Quantitative Finance

ISSN: 1469-7688 (Print) 1469-7696 (Online) Journal homepage: http://www.tandfonline.com/loi/rquf20

Estimation of tail-related value-at-risk measures: range-based extreme value approach

Heng-Chih Chou & David K. Wang

Quantitative Finance, 2014
Vol. 14, No. 2, 293–304, http://dx.doi.org/10.1080/14697688.2013.819113

Estimation of tail-related value-at-risk measures: range-based extreme value approach

HENG-CHIH CHOU† and DAVID K. WANG*‡

†Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung, Taiwan
‡Department of Finance, National University of Kaohsiung, Kaohsiung, Taiwan

(*Received 25 May 2010; accepted 11 June 2013*)

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 well-known phenomena found in most financial time series: stochastic volatility and the fat-tailedness of conditional distributions. This approach presents two main advantages over the McNeil and Frey 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 model in the McNeil and Frey 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 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 approach.

Keywords: Risk management; Value-at-risk (VaR); Asymmetric conditional autoregressive range (ACARR) model; Extreme value theory (EVT)

JEL Classifications: C22—Time-series models, C53—Forecasting and other model applications

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.

Previous studies analyse the extreme variations in financial markets, which are mostly because of currency crises, stock market crashes and recent large credit defaults. Many researchers have discussed the tail behaviour of financial time series: Koedijk *et al*. (1990), Loretan and Phillips (1994), Dacorogna *et al*. (1995), Longin (1996), Kuan and Webber (1998), Straetmans (1998), Jondeau and Rockinger (1999), McNeil (1999), Rootzuen and Klauppelberg (1999), Danielsson and de Vries (2000), McNeil and Frey (2000), Neftci (2000), Gencay *et al*. (2003) and Gilli and Kellezi (2006). Diebold *et al*. (1998) provided an interesting discussion on the potential of EVT in risk management.

This paper deals with the behaviour 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

*Corresponding author. Email: dwang@nuk.edu.tw

§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 1996 Capital Adequacy Directive by the Bank of International Settlement in Basle, 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.

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 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 behaviour 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 (POT) 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 POT 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 for 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 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 POT 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 well-known phenomena found in 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 POT 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 POT 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

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 EVT (i.e. the block maxima models) with a

†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.

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

§See Section 5.1.2 for details about the McNeil and Frey (2000) approach.

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

||There is a growing recognition of the fact that the range-based volatility models (e.g. the ACARR model) can provide sharper estimates and forecasts than the return-based volatility models (e.g. the GARCH model). Many insightful studies have provided powerful evidence to support this viewpoint including Parkinson (1980), Garman and Klass (1980), Wiggins (1991), Rogers and Satchell (1991), Kunitomo (1992), Gallant *et al*. (1999), Yang and Zhang (2000), Alizadeh *et al*. (2002), and more recently, Brandt and Jones (2006), Chou (2006), and Martens and van Dijk (2007).

††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 *et al*. (1987), Nelson (1991), Duan (1995), Engle and Ng (1993), Campbell (1999), Barberis and Huang (2000), and Tsay (2002).

range-based volatility model (i.e. the ACARR model). The following subsections provide details about the proposed approach.

3.1. Step 1: fit an ACARR-type model to the range data using the 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 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 UPR_t and the downward price range DNR_t at time t are

$$\mathrm{UPR}_t = \ln(P_t^{\mathrm{HIGH}}) - \ln(P_t^{\mathrm{OPEN}}), \tag{1}$$

$$\mathrm{DNR}_t = \ln(P_t^{\mathrm{LOW}}) - \ln(P_t^{\mathrm{OPEN}}), \tag{2}$$

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

$$\mathrm{UPR}_t = \lambda_t^u \varepsilon_t^u, \tag{3}$$

$$\lambda_t^u = \omega^u + \sum_{i=1}^{p} \alpha_i^u \mathrm{UPR}_{t-i} + \sum_{i=1}^{q} \beta_i^u \lambda_{t-i}^u, \tag{4}$$

$$\varepsilon_t^u \sim \mathrm{iid}\, f^u(.), \tag{5}$$

$$\mathrm{DNR}_t = -\lambda_t^d\, \varepsilon_t^d, \tag{6}$$

$$\lambda_t^d = \omega^d + \sum_{i=1}^{p} \alpha_i^d \mathrm{DNR}_{t-i} + \sum_{i=1}^{q} \beta_i^d\, \lambda_{t-i}^d, \tag{7}$$

$$\varepsilon_t^d \sim \mathrm{iid}\, f^d(.), \tag{8}$$

where λ_t^u (λ_t^d) is the conditional mean of UPR_t (DNR_t) based on all information up to time t and the distribution of the error term ε_t^u (ε_t^d), 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 ω^u and ω^d characterize the inherent uncertainty in ranges. The parameters α_i^u and α_i^d characterize the short-term impact of range shocks. The parameters β_i^u and β_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 UPR_t and DNR_t, i.e. $\overline{\omega^u}$ and $\overline{\omega^d}$, can be calculated as

$$\overline{\omega^u} = \omega^u \Big/ \left[1 - \left(\sum_{i=1}^{p} \alpha_i^u + \sum_{i=1}^{q} \beta_i^u\right)\right], \tag{9}$$

$$\overline{\omega^d} = \omega^d \Big/ \left[1 - \left(\sum_{i=1}^{p} \alpha_i^d + \sum_{i=1}^{q} \beta_i^d\right)\right]. \tag{10}$$

Moreover, for the ACARR model to be stationary, the parameters $\omega^u, \omega^d, \alpha_i^u, \alpha_i^d, \beta_i^u$ and β_i^d must meet the following conditions‡:

$$\sum_{i=1}^{p} \alpha_i^u + \sum_{i=1}^{q} \beta_i^u < 1,$$
$$\sum_{i=1}^{p} \alpha_i^d + \sum_{i=1}^{q} \beta_i^d < 1 \text{ and } \omega^u, \omega^d, \alpha_i^u, \alpha_i^d, \beta_i^u, \beta_i^d > 0. \tag{11}$$

Because the ACARR model assumes that the iid error terms ε_t^u and ε_t^d follow an exponential distribution with unit mean, Chou (2006) also showed that the log likelihood functions for UPR_t and DNR_t can be calculated as

$$L(\omega^u, \alpha_i^u, \beta_i^u; \mathrm{UPR}_1, \mathrm{UPR}_2, \ldots, \mathrm{UPR}_T) = -\sum_{t=1}^{T}\left[\ln(\lambda_t^u) + \frac{\mathrm{UPR}_t}{\lambda_t^u}\right], \tag{12}$$

$$L(\omega^d, \alpha_i^d, \beta_i^d; \mathrm{DNR}_1, \mathrm{DNR}_2, \ldots, \mathrm{DNR}_T) = -\sum_{t=1}^{T}\left[\ln(\lambda_t^d) + \frac{\mathrm{DNR}_t}{\lambda_t^d}\right]. \tag{13}$$

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 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

†Note that t is not a point in time, but a specified time interval, may be an exchange trading day, so the notions of highest, lowest, and opening prices apply. Thanks to the anonymous referees for pointing this out.

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

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.† 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 EVT to model the tail distribution of the error terms of the fitted ACARR model. Use this EVT model to estimate the VaR

The classical EVT (i.e. the block maxima model) deals with the study of the asymptotic behaviour 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 behaviour. These three distributions are: (1) Gumbel distribution, (2) Frechet distribution and (3) Weibull distribution. These distributions provide solutions for determining extreme value behaviour of all data generation processes and are useful because the distributional form of extreme values is same as the tail behaviour of the parent distribution.‡

To operationalize Fisher and Tippett's three distributions, Maritz and Munro (1967) extended the research of Jenkinson (1955) and suggested the following general relationships:

$$F^{\min} = 1 - \exp\left[-(1+\tau^{\min}x)^{1/\tau^{\min}}\right], \tag{14}$$

$$F^{\max} = \exp\left[-(1-\tau^{\max}x)^{1/\tau^{\max}}\right]. \tag{15}$$

Equation (14) describes the cumulative distribution for negative extreme values of x and equation (15) performs the same task for positive extreme values. The tail index τ 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 is obtained. Finally, when $\tau > 0$, the Weibull distribution is the limiting distribution.

Longin (1999) rewrote equations (14) and (15) such that

$$F^{\min} = 1 - \exp\left[-\left\{1+\tau^{\min}\frac{(r^{\min}-\beta^{\min})}{\alpha^{\min}}\right\}^{1/\tau^{\min}}\right], \tag{16}$$

$$F^{\max} = \exp\left[-\left\{1-\tau^{\max}\frac{(r^{\max}-\beta^{\max})}{\alpha^{\max}}\right\}^{1/\tau^{\max}}\right], \tag{17}$$

where the scale paremeter α and the location parameter β 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. Note that unlike price range UPR_t and DNR_t, for price change r there is no reference price—such as opening price—to measure the variations, but variations are obtained sequentially by ordering prices.§ Thus, the $(r-\beta)/\alpha$ variable that replaces x in equations (14) and (15) can be viewed as a deviation of r from its location standardized by its dispersion.

A straightforward manner of estimating $\tau, \alpha,$ and β is to apply non-linear least squares regression to the following equations¶:

$$-\ln\left[-\ln\left(\frac{m}{N+1}\right)\right] = \frac{1}{\tau^{\min}}\ln\alpha^{\min} - \frac{1}{\tau^{\min}}\ln\left[\alpha^{\min}-\tau^{\min}(\beta^{\min}-r_m^{\min})\right] + u_m^{\min}, \tag{18}$$

$$-\ln\left[-\ln\left(\frac{m}{N+1}\right)\right] = \frac{1}{\tau^{\max}}\ln\alpha^{\max} - \frac{1}{\tau^{\max}}\ln\left[\alpha^{\max}-\tau^{\max}(\beta^{\max}-r_m^{\max})\right] + u_m^{\max}, \tag{19}$$

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 (18) and (19) are empirical analogues of equations (16) and (17), respectively, with $m/(N+1)$ representing the observed cumulative probability. As Kinnison (1985) pointed out, this non-linear least squares regression fits the expected cumulative probabilities to their observed extreme values.

An alternative to the non-linear least squares regression method is to estimate the parameters of equations (16) and (17) 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 non-linear 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_t or the downward price range DNR_t indicates the maximal daily price change. Thus, the proposed approach applies non-linear least squares regression to equation (19) to estimate $\tau, \alpha,$ and β for the error terms of the fitted ACARR model obtained in Step 1. Following the McNeil and Frey (2000) approach (see Section 5.1.2 for details), we can assume that the upward price range UPR_t is associated with a short position and that the downward price range DNR_t is associated with a long position. This is because a short position is at risk if the futures price increases (measured by the upward price range UPR_t). Alternatively, a long

†See Chou (2006) for more details of this and other related issues.

‡See Gnedenko (1943), Gumbel (1958), Jenkinson (1955), and Kinnison (1985).

§Thanks to the anonymous referees for bringing this important clarification to our attention.

¶See Gumbel (1958).

position is at risk if the futures price decreases (measured by the downward price range DNR_t).†

According to equations (5) and (8), the error terms ε_t^u and ε_t^d in the ACARR model follow an iid distribution which 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.‡ Thus, the proposed approach extracts the ε_t^u and ε_t^d error terms in the fitted ACARR model obtained in Step 1 and uses the EVT to model their tail distribution.

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

$$\mathrm{VaR} = -\beta + \frac{\alpha}{\tau}[1 - (-\ln(p))^{\tau}], \tag{20}$$

where $\tau, \alpha,$ and β are the parameter estimates from equations (18) and (19) and p is the probability of an observation *not* exceeding the VaR. Note that, as aforementioned, τ denotes the fatness of the distribution (i.e. the weight of the tails), α is a scale parameter (i.e. the volatility of the obervations) and β is a location parameter (i.e. the average of the observations). By plugging in $\tau, \alpha, \beta,$ and p in equation (20), we are essentially looking at the left tail of the (loss) distribution (with p being the probability, say, of 95%) and $-\beta + \frac{\alpha}{\tau}[\ldots]$ actually yields a negative VaR number, where the negative sign simply indicates a loss. According to Longin (2000), equation (20) gives the VaR number in terms of a *percentage* of the value of the position. This is in contrast with the conventional VaR measures which are expressed as *dollars*. All these have shown the less conventional nature of Longin (2000)'s VaR measurement.§

According to McNeil and Frey (2000), the one-day dynamic VaR at time t can be computed as

$$\mathrm{VaR}_t = \mu_{t+1} + \sigma_{t+1}\mathrm{VaR}(\varepsilon_t), \tag{21}$$

where μ_{t+1} is the conditional mean at time $t+1$ and σ_{t+1} is the conditional volatility at time $t+1$ of a stochastic volatility model and $\mathrm{VaR}(\varepsilon_t)$ is the VaR for the error term ε_t. We can combine equations (20) and (21) to show that the one-day dynamic VaR at time t can be computed as

$$\mathrm{VaR}_t = \mu_{t+1} + \sigma_{t+1}\left\{-\beta + \frac{\alpha}{\tau}[1 - (-\ln(p))^{\tau}]\right\}, \tag{22}$$

where the fitted ACARR model estimates μ_{t+1} and σ_{t+1}, and applying non-linear least squares regression to equation (19) estimates $\tau, \alpha,$ and β.

4. Data

This data used in this study consists of 2251 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 1119 daily observations (from January 1997 through December 2001) were used for in-sample estimation and the remaining 1132 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 *et al.* (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 based on the full sample from January 1997 through December 2006. For each variable considered, this table reports the sample mean, median, maximum, minimum, standard deviation, skewness, kurtosis, Jarque–Bera statistic, Phillips–Perron statistic and the 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 $7\times$ programming environment.¶ The files with the data and the code can be obtained from the authors upon request.

---

†According to McNeil and Frey (2000), the upper tail (the right tail) of the *return* distribution 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 (see also Section 5.1.2). Therefore, following McNeil and Frey (2000), we can assume that the upward price range UPR_t is associated with a short position, and that the downward price range DNR_t is associated with a long position. This is because the upper tail of the *range* distribution (i.e. the range (price) increases as measured by the upward price ranges UPR_t) represents losses for an investor with a short position in futures, whereas the lower tail (i.e. the range (price) decreases as measured by the downward price ranges DNR_t) represents losses for an investor being long in futures. Thanks so much to one anonymous referee for bringing this important point to our attention.

‡See Embrechts *et al.* (1997) and McNeil and Frey (2000).

§We thank the anonymous referees for bringing this point to our attention.

¶Like other standard numerical or statistical software, Matlab now also provides functions or routines for extreme value analysis.

Table 1. Descriptive statistics.

	Upward price range	Downward price range	Return
Mean	0.4298	−0.4699	0.0317
Median	0.2100	−0.2150	0.0449
Maximum	7.4000	0.0000	0.5754
Minimum	0.0000	−10.0000	−0.7762
Standard deviation	0.6031	0.6829	1.0800
Skewness	2.4445	0.8039	−0.1686
Kurtosis	16.0978	3.7383	7.0019
Jarque–Bera statistic	3098.8900 (<0.0001)***	2555.5570 (<0.0001)***	496.0620 (<0.0001)***
Phillips–Perron statistic	74.1000 (<0.0001)***	69.1000 (<0.0001)***	84.8000 (<0.0001)***
Number of observations	2251	2251	2251

This table presents the descriptive statistics for the upward daily price ranges, the downward daily price ranges and the daily returns for the S&P 500 index futures contracts based on the full sample from January 1997 through December 2006. For each variable considered, this study reports the sample mean, median, maximum, minimum, standard deviation, skewness, kurtosis, Jarque–Bera statistic, Phillips–Perron statistic and the number of observations. The numbers in parentheses are *p*-values.
***Significance at 1% level.

5.1. In-sample estimation

A total of 1119 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. The first step of the proposed approach compares the performance of various 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. The (in-sample) estimation period is from January 1997 through December 2001. LLF is the log likelihood function. The terms $\omega, \alpha_1, \alpha_2, \beta_1$, and β_2 are the model

Table 2. In-sample estimation results based on the ACARR and GARCH models.

Panel A: The two-step dynamic range-based extreme value approach			
Upward range	ACARR (1,1)	ACARR (1,2)	ACARR (2,1)
LLF	−7236.4660	−7224.4510	−7227.4560
ω^u	0.0024 (0.0004)***	0.0025 (0.0802)*	0.0023 (0.0009)***
α_1^u	0.0475 (<0.0001)***	0.0484 (0.0850)*	0.0502 (0.0756)*
α_2^u			−0.0036 (0.8974)
β_1^u	0.9350 (<0.0001)***	0.9088 (0.1275)	0.9364 (<0.0001)***
β_2^u		0.0249 (0.9646)	
N	1119	1119	1119
Downward range	ACARR (1,1)	ACARR (1,2)	ACARR (2,1)
LLF	−7450.4670	−7440.4610	−7442.4630
ω^d	0.0019 (0.0006)***	0.0015 (0.1125)	0.0018 (0.0010)***
α_1^d	0.0474 (<0.0001)***	0.0491 (0.0968)*	0.0494 (0.0985)*
α_2^d			−0.0021 (0.9456)
β_1^d	0.9371 (<0.0001)***	0.8977 (0.1723)	0.9374 (<0.0001)***
β_2^d		0.0371 (0.9523)	
N	1119	1119	1119
Panel B: the McNeil and Frey (2000) approach			
	GARCH (1,1)	GARCH (1,2)	GARCH (2,1)
LLF	−5195.6370	−5180.5520	−5181.3520
ω	0.0062 (0.0076)***	0.0060 (0.0602)*	0.0067 (0.0070)***
α_1	0.0601 (<0.0001)***	0.0572 (0.0182)**	0.0504 (0.0419)**
α_2			0.0125 (0.6368)
β_1	0.9356 (<0.0001)***	0.9983 (0.0158)**	0.9324 (<0.0001)***
β_2		−0.0595 (0.8784)	
N	1119	1119	1119

This table presents the in-sample estimation results for various ACARR and GARCH model specifications for the S&P 500 index futures contracts. The in-sample estimation period is from January 1997 through December 2001. LLF is the log likelihood function. $\omega, \alpha_1, \alpha_2, \beta_1$, and β_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.
*Significance at 10% level.
**Significance at 5% level.
***Significance at 1% level.

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 $\alpha_1^u, \beta_1^u, \alpha_1^d$, and β_1^d from the ACARR (1,1) model are highly significant, while the model coefficients $\alpha_2^u, \beta_2^u, \alpha_2^d$, and β_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 downward ranges.†

The second step of our proposed approach uses the EVT to model the tail distribution of the error terms ε_t^u and ε_t^d.

Table 3 presents the estimation results (based on the in-sample estimation period from January 1997 through December 2001), showing R^2, τ, α, β, and N for the error terms ε_t^u and ε_t^d. The error terms ε_t^u and ε_t^d are extracted from the adopted ACARR (1,1) model.‡ R^2, τ, α, and β are used to define the type of (classical) EVT distribution that best describes the tail of the error terms ε_t^u and ε_t^d.

To estimate R^2, τ, α, and β, we follow Gumbel (1958) and Kinnison (1985) by applying non-linear least squares regression approach to equation (19). We solve the least squares problem using the Gauss–Newton algorithm (see Nocedal and Wright (1999)) which *iteratively* searches for the minimum of the sum of squares.

R^2 is only a descriptive statistic.§ τ is the tail index reflecting the fatness of the distribution (see equations (14) and (15)). According to the classical EVT (see Maritz and Munro (1967)), when $\tau = 0$, the Gumbel distribution is the best-fit distribution for the error terms (see equation (18)); when $\tau < 0$, the Frechet distribution obtains (see equation (19)); and, when $\tau > 0$, the Weibull distribution best fits the error terms (see equation (20)). α is the scale parameter, representing the volatility for the error terms, and β is the location parameter, representing the average for the error terms (see equations (16) and (17)). N refers to 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 τ, α, and β are statistically significant at p-values of less than 0.01. These measures suggest that the estimated non-linear equation (i.e. equation (19)) 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 τ. The τ value for the error term ε_t^u is smaller in absolute terms than that for the error term ε_t^d, indicating that the latter distribution has thicker tails (i.e. a higher probability of observing extreme price movements). The estimates for β are another indication that the error term ε_t^d is more extreme. The β in absolute terms for the error term ε_t^d is larger than that for the error term ε_t^u. This simply indicates that the extreme price movements are larger in the error term ε_t^d than in the error term ε_t^u.

5.1.2. The McNeil and Frey (2000) approach. McNeil and Frey (2000) proposed the following two-step approach to estimate the VaR:

(1) Fit a 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 best-fit GARCH model.
(2) Treat the error terms as a realization of a strict white noise process (i.e. iid) with zero mean, unit variance and marginal distribution function $F_Z(z)$. Use the modern EVT to model the tail of $F_Z(z)$ and 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.¶ The conditional mean of a GARCH (1,1) model is

$$\mu_t = \lambda X_{t-1}, \tag{23}$$

and the conditional variance is

$$\sigma_t^2 = \omega + \alpha_1 (X_{t-1} - \mu_{t-1})^2 + \beta_1 \sigma_{t-1}^2, \tag{24}$$

where $\omega, \alpha_1, \beta_1 > 0, \alpha_1 + \beta_1 < 1$, and $|\lambda| < 1$. This model is fitted using the PML method. This means that maximizing the likelihood for a GARCH (1,1) model with normal innovations obtains parameter estimates $\hat{\theta} = (\hat{\lambda}, \hat{\omega}, \hat{\alpha}_1, \hat{\beta}_1)^T$. While this amounts to fitting a model using an unrealistic

†Note that according to Chou (2006), the model coefficients α_i and β_i measure the short-term impact effect and the long-term impact effect of volatility shocks respectively. By comparing the model coefficients, Chou reported that there is an asymmetric relationship of volatility shocks between the upward ranges and the downward ranges. As one anonymous referee has pointed out, this inference on volatility asymmetry between the two ranges is just too strong by using point estimation alone. We totally agree with the referee's comment, and therefore, decide not to follow Chou's argument in our current study. We will leave this rather interesting issue for future research. We thank the referee so much for bringing this important point to our attention.

‡According to equations (5) and (8), the error terms ε_t^u and ε_t^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 (see Embrechts *et al*. (1997) and McNeil and Frey (2000)). Thus, we extract the ε_t^u and ε_t^d error terms from the adopted ACARR (1,1) model, and use (classical) EVT to model their tail distribution.

§In the nonlinear context, it is not possible to construct an overall goodness-of-fit statistic.

¶This model mimics many features of real financial return series.

distributional assumption, Gourieroux (1997) shows that the PML method does deliver reasonable 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. The (in-sample) estimation period is from January 1997 through December 2001. LLF is the log likelihood function. The terms $\omega, \alpha_1, \alpha_2, \beta_1$, and β_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 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 α_1 and β_1 from the GARCH (1,1) model are highly significant, but the model coefficients α_2 and β_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.

In the second step of McNeil and Frey's (2000) approach, one uses the modern EVT (i.e. the 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 GPD with the distribution function

$$G_{\xi,\beta}(y) = \begin{cases} 1-(1+\xi y/\beta)^{-1/\xi} & \text{if } \xi \neq 0, \\ 1-\exp(-y/\beta) & \text{if } \xi = 0, \end{cases} \quad (25)$$

where $\beta > 0$ and the support is $y \geqslant 0$ when $\xi \geqslant 0$ and $0 \leqslant y \leqslant -\beta/\xi$ when $\xi < 0$. The term ξ represents the shape parameter of the distribution and β 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 the modern EVT. Consider a general distribution function F and the corresponding excess distribution above the threshold u given by

$$F_u(y) = P\{X-u \leq y | X > u\} = \frac{F(y+u)-F(u)}{1-F(u)}, \quad (26)$$

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

$$\lim_{u \to x_0} \sup_{0 \leq y < x_0 - u} |F_u(y) - G_{\xi,\beta(u)}(y)| = 0. \quad (27)$$

That is, for a large class of underlying distributions F, the excess distribution F_u converges to a GPD as the threshold u rises.† For details, see Embrechts *et al*. (1997).

McNeil and Frey's (2000) approach assumes that the tail of the underlying distribution begins at the threshold u. From the data of n points, a random number $N = N_u > 0$ will exceed this threshold u. If the N excesses are iid with exact GPD distribution, Smith (1987) showed that maximum likelihood estimates of the GPD parameters ξ and β are consistent and asymptotically normal as $N \to \infty$. Under the weaker assumption that $F_u(y)$ 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 under the condition that $u \to x_0$ sufficiently fast. This necessary speed depends on the rate of convergence in equation (27). Practically speaking, this means that the best GPD estimator of the excess distribution is obtained by trading bias off against variance. To control the variance, 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

$$1 - F(x) = (1 - F(u))(1 - F_u(x - u)). \quad (28)$$

If we estimate the first term on the right hand side of equation (28) using the random proportion of the data in the tail N/n while fixing the number of data in the tail to be $N < n$,‡ 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 ξ and β fitted by maximum likelihood, the tail estimator for $F_Z(z)$ is

$$\hat{F}_Z(z) = 1 - \frac{N}{n}\left(1 + \hat{\xi}\frac{z-u}{\hat{\beta}}\right)^{-1/\hat{\xi}}. \quad (29)$$

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

$$\mathrm{VaR}(z)_q = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{1-q}{N/n}\right)^{-\hat{\xi}} - 1\right). \quad (30)$$

Assuming that the dynamics of return series X are given by

$$X_t = \mu_t + \sigma_t Z_t, \quad (31)$$

where μ_t is the conditional mean, σ_t is the conditional volatility and the residuals Z_t are a strict white noise

†In other words, the GDP is the *natural model* for the unknown excess distribution above sufficiently high thresholds.

‡This effectively gives us a random threshold at the $(N+1)$th order statistic.

Table 3. In-sample estimation results based on the EVT.

Panel A: the two-step dynamic range-based extreme value approach		
	ACARR (1,1)	ACARR (1,1)
	Upward range	Downward range
R^2	0.9771	0.9823
τ	−0.3769 (0.0040)***	−0.3811 (0.0030)***
α	0.2685 (0.0025)***	0.2437 (0.0019)***
β	0.2301 (<0.0001)***	−0.2432 (<0.0001)***
N	1119	1119
Panel B: the McNeil and Frey (2000) approach		
	GARCH (1,1)	GARCH (1,1)
	Upper tail	Lower tail
u	1.2893	1.2054
$\hat{\beta}$	0.6253 (<0.0001)***	0.6497 (<0.0001)***
$\hat{\xi}$	0.2809 (<0.0001)***	−0.6704 (<0.0001)***
N	100	100
n	1119	1119

This table presents the in-sample estimation results for both the classical EVT model (Panel A) and modern EVT model (Panel B) for the S&P 500 index futures contracts. The in-sample estimation period is from January 1997 through December 2001. Panel A reports the R^2, τ, α, β and N for the error terms extracted from the ACARR (1,1) model. R^2 is a descriptive statistic. τ is the tail index. α is the scale parameter. β is the location parameter. N is the number of observations. The numbers in parentheses are p-values. Panel B gives the threshold values u, the maximum likelihood GPD parameter estimates $\hat{\beta}$ and $\hat{\xi}$, the threshold exceedances N and the data points n for both tails of the residual distribution of the GARCH (1,1) model. The numbers in parentheses are p-values.
***Significance at 1% level.

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

$$\mathrm{VaR}_q^t = \mu_{t+1} + \sigma_{t+1}\mathrm{VaR}(z)_q, \tag{32}$$

where μ_{t+1} (i.e. the conditional mean at time $t+1$) and σ_{t+1} (i.e. the conditional volatility at time $t+1$) are estimated using the fitted GARCH (1,1) model† and $\mathrm{VaR}(z)_q$ (i.e. the VaR estimate for the residuals) is estimated using equation (30).

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 $n=1119$ and the threshold exceedances $N=100$.‡ The (in-sample) estimation period is from January 1997 through December 2001. 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 1132 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.

†See equations (23) and (24).

‡See McNeil and Frey (2000) for the discussion on the choice of $N=100$.

5.2.1. The VaR estimates. Table 4 reports the forecast results for the VaR estimates§. The (out-of-sample) forecasting period is from January 2002 through December 2006. This table also reports the historical VaR as a benchmark¶. 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||.

The results in table 4 indicate that the two-step dynamic range-based extreme value approach sometimes slightly underestimates or overestimates the historical VaR,†† 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

§The estimations obtained above are now used to compute VaR for our approach and the McNeil and Frey (2000) approach. See equations (22), (30) and (32) for details.

¶The historical VaR is obtained based on the historical distribution.

||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.

††In 7 (3) out of 10 cases, the historical VaR is slightly underestimated (overestimated) by our approach.

Table 4. Out-of-sample forecasts results on the VaR estimates.

Probability level	Historical VaR	The two-step dynamic range-based extreme value approach	The McNeil and Frey (2000) approach
Panel A: long position			
		Downward range	Lower tail
99.9%	9.6001%	9.7134% (+1.1802%)	8.5262% (−11.1863%)
99.5%	9.5534%	9.6563% (+1.0771%)	8.4560% (−11.4870%)
99%	9.1951%	9.2997% (+1.1376%)	8.1406% (−11.4681%)
95%	5.8026%	5.6198% (−3.1503%)	4.9643% (−14.4470%)
90%	4.3613%	4.2345% (−2.9074%)	3.6920% (−15.3463%)
Panel B: short position			
		Upward range	Upper tail
99.9%	7.0158%	6.6106% (−5.7755%)	5.9015% (−15.8827%)
99.5%	6.7520%	6.4227% (−4.8771%)	5.6213% (−16.7461%)
99%	6.5743%	6.2447% (−5.0135%)	5.5526% (−15.5408%)
95%	4.9706%	4.5916% (−7.6248%)	4.1251% (−17.0100%)
90%	4.1160%	3.8587% (−6.2512%)	3.2934% (−19.9854%)

This table presents the out-of-sample forecasts results on the VaR estimates for both the two-step dynamic range-based extreme value approach and McNeil and Frey (2000) approach for the S&P 500 index futures contracts. The out-of-sample forecasting period is from January 2002 through December 2006. This table also shows the historical VaR as a benchmark, expresses the VaR numbers as the percentage of the value of a long position (Panel A) and a short position (Panel B) and considers 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 the two-step dynamic range-based extreme value approach or the McNeil and Frey (2000) approach. Any difference should be attributed to estimation error.

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 (2000) 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

$$H_0{:}F(x) = G(x), \tag{33}$$

and the alternative hypothesis is

$$H_1{:}F(x) \neq G(x). \tag{34}$$

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

$$H'_0{:}F(x) = G'(x), \tag{35}$$

and the associated alternative hypothesis is

$$H'_1{:}F(x) \neq G'(x). \tag{36}$$

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

$$\mathrm{KS} = \sup_x |F(x) - G(x)|, \tag{37}$$

Table 5. The Kolmogorov–Smirnov test statistics.

	The two-step dynamic range-based extreme value approach	The McNeil and Frey (2000) approach
Panel A: long position		
Downward range/lower tail	0.5180**	0.7059**
Panel B: short position		
Upward range/upper tail	0.8575**	0.6521**

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).
**Significance at 5% level.

and

$$\mathrm{KS}' = \sup_x |F(x) - G'(x)|. \quad (38)$$

Table 5 presents the Kolmogorov–Smirnov test results. These results show that the Kolmogorov–Smirnov test statistics are significant at 5% level,† 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 VaR. This approach combines quasi-maximum-likelihood fitting of ACARR models to estimate the current volatility and classical EVT to estimate the tail of the innovation distribution of the ACARR model. The proposed approach reflects two well-known phenomena found in most financial time series: stochastic volatility and the fat-tailedness of conditional distributions.

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 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 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 editors, Michael Dempster and Jim Gatheral, the anonymous referees, and the participants at the 4th Singapore Economic Review (SER) conference, 18th conference on Pacific Basin Finance, Economics, Accounting, and Management (PBFEAM), 18th conference of the Multinational Finance Society (MFS), 4th Financial Management Association (FMA) Asian conference, 19th conference on the Theories and Practices of Securities and Financial Markets (SFM), and a seminar at Academia Sinica, for their valuable comments and suggestions on earlier versions of this paper.

†The critical value of the Kolmogorov–Smirnov test statistic at the 5% significance level is 0.4670.

References

Alizadeh, S., Brandt, M. and Diebold, F., Range-based estimation of stochastic volatility models. *J. Financ.*, 2002, **57**, 1047–1091.

Balkema, A. and de Haan, L., Residual life time at great age. *Ann. Probab.*, 1974, **2**, 792–804.

Barberis, N. and Huang, M., Mental accounting, loss aversion, and individual stock returns. GSB Working Paper, University of Chicago, 2000.

Bollerslev, T., Generalized autoregressive conditional heteroskedasticity. *J. Econom.*, 1986, **31**, 327–337.

Bollerslev, T., Chou, R. and Kroner, K., ARCH modeling in finance: A review of the theory and empirical evidence. *J. Econom.*, 1992, **52**, 5–59.

Bortkiewicz, L., Variationsbreite und mittlerer fehler [Variation and average error]. *Sitzungsberichte den Berliner Mathematik Gesellschaft*, 1922, **21**, 3–11.

Brandt, M. and Jones, C., Volatility forecasting with range-based EGARCH models. *J. Bus. Econ. Stat.*, 2006, **24**, 470–486.

Campbell, J., Asset prices, consumption and the business cycle. In *The Handbook of Macroeconomics*, edited by J.B. Taylor and M. Woodford, pp. 1231–1303, 1999 (Elsevier: Amsterdam).

Chou, R., Modeling the asymmetry of stock movements using price ranges. *Adv. Econom.*, 2006, **20**, 231–257.

Dacorogna, M., Mauller, U., Pictet, O. and de Vries, C., *The Distribution of Extremal Foreign Exchange Rate Returns in Extremely Large Data Sets*, 1995 (Tinbergen Institute: Amsterdam).

Danielsson, J. and de Vries, C., Tail index and quantile estimation with very high frequency data. *J. Empir. Financ.*, 1997, **4**, 241–257.

Danielsson, J. and de Vries, C., Value-at-risk and extreme returns. *Annales d'Economie et de Statistique*, 2000, **60**, 239–270.

Danielsson, J., de Vries, C., de Haan, L. and Peng, L., Using a boot-strap method to choose the sample fraction in tail index estimation. *J. Multivariate Anal.*, 2001, **76**, 226–248.

Diebold, F., Schuermann, T. and Stroughair, J., Pitfalls and opportunities in the use of extreme value theory in risk management. In *Decision Technologies for Computational Finance*, edited by A.-P.N. Refenes, A.N. Burgess and J.E. Moody, pp. 3–12, 1998 (Springer: London).

Duan, J., The GARCH option pricing model. *Math. Financ.*, 1995, **5**, 13–32.

Dupuis, D., Exceedances over high thresholds: A guide to threshold selection. *Extremes*, 1998, **1**, 251–261.

Embrechts, P., Klauppelberg, C. and Mikosch, T., *Modelling Extremal Events*, 1997 (Springer: Berlin).

Engle, R., Lillian, D. and Robbins, R., Estimating time varying risk premia in the term structure: The ARCH-M model. *Econometrica*, 1987, **55**, 391–408.

Engle, R. and Ng, V., Measuring and testing the impact of news on volatility. *J. Financ.*, 1993, **48**, 1749–1778.

Fisher, R. and Tippett, L., Limiting forms of the frequency distribution of the largest and smallest member of a sample. *Proc. Cambridge Philos. Soc.*, 1928, **24**, 180–190.

Frechet, M., Sur la loi de probabilite de lecart maximum [On the law of probability of Lecart maximum]. *Annals de la Societe polonaise de Mathematiques*, 1927, **6**, 93–116.

Gallant, R., Hsu, C. and Tauchen, G., Using daily range data to calibrate volatility diffusions and extracting integrated volatility. *Rev. Econ. Stat.*, 1999, **81**, 617–631.

Garman, M. and Klass, M., On the estimation of security price volatilities from historical data. *J. Bus.*, 1980, **53**, 67–78.

Gencay, R., Selcuk, F. and Ulugaulyagci, A., High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insur. Math. Econ.*, 2003, **33**, 337–356.

Gilli, M. and Kellezi, E., An application of extreme value theory for measuring financial risk. *Comput. Econ.*, 2006, **27**, 1–23.

Gnedenko, B., Sur la distribution limite du terme maximum dune se rie aleatoire [On the limit distribution of the maximum term in a random series]. *Ann. Math.*, 1943, **44**, 423–453.

Gourieroux, C., *ARCH-Models and Financial Applications, Springer Series in Statistics*, 1997 (Springer: New York).

Gumbel, E., *Statistics of Extremes*, 1958 (Columbia University Press: New York).

Jansen, D. and de Vries, C., On the frequency of large stock returns: Putting booms and busts into perspective. *Rev. Econ. Stat.*, 1991, **73**, 18–24.

Jenkinson, A., The frequency distribution of the annual maximum (or minimum) values of meteorological elements. *Q. J. R. Meteorol. Soc.*, 1955, **87**, 145–158.

Jondeau, E. and Rockinger, M., *The Tail Behavior of Stock Returns: Emerging Versus Mature Markets*, 1999 (Groupe HEC: Paris).

Kinnison, R., *Applied Extreme-value Statistics*, 1985 (Battell Press: Columbus, GA).

Koedijk, K., Schafgans, M. and de Vries, C., The tail index of exchange rate returns. *J. Int. Econ.*, 1990, **29**, 93–108.

Kofman, P., Optimizing futures margins with distribution tails. *Adv. Rev. Futures Markets*, 1993, **6**, 263–278.

Kofman, P. and de Vries, C., Potato futures returns: A tail investigation. *Rev. Futures Markets*, 1989, **8**, 244–258.

Kuan, C. and Webber, N., Valuing interest rate derivatives consistent with a volatility smile. Working Paper, University of Warwick, 1998.

Kunitomo, N., Improving the Parkinson method of estimating security price volatilities. *J. Bus.*, 1992, **65**, 295–302.

Levy, H., Equilibrium in an imperfect market: A constraint on the number of securities in a portfolio. *Am. Econ. Rev.*, 1978, **68**, 643–658.

Longin, F., Optimal margins in futures markets: a parametric extreme-based approach. Proceedings of 9th Chicago Board of Trade Conference on Futures and Options, Bonn, 1995.

Longin, F., The asymptotic distribution of extreme stock market returns. *J. Bus.*, 1996, **69**, 383–408.

Longin, F., Optimal margin levels in futures markets: Extreme price movements. *J. Futures Markets*, 1999, **19**, 127–152.

Longin, F., From value at risk to stress testing: The extreme value approach. *J. Bank. Finance*, 2000, **24**, 1097–1130.

Loretan, M. and Phillips, P., Testing the covariance stationarity of heavy-tailed time series. *J. Empir. Financ.*, 1994, **1**, 211–248.

Ma, C., Mercer, J. and Walker, M., Rolling over futures contracts: A note. *J. Futures Contracts*, 1992, **12**, 203–217.

Maritz, J. and Munro, A., On the use of the generalized extreme-value distribution in estimating extreme percentiles. *Biometrics*, 1967, **23**, 79–103.

Martens, M. and van Dijk, D., Measuring volatility with the realized range. *J. Econom.*, 2007, **138**, 181–207.

McNeil, A., Estimating the tails of loss severity distributions using extreme value theory. *ASTIN Bull.*, 1997, **27**, 117–137.

McNeil, A., Extreme value theory for risk managers. In *Internal Modelling and CAD II*, edited by British Bankers' Association, pp. 93–113, 1999 (Risk Books: London).

McNeil, A. and Frey, R., Estimation of tail-related risk measures for heteroscedastic financial time series: An extreme value approach. *J. Empir. Financ.*, 2000, **7**, 271–300.

Muller, O., Dacarogna, M., Dave, R., Olsen, R., Pictet, O. and von Weizsacker, J., Volatilities of different time resolutions—Analyzing the dynamics of market components. *J. Empir. Financ.*, 1997, **4**, 213–240.

Neftci, S., Value at risk calculations, extreme events, and tail estimation. *J. Deriv.*, 2000, **7**, 23–38.

Nelson, D., Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, 1991, **59**, 347–370.

Nocedal, J. and Wright, S., *Numerical Optimization*, 1999 (Springer: New York).

Parkinson, M., The extreme value method for estimating the variance of the rate of return. *J. Bus.*, 1980, **53**, 61–65.

Pickands, J., Statistical inference using extreme order statistics. *Ann. Stat.*, 1975, **3**, 119–131.

Rogers, L. and Satchell, S., Estimating variance from high, low and closing prices. *Ann. Appl. Probab.*, 1991, **1**, 504–512.

Rootzuen, H. and Klauppelberg, C., A single number can't hedge against economic catastrophes. *Ambio*, 1999, **28**, 550–555.

Shephard, N., Statistical aspects of ARCH and stochastic volatility. In *Time Series Models in Econometrics, Finance and Other Fields*, edited by D.R. Cox, D.V. Hinkley and O.E. Barndorff-Nielsen, pp. 1–55, 1996 (Chapman and Hall: London).

Smith, R., Estimating tails of probability distributions. *Ann. Stat.*, 1987, **15**, 1174–1207.

Straetmans, S., Extreme financial returns and their comovements. PhD Thesis, Tinbergen Institute Research Series, Erasmus University Rotterdam, 1998.

Tiago de Oliveira, J., *Statistical Extremes—A Survey*, 1973 (Faculty of Sciences, Center of Applied Mathematics: Lisbon).

Tsay, R., *Analysis of Financial Time Series*, 2002 (John Wiley: New York).

Wiggins, J., Empirical tests of the bias and efficiency of the extreme-value variance estimator for common stocks. *J. Bus.*, 1991, **64**, 417–432.

Yang, D. and Zhang, Q., Drift independent volatility estimation based on high, low, open, and close prices. *J. Bus.*, 2000, **73**, 477–491.