Estimation of Tail-Related Value at Risk Measures: Range-based Extreme Value Approach

Heng-Chih Chou
Professor, Department of Shipping and Transportation Management, National Taiwan Ocean University
2, Beining Road, Keelung 202, Taiwan
Phone: +886-2-24622192 Fax: +886-2-24631903
Email: hechou@mail.ntou.edu.tw

David Wang
Assistant Professor, Department of Finance, Chung Yuan Christian University
200, Chung Pei Road, Chung Li 320, Taiwan
Phone: +886-3-2655707 Fax: +886-3-2655749
Email: dwang@cycu.edu.tw

First version: February 2009; this version: Jan. 2011

Abstract

This study proposes a new approach for estimating value at risk (VaR). This approach combines quasi-maximum-likelihood fitting of asymmetric conditional autoregressive range (ACARR) models to estimate the current volatility and classical extreme value theory (EVT) to estimate the tail of the innovation distribution of the ACARR model. The proposed approach reflects two stylized facts exhibited by most financial time series: stochastic volatility and the fat-tailedness of conditional distributions over short time horizons. This approach presents two main advantages over the McNeil and Frey (2000) approach. First, the ACARR model in this approach is an asymmetric model that treats the upward and downward movements of the asset price asymmetrically, whereas the generalized autoregressive conditional heteroskedasticity (GARCH) model in the McNeil and Frey (2000) approach is a symmetric model that ignores the asymmetric structure of the asset price. Second, the proposed method uses classical EVT to estimate the tail of the distribution of the residuals to avoid the threshold issue in the modern EVT model. Since the McNeil and Frey (2000) approach uses modern EVT, it may estimate the tail of the innovation distribution poorly. Back testing of historical time series data shows that our approach gives better VaR estimates than the McNeil and Frey (2000) approach.

JEL classifications: C22; G10; G21
Keywords: Risk management; Value at risk (VaR); Asymmetric conditional autoregressive range (ACARR) model; Extreme value theory (EVT); S&P 500 index futures

1. Introduction

Financial markets around the world have experienced significant instabilities in recent years. This has led to numerous criticisms about existing risk management systems, and has
motivated the search for more appropriate approaches that are able to cope with rare events that have severe consequences.

It is often difficult to model rare phenomena that lie outside the range of available observations. In this case, it is necessary to rely on a well-established approach. Extreme value theory (EVT) provides a firm theoretical foundation on which to build statistical models describing extreme events.


This paper deals with the behavior of the tails of financial time series. More specifically, we propose a new approach that uses EVT to compute the primary tail risk measure, the value at risk (VaR). The rest of this paper is organized as follows. Section 2 discusses the motivation for this study. Section 3 presents the proposed approach. Section 4 describes the data used to evaluate the empirical performance of this approach. Section 5 provides the empirical results, while Section 6 concludes the paper.

2. Motivation

Broadly speaking, there are two types of extreme value models to estimate the value at risk (VaR) of financial assets. The oldest (classical) models are the block maxima models; these models fit the generalized extreme value (GEV) distribution to the extreme (maxima or minima) observations collected from large samples of identically distributed observations. The GEV distribution reflects the behavior of very high profits (in case of maxima) and very high losses (in case of minima) from the portfolio.

The second, and more modern, group of models includes the peaks-over-threshold models; these models attempt to estimate the tails of the underlying return distribution rather than model the distribution of extremes like the block maxima models. The peaks-over-threshold models identify a threshold to define the starting point of the tail of the return distribution before estimating the distribution of the excesses (i.e., the return observations that exceed the threshold).

There are two approaches to estimating the excess distribution: the semi-parametric model based on the Hill estimator (Danielsson and de Vries, 1997), and the fully parametric model based on the generalized Pareto distribution (GPD) (McNeil and Frey, 2000). McNeil and Frey’s (2000) simulation study shows that the GPD-based VaR estimate is more stable (in

1 VaR is generally defined as the capital sufficient to cover, in most instances, losses from a portfolio over a holding period of a fixed number of days. The Basle accord proposes a 1% VaR over a 10-day holding period. According to the Capital Adequacy Directive by the Bank of International Settlement (BIS) in Basle, (Basle Committee, 1996) the risk capital of a bank must be sufficient to cover losses on the bank’s trading portfolio over a 10-day holding period in 99% of occasions. For internal risk control purposes, most of the financial firms compute a 5% VaR over a one-day holding period.

2 Three forms of extreme value distributions represent the GEV distribution: the Gumbel distribution, the Frechet distribution, and the Weibull distribution. See Section 3.2 for details.
terms of mean squared error) than the Hill-based VaR estimate, and that the GPD method applies to both light-tailed and heavy-tailed data, whereas the Hill method is designed specifically for heavy-tailed data.

McNeil and Frey’s (2000) approach (or “the McNeil and Frey (2000) approach” hereafter) yields VaR estimates that reflect the current volatility background. They fitted generalized autoregressive conditional heteroskedasticity (GARCH) models to return data using pseudo maximum likelihood (PML) to obtain estimates of the conditional volatility, and used a GPD approximation suggested by the peaks-over-threshold models to estimate the tail of the distribution of the residuals. It is easy to estimate the conditional return distribution from the estimated distribution of the residuals and estimates of the conditional mean and volatility. The McNeil and Frey (2000) approach reflects two stylized facts exhibited by most financial return series: stochastic volatility and the fat-tailedness of conditional return distributions over short time horizons.

By comparing various methods for tail estimation for financial data, McNeil and Frey (2000) found that their approach is in general the best method for VaR estimation. However, the main weakness of the McNeil and Frey (2000) approach is that so far, there are no automatic algorithms with satisfactory performance for selecting the threshold. Choosing the threshold is the most important implementation issue in the peaks-over-threshold models. Theory tells us that the threshold should be high to satisfy the asymptotic theorem, but a higher threshold leaves fewer observations for estimating the parameters of the tail distribution function. McNeil (1997), Danielsson and de Vries (1997), Dupuis (1998), McNeil and Frey (2000), and Danielsson et al. (2001) examined the issue of determining the fraction of data belonging to the tail. However, these references do not provide a clear answer to the question of which method should be used.

This paper proposes a new approach for VaR estimation that overcomes the drawbacks of the McNeil and Frey (2000) approach. The proposed approach first uses the GEV distribution suggested by the block maxima models to estimate the tail of the distribution of the residuals. This method avoids dealing with the threshold issue posed by the peaks-over-threshold models. Second, we fit asymmetric conditional autoregressive range (ACARR) models to range data using quasi-maximum likelihood estimation (QMLE) to obtain estimates of the conditional volatility. Chou (2006) developed the ACARR model. In his empirical tests, Chou (2006) found that the ACARR model dominates the GARCH model in its ability to forecast volatility. This is because the ACARR model is an asymmetric model that treats the upward and downward movements of the asset price differently (asymmetrically), whereas the GARCH model is a symmetric model that ignores the asymmetric structure of the asset price. Thus, unlike the McNeil and Frey (2000) approach (which uses GARCH models to estimate the conditional volatility), the approach proposed in this study adopts Chou (2006)’s ACARR models.

### 3. Model

---

3 There are periods when the conditional distribution of financial returns appears light-tailed rather than heavy-tailed.


5 Their approach is vindicated by the very satisfying overall performance in various back testing experiments.

6 There are good reasons why the asset price should behave asymmetrically. For example, for investors the more relevant risk is generated by the downward movement rather than the upward movement of the asset price; the upward movement is important in generating investors’ expected returns. For more details of this and other related issues, see also Levy (1978), Engle, Lilian and Robbins (1987), Nelson (1991), Duan (1995), Engle and Ng (1993), Campbell (1997), Barberis and Huang (2000), and Tsay (2002).
This study refers to the proposed approach as the “the two-step dynamic (conditional) range-based extreme value approach” because it involves two steps and combines the classical extreme value theory (EVT) (i.e., the block maxima models) with a range-based volatility model (i.e., the asymmetric conditional autoregressive range (ACARR) model). The following subsections provide details about the proposed approach.

3.1 Step 1: Fit an asymmetric conditional autoregressive range (ACARR)-type model to the range data using the quasi-maximum likelihood estimation (QMLE) method.

Researchers have proposed many different models for volatility dynamics, including models from the ARCH/GARCH family (Bollerslev et al. (1992)), HARCH processes (Muller et al. (1997)), and stochastic volatility models (Shephard (1996)). The proposed approach uses the asymmetric conditional autoregressive range (ACARR) model developed by Chou (2006).

The ACARR model assumes that, given a speculative asset, the upward price range and the downward price range follow distinctly different dynamic processes. According to Chou (2006), the upward price range \( \text{UPR}_t \) and the downward price range \( \text{DNR}_t \) at time \( t \) are

\[
\text{UPR}_t = \ln(P_t^{\text{HIGH}}) - \ln(P_t^{\text{OPEN}}),
\]

\[
\text{DNR}_t = \ln(P_t^{\text{LOW}}) - \ln(P_t^{\text{OPEN}}),
\]

where \( P_t^{\text{HIGH}} \), \( P_t^{\text{LOW}} \) and \( P_t^{\text{OPEN}} \) are the highest, lowest, and opening prices at time \( t \). The ACARR model of order \((p,q)\), or ACARR \((p,q)\) is then

\[
\text{UPR}_t = \lambda^u_t \epsilon^u_t,
\]

\[
\lambda^u_t = \omega^u + \sum_{i=1}^{p} \alpha_{i}^u \text{UPR}_{t-i} + \sum_{i=1}^{q} \beta_{i}^u \lambda^u_{t-i},
\]

\[
\epsilon^u_t \sim iid f^u(.) ,
\]

\[
\text{DNR}_t = -\lambda^d_t \epsilon^d_t,
\]

\[
\lambda^d_t = \omega^d + \sum_{i=1}^{p} \alpha_{i}^d \text{DNR}_{t-i} + \sum_{i=1}^{q} \beta_{i}^d \lambda^d_{t-i},
\]

\[
\epsilon^d_t \sim iid f^d(.) ,
\]

where \( \lambda^u_t \) (\( \lambda^d_t \)) is the conditional mean of \( \text{UPR}_t \) (\( \text{DNR}_t \)) based on all information up to time \( t \), and the distribution of the error term \( \epsilon^u_t \) (\( \epsilon^d_t \)), or the normalized range, is assumed to be independent, identically distributed (iid) and have an exponential density function \( f^u(.) \) \( (f^d(.) \) with unit mean. The parameters \( \omega^u \) and \( \omega^d \) characterize the inherent uncertainty in ranges. The parameters \( \alpha_{i}^u \) and \( \alpha_{i}^d \) characterize the short-term impact of range shocks. The parameters \( \beta_{i}^u \) and \( \beta_{i}^d \) characterize the long-term impact of range shocks. The sums of the parameters, \( \sum_{i=1}^{p} \alpha_{i}^u + \sum_{i=1}^{q} \beta_{i}^u \) and \( \sum_{i=1}^{p} \alpha_{i}^d + \sum_{i=1}^{q} \beta_{i}^d \), determine the persistence of range shocks.

Chou (2006) showed that the unconditional long-term means of \( \text{UPR}_t \) and \( \text{DNR}_t \), i.e., \( \overline{\omega}^u \) and \( \overline{\omega}^d \), can be calculated as

\[
\overline{\omega}^u = \omega^u /[1 - (\sum_{i=1}^{p} \alpha_{i}^u + \sum_{i=1}^{q} \beta_{i}^u)],
\]
\[
\bar{\sigma}^d = \sigma^d / [1 - (\sum_{i=1}^{p} \alpha_i^d + \sum_{i=1}^{q} \beta_i^d)].
\]

Moreover, for the ACARR model to be stationary, the parameters \(\omega^u\), \(\omega^d\), \(\alpha_i^u\), \(\alpha_i^d\), \(\beta_i^u\) and \(\beta_i^d\) must meet the following conditions\(^7\): \(\sum_{i=1}^{p} \alpha_i^u + \sum_{i=1}^{q} \beta_i^u < 1, \sum_{i=1}^{p} \alpha_i^d + \sum_{i=1}^{q} \beta_i^d < 1\) and \(\omega^u, \omega^d, \alpha_i^u, \alpha_i^d, \beta_i^u, \beta_i^d > 0\). \(^{11}\)

Because the ACARR model assumes that the iid error terms \(\varepsilon_t^u\) and \(\varepsilon_t^d\) follow an exponential distribution with unit mean, Chou (2006) also showed that the log likelihood functions for \(U_{PR_t}\) and \(D_{NR_t}\) can be calculated as

\[
L(\omega^u, \alpha_i^u, \beta_i^u; U_{PR_1}, U_{PR_2}, \ldots, U_{PR_T}) = -\sum_{t=1}^{T} \left[ \ln(\lambda_t^u) + \frac{U_{PR_t}}{\lambda_t^u} \right],
\]

\[
L(\omega^d, \alpha_i^d, \beta_i^d; D_{NR_1}, D_{NR_2}, \ldots, D_{NR_T}) = -\sum_{t=1}^{T} \left[ \ln(\lambda_t^d) + \frac{D_{NR_t}}{\lambda_t^d} \right].
\]

Ease of estimation is one of the important properties of the ACARR model. Given that this model specifies the evolutions of the upward price range and the downward price range independently, the quasi-maximum likelihood estimation (QMLE) method can consistently estimate parameters independently. Specifically, the QMLE of the parameters in the ACARR model can be obtained by estimating a GARCH model by specifying a GARCH model for the square root of price range without a constant term in the conditional mean equation. The intuition behind this property is that with some simple adjustments on the specification of the conditional mean equation, the log likelihood function in the ACARR model with an exponential density function is identical to the GARCH model with a normal density function. Furthermore, all asymptotic properties of the GARCH model are applicable to the ACARR model. Given that, the ACARR model is a model for the conditional mean; its regularity conditions are less stringent than those of the GARCH model are\(^8\). Thus, following Chou (2006), the proposed approach employs the QMLE method to determine the best form of the ACARR model that fits our data.

3.2 Step 2: Use the classical extreme value theory (EVT) to model the tail distribution of the error terms of the fitted asymmetric conditional autoregressive range (ACARR) model. Use this EVT model to estimate the value at risk (VaR).

The classical extreme value theory (EVT) (i.e., the block maxima model) deals with the study of the asymptotic behavior of extreme (maxima and minima) observations of a random variable. Following Bortkiewicz (1922) and Frechet (1927), Fisher and Tippett (1928) derived three asymptotic distributions that describe extreme value behavior. These distributions provide solutions for determining extreme value behavior of all data-generation processes, and are useful because the distributional form of extreme values is the same as the tail behavior of the parent distribution\(^9\). The following analysis presents some general concepts related to these three asymptotic distributions.

Consider a sequence of stationary iid random variables \(X_1, X_2, \ldots, X_n\) with a cumulative probability density function \(F\). In determining the probability that the maximum value of the

---

\(^7\) See Bollerslev (1986) for a discussion of the parameters in the context of the GARCH model.

\(^8\) See Chou (2006) for more details of this and other related issues.

first $n$ variables

$$M_n = \max(X_1, X_2, \ldots, X_n)$$

is below a certain level $x$, the cumulative probability relationship is

$$P\{M_n \leq x\} = F^n(x).$$

EVT concentrates on the asymptotic distributional properties of the scaled order statistic $M_n$. Its objective is to find an appropriate situation that satisfies

$$P\{a_n(M_n - b_n) \leq x\} \xrightarrow{w} G(x),$$

or in terms of $F$,

$$F^n(x/a_n + b_n) \xrightarrow{w} G(x),$$

where $G(x)$ is one of Fisher and Tippett’s (1928) three distributional forms, $a_n$ is a dispersion parameter, $b_n$ is a location parameter, and $w$ means, “to weakly converge.” If such a situation exists, $F^n$ belongs to the domain of attraction of $G$.

$G$ takes one of the following three forms:

- **Gumbel distribution:**
  
  $$G(x) = \exp(-e^{-x}), \quad \text{if } -\infty < x < \infty;$$
  
  (18)

- **Frechet distribution:**
  
  $$G(x) = 0, \quad \text{if } x \leq 0,$$
  
  $$G(x) = \exp(-x^{-\lambda}), \quad \text{if } x > 0;$$
  
  (19)

- **Weibull distribution:**
  
  $$G(x) = \exp(-(x)^\lambda), \quad \text{if } x < 0,$$
  
  $$G(x) = 1, \quad \text{if } x \geq 0;$$
  
  (20)

where $\lambda$ is some constant greater than zero. The double exponential form of the Gumbel distribution represents cases where the tail behavior of the parent distribution is exponentially approaching its asymptote. This situation is representative of the tail behavior of the normal distribution. In Frechet distributions, the tail of the underlying distribution does not decay rapidly. Distributions with fat tails can generate extreme instances of Frechet behavior. Finally, the Weibull distributions represent extreme value behavior from truncated data-generation processes.

To operationalize Eq. (18)-(20), Maritz and Munro (1967) extended the research of Jenkinson (1955) and suggested the following general relationships:

$$F_{\min} = 1 - \exp[-(1 + \tau_{\min}x)^{1/\tau_{\min}}],$$

(21)

$$F_{\max} = \exp[-(1 - \tau_{\max}x)^{1/\tau_{\max}}].$$

(22)

Equation (21) describes the cumulative distribution for negative extreme values of $x$, and Eq. (22) performs the same task for positive extreme values. The tail index $\tau$ reflects the fatness of the distribution (i.e., the weight of the tails), and defines the type of distribution from which the extreme values are drawn. When $\tau = 0$, the Gumbel distribution is the limiting distribution. When $\tau < 0$, the Frechet distribution obtains. Finally, when $\tau > 0$, the Weibull distribution is the limiting distribution.

Longin (1999) rewrote Eq. (21) and (22) such that

$$F_{\min} = 1 - \exp[-\{1 + \tau_{\min}(r_{\min} - \beta_{\min})/\alpha_{\min}\}^{1/\tau_{\min}}],$$

(23)

$$F_{\max} = \exp[-\{1 - \tau_{\max}(r_{\max} - \beta_{\max})/\alpha_{\max}\}^{1/\tau_{\max}}],$$

(24)

where the scale parameter $\alpha$ and the location parameter $\beta$ represent the volatility and the average of the extreme values, respectively, and $r$ denotes the minimal or maximal price changes over a specified time interval. Thus, the $(r - \beta)/\alpha$ variable that replaces $x$ in Eq. (21)
and (22) can be viewed as a deviation of \( r \) from its location standardized by its dispersion.

A straightforward manner of estimating \( \tau, \alpha, \) and \( \beta \) is to apply nonlinear least squares regression to the following equations\(^{10}\):

\[
- \ln[-\ln \left(\frac{m}{N+1}\right)] \approx \frac{1}{\tau_{\min}} \ln \alpha_{\min} \max - \tau_{\min} (\beta_{\min} - r_m^{\min}) + u_m^{\min},
\]

(25)

\[
- \ln[-\ln \left(\frac{m}{N+1}\right)] \approx \frac{1}{\tau_{\max}} \ln \alpha_{\max} \max - \tau_{\max} (\beta_{\max} - r_m^{\max}) + u_m^{\max},
\]

(26)

Here, \( m \) is the ranking assigned to each price change after sorting all observations from the lowest to highest value, \( N \) is the number of observations for either the extreme positive or negative price changes, and \( u_m \) is the disturbance term. Equations (25) and (26) are empirical analogs of Eq. (23) and (24), respectively, with \( m/(N+1) \) representing the observed cumulative probability. As Kinnison (1985) pointed out, this nonlinear least squares regression fits the expected cumulative probabilities to their observed extreme values.

An alternative to the nonlinear least squares regression method is to estimate the parameters of Eq. (23) and (24) using the maximum likelihood method, as described by Tiago de Oliveira (1973). The maximum likelihood method provides asymptotically unbiased and minimum variance estimates. Since both these parametric methods provide consistent estimates, the proposed approach uses the nonlinear least squares regression method for parameter estimation.

Equations (1) and (2) show that, over a specified time interval (e.g. one day), either the upward price range \( UPR_i \) or the downward price range \( DNR_i \) indicates the maximal daily price change. Thus, the proposed approach applies nonlinear least squares regression to Eq. (26) to estimate \( \tau, \alpha, \) and \( \beta \) for the fitted asymmetric conditional autoregressive range (ACARR) model obtained in Step 1. Assume that the upward price range \( UPR_i \) is associated with a short position, and that the downward price range \( DNR_i \) is associated with a long position. This is because a short position is at risk if the future’s price increases. Alternatively, a long position is at risk if the futures price decreases.

According to Eq. (5) and (8), the error terms \( \varepsilon_i^u \) and \( \varepsilon_i^d \) in the ACARR model follow an iid distribution this is consistent with the basic EVT assumption that the random variables \( X_1, X_2, \ldots, X_n \) also follow an iid distribution. Empirical evidence also supports the idea that pre-whitening of data by fitting of a dynamic model is a sensible prelude to EVT analysis in practice\(^{11}\). Thus, the proposed approach extracts the \( \varepsilon_i^u \) and \( \varepsilon_i^d \) error terms in the fitted ACARR model obtained in Step 1, and uses EVT to model their tail distribution.

According to Longin (2000), the value at risk (VaR) expressed as a percentage of the value of the position can be computed as

\[
\text{VaR} = -\beta + \frac{\alpha}{\tau} [1 - (-\ln(p))^\tau],
\]

(27)

where \( \tau, \alpha, \) and \( \beta \) are the parameter estimates from Eq. (25) and (26), and \( p \) is the probability of an observation not exceeding the VaR. Moreover, according to McNeil and Frey (2000), the one-day dynamic VaR at time \( t \) can be computed as

\[
\text{VaR}_t = \mu_{t+1} + \sigma_{t+1} \text{VaR}(\varepsilon_t),
\]

(28)

---

\(^{10}\) See Gumbel (1958).

\(^{11}\) See Embrechts et al. (1997) and McNeil and Frey (2000).
where $\mu_{t+1}$ is the conditional mean at time $t+1$ and $\sigma_{t+1}$ is the conditional volatility at time $t+1$ of a stochastic volatility (SV) model, and $VaR(\varepsilon_t)$ is the VaR for the error term $\varepsilon_t$. Combining Eq. (27) and (28) shows that the one-day dynamic VaR at time $t$ can be computed as

$$VaR_t = \mu_{t+1} + \sigma_{t+1} \left\{ -\beta + \frac{\alpha}{\tau} \left[ 1 - (-\ln(p))^\tau \right] \right\},$$

where the fitted ACARR model estimates $\mu_{t+1}$ and $\sigma_{t+1}$, and applying nonlinear least squares regression to Eq. (26) estimates $\tau$, $\alpha$, and $\beta$.

4. Data

This data used in this study consists of 2,251 daily observations of the S&P 500 index futures contracts from January 1997 through December 2006. This data was retrieved from the Datastream International database. The first 1,119 daily observations (from January 1997 through December 2001) were used for in-sample estimation, and the remaining 1,132 daily observations (from January 2002 through December 2006) were used for out-of-sample forecasts to assess the performance of the proposed approach.

To generate the time series of futures prices, each futures contract is rolled over into the next contract on the last trading day of the month preceding each contract expiration month. This rollover period is well in advance of when most traders roll into the next contract, and therefore minimizes any contract expiration effects. Ma, Mercer, and Walker (1992) indicated that this rollover approach mitigates the potentially meaningless data often found in the expiration month of futures contracts.

Table 1 presents descriptive statistics for the upward daily price ranges, the downward daily price ranges, and the daily returns. For each variable considered, this table reports the sample mean, median, maximum, minimum, standard deviation, skewness, kurtosis, Jarque-Bera statistic, Phillips-Perron statistic, and number of observations. The numbers in parentheses are $p$-values. These results indicate that higher peakedness and fat tails relative to a normal distribution characterize all three variables. The highly significant Jarque-Bera statistics provide further evidence of non-normality for all three variables. The Phillips-Perron test tests the stationarity of each variable. The results show that the Phillips-Perron statistics are extremely significant for all three variables. Hence, we can reject the null hypothesis of non-stationarity and conclude that all three variables are stationary.

5. Empirical Results

Empirical analysis was performed in a Matlab 7.x programming environment\(^{12}\). The files with the data and the code can be obtained from the authors upon request.

5.1 In-sample estimation

A total of 1,119 daily observations (from January 1997 through December 2001) were used for in-sample estimation for our two-step dynamic range-based extreme value approach and the McNeil and Frey (2000) approach.

5.1.1 The two-step dynamic range-based extreme value approach

\(^{12}\) Like other standard numerical or statistical software, Matlab now also provides functions or routines for extreme value analysis.
The first step of the proposed approach compares the performance of various asymmetric conditional autoregressive range (ACARR) model specifications to determine the best form of the model that fits our data. Specifically, we considered three forms of the ACARR model: ACARR (1,1), ACARR (1,2), and ACARR (2,1).

Table 2 presents the estimation results, showing the LLF, \( \omega \), \( \alpha_1 \), \( \alpha_2 \), \( \beta_1 \), \( \beta_2 \), and \( N \) for each model specification considered. LLF is the log likelihood function. The terms \( \omega \), \( \alpha_1 \), \( \alpha_2 \), \( \beta_1 \), and \( \beta_2 \) are the model coefficients. \( N \) is the number of observations. The superscripts \( u \) and \( d \) represent the upward ranges and the downward ranges, respectively. The numbers in parentheses are \( p \)-values.

The log likelihood function indicates that the ACARR (1,1) model outperforms the ACARR (1,2) model and the ACARR (2,1) model for both the upward ranges and the downward ranges. This is consistent with the results from the likelihood ratio tests, which are omitted here for brevity. Moreover, the \( p \)-values show that the model coefficients \( \alpha_1^u \), \( \alpha_1^d \), \( \beta_1^u \), and \( \beta_1^d \) from the ACARR (1,1) model are highly significant, while the model coefficients \( \alpha_2^u \), \( \alpha_2^d \), \( \beta_2^u \), and \( \beta_2^d \) from the ACARR (1,2) model and the ACARR (2,1) model are insignificant. Therefore, this study adopts the ACARR (1,1) model for both the upward ranges and the downward ranges.

Table 2 also shows the asymmetric relationship of volatility shocks between the upward ranges and the downward ranges. According to Chou (2006), the model coefficients \( \alpha_i \) and \( \beta_i \) measure the short-term impact effect and the long-term impact effect of volatility shocks, respectively. The model coefficient \( \alpha_i^u \) from the ACARR (1,1) model has a value of 0.047502, and the corresponding model coefficient \( \alpha_i^d \) has a value of 0.047488. That is, \( \alpha_i^u > \alpha_i^d \). This suggests that the short-term impact effect of volatility shocks is greater for the upward ranges than for the downward ranges. Furthermore, the model coefficient \( \beta_i^u \) from the ACARR (1,1) model has a value of 0.935026, and the corresponding model coefficient \( \beta_i^d \) has a value of 0.937103. That is, \( \beta_i^u < \beta_i^d \). This suggests that the long-term impact effect of volatility shocks is smaller for the upward ranges than for the downward ranges. These findings agree with Chou’s (2006) findings.

The second step of our proposed approach begins by taking the logarithm of the error terms \( \varepsilon_i^u \) and \( \varepsilon_i^d \) from the adopted ACARR (1,1) model and constructing a Quantile-Quantile (Q-Q) plot (Fig. 1). The straight line represents the quantile plot of the normal distribution, whereas the curved line represents the quantile plot of the empirical distribution of the logarithmic error terms. If the logarithmic error terms follow a normal distribution, their quantile plot should match the quantile plot of the normal distribution and be a straight line. The extent to which the logarithmic error terms deviate from the straight line indicates the relative degree of non-normality.

According to Chou (2006), the error terms \( \varepsilon_i^u \) and \( \varepsilon_i^d \) follow an exponential distribution, meaning that the logarithmic error terms should follow a normal distribution. However, as Fig. 1 shows, the logarithmic error terms deviate far from the straight line, suggesting that the empirical distribution of the error terms \( \varepsilon_i^u \) and \( \varepsilon_i^d \) is in fact fat-tailed or leptokurtic. Thus, this study uses the extreme value theory (EVT) to model the tail distribution of the error terms.
Table 3 presents the estimation results, showing that $R^2$, $\tau$, $\alpha$, $\beta$, and $N$ for the error terms $\varepsilon_i^u$ and $\varepsilon_i^d$. $R^2$ is only a descriptive statistic\textsuperscript{13}. $\tau$ is the tail index. $\alpha$ is the scale parameter. $\beta$ is the location parameter. $N$ is the number of observations. The numbers in parentheses are $p$-values.

These $R^2$ results, with the lowest being 0.9771, indicate a close coherence between actual and fitted error terms. Moreover, in both cases, the parameter estimates for $\tau$, $\alpha$, and $\beta$ are statistically significant at $p$-values of less than 0.01. These measures suggest that the estimated nonlinear equation (i.e., Eq. (26)) fits the error terms well.

Similar to the results reported by Longin (1995, 1996) and assumed by Kofman and de Vries (1989), Jansen and de Vries (1991), and Kofman (1993), both of the extreme value error terms follow a Frechet limiting distribution, as evidenced by a negative $\tau$. The $\tau$ value for the error term $\varepsilon_i^u$ is smaller in absolute terms than that for the error term $\varepsilon_i^d$, indicating that the latter distribution has thicker tails (i.e., a higher probability of observing extreme price movements). The estimates for $\beta$ are another indication that the error term $\varepsilon_i^d$ is more extreme. The $\beta$ in absolute terms for the error term $\varepsilon_i^d$ is larger than that for the error term $\varepsilon_i^u$. This simply indicates that the extreme price movements are larger in the error term $\varepsilon_i^d$ than in the error term $\varepsilon_i^u$.

\textbf{5.1.2 The McNeil and Frey (2000) approach}

McNeil and Frey (2000) proposed the following two-step approach to estimate the value at risk (VaR):

1. Fit a generalized autoregressive conditional heteroskedasticity (GARCH)-type model to the return data by a pseudo-maximum-likelihood (PML) approach. The residuals (i.e., the error terms) are extracted from the fitted model.

2. Treat the residuals as a realization of a strict white noise process (i.e., iid) with zero mean, unit variance, and marginal distribution function $F_Z(z)$, and use the modern extreme value theory (EVT) to model the tail of $F_Z(z)$. Use this EVT model to estimate VaR.

The following discussion examines these steps in greater detail, and illustrates those using S&P 500 index futures data.

The first step of McNeil and Frey’s (2000) approach fits a GARCH (1,1) model to the return data\textsuperscript{14}. The conditional mean of a GARCH (1,1) model is

$$\mu_t = \lambda X_{t-1},$$ (30)

and the conditional variance is

$$\sigma^2_t = \omega + \alpha_1 (X_{t-1} - \mu_{t-1})^2 + \beta_1 \sigma^2_{t-1},$$ (31)

where $\omega, \alpha_1, \beta_1 > 0$, $\alpha_1 + \beta_1 < 1$, and $|\lambda| < 1$. This model is fitted using the PML method.

\textsuperscript{13} In the nonlinear context, it is not possible to construct an overall goodness-of-fit statistic.

\textsuperscript{14} This model mimics many features of real financial return series.
This means that maximizing the likelihood for a GARCH (1,1) model with normal innovations obtains parameter estimates \( \hat{\vartheta} = (\hat{\omega}, \hat{\alpha}_1, \hat{\beta}_1, \hat{\beta}_2) \). While this amounts to fitting a model using a distributional assumption we do not necessarily believe, Gourieroux (1997) for instance shows that the PML method delivers reasonable parameter estimates. Step 1 ends with extracting the residuals from the fitted GARCH (1,1) model.

Table 2 presents the estimation results for three forms of the GARCH model: GARCH (1,1), GARCH (1,2), and GARCH (2,1). This table shows the \( LLF \), \( \omega \), \( \alpha_1 \), \( \alpha_2 \), \( \beta_1 \), \( \beta_2 \), and \( N \) for each model specification considered. \( LLF \) is the log likelihood function. The terms \( \omega \), \( \alpha_1 \), \( \alpha_2 \), \( \beta_1 \), and \( \beta_2 \) are the model coefficients. \( N \) is the number of observations. The numbers in parentheses are \( p \)-values.

The log likelihood function indicates that the GARCH (1,1) model outperforms both the GARCH (1,2) model and the GARCH (2,1) model. This is consistent with the results from the likelihood ratio tests, which are omitted here for brevity. Moreover, the \( p \)-values show that the model coefficients \( \alpha_1 \) and \( \beta_1 \) from the GARCH (1,1) model are highly significant, but the model coefficients \( \alpha_2 \) and \( \beta_2 \) from the GARCH (1,2) model and the CARCH (2,1) model, on the contrary, are insignificant. These findings are in line with the common belief that the GARCH (1,1) model mimics many features of real financial return series.

The second step of McNeil and Frey’s (2000) approach begins by forming a Quantile-Quantile (Q-Q) plot of the residuals against the normal distribution to confirm that an assumption of conditional normality is unrealistic, and that the innovation process has fat tails or is leptokurtic (Fig. 1).

McNeil and Frey’s (2000) approach uses the modern EVT (i.e., the peaks-over-threshold (POT) model) to model the tail distribution of the residuals. Simply put, the POT model estimates the distribution of exceedances over a certain threshold. Fix a high threshold \( u \) and assume that excess residuals over this threshold have a generalized Pareto distribution (GPD) with the distribution function

\[
G_{\xi,\beta}(y) = \begin{cases} 
1 - (1 + \frac{\xi y}{\beta})^{-1/\xi} & \text{if} \xi \neq 0, \\
1 - \exp(-\frac{y}{\beta}) & \text{if} \xi = 0,
\end{cases} \tag{32}
\]

where \( \beta > 0 \), and the support is \( y \geq 0 \) when \( \xi \geq 0 \) and \( 0 \leq y \leq -\beta / \xi \) when \( \xi < 0 \). The term \( \xi \) represents the shape parameter of the distribution, and \( \beta \) is an additional scaling parameter. If \( \xi > 0 \), the GPD is heavy-tailed; the case \( \xi = 0 \) corresponds to an exponential distribution; and \( \xi < 0 \) corresponds to a short-tailed GPD.

This particular distributional choice is motivated by a limit result in EVT. Consider a general distribution function \( F \) and the corresponding excess distribution above the threshold \( u \) given by

\[
F_u(y) = P\{X - u \leq y \mid X > u\} = \frac{F(y + u) - F(u)}{1 - F(u)}, \tag{33}
\]

for \( 0 \leq y < x_0 - u \), where \( x_0 \) is the right endpoint of \( F \). The excess distribution represents the probability that an observation exceeds the threshold \( u \) by an amount \( y \) at most, given the information that it exceeds the threshold. For a large class of distributions \( F \), Balkema and de Haan (1974) and Pickands (1975) showed that it is possible to find a positive measurable function \( \beta(u) \) such that
\[
\lim_{u \to 0} \sup_{0 < y < y(u)} \left| F_u(y) - G_{\xi, \beta}(y) \right| = 0. \tag{34}
\]

That is, for a large class of underlying distributions \( F \), the excess distribution \( F_u \) converges to a GPD as the threshold \( u \) rises\(^{15}\). For more details, consult Embrechts et al. (1997).

McNeil and Frey’s (2000) approach assumes the tail of the underlying distribution begins at the threshold \( u \). From data of \( n \) points, a random number \( u = \frac{N}{n} > 0 \) will exceed this threshold. If the \( N \) excesses over the threshold are iid with exact GPD distribution, Smith (1987) showed that maximum likelihood estimates \( \hat{\xi} = \hat{\xi}_N \) and \( \hat{\beta} = \hat{\beta}_N \) of the GPD parameters \( \xi \) and \( \beta \) are consistent and asymptotically normal as \( N \to \infty \), provided \( \xi > -1/2 \). Under the weaker assumption that the excesses are iid from \( F_u(y) \) which is only approximately GPD, Smith also obtained asymptotic normality results for \( \hat{\xi} \) and \( \hat{\beta} \). By letting \( u = u_n \to x_0 \) and \( N = N_u \to \infty \) as \( n \to \infty \), he showed that the procedure is essentially asymptotically unbiased provided that \( u \to x_0 \) sufficiently fast. The necessary speed depends on the rate of convergence in Eq. (34). In practical terms, this means that the best GPD estimator of the excess distribution is obtained by trading bias off against variance. To control the variance of the parameter estimates, McNeil and Frey’s (2000) approach chooses a high \( u \) to reduce the chance of bias while keeping \( N \) large (i.e., \( u \) low).

Consider the following equality for points \( x > u \) in the tail of \( F \):

\[
1 - F(x) = (1 - F(u))(1 - F_u(x - u)). \tag{35}
\]

If we estimate the first term on the right hand side of Eq. (35) using the random proportion of the data in the tail \( N/n \) while fixing the number of data in the tail to be \( N < n \)\(^{16}\), and if we estimate the second term by approximating the excess distribution (i.e., the distribution for the excess amounts over the threshold \( u \) for all residuals \( z \) exceeding the threshold \( u \)) with a GPD with parameters \( \xi \) and \( \beta \) fitted by maximum likelihood, the tail estimator for \( F_z(z) \) is

\[
\hat{F}_z(z) = 1 - \frac{N}{n} \left( 1 + \frac{\hat{\xi} z - u}{\hat{\beta}} \right)^{-\frac{1}{\hat{\xi}}}. \tag{36}
\]

For a given probability \( q > 1 - N/n \), we can invert this tail estimation formula to get the VaR estimate for the residuals

\[
VaR(z)_q = u + \frac{\hat{\beta}}{\hat{\xi}} \left( \frac{1 - q}{N/n} \right)^{-\frac{1}{\hat{\xi}}} - 1. \tag{37}
\]

Assuming that the dynamics of return series \( X \) are given by

\[
X_t = \mu_t + \sigma_t Z_t, \tag{38}
\]

where \( \mu_t \) is the conditional mean, \( \sigma_t \) is the conditional volatility, and the residuals \( Z_t \) are a strict white noise process (i.e., iid) with zero mean, unit variance, and marginal distribution function \( F_z(z) \), the one-day VaR estimate for the return series at time \( t \) is

\[
VaR_t^q = \mu_{t+1} + \sigma_{t+1} VaR(z)_q, \tag{39}
\]

where \( \mu_{t+1} \) (i.e., the conditional mean at time \( t+1 \)) and \( \sigma_{t+1} \) (i.e., the conditional volatility at time \( t+1 \)) are estimated using the fitted GARCH (1,1) model\(^{17}\), and \( VaR(z)_q \) (i.e., the VaR

\(^{15}\) In other words, the GDP is the natural model for the unknown excess distribution above sufficiently high thresholds.

\(^{16}\) This effectively gives us a random threshold at the \((N+1)\)th order statistic.

\(^{17}\) See Equations (30) and (31).
estimate for the residuals) is estimated using Eq. (37).

Table 3 shows the threshold values \( u \) and maximum likelihood GPD parameter estimates \( \hat{\beta} \) and \( \hat{\xi} \) for both tails of the residual distribution of our data when the data points \( n = 1,119 \) and the threshold exceedances \( N = 100^{18} \). The numbers in parentheses are \( p \)-values.

Consider both tails of the residual distribution. This is because the upper tail (the right tail) represents losses for an investor with a short position in futures, whereas the lower tail (the left tail) represents losses for an investor being long in futures. Clearly, the upper tail is heavier than the lower tail. The estimated value of the shape parameter \( \hat{\xi} \) is significantly positive in the upper tail \( \hat{\xi} = 0.2809 \), but significantly negative in the lower tail \( \hat{\xi} = -0.6704 \).

5.2 Out-of-sample forecasts

This study uses 1,132 daily observations (from January 2002 through December 2006) for out-of-sample forecasts for assessing the relative performance of the proposed two-step dynamic range-based extreme value approach and the McNeil and Frey (2000) approach.

5.2.1 The value at risk (VaR) estimates

Table 4 reports the forecast results for the value at risk (VaR) estimates\(^{19} \). This table also reports the historical VaR as a benchmark\(^{20} \). We express the VaR numbers as the percentage of the value of a long position (Panel A) and a short position (Panel B), and consider five values of the probability level: 99.9%, 99.5%, 99%, 95%, and 90%. The parentheses next to the VaR numbers represent the percentage difference between the historical VaR and the VaR estimates obtained with our two-step dynamic range-based extreme value approach or the McNeil and Frey (2000) approach. Any difference should be attributed to estimation error\(^{21} \).

The results in Table 4 indicate that the two-step dynamic range-based extreme value approach sometimes slightly underestimate or overestimates the historical VaR\(^{22} \), while the McNeil and Frey (2000) approach always largely underestimates the historical VaR. For example, for a long (short) position and a 99% probability level, the percentage difference between the historical VaR and the VaR estimates is +1.1376% (-5.0135%) with our two-step dynamic range-based extreme value approach and -11.4681% (-15.5408%) with the McNeil and Frey (2000) approach. In most cases, the two-step dynamic range-based extreme value approach gives results similar to historical numbers, and always performs better than the McNeil and Frey (2000) approach.

5.2.2 The Kolmogorov-Smirnov test statistics

This study uses a non-parametric Kolmogorov-Smirnov test to test the statistical significance of the percentage difference between the historical VaR and the VaR estimates obtained with the proposed two-step dynamic range-based extreme value approach and the McNeil and Frey

\(^{18}\text{See McNeil and Frey (2000) for the discussion on the choice of }N = 100.\)

\(^{19}\text{The estimations obtained above are now used to compute VaR for our approach and the McNeil and Frey (2000) approach. See Equations (29), (37) and (39) for details.}\)

\(^{20}\text{The historical VaR is obtained based on the historical distribution.}\)

\(^{21}\text{To test for the statistical significance of any difference, we carry out a non-parametric Kolmogorov-Smirnov test. See Section 5.2.2 for details.}\)

\(^{22}\text{In 7 (3) out of 10 cases, the historical VaR is slightly underestimated (overestimated) by our approach.}\)
Suppose that $F(x)$ denotes the historical distribution function, and that $G(x)$ and $G'(x)$ denote the distribution functions in the two-step dynamic range-based extreme value approach and the McNeil and Frey (2000) approach, respectively. First, test whether the VaR estimates obtained with the two-step dynamic range-based extreme value approach are significantly different from the historical VaR. The first null hypothesis is

$$H_0 : F(x) = G(x),$$

and the alternative hypothesis is

$$H_1 : F(x) \neq G(x).$$

Second, test whether the VaR estimates obtained with the McNeil and Frey (2000) approach are significantly different from the historical VaR. The second null hypothesis is then

$$H'_0 : F(x) = G'(x),$$

and the associated alternative hypothesis is

$$H'_1 : F(x) \neq G'(x).$$

Test both null hypotheses for the long and the short position. Write the Kolmogorov-Smirnov test statistics (KS) for both null hypotheses as

$$KS = \sup_x |F(x) - G(x)|,$$

and

$$KS' = \sup_x |F(x) - G'(x)|.$$

Table 5 presents the Kolmogorov-Smirnov test results. These results show that the Kolmogorov-Smirnov test statistics are significant at the 5% level\(^23\), indicating that the VaR estimates are significantly different from the historical VaR. In other words, the percentage difference between the historical VaR and the VaR estimates obtained with our two-step dynamic range-based extreme value approach or the McNeil and Frey (2000) approach is statistically significant. This supports the statement in Section 5.2.1 that “any difference (between the historical VaR and the VaR estimates) should be attributed to estimation error.”

6. Conclusions

This paper proposes a new approach for estimating value at risk (VaR). This approach combines quasi-maximum-likelihood fitting of asymmetric conditional autoregressive range (ACARR) models to estimate the current volatility and classical extreme value theory (EVT) to estimate the tail of the innovation distribution of the ACARR model. The proposed approach reflects two stylized facts exhibited by most financial time series: stochastic volatility and the fat-tailedness of conditional distributions over short time horizons.

The proposed approach offers two main advantages over McNeil and Frey (2000)’s approach. First, the ACARR model in this approach is an asymmetric model that treats the upward and downward movements of the asset price asymmetrically, whereas the generalized autoregressive conditional heteroskedasticity (GARCH) model used by McNeil and Frey (2000) is a symmetric model that ignores the asymmetric structure of the asset price. (There are good reasons why the asset price should behave asymmetrically. For example, downward movement rather than the upward movement of the asset price generates the more relevant risk for investors; the upward movement is important in generating investors’ expected returns.) Second, because our approach uses classical EVT to estimate the tail of the distribution of the residuals, it avoids the threshold issue in the modern EVT model. (So far, no automatic

\(^{23}\) The critical value of the Kolmogorov-Smirnov test statistic at the 5% significance level is 0.4670.
algorithm with satisfactory performance for selecting the threshold is available.) Because McNeil and Frey’s (2000) approach uses modern EVT, it may badly estimate the tail of the innovation distribution.

In practice, VaR estimation often involves multivariate time series. We are optimistic that our approach (called the “the two-step dynamic range-based extreme value approach”) can be extended to multivariate series. However, a detailed analysis of this question is left for future research.

Acknowledgements

We are grateful to the anonymous referees and the seminar participants at the National Taiwan Ocean University and the Chung Yuan Christian University for their helpful comments and suggestions on earlier versions of this paper. All remaining errors are ours. We also gratefully acknowledge the financial support of the National Science Council of Taiwan.

References


The ACARR (1,1) model using upward ranges

The ACARR (1,1) model using downward ranges

The GARCH (1,1) model

Figure 1. Quantile-Quantile (Q-Q) Plots
This figure shows the Quantile-Quantile (Q-Q) plots of the error terms for the adopted ACARR (1,1) and GARCH (1,1) models for the S&P 500 index futures contracts. The sample period is from January 1997 through December 2001. The straight line represents the quantile plot of the normal distribution, whereas the curved line represents the quantile plot of the empirical distribution of the error terms. If the error terms follow a normal distribution, their quantile plot should match the quantile plot of the normal distribution and be a straight line. The extent to which the error terms deviate from the straight line indicates the relative degree of non-normality.
<table>
<thead>
<tr>
<th></th>
<th>Upward Price Range</th>
<th>Downward Price Range</th>
<th>Return</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>0.4298</td>
<td>-0.4699</td>
<td>0.0317</td>
</tr>
<tr>
<td>Median</td>
<td>0.2100</td>
<td>-0.2150</td>
<td>0.0449</td>
</tr>
<tr>
<td>Maximum</td>
<td>7.4000</td>
<td>0.0000</td>
<td>0.5754</td>
</tr>
<tr>
<td>Minimum</td>
<td>0.0000</td>
<td>-10.0000</td>
<td>-0.7762</td>
</tr>
<tr>
<td>Standard Deviation</td>
<td>0.6031</td>
<td>0.6829</td>
<td>1.0800</td>
</tr>
<tr>
<td>Skewness</td>
<td>2.4445</td>
<td>0.8039</td>
<td>-0.1686</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>16.0978</td>
<td>3.7383</td>
<td>7.0019</td>
</tr>
<tr>
<td>Jarque-Bera Statistic</td>
<td>3,098.8900</td>
<td>2,555.5570</td>
<td>496.0620</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(&lt;0.0001)***</td>
<td>(&lt;0.0001)***</td>
</tr>
<tr>
<td>Phillips-Perron Statistic</td>
<td>74.1000</td>
<td>69.1000</td>
<td>84.8000</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(&lt;0.0001)***</td>
<td>(&lt;0.0001)***</td>
</tr>
<tr>
<td>Number of Observations</td>
<td>2,251</td>
<td>2,251</td>
<td>2,251</td>
</tr>
</tbody>
</table>

*** denotes significance at the 1% level.
Table 2. In-Sample Estimation Results Based on the ACARR

<table>
<thead>
<tr>
<th>Upward Range:</th>
<th>ACARR (1,1)</th>
<th>ACARR (1,2)</th>
<th>ACARR (2,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LLF</strong></td>
<td>-7,236,4660</td>
<td>-7,236,4510</td>
<td>-7,236,4660</td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0024</td>
<td>0.0025</td>
<td>0.0023</td>
</tr>
<tr>
<td></td>
<td>(0.0004)***</td>
<td>(0.0802)*</td>
<td>(0.0009)***</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.0475</td>
<td>0.0484</td>
<td>0.0502</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.0850)*</td>
<td>(0.0756)*</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.0036</td>
<td></td>
<td>(0.8974)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.9350</td>
<td>0.9088</td>
<td>0.9364</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.1275)</td>
<td>(&lt;0.0001)***</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.0249</td>
<td></td>
<td>(0.9646)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Downward Range:</th>
<th>ACARR (1,1)</th>
<th>ACARR (1,2)</th>
<th>ACARR (2,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LLF</strong></td>
<td>-7,450,4670</td>
<td>-7,450,4610</td>
<td>-7,450,4630</td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0019</td>
<td>0.0015</td>
<td>0.0018</td>
</tr>
<tr>
<td></td>
<td>(0.0006)***</td>
<td>(0.1125)</td>
<td>(0.0010)***</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.0474</td>
<td>0.0491</td>
<td>0.0494</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.0968)*</td>
<td>(0.0985)*</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>-0.0021</td>
<td></td>
<td>(0.9456)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.9371</td>
<td>0.8977</td>
<td>0.9374</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.1723)</td>
<td>(&lt;0.0001)***</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.0371</td>
<td></td>
<td>(0.9523)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th>GARCH (1,1)</th>
<th>GARCH (1,2)</th>
<th>GARCH (2,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>LLF</strong></td>
<td>-5,195.6370</td>
<td>-5,195.5520</td>
<td>-5,195.3520</td>
</tr>
<tr>
<td>( \omega )</td>
<td>0.0062</td>
<td>0.0060</td>
<td>0.0067</td>
</tr>
<tr>
<td></td>
<td>(0.0076)***</td>
<td>(0.0602)*</td>
<td>(0.0070)***</td>
</tr>
<tr>
<td>( \alpha_1 )</td>
<td>0.0601</td>
<td>0.0572</td>
<td>0.0504</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.0182)***</td>
<td>(0.0419)***</td>
</tr>
<tr>
<td>( \alpha_2 )</td>
<td>0.0125</td>
<td></td>
<td>(0.6368)</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.9356</td>
<td>0.9983</td>
<td>0.9324</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)***</td>
<td>(0.0158)***</td>
<td>(&lt;0.0001)***</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>-0.0595</td>
<td></td>
<td>(0.8784)</td>
</tr>
<tr>
<td><strong>N</strong></td>
<td>1,119</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>

* denotes significance at the 10% level. ** denotes significance at the 5% level. *** denotes significance at the 1% level.
Table 3. In-Sample Estimation Results Based on the EVT

Panel A: The two-step dynamic range-based extreme value approach

<table>
<thead>
<tr>
<th></th>
<th>ACARR (1,1)</th>
<th>ACARR (1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upward Range</td>
<td>Downward Range</td>
</tr>
<tr>
<td>(R^2)</td>
<td>0.9771</td>
<td>0.9823</td>
</tr>
<tr>
<td>(\tau)</td>
<td>-0.3769</td>
<td>-0.3811</td>
</tr>
<tr>
<td></td>
<td>(0.0040)**</td>
<td>(0.0030)**</td>
</tr>
<tr>
<td>(\alpha)</td>
<td>0.2685</td>
<td>0.2437</td>
</tr>
<tr>
<td></td>
<td>(0.0025)**</td>
<td>(0.0019)**</td>
</tr>
<tr>
<td>(\beta)</td>
<td>0.2301</td>
<td>-0.2432</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)**</td>
<td>(&lt;0.0001)**</td>
</tr>
<tr>
<td>(N)</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th>GARCH (1,1)</th>
<th>GARCH (1,1)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Upper Tail</td>
<td>Lower Tail</td>
</tr>
<tr>
<td>(u)</td>
<td>1.2893</td>
<td>1.2054</td>
</tr>
<tr>
<td>(\hat{\beta})</td>
<td>0.6253</td>
<td>0.6497</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)**</td>
<td>(&lt;0.0001)**</td>
</tr>
<tr>
<td>(\hat{\xi})</td>
<td>0.2809</td>
<td>-0.6704</td>
</tr>
<tr>
<td></td>
<td>(&lt;0.0001)**</td>
<td>(&lt;0.0001)**</td>
</tr>
<tr>
<td>(N)</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>(n)</td>
<td>1,119</td>
<td>1,119</td>
</tr>
</tbody>
</table>

*** denotes significance at the 1% level.
Table 4. Out-of-Sample Forecasts Results on the Value at Risk (VaR) Estimates

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Panel A: Long position</strong></td>
<td></td>
<td><strong>Downward Range</strong></td>
<td><strong>Lower Tail</strong></td>
</tr>
<tr>
<td>99.9%</td>
<td>9.6001%</td>
<td>9.7134% (+1.1802%)</td>
<td>8.5262% (-11.1863%)</td>
</tr>
<tr>
<td>99.5%</td>
<td>9.5534%</td>
<td>9.6563% (+1.0771%)</td>
<td>8.4560% (-11.4870%)</td>
</tr>
<tr>
<td>99%</td>
<td>9.1951%</td>
<td>9.2997% (+1.1376%)</td>
<td>8.1406% (-11.4681%)</td>
</tr>
<tr>
<td>95%</td>
<td>5.8026%</td>
<td>5.6198% (-3.1503%)</td>
<td>4.9643% (-14.4470%)</td>
</tr>
<tr>
<td>90%</td>
<td>4.3613%</td>
<td>4.2345% (-2.9074%)</td>
<td>3.6920% (-15.3463%)</td>
</tr>
<tr>
<td><strong>Panel B: Short position</strong></td>
<td></td>
<td><strong>Upward Range</strong></td>
<td><strong>Upper Tail</strong></td>
</tr>
<tr>
<td>99.9%</td>
<td>7.0158%</td>
<td>6.6106% (-5.7755%)</td>
<td>5.9015% (-15.8827%)</td>
</tr>
<tr>
<td>99.5%</td>
<td>6.7520%</td>
<td>6.4227% (-4.8771%)</td>
<td>5.6213% (-16.7461%)</td>
</tr>
<tr>
<td>99%</td>
<td>6.5743%</td>
<td>6.2447% (-5.0135%)</td>
<td>5.5526% (-15.5408%)</td>
</tr>
<tr>
<td>95%</td>
<td>4.9706%</td>
<td>4.5916% (-7.6248%)</td>
<td>4.1251% (-17.0100%)</td>
</tr>
<tr>
<td>90%</td>
<td>4.1160%</td>
<td>3.8587% (-6.2512%)</td>
<td>3.2934% (-19.9854%)</td>
</tr>
</tbody>
</table>
Table 5. The Kolmogorov-Smirnov Test Statistics

This table presents the Kolmogorov-Smirnov test results. Suppose that $F(x)$ denotes the historical distribution function, and that $G(x)$ and $G'(x)$ denote the distribution functions used in the two-step dynamic range-based extreme value approach and the McNeil and Frey (2000) approach, respectively. First, test whether the VaR estimates obtained with the two-step dynamic range-based extreme value approach are significantly different from the historical VaR. The first null hypothesis is $H_0 : F(x) = G(x)$, and the alternative hypothesis is $H_1 : F(x) \neq G(x)$. Second, test whether the VaR estimates obtained with the McNeil and Frey (2000) approach are significantly different from the historical VaR. The second null hypothesis is $H_0' : F(x) = G'(x)$, and the associated alternative hypothesis is $H_1' : F(x) \neq G'(x)$. Both null hypotheses are tested for the long position (Panel A) and the short position (Panel B).

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A: Long position</td>
<td>Downward Range/Lower Tail</td>
<td>0.5180**</td>
</tr>
<tr>
<td>Panel B: Short position</td>
<td>Upward Range/Upper Tail</td>
<td>0.8575**</td>
</tr>
</tbody>
</table>

** denotes significance at the 5% level.