stochastic processes as alternative models for speculative return data. Both models share the fact that the unconditional distribution has fat tails and that the tail shape is invariant under addition. The differences between the models are as follows: (i) The stable model is an independently and identically distributed (i.i.d.) process and hence unable to display the volatility clusters which are typical for the GARCH process. (ii) A

---

<sup>\*</sup> Tinbergen Institute, Oostmaaslaan 950-952, 3063 DM Rotterdam, The Netherlands.

<sup>1</sup> Thanks to Henk Hoek for computational advice.

convolution of stable random variables with the same characteristic exponent is again stable, but GARCH variates are not self similar. Notice the difference between ‘tail additivity’, mentioned above, and the self similarity of the entire law. (iii) While both models are heavy tailed, the GARCH models allow for bounded second and higher order moments, but the non-normal stable laws necessarily display infinite variance. The authors use the differences (i) and (iii) to show that the GARCH or IGARCH model provides a better description of speculative return data.

On a first reading the authors’ conclusion may appear unsurprising given the prevalence of volatility clustering in financial data and the inability of an i.i.d. process to display such dependency. Ghose and Kroner rightly do not make too much out of the difference (i), however, because the stable model is often viewed as a model for the unconditional distribution. This leaves the difference (iii) as the main discriminatory device. The focus then is on the tail shape of the unconditional distribution of the GARCH variates vis-à-vis the tail shape of non-normal stable laws.<sup>2</sup> The larger part of the Ghose and Kroner paper is devoted to this issue. Their analysis involves a number of interesting intricacies which may not be apparent to every reader. Presumably with this in mind, the editors have asked us to write a note on the paper.

Because the analysis is competent and well balanced, the nature of our note is merely to provide some clarifications and extensions on some of the points made by the authors. This may help the reader to better judge the significance of the Ghose and Kroner contribution. Specifically, we address the following three issues:

- (1) Even though the tail shapes of the GARCH class processes and the stable laws may differ, and hence can be used to discriminate between the competing models, this is not necessarily the case. In fact, the two models partially overlap.
- (2) In order to develop a better understanding of the McCulloch (1986) procedure for estimating the characteristic exponent  $\alpha$  from a stable distribution, we provide a novel interpretation related to the tail shape of fat tailed distributions in general. This interpretation explains the spectacular failure of the  $\hat{\alpha}$  method under certain alternative distributional assumptions.
- (3) The new view of the  $\hat{\alpha}$  estimator is also used to clarify the puzzling 1.46 point estimates by the  $\hat{\alpha}$  procedure as reported in Fig. 1 and Tables 4 and 5 from Ghose and Kroner. The novel interpretation also shows how to fix this downward bias problem. This interpretation lends further credence to the main conclusion of the paper under review.

We conclude by listing some open problems.

---

<sup>2</sup> The tail shape of the conditional distribution of the GARCH process is not the critical issue. It can be either thin tailed or fat tailed.

## 2. The tail index of the ARCH process when $\beta_0 > 1$

The major discriminatory device exploited by Ghose and Kroner is the difference in the number of bounded (raw) moments that exist under the two alternative models for the unconditional distribution of speculative return series. The authors correctly argue that the ARCH-cum-normal-innovations class allows for any maximal finite moment of the stationary distribution in the interval  $[2, \infty)$  by choosing  $\beta_0$  from  $[0,1]$  appropriately (the lower  $\beta_0$ , the higher the number of finite moments). In contrast, the non-normal stable distributions necessarily have the maximal finite moment in the interval  $(0,2)$ . The authors then set out to measure the maximal finite moment as a way of choosing between the alternative models. What is left aside, though, is that the apparent non-nestedness is only a one-way street. We show that the two models partially overlap.

To do this we have to explain how the maximal finite moment is obtained. Consider the Pareto distribution

$$F(x) = 1 - s^\alpha x^{-\alpha}, \quad x \in [s, \infty). \quad (1)$$

It is easily shown that the exponent  $\alpha > 0$  determines the maximal finite moment, i.e.  $\mathcal{E}[x^k]$  is bounded if  $k < \alpha$  and unbounded if  $k \geq \alpha$ . The unboundedness of the higher moments is caused by the power decline  $x^{-\alpha}$  of the tail, rather than the exponential decline as in case of e.g. the normal distribution. Therefore the Pareto law is said to be fat tailed, while the normal is thin tailed. The following condition, called regular variation at infinity, is necessary and sufficient for a distribution to be fat tailed:  $\lim_{t \rightarrow \infty} (1 - F(tx))/(1 - F(t)) = x^{-\zeta}, \zeta > 0$ . The condition characterizes the tail behaviour, i.e. it states that  $F(x)$  eventually displays the same behaviour in the tails as the Pareto model. The exponent  $\zeta$  is called the tail index and is such that  $\mathcal{E}(|x|^\eta) < \infty$  for all  $0 \leq \eta < \zeta$ . The Pareto model (1) trivially satisfies the regular variation condition. For the stable model the regular variation property can be verified by using the Bergström-Feller expansions, and in this case the  $\zeta$  equals the characteristic exponent  $\alpha$ .

The verification of the regular variation property of the unconditional distribution of ARCH variates is, however, not straightforward, because no closed form expression for this distribution is known. A different approach based on the theory of stochastic difference equations is needed. To this end we rewrite the ARCH-cum-normal-innovations process from Ghose and Kroner (1995, Eqs. (3) and (4)) with  $u_t \sim N(0,1)$  and  $\beta_1 = 0$ . Taking squares of Eq. (3) and combining with Eq. (4) from Ghose and Kroner yields

$$\epsilon_t^2 = \omega u_t^2 + \beta_0 u_t^2 \epsilon_{t-1}^2 = B_t + A_t \epsilon_{t-1}^2, \quad (2)$$

say. This stochastic difference equation with i.i.d. pairs  $(B_t, A_t)$  of chi-squared random variables is equivalent with the ARCH process up to a coin flip process

for the sign. As is discussed in de Haan et al. (1989), we have from Kesten (1973, Theorem 5) that under the conditions stated below:

$$\epsilon_t^2 \xrightarrow{d} \epsilon_x^2 = \sum_{j=1}^{\infty} B_j \prod_{i=1}^{j-1} A_i,$$

i.e. Eq. (2) has a solution  $\epsilon_x$  which is unique in distribution. Moreover,  $\epsilon_x$  has a Pareto like tail

$$\lim_{x \rightarrow \infty} x^\zeta P\{\epsilon_x > x\} = c, \tag{3}$$

where  $c > 0$  is some constant and  $\zeta > 0$ .

The conditions under which these results hold are as follows:

$$\mathcal{E}[\log A_t] < 0 \Leftrightarrow \log 2\beta_0 + \psi\left(\frac{1}{2}\right) < 0; \tag{4}$$

there exists a  $\kappa > 0$  such that

$$\mathcal{E}[A_t^\kappa] = 1 \Leftrightarrow \Gamma\left(\kappa + \frac{1}{2}\right) = \sqrt{\pi} (2\beta_0)^{-\kappa}; \tag{5}$$

$$\mathcal{E}[A_t^\kappa \log^+ A_t] < \infty; \tag{6}$$

$$0 < \mathcal{E}[B_t^\kappa] < \infty; \tag{7}$$

$$B_t/(1 - A_t) \text{ is nondegenerate and the conditional distribution of } \log A_t \text{ given } A_t \neq 0 \text{ is nonlattice.} \tag{8}$$

We discuss conditions (4) and (5); (6), (7) and the first part of (8) are easily checked using the fact that  $A_t$  and  $B_t$  are  $\chi^2$  distributed. From this it also follows that the distribution of  $\log A_t$ , given  $A_t \neq 0$ , has no atoms, thus excluding the possibility of a lattice distribution [see Feller (1971, p. 138)]. Condition (4) also appears in Vervaat (1979) and Nelson (1990), and guarantees stationarity. For the process at hand, (4) is equivalent with  $\beta_0 < \bar{\beta} \approx 3.56856$ , where  $\bar{\beta} = 1/2 \exp^{\psi(1/2)}$ , and where  $\psi(\cdot)$  is the Euler psi function. Also recall the fact that for  $\beta_0 < 1$ ,  $\epsilon_x$  has finite variance. Thus we can distinguish four regions. The ARCH process is i.i.d.  $N(0, \omega)$  if  $\beta_0 = 0$ , is covariance stationary for  $0 < \beta_0 < 1$ , is stationary when  $1 \leq \beta_0 < \bar{\beta}$ , and nonstationary if  $\beta_0 > \bar{\beta}$  [for the borderline case  $\beta_0 = \bar{\beta}$ , see Vervaat (1979)].

Condition (5) yields the tail index as a function of  $\beta_0$ . The tail index of the ARCH process  $\zeta$  is related to the tail index  $\kappa$  of the squared process as follows:  $\zeta = 2\kappa$ , where the factor 2 stems from squaring the ARCH process [see de Haan et al. (1989) for further details]. The dashed curve in Fig. 1 plots  $\zeta$  as a function of  $\beta_0$  ( $\beta_0$  varies with a step size of 0.05). Note that  $\zeta = 2$  at  $\beta_0 = 1$ ,  $\zeta = \infty$  at  $\beta_0 = 0$ , and at the upper limit  $\bar{\beta} = 2 \exp(\gamma) \approx 3.56856$  (where  $\gamma$  is Euler's constant),

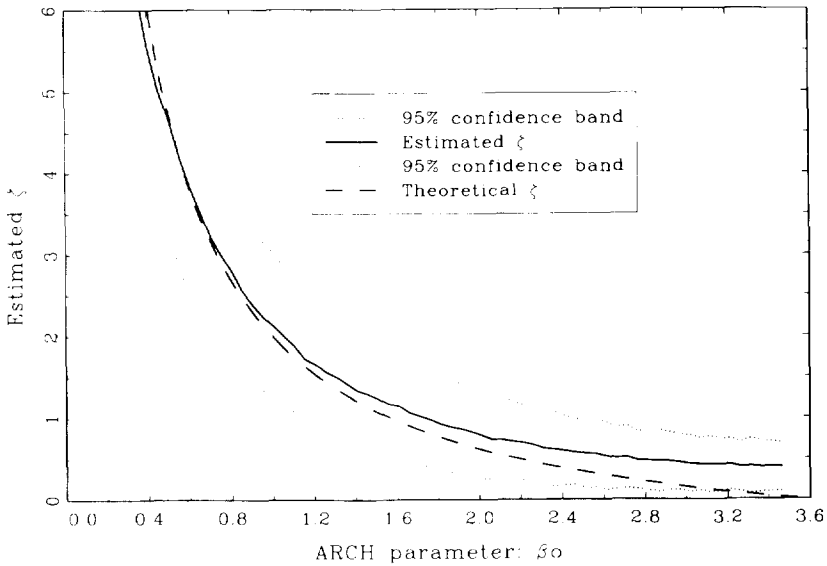


Fig. 1. Tail index estimator for ARCH.

$\zeta = 0.00279$ . It follows that the stable model and the stationary ARCH distribution exhibit the same tail behaviour in the sense of the tail index  $\zeta$  when  $\bar{\beta} > \beta_0 > 1$  and hence it is possible that

$$0.00279 < \zeta = 2\kappa = \alpha < 2. \tag{9}$$

The interesting result is that the two models are only partially nested with respect to their tail behaviour: At  $\beta_0 = 0$  and  $\alpha = 2$  the two models have normal tails; for  $0 < \beta_0 < 1$ , then  $\zeta > 2$  and there is no counterpart in the stable class; for  $\bar{\beta} > \beta_0 > 1$  the tail indices can be equal; for  $\alpha < 0.00279$  there is no counterpart in the ARCH class. It follows that the tail method as a discriminatory device only works in the regions where the two models do not overlap.

The above arguments can be extended to the GARCH(1,1) case. This can be done by viewing the variance Eq. (4) from Ghose and Kroner as a stochastic difference equation in  $\sigma_t^2$ . It can be shown that this difference equation satisfies conditions (4)–(8) above. Hence, again we can distinguish a region  $1 < \beta_0 + \beta_1 < f(\beta_0, \beta_1)$ , with  $f(\cdot, \cdot)$  a real-valued function, where the tail index cannot be used as a discriminatory device.<sup>3</sup> This point is mainly theoretical, although empirical examples of  $\beta_0 + \beta_1 > 1$  are found in, e.g., Engle et al. (1987), Hong (1988), and Pagan (1993).

<sup>3</sup> See Nelson (1990, Fig. 1) for the form of the function  $f(\beta_0, \beta_1)$ .

Interestingly, Ghose and Kroner and others have found that  $\zeta > 2$ . The way in which this conclusion is obtained is by means of the semiparametric Hill estimator  $\hat{\zeta}$  [Eq. (8) in Ghose and Kroner (1995)]. The motivation behind the Hill estimator is easy to obtain from the Pareto model. The log density from (1) reads

$$\log \alpha + \alpha \log s - (\alpha + 1) \log x.$$

Differentiating this with respect to  $\alpha$  yields  $1/\alpha - \log(x/s)$ . Summing over different observations and equating to zero then yields the Hill estimator  $\hat{\zeta}$ . When applied to other distributions that only exhibit Pareto like behaviour in the tails, one takes  $s$  sufficiently high (low) and employs only the highest (lowest) order statistics. In Fig. 2 of Ghose and Kroner  $\hat{\zeta}$  is plotted as a function of  $\beta_0 \leq 1$ . In our Fig. 1, and on basis of the analysis above, this plot is extended to the interval  $\bar{\beta} > \beta_0 \geq 1$ , see the solid curve. The estimator performs well in the region  $2.5 \geq \beta_0 > 0.5$  say, but is downward biased for low  $\beta_0$  and upward biased for high  $\beta_0$ . The same phenomenon is present in the right-hand side of Ghose and Kroner's Fig. 3. Note that the figures are based on using a fixed percentage (1%) of the extreme order statistics. What the plots indicate is that one should go deeper into the tails for low  $\beta_0$  (close to normality the percentage of extremes decreases), and vice versa for high  $\beta_0$ .

### 3. Estimation of the characteristic exponent

The characteristic function of the symmetric stable distributions is  $e^{-|t|^\alpha}$ ,  $0 < \alpha \leq 2$ , and where  $\alpha$  is called the characteristic exponent. When  $\alpha = 2$ , then the distribution is normal and has all moments bounded; if  $\alpha < 2$ , then all moments greater than  $\alpha$  are unbounded. As we noted in the previous section, the stable laws with  $\alpha < 2$  are regularly varying at infinity, i.e. exhibit Pareto like tails.<sup>4</sup> The tail index equals  $\zeta = \alpha$ . The characteristic exponent can thus be estimated by the Hill estimator that was discussed above. Alternatively, one can employ McCulloch's (1986) procedure for estimating  $\alpha$ , see Ghose and Kroner Eqs. (1a,b). Unlike the Hill estimator, this latter procedure confines the estimator to the interval (0,2]. In addition, this procedure is only valid if the observations come from a stable law and not from any other distribution with  $\zeta < 2$ . This is vividly illustrated in Fig. 1 in Ghose and Kroner. One clearly sees that the estimator  $\hat{\alpha}$  as an estimator for  $\zeta$  from the IGARCH(1,1) process is downward biased to approximately 1.46. Note that the Hill estimator in Fig. 3, instead, correctly locates  $\zeta$  at 2 (except when the normal distribution becomes a local alternative close to  $\beta_0 = 0$ ). Also note that a very similar discrepancy between the Hill  $\zeta$  estimator and the McCulloch  $\alpha$  estimator arises from the real data (Tables 4

<sup>4</sup> In economics, the terminology 'stable Pareto distribution' is often used.

Table 1  
The  $\alpha$  and  $\gamma$  estimators for alternative distributions

$\alpha$	Symmetric stable distribution		Symmetric extremal type distribution		Student-t distribution		Symmetric Pareto distribution	
	$Q = v_\alpha$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$
0.50	44.636	0.424	43.281	0.427	26.475	0.491	33	0.460
0.60	23.612	0.509	23.098	0.513	16.016	0.580	20.883	0.530
1.00	6.314	0.873	6.579	0.854	6.314	0.873	9	0.732
1.50	3.150	1.403	3.511	1.281	4.246	1.113	6.199	0.882
1.90	2.513	1.747	2.695	1.623	3.670	1.238	5.360	0.959
2.00	2.439	1.805	2.565	1.709	3.576	1.263	5.220	0.974
3.00	-	-	1.874	2.563	3.077	1.432	4.441	1.079
5.00	-	-	1.458	4.272	2.773	1.578	3.933	1.175
30.00	-	-	1.065	25.630	2.486	1.767	3.413	1.311

Notes: The theoretical value of  $\gamma$  is for each distribution the  $\alpha$  value in the first column. The  $Q = x_{0.95}/x_{0.75}$ , i.e. the 95% to 75% quantile ratio; and  $\hat{\gamma}$  is the theoretical de Haan and Resnick estimate using  $Q$ .

and 5 in Ghose and Kroner). We first obtain the intuition behind the McCulloch  $\alpha$  estimator, and then, in the next section, we turn to explaining the apparent downward bias in Fig. 1.

The McCulloch estimator  $\hat{\alpha}$  for symmetric distributions is based on computing the statistic  $v_\alpha = x_{0.95}/x_{0.75}$ , i.e. the ratio of the 95% and 75% quantiles respectively (cf. Eqs. (1a) and (1b) in Ghose and Kroner, which also cover the asymmetric cases). The conventional motivation behind this estimator is the fact that  $x_{0.75} \approx 1.00$  when  $0.50 < \alpha \leq 2.00$ , while  $x_{0.95}$  increases monotonically when  $\alpha$  decreases from 2.00 to 0.50.<sup>5</sup> Thus  $v_\alpha$  also increases monotonically when  $\alpha$  decreases. A simple conversion table in McCulloch (1986) then maps  $v_\alpha$  into  $\alpha$ . This table is partly reproduced in columns 1 and 2 of Table 1.

The estimator  $\hat{\alpha}$  works well if one is ensured that the observations come from a stable distribution, or that a version of the generalized central limit theorem applies such that appropriately normed sums are in the domain of attraction of a stable law. But the procedure can fail spectacularly if these conditions are not met. We provide three examples.

The first example is based on the following symmetric extremal type distribution.

$$F(x) = \begin{cases} \frac{1}{2} - \frac{1}{2} \exp\left[-(-a/x)^\alpha\right], & x \leq 0 \\ \frac{1}{2} + \frac{1}{2} \exp\left[-(a/x)^\alpha\right], & x \geq 0 \end{cases} \quad (10)$$

and where  $a > 0$ ,  $\alpha > 0$ . Let  $p = F(x_p)$  when  $x \geq 0$  and write  $q = 2p - 1$ ; here

<sup>5</sup> The constancy of the 75% quantile also motivated the earlier Fama and Roll (1971) estimator.

$x_p$  is the  $p$ -th quantile. It is easy to show that  $Q = x_{0.95}/x_{0.75} = [\log(0.50)/\log(0.90)]^{1/\alpha} = [6.579]^{1/\alpha}$ .<sup>6</sup> Comparing these  $Q$ -values for the range  $0 < \alpha < 2$ , see column 4 in Table 1, with the  $\nu_\alpha$  values for the stable model, column 2 of Table 1, it is clear that McCulloch’s procedure performs well for this alternative distribution.

Quite different in this respect is the second example which is based on the Student-t distribution, see column 6 of Table 1, where  $\alpha$  equals the degrees of freedom. The stable and Student-t model coincide at  $\alpha = 1$ , i.e. the Cauchy distribution, and  $\alpha = \infty$ , the normal case. Therefore the quantile ratio of the Student model ranges between [6.314, 2.439] as  $\alpha$  increases from 1 to infinity, whereas the stable model covers the same range when  $\alpha$  varies from 1 to 2. Hence, if the McCulloch estimator is employed for estimating the Student degrees of freedom in the range  $(2, \infty)$ , these are heavily downward biased to the range  $(1, 2)$ .

Even more dramatic is the third example that is based on the ‘symmetric Pareto’ law:

$$F(x) = \begin{cases} \frac{1}{2}[1 - x]^{-\alpha}, & x \leq 0 \\ 1 - \frac{1}{2}[1 + x]^{-\alpha}, & x \geq 0 \end{cases} \tag{11}$$

and where  $\alpha > 0$ . Using the McCulloch  $\hat{\alpha}$ , one never gets  $\hat{\alpha} > 1.40$ ; and the bias is reversed for  $\alpha$  around 0.50, see Table 1, column 8.

To get a better understanding of the McCulloch procedure and its applicability to non-stable distributions, we provide a novel interpretation of the estimator. For this purpose, we first discuss an alternative tail index and characteristic exponent estimator developed by de Haan et al. (1980). A heuristic derivation is as follows. Consider again the Pareto law (1). By the Glivenko-Cantelli theorem we may replace the left hand side probability  $F(x)$  by its empirical counterpart  $(n - m)/n$ , where  $m$  refers to the  $m$ -th descending order statistic,  $X_{(1)} \geq \dots \geq X_{(m)} \geq \dots \geq X_{(n)}$ , in a sample from  $F(X)$  of size  $n$ . This yields  $m/n = s^\alpha X_{(m)}^{-\alpha}$ , and in logarithms  $\log m - \log n = \alpha \log s - \alpha \log X_{(m)}$ . Similarly, for the  $k$ -th order statistic, where  $1 \leq k < m$ , we have  $\log(k/n) = \alpha \log s - \alpha \log X_{(k)}$ . Subtracting this latter equation from the former and rearranging yields the estimator  $\hat{\gamma}$  of  $1/\alpha$

$$\hat{\gamma} = \frac{\log X_{(k)}/X_{(m)}}{\log m/k} \tag{12}$$

De Haan and Resnick take  $k = 1$ , but this is not necessary.

How is this estimator related to the  $\hat{\alpha}$  procedure? The McCulloch procedure relies on computing the statistic  $\nu_\alpha$ , which in case of symmetry can be achieved by

<sup>6</sup> If  $a = -\log q$ ,  $x_p$  is constant, i.e. independent from  $\alpha$ . This implies  $p = [1 + \exp(a)]/2 \exp(a)$ . But this means that  $p$  can be anywhere in the interval  $(1/2, 1)$  by choosing  $a$  appropriately from  $(0, \infty)$ .



Table 2  
 McCulloch ( $\hat{\alpha}$ ), de Haan and Resnick ( $(1/\hat{\gamma})$ ), and Hill ( $\hat{\xi}$ ) estimator of the tail index for ARCH(1)

$\beta_0$	$\zeta$	$Q = x_{0.95} / x_{0.75}$	$\hat{\alpha}$	$1/\hat{\gamma}$	$\hat{\xi}$	$Q = x_{0.995} / x_{0.975}$	$\hat{\alpha}$	$1/\hat{\gamma}$	$\hat{\xi}$
0.000	$\infty$	2.438	1.984	1.807	4.718	1.314	1.943	5.924	7.694
0.476	5.000	2.600	1.811	1.686	3.468	1.519	1.865	3.876	4.276
0.900	2.304	3.070	1.532	1.436	2.131	2.069	1.685	2.244	2.334
1.000	2.000	3.266	1.462	1.362	1.892	2.290	1.611	1.976	2.039
1.038	1.900	3.355	1.435	1.331	1.816	2.388	1.581	1.884	1.945
2.000	0.620	14.660	0.715	0.606	0.640	18.924	0.645	0.652	0.715

Notes: The entries are the Monte-Carlo means of the various estimates over 1,000 iterations using series of length 10,000. The  $\hat{\alpha}$  estimate in column 8 was computed by constructing an inversion table from  $\hat{\nu}_\alpha$  to  $\alpha$ , as in McCulloch (1986).  $\hat{\nu}_\alpha$  was taken to be  $(x_{0.995} - x_{0.005}) / (x_{0.975} - x_{0.025})$ .  $\beta_0$  is the parameter of the ARCH process, and  $\zeta$  is the corresponding theoretical tail index, computed from (5).

taking  $X_{(k)}/X_{(m)}$ , with  $k$  equal to the integer closest to  $0.05 n$  and  $m$  equal to the integer closest to  $0.25 n$  in a sample of size  $n$ . Hence, the map of  $\nu_\alpha$  into  $\alpha$ , by going from column 2 to 1 in Table 1, can be represented analytically by Eq. (12). Column 3 of Table 1 shows that this representation of the conversion map is fairly decent, albeit some downward bias is present. The  $\hat{\gamma}$  statistic, however, also explains the failure of the  $\hat{\alpha}$ -method under non-stable alternatives. Assuming either a Student-t or a symmetric Pareto with  $\alpha \in (2, \infty)$ , one sees from columns 7 and 9 in Table 1 that the  $1/\hat{\gamma}$  estimates are approximately the same as the  $\hat{\alpha}$  estimates from the McCulloch procedure. Apparently the conversion procedure fails because  $1/\hat{\gamma}$  does not capture  $\alpha$  correctly for non-stable distributions. The next section shows how this failure can be partially repaired.<sup>7</sup>

#### 4. An explanation for $\hat{\alpha} = 1.46$ under IARCH

Return to the puzzling  $\hat{\alpha} = 1.46$  estimates for the IGARCH(1,1) model reported in Ghose and Kroner's Fig. 1, and the empirical estimates in their Tables 4 and 5 of about the same value. In Table 2 column 4, we show that this observation is also obtained under the IARCH model. This phenomenon is now readily explained by analogy. Return to Table 1 and consider the  $Q$ -values obtained under the Student-t and symmetric Pareto distribution when  $\alpha = 2$ . Using the  $\hat{\alpha}$  procedure, i.e. using the conversion from column 2 to 1, yields  $\alpha$  estimates of say 1.38 in case of the Student-t and 1.10 for the symmetric Pareto law. Only in case of the

<sup>7</sup> One may wonder how  $\hat{\gamma}$  is related to  $\hat{\xi}$ . Loosely speaking one may interpret  $\hat{\xi}$  as a weighted average of  $\hat{\gamma}$  evaluated at different  $k$ .

Table 3  
Bias elimination

$\alpha$	Symmetric stable distribution		Symmetric extremal type distribution		Student-t distribution		Symmetric Pareto distribution	
	$Q = v_\alpha$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$	$Q$	$1/\hat{\gamma}$
0.50	26.052	0.494	26.047	0.494	25.000	0.500	25.060	0.500
0.60	15.104	0.593	15.129	0.592	14.621	0.600	14.713	0.599
1.00	5.010	0.999	5.104	0.987	5.010	0.999	5.211	0.975
1.50	2.674	1.636	2.964	1.481	2.962	1.482	3.226	1.374
1.90	1.489	4.043	2.358	1.876	2.397	1.841	2.680	1.633
2.00	1.314	5.894	2.259	1.975	2.307	1.925	2.592	1.690
3.00	-	-	1.722	2.961	1.835	2.649	2.124	2.137
5.00	-	-	1.385	4.941	1.569	3.573	1.842	2.635
30.00	-	-	1.056	29.537	1.347	5.403	1.580	3.518

Notes: See explanatory note to Table 1. The quantile ratio in this case is  $Q = x_{0.995} / x_{0.975}$ .

symmetric extremal type distribution is the  $Q$ -value close to the  $v_\alpha$  of the stable model.<sup>8</sup>

How can this downward bias be circumvented? Recall the alternative explanation of  $\hat{\alpha}$  through the  $\hat{\gamma}$  estimator. What this estimator does, is to exploit the shape of the Pareto law. By logarithmic transformation the Pareto curve becomes linear. Then taking two points on the line and using simple trigonometry yields  $\hat{\gamma}$  as the size of the angle. At the 75% and 95% quantiles of the stable model this approximates  $1/\hat{\alpha}$ . The  $\hat{\gamma}$ -procedure should, however, also apply to the other distributions, because it is based on the shape of the Pareto distribution. All fat tailed distributions eventually display Pareto like behaviour as the quantiles are increased. Thus by going sufficiently far into the tails the  $\hat{\gamma}$ -procedure should work well. For this reason we have recomputed  $\hat{\gamma}$  for the different models by using the  $x_{0.975}$  and  $x_{0.995}$  quantiles. Results are in Table 3 and the last 3 columns of Table 2. Clearly, by going deeper into the tails a significant improvement in the tail index estimates is achieved. Hence, the McCulloch  $\hat{\alpha}$  procedure can be robustified by going deeper into the tails and using the transformation (12).

Even though going into the tails helps, the downward bias at  $\alpha = 30$  for the Student and symmetric Pareto model is still considerable. For the stable model, going into the tails also helps for the lower  $\alpha$ -values. But the downward bias in  $\hat{\gamma}$  for  $1.50 \leq \alpha \leq 2.00$  from Table 2, column 2, is reversed and worsens into a considerable upward bias. Should one conclude that going into the tails does not work in this case? The answer is that as  $\alpha$  approaches 2 under the stable model,

<sup>8</sup> These findings illustrate that the downward bias in the McCulloch estimator may be due to the non-stable behaviour of the distribution underlying the observations. It appears that the unconditional distribution of the IARCH model displays such non-stable behaviour.

the normal distribution becomes a local alternative and biases the  $\hat{\gamma}$  procedure. The same argument can be used to explain the  $\hat{\gamma} = 5.403$  value under the Student-t model with 30 degrees of freedom. The remedy is to go still deeper into the tails.<sup>9</sup> The stable model is just special in the sense that there happen to be two regions where  $\hat{\gamma}$  yields good estimates.

## 5. Conclusions

To summarize, we first showed that in contradiction with the presentation in Ghose and Kroner the stable laws and the unconditional distribution of the stationary ARCH models are partially nested with respect to their tail index. If  $\zeta > 2$ , then this can only be the case if the model is ARCH. In case  $\beta_0 > 2e^\gamma$ , then the model is necessarily stable. In the interval  $\beta_0 \in [1, 2 \exp(\gamma)]$  both models are candidates. Then we reviewed the two approaches for measuring the tail index value  $\zeta$ : McCulloch's method  $\hat{\alpha}$  which is conditional on the stable model, and the tail index estimators  $\hat{\zeta}$  and  $\hat{\gamma}$ . It was shown how the  $\hat{\alpha}$  procedure can be viewed as a particular case of the  $\hat{\gamma}$  estimator. This view explains the downward bias in the  $\hat{\alpha}$  procedure if the innovations are not stable, but Student-t or IARCH. The reason is that  $\hat{\alpha}$  does not go deep enough in the tails, whereas this is possible with the  $\hat{\gamma}$  procedure. On the one hand,  $\hat{\gamma}$  and  $\hat{\zeta}$  are upward biased if the data are generated by a stable law with  $\alpha$  close to 2. On the other hand if the true data exhibit a tail index  $\zeta$  of about 2, the  $\hat{\gamma}$  and  $\hat{\zeta}$  have hardly any bias. When applied to real data as in Ghose and Kroner the  $\hat{\alpha}$  procedure yields values around 1.46, while  $\hat{\zeta}$  gives values slightly above 2. Combining these two findings with the above, it follows that our results give further credence to Ghose and Kroner's main conclusion.

For the interested researcher there are still some nagging questions in this area. It is known, see Diebold (1988), that appropriately normed sums of covariance stationary ARCH variates adhere to the central limit law. An open issue, which was touched upon in an earlier version of the Ghose and Kroner paper, is whether appropriately normed sums of stationary ARCH variates with  $\beta_0 > 1$  are in the domain of attraction of a non-normal stable distribution. The difficulty is that theorems like in Ibragimov and Linnik (1971) presuppose strong mixing. As of today this property has not been established for ARCH processes. An alternative route is to use the theory on the point process character of exceedances, because only the highest realizations contribute to the sum as the sample gets large. But so far this route has not been taken.

<sup>9</sup> The selection of the optimal number of order statistics is pursued in Hall (1982), Goldie and Smith (1987), Dekkers and de Haan (1991), Dacorogna et al. (1994), and Danielsson and de Vries (1994). This work is relevant for simultaneously reducing the bias and variance of the estimators and for providing guidelines to select the number of order statistics to be used in relatively small samples.

Another open problem is as follows. De Haan et al. (1989) obtained the limit law for the maximum order statistic from a sample of ARCH(1) realizations. Using this result, the tail index for a GARCH(1,1) process can be derived. However, Kesten's (1973) theory cannot be used for higher order ARCH and GARCH processes.

## References

- Dacorogna, M.M., U.A. Müller, O.V. Pictet and C.G. de Vries, 1994, The distribution of extremal foreign exchange rate returns in extremely large data sets, Tinbergen Institute Rotterdam, Discussion Paper 95-70.
- Danielsson, J. and C.G. de Vries, 1994, Robust tail index and quantile estimation, Mimeo, Tinbergen Institute Rotterdam.
- Dekkers, A.L.M. and L. de Haan, 1991, Optimal choice of sample fraction in extreme value estimation, in: A.L.M. Dekkers, *On Extreme-Value Estimation*, Ph.D. thesis (Erasmus University Rotterdam) 73–95.
- Diebold, F.X., 1988, *Empirical Modelling of Exchange Rate Dynamics* (Springer-Verlag, Berlin).
- Engle, R.F., D.M. Lilen and R.P. Robbins, 1987, Estimating time varying risk premia in the term structure: The ARCH-M model, *Econometrica*, 55, 391–407.
- Fama, E.F. and R. Roll, 1971, Parameter estimates for symmetric stable distributions, *Journal of the American Statistical Association*, 66, 331–338.
- Feller, W., 1971, *An Introduction to Probability Theory and Its Applications*, Vol. 2, second edition (Wiley, New York).
- Ghose, D. and K.F. Kroner, 1995, The relationship between GARCH and symmetric stable processes: Finding the source of fat tails in financial data, *Journal of Empirical Finance*, this issue.
- Goldie, C.M. and R.L. Smith, 1987, Slow variation with remainder: Theory and applications, *Quarterly Journal of Mathematics (Oxford 2nd Series)*, 38, 45–71.
- Haan, L. de and S.I. Resnick, 1980, A simple asymptotic estimate for the index of a stable distribution, *Journal of the Royal Statistical Society (Series B)*, 42, 83–87.
- Haan, L. de, S.I. Resnick, H. Rootzén and C.G. de Vries, 1989, Extremal behaviour of solutions to a stochastic difference equation with applications to ARCH processes, *Stochastic Processes and their Applications*, 32, 213–224.
- Hall, P., 1982, On some simple estimates of an exponent of regular variation, *Journal of the Royal Statistical Society (Series B)*, 44, 37–42.
- Hong, C.-H., 1988, The integrated generalized autoregressive conditional heteroskedastic model: The process, estimation and Monte Carlo experiments, Unpublished manuscript (Department of Economics, University of California, San Diego, CA).
- Ibragimov, I.A. and Yu V. Linnik, 1971, *Independent and Stationary Sequences of Random Variables* (Wolters-Noordhoff, Groningen).
- Kesten, H., 1973, Random difference equations and renewal theory for products of random matrices, *Acta Mathematica*, 131, 207–248.
- McCulloch, J.H., 1986, Simple consistent estimators of stable distribution parameters, *Communications in Statistics – Simulation*, 15, 1109–1136.
- Nelson, D.B., 1990, Stationarity and persistence in the GARCH(1,1) model, *Econometric Theory*, 6, 318–334.
- Pagan, A., 1993, *The econometrics of financial markets*, Mimeo, The Australian National University.
- Vervaat, W., 1979, On a stochastic difference equation and a representation of non-negative infinitely divisible random variables, *Advances in Applied Probability*, 11, 750–783.