Measuring Tail Thickness under GARCH and an Application to Extreme Exchange Rate Changes

Niklas Wagner
and
Terry A. Marsh

Working Paper RPF-297
Version: September 2003

This paper is partly based on the Berkeley Research Program in Finance Working Paper RPF-295. The authors would like to thank Frank Diebold and participants at the following meetings for their helpful comments: IAFE 2001, Operations Research 2002, European Investment Review 2002, DGF 2002 and CFS 2003 workshop on risk management. All errors are the authors' sole responsibility. Niklas Wagner acknowledges financial support from DFG research grant Wa1380/1-1.

Niklas Wagner, Munich University of Technology, 80290 Munich, Germany, Phone: +49 89 289-25486, Fax: +49 89 289-25488. E-mail: nwagner@alum.calberkeley.org

Terry A. Marsh, Walter A. Haas School of Business, U.C. Berkeley, Berkeley CA 94720-1900, USA. Phone: +1 510 642-1651, Fax: +1 510 643-8460.

Abstract

Accurate modeling of extreme price changes is vital to financial risk management. We examine the small sample properties of adaptive tail index estimators under the class of student-t marginal distribution functions including GARCH and propose a model-based bias-corrected estimation approach. Our simulation results indicate that bias strongly relates to the underlying model and may be positively as well as negatively signed. The empirical study of daily exchange rate changes reveals substantial differences in measured tail-thickness due to small sample bias. As a consequence, high quantile estimation may lead to a substantial underestimation of tail risk.

Key Words: fat tails, tail index, stationary marginal distribution, GARCH, Hill estimator, foreign exchange

JEL Classification: C13, C14, F31

1. Introduction

Events such as market crashes or cases of individual financial distress regularly point out the potential effects of fat-tails in unconditional return distributions. Empirical research in finance aims at a careful modeling of such extreme events and at the same time provides a basis for financial risk management. To this end, extreme value theory models the asymptotic extreme characteristics of distributions of stationary return series. Theory allows us to make inferences about return distributions not only within, but also beyond, the observed range of sample returns and to obtain an adequate characterization of the extreme behavior of returns. Estimation of the so-called tail index turns out to be essential and theory offers a variety of approaches; see for example Embrechts et al. (1997) and Coles (2001). Previous studies of the tail index of return distributions include Dacorogna et al. (1995), Danielsson and de Vries (1997), Huisman et al. (1998, 2001), Jondeau and Rockinger (1999), Lauridsen (2000), and McNeil and Frey (2000), among others.

This paper contributes to the empirical finance literature on modeling fat-tailed distributions of exchange rate changes. It thereby follows the seminal papers by Koedijk et al. (1990) and Loretan and Phillips (1994) who give results on weekly and daily exchange rate data, respectively. Based on earlier studies, the student-t specification as proposed by Clark (1973) was found to provide a suitable model for the distribution of exchange rate changes; recent empirical results e.g. by Bond (2000) indicate that generalized autoregressive conditional heteroskedastic (GARCH) models with student-t innovations (GARCH-t) are a good benchmark model which is hardly outperformed by an asymmetric distributional alternative. Hence, we examine the small sample properties of adaptive tail index estimators under the class of student-t marginal distribution functions including the ARCH-type dependent case; Kearns and Pagan (1997) point out that the latter can have a substantial impact on tail index estimation. Based on our simulation results, we then derive a bias-corrected, adaptive tail index estimation approach. This approach is based on the GARCH-t model as a prior and has two steps. In the first step, the GARCH model is estimated as a benchmark model for the overall exchange rate data, yielding a prior tail index estimate together with its simulated small sample properties under the model. In the second step, we follow the philosophy of extreme value theory to "let the tails speak for themselves" and derive the actual estimate of the tail index of the exchange rate changes only from a subset of extreme sample observations.

The approaches to tail index estimation which we examine go back to the proposal by Hill (1975). The Hill estimator has been used widely in financial applications. The approach is based on the assumption that the underlying distribution is in the maximum domain of attraction of the Fréchet extreme value distribution. This generally holds for fat-tailed distributions as they are usually analyzed in finance. Considering alternative estimation methodologies, we point out that, tail index estimation is typically based on maximum likelihood or other methods of statistical inference jointly with asymptotic arguments from extreme value theory. The choice of a threshold which yields a sample fraction of extreme observations is essential to the bias/variance trade-off for all estimation procedures. Comprehensive theoretical arguments for the case of the Hill estimator are given for example in Segers (2001).

Our simulation results indicate that the interpretation of tail index estimation results in small samples heavily relies on a what is assumed as a suitable model for the underlying series. The novelty of our approach is to study effects under general parametrizations of GARCH(1,1)-t models and to compare them to the independent identically distributed (iid) student-t case. Recent findings by Gomes and Oliveira (2001) and Matthys and Beirlant (2000) document the small sample properties of adaptive estimators under various iid distributional model assumptions also pointing out small sample bias. Based on our results, we propose a bias-corrected, two-step, GARCH-t model based, adaptive estimation approach. In the first step, the GARCH-t model parameters are estimated based on the *overall* dataset. Numerical integration thereby yields an assessment of the tail index which may be considered as a raw prior not derived from the tails of the distribution only. However, this prior together with the underlying GARCH-t model allows us to simulate the small sample properties of tail index estimates and to derive an optimal estimator under the model. In the second step, we use a *subsample* of extreme sample observations and derive the estimate of the tail index from a series of Hill estimates given the information from our first step. Using several Hill estimates instead of just one estimate is also done e.g. in the regression approach by Huisman et al. (2001). We follow the approach by Drees et al. (2000) which aims at choosing an estimate of the tail index within a "stable" region of the so-called *Hill plot*; see also Embrechts et al. (1997).

There are two important points which we have to point out regarding related work in the literature. (i) The proposed procedure allows us to derive *unconditional* tail estimates for the marginal distribution of exchange rate returns. An alternative methodology is to model *conditional* returns and estimate the model innovations based on a GARCH specification; see Diebold et al. (1998), McNeil and Frey (2000) and Lauridsen (2000). With our unconditional approach, we refer to the small sample distributional properties from our simulations as well as to theoretical results. The theoretical robustness results in Hsing (1991) and Resnick and Stărică (1998) show that consistency of the Hill estimator is given not only under independence but also under quite general forms of dependence including ARCH-type dependence. Ignoring dependence, thereby fitting the tail of the marginal distribution, has the advantage that the tail estimator is derived from the original dataset, without the disturbing effects of potential mis-specification of the conditional model which may include estimation error and questions of robustness. Furthermore, the vast empirical findings on ARCH modelling in finance point out that the usefulness of those models lies in predicting periods of increased, decaying return variance *after* exogenous shocks to return variance have occurred. However, risk management in face of extreme events will also be concerned with unpredictable, i.e. unconditional, shocks to return variance that may cause large portfolio losses. (ii) The recent literature suggests an increasing number of methods to reduce estimation bias in Hill estimation applications. An alternative Hill plot based estimation approach is the Huisman et al. (1998, 2001) regression approach. Although our method is not restricted to applications under the GARCH-t model, it is indeed more model-specific; i.e. it relies on a pre-assumed suitable model for the underlying series. Huisman et al. (2001) document favorable small sample properties for their estimator under a particular given GARCH(1,1) model. Our approach may account for small sample bias *explicitly* by taking model parametrization and hence also taking the magnitude of the tail index into account; such a procedure appears appropriate since our simulation results indicate that, for the estimators studied, small sample bias is sensitive with respect to the parameters of a particular GARCH specification.

In our empirical investigation we measure tail-thickness for a given sample of daily US-Dollar/Deutschemark exchange rate changes. Simulating small sample bias under the fitted GARCH model, we find that bias is a dominant issue to cope with in financial risk assessment and therefore we apply our model-based adaptive tail index estimation approach. Compared to more traditional approaches, our inference indicates a 40 percent deviation in the estimate of tail-thickness; given the assumption that the class of student-t distributions is a suitable model, it also implies that the existence of the fourth and even the third moment of the underlying distribution is questionable; this adds to the earlier findings by Loretan and Phillips (1994) who used a conventional Hill estimator. Whereas Huisman et al. (2001) find that tail fatness tends to be overestimated for weekly exchange rate data, we find that the direction of bias is generally *ambiguous*; for our given sample of daily exchange rates, tail fatness turns out to be underestimated by conventional methods and our method indicates substantially higher risk in the tails of the exchange rate distribution. This finding also has implications for high quantile estimation.

Following this introduction, the remainder of the present paper is divided into four main sections. Section 2 presents an outline of the extreme value methodology. This includes results on the asymptotics of extremes of iid and ARCH series, the Hill estimator, a selection of adaptive methods for Hill estimation as well as our bias-corrected extension. The adaptive methods used in our study include sequential, bootstrap and Hill plot-based approaches. Section 3 documents the results of a simulation study for the estimators' small sample properties under the class of marginal student-t distributions. Section 4 contains our empirical study of extreme daily exchange rate changes. Section 5 concludes.

2. The Methodological Background

2.1. Extreme Value Theory: The Classical Case

Classical extreme value theory is concerned with the asymptotic distribution of standardized maxima from a series of iid random variables $(R_t)_{1\leq t\leq T}$ with a common distribution function F. Without loss of generality, we assume that the variables R_t denote time-t log-returns of some financial asset in the following. For given normalizing constants, $a_T > 0$, $b_T \in \mathbb{R}$, and $M_T = \max(R_1, ..., R_T)$, the classical result by Fisher/Tippett and Gnedenko states that if H exists as the non-degenerate distributional limit of the standardized maximum

$$\Pr\{a_T^{-1}(M_T - b_T) \leq r\} = F^T(a_T r + b_T) \longrightarrow H(r) \quad \text{as} \quad T \longrightarrow \infty, \tag{2.1}$$

then $H(r)$ is equal to one of three different types of extreme value distributions. The latter are nested within the so-called generalized extreme value distribution

$$H_\xi(r) = \begin{cases} \exp\left(-(1+\xi r)^{-1/\xi}\right), & \text{for} \quad \xi \neq 0 \\ \exp\left(-\exp(-r)\right), & \text{for} \quad \xi = 0 \end{cases}, \tag{2.2}$$

where $1 + \xi r > 0$. The shape parameter, $\xi \in \mathbb{R}$, also denoted as tail index, characterizes the extreme behavior of the distribution function.

Condition (2.1) states that F belongs to the maximum domain of attraction of H_ξ, $F \in MDA(H_\xi)$. Fat-tailed distribution functions, which are of interest in financial applications, particularly belong to the maximum domain of attraction of the Fréchet type extreme value distribution, $F \in MDA(\Phi_\xi)$, where: $\Phi_\xi(r) = \exp(-r^{-1/\xi}),\ r > 0,\ \xi > 0$. From a theorem by Gnedenko, it is well-known that the condition $F \in MDA(\Phi_\xi)$ is satisfied if and only if the tail $\overline{F}(r) = 1 - F(r)$ of the distribution function F is regularly varying at infinity with parameter $-1/\xi < 0$, i.e.

$$\overline{F}(r) = L(r) r^{-1/\xi}, \quad r > 0, \tag{2.3}$$

where the function $L(r)$ is slowly varying at infinity:

$$\lim_{r \to \infty} \frac{L(sr)}{L(r)} = 1, \quad s > 0.$$

2.2. Extreme Value Theory: ARCH-type Dependence

Extreme value theory for iid series can be extended to the case where the variables $(R_t)_{1 \leq t \leq T}$ have a stationary marginal distribution F. We consider the case of ARCH which is a well-known model class for non-linear dependence in financial time series (see e.g. Bollerslev et al. (1992)). A detailed discussion of ARCH as a model for exchange rate changes is given in Diebold (1988).

Assuming a constant expectation and excess returns R_t with a conditional time-varying variance σ_t^2, the prominent GARCH(1,1) process has the representation:

$$\begin{aligned} R_t &= \sigma_t Z_t, \\ \sigma_t^2 &= \beta_0 + \beta_1 R_{t-1}^2 + \beta_2 \sigma_{t-1}^2, \quad \beta_0, \beta_1, \beta_2 \geq 0. \end{aligned} \tag{2.4}$$

The random variables Z_t are standardized iid draws from some symmetric, possibly fat-tailed, distribution function with density $g(z) : \mathbb{R} \mapsto \mathbb{R}^+$. In our simulation study and in the empirical application, $g(z)$ will be given by the student-t density with $\nu \in \mathbb{N}$ degrees of freedom. This yields the GARCH(1,1)-t(ν) model.

As outlined for example in Mikosch and Stărică (2000) and the literature given therein, a stationary marginal distribution F for the GARCH(1,1) process (2.4) exists if

$$\int_{-\infty}^{\infty} \ln \left|\beta_1 z^2 + \beta_2\right| g(z) dz < 0, \quad \beta_0 > 0. \tag{2.5}$$

It can then be shown that the tail index ξ of the stationary marginal distribution F is given as a solution to an integral equation such that

$$\overline{F}(r) \sim cr^{-1/\xi}, \quad r > 0,\ c > 0, \quad \text{as } r \to \infty,$$

which is equivalent to the implication of result (2.3) from the previous section. For the ARCH(1) process, solutions for the tail index are discussed in de Haan et al. (1989). For the GARCH(1,1) process (2.4), the tail index ξ can be determined numerically as a solution to the following integral equation:[1]

$$I(\xi) \equiv \int_{-\infty}^{\infty} \left|\beta_1 z^2 + \beta_2\right|^{\frac{1}{2\xi}} g(z)dz - 1 = 0, \quad \xi > 0. \tag{2.6}$$

2.3. Semi-parametric Tail Index Estimation

The above results show that the upper[2] tail of a fat-tailed marginal distribution F behaves asymptotically like the tail of the Pareto distribution. For the latter, given the order statistics $R_{T,T} \leq \cdots \leq R_{k,T} \leq \cdots \leq R_{1,T}$, the maximum likelihood estimator (MLE) of ξ is given by

$$\widehat{\xi}_{k,T} = \frac{1}{k}\sum_{i=1}^{k} \ln R_{i,T} - \ln R_{k,T}, \tag{2.7}$$

where $k = T$. As the series $(R_t)_{1\leq t\leq T}$ has a stationary marginal distribution which behaves only asymptotically Pareto-like, $\overline{F}(r) \sim cr^{-1/\xi}$ as $r \to \infty$, optimality of the MLE (2.7) does not apply.

2.3.1. Asymptotic Tail Behavior: Where does the Tail begin?

Hill (1975) proposed a conditional MLE approach. Based on a known high threshold $u > 0$ let model (2.3) hold for $r > u$. Given that k is a realization of $K = \#\{i : R_{i,T} > u\}$, select a subsample of $k < T$ largest observations in (2.7) and the Hill estimator is given as a MLE of ξ conditional on u. The number $k = k(T)$ of order statistics should increase with the overall sample size T, while

[1] This result is based on a set of regularity conditions as well as on the assumption that the tail of the innovations is thinner than that of the conditional standard deviation of the GARCH process; see e.g. Mikosch and Stărică (2000) for details.

[2] The focus on the upper tail is an arbitrary choice. The results hold equivalently for the lower tail since $\min(R_1, ..., R_T) = -\max(-R_1, ..., -R_T)$.

on the other hand, it should be small relative to the overall sample size T. In the literature this is frequently made precise by the condition

$$k(T) \longrightarrow \infty, \quad k(T)/T \longrightarrow 0, \quad \text{as} \quad T \longrightarrow \infty. \tag{2.8}$$

Under model (2.3) and condition (2.8), the Hill estimator can be shown to have following properties: (i) the estimator is consistent (see e.g. Embrechts et al. (1997), Example 4.1.12); (ii) under additional assumptions on the asymptotic tail behavior of F, asymptotic normality follows

$$\sqrt{k}(\widehat{\xi}_{k,T} - \xi) \xrightarrow{d} N(B_\xi; \xi^2), \tag{2.9}$$

where B_ξ denotes some asymptotic bias term, see e.g. de Haan and Peng (1998) and Segers (2001); (iii) the Hill estimator obtains a theoretically derived optimal rate of convergence, being superior to other popular estimators proposed in the literature; see Drees (1998); (iv) the results by Hsing (1991) indicate that the Hill estimator is asymptotically quite robust with respect to deviations from independence; Resnick and Stărică (1998) prove consistency under ARCH-type dependence.

Several approaches to the determination of an optimal sample fraction $k(T)$ for the Hill estimator have been studied. Theoretical results on an optimal bias/variance trade-off can be derived using the asymptotic mean squared error (AMSE) as the optimality criterion. The results are based on the so-called Hall model, which forms a generalization of the Pareto model. It imposes a second order condition on the asymptotic behavior of the tail of the distribution function F. By assuming $L(r) = c_1 \left(1 + c_2 r^{-\rho/\xi} + o(r^{-\rho/\xi})\right)$ in (2.3) it follows that

$$\overline{F}(r) = c_1 r^{-1/\xi} \left(1 + c_2 r^{-\rho/\xi} + o(r^{-\rho/\xi})\right), \quad \text{as } r \to \infty, \tag{2.10}$$

where $c_1 > 0,\ c_2 \in \mathbb{R}$ and $\rho > 0$. The above model gives an asymptotic characterization of the tail of the underlying distribution and at the same time robustifies the semi-parametric estimation approach against deviations from the exact Pareto tail. The model holds for distribution functions such as the Fréchet and the student-t.

We consider two Hill-based adaptive tail index estimation approaches. Both approaches rely on model (2.10) and select a MSE-optimal sample fraction k.

- Drees and Kaufmann (1998) derive a sequential estimator of the optimal k by extending previous theoretical work on the asymptotic bias and variance of the Hill estimator. A stopping time criterion for a sequence of Hill estimators is used in order to approximate the number of upper order statistics k under which the bias in the Hill estimator starts to dominate the maximum random fluctuation of the series $\sqrt{i}\left|\widehat{\xi}_{i,T}-\xi\right|$, $2 \leq i \leq k$, given some threshold $u_T > 0$:

$$\overline{k}(u_T) = \min\left\{k \in \{2, ..., T\} : \max_{2\leq i\leq k}\left(\sqrt{i}\left|\widehat{\xi}_{i,T}-\widehat{\xi}_{k,T}\right|\right) > u_T\right\}.$$

The second order parameter ρ in (1.7) is either set equal to a constant or given by a consistent estimator such as

$$\widehat{\rho}_T(u_T,\lambda) = \frac{1}{\ln\lambda}\ln\frac{\max\limits_{2\leq i\leq[\lambda\overline{k}(u_T)]}\left(\sqrt{i}\left|\widehat{\xi}_{i,T}-\widehat{\xi}_{[\lambda\overline{k}(u_T)],T}\right|\right)}{\max\limits_{2\leq i\leq\overline{k}(u_T)}\left(\sqrt{i}\left|\widehat{\xi}_{i,T}-\widehat{\xi}_{\overline{k}(u_T),T}\right|\right)} - \frac{1}{2}, \qquad (2.11)$$

where $\lambda \in (0,1)$ and $[x]$ denotes the largest integer smaller or equal to x. Then, for a given ρ and a consistent initial estimator of ξ, a consistent estimator of the optimal k is derived. The procedure yields a Hill estimator with minimal asymptotic MSE (Drees and Kaufmann (1998), Theorem 1).

- Dacarogna et al. (1995) as well as Danielsson and de Vries (1997) and Danielsson et al. (2001) use a subsample bootstrap approach derived from Hall (1990) to estimate the optimal sample fraction. The approach is based on bootstrap subsamples of size $T_1 < T$. The estimate of the optimal subsample fraction can be derived from

$$\widehat{k}_1 = \arg\min_{1\leq k\leq T_1}\widehat{E}\left((\widehat{\xi}^*_{k,T_1}-\widetilde{\xi}_T)^2\middle|(R_1,...,R_T)\right), \qquad (2.12)$$

where * denotes estimates based on resamples drawn from $(R_1, ..., R_T)$. $\widetilde{\xi}_T$ denotes a consistent initial Hill estimator. The expectation operator is approximated by a given number of bootstrap runs. Instead of applying the bootstrap to the MSE of the Hill estimator directly as in (2.12), an asymptotically equivalent criterion which does not depend on the choice of an initial estimator is proposed in Danielsson et al. (2001). Their alternative bootstrap estimate is given by[3]

$$\widehat{k}_1 = \arg \min_{1 \leq k < T_1} \widehat{E} \left((M^*_{k,T_1} - 2(\widehat{\xi}^*_{k,T_1})^2)^2 \middle| (R_1, ..., R_T) \right), \tag{2.13}$$

where $M_{k,T},\ k < T$, is a second order moment estimator of the form:

$$M_{k,T} = \frac{1}{k} \sum_{i=1}^{k} (\ln R_{i,T} - \ln R_{k+1,T})^2.$$

The estimate for the optimal overall sample fraction k_{opt} can then, most simply as in Hall (1990) and Danielsson and de Vries (1997), be derived from:

$$\widehat{k}_{opt} = \left[\widehat{k}_1 (T/T_1)^{2\rho/(2\rho+1)} \right]. \tag{2.14}$$

Again, the second order parameter ρ is either e.g. estimated by equation (2.11) or set equal to a constant.

2.3.2. Hill Plot based Estimation and Optimal Scaling

A different class of Hill-based adaptive tail index estimation approaches is based on a series of Hill estimates

$$\{(k, \widehat{\xi}_{k,T}) : 1 \leq k \leq T\}, \tag{2.15}$$

also denoted as *Hill plot*. We apply the consistent estimator by Drees et al. (2000) which is based on the frequent recommendation of choosing an estimate of the tail index within a "stable" region of the Hill plot. Rescaling the axis of the number of upper order statistics k by a continuous parameter φ, yields the so-called *alternative Hill plot* proposed by Resnick and Stărică (1997)

$$\{(\varphi, \widehat{\xi}_{[T^\varphi],T}) : 0 \leq \varphi \leq 1\}. \tag{2.16}$$

[3] In case the maximum sample realization is drawn at least twice in a bootstrap resample, the MSE equivalent criterion yields a zero value for $k = 1$ (and possibly for $k > 1$), whereas the MSE criterion yields the squared initial estimator. We enter the squared initial estimator value in this case.

When the underlying distribution is not exactly Paretian, Drees et al. show that superior properties of the estimator result under the alternative Hill plot (2.16) as compared to the standard plot (2.15); hence we use the corresponding estimator

$$\widehat{\xi}_{m,T} = \arg\max_{\xi>0} \int_0^{\overline{\varphi}} \mathbf{1}_{\{[T^\varphi]^{1/2}\left|\widehat{\xi}_{[T^\varphi],T}-\xi\right|\leq m\widetilde{\xi}_T\}} d\varphi, \tag{2.17}$$

with an upper boundary $0 < \overline{\varphi} < 1$ and a scaling constant $m > 0$. The parameter $\widetilde{\xi}_T$ denotes a consistent initial Hill estimator which is used as an estimate of the standard deviation of the alternative Hill plot, $[T^\varphi]^{-1/2}\xi$, for varying φ; see also (2.9).

The estimator defined by (2.17) is denoted *maximal occupation time (MOT) estimator*. Thereby occupation in a neighborhood of some tail index value is defined by an interval, which is constructed from the standard deviation of the Hill estimator times the scaling constant m. The sensitivity of the tail index estimator (2.17) with respect to m can be visualized by what may be labelled a *MOT plot*

$$\{(m, \widehat{\xi}_{m,T}) : 0 < m < \infty\}. \tag{2.18}$$

As we are facing a bias/variance trade-off in (2.9), an optimal choice of m will not only consider estimation variance but also estimation bias. We therefore propose a MSE-based method for choosing m given model estimates which provide information about the underlying distribution function F.

In particular, assume that the marginal distribution function F is given by the dynamics of a GARCH(1,1) model according to (2.4) where the stationarity condition (2.5) holds. Given the parameters θ of the GARCH model, let the variable $\overline{\xi}_T > 0$ denote the tail index which follows as a solution to $I(\xi) = 0$ according to equation (2.6). A MSE-optimal scaling constant m conditional on the model point estimates is then given as:

$$\overline{m} = \arg\min_{m\in\mathcal{G}} \widehat{E}\left((\widehat{\xi}_{m,T} - \overline{\xi}_T)^2 \middle| \widehat{\theta}_T \right). \tag{2.19}$$

In applications, the optimal $m > 0$ can be determined by a grid search within a set $\mathcal{G} \subset \mathbb{R}^+$ and a set of discrete values for the parameter φ in (2.17). The expectation operator is approximated by a given number of independent Monte Carlo simulation runs. Based on the MSE-optimal scaling constant, $\widehat{\xi}_{\overline{m},T}$ serves as a *model-based optimal MOT estimator* of the tail index.

3. Small Sample Performance of Tail Index Estimators

In the previous section we gave a summary of theoretical results on the Hill estimator showing that the estimator has favorable asymptotic properties under quite general model assumptions. However, there remain open questions in small samples where the asymptotic results may only give a crude approximation to the estimator's true sampling properties. In particular, little is known about how ARCH-type dependence influences the bias component and the bias/variance trade-off under adaptive subsample selection criteria. In this section, we perform a simulation study on the small sample performance of the Hill estimators outlined above.

Previous simulation studies e.g. by Gomes and Oliveira (2001) and Matthys and Beirlant (2000) compare different adaptive estimation procedures under iid variables. They document the small sample properties of the procedures pointing out small sample bias. Turning to dependent series, the special case of an integrated GARCH (IGARCH) process which implies infinite unconditional second moments is discussed in Kearns and Pagan (1997). Huisman et al. (2001) give simulation results on their estimator under a GARCH(1,1) model with typical parameterization and normal innovations.

The general class of "nearly integrated", stationary GARCH(1,1) models with student-t innovations is studied in the following. Table 3.1 contains an overview of the return models for our simulation study. The models include the iid symmetric student-t(ν) return model ("stud/ξ") as well as the GARCH(1,1) return generating process (2.4) with student-t(ν) innovations ("arch/ξ"). The tail indices ξ for the latter model are calculated by solving equation (2.6). Note that the model parameters are chosen such that the tail indices of the marginal distributions are approximately equal to those of the iid student models. When studying the properties of the tail index estimators, this allows us to distinguish between the effects of the magnitude of the tail index on the one hand and ARCH-type dependence on the other hand. Referring to empirical evidence from GARCH estimation in finance, the parameters β_1 and β_2 are chosen such that $\beta_1 + \beta_2 \lessapprox 1$, i.e. the models are what may be called "nearly integrated".

3.1. Estimator Settings and Definitions

In order to characterize the small sample properties of the various adaptive Hill estimators presented above, we study the estimators' error distributions and par-

Table 3.1
Student-t models, model parameters and the corresponding tail index of the stationary marginal distribution.

Parameter Model:	Label:	ξ	$1/\xi$	ν	β_0	β_1	β_2
iid student-t(ν)	stud/0.17	0.17	6	6	—	—	—
	stud/0.25	0.25	4	4	—	—	—
	stud/0.33	0.33	3	3	—	—	—
GARCH(1,1)-t(ν)	arch/0.17	0.17	6	9	10^{-6}	0.05	0.92
	arch/0.25	0.25	4	5	10^{-6}	0.03	0.94
	arch/0.33	0.33	3	4	10^{-6}	0.03	0.93

ticularly their root mean squared error in a series of Monte Carlo simulations. Without loss of generality, the sample size T will be defined as the number of positive sample observations from the upper distribution tail. We exploit the property that all our return models are based on distribution functions which are symmetric around zero by simulating T observations and taking absolute values. The corresponding overall size of a sample of observed returns will then equal T if one assumes symmetry or $N = 2T$ if one is interested in inferences about each of both tails separately.

The adaptive estimators that were outlined in Section 2.3 above require the choice of several variables, where we rely on the suggestions made in the literature. Altogether, four different estimators are calculated: *H-INI*, *H-DK*, *H-BS*, and *H-MOT*. The initial estimator, $\widetilde{\xi}_T = \widehat{\xi}_{[2\sqrt{T}],T}$, is given as a naive estimator labelled *H-INI*. Following Drees and Kaufmann (1998), the second order parameter ρ in the asymptotic tail expansion (2.10) is estimated by (2.11) with $\lambda = 0.6$, giving the estimator *H-DK*. For the bootstrap approach we choose the subsample sizes $T_1 = [T/10]$ as in Danielsson and de Vries (1997). The asymptotically MSE-equivalent criterion proposed by Danielsson et al. (2001) is chosen, which yields an optimal subsample fraction according to (2.13). Handling the second order parameter ρ as in the Drees/Kaufmann approach, (2.14) yields the estimated optimal sample fraction and the estimator *H-BS*. For the Drees et al. (2000) maximum occupation time estimator (2.17), the continuous parameter φ is approximated by a grid with stepsize 0.05, where the upper bound $\overline{\varphi} = \ln([T/2])/\ln(T)$ corresponds to $k \leq [T/2]$. Together with the standard scaling constant $m = 1$, this defines the estimate *H-MOT*.

3.2. Simulation Results

Given the six return models from Table 3.1, our simulations are based on time series samples of two different sizes: (i) a shorter sample with $T = 500$ upper tail observations (corresponds to $N = 1000$ observations) and (ii) a longer sample with $T = 1500$ upper tail observations (corresponds to $N = 3000$ observations). We report mean error (ME), standard deviation (STD) and root mean squared error (RMSE) where all our simulation results are based on 500 independent simulation runs, under each of which 100 bootstrap runs are performed for calculating the bootstrap estimator *H-BS*. The results for $T = 500$ are given in Table 3.2, the results for $T = 1500$ are given in Table 3.3. Our conclusions from the simulations are given in the following subsections.

3.2.1. Magnitude of the Tail Index and Estimation Precision

From the result on the asymptotic sampling distribution of the Hill estimator it is known that the asymptotic bound to its standard deviation, $\xi/\sqrt{k(T)}$, is linearly increasing in ξ. Interestingly however, we observe in Tables 3.2 and 3.3 that higher degrees of fat-tailedness (i.e. larger values of ξ) frequently show quite stable or even decreased RMSE statistics for the estimators.[4] This decrease in RMSE with an increase in ξ deserves further consideration recalling that $\text{MSE} = \text{STD}^2 + \text{ME}^2$. With, for example $k(T) = [2\sqrt{T}]$ and sample size $T = 1500$, the asymptotic standard deviations are 0.019 for $\xi = 0.17$, 0.028 for $\xi = 0.25$ and 0.038 for $\xi = 0.33$, respectively. In Table 3.3, the corresponding simulated standard deviations for the initial estimator *H-INI* are 0.028, 0.033 and 0.041 respectively; i.e. particularly under the smaller tail indices, the standard deviations have not yet reached their asymptotic bounds. The observation that the RMSE remains stable or even decreases for increasing tail index ξ is therefore partly explainable by a slower convergence to the asymptotic standard deviations for models with small tail index. The main explanation of our findings, however, is the behavior of

[4]For example, in Table 3.2, the RMSE for the bootstrap estimator *H-BS* is 0.080 for the student-t model with $\xi = 0.17$ and 0.071 for the student-t model with $\xi = 0.33$. The RMSE for the initial Hill estimator *H-INI* is 0.091 for the student-t model with $\xi = 0.17$ and declines by roughly a third to 0.063 when the tail index is nearly doubled to $\xi = 0.33$.

Table 3.2

Small sample properties of the Hill estimators: mean error (ME), standard deviation (STD), and root mean squared error (RMSE), $T = 500$

	H-INI	*H-DK*	*H-BS*	*H-MOT*
	ME STD RMSE	ME STD RMSE	ME STD RMSE	ME STD RMSE
Model:				
stud/0.17	0.12 0.039 0.12	0.10 0.061 0.12	0.081 0.062 0.10	0.074 0.059 0.095
stud/0.25	0.093 0.050 0.11	0.092 0.068 0.11	0.072 0.074 0.10	0.066 0.068 0.094
stud/0.33	0.072 0.057 0.092	0.079 0.072 0.11	0.057 0.083 0.10	0.038 0.069 0.079
arch/0.17	0.14 0.055 0.15	0.12 0.084 0.15	0.096 0.083 0.13	0.081 0.070 0.11
arch/0.25	0.086 0.055 0.10	0.074 0.079 0.11	0.046 0.081 0.093	0.036 0.069 0.078
arch/0.33	0.049 0.065 0.081	0.038 0.098 0.11	0.015 0.093 0.094	−0.0097 0.072 0.072

Table 3.3

Small sample properties of the Hill estimators: mean error (ME), standard deviation (STD), and root mean squared error (RMSE), $T = 1500$

	H-INI	*H-DK*	*H-BS*	*H-MOT*
	ME STD RMSE	ME STD RMSE	ME STD RMSE	ME STD RMSE
Model:				
stud/0.17	0.087 0.028 0.091	0.080 0.041 0.090	0.067 0.047 0.082	0.053 0.046 0.070
stud/0.25	0.064 0.033 0.072	0.067 0.045 0.081	0.052 0.056 0.077	0.039 0.058 0.070
stud/0.33	0.048 0.041 0.063	0.060 0.047 0.076	0.048 0.056 0.074	0.030 0.064 0.070
arch/0.17	0.10 0.043 0.11	0.090 0.063 0.11	0.069 0.069 0.098	0.049 0.058 0.075
arch/0.25	0.069 0.052 0.086	0.060 0.069 0.092	0.041 0.078 0.088	0.016 0.066 0.067
arch/0.33	0.030 0.062 0.069	0.027 0.087 0.092	0.0001 0.098 0.098	−0.030 0.071 0.077

the sample bias (ME). A decreasing absolute bias tends to more than compensate for increasing estimation variance.[5]

In summary, these results indicate that the bias/variance trade-off is complex with small sample bias playing an important role in overall small sample estimation precision. While estimation STD increases for larger values of the tail index, this is not necessarily the case for estimation MSE.

3.2.2. GARCH versus iid Student Models

In comparing the results for the iid student-t and GARCH models, a first glance at the tables may lead to the conclusion that there is a relatively high overall mean squared error-robustness of the tail index estimators with respect to heteroskedasticity. Again, as in the case of increases in the magnitude of the tail index ξ in the preceding section, unchanged or even improved RMSE results under GARCH come from a compensating effect of a reduction in estimation bias ME. ARCH-type volatility clustering leads to clustering in the extremes which increases estimation error for the tail index. This is confirmed by our results when we look at the STD statistics of the estimators which mostly show a notable increase in estimation variance under GARCH.

An interesting finding in the tables is that the MOT estimator's sample variance is relatively large under iid student-t data, but is hardly affected by ARCH-type dependence. This robustness characteristic under dependence appears to be due to the construction of the estimator based on an entire series of Hill estimators.

3.2.3. Increasing the Sample Size

Comparing the results for the sample size $T = 500$ in Table 3.2 with those for the sample size $T = 1500$ in Table 3.3 indicates that the relative performance ranking of the estimation approaches under the student-t and GARCH models is basically unaffected by an increase in the sample size; *M-MOT* and *M-BS* achieve the best RMSE results irrespective of sample size. Apart from that, the results suggest that an increased sample size yields larger performance improvements for the bootstrap as compared to the maximum occupation time estimator.

[5] For example, in the case mentioned in the previous footnote, when RMSE falls from 0.091 to 0.063, we see that the bias falls from 0.087 to 0.048, more than offsetting the increase in standard deviation from 0.028 to 0.041.

Considering estimation precision as measured by STD, an increase in the sample size causes improvements in the estimators' standard deviations, especially under the iid student-t model with a small tail index. Interestingly, there is a notable difference in the reduction of estimation variance between the GARCH and the iid models. Under the GARCH models with $\xi = 0.25$ and $\xi = 0.33$ the standard deviation of the Hill estimator can hardly be improved by increasing the sample size. This shows that the convergence to the asymptotic lower bounds for the estimator's standard deviation can be very slow under ARCH-type dependence in the data. Hence, large sample sizes may be considered a necessary although not sufficient condition for a successful application of tail estimation procedures.

3.2.4. Comparing the Estimation Approaches

Although, it is impossible to conclude from our simulation results that a particular estimator under consideration is dominant, the bootstrap and the maximum occupation time approach seem to be appropriate choices for parameter estimation in small samples. The *H-BS* shows advantages under the increased sample size. The *H-MOT* estimation approach is less sensitive with respect to sample size and ARCH-type dependence.[6]

Interestingly, it turns out that all the approaches show a common tendency to be positively biased for smaller values of the tail index. Under the GARCH models there is also a tendency towards negative bias for larger values of the tail index especially for the MOT estimator.[7] Hence, bias as a function of ξ appears to be a critical issue in tail index estimation in that it causes the estimated degree of fat-tailedness to be systematically biased. In order to reduce such potential small sample bias, the model-based optimal MOT estimator defined by equations (2.17) and (2.19) will be applied to exchange rate data in the following. Inferring a priori knowledge about the approximate magnitude of the tail index, we thereby simulate the small sample properties of our estimation approach under GARCH.

[6] The simulation results in Wagner and Marsh (2000) indicate that slightly dominant overall small sample results can be obtained by replacing the Hill estimator with a Hill-derived moment ratio estimator. A detailed discussion is beyond the scope of this paper.

[7] The simulation results for a GARCH model with $\xi > 0.33$ in the following section confirm this conclusion.

4. Empirical Study of Extreme Exchange Rate Changes

The preceding simulations made obvious that the small sample properties of the Hill estimator depend critically on our assumption of the underlying return model. In particular, not only estimation variance but also bias depend on the magnitude of the tail index both being also affected by potential ARCH-effects.

In this section, we study these effects for a given dataset of daily changes in the exchange rate of one US-dollar versus Deutschemarks (USD/DEM). We proceed in three steps. First, based on the prior assumption that a GARCH(1,1)-t(ν) specification is a suitable model for the given exchange rate dynamics, GARCH model parameter estimation is carried out. Second, we perform a simulation study of tail index estimation small sample properties under the specific model and derive the model-based optimal MOT estimator. Finally, the tail index of the marginal distribution of our exchange rate changes is estimated. The three-step procedure allows us to calibrate the MOT estimator under the assumed model while maintaining a general approach to tail index estimation based on extreme value theory. The estimation results obtained are compared to those of a bootstrap approach.

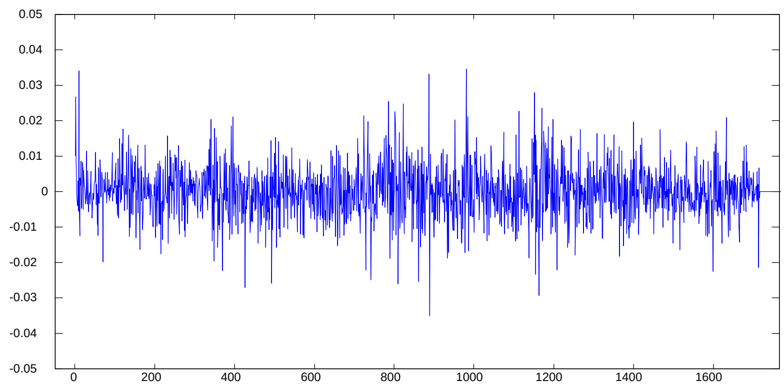

Figure 4.1: Plot of the logarithmic USD/DEM exchange rate changes r_t, $t = 1, ..., 1716$; January 2, 1988 to December 30, 1994.

4.1. The Data and GARCH Model Estimation

The empirical investigation is based on logarithmic changes r_t, $t = 1, ..., T$, in daily USD/DEM exchange rates during the period 1988 to 1994. This series contains $T = 1716$ non-zero observations as obtained from Datastream. The pre-EURO

Table 4.1

MLE point estimates of GARCH(1,1)-t(ν) model parameters for the exchange rate changes, $T = 1716$; start values were obtained from fitting the model under the assumption of normally distributed innovations; standard errors in parenthesis.

Parameter	$\overline{\xi}_T$	ν	β_0	β_1	β_2
GARCH(1,1)-t(ν)	0.41	6.01	$1.28 \cdot 10^{-6}$	0.0410	0.937
		(0.89)	$(6.17 \cdot 10^{-7})$	(0.0108)	(0.0184)

sample period was chosen as a period of exchange rates without severe exogenous policy impact or structural breaks which makes the assumption of a stationary underlying distribution plausible. Figure 4.1 plots the logarithmic exchange rates changes in the sample period.

Estimation of the GARCH model with student-t innovations (2.4) is carried out by maximum likelihood as outlined in Hamilton (1994), where the parameter vector is given by $\theta = (\beta_0, \beta_1, \beta_2, \nu)$. The parameter estimation results are given in Table 4.1. As typically observed for GARCH estimation in finance, all estimates differ from zero at high confidence levels and the sum of the point estimates is close to, but smaller than, one. The maximum likelihood estimate of ν turns out to be close to six.[8]

4.2. Small Sample Error Simulation

The simulation of the sample properties of the Hill estimators is carried out for the estimated GARCH(1,1)-t(ν) model. Given our above point estimates and setting $\nu = 6$, the stationarity condition (2.5) for the GARCH model can be verified numerically. Solving the integral equation (2.6) results in a theoretical model tail index $\overline{\xi}_T = 0.41$. This theoretical tail index gives an approximate prior estimate of the tail index of the marginal distribution of the exchange rate changes. Under a sample size of $T = 1716$ we then simulate the error distributions

[8] GARCH model estimation diagnostics indicate a good fit of the GARCH model; these include QQ-plots as well as autocorrelation, partial autocorrelation and Ljung-Box statistics for the model residuals and the squared model residuals; note that in the latter case standard confidence levels are conservative under nonexistence of the forth moment of the underlying return distribution. Detailed results are available from the authors upon request.

Table 4.2

Small sample properties of Hill estimators under the estimated GARCH(1,1)-t(ν) model: mean error (ME), standard deviation (STD), and root mean squared error (RMSE), $T = 1716$; *H-BS* with 100 bootstrap runs and second order parameter set equal to one; *H-MOT* with $\overline{m}$ obtained from the grid $\{1, 1.5, ..., 10\}$ under 250 simulation runs in (2.19).

	H-INI	*H-BS*	*H-MOT*			
	ME STD RMSE	ME STD RMSE	ME STD RMSE	ME STD RMSE	ME STD RMSE	ME STD RMSE
Model:			$m = 1$	$m = 3$	$\overline{m} = 5.5$	$m = 10$
arch/0.41	−0.10 0.052 0.12	−0.15 0.073 0.17	−0.17 0.061 0.18	−0.077 0.059 0.096	−0.035 0.040 0.054	−0.059 0.043 0.076

of the bootstrap estimator *H-BS* and the maximum occupation time estimator *MOT*.[9] Additionally, we determine the optimal scaling factor $\overline{m}$ for the optimal *H-MOT* estimator by minimizing (2.19) via a grid search. The grid for varying m is thereby set as $\mathcal{G} = \{1, 1.5, ..., 10\}$ where 250 Monte Carlo runs approximate MSE in (2.19) for each grid element.

As in Section 3, Table 4.2 reports the mean error (ME), the standard deviation of the errors (STD) and the square root of the mean squared errors (RMSE) of the estimation approaches under 500 independent Monte Carlo runs. The simulation results reveal a negative bias for all estimation approaches; i.e. all estimators underestimate tail thickness for the given GARCH(1,1)-t(ν) model. The maximum occupation time estimator with MSE-optimal scaling factor $\overline{m} = 5.5$ substantially reduces bias under the simulated model.

4.3. Estimation Results

We now turn back to our sample of exchange rate changes and examine the tail behavior for the absolute changes $|r_t|$, $t = 1, ..., 1716$. A classical Hill plot (2.15)

[9] All estimator and simulation settings remain unchanged from Section 3. Only for computational simplification, the second order parameter ρ was set equal to one in simulating the bootstrap estimation approach. Simulation results unreported here indicate equivalent small sample performance. Refer to Wagner and Marsh (2000) for details.

for the series is given in Figure 4.2. As is well-known, time series dependence strongly influences the properties of the Hill plot. Obviously the plot in Figure 4.2 does not converge against some limit as k becomes larger and positive bias enters in the estimates of ξ. As our simulation results indicate, the Hill estimates may also exhibit bias in that the entire Hill plot is shifted up- or downward which may well be due to the non-linear dependence in the data.[10] Apart from these difficulties, visual inspection of the plot reveals a nearly flat area of Hill estimates with moderate variability in an interval roughly between $k = 60$ and $k = 90$. A classical approach would be to choose an estimator from that range which would imply a tail index estimate in the interval $[0.25; 0.27]$. A glance at the alternative Hill plot (2.16) in Figure 4.3 indicates that rescaling the Hill plot for small values of k reveals a flat area for $\varphi \in [0.2; 0.4]$ which relates to $k \in [4; 20]$. Note that choosing $k = 20$ would yield a tail index estimate of about 0.18.

The above discussion makes obvious that it seems quite difficult to decide on the choice of a subsample fraction for the given dataset via graphical methods alone. The results for the adaptive, i.e. data-driven, Hill estimation results are summarized in Table 4.3. The bootstrap estimator *H-BS* yields results similar to our visual inspection for the standard Hill plot. The choice of an optimal subsample fraction $k = 84$ yields an estimate of 0.25; a result which happens to be nearly identical to that by the *H-INI* estimator. The maximum occupation time estimators' behavior corresponds well to our previous simulation results: *H-MOT* with $m = 10$ for example, yields a larger estimate of a ξ than *H-MOT* with $m = 3$ which again is larger than the estimate by *H-MOT* with $m = 1$. This is due to the observation that, as k becomes larger, positive bias enters in the Hill plot. The application of our model-based MSE-optimal MOT estimator yields a tail index estimate of 0.35 indicating a 40 percent deviation in the estimate of tail-thickness as compared to the results mentioned above. For a sensitivity analysis of our result we refer to Figure 4.4. The given MOT plot appears to be quite stable for a wide range of scaling factors $m \in [2; 8]$ including our choice of $\overline{m}$. All these scaling factors yield tail index estimates equal to or, mostly, substantially *larger than* 0.25.

When using estimates of $1/\xi$ as an indication of the existence of higher order moments, one should therefore rather doubt the existence of the forth moment of the underlying distribution F. Once we conclude that the tail index for the exchange rate changes is potentially larger than 0.33, we would even doubt the existence of the third moment of the underlying distribution, although an inter-

[10] See also the discussion e.g. in Embrechts et al. (1997), p. 343-344.

pretation of traditional results would not yield such implication.

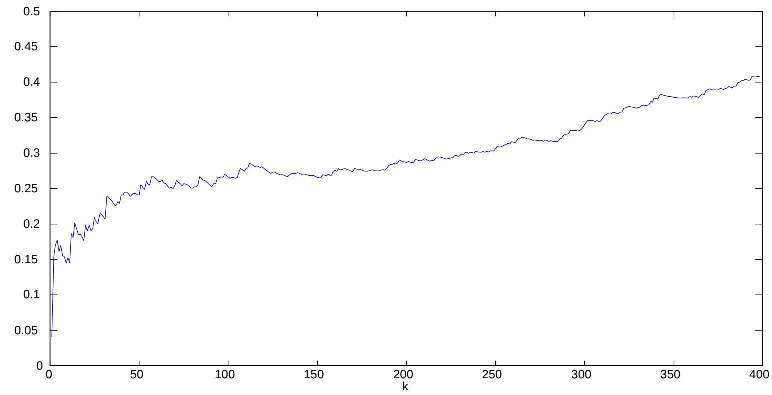

Figure 4.2: Hill plot, $\{(k, \widehat{\xi}_{k,T}) : 1 < k \leq 400\}$, for the sample of exchange rate changes, $T = 1716$.

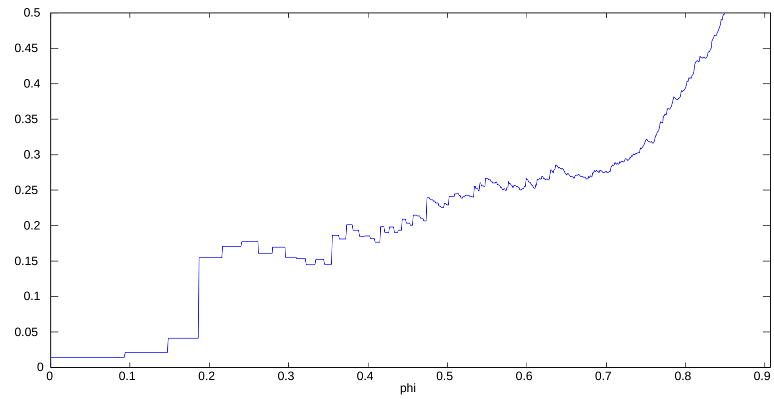

Figure 4.3: Alternative Hill plot, $\{(\varphi, \widehat{\xi}_{[T\varphi],T}) : 0 < \varphi \leq 0.9\}$, for the sample of exchange rate changes, $T = 1716$.

Table 4.3

Tail index estimation results for the exchange rate changes, symmetric tail estimated from $|R_t|$: naive estimator *H-INI*, *H-BS* with 100 bootstrap runs and second order parameter set equal to one, *H-MOT* with different scaling parameters m. Estimated asymptotic standard errors in parenthesis.

	H-INI	*H-BS*	*H-MOT*			
	$k = 82$	$k = 84$	$m = 1$	$m = 3$	$\overline{m} = 5.5$	$m = 10$
$\widehat{\xi}$	0.25 (0.028)	0.25 (0.028)	0.18 –	0.30 –	0.35 –	0.31 –

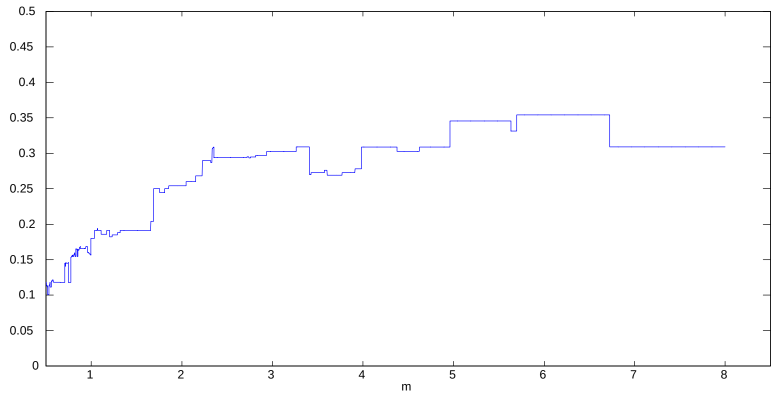

Figure 4.4: MOT plot, $\{(m, \widehat{\xi}_{m,T}) : 0 < m \leq 8\}$, for the sample of exchange rate changes, $T = 1716$.

Our findings also have implications for high quantile estimation: Under tail symmetry, which was assumed throughout, let the p-percent quantile r_p be defined by $P(|R_t| \leq r_p) = p$. Then, under the Pareto tail, the corresponding quantile estimate is (e.g. Embrechts et al. 1997)

$$\widehat{r}_p = \left(\frac{T+c}{k}(1-p)\right)^{-\widehat{\xi}} |R_t|_{k,T},$$

where the sample size T of upper tail observations is corrected by the number $c > 0$ of zero returns in the sample. $|R_t|_{k,T}$ denotes the kth upper order statistic from the absolute exchange rate changes. For the subsample size we choose $k = 84$ as given by *H-BS* (which could be replaced by other choices); i.e. we model a fraction

Table 4.4

Quantile estimation results $\widehat{r}_p \cdot 100$ for the exchange rate changes: $k = 84$, alternative probability levels and tail index estimates.

	$\widehat{\xi}$	0.25	0.30	0.35
$p = 0.95$		1.41	1.39	1.37
$p = 0.99$		2.10	2.25	2.41
$p = 0.999$		3.74	4.49	5.39

of approximately five percent of the largest absolute exchange rate changes. Table 4.4 reports quantile estimates $\widehat{r}_p$ for different probabilities and tail index estimates. The quantile estimates indicate that the 95 percent quantile estimate is robust under differences in estimates of the tail index: the five in a hundred days quantile is estimated to equal roughly 1.4 percent. Moving to the higher probability levels indicates economically significant differences in estimated quantiles: under a tail estimate of 0.35 the one in a hundred days estimated quantile increases from a 2.1 percent change to a 2.4 percent change, the one in a thousand days estimated quantile increases from a 3.7 percent change to a 5.4 percent change in asset value.

5. Conclusion

This paper focused on measuring tail thickness of the marginal distribution of stationary ARCH-type financial time series. By the very nature of extreme value theory, small sample bias and variance play an important role in estimation. Our simulation study documents that measuring tail thickness relies on proper assumptions about the underlying return model and its time series properties; i.e. what tail index estimation results tell us about the extreme behavior of the returns cannot be separated from a priori assumptions. An example is the interpretation of classical methods such as the Hill plot which should be treated with care given the dependence structure in financial data.

Hence, while the extreme value theory offers a general methodology for specifically modeling the distribution tails, suitable modeling of the overall data remains essential in practical applications where a limited amount of independent data is available. One practical approach was outlined here. A drawback of the approach is that it is relatively cumbersome, since it requires a pre-analysis of the financial time series which may even be extended by additional exploratory methods. It follows that the estimation results will not be readily available; once they are available, they are subject to prior judgement. Further integration of models of the overall time series with models of the tails seems an interesting issue to address in future work on extremes in finance.

References

[1] **Bollerslev, T., Chou, R. Y., Kroner, K. F. (1992)**: ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence, *Journal of Econometrics* 52, 5-59

[2] **Bond, S. A. (2000)**: Asymmetry and Downside Risk in Foreign Exchange Markets, Working Paper, University of Cambridge

[3] **Clark, P. K. (1973)**: A Subordinate Stochastic Process Model with Finite Variance for Speculative Prices, *Econometrica* 41: 135-155

[4] **Coles, S. (2001)**: An Introduction to Statistical Modeling of Extreme Values, Springer, London

[5] **Dacorogna, M. M., Müller, U. A., Pictet, O. V., de Vries, C. G. (1995)**: The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets, O&A Preprint, Zürich

[6] **Danielsson, J., de Haan, L., Peng, L., de Vries, C. G. (2001)**: Using a Bootstrap Method to Choose the Sample Fraction in Tail Index Estimation, *Journal of Multivariate Analysis* 76: 226-248

[7] **Danielsson, J., de Vries, C. G. (1997)**: Tail Index and Quantile Estimation with Very High Frequency Data, *Journal of Empirical Finance* 4: 241-257

[8] **Diebold, F. X. (1988)**: Empirical Modeling of Exchange Rate Dynamics, Springer, New York

[9] **Diebold, F. X., Schuermann, T., Stroughair, J. D. (1998)**: Pitfalls and Opportunities in the Use of Extreme Value Theory in Risk Management, in: **Refenes, A.-P. N., Burgess, A. N., Moody, J. E. (eds.)**: Decision Technologies for Computational Finance, Kluwer, Dordrecht, pp. 3-12

[10] **Drees, H. (1998)**: Optimal Rates of Convergence for Estimates of the Extreme Value Index, *Annals of Statistics* 26: 434-448

[11] **Drees, H., de Haan, L., Resnick, S. (2000)**: How to Make a Hill Plot, *Annals of Statistics* 28: 254-274

[12] **Drees, H., Kaufmann, E. (1998)**: Selecting the Optimal Sample Fraction in Univariate Extreme Value Estimation, *Stochastic Processes and their Application* 75: 149-172

[13] **Embrechts, P., Klüppelberg, C., Mikosch, T. (1997)**: Modelling Extremal Events for Insurance and Finance, Springer, New York

[14] **Gomes, M. I., Oliveira, O. (2001)**: The Bootstrap Methodology in Statistics of Extremes–Choice of the Optimal Sample Fraction, *Extremes* 4: 331-358

[15] **de Haan, L., Peng, L. (1998)**: Comparison of Tail Index Estimators, *Statistica Neerlandica* 52: 60-70

[16] **de Haan, L., Resnick, S., Rootzén, H., de Vries, C. G. (1989)**: Extremal Behavior of Solutions to a Stochastic Difference Equation with Applications to ARCH Processes, *Stochastic Processes and their Application* 32: 213-224

[17] **Hall, P. (1990)**: Using the Bootstrap to Estimate Mean Squared Error and Select Smoothing Parameters in Nonparametric Problems, *Journal of Multivariate Analysis* 32: 177-203

[18] **Hamilton, J. D. (1994)**: Time Series Analysis, Princeton University Press, Princeton

[19] **Hill, B. M. (1975)**: A Simple General Approach to Inference about the Tail of a Distribution, *Annals of Statistics* 3: 1163-1174

[20] **Hsing, T. (1991)**: On Tail Index Estimation using Dependent Data, *Annals of Statistics* 19: 1547-1569

[21] **Huisman, R., Koedijk, K. G., Kool, C., Palm, F. C. (2001)**: Tail-Index Estimates in Small Samples, *Journal of Business and Economic Statistics* 19: 208-216

[22] **Huisman, R., Koedijk, K. G. Pownall, R. (1998)**: Fat Tails in Financial Risk Management, *Journal of Risk* 1: 47-62

[23] **Jondeau, E., Rockinger, M. (1999)**: The Tail Behavior of Stock Returns: Emerging versus Mature Markets, Working Paper, HEC-School of Management

[24] **Kearns, P. Pagan, A. (1997)**: Estimating the Density Tail Index for Financial Time Series, *Review of Economics and Statistics* 79: 171-175

[25] **Koedijk, K. G., Schafgans, M. M., de Vries, C. G. (1990)**: The Tail Index of Exchange Rate Returns, *Journal of International Economics* 29: 93-108

[26] **Lauridsen, S. (2000)**: Estimation of Value at Risk by Extreme Value Methods, *Extremes* 3: 107-144

[27] **Loretan, M., Phillips, P. C. B. (1994)**: Testing the Covariance Structure of Heavy-Tailed Time Series, *Journal of Empirical Finance* 1: 211-248

[28] **Matthys, G., Beirlant, J. (2000)**: Adaptive Threshold Selection in Tail Index Estimation, in: **Embrechts, P. (ed.)**: Extremes and Integrated Risk Management, Risk Books, London, pp. 37-49

[29] **McNeil, A. J., Frey, R. (2000)**: Estimation of Tail-Related Measures for Heteroscedastic Financial Time Series: An Extreme Value Approach, *Journal of Empirical Finance* 7: 271-300

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

[31] **Resnick, S., Stărică, C. (1997)**: Smoothing the Hill Estimator, *Advances in Applied Probability* 29: 271-293

[32] **Resnick, S., Stărică, C. (1998)**: Tail Index Estimation for Dependent Data, *Annals of Applied Probability* 8: 1156-1183

[33] **Segers, J. (2001)**: Abelian and Tauberian Theorems on the Bias of the Hill Estimator, Working Paper, University of Leuven

[34] **Wagner, N., Marsh, T. A. (2000)**: On Adaptive Tail Index Estimation for Financial Return Models, Working Paper No. RPF-295, U.C. Berkeley