

Mittnik, Stefan; Paoella, Marc S.

**Working Paper**

## Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions

CFS Working Paper, No. 2003/04

**Provided in Cooperation with:**

Center for Financial Studies (CFS), Goethe University Frankfurt

*Suggested Citation:* Mittnik, Stefan; Paoella, Marc S. (2003) : Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions, CFS Working Paper, No. 2003/04, Goethe University Frankfurt, Center for Financial Studies (CFS), Frankfurt a. M., <https://nbn-resolving.de/urn:nbn:de:hebis:30-10106>

This Version is available at:

<https://hdl.handle.net/10419/72636>

**Standard-Nutzungsbedingungen:**

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

**Terms of use:**

*Documents in EconStor may be saved and copied for your personal and scholarly purposes.*

*You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.*

*If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.*



---

No. 2003/04

**Prediction of Financial Downside-Risk  
with Heavy-Tailed Conditional  
Distributions**

Stefan Mittnik / Marc S. Paolella

---



Center for Financial Studies

an der Johann Wolfgang Goethe-Universität [Taunusanlage 6](#) [D-60329 Frankfurt am Main](#)

Tel: (+49)069/242941-0 [Fax: \(+49\)069/242941-77](#) [E-Mail: ifk@ifk-cfs.de](#) [Internet: http://www.ifk-cfs.de](#)

# Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions

Stefan Mittnik / Marc S. Paolella \*

March 2003

**Abstract:**

The use of GARCH models with stable Paretian innovations in financial modeling has been recently suggested in the literature. This class of processes is attractive because it allows for conditional skewness and leptokurtosis of financial returns without ruling out normality. This contribution illustrates their usefulness in predicting the downside risk of financial assets in the context of modeling foreign exchange-rates and demonstrates their superiority over use of normal or Student's  $t$  GARCH models.

**JEL Classification:** C22, C51, G10

**Keywords:** Risk Management, Value at Risk, Density Forecasting, Predictive Likelihood

# 1 Introduction

Risk managers of financial institutions are particularly interested in the left—i.e., downside—tail of the return distribution of financial assets. To assess the short-term exposure to market risks, they are required to evaluate future shortfall probabilities or value-at-risk levels of financial investments. Such estimates can be based on the distribution of the returns themselves. For example, ever since the pioneering works of Mandelbrot (1963) and Fama (1965), there have been numerous studies investigating the appropriateness of the stable Paretian distribution for modeling the unconditional distribution of asset returns (for an overview, see, for example, Mittnik and Rachev, 1993; and McCulloch, 1997).

However, short-term prediction often benefits substantially when taking conditional volatility into account. The GARCH class of conditional models has been widely and—both from an academic and applied perspective—successfully used to model returns on financial assets (see Palm, 1997 and Gouriéroux, 1997, for surveys). Although a stationary GARCH model with normally distributed innovations gives rise to an unconditional distribution with higher (possibly nonexistent) kurtosis than the normal, it is often found that residuals from estimated GARCH models of financial return data still tend to exhibit nonnegligible kurtosis. To allow for this, other fatter tailed distributions for GARCH innovations have been considered in the literature, most notably the Student's  $t$ . Only very recently has the stable Paretian distribution been considered in the context of modeling the conditional heteroscedastic distribution of asset returns. Special cases of the model considered herein were developed by McCulloch (1985), Nelson (1990), Panorska *et al.* (1995), and Mittnik *et al.* (1998), while a more general case was examined in Liu and Brorsen (1995), Paoletta (1999) and Mittnik, Paoletta and Rachev (2000, 2002).

Like the Student's  $t$ , the stable Paretian distribution includes the normal distribution as a special, limiting case and permits heavy-tailed distributions for GARCH innovations. However, the stable Paretian distribution allows for skewness, an attractive property in financial applications not shared by the Student's  $t$ . In addition to this practical aspect, the stable Paretian distribution also has the appealing theoretical property that it is the only valid distribution that arises as a limiting distribution of sums of independently, identically distributed (iid) random variables. This is highly desirable, given that error terms in econometric models are usually interpreted as random variables that represent the sum of the external effects not being captured by the model.

This contribution investigates the use of asymmetric stable Paretian power GARCH models for modeling downside risk and demonstrates that this model class is more suitable than the class of Student's  $t$  GARCH models, particularly when one uses a goodness-of-fit criterion that focuses on the tails of the conditional distribution.

The remainder is organized as follows. Section 2 discusses GARCH processes with stable Paretian innovations and stationarity conditions. Section 3 reconsiders the empirical analysis of the five exchange-rate series in Liu and Brorsen (1995) using the appropriate measure for persistence of volatility and compares the goodness of fit of the estimated stable Paretian and Student's  $t$  GARCH models. The problem of out-of-sample conditional density prediction with particular focus on predicting downside market risk is considered in Section 4. Section 5 concludes.

## 2 GARCH–Stable Processes

Sequence  $y_t$  is said to be a stable Paretian power GARCH process or, in short, an  $S_{\alpha,\beta}^\delta$  GARCH( $r, s$ ) process (see Panorska *et al.*, 1995; Paoletta, 1999; Rachev and Mittnik, 2000), if

$$y_t = \mu + c_t \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} S_{\alpha,\beta}(0, 1) \quad (1)$$

and

$$c_t^\delta = \theta_0 + \sum_{i=1}^r \theta_i |y_{t-i} - \mu|^\delta + \sum_{j=1}^s \phi_j c_{t-j}^\delta, \quad (2)$$

where  $S_{\alpha,\beta}(0, 1)$  denotes the standard asymmetric stable Paretian distribution with stable index  $\alpha$ , skewness parameter  $\beta \in [-1, 1]$ , zero location parameter, and unit scale parameter. There exist several notational varieties of the stable Paretian distribution; we use the same as in Samorodnitsky and Taqqu (1994) and Rachev and Mittnik (2000), whereby

$$\int_{-\infty}^{\infty} e^{itx} dH(x) = \begin{cases} \exp\{-c^\alpha |t|^\alpha [1 - i\beta \text{sign}(t) \tan \frac{\pi\alpha}{2}] + i\delta t\}, & \text{if } \alpha \neq 1, \\ \exp\{-c|t|[1 + i\beta \frac{2}{\pi} \text{sign}(t) \ln |t|] + i\delta t\}, & \text{if } \alpha = 1, \end{cases} \quad (3)$$

is the characteristic function and  $H$  denotes the distribution function corresponding to  $S_{\alpha,\beta}(\delta, c)$ . The density is symmetric for  $\beta = 0$  and skewed to the right (left) for  $\beta > 0$  ( $\beta < 0$ ). Stable index  $\alpha$ , which, in general, assumes values in interval  $(0, 2]$ , determines the tail-thickness of the distribution. As  $\alpha$  approaches 2, tails become thinner; and for  $\alpha = 2$  the standard stable Paretian distribution coincides with normal distribution  $N(0, 2)$ . For  $\alpha < 2$ ,  $\varepsilon_t$  does not possess moments of order  $\alpha$  or higher.

Mittnik *et al.* (2002) derived sufficient conditions under which the  $S_{\alpha,\beta}^\delta \text{GARCH}(r,s)$  process has a unique strictly stationary solution. These are given by  $1 < \alpha \leq 2$ ,  $0 < \delta < \alpha$ ,  $c_0 > 0$ ,  $c_i \geq 0$ ,  $i = 1, \dots, r$ ,  $r \geq 1$ ,  $d_j \geq 0$ ,  $j = 1, \dots, s$ ,  $s \geq 0$ , and that the *volatility persistence*,  $V_S$ , defined by

$$V_S := \mathbb{E}|Z|^\delta \sum_{i=1}^r \theta_i + \sum_{j=1}^s \phi_j \quad (4)$$

for  $Z \sim S_{\alpha,\beta}(0,1)$ , satisfies

$$V_S \leq 1. \quad (5)$$

If  $1 < \alpha \leq 2$  and  $0 < \delta < \alpha$ , they also showed that

$$\lambda_{\alpha,\beta,\delta} := \mathbb{E}|Z|^\delta = \frac{1}{\psi_\delta} \Gamma\left(1 - \frac{\delta}{\alpha}\right) (1 + \tau_{\alpha,\beta}^2)^{\frac{\delta}{2\alpha}} \cos\left(\frac{\delta}{\alpha} \arctan \tau_{\alpha,\beta}\right), \quad (6)$$

where  $\tau_{\alpha,\beta} := \beta \tan(\alpha\pi/2)$  and

$$\psi_\delta = \begin{cases} \Gamma(1 - \delta) \cos \frac{\pi\delta}{2}, & \text{if } \delta \neq 1, \\ \pi/2, & \text{if } \delta = 1. \end{cases} \quad (7)$$

Restrictions  $1 < \alpha \leq 2$  and  $0 < \delta < \alpha$  not only appear to be satisfied for the data sets used below, but also for other, even more volatile series, such as stock price indices and East Asian currencies (see Mittnik *et al.*, 1998 and Mittnik *et al.*, 2000, respectively).

Analogous to the ordinary normal GARCH model (Engle and Bollerslev, 1986), we say that  $y_t$  is an *integrated*  $S_{\alpha,\beta}^\delta \text{GARCH}(r,s)$  process, denoted  $S_{\alpha,\beta}^\delta \text{IGARCH}(r,s)$ , if, in (5),  $V_S = 1$ . In practice, the estimated volatility persistence,  $\widehat{V}_S$ , tends to be quite close to one for highly volatile series, so that an integrated model might offer a reasonable data description. Because both finite sample and even asymptotic properties of  $\widehat{V}_S$  and the associated likelihood ratio test statistics are not known (see, however, Mittnik *et al.*, 2000), it is not immediately clear how one can test for an integrated process. Instead of formally testing, we suggest fitting both models and examining the change in various goodness-of-fit statistics, most notably the Anderson-Darling statistic, which is particularly relevant for assessing the models' ability to successfully model the value-at-risk (see Section 3.3 below).

### 3 Modeling Exchange–Rate Returns

To examine the appropriateness of the stable GARCH hypothesis, we model returns<sup>1</sup> on five daily spot foreign exchange rates against the U.S. dollar, namely the British pound, Canadian dollar, German mark, Japanese yen, and the Swiss franc. The choice of exchange rate allows us to compare our more general GARCH specification to that used by Liu and Brorsen (1995), who set  $\alpha = \delta$  in (2). However, our sample is somewhat larger than theirs, covering the period January 2, 1980 to July 28, 1994, yielding series of lengths 3681, 3682, 3661, 3621, and 3678, respectively. Serial correlation was found to be negligible, and, as is common in practice, a GARCH( $r, s$ ) specification with  $r = s = 1$  was sufficient to capture serial correlation in the absolute returns. Therefore, we specify a model of the form

$$r_t = \mu + c_t \varepsilon_t \quad (8)$$

$$c_t^\delta = \theta_0 + \theta_1 |r_{t-1} - \mu|^\delta + \phi_1 c_{t-1}^\delta \quad (9)$$

for each of the five currencies.

#### 3.1 Approximate Maximum Likelihood Estimation

Evaluation of the probability density function (pdf) and, thus, the likelihood function of the  $S_{\alpha,\beta}$  distribution is nontrivial, because it lacks an analytic expression. The maximum likelihood (ML) estimate of parameter vector  $\boldsymbol{\theta} = (\mu, c_0, \theta_0, \theta_1, \phi_1, \alpha, \beta, \delta)'$  for the  $S_{\alpha,\beta}^\delta$  GARCH(1, 1) models (8) – (9) is obtained by maximizing the logarithm of the likelihood function

$$L(\boldsymbol{\theta}; r_1, \dots, r_T) = \prod_{t=1}^T c_t^{-1} S_{\alpha,\beta} \left( \frac{r_t - \mu}{c_t} \right), \quad (10)$$

where  $c_0$  denotes the unknown initial value of  $c_t$ .

The ML estimation we conduct is *approximate* in the sense that the stable Paretian density function  $S_{\alpha,\beta}((r_t - \mu)/c_t)$  needs to be approximated. To do so, we follow the algorithm of Mittnik *et al.* (1999), which approximates the stable Paretian density via fast Fourier transform of the characteristic function. DuMouchel (1973) shows that the ML estimator of the parameters of the stable density is consistent and asymptotically normal with the asymptotic covariance matrix

---

<sup>1</sup>We define the return  $r_t$  in period  $t$  by  $r_t = 100 \times (\ln P_t - \ln P_{t-1})$ , where  $P_t$  is the exchange rate at time  $t$ .

being given by the inverse of the Fisher information matrix. Approximate standard errors of the estimates can be obtained via numerical approximation of the Hessian matrix.

Below, we will demonstrate that—for the five series under consideration—the  $S_{\alpha,\beta}^\delta$ GARCH( $r, s$ ) model outperforms its Student’s  $t$  counterpart. However, it is of practical interest to know at least three things before adopting a new and more complex method: First, how easy the stable ML estimation routine is to implement; second, whether it is numerically well-behaved; and third, how fast it performs. When implemented in high-level software which provide both FFT and linear interpolation routines (such as Matlab and Splus), the algorithm becomes a straightforward programming exercise. Our experience has shown that the method is extremely well behaved, giving rise to numerical problems only for grossly misspecified and/or overspecified models (for which the Student’s  $t$  GARCH model also has difficulties) or, in the case of the more general class of ARMA-GARCH models, when there is near zero-pole cancellation in the ARMA structure—a well-known difficulty in ARMA estimation.

The satisfactory behavior of the algorithm is actually not surprising for at least two reasons. First, there is no explicit numerical integration involved (as in the approach of Liu and Brorsen, 1995) and, second, the method can be made arbitrarily accurate by the choice of several tuning constants (recommendations for which are given in Mittnik *et al.*, 1999). Nevertheless, it is clear that the method will take longer than the (essentially closed form) evaluation of the Student’s  $t$  density. For the series considered in this paper, use of a quasi-Newton minimization algorithm (BFGS, as implemented in Matlab 5.2) with convergence tolerance of  $10^{-4}$  resulted in convergence after about 150 to 350 function evaluations (including gradient calculations). Rather contrary to our initial expectations—and fears—the choice of initial values is of surprisingly little importance. Given any “reasonable” set of values, say  $\alpha > 1.4$ ,  $|\beta| < 0.7$ ,  $|\mu| < 0.2$ ,  $\theta_0 > 0$ ,  $\theta_1 > 0$  and  $\phi_1 > 0.2$ , convergence to the same respective maxima occurred for all five exchange-rate series under consideration, and also for the vast majority of trials from simulation experiments. From a purely numerical standpoint then, the method appears both highly reliable and “stable”.

Evaluation of the GARCH recursion requires presample values  $\varepsilon_0$  and  $c_0$ . Following Nelson and Cao (1992), one could set those to their unconditional expected values, i.e.,

$$\hat{c}_0 = \frac{\hat{\theta}_0}{1 - \lambda_{\hat{\alpha},\hat{\beta},\hat{\delta}} \sum_{i=1}^r \hat{\theta}_i - \sum_{j=1}^s \hat{\phi}_j} \quad \text{and} \quad \hat{\varepsilon}_0 = \hat{\lambda} \hat{c}_0. \quad (11)$$

In the IGARCH case, (11) will be invalid, so we instead *estimate*  $c_0$  as an additional parameter.

In fact, we chose to do this for all models considered here, as (11) will clearly be problematic for nearly integrated GARCH models.

For the integrated model  $S_{\alpha,\beta}^\delta$ IGARCH(1,1), the restriction  $\phi_1 = 1 - \lambda_{\alpha,\beta,\delta}\theta_1$  needs to be imposed. Notice that this entails evaluation of (4) at each iteration, as  $\phi_1$  is also dependent on values  $\hat{\alpha}$ ,  $\hat{\beta}$  and  $\hat{\delta}$ .

We compare the  $S_{\alpha,\beta}^\delta$ GARCH model to the most commonly used heavy-tailed variant of the GARCH model, the Student's  $t$ -GARCH models in power form, say  $t_\nu^\delta$ -GARCH( $r, s$ ), given by

$$r_t = \mu + c_t \varepsilon_t, \quad \varepsilon_t \stackrel{\text{iid}}{\sim} t(\nu), \quad (12)$$

$$c_t^\delta = \theta_0 + \sum_{i=1}^r \theta_i |r_{t-i} - \mu|^\delta + \sum_{j=1}^s \phi_j c_{t-j}^\delta, \quad (13)$$

where  $t(\nu)$  refers to the Student's  $t$  distribution with  $\nu$  degrees of freedom, i.e.,

$$f(x; \nu) = K_\nu \left(1 + \frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}} \quad (14)$$

and

$$K_\nu = \frac{\Gamma\left(\frac{\nu+1}{2}\right) \nu^{-1/2}}{\sqrt{\pi} \Gamma\left(\frac{\nu}{2}\right)}. \quad (15)$$

Assuming  $0 < \delta < \nu$  and  $\nu > 1$ ,<sup>2</sup> taking unconditional expectations of  $c_t^\delta$  in (13) shows that  $E c_t^\delta$  exists if  $E |T|^\delta \sum_{i=1}^r \theta_i + \sum_{j=1}^s \phi_j < 1$ , where  $T \sim t(\nu)$  and

$$\lambda_{\nu,\delta} := E |T|^\delta = \sqrt{\frac{\nu^\delta}{\pi}} \Gamma\left(\frac{\delta+1}{2}\right) \Gamma\left(\frac{\nu-\delta}{2}\right) \Gamma^{-1}\left(\frac{\nu}{2}\right). \quad (16)$$

Analogous to (4), the measure of volatility persistence for  $t_\nu^\delta$ -GARCH( $r, s$ ) models is defined to be

$$V_t := \lambda_{\nu,\delta} \sum_{i=1}^r \theta_i + \sum_{j=1}^s \phi_j. \quad (17)$$

Similar remarks regarding treatment of presample values and the imposing of the IGARCH constraint apply to the Student's  $t$  model as well.

---

<sup>2</sup>The condition  $\nu > 1$  is analogous to requiring  $\alpha > 1$  in the stable Paretian case and implies existence of a finite first moment of the innovations.

### 3.2 Estimation Results and Volatility Persistence

The parameter estimates of the models are presented in Table 1. Noteworthy are the estimates of the skewness parameter  $\beta$ : All  $\hat{\beta}$  values are (statistically) significantly different from zero, although those for the British pound and German mark series are quite close to zero. In addition, when  $|\beta| < 0.3$  and  $\alpha$  is over 1.8, the amount of skewness is, for practical purposes, slight. Skewness is most pronounced for the Japanese yen, for which  $\hat{\alpha} = 1.81$  and  $\hat{\beta} = -0.418$ .

The persistence-of-volatility measure given in the last column of Table 1 reflects the speed with which volatility shocks die out. A  $\hat{V}$ -value near one is indicative of an integrated GARCH process, in which volatility shocks have persistent effects. Under the  $S_{\alpha,\beta}$  assumption, the models for the Canadian dollar ( $\hat{V}_S = \lambda_{\hat{\alpha},\hat{\beta},\hat{\delta}}\hat{\theta}_1 + \hat{\phi}_1 = 1.001$ ) and Japanese yen ( $\hat{V}_S = 1.002$ ) series would suggest that they are very close to being integrated. Under the Student's  $t$  assumption,  $\hat{V}_t = \lambda_{\hat{\nu},\hat{\delta}}\hat{\theta}_1 + \hat{\phi}_1 = 0.992$  for the Canadian dollar, which is also rather close to being integrated, while  $\hat{V}_t$  is only 0.972 for the Japanese yen. Thus, for these two currencies, the indications regarding persistence of volatility differ under the two distributional assumptions. For the other currencies, the measures are strikingly close, most notably for the German mark ( $\hat{V}_S = \hat{V}_t = 0.969$ ) and the Swiss franc ( $\hat{V}_S = 0.971$ ,  $\hat{V}_t = 0.968$ ). It is interesting to note that, for each of these two currencies, the log-likelihood values  $\mathcal{L}_t$  and  $\mathcal{L}_S$  are also extremely close. These are discussed further in the next section.

For all five series, we also estimated the models with the IGARCH condition imposed. Table 2 shows the resulting parameter estimates. Not surprisingly, for those models for which the persistence measure was close to unity, the IGARCH-restricted parameter estimates differ very little. For the remaining models, the greatest changes occur with the power parameter  $\delta$  and, to a lesser extent, the shape parameters  $\alpha$  and  $\nu$ . The former increase, while the latter decrease under IGARCH restrictions.

It should also be noted that the restriction  $\alpha = \delta$ , imposed by Liu and Brorsen (1995) when estimating GARCH-stable models for the same five currencies, is not supported by our results. This is important because, if  $\delta \geq \alpha$ , the unconditional first moments of  $c_t$  is infinite for any  $\alpha < 2$ . The knife-edge specification  $\delta = \alpha$  does not only induce conceptual difficulties, but also leads to a highly volatile evolution of the  $c_t$  series in practical work. For our estimates, we obtain  $\hat{\delta} < \hat{\alpha}$ , which suggest that conditional volatility  $c_t^\delta$  is a well-defined quantity in the sense that  $E(c_t^\delta | r_{t-1}, r_{t-2}, \dots) < \infty$  for  $V_S < 1$ .

Table 1: GARCH Parameter Estimates<sup>a</sup>

	Intercept	GARCH Parameters				Distribution Parameters		Persistence Measure <sup>b</sup>
	$\mu$	$\theta_0$	$\theta_1$	$\phi_1$	$\delta$	Shape	Skew	$\hat{V}$
British								
$S_{\alpha,\beta}$	-9.773e-3 (0.012)	8.085e-3 (2.39e-3)	0.04132 (6.42e-3)	0.9171 (0.0118)	1.359 (0.0892)	1.850 (0.0245)	-0.1368 (0.0211)	0.984
$t$	-2.312e-3 (0.010)	0.01190 (3.56e-3)	0.06373 (0.0115)	0.9071 (0.0200)	1.457 (0.167)	6.218 (0.615)	—	0.976
Canadian								
$S_{\alpha,\beta}$	5.167e-3 (0.0614)	1.034e-3 (3.12e-4)	0.04710 (6.63e-3)	0.9164 (0.0118)	1.404 (0.0143)	1.823 (0.0104)	0.3577 (0.0209)	1.001
$t$	-2.240e-3 (3.83e-3)	7.774e-4 (6.90e-4)	0.06112 (5.98e-3)	0.9118 (7.27e-3)	1.793 (0.0150)	5.900 (0.0801)	—	0.992
German								
$S_{\alpha,\beta}$	2.580e-3 (0.016)	0.01525 (1.61e-3)	0.05684 (3.44e-3)	0.8971 (7.42e-3)	1.101 (9.78e-3)	1.892 (0.0216)	-0.06779 (0.0184)	0.969
$t$	6.643e-3 (9.21e-4)	0.01812 (2.25e-3)	0.07803 (6.45e-3)	0.8938 (4.43e-3)	1.261 (0.147)	7.297 (0.186)	—	0.969
Japanese								
$S_{\alpha,\beta}$	-0.01938 (0.0166)	4.518e-3 (1.12e-3)	0.06827 (7.91e-3)	0.8865 (0.0124)	1.337 (0.0132)	1.814 (0.0107)	-0.4175 (8.80e-3)	1.002
$t$	5.318e-3 (8.87e-3)	9.949e-3 (3.03e-3)	0.07016 (0.0119)	0.8756 (0.0205)	1.816 (0.162)	5.509 (0.461)	—	0.972
Swiss								
$S_{\alpha,\beta}$	-2.677e-3 (0.0124)	0.01595 (3.30e-3)	0.04873 (6.84e-3)	0.9115 (0.0132)	1.041 (0.144)	1.902 (0.0206)	-0.2836 (0.0722)	0.971
$t$	8.275e-3 (0.0118)	0.02099 (3.91e-3)	0.06825 (6.85e-3)	0.9061 (7.25e-3)	1.159 (0.179)	8.294 (0.933)	—	0.968

<sup>a</sup>Estimated models:  $r_t = \mu + c_t \varepsilon_t$ ,  $c_t^\delta = \theta_0 + \theta_1 |r_{t-1} - \mu|^\delta + \phi_1 c_{t-1}^\delta$ . “Shape” denotes the degrees of freedom parameter  $\nu$  for the Student’s  $t$  distribution and stable index  $\alpha$  for the stable Paretian distribution; “Skew” refers to the stable Paretian skewness parameter  $\beta$ . Standard deviations resulting from ML estimation are given in parentheses.

<sup>b</sup>  $\hat{V}$  corresponds to  $\hat{V}_S$  in the stable Paretian and  $\hat{V}_t$  in the Student’s  $t$  case.  $V = 1$  implies an IGARCH model.

Table 2: IGARCH Parameter Estimates<sup>a</sup>

	Intercept	IGARCH Parameters				Distribution Parameters	
	$\mu$	$\theta_0$	$\theta_1$	$\phi_1$	$\delta$	Shape	Skew
British							
$S_{\alpha,\beta}$	-0.01023 (0.0103)	7.050e-3 (1.79e-3)	0.03781 (5.64e-3)	0.9114 —	1.598 (0.0677)	1.846 (0.0224)	-0.1340 (0.0147)
$t$	-3.033e-3 (0.0101)	4.237e-3 (1.68e-3)	0.05774 (9.83e-3)	0.9130 —	1.949 (0.264)	5.543 (0.484)	—
Canadian							
$S_{\alpha,\beta}$	5.148e-3 (3.65e-3)	1.115e-3 (2.14e-4)	0.04689 (5.71e-3)	0.9154 —	1.404 (0.0143)	1.823 (0.0105)	0.3578 (0.0209)
$t$	-2.098e-3 (3.48e-3)	4.998e-4 (1.37e-4)	0.06468 (7.54e-3)	0.9146 —	1.796 (0.0226)	5.890 (0.0838)	—
German							
$S_{\alpha,\beta}$	8.959e-3 (0.0113)	9.666e-3 (1.85e-3)	0.04518 (6.10e-3)	0.8896 —	1.676 (0.0662)	1.881 (0.0217)	0.03944 (0.0930)
$t$	8.851e-3 (0.0106)	5.505e-3 (1.60e-3)	0.08124 (0.0106)	0.9003 —	1.741 (0.231)	6.560 (0.676)	—
Japanese							
$S_{\alpha,\beta}$	-0.01932 (8.44e-3)	4.814e-3 (9.75e-4)	0.06768 (7.68e-3)	0.8858 —	1.336 (0.0751)	1.814 (0.0226)	-0.4175 (0.0151)
$t$	6.136e-3 (8.57e-3)	5.611e-3 (1.31e-3)	0.06036 (0.0112)	0.8689 —	2.314 (0.224)	5.066 (0.410)	—
Swiss							
$S_{\alpha,\beta}$	3.823e-3 (0.0127)	0.01111 (2.65e-3)	0.03700 (5.40e-3)	0.9009 —	1.724 (0.0419)	1.889 (0.0169)	-0.1703 (0.137)
$t$	9.130e-3 (0.0119)	2.047e-3 (8.34e-4)	0.07125 (9.13e-3)	0.9347 —	1.166 (9.79e-3)	8.194 (0.0996)	—

<sup>a</sup> Estimated models:  $r_t = \mu + c_t \varepsilon_t$ ,  $c_t^\delta = \theta_0 + \theta_1 |r_{t-1} - \mu|^\delta + (1 - \lambda \theta_1) c_{t-1}^\delta$  with IGARCH condition  $\hat{\phi}_1 = 1 - \hat{\lambda} \hat{\theta}_1$  imposed. See footnote to Table 1 for further details.

### 3.3 Goodness of Fit

We employ three likelihood-based and one empirical CDF-based criteria for comparing the goodness of fit of the candidate models. The first is the maximum log-likelihood value obtained from ML estimation. This value may be viewed as an overall measure of goodness of fit and allows us to judge which candidate is more likely to have generated the data. The second is the AICC (Hurvich and Tsai, 1989; see also Brockwell and Davis, 1991, eq. 9.3.7) given by

$$\text{AICC} = -2\mathcal{L} + \frac{2T(k+1)}{T-k-2}, \quad (18)$$

where  $k$  denotes the number of estimated parameters and  $T$  the number of observations. This is the bias-corrected information criterion of Akaike (1973), which corrects the latter's tendency to overfit. Similarly, the SBC (Schwartz, 1978), defined as

$$\text{SBC} = -2\mathcal{L} + \frac{k \log(T)}{T}, \quad (19)$$

is a similar penalizing strategy which is commonly used.

The fourth criterion is the Anderson-Darling statistic (Anderson and Darling, 1952; see also Press *et al.*, 1991; and Tanaka, 1996), given by

$$\text{AD} = \sup_{x \in \mathbf{R}} \frac{|F_s(x) - \hat{F}(x)|}{\sqrt{\hat{F}(x)(1 - \hat{F}(x))}}, \quad (20)$$

where  $\hat{F}(x)$  denotes the cdf of the estimated parametric density, and  $F_s(x)$  is the empirical sample distribution, i.e.,

$$F_s(x) = \frac{1}{T} \sum_{t=1}^T \mathcal{I}_{(-\infty, x]} \left( \frac{r_t - \hat{\mu}}{\hat{c}_t} \right),$$

where  $\mathcal{I}(\cdot)$  is the usual indicator function. The AD statistic weights discrepancies appropriately across the whole support of the distribution. This is especially important if one is interested in determining conditional shortfall probabilities, i.e., the probability of large investment losses, or so-called value-at-risk measures, where one focuses on the left tail of the conditional return distribution.

Table 3 displays the aforementioned goodness-of-fit measures for the estimated models. In both the unrestricted and IGARCH restricted cases, the inference suggested from the maximum log-likelihood value  $\mathcal{L}$ , and the AICC and SBC are identical. This is not too surprising, given the large ratio of observations to parameters, and the fact that there is only one parameter difference between the Student's  $t$  and stable Paretian GARCH models.

It appears that  $\mathcal{L}$  significantly favors the Student's  $t$  distribution for the British pound (with values, in obvious notation,  $\mathcal{L}_t = -3828.6$  and  $\mathcal{L}_S = -3842.0$ ) and the Canadian dollar ( $\mathcal{L}_t =$

$-152.25$ ,  $\mathcal{L}_S = -159.92$ ). For the German mark ( $\mathcal{L}_t = -3896.2$ ,  $\mathcal{L}_S = -3896.5$ ) and the Swiss franc ( $\mathcal{L}_t = -4308.1$ ,  $\mathcal{L}_S = -4308.6$ ), the log-likelihood values, AICC and SBC are very close, albeit larger for the Student's  $t$ . On the other hand, the  $S_{\alpha,\beta}$  assumption is favored quite strongly for the Japanese yen with  $\mathcal{L}_S = -3178.7$  as compared to  $\mathcal{L}_t = -3331.7$ .

For the British pound, the AD statistic ( $AD_t = 0.0244$ ,  $AD_S = 0.0375$ ) slightly favors the Student's  $t$  model, in agreement with  $\mathcal{L}$ , although the difference is relatively small. The AD statistics for the remaining countries all favor the stable Paretian model, particularly for the German mark ( $AD_t = 0.345$ ,  $AD_S = 0.0368$ ), the Japanese yen ( $AD_t = 0.0986$ ,  $AD_S = 0.0401$ ) and the Swiss franc ( $AD_t = 0.287$ ,  $AD_S = 0.0457$ ). The usual caveat applies, in that, statistically speaking, it is not clear to what extent these differences are significant. However, given virtually identical log-likelihood values, but AD statistics which are several times smaller for the  $S_{\alpha,\beta}$  distribution, one might safely conclude that, particularly in the tails of the conditional distribution, the  $S_{\alpha,\beta}$  model offers a distinct advantage, irrespective of its desirable theoretical properties which are not shared by the Student's  $t$  distribution.

For each currency and both distributional assumptions, Figure 1 plots the values

$$AD_t = \frac{|F_s(\hat{\varepsilon}_{t:T}) - \hat{F}(\hat{\varepsilon}_{t:T})|}{\sqrt{\hat{F}(\hat{\varepsilon}_{t:T})(1 - \hat{F}(\hat{\varepsilon}_{t:T}))}}$$

$t = 1, \dots, T$ , where  $T$  is the sample size and  $\hat{\varepsilon}_{t:T}$  denotes the *sorted* GARCH-filtered residuals. In most cases, most notably for the Student's  $t$  GARCH model of the German, Japanese and Swiss currency returns, the maximum absolute value of the  $AD_t$  occurs in the (left) tail of the distribution.

Turning now to the IGARCH-restricted fits, it is clear that the log-likelihood values must necessarily decrease, since none of the unrestricted GARCH models precisely satisfied the IGARCH restrictions. However, for the  $S_{\alpha,\beta}$  model of the Canadian dollar ( $\mathcal{L} = 159.97$ ) and Japanese yen ( $\mathcal{L} = 3178.8$ ), the log-likelihoods are very close to their unrestricted counterparts. This was expected, as the IGARCH condition for the unrestricted models of these two currencies were nearly met. Somewhat surprising, however, is the small *decrease* in AD values for the  $S_{\alpha,\beta}$  model of the Canadian dollar ( $AD_S = 0.0529$ ) and Japanese yen ( $AD_S = 0.0394$ ). Particularly for the latter two currencies, stable IGARCH models appear to describe the daily returns quite plausibly.

## 4 Prediction of Densities and Downside Risk

Decisions on financial investments are typically based on the expected return and the expected risk of the assets under consideration. Rather than adhering to the conventional mean-variance

Table 3: Goodness-of-Fit Measures of Estimated Models <sup>a</sup>

	$\mathcal{L}$		AICC		SBC		AD	
	$S_{\alpha,\beta}$	$t$	$S_{\alpha,\beta}$	$t$	$S_{\alpha,\beta}$	$t$	$S_{\alpha,\beta}$	$t$
Britain:								
GARCH	-3842.0	-3828.6	7700.0	7671.2	7684.0	7657.2	0.0375	0.0244
IGARCH	-3842.3	-3837.1	7698.6	7686.2	7684.6	7674.2	0.0417	0.0420
Canada:								
GARCH	-159.92	-152.25	0335.9	0318.5	0319.9	0304.5	0.0532	0.0571
IGARCH	-159.97	-153.71	0334.0	0319.4	0320.0	0307.4	0.0529	0.0633
Germany:								
GARCH	-3986.5	-3986.2	7989.0	7986.4	7973.0	7972.4	0.0368	0.345
IGARCH	-3989.9	-3999.4	7993.8	8010.8	7979.8	7998.8	0.0506	0.200
Japan:								
GARCH	-3178.7	-3333.7	6373.4	6681.4	6357.4	6667.4	0.0401	0.0986
IGARCH	-3178.8	-3334.6	6371.6	6681.2	6357.6	6669.2	0.0394	0.0793
Switzerland:								
GARCH	-4308.6	-4308.1	8633.2	8630.2	8617.2	8616.2	0.0457	0.287
IGARCH	-4314.2	-4325.0	8642.4	8662.0	8628.4	8650.0	0.0460	0.278

<sup>a</sup> $\mathcal{L}$  refers to the maximum log-likelihood value; AICC is the corrected AIC criteria (18); SBC is the Schwarz Bayesian Criteria (19); and AD is the Anderson-Darling statistic (20).

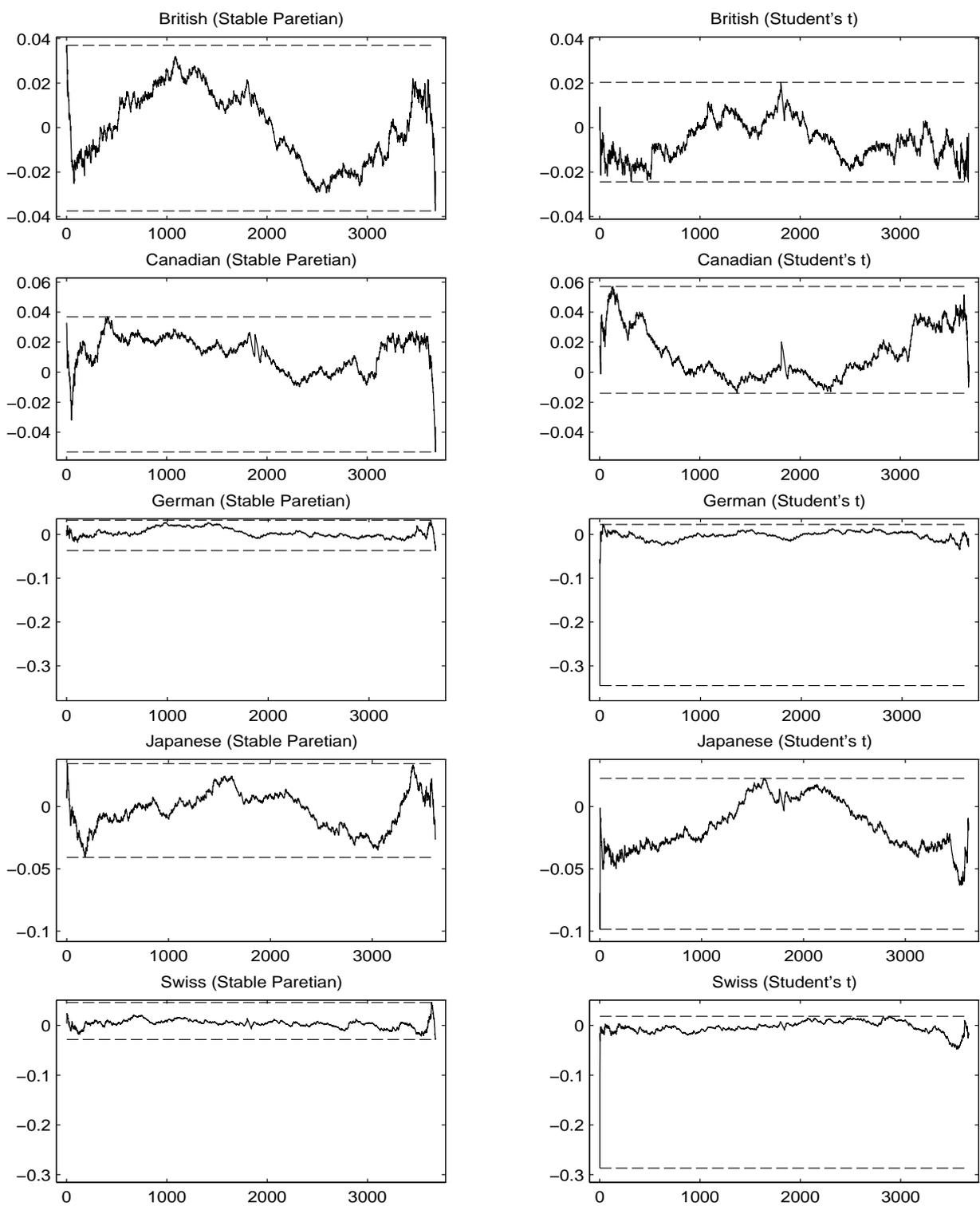


Figure 1: Comparison of the variance adjusted differences between the sample and fitted distribution functions.

criterion, recent risk management concepts for financial institutions focus on the downside risk or the value-at-risk of a financial position due to market movements. In this context, a typical question would be: what is the probability that the value of a financial position will drop by 50% or more over the next period, i.e.,  $\Pr(r_{t+1} < -0.50)$ ? Alternatively, one may ask what is the threshold or value-at-risk,  $-z(\gamma)$ , under which a position will not fall with a probability of  $100(1 - \gamma)\%$ ; i.e., find  $-z(\gamma)$  such that  $\Pr(r_{t+1} < -z(\gamma)) = \gamma$ .

Under unconditional normality, it would be sufficient to simply predict the conditional mean and variance to answer such questions. However, for GARCH processes driven by nonnormal, asymmetric and, possibly, infinite-variance innovations, the predictive conditional density

$$\hat{f}_{t+1|t}(r_{t+1}) = f\left(\frac{r_{t+1} - \mu(\hat{\theta}_t)}{c_{t+1}(\hat{\theta}_t)} \middle| r_t, r_{t-1}, \dots\right), \quad (21)$$

needs to be computed. In (21),  $\hat{\theta}_t$  refers to the estimated parameter vector using the sample information up to and including period  $t$ ; and  $c_{t+1}(\cdot)$  is obtained from the conditional-scale recursion (2) using  $\hat{\theta}_t$ .<sup>3</sup> Multistep density predictions,

$$\hat{f}_{t+n|t}(r_{t+n}) = f\left(\frac{r_{t+n} - \mu(\hat{\theta}_t)}{c_{t+n}(\hat{\theta}_t)} \middle| r_t, r_{t-1}, \dots\right), \quad (22)$$

are obtained by recursive application of (2) with unobserved quantities being replaced by their conditional expectations.

For each of the five currencies under consideration, we evaluate  $\hat{f}_{t+1|t}(r_{t+1})$ ,  $t = 2000, \dots, T - 1$ , for the  $S_{\alpha,\beta}^\delta$ GARCH(1,1) and  $t_\nu^\delta$ GARCH(1,1) models, as well as the conventional GARCH(1,1) model with normal innovations.<sup>4</sup> We re-estimate (via ML estimation) the model parameters at each step, as would typically be done in actual applications.

The overall density forecasting performance of competing models can be compared by evaluating their conditional densities at the future observed value  $r_{t+1}$ , i.e.,  $\hat{f}_{t+1|t}(r_{t+1})$ . A model will fare well in such a comparison if realization  $r_{t+1}$  is near the mode of  $\hat{f}_{t+1|t}(\cdot)$  and if the mode of the conditional density is more peaked. The conditional densities are determined not only by the specification of the mean and GARCH equations, but also by the distributional choice for the innovations.

Table 4 presents the means, standard deviations and medians of the density values  $\hat{f}_{t+1|t}(r_{t+1})$ ,  $t = 2000, \dots, T - 1$ , for each currency. Based on the means, values corresponding to the  $S_{\alpha,\beta}$  and Student's  $t$  assumptions are extremely close, with the Student's  $t$  values nevertheless larger

<sup>3</sup>A conditionally varying location parameter,  $\mu_t$ , would be handled analogously.

<sup>4</sup>Since the sample sizes,  $T$ , of the five currencies vary, the number of forecasts ranges from 1,621 to 1,682.

in each case. Based on the medians, however, the stable Paretian model is (slightly) favored by the British, Canadian and German currencies. Notice that this is contrary to the model selection based on the goodness of fit measures; both AICC and AD statistics favored use of stable Paretian innovations for the Japanese yen and Student's  $t$  innovations for the British pound.

Next, we examine how well the models predict the downside risk. Consider the value-at-risk implied by a particular model,  $M$ , namely

$$\Pr(r_{t+1} \leq -z_{t+1}^M(\gamma)) = \gamma, \quad t = 2000, \dots, T-1. \quad (23)$$

For a correctly specified model we expect  $100\gamma\%$  of the observed  $r_{t+1}$ -values to be less than or equal to the implied threshold-values  $-z_{t+1}^M(\gamma)$ . If the observed frequency

$$\hat{\gamma}^M := \frac{1}{T-2000} \sum_{t=2000}^{T-1} \mathcal{I}_{(-\infty, -z_{t+1}^M(\gamma)]}(r_{t+1})$$

is less (higher) than  $\gamma$ , then model  $M$  tends to overestimate (underestimate) the risk of the currency position; i.e., the implied absolute  $z_{t+1}^M(\gamma)$ -values tend to be too large (small).

The predictive performance for assessing the downside risk achieved by the normal, Student's  $t$  and stable Paretian models are compared in Table 5 for the shortfall probabilities  $\gamma = 0.01, 0.025, 0.05, 0.10$ . A comparison of the stable Paretian and Student's  $t$  GARCH models over the five currencies and four cutoff values,  $\gamma$ , shows that, in 4 out of the 20 cases, the Student's  $t$  GARCH model outperforms that of the stable Paretian, while the latter is more accurate in 11 cases, sometimes considerably so (as for the Canadian dollar with  $\gamma = 0.025$  and  $0.05$ ). The remaining 5 cases are tied.

Table 6 presents summary measures<sup>5</sup> for the predictive performance of the three models across all five currencies in the form of the mean error

$$ME_M(\gamma) = \frac{1}{5} \sum_{i=1}^5 100 (\hat{\gamma}_i^M - \gamma),$$

mean absolute error

$$MAE_M(\gamma) = \frac{1}{5} \sum_{i=1}^5 100 |\hat{\gamma}_i^M - \gamma|$$

and the mean squared error

$$MSE_M(\gamma) = \frac{1}{5} \sum_{i=1}^5 100^2 (\hat{\gamma}_i^M - \gamma)^2.$$

---

<sup>5</sup>The measures are evaluated for  $100\gamma$  rather than  $\gamma$  because the resulting scales of the reported values enhance readability.

Table 4: Comparison of Overall Predictive Performance<sup>a</sup>

	British	Canadian	German	Japanese	Swiss
	Mean				
Normal	0.4198	1.1248	0.4064	0.4796	0.3713
$t$	0.4429	1.1871	0.4258	0.5207	0.3851
$S_{\alpha,\beta}$	0.4380	1.1798	0.4213	0.5173	0.3820
	Standard Deviation				
Normal	0.1934	0.5697	0.1888	0.1988	0.1620
$t$	0.2325	0.6802	0.2151	0.2782	0.1840
$S_{\alpha,\beta}$	0.2189	0.6482	0.2016	0.2662	0.1771
	Median				
Normal	0.4291	1.0824	0.4178	0.5172	0.3942
$t$	0.4483	1.1500	0.4452	0.5261	0.4069
$S_{\alpha,\beta}$	0.4493	1.1730	0.4477	0.5242	0.4041

<sup>a</sup>The entries represent average predictive likelihood values,  $\sum_{t=2000}^{T-1} \hat{f}_{t+1|t}(r_{t+1})$ .

Table 5: Comparison of Predictive Performance for Downside Risk<sup>a</sup>

$100\gamma$	Model	British	German	Canadian	Japanese	Swiss
1.0	Normal	1.9036	1.5051	1.3674	1.9124	1.4899
	$t$	1.3682	0.9031	0.7134	1.4189	1.3707
	$S_{\alpha,\beta}$	1.3682	0.9031	1.3080	1.3572	1.2515
2.5	Normal	3.0339	2.6490	2.3187	2.8994	3.2777
	$t$	2.8554	2.9500	2.1403	3.2079	3.3969
	$S_{\alpha,\beta}$	2.9149	2.9500	2.4970	2.5910	3.1585
5.0	Normal	4.7591	4.5756	3.6266	4.9969	4.7676
	$t$	5.1160	5.2378	3.9834	5.7372	5.0656
	$S_{\alpha,\beta}$	5.1160	5.2378	5.0535	5.2437	5.0656
10.0	Normal	8.3879	9.2113	8.5612	8.0814	8.9392
	$t$	9.8751	10.6562	9.9287	10.3023	10.8462
	$S_{\alpha,\beta}$	9.6966	10.4154	10.2259	9.8088	10.2503

<sup>a</sup>The entries show the observed frequencies  $\hat{\gamma}^M = (T - 2000)^{-1} \sum_{t=2000}^{T-1} \mathcal{I}_{(-\infty, -z_{t+1}^M(\gamma)]}(r_{t+1})$  multiplied by 100. For a correctly specified model, we expect  $\hat{\gamma}^M \approx \gamma$ .

Table 6: Summary Measures for the Predictive Performance<sup>a</sup>

$100\gamma$	Model	$ME(\gamma)$	$MAE(\gamma)$	$MSE(\gamma)$
1.0	Normal	0.6357	0.6357	0.4558
	$t$	0.1549	0.3083	0.1080
	$S_{\alpha,\beta}$	0.2376	0.2764	0.0861
2.5	Normal	0.3357	0.4083	0.2209
	$t$	0.4101	0.5540	0.3527
	$S_{\alpha,\beta}$	0.3223	0.3235	0.1633
5.0	Normal	-0.4548	0.4548	0.4357
	$t$	0.0280	0.4346	0.3302
	$S_{\alpha,\beta}$	0.1433	0.1433	0.0273
10.0	Normal	-1.3638	1.3638	2.0195
	$t$	0.3217	0.4002	0.2517
	$S_{\alpha,\beta}$	0.0794	0.2772	0.0830
Aggregate	Normal	-0.2118	0.7156	0.7830
	$t$	0.2287	0.4243	0.2607
	$S_{\alpha,\beta}$	0.1956	0.2551	0.0899

<sup>a</sup>Shown are the mean error ( $ME$ ), mean absolute error ( $MAE$ ) and mean squared error ( $MSE$ ) of the observed extreme-tail frequencies from Table 5 across the five currencies. The bottom panel is the aggregate over all  $\gamma$ -values considered.

The  $ME$ 's for the normal show that it underestimates the probability of extreme downturns ( $ME_{\text{Normal}}(\gamma) > 0$  for  $\gamma = 0.01, 0.025$ ) and overestimates the probability of moderate downturns ( $ME_{\text{Normal}}(\gamma) < 0$  for  $\gamma = 0.05, 0.10$ ). With one exception, the  $ME$ 's of the stable Paretian and Student's  $t$  GARCH models are smaller (in absolute terms) than those for the normal. However, they are always positive, indicating, on average, slight underprediction of the downturn probabilities. For  $\gamma = 0.01$  and  $\gamma = 0.05$ , the Student's  $t$  model has smaller  $ME$  than the stable Paretian model. This is due to the Student's  $t$  model's offsetting prediction error for the Canadian dollar for these  $\gamma$ -values.

While the  $ME$ 's indicate possible systematic prediction bias, the  $MAEs$  and  $MSEs$  reflect the size of the prediction error. With respect to both measures, the stable Paretian model dominates those of both the normal and the Student's  $t$  for all  $\gamma$ -values considered. This is also evident from the bottom panel of Table 6, which aggregates the summary measures over all  $\gamma$ -values considered. In the aggregate, the model using the stable Paretian innovation assumption outperforms those using the normal and Student's  $t$  in terms of all three summary measures.

## 5 Conclusions

Power GARCH processes driven by either stable Paretian or Student's  $t$  innovations have been evaluated and compared in the context of predicting downside market risk, an activity which is particularly important for risk managers of financial institutions. For all five exchange-rate series considered, the asymmetric stable Paretian distributional assumption was found to be superior.

While there exist several popular model classes designed to parsimoniously and effectively fit financial return data, the GARCH class of models is arguably the most common. Furthermore, the usual assumption, and that which is implemented in popular software packages, is that the driving innovations are either normally or Student's  $t$  distributed. The former is the "standard" assumption in financial and even most econometric or statistical models, but fails demonstrably in empirical applications (see, e.g., Palm, 1997; Gouriéroux, 1997; and the references therein). Indeed, normality is a special, limiting case of the stable Paretian distribution, which, otherwise, allows for fatter-than-normal tails and skewness, these being precisely two of the typical "stylized facts" associated with financial returns data. The Student's  $t$  assumption does allow for fatter tails, but is restricted to being symmetric. The latter restraint can actually be overcome if more general Student's  $t$ -like distributions are used (Paoletta, 1999; Mittnik and Paoletta, 2000), but these suggestions, while often providing admirable in- and out-of-sample fits, do not possess the theoretical property of summability, common only to the stable Paretian (and, thus, normal) class of distributions.

With respect to the summability property, one might argue that the value of stable Paretian models is, as shown here, their improved forecasting ability as compared to competing models, with such “theoretical niceties” as summability being largely irrelevant. In a larger context, however, the summability property can often be judiciously used when building more complex financial models such as those used in portfolio analysis. In such models, the ad hoc nature of, say, the Student’s  $t$  distribution can become quite problematic. Further discussion along these lines and a test for the summability property in the context of GARCH models has been proposed in Paoletta (2001) and further applied in Mittnik *et al.* (2000).

## References

- Akaike, H. (1973). Information Theory and an Extension of the Maximum Likelihood Principle. In Petrov, B. and Csaki, F., editors, *2nd International Symposium on Information Theory*, pages 267–281, Akademiai Kiado, Budapest.
- Anderson, T. and Darling, D. (1952). Asymptotic Theory of Certain “Goodness of Fit” Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics*, 23:193–212.
- Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Springer-Verlag.
- DuMouchel, W. (1973). On the Asymptotic Normality of the Maximum–Likelihood Estimate when Sampling from a Stable Distribution. *Annals of Statistics*, 1:948–957.
- Engle, R. F. and Bollerslev, T. (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, 5:1–50; 81–87.
- Fama, E. (1965). The Behavior of Stock Market Prices. *Journal of Business*, 38:34–105.
- Gourieroux, C. (1997). *ARCH Models and Financial Applications*. Springer, New York.
- Hurvich, C. and Tsai, C. (1989). Regression and Time Series Model Selection in Small Samples. *Biometrika*, 76:297–307.
- Liu, S. and Brorsen, B. W. (1995). Maximum Likelihood Estimation of a GARCH-Stable Model. *Journal of Applied Econometrics*, 10:273–285.
- Mandelbrot, B. (1963). The Variation of Certain Speculative Prices. *Journal of Business*, 36:394–419.
- McCulloch, J. H. (1985). Interest-Risk Sensitive Deposit Insurance Premia: Stable ACH Estimates. *Journal of Banking and Finance*, 9:137–156.
- McCulloch, J. H. (1997). Financial Applications of Stable Distributions. In Maddala, G. and Rao, C., editors, *Handbook of Statistics*, volume 14. Elsevier Science.
- Mittnik, S. and Paoletta, M. S. (2000). Conditional Density and Value-at-Risk Prediction of Asian Currency Exchange Rates. *Journal of Forecasting*, 19:313–333.

- Mittnik, S., Paoella, M. S., and Rachev, S. T. (2000). Diagnosing and Treating the Fat Tails in Financial Returns Data. *Journal of Empirical Finance*, 7:389–416.
- Mittnik, S., Paoella, M. S., and Rachev, S. T. (2002). Stationarity of Stable Power–GARCH Processes. *Journal of Econometrics*, 106:97–107.
- Mittnik, S. and Rachev, S. (1993). Modeling Asset Returns With Alternative Stable Models. *Econometric Reviews*, 12:261–330.
- Mittnik, S., Rachev, S. T., Doganoglu, T., and Chenyao, D. (1999). Maximum Likelihood Estimation of Stable Paretian Models. *Mathematical and Computer Modelling*, 29:275–293.
- Mittnik, S., Rachev, S. T., and Paoella, M. S. (1998). Stable Paretian Modeling in Finance. In Adler, R. J., Feldman, R. E., and Taqqu, M. S., editors, *A Practical Guide to Heavy Tails: Statistical Techniques for Analyzing Heavy Tailed Distributions*. Birkhäuser, Boston.
- Nelson, D. (1990). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Theory*, 6:318–344.
- Nelson, D. B. and Cao, C. Q. (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, 10(2):229–235.
- Palm, F. C. (1997). GARCH Models of Volatility. In Maddala, G. and Rao, C., editors, *Handbook of Statistics*, volume 14. Elsevier Science.
- Panorska, A. K., Mittnik, S., and Rachev, S. T. (1995). Stable GARCH Models for Financial Time Series. *Applied Mathematics Letters*, 8(5):33–37.
- Paoella, M. (1999). *Tail Estimation and Conditional Modeling of Heteroscedastic Time-Series*. Ph.D. Thesis, Institute of Statistics and Econometrics, Christian Albrechts University at Kiel. Pro Business, Berlin.
- Paoella, M. S. (2001). Testing the Stable Paretian Assumption. *Mathematical and Computer Modelling*, 34:1095–1112.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1991). *Numerical Recipes in Fortran: The Art of Scientific Computing, 2nd ed.* Cambridge University Press, New York.
- Rachev, S. T. and Mittnik, S. (2000). *Stable Paretian Models in Finance*. Wiley & Sons, Chichester.
- Samorodnitsky, G. and Taqqu, M. (1994). *Stable Non-Gaussian Random Processes, Stochastic Models with Infinite Variance*. Chapman & Hall, London.
- Schwartz, G. (1978). Estimating the Dimension of a Model. *Annals of Statistics*, 6:461–464.
- Tanaka, K. (1996). *Time Series Analysis, Nonstationary and Noninvertible Distribution Theory*. John Wiley & Sons, New York.

## CFS Working Paper Series:

No.	Author(s)	Title
2002/10	<b>Markus Haas</b> <b>Stefan Mittnik</b> <b>Marc S. Paoella</b>	Mixed Normal Conditional Heteroskedasticity
2002/11	<b>Elke Hahn</b>	Core Inflation in the Euro Area: An Application of the Generalized Dynamic Factor Model
2002/12	<b>Yunus Aksoy</b> <b>Hanno Lustig</b>	On the Short and Long Term Real Effects of Nominal Exchange Rates
2002/13	<b>Andreas A. Jobst</b>	Collateralised Loan Obligations (CLOs) - A Primer
2002/14	<b>Andreas A. Jobst</b>	The Pricing Puzzle: The Default Term Structure of Collateralised Loan Obligations
2002/15	<b>Athanasios Orphanides</b>	Activist Stabilization Policy and Inflation: The Taylor Rule in the 1970s
2002/16	<b>Stefanie A. Franzke</b> <b>Christian Schlag</b>	Over-Allotment Options in IPOs on Germany's Neuer Markt - An Empirical Investigation -
2003/01	<b>Klaus Adam</b>	Learning to Forecast and Cyclical Behavior of Output and Inflation
2003/02	<b>Klaus Adam</b>	On the Relation between Robust and Bayesian Decision Making
2003/03	<b>Klaus Adam</b>	Learning and Equilibrium Selection in a Monetary Overlapping Generations Model with Sticky Prices
2003/04	<b>Stefan Mittnik</b> <b>Marc S. Paoella</b>	Prediction of Financial Downside-Risk with Heavy-Tailed Conditional Distributions

Copies of working papers are available at the Center for Financial Studies or can be downloaded (<http://www.ifk-cfs.de>).