International Review of Financial Analysis 44 (2016) 34–55

Contents lists available at ScienceDirect

International Review of Financial Analysis

Intraday risk management in International stock markets: A conditional EVT approach

Madhusudan Karmakar *, Samit Paul [1]

Indian Institute of Management Lucknow, Prabandh Nagar, IIM Road, Lucknow, 226 013, UP, India

A R T I C L E I N F O

Article history:
Received 18 April 2015
Received in revised form 5 November 2015
Accepted 19 November 2015
Available online 28 November 2015

JEL classification:
G 15
G 17

Keywords:
Value-at-risk
Peak over threshold method
Conditional EVT
Deseasonalized returns
High frequency data
Forecasting

A B S T R A C T

The study compares the predictive ability of various models in estimating intraday Value-at-Risk (VaR) and Expected Shortfall (ES) using high frequency share price index data from sixteen different countries across the world for a period of seven and half months from September 20, 2013 to May 07, 2014. The main emphasis of the study has been given to Extreme Value Theory (EVT) and to evaluate how well Conditional EVT model performs in modeling tails of distributions and in estimating and forecasting intraday VaR and ES measures. We have followed McNeil and Frey's (2000) two stage approach called Conditional EVT to estimate dynamic intraday VaR and ES. We have compared the accuracy of Conditional EVT approach to intraday VaR and ES estimation with other competing models. The best performing model is found to be the Conditional EVT in estimating both the quantiles for the entire sample. The study is useful for market participants (such as intraday traders and market makers) involved in frequent intraday trading in such equity markets.

1. Introduction

One of the most interesting developments in empirical finance is given by the availability of tick-by-tick data from a variety of liquid financial markets. This development of high frequency data bases allows researchers to investigate a wide range of issues in the financial markets. One such important issue is intraday risk management which is connected with the use of high frequency data. For active market participants such as high frequency traders, day traders or market makers, trading risk should be evaluated on shorter-than-daily intervals since the horizon of their investments is generally less than a day. For example, day traders liquidate any open positions at closing, in order to preempt any adverse overnight moves resulting in large gap openings. Brokers must also be able to calculate trading limits as fast as clients place their orders. Significant intraday variations in asset prices affect the margins a client has to deposit with a clearing firm, and this should be taken into consideration while designing an appropriate model to estimate the margins. Sometimes banks also use intraday risk analysis for internal control of their trading desk.

Value-at-risk (VaR) has become a widely used tool in risk management of financial institutions and regulators. A VaR model measures market risk by determining how much the value of a portfolio could decline with α% probability over a certain time horizon τ as a result of changes in market prices or rates. Another useful measure of risk is the expected shortfall (ES) which is defined as the expected size of a loss that exceeds VaR. Where VaR addresses the question: "How bad can things get?", the ES addresses the question: "If things go bad, what is the expected loss?" Much effort has been spent on developing increasingly sophisticated risk models of VaR type for daily data and/or longer horizon,[2] but the issue of intraday market risk measurement has been less explored. With increased access to intraday financial data bases and advanced computing power, it has now become possible to address the question of how to define practical risk measures for investors or market makers operating on an intraday basis. This paper examines market risk at very short time horizon with intraday VaR and ES, using high frequency data of various stock markets across the globe.

The most commonly used VaR and ES models assume that the probability distribution of the daily/intraday financial asset return is normal, an assumption that is far from reality. Many of the asset returns exhibit significant amounts of excess kurtosis. This means that the probability distributions of these returns have "fat tails" so that extreme outcomes

* Corresponding author. Tel.: +91 522 6696540 (R), +91 522 6696624 (office).
E-mail addresses: madhu@iiml.ac.in (M. Karmakar), fpm13005@iiml.ac.in (S. Paul).
[1] Tel.: +91 8009982234 (mobile).

[2] This is motivated by the fact that financial institutions generally produce their market VaR at the end of the business day to measure their total risk exposure over the next day. For regulated capital adequacy purposes, banks usually compute the market VaR daily and then re-scale it to a 10-day horizon.

http://dx.doi.org/10.1016/j.irfa.2015.11.008
1057-5219/

happen much more frequently than that would be predicted by the normal distribution assumption. In this article, we show how the normal distribution assumption can be relaxed. We propose an extreme value approach popularly known as extreme value theory (EVT) to calculate intraday VaR and ES that allows the user to choose more generalized fat tailed distributions for the high frequency stock market returns.

Classical EVT typically relies on an important assumption of independent and identically distributed (*iid*) observations, which obviously does not match the actual situation of financial return series. In order to overcome the drawbacks the immediate solution is provided by McNeil and Frey (2000). Using a two stage approach, McNeil and Frey estimate a GARCH model in stage one with a view to filtering the return series to obtain (nearly) *iid* residuals. In stage two, the EVT framework is applied to the standardized residuals. The advantage of this GARCH–EVT combination lies in its ability to capture conditional heteroscedasticity in the data through the GARCH framework, while at the same time modeling the extreme tail behavior through the EVT method.

Following McNeil and Frey (2000), many researchers use GARCH–EVT approach along with other traditional models on different data sets and find that this approach performs better than other competing models for VaR estimations. Bali and Neftci (2003) apply the GARCH–EVT model to U.S. short-term interest rates and show that the model yields more accurate estimates of VaR than that obtained from a Student *t*-distributed GARCH model. Bystrom (2004) and Fernandez (2005) also find that the GARCH–EVT model performs better than the parametric models in forecasting VaR for various international stock markets. In an energy application, Bystrom (2005) employs GARCH–EVT framework to Nord Pool returns and observes that the model produces more accurate estimates as well as forecasts of extreme quantiles than a pure GARCH model. Recently, Cotter (2007); Ghorbel and Trabelsi (2008); Marimoutou, Raggad, and Trabelsi (2009) etc., use GARCH–EVT model to measure VaR in different markets and find that the model predicts better estimates of VaR than that of other well-known modeling techniques. Very recently, Karmakar and Shukla (2015) have compared the accuracy of GARCH–EVT approach for VaR calculation with other competing models using data from six emerging as well as developed stock markets of the world. They also observe that the GARCH–EVT approach performs the best in estimating VaR.

In all these papers the GARCH–EVT model was applied on daily data to forecast daily VaR. Intraday VaR was initially discussed by Giot (2005) and the study was followed by a growing number of research papers focusing on intraday VaR with the increasing availability of high frequency financial databases. However, while there exists a number of studies on intraday VaR measures using traditional GARCH-based models (Beck, Kim, Rachev, Feindt, & Fabozzi, 2013; Colletaz, Hurlin, & Tokpavi, 2007; Dionne, Duchesne, & Pacurar, 2009; Mike, So, & Xu, 2013; Morimoto & Kawasaki, 2008; Qi & Lon Ng, 2009 among others), to the best of our knowledge, a limited work has been done on forecasting intraday VaR based on sophisticated GARCH–EVT model. Very recently some researchers have concentrated to estimate GARCH–EVT based intraday VaR on high frequency data. Ergun and Jun (2010) estimate several GARCH and EVT based models to forecast intraday VaR for S&P 500 stock index futures returns. They find that the EVT based model and the GARCH based models which take conditional skewness and kurtosis into account provide accurate VaR forecast. Chavez-Demoulin and McGill (2012) measures intraday VaR using EVT–Hawkes process. They observe that the process provides a suitable estimate of risk measures at high quantile for financial time series in the US market.

While there are voluminous studies on VaR forecasting based on GARCH as well as sophisticated EVT models, the literature on ES forecasting is limited to a few studies which use mainly traditional GARCH type models. Embrechts, Kaufmann, and Patie (2005) test different traditional models on daily data taken from different markets by performing backtesting on ES prediction. Watanabe (2012) applies realized GARCH model to forecast ES along with VaR using daily return of S&P 500 stock index.

In this paper we compute intraday VaR and ES for sixteen stock markets across Asia, Europe, North America, Latin America, Africa and Australia. For robustness purposes, we have used sixteen different stock indices, to avoid results dependent on a specific financial market. In this respect the present paper differs from earlier studies which have measured EVT based intraday VaR using high frequency data of US market only.

A plethora of models used for forecasting VaR and ES are examined. However, the primary focus of this paper is to compare the accuracy of GARCH–EVT approach for intraday VaR and ES calculation with other competing approaches. We have followed McNeil and Frey's (2000) two stage approach and demonstrated in detail how the approach combines the simple EVT approach with the appropriate GARCH model to accommodate both conditional volatility and fat tailed return distribution to estimate the tail related risk measures in different stock markets across the world.

Although both the negative and the positive tails of stock return distributions are interesting from a risk management perspective, most studies of extreme stock returns focus on losses (as opposed to gains), and large crashes are generally considered more important than large booms. In this paper, we look at sixteen countries' aggregate stock markets and we expect losses to be of more general interest than gains. To save space, we therefore chose to focus solely on the negative tails of the distribution of the sixteen return series.

The remainder of the paper is structured as follows. Section 2 presents a brief overview of EVT, describes the estimation of VaR and ES, and then explains McNeil and Frey's (2000) two stage approach called Conditional EVT to estimate dynamic intraday VaR and ES. Section 3 focuses on the data used in the study. Section 4 presents the empirical findings from both in-sample and out-of sample evaluations of intraday VaR and ES measures. Finally, Section 5 concludes the study.

2. Modeling the tails of stock return distributions

In the following subsections, we present a brief overview of the theoretical framework of EVT, describe VaR and ES and explain how conditional EVT is applied to VaR and ES.

2.1. Extreme value theory

Extreme value theory provides the fundamentals for the statistical modeling of rare (extreme) events, and is used to compute tail related measures. The extreme values can be modeled by the block maxima or the peak over threshold (POT).[3] The first approach defines extreme events as the maximum (minimum) value in each sub-period. Therefore, it tends to abandon a great deal of data. The second approach considers the sort of clustering phenomena frequently found in financial data. Hence, we would use here the second approach which is more appropriate for the current study.

We would consider here a sequence of n iid random variables X $(x_1, x_2, \ldots . x_n)$ that represents the residuals of the intraday return series. The excess distribution F(x), which is the probability that X exceeds a fixed threshold u, can be estimated using a generalized Parato distribution (GPD) fitted by the maximum likelihood method. The tail estimator is as follows:

$$F(x) = 1 - \frac{k}{n}\left[1 + \xi\frac{(x-u)}{\psi}\right]^{-\frac{1}{\xi}} \quad , \text{for } X > u, \tag{1}$$

where ξ is the shape parameter, and ψ is the scale parameter, n is the total number of observations, and k is the number of observations above the threshold u. For a given probability, $q > F(u)$, the tail quantile can be obtained

[3] See Coles (2001) and Beirlant et al. (2004) for detailed treatments of EVT.

by inverting the tail estimation formula above to get (see Embrechts, Kluppelberg, & Mikosh, 1997)

$$x_q = u + \frac{\psi}{\xi}\left[\left(\frac{1-q}{k/n}\right)^{-\xi} - 1\right] \tag{2}$$

2.2. Estimation of VaR and ES

As referred in the introduction, two important measures of market risk are the VaR, and the ES which are mathematically defined as follows: Suppose a random variable X with continuous distribution function F models the return distribution of a risky financial portfolio over the specified time horizon. For a given probability q, VaR can be defined as the qth quantile of the distribution F

$$VaR_q = F^{-1}(1-q) \tag{3}$$

where F^{-1} is the so-called quantile function defined as the inverse of the distribution function F. As VaR is exactly the same extreme quantile defined earlier by Eq. (2), it can be estimated by

$$V\hat{a}R_q = x_q = u + \frac{\psi}{\xi}\left[\left(\frac{1-q}{k/n}\right)^{-\xi} - 1\right] \tag{4}$$

The ES for risk X at given probability level q is formally defined as

$$ES_q = E(X|X > VaR_q) \tag{5}$$

The ES is estimated by the following equation

$$ES_q = \frac{VaR_q}{1-\xi} + \frac{\psi - \xi u}{1-\xi} \tag{6}$$

2.3. Conditional EVT applied to VaR and ES

So far we have described the Unconditional EVT approach that focuses directly on the tail but does not acknowledge the fact that financial asset returns are non-iid. McNeil and Frey (2000) recognize that most financial return series exhibit stochastic volatility and fat-tailed distributions. While the fat tails might be modeled directly with EVT, the lack of *iid* returns is problematic. One approach to this problem is provided by McNeil and Frey. Using a two-stage approach they estimate the conditional volatility using a GARCH model in stage one. The GARCH model serves to filter the return series such that GARCH residuals are closer to *iid* than the raw return series. Even so, GARCH residuals have been shown to exhibit fat tails. In stage two, they apply EVT to the GARCH residuals. As such, the GARCH–EVT combination accommodates both time-varying volatility and fat-tailed return distributions. In this paper, we follow McNeil and Frey (2000) in combining the extreme value approach with appropriate GARCH specification. This approach is denoted as Conditional EVT.

We assume that the dynamics of conditional mean returns can be represented by the following ARMA (p_1,q_1) model

$$r_t = a_0 + \sum_{i=1}^{p_1} a_i r_{t-i} + \sum_{j=1}^{q_1} b_j \varepsilon_{t-j} + \varepsilon_t = \mu_t + \sqrt{h_t} Z_t \tag{7}$$

where $\mu_t = a_0 + \sum_{i=1}^{p_1} a_i r_{t-i} + \sum_{j=1}^{q_1} b_j \varepsilon_{t-j}$, a_i and b_j are parameters, r_{t-i} are lagged returns, ε_t is residual which follows Student-t distribution with mean = 0, and variance = h_t, Z_t is the standardized residual which is defined by $\varepsilon_t/\sqrt{h_t}$ and h_t is conditional variance of ε_t. We also assume that the conditional variance h_t follows any one of the GARCH (p_2,q_2) processes including symmetric as well as asymmetric ones such as GARCH, EGARCH, TGARCH and APARCH which are briefly explained below:

Bollerslev (1986) proposes a generalized autoregressive conditional heteroskedasticity, GARCH (p_2,q_2) model:

$$h_t = \omega + \sum_{i=1}^{p_2} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j h_{t-j} \tag{8}$$

Nelson (1991) proposes the following Exponential GARCH (EGARCH) model to allow for leverage effects:

$$\log h_t = \omega + \sum_{i=1}^{p_2} \alpha_i \frac{|\varepsilon_{t-i}| + \gamma_i \varepsilon_{t-i}}{\sqrt{h_{t-i}}} + \sum_{j=1}^{q_2} \beta_j \log h_{t-j} \tag{9}$$

Another GARCH model that is capable of modeling asymmetric effects is the Threshold GARCH (TGARCH) model or also known as the GJR model (Glosten, Jagannathan, & Runkle, 1993), which has the following forms:

$$h_t = \omega + \sum_{i=1}^{p_2} (\alpha_i + \gamma_i D_{t-i}) \varepsilon_{t-i}^2 + \sum_{j=1}^{q_2} \beta_j h_{t-j} \tag{10}$$

Ding, Granger, and Engle (1993) propose the asymmetric power ARCH (APARCH) model where the power parameter δon the standard deviation is estimated rather than imposed:

$$\sigma_t^\delta = \omega + \sum_{i=1}^{p_2} \alpha_i (|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i})^\delta + \sum_{j=1}^{q_2} \beta_j \sigma_{t-j}^\delta \tag{11}$$

The above mentioned GARCH models would be fitted here with Student-t distribution instead of normal distribution. This is because empirical evidence strongly rejects the idea that financial returns are normally distributed. In fact, it is well established that the stock returns are fat-tailed. Hence we use the fat-tailed density, e.g., Student-t with GARCH in order to better account for heavy-tailedness.

Standardized residuals or innovations can be computed after maximizing log-likelihood function of Student-t distribution with respect to the unknown parameters:

$$Z_t = \frac{r_t - \mu_t}{\sqrt{h_t}} \tag{12}$$

If the standardized residuals are *iid* and the fitted model is well-specified, we end stage 1 by estimating the conditional mean (μ_{t+1}) and variance (h_{t+1}) for interval $t+1$ by using standard 1-step ahead forecasts.

The 1-step ahead conditional mean forecast is given by

$$\hat{\mu}_{t+1} = \hat{a}_0 + \sum_{i=1}^{p_1} \hat{a}_i r_{t-i+1} + \sum_{j=1}^{q_1} \hat{b}_j \varepsilon_{t-j+1} \tag{13}$$

and the 1-step ahead conditional variance forecast is given by Eqs. (14), (15), (16), and (17) for GARCH (p_2,q_2), EGARCH (p_2,q_2), TGARCH (p_2,q_2) and APARCH (p_2,q_2) models, respectively.

$$\hat{h}_{t+1} = \hat{\omega} + \sum_{i=1}^{p_2} \hat{\alpha}_i \varepsilon_{t-i+1}^2 + \sum_{j=1}^{q_2} \hat{\beta}_j \hat{h}_{t-j+1}] \tag{14}$$

$$\hat{h}_{t+1} = \exp\left[\hat{\omega} + \sum_{i=1}^{p_2} \hat{\alpha}_i \frac{|\varepsilon_{t-i+1}| + \hat{\gamma}_{i+1} \varepsilon_{t-i+1}}{\sqrt{\hat{h}_{t-i+1}}} + \sum_{j=1}^{q_2} \hat{\beta}_j \log \hat{h}_{t-j+1}\right] \tag{15}$$

Table 1
Name of countries and the corresponding indices along with other relevant information.

Country	Name of indices	Short forms of indices	Trading period (local time) during which price points are available	Obs./day	No. of days	Total obs.
India	S&P CNX Nifty Index	Nifty	9:10 h–15:30 h	76	152	11,552
Japan	Nikkei 225 Index	Nikkei	9:00 h–11:30 h; 12:30 h–15:05 h	61	150	9150
Taiwan	TSEC weighted index	TWSE	9:00 h–13:25 h	53	152	8056
Korea	Korea Stock Exchange KOSPI Index	KOSPI	9:00 h–15:00 h	70	151	10,570
Hong Kong	Hang Seng Index	HSI	9:20 h–12:00 h; 13:00 h–16:00 h	68	148	10,064
Philippines	Philippines Stock Exchange PSEi Index	Pcomp	9:30 h–12:00 h 13:30 h–15:15 h	51	150	7650
Singapore	Singapore Stock Market Index	STI	9:00 h–17:00 h	96	155	14,880
UK	FTSE 100 Index	FTSE	8:00 h–16:30 h	102	156	15,912
Germany	Deutsche Boerse AG German Stock Index	DAX	9:00 h–17:30 h	102	155	15,810
Turkey	XU100 Index	XU100	9:30 h–12:30 h 14:00 h–17:30 h	78	152	11,856
US	S&P 500 Index	SPX	9:30 h–16:00 h	78	155	12,090
Mexico	Mexican Bolsa IPC Index	Mexbol	8:30 h–14:55 h	77	151	11,627
Brazil	Ibovespa Brasil Sao Paulo Stock Exchange Index	Ibov	10:00 h–17:05 h	85	151	12,835
Argentina	Buenos Aires Stock Exchange Index	Merval	11:00 h–16:55 h	71	148	10,508
S Africa	Johanesburg Stock Exchange Index	Jalsh	9:00 h–16:55 h	95	151	14,345
Australia	ASX 200 Index	ASX	10:00 h–16:05 h	73	155	11,315

Note: The last three columns of Table 1 respectively, show number of sample observations in the return series per day, number of trading days and total number of sample observations finally after removing certain figures from the raw price set for each country. Let us see how we have arrived at the final number of observations for India. Consistent with the literature, overnight return is excluded. Again consistent with the literature, days which do not have 76 five-minute intervals are also excluded, which finally leaves us with $r_{t,v}$ $t = 1,..., 152$, $v = 1,...,$ 76, for a total of 11, 552 observations for India. Following the same procedure we have got the final numbers of sample observations in the return series for the rest of the other countries.

$$\hat{h}_{t+1} = \hat{\omega} + \sum_{i=1}^{p_2} (\hat{\alpha}_i + \hat{\gamma}_i D_{t-i+1}) \varepsilon_{t-i+1}^2 + \sum_{j=1}^{q_2} \hat{\beta}_j \hat{h}_{t-j+1} \tag{16}$$

$$\hat{\sigma}_{t+1}^{\delta} = \hat{\omega} + \sum_{i=1}^{p_2} \hat{\alpha}_i (|\varepsilon_{t-i+1}| - \hat{\gamma}_i \varepsilon_{t-i+1})^{\delta} + \sum_{j=1}^{q_2} \hat{\beta}_j \hat{\sigma}_{t-j+1}^{\delta} \tag{17}$$

In stage 2 we apply the EVT tools kit to the standardized residuals (Z_t) and estimate the $V\hat{a}R_q$ and ES_q quantiles defined by Eqs. (4) and (6), respectively. An estimate of the Conditional VaR is

$$VaR_q^{t+1} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}}\ V\hat{a}R_q \tag{18}$$

And an estimate of the Conditional ES is

$$ES_q^{t+1} = \hat{\mu}_{t+1} + \sqrt{\hat{h}_{t+1}} ES_q \tag{19}$$

where $\hat{\mu}_{t+1}$ is given by Eq. (13) and $\hat{h}_{t+1}$ is given by Eqs. ((14) or (15) or (16) or (17), whichever is appropriate).

3. Data, properties and the stylized facts

The data set used in this study includes sixteen high frequency stock price indices across Asia, Europe, the US, Latin America, Africa and Australia. The sixteen indices have been taken from sixteen countries, one each from every country. The names of the countries and their corresponding indices along with short forms are reported in first three columns of Table 1. The price index data have been extracted from Bloomberg for a period of seven and half months from September 20, 2013 to May 07, 2014 at 5 min interval. It is noted that stock trading hours vary from country to country and the specific trading periods during which the price records are available are reported for different countries in column 4 of Table 1.

As high-frequency data carry more information, using data with the highest possible frequency theoretically optimizes the accuracy of the intraday VaR estimation. However, many researchers have expressed concerns over the adverse effect of microstructure noise of the market on the high frequency estimator (e.g., Andersen & Bollerslev, 1998; Andersen, Bollerslev, Diebold, & Labys, 1999; Alizadeh, Brandt, & Diebold, 2002; Bandi & Rusell, 2005; Zhang, Mykland, & Ait-Sahalia, 2005). Andersen and Bollerslev (1998) choose 5 min frequency and suggest that the 5 min frequency is the best way to avoid microstructure error of the market.[4] So we choose the 5-min frequency data of each of the sixteen indices. For each series we obtain 5 min continuously compounded returns ($r_{t,v}$) for each interval v on day t and the return is calculated as the logarithmic difference of prices, i.e., $r_{t,v} = \log(P_{t,v}/P_{t,v-1})$, where $P_{t,v}$ is the closing price for interval v on day t and $P_{t,v\text{-}1}$ is the opening price for interval v on day t. The last three columns of Table 1 respectively, show number of sample observations in the return series per day, number of trading days and total number of sample observations finally after removing certain figures from the raw price set for each country. Let us see how we have arrived at the final number of observations for India. Consistent with the literature, overnight return is excluded. Again consistent with the literature, days which do not have 76 five-minute intervals are also excluded, which finally leaves us with $r_{t,v}$, $t = 1,..., 152$, $n = 1,..., 76$, for a total of 11,552 observations for India. Following the same procedure we have got the final numbers of sample observations in the return series for the rest of the other countries.

Table 2 presents the summary statistics of the distribution of returns for each series. The estimated results suggest that the characteristics of 5 min equity returns can be consistent with the stylized properties of high frequency financial time series returns documented in the literature. For example, the mean values of the series are all approximately zero, as is the case for the returns of other financial assets. The sample skewness in most of the countries, is negative which suggests that the negative shocks are more frequent than the positive ones. The skewness is positive only for Japan, Singapore, Mexico and Argentina. The excess-kurtosis estimate is very high in majority of the countries and there is a wide variation in the estimates with the highest value of 65.249 in South Africa and the lowest value of 4.709 in Singapore. The average excess-kurtosis is very high (18.25) which means that return distributions are leptokurtic, with much heavier tails than the normal distribution.

[4] Andersen and Bollerslev (1998) explore the effective use of time series data with different frequencies in constructing accurate ex-post volatility measurements. They observe the least measurement error in case of 5-min frequency data and hence suggest that 5-min frequency data is the best way to avoid microstructure noise.

Table 2
Descriptive statistics of original returns (full sample from September 20, 2013 to May 07, 2014).

	Total obs.	Mean	Std. Dev.	Skewness	Kurtosis	Jarque–Bera	Q (16)	Q^2 (16)
India	11,552	−0.000	0.001	−0.739	17.291	99,359.97 (0.000)	15.197 (0.510)	808.030 (0.000)
Japan	9150	−0.000	0.001	0.084	9.419	15,719.44 (0.000)	25.568 (0.060)	119.260 (0.000)
Taiwan	8056	−0.000	0.001	−0.025	8.235	9200.24 (0.000)	37.648 (0.002)	374.520 (0.000)
Korea	10,570	−0.000	0.001	−0.258	8.505	13,465.89 (0.000)	13.924 (0.604)	757.840 (0.000)
Hong Kong	10,064	−0.000	0.001	−1.036	20.507	130,321.20 (0.000)	40.948 (0.001)	153.250 (0.000)
Philippines	7650	−0.000	0.001	−1.730	34.042	310,962.20 (0.000)	514.780 (0.000)	137.770 (0.000)
Singapore	14,880	−0.000	0.000	0.014	4.709	1812.19 (0.000)	401.500 (0.000)	929.860 (0.000)
UK	15,912	0.000	0.001	−0.399	13.899	79,184.01 (0.000)	25.339 (0.064)	961.100 (0.000)
Germany	15,810	0.000	0.001	−0.322	11.409	46,848.90 (0.000)	42.470 (0.000)	1249.500 (0.000)
Turkey	11,856	−0.000	0.001	−1.094	27.721	304,255.40 (0.000)	50.274 (0.000)	906.821 (0.000)
US	12,090	0.000	0.001	−0.115	7.046	8274.72 (0.000)	18.239 (0.310)	2394.100 (0.000)
Mexico	11,627	0.000	0.001	0.291	9.929	23,424.37 (0.000)	87.929 (0.000)	916.930 (0.000)
Brazil	12,835	−0.000	0.001	−0.143	7.455	10,658.21 (0.000)	28.244 (0.030)	506.410 (0.000)
Argentina	10,508	−0.000	0.001	0.575	24.054	194,665.40 (0.000)	658.720 (0.000)	422.200 (0.000)
S Africa	14,345	0.000	0.001	−0.936	65.249	2,318,191.00 (0.000)	23.953 (0.091)	8.090 (0.946)
Australia	11,315	−0.000	0.001	−0.558	22.710	183,730.59 (0.000)	30.905 (0.014)	52.522 (0.000)

Note: The table reports summary statistics for the 5-min interval stock market returns (r_t) of the sixteen countries. The p values are given in the parentheses which indicate that the Ljung–Box Q and Q^2 statistics are significant at better than 1% levels except Q(16) for India, Japan, South Korea, UK, US, and South Africa and Q^2(16) for South Africa.

The non-normality of the distribution is also confirmed by the high Jarque–Bera statistics. On the basis of Ljung–Box Q statistic, the hypothesis that all correlation coefficients up to 16 are jointly zero is rejected for all countries. Therefore, we can conclude that return series in each country present some linear dependence in returns. In addition, the statistically significant serial correlations in squared returns [Q^2 (16)] imply that there are non-linear dependences in all return series. This indicates volatility clustering which is clearly observed in each return series plotted in Fig. 1. Together with Table 2, Fig. 1 demonstrates the defining characteristics of high frequency stock returns of sixteen countries: occasional extreme movements, volatility clustering and fat-tailed distributions.

Another stylized property of high-frequency returns, which needs a special mention and has been documented in many studies, is that most intraday equity return volatilities exhibit strong periodicity (see Andersen & Bollerslev, 1997; Aradhyula & Ergun, 2004: Bollerslev & Ghysels, 1996; Goodhart & O'Hara, 1997; Martens, Chang, & Taylor, 2002). Volatility is typically higher at the opening and towards the close of trade and lower during midday. To investigate the periodicity of the intraday volatility, we have estimated ACF of the absolute returns for each return series and plotted the same in Fig. 2. The apparent U-shaped periodicity recurs every day for almost all the series. Consider the case for India where there are 76 observations per day for the 5 min return series. The U-shaped pattern can be observed in India at every 76 lags, which strongly indicates periodicity with a period of one day. The autocorrelation is highest at the beginning and end of one day interval, and lowest in the middle. For other countries too, the same pattern can be observed at every v_i lags where v_i is the number of observations per day of the *ith* country, indicating periodicity with a period of one day. The periodicity of the intraday volatility observed in the present study is consistent with the findings of various other studies cited earlier. The evidence of this intraday periodicity reflects the behavior of traders who are very active at the beginning of the trading session and adjust their positions to incorporate the overnight change in information. Towards the end of the day, traders are changing their positions in anticipation of the close and to pre-empt the risk posed by any information that could arrive during the night.

In summary, the 5 min returns series used in this study have properties that are consistent with the stylized facts of high frequency financial returns reported in the literature. They are all fat-tailed, slightly skewed and have a zero mean. Furthermore, there are linear dependence in returns and the series exhibit volatility clustering. Most importantly, the series displays strong periodicity patterns in intraday volatility. While the evidence of volatility clustering needs an appropriate GARCH model to filter the return series of each market separately, the existence of occasional extreme movements and fat-tailed distribution further motivate the exploration of conditional EVT to estimate intraday VaR and ES. The evidence of intraday periodicity also requires an adjustment for intraday risk measurement.

4. Empirical findings

We divide the empirical study into two parts: an in-sample study where the sample data are used for model estimation, and an out-of-sample study to compare the accuracy of Conditional EVT approach to VaR and ES calculation with other competing models.[5] To do this, the full data sample in each market is divided into an in-sample period from September 20, 2013 to March 19, 2014 on which models are based and an out-of-sample period from March 20, 2014 to May 07, 2014 over which forecasting performance of VaR and ES is measured. All through this section, we look at losses; that is, we have chosen to look at extremes in the negative part of the original return distribution.

4.1. In-sample evidence

The first step is to model the conditional volatility of in-sample intraday return series using an appropriate GARCH approach. The volatilities of intraday return series typically display a strong periodicity in intervals of 24 h, as has been demonstrated in the previous section. Andersen and Bollerslev (1997) and Martens et al. (2002), among others, show that the estimates of traditional time-series models (e.g., GARCH-type models) can be ruined by intraday periodic patterns. Therefore, to prevent distortion of the results, the intraday seasonality must be taken out prior to estimating any model. Hence, we first remove the seasonality from the return series and then use the deseasonalized filtered returns to estimate the traditional time series models. Following a method proposed by Taylor and Xu (1997) and subsequently used by others, we describe the deseasonalized filtered return as the nth intraday return divided by an estimated seasonality term,

$$\tilde{r}_{t,v} = r_{t,v}/S_{t,v} \quad (v = 1, 2, \ldots\ldots\ldots, V), \tag{20}$$

[5] The other competing models include Unconditional EVT, Static Normal, Conditional Normal, and RiskMetrics. The Unconditional EVT has already been discussed in the text and the rest of the three models are well known and are explained in different studies (See Gencay & Selcuk, 2004; Karmakar & Shukla, 2015).

where $r_{t,v}$ is the vth intraday return on day t and $S_{t,v}$ is the respective seasonality term, for V intraday periods. The seasonality term includes averaging the squared returns for each intraday period, i.e.:

$$\hat{S}_{t,v}^2 = \frac{1}{T}\sum_{t=1}^{T} r_{t,v}^2 \quad (v = 1, 2, \ldots, V), \tag{21}$$

where T is the total number of days in the sample. This method seems to be quite effective as it almost removes the U-shaped pattern from all series shown in Fig. 2. Thus, we use the deseasonalized returns to estimate the intraday GARCH models.

The appropriate model for each country is selected by adopting the following procedure. First, various GARCH models i.e., ARMA (0,0)-GARCH (1,1) models are estimated and compared using the usual information criteria such as AIC, BIC and Log-likelihood statistics. Once the model is selected, the ARMA–GARCH specifications are augmented with additional AR, MA and ARCH, GARCH lagged terms when necessary to eliminate autocorrelation in the standardized and squared standardized residuals, respectively. Based on this procedure, we have selected the appropriate ARMA (p_1, q_1)–GARCH (p_2, q_2) model on different deseasonalized return series for each country and the selected models are reported in Table 3. It appears from the table that while asymmetric GARCH type model has fitted

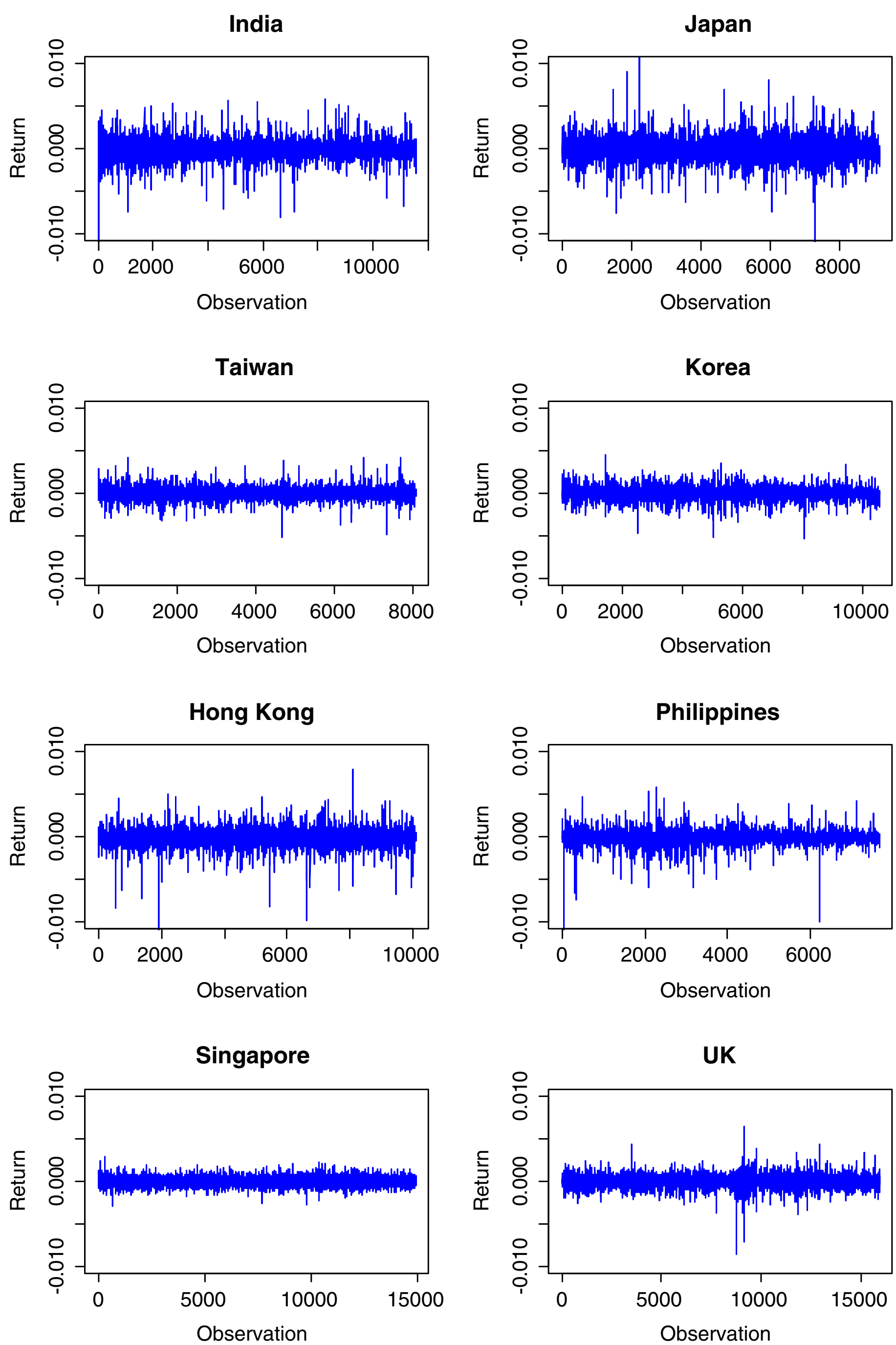

Fig. 1. Plot of return series of the indices over the time period.

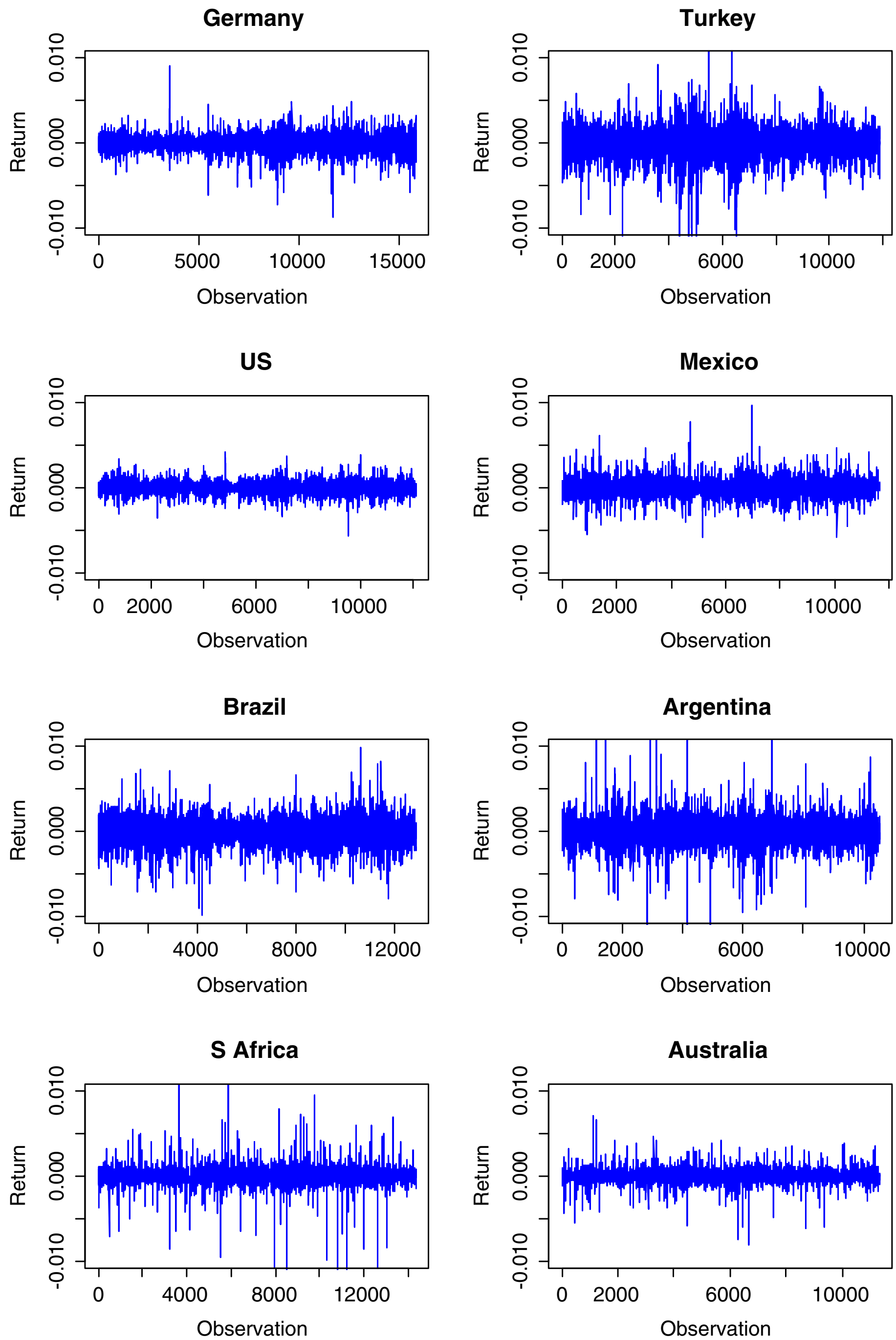

Fig. 1 (*continued*).

volatilities of most of the countries, symmetric GARCH has modeled volatilities of only two countries: Hong Kong and Singapore.

Table 3 also presents the estimated parameters of the mean and variance equations of the selected models. The constant term, AR (p_1), and MA (q_1) coefficients in the mean equation are mostly significant. Similarly, the parameters in the variance equation: the constant, the ARCH (p_2) coefficients, and the GARCH (q_2) coefficients are significant in majority of the cases. The values of γ_1 are also mostly significant, which suggests that the conditional variance in majority of the countries is an asymmetric function of past innovations, rising proportionately more during market declines. Estimates of the conditional mean and standard deviation series $(\hat{\mu}_{t-v+1},\ldots,\hat{\mu}_t)$ and $(\sqrt{\hat{h}_{t-v+1}},\ldots,\sqrt{\hat{h}_t})$ can be calculated recursively from Eqs. (13), and (14) or (15) or (16) or (17), whichever is appropriate, respectively after substitution of sensible starting values.[6]

[6] Here it needs to clarify how to choose the starting values when estimating conditional means and variances in the present study. We have chosen the starting values as the unconditional estimates of means and variances of the in-sample return series. This is indeed the choice of almost all software packages that estimate ARMA-GARCH type models. In fact, the choice of alternative starting values does not matter much on mean and variance forecasting for a larger sample, which can be evidenced from the illustration given in Appendix A.

Standardized residuals are calculated both to check adequacy of the selected models and to use in stage 2 of the method. They are calculated as

$$z_{t-v+1} = \frac{r_{t-v+1} - \hat{\mu}_{t-v+1}}{\sqrt{\hat{h}_{t-v+1}}}, \ldots\ldots\ldots\ldots, z_t = \frac{r_t - \hat{\mu}_t}{\sqrt{\hat{h}_t}} \tag{22}$$

and should be *iid* if the fitted model is tenable.

Panels A and B of Table 4, present diagnostic statistics of deseasonalized returns and their standardized residuals, respectively. The Ljung–Box Q and Q^2 statistics respectively provide an indication of whether any serial correlation and heteroscedasticity are present in the data series. The results in Panel A strongly suggests that the deseasonalized returns are not *iid* as required by EVT. In contrast, their standardized residuals in Panel B are close to *iid*. Thus the filtering procedure advocated by McNeil and Frey (2000) has been effective in producing *iid* residuals on which EVT can be implemented. Q^2 (16) statistic of standardized residuals in all series failed to detect serial correlations in squared standardized residuals, suggesting that the selected GARCH models are well specified. However, it appears from Panel B that skewness and excess kurtosis remain in the standardized residuals. It is also noted that neither the deseasonalized return series nor their standardized residual series are normally distributed as suggested by Jarque–Bera statistics.

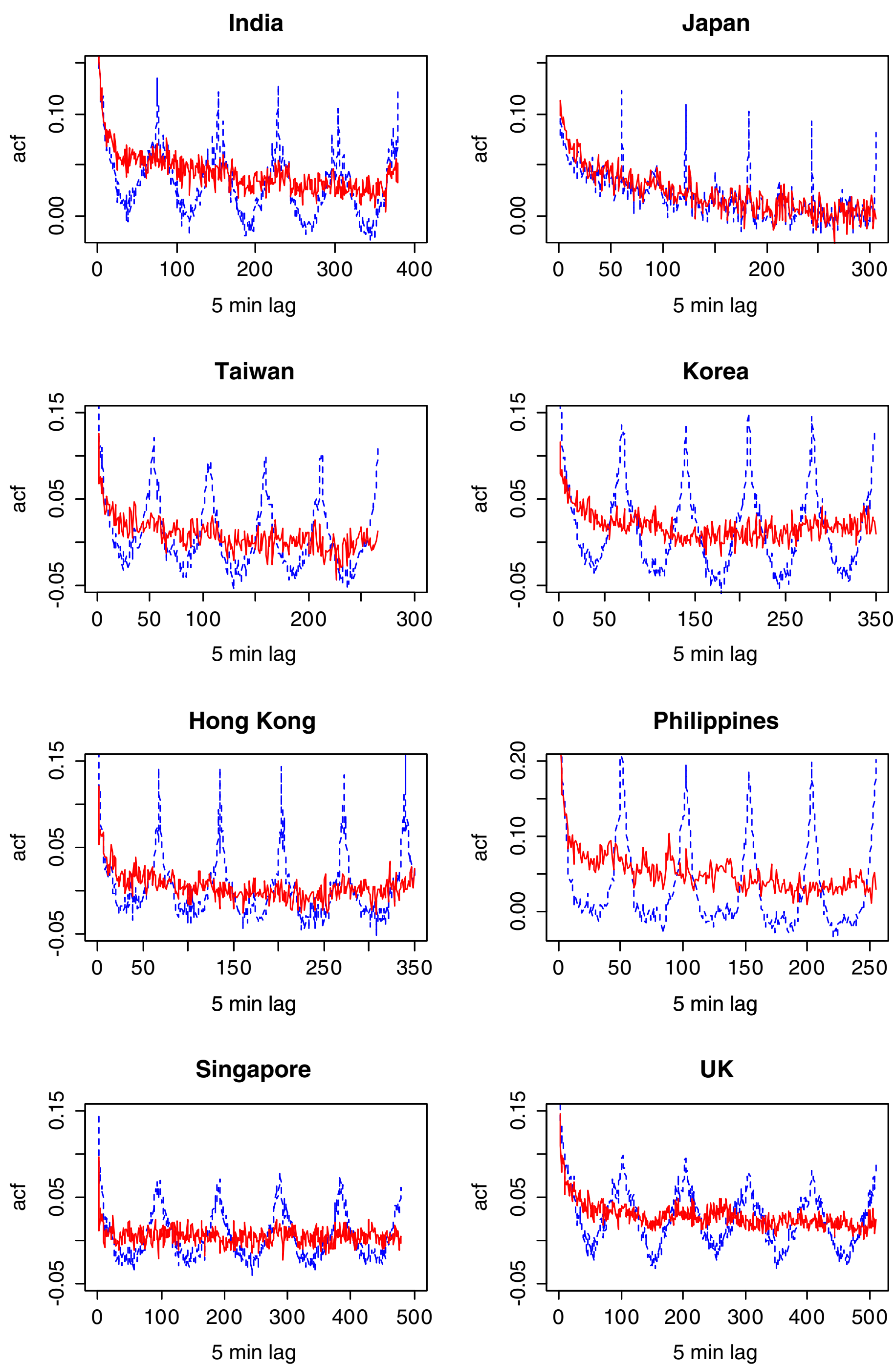

Fig. 2. Intraday return (dotted) and deseasonalized return series (line) for five consecutive days.

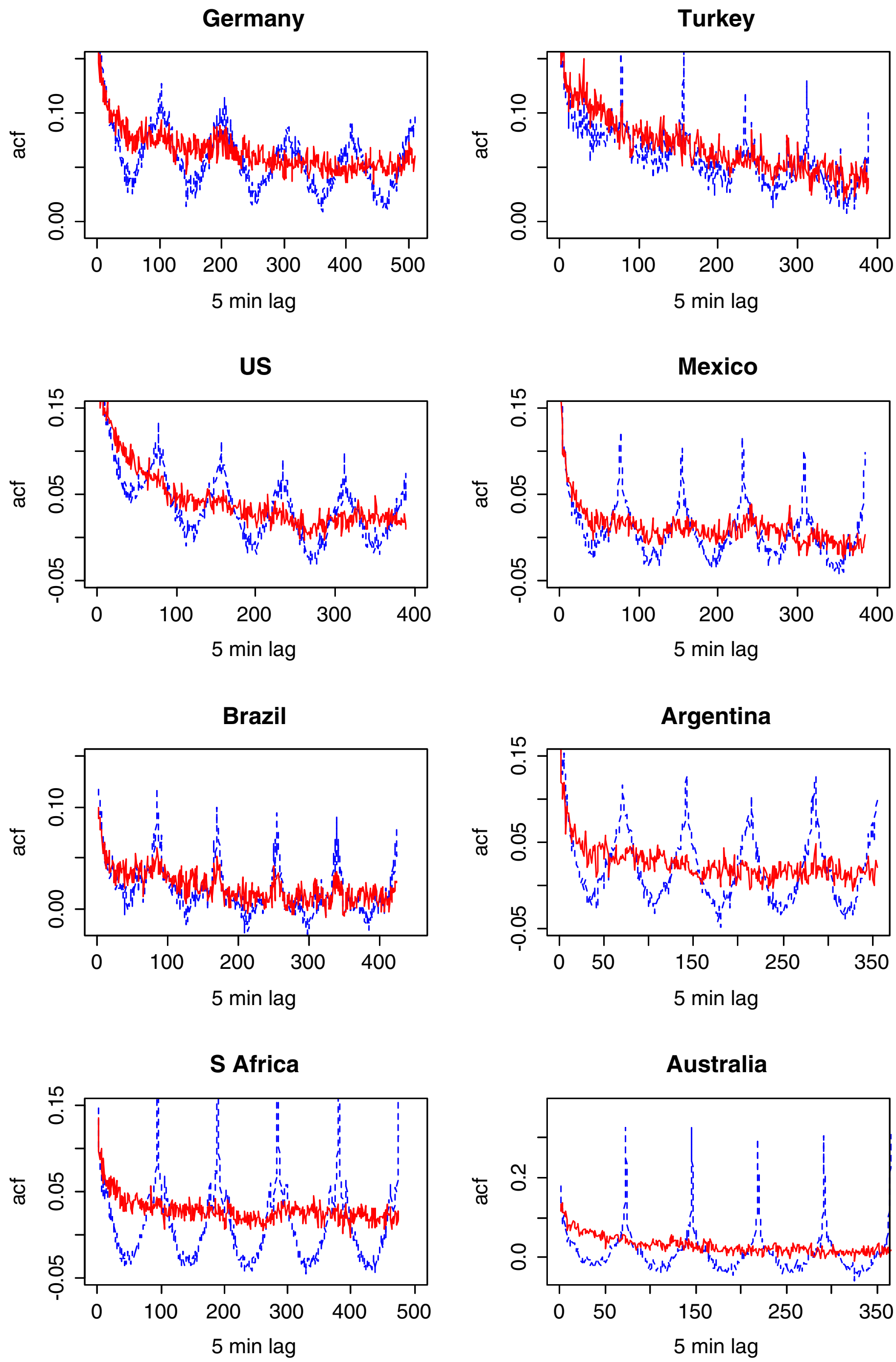

Fig. 2 (*continued*).

All these findings motivate the second stage of McNeil and Frey's EVT implementation, where fat tails of the standardized residuals are explicitly modeled.

As mentioned earlier, we would employ the POT method using GPD for tail estimation of the standardized residual series. The first step in this modeling is to choose the threshold for identifying the relevant tail region. However, this choice is subject to a trade-off between variance and bias. By increasing the number of observations for the series of maxima (a lower threshold), some observations from the centre of the distribution are introduced in the series, and the tail index is more precise but biased (i.e., there is less variance). On the other hand, choosing a high threshold reduces the bias but makes the estimator more volatile (i.e., there are fewer observations). The problem of finding an optimal threshold is very subjective: we need to find a sufficiently high threshold u, above which the GPD is a reasonable model of exceedances. However, the threshold must also be chosen such that we have sufficient data to accurately estimate parameters of the distribution.

There is no unique choice of the threshold level. A number of diagnostic techniques exist for this purpose, including graphical bootstrap methods (see Embrechts et al., 1997; Reiss & Thomas, 1997). To optimize this trade-off between variance and bias inefficiency, we perform

Table 3
Parameter estimates for the ARMA–GARCH model (in-sample from September 20, 2013 to March 19, 2014).

	Model fitted	Mean equation		Variance equation	
		Coefficient	Probability	Coefficient	Probability
India	ARMA (1,1)–APARCH (2,3)	$a_0 = 0.010$	$a_0 = 0.138$	$\omega = 0.001$	$\omega = 0.034$
		$a_1 = 0.481$	$a_1 = 0.000$	$\alpha_1 = 0.108$	$\alpha_1 = 0.000$
		$b_1 = -0.572$	$b_1 = 0.000$	$\alpha_2 = -0.101$	$\alpha_2 = 0.000$
				$\gamma_1 = 0.005$	$\gamma_1 = 0.170$
				$\beta_1 = 1.819$	$\beta_1 = 0.000$
				$\beta_2 = -1.034$	$\beta_2 = 0.000$
				$\beta_3 = 0.208$	$\beta_3 = 0.055$
				$\delta_1 = 1.591$	$\delta_1 = 0.000$
Japan	ARMA (3,1)–TARCH (2,1)	$a_0 = -0.001$	$a_0 = 0.946$	$\omega = 0.018$	$\omega = 0.000$
		$a_1 = 0.490$	$a_1 = 0.000$	$\alpha_1 = 0.043$	$\alpha_1 = 0.004$
		$a_2 = -0.053$	$a_2 = 0.000$	$\alpha_2 = -0.012$	$\alpha_2 = 0.439$
		$a_3 = -0.031$	$a_3 = 0.060$	$\gamma_1 = 0.034$	$\gamma_1 = 0.000$
		$b_1 = -0.486$	$b_1 = 0.000$	$\beta_1 = 0.935$	$\beta_1 = 0.000$
Taiwan	ARMA (0,1)–EGARCH (1,1)	$a_0 = -0.026$	$a_0 = 0.007$	$\omega = -0.109$	$\omega = 0.000$
		$b_1 = -0.157$	$b_1 = 0.000$	$\alpha_1 = 0.138$	$\alpha_1 = 0.000$
				$\gamma_1 = -0.037$	$\gamma_1 = 0.000$
				$\beta_1 = 0.912$	$\beta_1 = 0.000$
Korea	ARMA (2,1)–EGARCH (2,1)	$a_0 = -0.006$	$a_0 = 0.554$	$\omega = -0.078$	$\omega = 0.000$
		$a_1 = 0.516$	$a_1 = 0.033$	$\alpha_1 = 0.206$	$\alpha_1 = 0.000$
		$a_2 = 0.019$	$a_2 = 0.085$	$\alpha_2 = -0.103$	$\alpha_2 = 0.000$
		$b_1 = -0.537$	$b_1 = 0.027$	$\gamma_1 = -0.017$	$\gamma_1 = 0.008$
				$\beta_1 = 0.975$	$\beta_1 = 0.000$
Hong Kong	ARMA (0,2)–GARCH (1,1)	$a_0 = -0.006$	$a_0 = 0.440$	$\omega = 0.071$	$\omega = 0.000$
		$b_1 = -0.050$	$b_1 = 0.000$	$\alpha_1 = 0.059$	$\alpha_1 = 0.000$
		$b_2 = -0.070$	$b_2 = 0.000$	$\beta_1 = 0.875$	$\beta_1 = 0.000$
Philippines	ARMA (1,2)–APARCH (2,2)	$a_0 = -0.008$	$a_0 = 0.559$	$\omega = 0.002$	$\omega = 0.036$
		$a_1 = 0.573$	$a_1 = 0.000$	$\alpha_1 = 0.157$	$\alpha_1 = 0.000$
		$b_1 = -0.468$	$b_1 = 0.000$	$\alpha_2 = -0.146$	$\alpha_2 = 0.000$
		$b_2 = 0.042$	$b_2 = 0.022$	$\gamma_1 = 0.004$	$\gamma_1 = 0.284$
				$\beta_1 = 1.582$	$\beta_1 = 0.000$
				$\beta_2 = -0.593$	$\beta_2 = 0.000$
				$\delta_1 = 1.689$	$\delta_1 = 0.000$
Singapore	ARMA (0,1)–GARCH (2,1)	$a_0 = -0.018$	$a_0 = 0.008$	$\omega = 0.087$	$\omega = 0.001$
		$b_1 = -0.224$	$b_1 = 0.000$	$\alpha_1 = 0.078$	$\alpha_1 = 0.000$
				$\alpha_2 = -0.039$	$\alpha_2 = 0.003$
				$\beta_1 = 0.870$	$\beta_1 = 0.000$
UK	ARMA (0,0)–TARCH (2,2)	$a_0 = 0.014$	$a_0 = 0.053$	$\omega = 0.010$	$\omega = 0.002$
				$\alpha_1 = 0.103$	$\alpha_1 = 0.000$
				$\alpha_2 = -0.082$	$\alpha_2 = 0.000$
				$\gamma_1 = 0.013$	$\gamma_1 = 0.010$
				$\beta_1 = 1.275$	$\beta_1 = 0.000$
				$\beta_2 = -0.312$	$\beta_2 = 0.024$
Germany	ARMA (1,2)–EGARCH (2,3)	$a_0 = 0.011$	$a_0 = 0.094$	$\omega = -0.008$	$\omega = 0.000$
		$a_1 = -0.665$	$a_1 = 0.000$	$\alpha_1 = 0.220$	$\alpha_1 = 0.000$
		$b_1 = 0.622$	$b_1 = 0.000$	$\alpha_2 = -0.209$	$\alpha_2 = 0.000$
		$b_2 = -0.058$	$b_2 = 0.000$	$\gamma_1 = -0.004$	$\gamma_1 = 0.000$
				$\beta_1 = 1.770$	$\beta_1 = 0.000$
				$\beta_2 = -0.713$	$\beta_2 = 0.000$
				$\beta_3 = -0.058$	$\beta_3 = 0.000$
Turkey	ARMA (2,2)–TARCH (3,3)	$a_0 = 0.011$	$a_0 = 0.119$	$\omega = 0.002$	$\omega = 0.000$
		$a_1 = 1.291$	$a_1 = 0.000$	$\alpha_1 = 0.093$	$\alpha_1 = 0.000$
		$a_2 = -0.495$	$a_2 = 0.000$	$\alpha_2 = -0.039$	$\alpha_2 = 0.001$
		$b_1 = -1.342$	$b_1 = 0.000$	$\alpha_3 = -0.049$	$\alpha_3 = 0.000$
		$b_2 = 0.525$	$b_2 = 0.000$	$\gamma_1 = 0.007$	$\gamma_1 = 0.000$
				$\beta_1 = 0.831$	$\beta_1 = 0.000$
				$\beta_2 = 0.770$	$\beta_2 = 0.000$
				$\beta_2 = -0.611$	$\beta_2 = 0.000$
US	ARMA (0,0) – EGARCH (1,1)	$a_0 = 0.022$	$a_0 = 0.007$	$\omega = -0.092$	$\omega = 0.000$
				$\alpha_1 = 0.120$	$\alpha_1 = 0.000$
				$\gamma_1 = -0.043$	$\gamma_1 = 0.000$
				$\beta_1 = 0.990$	$\beta_1 = 0.000$
Mexico	ARMA (1,1)–APARCH (2,1)	$a_0 = -0.003$	$a_0 = 0.762$	$\omega = 0.047$	$\omega = 0.000$
		$a_1 = -0.324$	$a_1 = 0.007$	$\alpha_1 = 0.189$	$\alpha_1 = 0.000$
		$b_1 = 0.395$	$b_1 = 0.001$	$\alpha_2 = -0.0923$	$\alpha_2 = 0.000$
				$\gamma_1 = 0.040$	$\gamma_1 = 0.032$
				$\beta_1 = 0.870$	$\beta_1 = 0.000$
				$\delta_1 = 1.658$	$\delta_1 = 0.000$
Brazil	ARMA (1,0)–EGARCH (2,1)	$a_0 = -0.018$	$a_0 = 0.027$	$\omega = -0.058$	$\omega = 0.000$
		$a_1 = -0.060$	$a_1 = 0.000$	$\alpha_1 = 0.164$	$\alpha_1 = 0.000$
				$\alpha_2 = -0.087$	$\alpha_2 = 0.000$
				$\gamma_1 = -0.003$	$\gamma_1 = 0.423$
				$\beta_1 = 0.993$	$\beta_1 = 0.000$

(continued on next page)

Table 3 (*continued*)

	Model fitted	Mean equation		Variance equation	
		Coefficient	Probability	Coefficient	Probability
Argentina	ARMA (1,2)–TARCH (2,1)	$a_0 = -0.016$	$a_0 = 0.299$	$\omega = 0.026$	$\omega = 0.000$
		$a_1 = 0.851$	$a_1 = 0.000$	$\alpha_1 = 0.114$	$\alpha_1 = 0.000$
		$b_1 = -0.803$	$b_1 = 0.000$	$\alpha_2 = -0.073$	$\alpha_2 = 0.000$
		$b_2 = 0.046$	$b_2 = 0.000$	$\gamma_1 = 0.020$	$\gamma_1 = 0.014$
				$\beta_1 = 0.921$	$\beta_1 = 0.000$
S Africa	ARMA (0,2)–EGARCH (1,1)	$a_0 = 0.005$	$a_0 = 0.516$	$\omega = -0.088$	$\omega = 0.000$
		$b_1 = -0.012$	$b_1 = 0.187$	$\alpha_1 = 0.113$	$\alpha_1 = 0.000$
		$b_2 = -0.023$	$b_2 = 0.013$	$\gamma_1 = -0.017$	$\gamma_1 = 0.001$
				$\beta_1 = 0.982$	$\beta_1 = 0.000$
Australia	ARMA (2,2)–APARCH (2,1)	$a_0 = -0.005$	$a_0 = 0.617$	$\omega = 0.016$	$\omega = 0.000$
		$a_1 = -0.024$	$a_1 = 0.870$	$\alpha_1 = 0.099$	$\alpha_1 = 0.000$
		$a_2 = -0.701$	$a_2 = 0.000$	$\alpha_2 = -0.042$	$\alpha_2 = 0.007$
		$b_1 = 0.044$	$b_1 = 0.762$	$\gamma_1 = 0.113$	$\gamma_1 = 0.007$
		$b_2 = 0.696$	$b_2 = 0.000$	$\beta_1 = 0.938$	$\beta_1 = 0.000$
				$\delta_1 = 1.452$	$\delta_1 = 0.000$

Note: The table reports ML estimates of the fitted ARMA–GARCH models with Student-*t* distribution governing the error terms. For each data series, parameter estimates are based on the in-sample period from September 20, 2013 to March 19, 2014. The majority of parameter estimates are statistically significant at better than 1% level.

a Monte Carlo simulation study. Return time-series are simulated from a known distribution for which the tail index can be computed. For each time series, the tail index value is estimated for different threshold levels. The choice of the optimal value is based on the bias and the mean squared error (MSE) criteria which allow one to take into account the trade-off between bias and inefficiency. The procedure is detailed in Appendix B.

Typically the threshold is chosen subjectively by looking at certain plots such as the Mean Excess plot or the Hill plot, which are standard practices in EVT. Here we would select the threshold using the Mean Excess plot which is a plot of Mean Excess Function (MEF). The MEF is defined by

$$e(u) = \frac{\sum_{i=1}^{n}(X_i - u)}{\sum_{i=1}^{n} I_{\{X_i > u\}}} \tag{23}$$

Table 4
Diagnostic statistics of deseasonalized returns and ARMA–GARCH standardized residuals (in-sample from September 20, 2013 to Mrach 19, 2014).

	Skewness	Kurtosis	Jarque–Bera	Q (16)	Q^2 (16)
Panel A: In-sample return series					
India	−0.337	6.718	5515.23 (0.000)	29.709 (0.020)	1368.400 (0.000)
Japan	−0.022	4.869	1056.54 (0.000)	44.398 (0.000)	596.760 (0.000)
Taiwan	0.013	4.329	464.31 (0.000)	101.960 (0.000)	141.730 (0.000)
Korea	−0.190	6.817	5251.75 (0.000)	12.763 (0.690)	412.150 (0.000)
Hong Kong	−0.238	7.959	8228.52 (0.000)	33.427 (0.000)	241.020 (0.000)
Philippines	0.104	6.148	2538.24 (0.000)	456.790 (0.000)	1444.000 (0.000)
Singapore	−0.030	3.504	125.58 (0.000)	472.070 (0.000)	179.980 (0.000)
UK	−0.164	7.223	9456.64 (0.000)	17.974 (0.325)	1222.100 (0.000)
Germany	−0.153	8.291	14,682.56 (0.000)	51.788 (0.000)	1559.100 (0.000)
Turkey	−0.440	7.433	7900.46 (0.000)	35.719 (0.000)	2638.300 (0.000)
US	−0.089	5.931	3417.78 (0.000)	14.248 (0.580)	2246.200 (0.000)
Mexico	0.033	6.750	5415.26 (0.000)	105.800 (0.000)	1232.400 (0.000)
Brazil	−0.123	4.828	1445.63 (0.000)	43.927 (0.000)	488.640 (0.000)
Argentina	−0.097	6.510	4387.55 (0.000)	589.600 (0.000)	1067.100 (0.000)
S Africa	−0.274	5.296	2691.51 (0.000)	32.964 (0.007)	1209.800 (0.000)
Australia	−0.187	4.769	1223.04 (0.000)	41.578 (0.000)	1343.100 (0.000)
Panel B: Standardized residuals					
India	−0.271	5.910	3385.13 (0.000)	36.600 (0.001)	22.139 (0.076)
Japan	−0.004	4.860	1045.96 (0.000)	35.210 (0.000)	13.718 (0.319)
Taiwan	0.029	4.290	438.31 (0.000)	16.380 (0.357)	19.404 (0.196)
Korea	0.005	5.094	1521.22 (0.000)	14.475 (0.341)	11.433 (0.575)
Hong Kong	−0.262	7.984	8326.24 (0.000)	33.947 (0.002)	20.715 (0.109)
Philippines	0.052	4.540	607.00 (0.000)	23.517 (0.036)	20.657 (0.080)
Singapore	−0.051	3.525	139.44 (0.000)	19.931 (0.175)	14.347 (0.499)
UK	−0.101	4.865	1854.78 (0.000)	17.467 (0.356)	21.584 (0.157)
Germany	−0.112	6.970	8264.67 (0.000)	19.281 (0.115)	22.079 (0.054)
Turkey	−0.383	6.331	4517.08 (0.000)	45.201 (0.000)	13.470 (0.490)
US	−0.225	5.002	1668.73 (0.000)	9.228 (0.904)	14.399 (0.276)
Mexico	−0.018	5.830	3084.33 (0.000)	34.054 (0.002)	10.777 (0.703)
Brazil	−0.256	6.384	4978.08 (0.000)	17.588 (0.285)	12.587 (0.634)
Argentina	−0.047	5.686	2563.79 (0.000)	15.813 (0.259)	14.899 (0.314)
S Africa	−0.306	7.217	8769.10 (0.000)	17.047 (0.254)	0.158 (1.000)
Australia	−0.189	4.765	1218.20 (0.000)	23.897 (0.021)	19.498 (0.077)

Note: The table reports summary statistics for the in-sample deseasonalized returns and standardized residuals from the fitted ARMA–GARCH models with Student-*t* distribution governing the error terms. Panels A and B report diagnostics for the deseasonalized returns and standardized residuals, respectively. The latter are the basis of the EVT estimation. The *p* values of Ljung–Box Q (16) and Q^2 (16) statistics are given in the parentheses.

Table 5
Parameter estimates for the ARMA–GARCH–EVT called Conditional EVT model (in-sample from September 20, 2013 to March 19, 2014).

	T	u	k	k/T (%)	ξ	ψ	VaR quantile			ES quantile		
							0.95	0.99	0.995	0.95	0.99	0.995
India	9272	0.985	1282	13.83	0.079** (2.812)	0.605** (25.311)	1.608 (0.02)	2.825 (0.04)	3.405 (0.08)	2.318 (0.03)	3.640 (0.09)	4.269 (0.11)
Japan	7259	1.251	667	9.19	0.040 (1.039)	0.568** (18.190)	1.600 (0.02)	2.567 (0.04)	3.003 (0.06)	2.206 (0.03)	3.213 (0.08)	3.667 (0.12)
Taiwan	6307	1.365	507	8.04	0.043 (0.958)	0.511** (15.812)	1.610 (0.02)	2.479 (0.04)	2.872 (0.06)	2.155 (0.03)	3.063 (0.09)	3.473 (0.13)
Korea	8330	1.150	864	10.37	0.087* (2.23)	0.553** (19.424)	1.567 (0.05)	2.585 (0.09)	3.069 (0.11)	2.212 (0.07)	3.327 (0.11)	3.857 (0.13)
Hong Kong	7956	0.925	1088	13.68	0.105** (3.273)	0.606** (22.680)	1.568 (0.02)	2.747 (0.05)	3.320 (0.08)	2.319 (0.04)	3.637 (0.10)	4.276 (0.13)
Philippines	6120	1.175	587	9.59	−0.019 (−0.471)	0.621** (17.250)	1.577 (0.03)	2.559 (0.05)	2.971 (0.08)	2.185 (0.04)	3.146 (0.11)	3.549 (0.16)
Singapore	11,712	1.365	959	8.19	−0.059* (−1.880)	0.559** (22.259)	1.637 (0.01)	2.470 (0.03)	2.806 (0.04)	2.149 (0.02)	2.937 (0.04)	3.254 (0.06)
UK	12,648	1.150	1333	10.54	0.0141 (0.481)	0.634** (24.770)	1.625 (0.05)	2.669 (0.09)	3.125 (0.11)	2.275 (0.07)	3.333 (0.11)	3.796 (0.13)
Germany	12,546	0.985	1528	12.18	0.086** (3.142)	0.613** (26.772)	1.552 (0.02)	2.695 (0.04)	3.237 (0.07)	2.276 (0.03)	3.526 (0.10)	4.120 (0.14)
Turkey	9282	1.200	920	9.91	0.081** (2.551)	0.624** (21.825)	1.639 (0.03)	2.774 (0.06)	3.311 (0.09)	2.361 (0.05)	3.592 (0.11)	4.174 (0.14)
US	9516	1.215	953	953	0.047 (1.382)	0.617** (21.408)	1.647 (0.02)	2.711 (0.04)	3.195 (0.06)	2.315 (0.03)	3.432 (0.08)	3.939 (0.12)
Mexico	9240	1.125	985	10.66	0.015 (0.44)	0.648** (21.889)	1.618 (0.02)	2.685 (0.04)	3.152 (0.07)	2.283 (0.03)	3.365 (0.09)	3.839 (0.13)
Brazil	10,200	1.200	1008	10.01	0.043 (1.55)	0.611** (23.961)	1.622 (0.05)	2.670 (0.08)	3.144 (0.10)	2.280 (0.07)	3.374 (0.12)	3.869 (0.15)
Argentina	8520	1.250	782	9.18	0.053* (1.702)	0.556** (21.131)	1.593 (0.02)	2.557 (0.04)	2.998 (0.07)	2.199 (0.03)	3.217 (0.11)	3.683 (0.14)
S Africa	11,590	1.100	1395	12.04	0.103** (4.834)	0.545** (29.270)	1.601 (0.05)	2.644 (0.09)	3.149 (0.11)	2.265 (0.07)	3.427 (0.11)	3.990 (0.14)
Australia	8979	1.350	724	8.06	0.032 (0.890)	0.558** (19.243)	1.619 (0.02)	2.555 (0.04)	2.974 (0.06)	2.205 (0.03)	3.172 (0.08)	3.605 (0.12)

Note: The table reports in-sample ML estimates of the GPD for the ARMA–GARCH–EVT model. T = total no. of observations, u = threshold level, k = No. of exceedances, k/T = percentage of exceedances. The *t*-statistics are given in the parenthesis in columns 6 and 7, while asterisks demonstrate the level of statistical significance. The asterisks (*) and (**) denote significance at 5% and 1% levels, respectively. The standard errors of quantile estimates are calculated by bootstrapping of 1000 samples for each index and are listed in parenthesis of columns 8–13. It may be noted that ξ value is significant for 50% of the countries and insignificant for another 50% of the countries. So one should take the fact into consideration and interpret the result cautiously.

where I is an indicator function and

$$\begin{aligned} I &= 1 \ \textit{if}\ X_t > u \\ &= 0 \ \textit{otherwise} \end{aligned}.$$

The MEF is the sum of the excesses over the threshold u divided by the number of data points which exceed the threshold u. It is an estimate of the mean excess function which describes the expected overshoot of a threshold once an exceedance occurs.

The interpretation of the mean excess plot is given in Beirlant, Teugels, and Vynckier (1996), Embrechts et al. (1997) and Hogg and Klugman (1984). If the empirical MEF is a positively sloped straight line above a certain threshold u, it is an indication that the data follow the GPD with a positive shape parameter ξ. On the other hand, exponentially distributed data would show a horizontal MEF while short-tailed data would have a negatively sloped line. The MEF of negative returns in each country is estimated to choose thresholds. From the MEF, the thresholds can be selected on the criterion of linearity in MEF plots.[7] While choosing threshold level subjectively from MEF plot, we make sure that the number of exceedances does not fall beyond a range in which bias and MSE are minimized as explained in Appendix B.

Based on the MEF plots we have chosen the thresholds for each market which along with its related statistics are reported in Table 5. The value of threshold for each market ranges from 0.925 to 1.365 and the number of exeedences, k (the number of points above the threshold) in each country is found to vary from 507 to 1528 which is large enough to facilitate a good estimation. In each case, the resulting exceedances k total roughly 10% of the sample, which is consistent with percentages reported by McNeil and Frey (2000).

As mentioned in Sub-section 2.1, the shape (ξ) and scale (ψ) parameters have been estimated by using the log-likelihood function:

$$log\,L(\xi,\psi,y_1,\ldots,y_k) = \sum_{j=1}^{k} \log G_{\xi\psi}\left(y_j\right) = -k\log\psi - \left(1+\frac{1}{\xi}\right)\sum_{j=1}^{k}\log\left(1+\xi\frac{y_j}{\psi}\right) \quad (24)$$

where $y_j = Z_j - u$ and Zj defines the standardized residuals exceeding the identified threshold value of u. Maximizing the log likelihood of Eq. (24) s. t. $\psi > 0$, $1+\xi y_j/\psi > 0$, the most likely values of ξ and ψ for each market are obtained and also reported in Table 5. Recall that value of $\xi > 0$ reflects heavy-tailed distributions. In fourteen out of sixteen countries the ξ estimate is positive suggesting that the left tail of the distribution of standardized residuals is mostly characterized by heavy tailedness. The table further documents the EVT tail quantiles: $V\hat{a}R_q$ and ES_q for each country which are obtained from Eqs. (4) and (6) respectively, using the values of n, u, k, ξ and ψ of the respective country at the specified end tail of α%. For example, the tail quantiles $VaR_{0.95}$ and $ES_{0.95}$ for deseasonalized returns of India are calculated as

$$VaR_{0.95} = 0.985 + \frac{0.605}{0.079}\left[\left(\frac{1-0.95}{1282/9272}\right)^{-0.079} - 1\right] = 1.608 \quad (25)$$

and

$$ES_{0.95} = \frac{1.608}{1-0.079} + \frac{0.605-0.079*0.985}{1-0.079} = 2.318 \quad (26)$$

Though the $V\hat{a}R_q$ quantile at $\alpha = 5\%$ is less than that under a normal distribution, the fatness of the tail is readily apparent, especially as we move to more extreme quantiles.

The EVT suggests that the excess distribution above a suitable threshold of intraday returns should follow a GPD. To determine how the GPD fits the tails of the return distribution, we plot the empirical distribution of exceedances along with the cumulative distribution simulated by a GPD in Fig. 3 and compare the results visually for each country. The empirical excess distribution function follows closely the trace of a corresponding GPD, implying that the GPD models the exceedances very well for each market.

After specifying our model completely by estimating the parameters and subsequently verifying the fitness of the model, we can now calculate the robust VaR and ES estimates based on Eqs. (18) and (19) respectively, where we multiply the GARCH volatilities with respective quantiles and finally add the conditional means. We report below the 95 percentile intraday *VaR* and *ES* for India. As shown earlier, the $VaR_{0.95}$ and $ES_{0.95}$ quantiles for India are found to be 1.608 and 2.318, respectively. For a five minute horizon, the intraday *VaR* and *ES* specifications of India are:

$$VaR_{0.95}^{t+1} = \hat{\mu}_{t+1} + 1.608\sqrt{\hat{h}_{t+1}} \quad (27)$$

[7] The MEF of negative returns has been plotted but not shown in the paper for brevity. To get the mean excess function of negative returns (left tail) we have transformed the residual series z_t in to $-z_t$ then the results for the minimum can be directly deduced from those of maximum.

and

$$ES_{0.95}^{t+1} = \hat{\mu}_{t+1} + 2.318\sqrt{\hat{h}_{t+1}} \tag{28}$$

Because intraday seasonality has been taken into account, the intraday forecasts of conditional mean ($\hat{\mu}_{t+1}$) and variances ($\hat{h}_{t+1}$) are based on deseasonalized filtered returns. So to compute VaR_q^{t+1} and ES_q^{t+1} for the original returns, it requires to re-include the seasonal component to

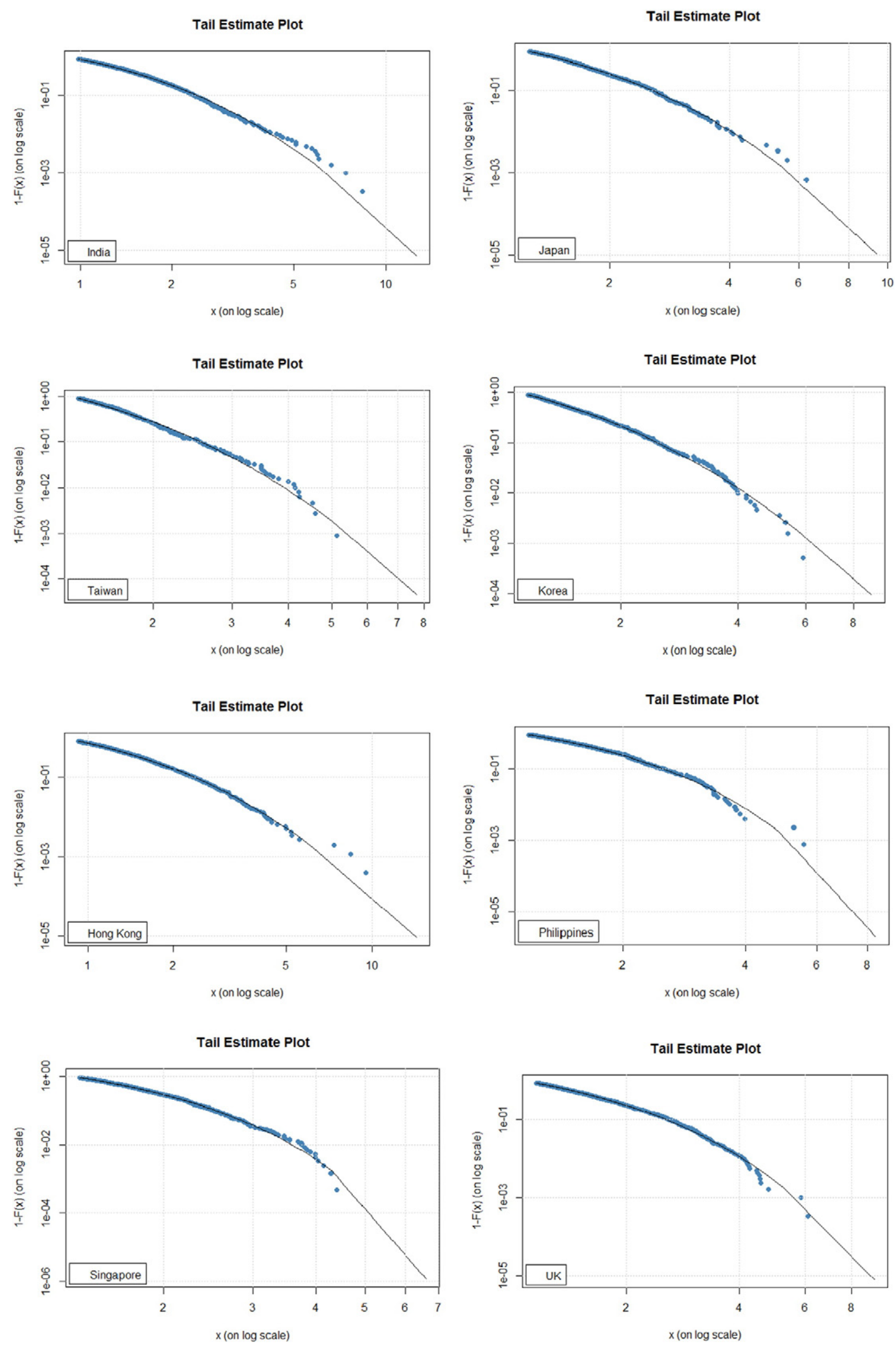

Fig. 3. Tail estimate plot for GPD fit.

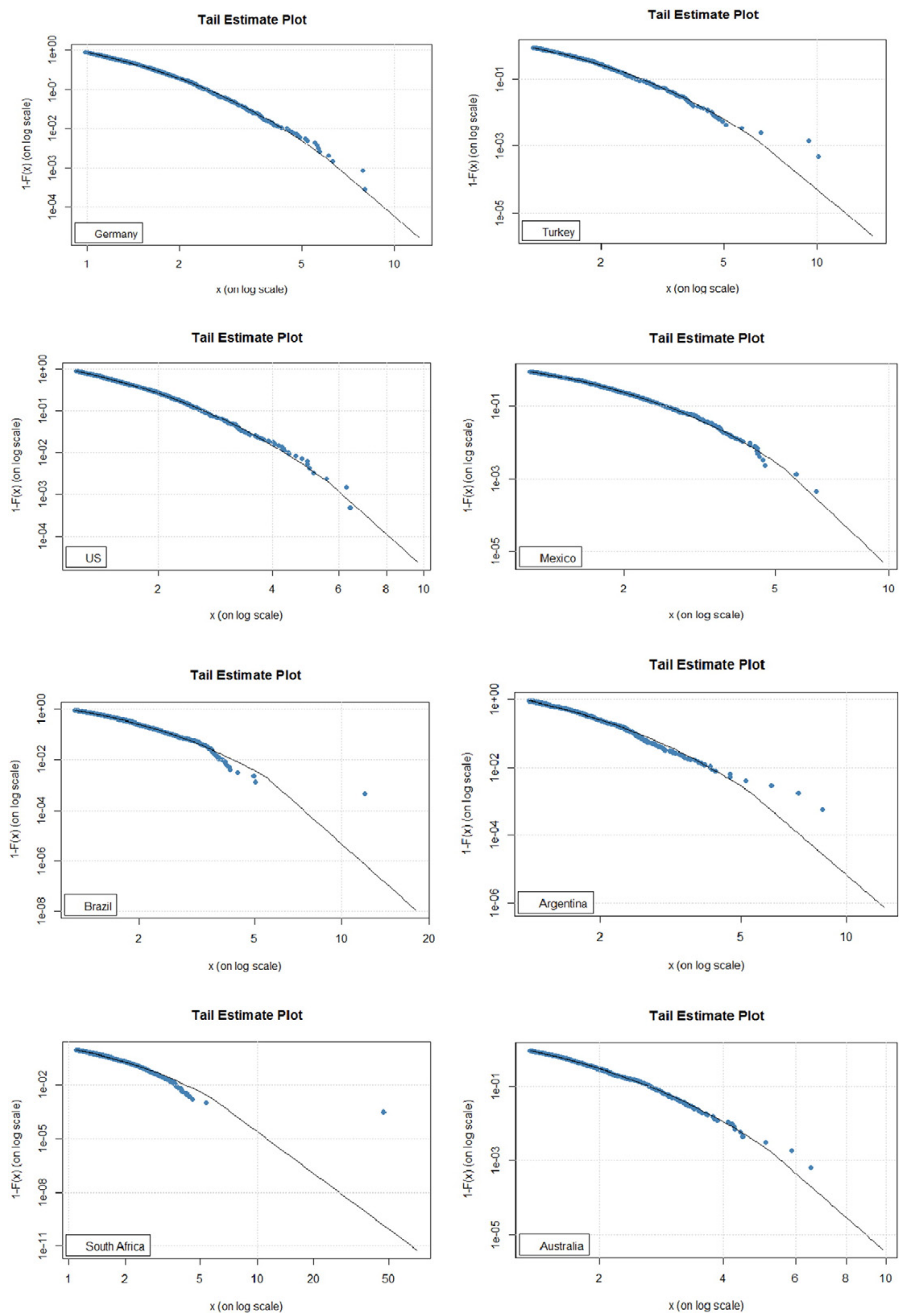

Fig. 3 (*continued*).

the intraday forecasts of condition mean and variance based on the deseasonalized filtered returns. To do so, we multiply $\hat{\mu}_{t+1}$ and $\hat{h}_{t+1}$ by the appropriate seasonal term $\hat{S}_{t,v}$ and its square $\hat{S}^2_{t,v}$ respectively, i.e.

$$\tilde{\mu}_{t+1} = (\hat{\mu}_{t+1})\left(\hat{S}_{t,v}\right) \tag{29}$$

$$\tilde{h}_{t+1} = \left(\hat{h}_{t+1}\right)\left(\hat{S}^2_{t,v}\right) \tag{30}$$

where, $\tilde{\mu}_{t+1}$ and $\tilde{h}_{t+1}$ are the transformed forecast of conditional mean and variance, respectively for the original returns, and $\hat{S}_{t,v}$ is estimated by the method described in the previous section.

Thus an estimate of the VaR and ES for the original returns are

$$VaR_q^{t+1} = \tilde{\mu}_{t+1} + \sqrt{\tilde{h}_{t+1}}\ V\hat{a}R_q \tag{31}$$

and

$$ES_q^{t+1} = \tilde{\mu}_{t+1} + \sqrt{\tilde{h}_{t+1}}ES_q \tag{32}$$

Using Eqs. (31) and (32), our intraday VaR and ES specifications for original returns of India are:

$$VaR_{0.95}^{t+1} = \tilde{\mu}_{t+1} + 1.608\ \sqrt{\tilde{h}_{t+1}} \tag{33}$$

and

$$ES_{0.95}^{t+1} = \tilde{\mu}_{t+1} + 2.318\sqrt{\tilde{h}_{t+1}} \tag{34}$$

4.2. Out-of-sample evidence

Till now, we have only looked in-sample, essentially fitting the Conditional EVT model to (extreme) data. In practice however, a risk manager is probably more interested in how well he or she can predict future extreme movements than in accurately modeling the past. To compare the accuracy of Conditional EVT for VaR and ES calculation with other alternatives, we have done backtesting of each method on out-of-sample return series using the following procedure.

In the beginning, the most recent n returns are used to estimate model parameters for each approach. The magnitude of n is set to be equal to the length of the in-sample period. That is, $n = 9272$ in Indian market, $n = 7259$ in Japanese market and so on as reported in Table 6. From the parameter estimates, the next interval VaR and ES are computed. Thus if we have the return series $r_1, r_2, \ldots\ldots\ldots\ldots, r_n, \ldots\ldots, r_m$, the conditional VaR_q^t and ES_q^t are computed on t intervals in the set $T = \{n + 1, \ldots\ldots, m\}$ using an n interval window each time.[8] In other words, keeping the size of the window n constant, the estimation procedure is rolled forward one interval and repeated to calculate the next interval VaR and ES. The main advantage of this rolling window technique is that it allows us to capture dynamic time-varying characteristics of the data in different time periods. As documented by McNeil and Frey (2000) and Gencay, Selcuk, and Ulugulyagci (2003), within the backtest period, it is difficult to choose the best parameterization of ARMA–GARCH model every time, so it is assumed that the model selected to the in-sample return series, as reported in Table 3 is adequate for each of the v_i (recall that v_i is defined as the number of observations per day of the *ith* country) rolling windows of first day of the out of sample period. A similar constraint is also related to the GPD modeling. In a long back test it is less feasible to examine the fitted model carefully every interval and to choose a threshold value (to determine a new value of k) for the tail estimator each time. For this reason the percentage of exceedances determined from the threshold value chosen for the first n sample, is set fixed for each of v_i rolling windows of the first day of the out of sample period. Thus for each of the v_i rolling windows of the first day of the out of sample period, we fit the same ARMA–GARCH model to generate a new set of standardized residuals and determine a new GPD quantile estimated on each set of the standardized residuals to calculate v_i one step ahead intraday VaR and ES. Thereafter, we re-estimate the whole model (ARMA–GARCH–EVT) and follow the same steps to forecast next v_i one step ahead intraday VaR and ES for the second day of the out of sample period. This process is repeated until the v_i interval returns of the last day.

Table 6
Relevant information of in-sample and out-of-sample periods.

	In-sample			Out–of-sample		
	Obs./day	No. of days	Total obs.	Obs./day	No. of days	Total obs.
India	76	122	9272	76	30	2280
Japan	61	119	7259	61	31	1891
Taiwan	53	119	6307	53	33	1749
Korea	70	119	8330	70	32	2240
Hong Kong	68	117	7956	68	31	2108
Philippines	51	120	6120	51	30	1530
Singapore	96	122	11,712	96	33	3168
UK	102	124	12,648	102	32	3264
Germany	102	123	12,546	102	32	3264
Turkey	78	119	9282	78	33	2574
US	78	122	9516	78	33	2574
Mexico	77	120	9240	77	31	2387
Brazil	85	120	10,200	85	31	2635
Argentina	71	120	8520	71	28	1988
S Africa	95	122	11,590	95	29	2755
Australia	73	123	8979	73	32	2336

Note: Three columns of both in-sample and out-of-sample periods, show number of sample observations in the return series per day, number of trading days and total number of sample observations finally after removing certain figures from the raw price set for each country. If number of days and total observations for both in-sample and out-of-sample periods are added, we get the number of days and total observations for the total period shown in Table 1.

Thus using the above procedure we estimate one step ahead intraday VaR and ES for each interval of the total out of sample period. Such procedure, as mentioned above, is called Conditional EVT. The Conditional EVT based intraday VaR and ES for the original returns are computed based on Eqs. (31) and (32) respectively, which have re-included the intraday periodicity component. In addition, we estimate the Unconditional EVT qth quantiles for VaR and ES using Eqs. (4) and (6) respectively, applied to the deseasonalized return series. To get the Unconditional EVT based intraday VaR and ES for original return series the Unconditional EVT qth quantiles calculated above have been multiplied by the seasonality component $\hat{S}_{t,v}$ defined earlier. The intraday VaR and ES for three other models (i.e., Static Normal, Conditional Normal and Risk Metrics) are also computed on rolling basis and the seasonality is adjusted in the same way.

4.2.1. Backtesting of VaR

Various methods and tests have been suggested for evaluating VaR model accuracy. Here, we first use Binomial test and then apply different Likelihood ratio tests for coverage probability.

4.2.1.1. Binomial test. The following step consists of comparing the quantile estimate in $t + 1$, $V\hat{a}R_q^{t+1}$ given by each method, with r_{t+1}, the log-negative return in $t + 1$, for $q \in \{0.95, 0.99, 0.995\}$. A violation is said to take place whenever $r_{t+1} < V\hat{a}R_q^{t+1}$. We can test whether the number of violations is statistically significant. In particular, let us consider the following statistic based on the binomial distribution:

$$Z = \frac{\frac{Y}{T} - p}{\sqrt{\frac{p(1-p)}{T}}} \quad \rightarrow \quad N(0,1) \tag{35}$$

where $T = (m - n)$ is the number of intervals of the out-of-sample and Y is the number of violations, so that Y/T is the actual proportion of violations in the T period. The proportion of p is the expected number of violations under the assumption that $Y = \sum_{t \in T} I_t \sim \mathbf{B}(T,p)$, where $I_t = 1$ {if $r_{t+1} < V\hat{a}R_q^{t+1}$}, I_t and I_s are independent for $t, s \in T$, $t \neq s$, and $\mathbf{B}(T,p)$ is binomial distribution.

[8] Here T is the total number of intervals of out-of-sample period i.e., $T = 2280$ for India, $T = 1891$ for Japan and so on as reported in Table 6. Thus for example, the return series for India is $r_1, r_2, \ldots\ldots\ldots\ldots, r_{9272}, \ldots\ldots, r_{11552(=9272+2280)}$ and for Japan is $r_1, r_2, \ldots\ldots\ldots\ldots, r_{7259}, \ldots\ldots, r_{9150(=7259+1891)}$.

Table 7
Backtesting of VaR (statistics of binomial test).

	S Norm	Cond Norm	RM	Unc. EVT	Cond. EVT
Panel A: $\alpha = 5\%$					
India	−3.075**	1.441	−0.577	−2.592**	−0.672
	(0.001)	(0.075)	(0.282)	(0.005)	(0.251)
Japan	−2.379**	0.892	0.259	−1.849*	0.470
	(0.0089)	(0.186)	(0.398)	(0.033)	(0.681)
Taiwan	−1.476	0.499	−1.256	−0.594	−0.598
	(0.070)	(0.309)	(0.105)	(0.275)	(0.275)
Korea	−4.169**	−0.679	0.097	−3.293**	−0.776
	(0.000)	(0.249)	(0.461)	(0.001)	(0.219)
Hong Kong	0.260	4.257**	0.360	1.161	0.660
	(0.398)	(0.000)	(0.360)	(0.121)	(0.255)
Philippines	−6.041**	−0.880	1.349	−5.568**	−1.114
	(0.000)	(0.190)	(0.089)	(0.000)	(0.133)
Singapore	−1.418	−0.521	−0.277	−1.416	−1.255
	(0.078)	(0.301)	(0.391)	(0.078)	(0.105)
UK	−2.104*	1.510	1.269	−0.738	−0.739
	(0.0177)	(0.066)	(0.102)	(0.231)	(0.230)
Germany	3.437**	5.525**	−0.498	5.445**	0.787
	(0.000)	(0.000)	(0.309)	(0.000)	(0.216)
Turkey	−5.218**	0.389	0.479	−4.676**	−1.149
	(0.000)	(0.349)	(0.316)	(0.000)	(0.125)
US	5.182**	3.735**	1.293	4.736**	0.027
	(0.000)	(0.000)	(0.098)	(0.000)	(0.489)
Mexico	−2.662**	1.000	−0.315	−2.097*	−1.535
	(0.004)	(0.159)	(0.377)	(0.018)	(0.062)
Brazil	0.469	0.112	−2.033*	1.048	−0.782
	(0.319)	(0.456)	(0.021)	(0.152)	(0.217)
Argentina	−4.055**	2.634**	1.914*	−3.433**	0.062
	(0.000)	(0.004)	(0.028)	(0.000)	(0.475)
S Africa	−5.748**	−1.115	−0.153	−4.741**	−1.464
	(0.000)	(0.133)	(0.439)	(0.000)	(0.072)
Australia	−1.880*	0.114	0.209	−1.878*	−0.646
	(0.030)	(0.455)	(0.417)	(0.030)	(0.259)
No. of rejections	12	4	2	11	0
Panel B: $\alpha = 1\%$					
India	1.726*	1.095	4.041**	−0.378	1.095
	(0.042)	(0.137)	(0.000)	(0.352)	(0.137)
Japan	1.870*	0.714	4.643**	−0.671	0.945
	(0.031)	(0.238)	(0.000)	(0.250)	(0.172)
Taiwan	2.285*	−0.118	1.564	1.327	0.843
	(0.011)	(0.453)	(0.059)	(0.0923)	(0.200)
Korea	0.552	0.552	6.286**	−2.204*	0.340
	(0.290)	(0.290)	(0.000)	(0.013)	(0.367)
Hong Kong	2.828**	−0.018	4.798**	−0.018	−0.674
	(0.002)	(0.493)	(0.000)	(0.493)	(0.250)
Philippines	−1.619	−1.876*	2.235*	−2.643**	−0.848
	(0.053)	(0.030)	(0.013)	(0.004)	(0.198)
Singapore	1.486	−0.836	1.486	0.059	−0.836
	(0.069)	(0.202)	(0.069)	(0.477)	(0.202)
UK	2.174*	2.400**	3.582**	−2.104*	−0.640
	(0.015)	(0.008)	(0.000)	(0.012)	(0.261)
Germany	9.915**	1.471	6.396**	−4.508**	−0.345
	(0.000)	(0.071)	(0.000)	(0.000)	(0.475)
Turkey	−0.543	1.731*	4.806**	−4.504**	−0.345
	(0.294)	(0.042)	(0.000)	(0.000)	(0.365)
US	10.154**	−1.137	4.806**	3.619**	−1.533
	(0.000)	(0.128)	(0.000)	(0.000)	(0.063)
Mexico	−0.385	−2.030*	1.056	−1.824*	−2.442**
	(0.350)	(0.021)	(0.146)	(0.034)	(0.007)
Brazil	4.826**	1.498	3.064**	0.324	1.106
	(0.000)	(0.067)	(0.001)	(0.374)	(0.134)
Argentina	−1.776*	−0.875	3.859**	−2.450**	−0.424
	(0.038)	(0.191)	(0.000)	(0.007)	(0.336)
S Africa	−1.446	−2.212*	2.384**	−3.358**	−2.212*
	(0.074)	(0.014)	(0.009)	(0.001)	(0.014)
Australia	−1.115	−1.323	2.836**	−3.194**	−1.530
	(0.133)	(0.093)	(0.002)	(0.000)	(0.063)
No. of rejections	9	5	13	10	2
Panel C: $\alpha = 0.5\%$					
India	3.444**	1.663*	7.007**	0.775	1.663*
	(0.000)	(0.048)	(0.000)	(0.221)	(0.048)
Japan	3.112**	−1.452	4.416**	−0.148	−1.126
	(0.001)	(0.073)	(0.000)	(0.441)	(0.130)

(continued on next page)

Table 7 (*continued*)

	S Norm	Cond Norm	RM	Unc. EVT	Cond. EVT
Taiwan	4.833**	−0.253	3.477**	2.463**	0.764
	(0.000)	(0.400)	(0.000)	(0.005)	(0.222)
Korea	0.539	−0.360	5.932**	−2.453**	−0.060
	(0.295)	(0.360)	(0.000)	(0.006)	(0.476)
Hong Kong	5.392**	−1.093	4.774**	−1.093	−1.093
	(0.000)	(0.137)	(0.000)	(0.137)	(0.137)
Philippines	−0.961	−2.048*	4.839**	−1.683*	−0.236
	(0.168)	(0.020)	(0.000)	(0.045)	(0.407)
Singapore	1.552	1.472	1.552	−0.714	−1.471
	(0.060)	(0.071)	(0.060)	(0.237)	(0.071)
UK	3.395**	−2.809**	4.636**	−1.816*	−1.072
	(0.000)	(0.003)	(0.000)	(0.034)	(0.142)
Germany	12.328**	0.665	7.862**	2.413**	1.161
	(0.000)	(0.253)	(0.000)	(0.008)	(0.123)
Turkey	−0.802	−1.920*	6.184**	−3.038**	−0.243
	(0.211)	(0.027)	(0.000)	(0.001)	(0.404)
US	12.891**	−1.920*	6.743**	2.831**	−1.920*
	(0.000)	(0.027)	(0.000)	(0.002)	(0.027)
Mexico	1.760*	−2.012*	1.760*	−2.010*	−1.722*
	(0.039)	(0.022)	(0.039)	(0.022)	(0.043)
Brazil	6.304**	−0.048	4.923**	0.505	−0.048
	(0.000)	(0.481)	(0.000)	(0.307)	(0.481)
Argentina	0.019	−2.207*	4.789**	−2.205*	−1.253
	(0.492)	(0.014)	(0.000)	(0.014)	(0.105)
S Africa	−0.209	−1.290	3.032**	−2.908**	−1.290
	(0.417)	(0.099)	(0.001)	(0.002)	(0.099)
Australia	−0.493	−1.079	3.907**	−1.662*	−1.079
	(0.311)	(0.140)	(0.000)	(0.049)	(0.140)
No. of rejections	9	7	15	11	3

Note: The table presents binomial test statistics VaR violation ratio under each competing approach. *p* values are given in the parentheses. The asterisks (*) and (**) denote significance at 5% and 1% levels, respectively.

Expression (35) is a one-tailed test that is asymptotically distributed as standard normal (see, e.g., Larsen & Marx, 1986, chapter 5). If $Y/T < p$, we test the null hypothesis of estimating correctly the conditional quantile against the alternative that the method systematically underestimates it. Otherwise, we test the null hypothesis against the alternative that the method systematically overestimates the conditional quantile.

Table 7 presents our backtesting results of binomial test for each individual market. The table shows the binomial test statistics with *p* values for 95%, 99%, and 99.5% quantiles, respectively. As a decision rule, we take a *p* value less than 5% to be evidence against the null hypothesis. It appears from the table that out of 48 cases (16 markets × 3 quantiles) analyzed, the null hypothesis is rejected 30 times under the Static Normal, 16 times under the Conditional Normal, 30 times under the Risk Metrics, 32 times under the Unconditional EVT and only 5 times under the Conditional EVT. The null hypothesis is accepted for the rest of the cases of different models where there is no significant difference between empirical and theoretical violations. In terms of number of rejection of null hypothesis, the Conditional EVT performs the best and the Unconditional EVT performs the least.

4.2.1.2. Likelihood ratio tests. To assess the forecasting performance of the VaR methods more precisely, we have also adopted the likelihood ratio tests for unconditional coverage, independence and conditional coverage, which are explained below.

At each time $t+1$, we introduce the violation indicator variable I_{t+1}, which compares the $V\hat{a}R_q^{t+1}$ with the r_{t+1}:

$$I_{t+1} = \begin{cases} 1, \mathit{if}\ r_{t+1} < V\hat{a}R_q^{t+1} \\ 0, \mathit{otherwise} \end{cases} \tag{36}$$

The test of unconditional coverage consists of examining if the realized coverage $p = \sum_{t=0}^{T} I_{t+1}/T$ equals to the theoretical coverage p. This is equivalent to testing if the indicator variable I_{t+1} follows an *iid* Bernoulli process with parameters p; where p equals to VaR's theoretical coverage rate α. The likelihood ratio (LR) test statistic for the unconditional coverage test follows a χ^2 distribution with one degree of freedom. That is,

$$LR_{uc} = 2\log\left[\frac{(1-p)^{T_0} p^{T_1}}{(1-T_1/T)^{T_0}(T_1/T)^{T_1}}\right] \sim \chi^2(1) \tag{37}$$

where T_0 and T_1 are the number of zeros and ones, respectively in the violation sequence.

The drawback of this unconditional coverage testing method is that it fails to properly characterize the behavior of the model in the presence of clustering. It may happen that, although the number of violation is correct, they might occur in clusters. To take into account that possibility, we also perform the test of independence and the test of conditional coverage suggested in Christoffersen (1998). The first tests for independence and the second tests for both independence and correct coverage leading to a complete test of correct conditional coverage, without making any hypothesis about the underlying true conditional distribution. Christoffersen's (1998) *LR* test for independence, against an explicit first-order Markov alternative is given by:

$$LR_{ind} = -2\log\left[\frac{(1-T_1/T)^{T_0}(T_1/T)^{T_1}}{(1-\pi_{01})^{T_{00}}\pi_{01}^{T_{01}}}\right] \sim \chi^2(1) \tag{38}$$

where, $T_{ij}, i, j = 0, 1$ is the number of observations with a j following an i in I_t sequence, and $\pi_{01} = T_{01}/(T_{00} + T_{01})$. Here too the LR test for independence follows a χ^2 distribution of one degree of freedom.

Christoffersen's (1998) conditional coverage test is thus a joint test for independence and correct unconditional coverage, which involves estimation of the following *LR* statistic:

$$LR_{cc} = LR_{uc} + LR_{ind} \sim \chi^2(2) \tag{39}$$

This statistic follows a χ^2 distribution with two degree of freedom.

Table 8
Backtesting of VaR (Statistical tests of unconditional coverage).

	S norm	Cond norm	RM	Unc. EVT	Cond. EVT
$\alpha = 5\%$					
India	10.436**	1.998	0.338	7.279**	0.462
Japan	6.147*	0.773	0.066	3.632	0.217
Taiwan	2.292	0.245	1.648	0.358	0.365
Korea	20.017**	0.470	0.009	12.069**	0.616
Hong Kong	0.067	16.190*	0.128	1.311	0.427
Philippines	48.883**	0.799	1.739	40.212**	1.294
Singapore	2.086	0.276	0.077	2.074	1.627
UK	4.672*	2.201	1.563	0.550	0.556
Germany	10.976**	27.200**	0.251	26.456**	0.608
Turkey	32.289**	0.150	0.227	25.379**	1.362
US	23.747**	12.728**	1.617	20.017**	0.001
Mexico	7.694**	0.973	0.100	4.674*	2.467
Brazil	0.217	0.012	4.382*	0.828	0.625
Argentina	19.039**	6.439*	3.465	13.299**	0.004
S Africa	39.715**	1.280	0.023	29.775**	2.232
Australia	3.740	0.013	0.043	3.722	0.425
$\alpha = 1\%$					
India	2.678	1.117	13.080**	0.146	1.118
Japan	3.087	0.485	16.498**	0.474	0.838
Taiwan	4.480*	0.014	2.193	1.601	0.669
Korea	0.294	0.294	28.784**	5.860*	0.113
Hong Kong	11.081**	0.000	17.698**	0.000	0.478
Philippines	3.075	4.261*	4.260*	9.472**	0.776
Singapore	2.038	0.735	2.038	0.014	0.735
UK	4.228*	6.776**	10.792**	4.792*	0.426
Germany	66.820**	2.000	30.993**	10.805**	0.004
Turkey	0.306	3.406	18.109**	32.768**	0.122
US	67.265**	1.400	18.109**	10.806**	2.627
Mexico	0.152	4.841*	1.042	3.829	7.294**
Brazil	18.290**	2.055	7.955**	0.104	1.145
Argentina	3.676	0.820	11.878**	7.544**	0.185
S Africa	2.310	5.760*	4.987*	14.932**	5.760*
Australia	1.349	1.932	6.833**	13.676**	2.633
$\alpha = 0.5\%$					
India	9.146**	2.400	31.568**	0.558	2.400
Japan	7.479**	2.549	13.899**	0.022	1.459
Taiwan	16.090**	0.066	9.037**	4.860*	0.540
Korea	0.276	0.134	23.697**	8.519**	0.004
Hong Kong	21.946**	1.356	16.146**	1.356	1.356
Philippines	1.052	5.190*	15.830**	3.691	0.057
Singapore	2.146	2.492	2.146	0.543	2.492
UK	9.226**	10.850**	16.155**	3.939*	1.266
Germany	85.843**	0.420	40.516**	5.868*	1.236
Turkey	0.697	4.601*	25.963**	14.331**	0.060
US	88.247**	4.595*	30.113**	6.488*	4.600*
Mexico	2.678	5.184*	2.678	5.184*	3.632
Brazil	26.924**	0.002	17.523**	0.245	0.002
Argentina	0.000	6.717**	16.110**	6.710**	1.830
S Africa	0.045	1.897	7.397**	12.439**	1.897
Australia	0.255	1.311	11.486**	3.376	1.311
No. of violations	25	14	28	27	3

Note: The table presents statistical tests of unconditional coverage (uc) of the intraday VaR forecasts under each competing approach. The test is asymptotically distributed as χ^2 with d.f. one. The asterisks (*) and (**) denote significance at 5% and 1% levels, respectively.

Table 9
Backtesting of VaR (statistical tests of independence).

	S Norm	Cond Norm	RM	Unc. EVT	Cond. EVT
$\alpha = 5\%$					
India	5.604*	0.000	3.823	6.207*	0.246
Japan	0.000	0.517	5.975*	0.038	0.814
Taiwan	15.207**	16.204**	17.583**	12.973**	19.731**
Korea	6.597*	0.227	1.707	8.573**	0.091
Hong Kong	3.621	1.336	4.135*	5.823*	1.920
Philippines	9.085**	0.228	9.981**	8.818**	1.136
Singapore	0.128	0.078	2.125	0.128	0.073
UK	26.535**	8.714**	13.305**	24.127**	8.020**
Germany	7.325**	6.161*	8.734**	9.692**	6.042*
Turkey	20.386**	4.203*	21.248**	17.779**	2.066
US	14.218**	0.908	5.974*	13.319**	2.433
Mexico	0.503	1.469	0.113	1.153	5.664*
Brazil	10.571**	3.178	6.441*	12.433**	2.993
Argentina	0.176	0.146	1.059	1.773	0.282
S Africa	2.211	1.622	2.388	2.527	1.886
Australia	11.434**	2.474	10.149**	11.433**	3.318
$\alpha = 1\%$					
India	5.533*	0.342	2.984	4.532	0.342
Japan	0.194	0.000	1.226	0.325	0.734
Taiwan	33.327**	14.957**	16.413**	28.760**	15.266**
Korea	2.646	2.301	6.106*	0.522	2.118
Hong Kong	2.061	0.035	0.123	0.040	1.259
Philippines	1.759	0.340	0.233	4.134*	0.770
Singapore	1.825	0.012	0.000	0.062	0.012
UK	12.268**	0.551	3.483	0.449	0.845
Germany	14.651**	1.677	6.531*	23.259**	0.097
Turkey	11.548**	3.152	4.420*	7.017**	1.119
US	12.762**	2.267	4.578*	20.303**	0.506
Mexico	1.494	0.668	2.913	0.767	0.490
Brazil	0.397	0.061	0.155	2.449	0.173
Argentina	1.109	1.054	0.044	0.330	0.188
S Africa	2.430	0.755	10.781**	0.293	0.755
Australia	2.430	1.010	5.083*	2.768	0.894
$\alpha = 0.5\%$					
India	3.602	2.965	1.738	4.382	2.965
Japan	0.069	0.107	0.688	1.967	0.154
Taiwan	28.760**	2.427	23.978**	22.227**	1.289
Korea	0.614	0.362	2.849	0.032	0.438
Hong Kong	3.235	0.188	0.427	0.188	0.188
Philippines	4.134*	0.021	0.013	6.583*	0.260
Singapore	0.244	0.255	0.244	0.432	0.255
UK	8.135**	0.062	3.230	0.200	0.357
Germany	21.921**	0.899	7.489**	7.028**	1.100
Turkey	2.125	4.072*	6.690**	8.907**	1.505
US	16.408**	4.334*	3.589	7.267**	4.334*
Mexico	0.362	0.084	1.109	0.084	0.122
Brazil	4.081*	0.520	0.150	0.694	0.520
Argentina	0.408	0.036	2.603	0.036	0.146
S Africa	0.497	0.237	0.016	0.026	0.237
Australia	1.959	0.221	8.738**	3.887*	0.221
No. of violations	22	7	21	22	6

Note: The table presents statistical test of independence (ind) of the intraday VaR forecasts under each competing approach. Following Chen et al. (2012), we have extended the test up to four lags but reported the results for 4th lag. The test is asymptotically distributed as χ^2 with d.f. one. The asterisks (*) and (**) denote significance at 5% and 1% levels, respectively.

Although in the original setting of Christoffersen's (1998) LR test, only the first lag is considered, however, as for high frequency data used in the present study, a few further lags are worth to be tested. Hence, in line with study made by Chen, Gerlach, and Lu (2012) we extended the test up to four lags.

Tables 8, 9, and 10 show statistics of unconditional, independence and conditional coverage tests, respectively for the different models at $p = 95\%$, $p = 99\%$ and $p = 99.5\%$. As a quick reference guide, the absence of 'asterisks' in the tables indicates that the difference between theoretical and empirical violation ratios is not statistically significant. For the unconditional coverage test reported in Table 8 out of 48 cases (16 markets $\times$ 3 quantiles) analyzed, the Standard Normal fails 25 times, the Conditional Normal fails 14 times, the RiskMetrics fails 28 times, the Unconditional EVT fails 27 times and the Conditional EVT fails only 3 times. Thus under this test the Conditional EVT model dominates other models in VaR forecasting. For the independence test reported in Table 9, the Conditional EVT model performs the best and is closely followed by the Conditional Normal model. The performance of the remaining three models is far behind than the former two models. For the conditional coverage test reported in Table 10 the Conditional EVT again performs the best. Out of 48 cases, the model fails 8 times while the Risk Metrics model fails as high as 36 times and becomes the least performing model.

So far we have done the backtesting analysis of the different models using various tests. The Conditional EVT model consistently performs the best in estimating and forecasting VaR. However, less consistency in relative performance is evident for other models.

Table 10
Backtesting of VaR (statistical tests of conditional coverage).

	S norm	Cond norm	RM	Unc. EVT	Cond. EVT
$\alpha = 5\%$					
India	16.039**	1.997	4.161	13.486**	0.707
Japan	6.147*	1.290	6.041*	3.670	1.031
Taiwan	17.499**	16.450**	19.231**	13.331**	20.096**
Korea	26.614**	0.697	1.717	20.642**	0.706
Hong Kong	3.688	17.526**	4.263	7.134*	2.347
Philippines	57.968**	1.028	11.721**	49.029**	2.430
Singapore	2.214	0.354	2.203	2.202	1.170
UK	31.207**	10.915**	14.868**	24.677**	8.577*
Germany	18.301**	33.361**	8.985*	36.148**	6.650*
Turkey	52.675**	4.353	21.475**	43.158**	3.428
US	37.965**	13.637**	7.591*	33.336**	2.433
Mexico	8.197*	2.443	0.213	5.826	8.132*
Brazil	10.789**	3.190	10.824**	13.261**	3.618
Argentina	19.215**	6.585*	4.525	15.072**	0.285
S Africa	41.926**	2.903	2.411	32.302**	4.117
Australia	15.174**	2.487	10.192**	15.156**	3.742
$\alpha = 1\%$					
India	8.210**	1.456	16.064**	4.679	1.459
Japan	3.281	0.485	17.725**	0.799	1.570
Taiwan	37.807**	14.971**	18.606**	30.361**	15.935**
Korea	2.940	2.595	34.890**	6.382*	2.231
Hong Kong	13.142**	0.035	17.821**	0.040	1.737
Philippines	4.834	4.601	4.493	13.606**	1.547
Singapore	3.863*	0.747	2.038	0.077	0.747
UK	16.497**	7.327*	14.276**	5.241	1.271
Germany	81.471**	3.677	37.524**	34.064**	0.101
Turkey	11.853**	6.558*	22.530**	39.823**	1.240
US	80.027**	3.668	22.688**	31.109**	3.133
Mexico	1.646	5.509	3.955	4.597	7.784*
Brazil	18.688**	2.117	8.110**	2.554	1.318
Argentina	4.785	1.873	11.923**	7.875*	0.373
S Africa	4.740	6.515*	15.768**	15.226**	6.515*
Australia	3.779	2.942	11.916**	16.445**	3.527
$\alpha = 0.5\%$					
India	12.748**	5.366	33.306**	4.940	5.366
Japan	7.548*	2.656	14.587**	1.990	1.613
Taiwan	44.850**	2.493	33.015**	27.087**	1.829
Korea	0.890	0.496	26.546**	8.552**	0.442
Hong Kong	25.181**	1.544	16.573**	1.544	1.544
Philippines	5.186	5.976	15.843**	10.274**	0.317
Singapore	2.390	2.747	2.390	0.975	2.747
UK	17.361**	10.911**	19.385**	4.140	1.622
Germany	107.764**	1.319	48.005**	12.896**	2.336
Turkey	2.822	8.673*	32.653**	23.239**	1.565
US	104.656**	8.935*	33.702**	13.756**	8.935*
Mexico	3.040	5.274	3.787	5.268	3.754
Brazil	31.005**	0.523	17.673**	0.939	0.523
Argentina	0.409	6.753*	18.713**	6.746*	1.976
S Africa	0.542	2.134	7.413*	12.465**	2.134
Australia	2.214	1.532	20.225**	7.264*	1.532
No. of violations	31	14	36	31	8

Note: The table presents statistical test of conditional coverage (cc) of the intraday VaR forecasts under each competing approach. The test is asymptotically distributed as χ^2 with d.f. two. The asterisks (*) and (**) denote significance at 5% and 1% levels, respectively.

4.2.2. Backtesting of ES

To backtest the estimated ES_q value, we use the measure proposed by Embrechts et al. (2005). The standard backtesting measure for the ES_q estimates is

$$E_1 = \frac{1}{c} \sum_{t \in \kappa} \phi_t \tag{40}$$

where, $\phi_t = r_{t+1} - ES_q^{t+1}$, c is the number of intervals for which a violation of VaR, i.e., $r_{t+1} < V\hat{a}R_q^{t+1}$ occurs and κ is the set of intervals for which it happens.

Weakness of this measure is that it depends strongly on the VaR estimates without adequately reflecting the correctness of these values. To correct for this, it is combined with the following measure, where the empirical quantile of ϕ is used in place of the VaR estimates.

$$E_2 = \frac{1}{d} \sum_{t \in \eta} \phi_t \tag{41}$$

where, d is the number of periods for which ϕ_t is less than the empirical quantile and η is the set of periods for which it happens.

Thus, the Embrechts et al. (2005) measure is given by

$$E = (|E_1| + |E_2|)/2 \tag{42}$$

Table 11
Backtesting of ES (Embrechts et al. measure).

	S norm	Cond norm	RM	Unc. EVT	Cond. EVT
$\alpha = 5\%$					
India	0.210	0.044*	0.273	0.069	0.093
Japan	0.258	0.051	0.255	0.046	0.036*
Taiwan	0.180	0.018*	0.159	0.100	0.052
Korea	0.067	0.020	0.164	0.126	0.016*
Hong Kong	0.172	0.115	0.180	0.059	0.055*
Philippines	0.198	0.088	0.122	0.158	0.006*
Singapore	0.025	0.020	0.017	0.009*	0.012
UK	0.060	0.023*	0.088	0.088	0.033
Germany	0.406	0.122	0.288	0.142	0.071*
Turkey	0.191	0.089	0.320	0.537	0.056*
US	0.238	0.060	0.134	0.079	0.048*
Mexico	0.046*	0.140	0.087	0.166	0.099
Brazil	0.318	0.033	0.219	0.053	0.014*
Argentina	0.128	0.165	0.218	0.366	0.038*
S Africa	0.073	0.079	0.062*	0.203	0.087
Australia	0.016*	0.026	0.094	0.120	0.020
$\alpha = 1\%$					
India	0.669	0.317	0.716	0.144*	0.266
Japan	0.643	0.295	0.405	0.129	0.104*
Taiwan	0.459	0.121	0.466	0.212	0.075*
Korea	0.097	0.077	0.250	0.205	0.055*
Hong Kong	0.312	0.477	0.243	0.190*	0.197
Philippines	0.299	0.341	0.384	0.211	0.060*
Singapore	0.015*	0.090	0.024	0.060	0.054
UK	0.170	0.172	0.212	0.163	0.032*
Germany	0.797	0.305	0.544	0.197	0.140*
Turkey	0.130	0.107	0.709	0.688	0.081*
US	0.367	0.276	0.208	0.079*	0.151
Mexico	0.359	0.304	0.282	0.094*	0.121
Brazil	0.596	0.289	0.468	0.061*	0.194
Argentina	0.258	0.658	0.383	0.678	0.176*
S Africa	0.051*	0.154	0.148	0.257	0.141
Australia	0.076	0.133	0.173	0.134	0.045*
$\alpha = 0.5\%$					
India	0.932	0.421	0.906	0.160*	0.335
Japan	0.917	0.135	0.583	0.196	0.107*
Taiwan	0.537	0.185	0.501	0.191	0.079*
Korea	0.187	0.086	0.313	0.216	0.051*
Hong Kong	0.300	0.656	0.281	0.128*	0.274
Philippines	0.510	1.195	0.461	0.217	0.044*
Singapore	0.020	0.118	0.019*	0.081	0.055
UK	0.228	0.267	0.278	0.233	0.028*
Germany	1.000	0.481	0.756	0.247	0.206*
Turkey	0.339	0.315	0.913	0.799	0.071*
US	0.387	0.276	0.217	0.094	0.081*
Mexico	0.463	0.356	0.419	0.231	0.133*
Brazil	0.702	0.493	0.474	0.090*	0.321
Argentina	0.320	0.865	0475	0.928	0.280*
S Africa	0.085*	0.219	0.210	0.254	0.181
Australia	0.161	0.163	0.198	0.189	0.085*
Min value occurences	5	3	2	9	29

Note: The table presents Embrechts et al. measure (scaled up by $\times 10^3$) of the intraday ES forecasts under each competing approach. The presence of (*) represents the minimum value of the measure among the approaches for each stock index under a given confidence level.

A good estimation of ES will lead to a low value of E. Readers are referred to Embrechts et al. (2005) for further technical details on the test statistics.

Table 11 reports backtesting results of ES for each individual market. It shows the value of the measure (E) for 95%, 99% and 99.5% quantiles. The minimum value of the same is marked with 'asterisks' for each of the cases. It appears from the table that out of 48 cases (16 markets × 3 quantiles) analyzed, minimum value of E has been achieved 5 times by Static Normal, 3 times by Conditional Normal, 2 times by Risk Metrics, 9 times by Unconditional EVT and as high as 29 times by Conditional EVT. Thus in terms of number of minimum value of the measure, Conditional EVT once again outperforms other competing models.

It is evident from the backtesting analysis that the Conditional EVT model performs the best in forecasting both intraday VaR and ES. The superiority of the Conditional EVT model should come as no surprise. The ARMA (p_1, q_1) mean equation accommodates the autoregression in returns. The GARCH (p_2, q_2) component captures conditional volatility clustering. The EVT component explicitly models the heavy tails of the standardized residuals. Taken together, the features ensure that quantile estimates from the ARMA–GARCH–EVT model alternatively called Conditional EVT model at any given time reflect the most recent and relevant information.

5. Conclusion

The purpose of this paper has been to make a comparative study of predictive ability of various models in estimating intraday VaR and ES. The main emphasis has been given to the Extreme Value methodology and to evaluate how well the Conditional EVT model performs in modeling the tails of distributions and in estimating and forecasting intraday VaR and ES measures. In order to investigate the same, the 5 min price series of sixteen stock markets across Asia, Europe, the United States, Latin America, Africa and Australia have been considered. The preliminary analysis of the data shows the 5 min returns series used in this study have properties that are consistent with the stylized facts of high frequency financial returns reported in the literature. They are all fat-tailed, slightly skewed and have a zero mean. Furthermore, there are linear dependence in returns and the series exhibit high volatility, and volatility clustering. Most importantly, the series displays strong periodicity patterns in intraday volatility. The findings suggest the exploration of the ARMA–GARCH–EVT to forecast intraday VaR and ES. Since the GARCH-type models can be corrupted by intraday periodic patterns, we use the deseasonalized filtered returns to estimate the volatility model instead of considering raw returns. However the VaR and ES are later computed for the original returns by re-including the intraday periodicity component. To compare the accuracy of the Conditional EVT with other alternatives, we have done backtesting analysis on out-of-sample return series. The best performing model in estimating VaR is found to be the Conditional EVT and interestingly the same model performs the best in forecasting ES too. The Conditional EVT model that captures the time series properties of both mean and volatility of returns, as well as explicitly modeling the tails of the distribution, may offer advantages during the period of market turmoil. Thus risk managers may benefit from adopting the sophisticated ARMA–GARCH–EVT, alternatively called the Conditional EVT model. The study is useful for market participants (such as intraday traders and market makers) involved in frequent intraday trading in such equity markets.

Appendix A. Choice of starting values for conditional mean and variance estimation

As mentioned in note 4, we have chosen the starting values as the unconditional estimates of mean and variances of the in-sample return series. It can be empirically shown that the estimates of conditional means and variances are robust enough with respect to the choice of any starting values. This has been illustrated using different starting values to estimate mean and variance for India. We consider four different initial values for mean and variance as stated below:

i) Unconditional mean and variance of entire in-sample return series
ii) Zero mean and zero variance
iii) Unconditional mean and variance of first 10 observations
iv) Unconditional mean and variance of first 100 observations.

We have plotted the conditional mean and variance estimates in Fig. 4.1 and 4.2 respectively, with different initial values. It appears from both the figures that as we move forward from the starting period, the deviation between the estimates for alternative starting values diminishes exponentially and after a certain period the estimates converge to a single line, hence the importance of the initial values chosen declines to zero. We have performed the same exercise for other return series as well leading to same conclusion, but not shown here for brevity.

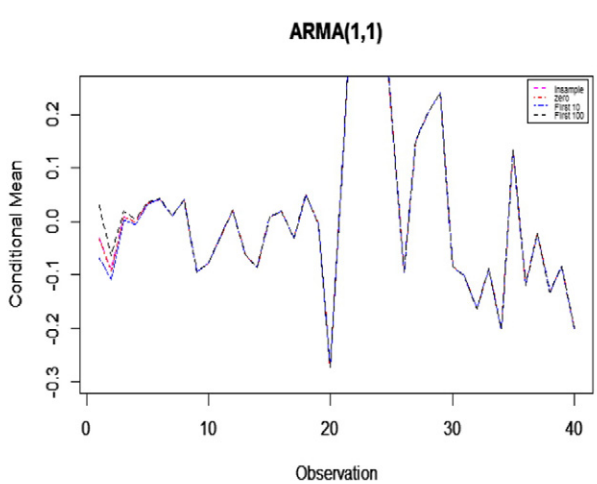

Figure 4.1

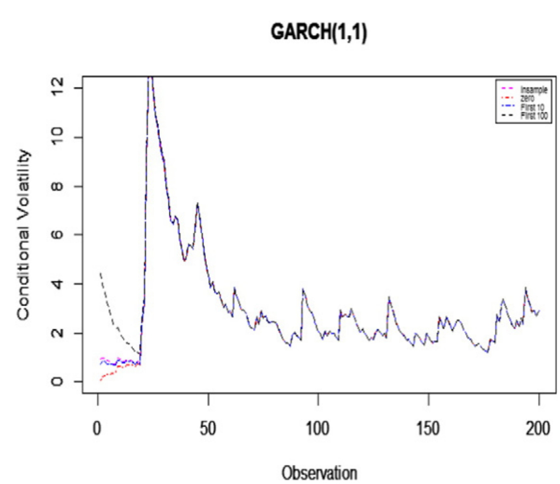

Figure 4.2

Appendix B. Threshold selection for the GPD approach

To select the threshold (i.e. choice of k) optimally, we perform a small simulation study following McNeil and Frey (2000). Since, the observed distribution of model residuals in every return series exhibits fat-tails and excess kurtosis, we generate random samples from Student-t distribution where sample size corresponds to the window length we use for in-sample estimation. The degrees of freedom are calculated from the moments of model residuals. Now, we estimate the quantiles from the series with various values of k using GPD. We restrict our attention to values of k such that $k >$ window length*$(1\text{-}q)$, so that the target quantile is beyond the threshold. For each return series, we estimate bias and the mean squared errors (MSE) using Monte Carlo estimates based on 1000 independent samples. For example, we estimate bias and MSE $(\hat{z}_{q,k})$ by

$$Bias = \frac{1}{1000}\sum_{j=1}^{1000} \hat{z}_{q,k}^{(j)} - z_q$$

$$MSE(\hat{z}_{q,k}) = \frac{1}{1000}\sum_{j=1}^{1000}\left(\hat{z}_{q,k}{}^{(j)} - z_q\right)^2$$

where, $\hat{z}_{q,k}{}^{(j)}$ represents the quantile estimate obtained from the j-th sample and z_q is the theoretical quantile estimates for Student-t distribution.

We have calculated the bias and MSE of GPD estimator of the 0.95 quantile against k for all the time series but for brevity the results are reported only for India. The results for the t-distribution with 5 degrees of freedom (calculated based on residual series of India) are depicted in Fig. 5.1 and 5.2 which correspond to bias-exceedances and MSE-exceedances, respectively.

From the figures it is quite evident that minimum of bias and MSE could be achieved when number of exceedances (k) varies from 750 to 1300 (roughly, 8% to 14). Therefore, while choosing threshold level subjectively from MEF plot, we make sure that the number of exceedances does not fall beyond this range.

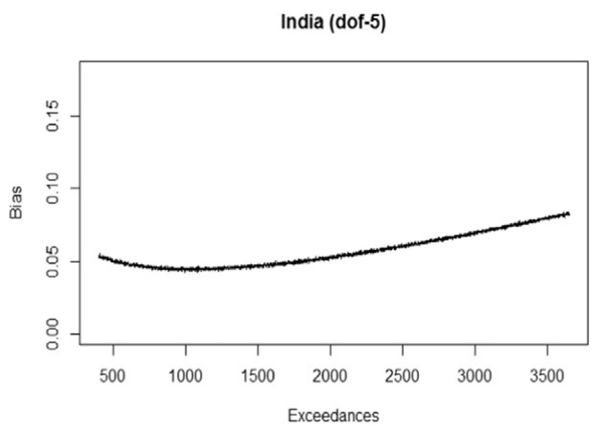

Figure 5.1

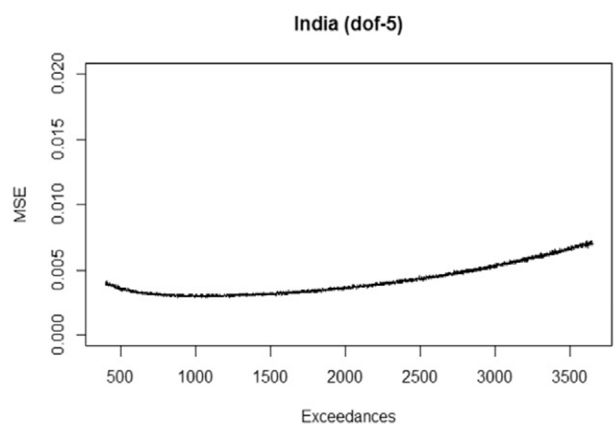

Figure 5.2

References

Alizadeh, S., Brandt, M., & Diebold, F. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance*, *57*(3), 1047–1091.

Andersen, T. G., & Bollerslev, T. (1997). Intraday periodicity and volatility persistence in financial markets. *Journal of Empirical Finance*, *4*(2), 115–158.

Andersen, T. G., & Bollerslev, T. (1998). Answering the skeptics: Yes, standard volatility models do provide accurate forecasts. *International Economic Review*, *39*(4), 885–905.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (1999). (Understanding, optimizing, using and forecasting) realizes volatility and correction. Manuscript, Northwestern University, Duke University and University of Pennsylvania. Published in revised form as "Great realizations" *Risk*, 105–108 (March 2000).

Aradhyula, S. V., & Ergun, A. T. (2004). Trading collar, intraday periodicity and stock market volatility. *Applied Financial Economics*, *14*(13), 909–913.

Bali, T. G., & Neftci, S. N. (2003). Disturbing extremal behavior of spot price dynamics. *Journal of Empirical Finance*, *10*(4), 455–477.

Bandi, F. M., & Rusell, J. R. (2005). Separating microstructure noise from volatility. *Journal of Financial Economics*, *79*(3), 655–692.

Beck, A., Kim, Y. S. A., Rachev, S., Feindt, M., & Fabozzi, F. (2013). Empirical analysis of ARMA–GARCH models in market risk estimation on high-frequency US data. *Studies in Nonlinear Dynamics & Econometrics*, *17*(2), 167–177.

Beirlant, J., Teugels, J., & Vynckier, P. (1996). *Practical analysis of extreme values*. Leuven: Leuven University Press.

Beirlant, J., Goegebeur, Y., Teugels, J., Seger, J., De Waal, D., & Ferro, C. (2004). *Statistics of extremes: theory and applications*. Chichester: John Wiley & Sons.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics*, *31*(3), 307–327.

Bollerslev, T., & Ghysels, E. (1996). Periodic autoregressive conditional heteroscedasticity. *Journal of Business and Economic Statistics*, *14*(2), 139–151.

Bystrom, H. N. (2004). Managing extreme risks in tranquil and volatile markets using conditional extreme value theory. *International Review of Financial Analysis*, *13*(2), 133–152.

Bystrom, H. N. (2005). Extreme value theory and extremely large electricity price changes. *International Review of Economics and Finance*, *14*(1), 41–55.

Chavez-Demoulin, V., & McGill, J. A. (2012). High-frequency financial data modeling using Hawkes processes. *Journal of Banking & Finance*, *36*(12), 3415–3426.

Chen, Q., Gerlach, R., & Lu, Z. (2012). Bayesian value-at-risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics and Data Analysis*, *56*(11), 3498–3516.

Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, *39*(4), 841–862.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.

Colletaz, G., Hurlin, C., & Tokpavi, S. (2007). Irregularly spaced intraday value at risk (ISIVaR) models: Forecasting and predictive abilities. *Proceedings of the methods in international finance network (1st workshop), Maastricht, 24–25 September, 2007*.

Cotter, J. (2007). Varying the VaR for unconditional and conditional environments. *Journal of International Money and Finance*, *26*(8), 1338–1354.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, *1*(1), 83–106.

Dionne, G., Duchesne, P., & Pacurar, M. (2009). Intraday value at risk (IvaR) using tick-by-tick data with application to the Toronto stock exchange. *Journal of Empirical Finance*, *16*(5), 777–792.

Embrechts, P., Kaufmann, R., & Patie, P. (2005). Strategic long-term financial risks: Single risk factors. *Computational Optimization and Applications*, *32*(1–2), 61–90.

Embrechts, P., Kluppelberg, C., & Mikosh, T. (1997). *Modelling extremal events for insurance and finance*. Berlin: Springer.

Ergun, A. T., & Jun, J. (2010). Time-varying higher-order conditional moments and forecasting intraday VaR and expected shortfall. *Quarterly Review of Economics and Finance*, *50*(3), 264–272.

Fernandez, V. (2005). Risk management under extreme events. *International Review of Financial Analysis*, *14*(2), 113–148.

Gencay, R., & Selcuk, F. (2004). Extreme value theory and value-at-risk: relative performance in emerging markets. *International Journal of Forecasting*, *20*(2), 287–303.

Gencay, R., Selcuk, F., & Ulugulyagci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance, Mathematics and Economics*, *33*(2), 337–356.

Ghorbel, A., & Trabelsi, A. (2008). Predictive performance of conditional extreme value theory in value-at-risk estimation. *International Journal of Monetary Economics and Finance*, *1*(2), 121–148.

Giot, P. (2005). Market risk models for intraday data. *Journal of Finance*, *11*(4), 309–324.
Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance*, *68*(5), 1779–1801.
Goodhart, C. A. E., & O'Hara, M. (1997). High frequency data in financial markets: issues and applications. *Journal of Empirical Finance*, *4*(2), 73–114.
Hogg, R., & Klugman, S. (1984). *Loss distribution*. New York: Wiley.
Karmakar, M., & Shukla, G. K. (2015). Managing extreme risk in some major stock markets: An extreme value approach. *International Review of Economics and Finance*, *35*, 1–25.
Larsen, R., & Marx, M. (1986). *An introduction to mathematical statistics and its application* (2nd ed.). New Jersey: Prentice Hall.
Marimoutou, V., Raggad, B., & Trabelsi, A. (2009). Extreme value theory and value at risk: application to oil market. *Energy Economics*, *31*(4), 519–530.
Martens, M., Chang, Y. C., & Taylor, S. (2002). A comparison of seasonal adjustment methods when forecasting intraday volatility. *Journal of Financial Research*, *25*(2), 283–299.
McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedasticity financial time series: An extreme value approach. *Journal of Empirical Finance*, *7*(3), 271–300.
Mike, K., So, P., & Xu, R. (2013). Forecasting intraday volatility and value at risk with high-frequency data. *Asia-Pacific Financial Markets*, *20*(1), 83–111.
Morimoto, T., & Kawasaki, Y. (2008). *Empirical comparison of multivariate GARCH models for estimation of intraday value at risk.* (Available at SSRN: http://ssrn.com/abstract=109080 or http://dx.doi.org/10.2139/ssrn.1090807).
Nelson, D. B. (1991). Conditional heteroscadasticity in asset returns: A new approach. *Econometrica*, *59*(2), 347–370.
Qi, J., & Lon Ng, W. (2009). Liquidity adjusted intraday value at risk. *Proceedings of the world congress on engineering, Vol II, WCE, July 1–3, London, UK*.
Reiss, R. D., & Thomas, M. (1997). *Statistical analysis of extreme values with applications to insurance, finance, hydrology and other fields*. Basel: Birkhauser Verlag.
Taylor, S., & Xu, X. (1997). The incremental volatility information in one million foreign exchange quotations. *Journal of Empirical Finance*, *4*(4), 317–340.
Watanabe, T. (2012). Quantile forecasts of financial returns using realized GARCH models. *Japanese Economic Review*, *63*(1), 68–80.
Zhang, L., Mykland, P. A., & Ait-Sahalia, Y. (2005). A tale of two time scales: Determining integrated volatility with noisy high frequency data. *Journal of the American Statistical Association*, *100*(472), 1394–1411.