Extreme Market Risk-An Extreme Value Theory Approach

David E Allen, Abhay K Singh & Robert Powell

School of Accounting Finance & Economics
Edith Cowan University

Abstract

The phenomenon of the occurrence of rare yet extreme events, “Black Swans” in Taleb’s terminology, seems to be more apparent in financial markets around the globe. This means there is not only a need to design proper risk modelling techniques which can predict the probability of risky events in normal market conditions but also a requirement for tools which can assess the probabilities of rare financial events; like the recent Global Financial Crisis (2007-2008). An obvious candidate, when dealing with extreme financial events and the quantification of extreme market risk is Extreme Value Theory (EVT). This proves to be a natural statistical modelling technique of relevance. Extreme Value Theory provides well established statistical models for the computation of extreme risk measures like the Return Level, Value at Risk and Expected Shortfall. In this paper we apply Univariate Extreme Value Theory to model extreme market risk for the ASX-All Ordinaries (Australian) index and the S&P-500 (USA) Index. We demonstrate that EVT can be successfully applied to financial market return series for predicting static VaR, CVaR or Expected Shortfall (ES) and expected Return Level and also daily VaR using a GARCH(1,1) and EVT based dynamic approach.

Keywords: Risk Modelling, Value at Risk, Expected Shortfall, Extreme Value Theory, GARCH

Acknowledgements: We are thankful to Manfred Gilli and Evis Këllezi for making available their MATLAB code online. Allen and Powell thank the Australian Research Council for funding support.
1 Introduction

One of the major challenges in modelling VaR is the distributional assumption made for the return data series of the asset or portfolio, which is taken to be normal in most of the quantification approaches. The assumption of normality is not valid when the data series have heavy tails, which are characterised by extreme events left outside the bounds of a normal distribution when modelling VaR. The problem of the normality assumption of the return series, can be addressed by using the distribution free assumption of quantile modelling statistics, and tools such as quantile regression (Koenker and Bassett, 1978) or by applying extreme distribution based methods such as Extreme Value Theory (EVT).

With growing turbulence in the financial markets worldwide, evaluating the probability of extreme events like the GFC, has become an important issue in financial risk management. Quantification of the extreme losses in a financial market is important in current market conditions. EVT provides a comprehensive theoretical base on which statistical models describing extreme scenarios can be formed. The distinguishing feature of EVT is that it provides quantification of the stochastic behavior of a process at unusually large or small levels. Specifically, EVT usually requires estimation of the probability of events that are more extreme than any other that has been previously observed.

EVT, refers to the branch of statistics which deals with the extreme deviations from the mean of a probability distribution. EVT assesses the type of limiting probability distributions for the processes. In broad terms, EVT has two substantial ways of obtaining results or principal models: viz. the Block Maxima model (BMM) and Peak Over Threshold model (POT). Through the block maxima method, the asymptotic distribution of a series of maxima (minima) is modelled and the distribution of the standardized maximum is shown to follow extreme value distributions of Gumbel, Fréchet or Weibull distributions. The generalized extreme value distribution (GEV) is a standard form of these three distributions, and hence the series is shown to converge to GEV. To analyse extreme market events, we are not always interested in maxima or minima of observations, but also in the behaviour of a large exceedance over a given threshold. The Peak over threshold method models a distribution of excess over a given threshold. EVT shows that the limiting distribution of exceedance is a generalized Pareto distribution or GPD (Coles, 2001; Coles and Tawn, 1991; 1994, Franke, Härdle and Hafner, 2008 and Gilli and Këllezi, 2006).

EVT techniques are widely used in areas of hydrology (Coles and Tawn, 1996; Tancredi, Anderson and O’Hagan, 2006; Katz, Parmange, Marc and Naveau, 2002), weather and environment (Pielke et. al., 2000; Pielke et. al., 1998; Smith, 1989; Tarleton and Katz, 1995; Wettstein and Mearns, 2002). EVT is also a well known technique in many fields of applied sciences including engineering and insurance (McNeil, 1999; Embrechts et al., 1999; Reiss and Thomas, 1997 and Giesecke & Goldberg, 2005). Numerous research studies surfaced recently which analyse the extremes in the financial markets due to currency crises, stock market turmoil and credit defaults. The behaviour of financial series tail distributions has, among others, been discussed in Mancini and Trojani (2010), Onour (2010), Gilli and Këllezi (2006), Koedijk et al. (1990), Dacorogna et al. (1995), Loretan and Phillips (1994), Longin (1996), Daniels-son and de Vries (2000), Kuan

Despite the promise of useful implementation of EVT in financial market analysis, it has only recently gained the attention of researchers in Australia. Chan and Gray (2006), Thomas et al. (2006) and Jeyasreeharan et al. (2009) are amongst the few studies to have used the technique. The lack of implementation of EVT methods on Australian markets act as our motivation to test it further on Australian market. This particular research paper also targets the United States market to analyse the recent GFC and the crash of 1987 as natural comparators.

In this paper we model and estimate the static next day VaR, ES and different return period Risk Levels for the ASX-All Ordinaries and the S&P-500 stock exchange index series of daily log-return data using EVT. We follow the method of Gilli and Këllezi (2006) to model Return level, VaR and ES using their MATLAB code to generate the results and also to test the models for different block lengths and threshold values. We also model these extreme risk measurements in a dynamic two stage extreme value process with a GARCH (1,1) model (McNeil and Frey, 2000), to forecast daily VaR and ES with historical data in a moving window. We also add return period calculations to our BMM analysis and pick the two major financial crashes of 1987 and 2008 from our historical data and forecast the return period for the same, which is helpful in assessing the likelihood of these crashes in the future.

The rest of the paper is designed as follows; in section-2 we give more details about EVT and the associated risk measures, in section-3 we outline the dynamic-EVT method for VaR and ES estimation. In Section-4 we discuss our empirical design, and provide a data description together with our research design and methodology. We discuss the results in section-5 and conclude in section-6.

2 Extreme Value Theory and Extreme Risk Modelling

EVT provides simple parametric models to capture the extreme tails of a distribution and to forecast risk. Mainly there are two broad methods of applying EVT: the first of which is based on the extreme value distributions of the Gumbel, Fréchet or Weibull distributions which are generalized as the Generalized extreme value distribution (GEV) and known as the Block Maxima (Minima) (BMM) approach, whilst the second is based on the Generalized Pareto Distribution (GPD) and is known as the peak over threshold (POT) approach. The BMM models are the most traditional of the two, and the BMM approach fits a block of maxima (minima) (extreme events) in a data series of independent and identically distributed observations (iid) to GEV using different statistical methods; the most common of which is via Maximum Likelihood Estimation (MLE). POT is considered more efficient in modelling limited data (Gilli and Këllezi, 2006; McNeil, Frey and Embrecht, 2005) as it fits the exceedances over a given threshold in a data set to GPD and hence is not as dependent on the requirement for large data sets as BMM. Our discussion of EVT in this paper is adopted from Embrechts, Klüppelberg & Mikosch (1997), Coles (2001), McNeil and Frey (2000), Gilli and Këllezi (2006), McNeil, Frey & Embrechts (2005), Franke, Härdle and Hafner (2008).
2.1 The Generalized Extreme Value Distribution (GEV) & Block Maxima Method

Consider $X_n$ as a series of random iid variables $X_1, \ldots, X_n$ with cumulative distribution function (cdf) $F(x)$ with a stochastic maximum $M_n = \max(X_1, \ldots, X_n)$. When dealing with financial risk, $X_t = -r_t$, i.e., the negative return at day $t$. The cdf of $M_n$ is given by (Franke, Härdle and Hafner, 2008; McNeil, Frey and Embrechts, 2005; Embrechts, Klüppelberg and Mikosch, 1997)

$$P(M_n \leq x) = P(X_1 \leq x, \ldots, X_n \leq x) = \prod_{t=1}^{n} P(X_t \leq x) = F^n(x). \quad (1)$$

Considering only unbounded random variables $X_t$ i.e. $F(x) < 1 \forall x < \infty$, it holds that $F^n(x) \to 0 \forall x$, if $n \to \infty$ and hence $M_n \to \infty$. $M_n$ has to be standardised to achieve a non-degenerate behaviour limit.

**Theorem 1. (Fisher and Tippett, 1928; Gnedenko, 1943).** If $X_n$ is a series of i.i.d. random variables. If for a non-degenerate distribution function $H$, there exist a constant $c_n > 0$ and $d_n \in \mathbb{R}$, then

$$\frac{M_n - d_n}{c_n} \xrightarrow{d} H, \quad (2)$$

$H$ here belongs to a GEV distribution.

GEV is generalised representation of the following three distributions:

Fréchet:

$$\Phi_\alpha(x) = \begin{cases} 0, & x \leq 0 \\ e^{-x^{-\alpha}}, & x > 0 \end{cases} \quad \alpha > 0 \quad (3)$$

Weibull:

$$\Psi_\alpha(x) = \begin{cases} e^{-(x)^\alpha}, & x \leq 0 \\ 1, & x > 0 \end{cases} \quad \alpha > 0 \quad (4)$$

Gumbell:

$$\Lambda(x) = e^{-e^{-x}}, \quad x \in \mathbb{R}. \quad (5)$$

The distribution function (df) of the standard GEV (Jenkinson, 1955; von Mises, 1954) is given by

$$H_\xi(x) = \begin{cases} e^{-(1+\xi x)^{-1/\xi}}, & \text{if } \xi \neq 0 \\ e^{-e^{-x}}, & \text{if } \xi = 0 \end{cases} \quad (6)$$

Here $x$ is such that $1 + \xi x > 0$. We obtain the three parameter family by defining $H_{\xi,\mu,\sigma}(x) :=$
$H_{\xi}(x - \mu)/\sigma)$ for a location parameter $\mu \in \mathbb{R}$ and a scale parameter $\sigma > 0$. With this generalization, $\xi$ is known as the shape parameter of the GEV distribution and $H_{\xi}$ gives the type of distribution. $\xi > 0$ for Fréchet distribution, $\xi < 0$ for the Weibull distribution and $\xi = 0$ for the Gumbel distribution. Figure 1.1 gives the probability density functions of these three distributions.

![Density Fréchet Distribution](image1)
![Density Weibull Distribution](image2)
![Density Gumbell Distribution](image3)

(a) Fréchet Distribution  
(b) Weibull Distribution  
(c) Gumbell Distribution

Figure 1: Probability density functions

2.1.1 The Block Maxima Method (BMM)

In Block Maxima Method (McNeil, Frey and Embrechts, 2005), suppose we have a data series typically having series of maxima for a fixed block size $n$ from an underlying distribution $F$, which is supposed to lie in domain of attraction of a GEV $H_{\xi}$ for some $\xi$. If the data is series of iid variables, it can be implied that the true distribution of $n$-block maximum $M_n$ can be approximated for large enough $n$ by a GEV, $H_{\xi,\mu,\sigma}$.

BMM uses this idea to fit the GEV distribution $H_{\xi,\mu,\sigma}$ to a data series containing block maximum for an equal period $n$. The parameters for the GEV fit $(\hat{\xi}, \hat{\mu}, \hat{\sigma})$ are estimated by maximum likelihood estimation (MLE), the confidence interval estimates for the parameters are estimated by profile likelihood estimation (Barndorff-Nielson and Cox, 1994). BMM originated in hydrology for extreme modelling e.g., annual maxima of rain fall, yet it can be analogously applied to financial daily return data by dividing the datasets into yearly, semester, quarterly or monthly blocks. The daily maximum in these blocks can be analyzed using BMM as we will see later in the empirical study of the S&P-500 and the ASX-All Ordinaries stock indices.

2.1.2 Return Level

The Generalized Extreme Value model fitted to the data of block maxima (minima) can be used to analyse extreme losses. This can be approached in two ways; in the first approach, known as return level estimation, we can define the return period of the occurrence of the extreme event and predict its magnitude and in the second, the return period estimation approach, we can calculate the return period for a given return level.
If $H$ is the distribution of maxima (minima) observed over period of time (non overlapping and equal periods), the return level is given as

$$R_n^k = H_{\xi,\mu,\sigma}^{-1} \left( 1 - \frac{1}{k} \right)$$

which is the return level expected to be exceeded in one out of $k$ periods ($k = 1/p$) of length $n$. This is a conservative measure of severe loss of a portfolio or an asset in financial risk.

### 2.2 Generalized Pareto Distribution & Peak Over Threshold (POT)

There are two major results of EVT, first Block Maxima Model (Section -) which fits a series of block maximas to the GEV distribution and the second based on threshold exceedances known as Peak Over Threshold which fits the excess distribution to the Generalized Pareto distribution (GPD). The POT method uses available data more efficiently which is an obvious advantages over BMM, in POT we use all the data which exceeds a particular threshold level while in BMM only the maximum from a block length is retained for distribution estimation.

**Theorem 2. (Pickands (1975), Balkema and de Haan (1974)).** For a large class of underlying distributions $F$, the excess distribution function $F_u$ can be approximated by GPD for an increasing threshold $u$.

$$F_u(y) \approx G_{\xi,\sigma}(y), \quad u \to \infty$$

$G_{\xi,\sigma}$ in theorem-2 is the Generalized Pareto Distribution (GPD) which is given by

$$G_{\xi,\sigma}(y) = \begin{cases} 
(1 + \frac{\xi}{\sigma} y)^{-1/\xi} & \text{if } \xi \neq 0 \\
1 - e^{-y/\sigma} & \text{if } \xi = 0
\end{cases}$$

for $y \in [0, (x_F - u)]$ if $\xi \geq 0$ and $y \in [0, -\frac{\sigma}{\xi}]$ if $\xi < 0$. Here $\xi$ is the shape parameter and $\sigma$ is the scale parameter for GPD.

Figure-2 gives the density plots for different values of $\xi$, the shape parameter in GPD.
Definition 3. (Excess Distribution). For a random variable \( X \) with df \( F \), the excess distribution over a threshold \( u \) is given by

\[
F_u(y) = P(X - u \leq y|X > u) = \frac{F(x) - F(u)}{1 - F(u)},
\]  

(9)

for \( 0 < y < x_F - u \) where \( x_F \leq \infty \) is the right endpoint of \( F \) and \( y = x - u \). \( F_u \) is the conditional excess distribution function.

2.2.1 VaR and Expected Shortfall

If there is an extreme distribution \( F \) with right endpoint \( x_F \), we can assume that for some threshold \( u \), \( F_u(x) = G_{\xi,\sigma}(x) \) for \( 0 \leq x < x_F - u \) and \( \xi \in \mathbb{R} \) and \( \sigma > 0 \). For \( x \geq u \),

\[
\tilde{F}(x) = P(X > u)P(X > x|X > u)
= \tilde{F}(u)P(X - u > x - u|X > u)
= \tilde{F}(u)\tilde{F}_u(x - u)
= \tilde{F}(u)\left(1 + \frac{\xi (x - u)}{\sigma}\right)^{-1/\xi},
\]  

(10)

given \( F(u) \), this gives a formula for tail probabilities. The inverse of (2.10) gives the high quantile of the distribution or VaR. For \( \alpha \geq F(u) \), VaR is given by
\[ VaR_\alpha = q_\alpha (F) = u + \frac{\sigma}{\xi} \left( \frac{1 - \alpha}{F(u)} \right)^{-\xi} - 1 \] (11)

For \( \xi < 1 \) the ES is given by

\[ ES_\alpha = \frac{1}{1 - \alpha} \int_0^1 q_x (F) dx = \frac{VaR_\alpha}{1 - \xi} + \frac{\sigma - \xi u}{1 - \xi} \] (12)

Analytical expressions for VaR and ES can also be defined as a function of estimated GPD parameters. Using (9)

\[ F(x) = (1 - F(u))F_u(y) + F(u), \]

if \( n \) is the total observations and \( N_u \) the number of observations above \( u \) and we replace \( F_u \) by the GPD and \( F(u) \) by \( (n - N_u)/n \), we get an estimator for tail probabilities (Smith, 1987)

\[ \hat{F}(x) = 1 - \frac{N_u}{n} \left( 1 + \frac{\hat{\xi}}{\hat{\sigma}} (x - u) \right)^{-1/\hat{\xi}} \] (13)

The inverse of (13) with a probability \( p \) gives the VaR

\[ \hat{VaR}_p = u + \frac{\hat{\sigma}}{\hat{\xi}} \left( \left( \frac{n}{N_u} p \right)^{-\hat{\xi}} - 1 \right) \] (14)

Using (12) the ES is given by

\[ \hat{ES}_p = \hat{VaR}_p \frac{1}{1 - \xi} + \frac{\hat{\sigma} - \hat{\xi}u}{1 - \xi} \] (15)

In POT method GPD is fitted to the excess distribution (value above threshold \( a \)) by MLE and the confidence interval estimates are calculated by profile likelihood and then the unconditional or static estimates for VaR and ES are calculated.

3 EVT VaR and ES-A Dynamic Approach

In VaR and ES calculations using POT when EVT is applied directly to raw return data assuming the distribution to be stationary or unconditional, the EVT model can be termed a static model (McNeil and Frey, 2000). EVT can also be used in a dynamic model, where the conditional distribution of \( F \) is taken into account and the volatility of returns is captured. The dynamic model uses an ARCH/GARCH type process along with the POT to model VaR and ES which reacts to fluctuations in market and hence captures current risk (McNeil and Frey, 2000), we discuss this dynamic method in this section.

McNeil and Frey (2000), proposed a dynamic VaR forecasting method based using EVT, their method makes use of GARCH modelling to model the current market volatility background which is further fed into VaR estimates obtained from the POT model fitted to residuals of a GARCH model. By use of GARCH models to forecast the estimates of conditional volatility the
model provides dynamic one day ahead forecasts for VaR and ES for the financial time series.

Let $R_t$ the return at time $t$ be defined by the following stochastic volatility (SV) model

$$R_t = \mu_t + \sigma_t Z_t,$$  \hspace{1cm} (16)

where $\mu_t$ is the expected return on day $t$ and $\sigma_t$ is the volatility and $Z_t$ gives the noise variable with a distribution $F_Z(z)$ (commonly assumed to be standard normal). We assume that $R_t$ is a stationary process.

The most widely used suitable models are drawn from the ARCH/GARCH family. An autoregressive GARCH(1,1) process is given by

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2,$$  \hspace{1cm} (17)

where $\varepsilon = R_{t-1} - \mu_{t-1}$, $\mu_t = \lambda R_{t-1}$, $\alpha_0, \alpha_1, \beta > 0$, $\beta + \alpha_1 < 1$ and $|\lambda| < 1$.

In contrast to static risk modelling using EVT, where we model the unconditional distribution $F_X(x)$ and are interested in loss for $k$ days in general, the dynamic approach models the conditional return distribution conditioned on the historical data to forecast the loss over the next $k \geq 1$ days. If we follow the GARCH(1,1) model the one day ahead forecast of VaR and ES are calculated as:

$$\text{VaR}_q = \mu_{t+1} + \sigma_{t+1} \text{VaR}(Z_q)$$

$$\text{ES}_q = \mu_{t+1} + \sigma_{t+1} \text{ES}(Z_q)$$  \hspace{1cm} (18)

With the assumption that $F_Z(z)$ is a known standard distribution, typically a normal distribution $Z_q$ can be easily calculated. The EVT approach (McNeil and Frey, 2000), instead of assuming $F_Z(z)$ to be normal applies the POT estimation procedure to this distribution of residuals.

For a return series at the close of day $t$ with time window of last $n$ returns ($R_{t-n+1}, \ldots, R_t$) the method is implemented in following two steps.

1. A GARCH(1,1) model is fitted to the historical data by pseudo maximum likelihood estimation (PML) also known as Quasi-maximum likelihood estimation. The GARCH (1,1) model in this step gives the residuals for step-2 and also 1 day ahead predictions of $\mu_{t+1}$ and $\sigma_{t+1}$.

2. EVT (POT method) is applied to the residuals extracted from step-1 for a constant choice of threshold $u$ to estimate $\text{VaR}(Z)_q$ and $\text{ES}(Z)_q$ to calculate the risk measures using equation-18.

The parameters of the GARCH model are estimated by the pseudo-maximum likelihood (PML) method. The likelihood of GARCH with a normality assumption is maximised to obtain parameter estimates $\hat{\theta} = (\hat{\lambda}, \hat{\alpha}_0, \hat{\alpha}_1, \hat{\beta})^T$. Although this means fitting the model with a normality
assumption, which is not always true for financial return data, PML usually generates fair es-
timates which are consistent and asymptotically normal (Gouriéroux, 1997). The POT method
in step-2 is fitted using MLE.

We will implement this method to forecast one day ahead VaR for ASX-All ordinaries and
S&P-500 indices and will compare the results with standard GARCH(1,1) and RiskMetrics\textsuperscript{TM}
based estimates later in the empirical exercise. The results will be backtested by application of
the binomial method based on the number of daily violations above VaR.

4 Data & Research Methodology

4.1 Description of Data

The objective here is to implement various EVT based risk modelling methods to model
extreme market risk. The focus of this empirical study is to model the market risk of the ASX-
All Ordinaries stock index’s daily log return data, this particular index is chosen as the historical
data available for this index dates back to 1973 and we require large data sets for implementing
EVT based methods, particularly BMM. We also model market risk using EVT for the S&P-500
index for the same data period.

The data period used here is from 03/01/1973-03/12/2010, which gives us approximately
38 year blocks and 155 quarterly blocks for BMM calculations. Also this data period includes
the crash of 1987 and recent GFC (2007-2008) for both stock markets. The data is downloaded
from Reuters Datastream (DS) and the series used are DS calculated series for both the indices
in US Dollars. We use DS calculated return series for the indices as it provides daily returns for
a longer period than the original.

Table-1 gives the summary statistics for the daily logarithmic returns in US Dollars for our
two datasets, it can be noted that the ASX-All Ordinaries is slightly more volatile than S&P-500
for the given period.

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Min.</td>
<td>-30.15918</td>
<td>-21.22849</td>
</tr>
<tr>
<td>Median</td>
<td>0.0279</td>
<td>0.0143</td>
</tr>
<tr>
<td>Mean</td>
<td>0.02523</td>
<td>0.02571</td>
</tr>
<tr>
<td>Max.</td>
<td>8.34088</td>
<td>10.90037</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>1.369605</td>
<td>1.092627</td>
</tr>
<tr>
<td>1% Quantile</td>
<td>-3.554387</td>
<td>-2.872808</td>
</tr>
<tr>
<td>5% Quantile</td>
<td>-2.002421</td>
<td>-1.614634</td>
</tr>
<tr>
<td>95% Quantile</td>
<td>1.98204</td>
<td>1.627681</td>
</tr>
<tr>
<td>99% Quantile</td>
<td>3.349155</td>
<td>2.902302</td>
</tr>
<tr>
<td>No. Of observations</td>
<td>9893</td>
<td>9893</td>
</tr>
</tbody>
</table>

Table 1: Summary Statistics for Data
4.2 Research Design & Methodology

We implement BMM, POT and a two step dynamic POT method to model extreme market risk for ASX-All Ordinaries and S&P-500 daily log return data. The methodology is outlined as follows:

1. First we quantify the Return level for various return periods and yearly Return Periods for the two major financial crashes of 1987 and 2008 using the BMM approach. We implement BMM on the data in yearly and quarterly block sizes. We generate both point and 95% confidence interval values for the estimated parameters and risk measures.

2. After BMM we use the same daily return dataset to quantify static VaR and ES for both the indices using POT method. We model POT for two different threshold ($u$) values; the lower 5% and 10% quantile of the distribution. The threshold can also be selected by use of a sample mean excess plot and we will show that the quantile threshold values in fact satisfies the sample mean excess plot criteria. Again point and interval estimates for both the indices are generated.

3. Finally we implement the third method of dynamic EVT to model daily VaR for both the indices by use of the last 1000 log returns (approximately 4 years) of data in a daily moving window, we use last 10 years (approximately) data from January-2000 to December-2010. The VaR estimates generated from the two step method will also be compared with normal GARCH(1,1) and RiskMetrics$^{TM}$ method by backtesting.

For the calculations we will use the MATLAB code of Gilli and Këllezi (2006) which is freely available from their website and self-coded R routines, both MATLAB and R codes use MLE for parameter estimation of EVT.

The results are discussed in the next section.

5 Discussion of the Results

5.1 The Block Maxima Method

When fitting block maxima to GEV we extract the $n$—period maximums from a sample, which is then fitted to the GEV distribution, the fit is finally used to compute point and interval estimates of return levels and return periods. We generate results for both the left and right tails of the return distribution, in financial applications to model the left tail of the distribution we change the signs of the return data such that $\hat{r}_t = -r_t$, for the right tail data is used as it is.

The main implementation of BMM model in this paper use two block lengths, yearly and quarterly, we do not quantify the model using a monthly block length as the quantile-quantile (Q-Q) plot of the block of maxima and the theoretical GEV shows that the monthly block
length does not necessarily follow GEV for left tail of our return data series. Figure-3 gives four different plots; a time series plot of block maxima, a plot of GEV residual density, a scatterplot of residuals and Q-Q plot of residuals, for ASX-All Ordinaries left tail data. The Q-Q plot from the figures clearly illustrates that the block size of a month is not a good fit to GEV distribution and hence we do not proceed with the monthly block length any further.

Figure 3: GEV Fit Plots for ASX-All Ordinaries-Monthly blocks

5.1.1 Yearly BMM

The return data is divided in 33 yearly overlapping sub-samples as the number of trading days in our yearly samples are not equal the sample blocks are not of equal size. The block maxima data consists of the maximum return for each block, which is used to estimate GEV. The yearly block choice is good for the purpose of avoiding seasonal effects in financial data. Figure-4 plots the yearly maxima for the left and right tails of the ASX-All Ordinaries return data. The block maxima data is fitted by MLE to get the point estimates and the interval estimates are obtained by profile log likelihood. As our focus in this study is Australian market, we will give graphical results only for ASX-All Ordinaries.

Figure 4: Yearly minima and maxima of the daily returns of the ASX-All Ordinaries

Figure-5 (a) and (b) give the fitted GEV distribution with sample distribution of ASX-All
Ordinaries for minima and maxima, with these figures it is safe to say that the estimated models fit the data.

![Figure 5: Fitted GEV with sample distribution-ASX All Ordinaries](image)

The ten year return level $R^{10}$ for the minima of ASX-All ordinaries is shown in figure-6(a) and for the maxima in figure-6(b), the return level is plotted against profile log likelihood in the return level graphs.

![Figure 6: $R^{10}$-ASX-All Ordinaries](image)

Table-2 gives the point and interval estimates (95% confidence intervals) for the parameters of both left tails and rights tail along with ten year return levels. The point estimates of $\xi$ for both the tail of both the indices indicate the distribution of minima (left tail) and maxima (right tail) follow the Fréchet distribution family. Return level shows that the ASX-All Ordinaries is more prone to losses (left tail) on average in a year than S&P-500, but if we look at the right tail return levels for ten years the difference is not much. According to point estimates of $R^{10}$ the ASX-All Ordinaries is prone to exceed a negative return of 9.32 at least in one year on average in ten years, whereas this value for S&P-500 is 7.66. For right tails (positive returns) these values are 6.88 and 6.19 for ASX-All Ordinaries and S&P-500 respectively.
<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Point Estimate</td>
</tr>
<tr>
<td><strong>Left Tail</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi$</td>
<td>0.207</td>
<td>0.522</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>1.06</td>
<td>1.305</td>
</tr>
<tr>
<td>$\tau_{10}$</td>
<td>6.974</td>
<td>9.32</td>
</tr>
<tr>
<td><strong>Right Tail</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\xi$</td>
<td>-0.051</td>
<td>0.3</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>0.829</td>
<td>1.084</td>
</tr>
<tr>
<td>$\tau_{10}$</td>
<td>5.608</td>
<td>6.88</td>
</tr>
</tbody>
</table>

Table 2: Point and interval estimates for yearly BMM

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries Return Period (Years)</th>
<th>S&amp;P-500 Return Period (Years)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Return Level</td>
<td>Lower Bound</td>
</tr>
</tbody>
</table>

Table 3: Return Periods for different return levels (Crash of 1987 and 2008)

In table-3 we give the point and interval estimates for the Return periods of the two financial crashes of 1987 and 2008. The point estimates for ASX-All Ordinaries show that a crash similar to the crash of 1987 is likely to occur in one out of 109.35 years and loss that occurred during the GFC can repeat itself in 28.25 years. For S&P-500 the crash of 1987 is likely to occur in one out of 55.78 years and S&P-500 can loose as much as the loss of GFC in one out of 14.2 years.

5.1.2 Quarterly BMM

We now take a smaller block size to evaluate BMM, here we take the maximum from quarter year blocks from our two datasets to finally fit them to GEV. Figure-7 gives the four graphs showing block maxima value plot, GEV residual histogram, scatterplot of residuals and Q-Q plot of residuals for the GEV fit for left tail of ASX-All Ordinaries, with the plot it is safe to assume that the limiting distribution of quarterly maxima for the indices return follow the GEV distribution.
Figure 7: GEV fit plots for ASX-All Ordinaries (Left Tail)-Quarterly BMM

Figure 8 gives the return level plot for left and right tail of the ASX-All Ordinaries, here the return levels are plotted against return periods.

Figure 8: Return Level Plots-ASX-All Ordinaries (Left & Right Tails)

Table-4 gives the point and interval estimates of the parameters for both left and right tail
GEV fits. Finally table-5 gives the return levels (point and interval estimates) for 4, 40 and 400 quarter return periods for both right and left tails of our two datasets.

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Point Estimate</td>
</tr>
<tr>
<td>Left Tail</td>
<td>( \hat{\xi} )</td>
<td>0.1764881</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma} )</td>
<td>0.8219721</td>
</tr>
<tr>
<td>Right Tail</td>
<td>( \hat{\xi} )</td>
<td>0.1736475</td>
</tr>
<tr>
<td></td>
<td>( \hat{\sigma} )</td>
<td>0.6261321</td>
</tr>
</tbody>
</table>

Table 4: Point and interval estimates for quarterly BMM

<table>
<thead>
<tr>
<th>Period</th>
<th>ASX-All Ordinaries Return Level</th>
<th>S&amp;P-500 Return Level</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Point Estimate</td>
</tr>
<tr>
<td>Left Tail</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.66</td>
<td>3.93</td>
</tr>
<tr>
<td>40</td>
<td>7.49</td>
<td>8.78</td>
</tr>
<tr>
<td>400</td>
<td>13.42</td>
<td>17.92</td>
</tr>
<tr>
<td>Right Tail</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3.36</td>
<td>3.58</td>
</tr>
<tr>
<td>40</td>
<td>6.50</td>
<td>7.72</td>
</tr>
<tr>
<td>400</td>
<td>11.25</td>
<td>16.26</td>
</tr>
</tbody>
</table>

Table 5: Return Levels for Quarterly-BMM

The quarterly results also show that the distribution of positive and negative maxima for the two indices follow the Fréchet family of distributions. The 40 period return levels for both the indices are also close to the 10 year return level calculated from yearly BMM.

5.2 The POT Method

The POT method involves the selection of a threshold \( u \), the exceedance above which are then fitted to the GPD function after which point and interval estimates for VaR and ES are calculated. The threshold can be selected by using a mean excess plot which is plotted by using GPD mean excess function. If we have a positive-valued extreme data (loss data) \( X_1, \ldots, X_n \), the estimator \( e(u) \) is given by

\[
e_n(v) = \frac{\sum_{i=1}^{n}(X_i - v)I_{\{X_i > v\}}}{\sum_{i=1}^{n}I_{\{X_i > v\}}} \tag{19}
\]

where \( u \leq v < \infty \) and \( I_{\{X_i > v\}} \) are the values exceeding threshold \( v \). This function is explored by mean excess plot

\[\{(X_{i,n}, e_n(X_{i,n})) : 2 \leq i \leq n\}\]

\( X_{i,n} \) is the \( i \)-th order statistic. The plot shows linearity in a region where above the threshold \( v \) the data supports GPD model. In ideal situations the linearity can be interpreted as

16
• Upward linear trend indicates a positive shape parameter ($\xi$) for the GPD.
• Horizontal linear trend indicates a GPD with $\xi \approx 0$.
• Linear downward trend can be interpreted as GPD with negative $\xi$.

Figure-9 gives a mean excess plot for ASX-All ordinaries daily log return data with a threshold $u=2.01$ and an upward linear trend and hence a positive $\xi$.

![Figure 9: Mean Excess Plot ASX-All Ordinaries](image)

The first step, i.e. the selection of $u$ is critical, $u$ should be high enough to satisfy the condition of GPD but not too high to decrease the number of observations significantly. Here we use a particular lower quantile of the daily return data as the threshold which can be shown agreeing to the mean excess plot method of selecting $u$. In figure-9, $u = 2.01$ is the 95% quantile of negative log return data of ASX-All Ordinaries and it well lies on the acceptable linear region on the plot. We will model POT using two different thresholds 95% and 90% of $-r_t$ for the left tail and 95% and 90% of $r_t$ for the right tail.

Figure-10 gives the plot of excesses above lower 5% (95% of $-r_t$) quantile of the ASX-All Ordinaries, it plots the timeseries of returns above $u$ for the left tail.

![Figure 10: Excess plot-ASX-All Ordinaries](image)
Figure-11(a) gives the fitted GPD model with the exceedances above the threshold for the left tail of ASX-All Ordinaries, figure-11(b) plots the same for the right tail. The plot shows that the estimated GPD fits the exceedances.

![Figure 11: GPD fitted to tail exceedances.](image)

The point and interval estimates of the parameters of the fitted GPD model for both tails along with 1% VaR and 1% ES values are given in table-6.

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Point Estimate</td>
<td>Upper Bound</td>
</tr>
<tr>
<td><strong>Left Tail</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\xi}$</td>
<td>-0.1</td>
<td>0.313</td>
<td>0.411</td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.721</td>
<td>0.805</td>
<td>0.901</td>
</tr>
<tr>
<td>$VaR_{1%}$</td>
<td>3.516</td>
<td>3.693</td>
<td>3.899</td>
</tr>
<tr>
<td>$ES_{1%}$</td>
<td>5.098</td>
<td>5.635</td>
<td>6.492</td>
</tr>
<tr>
<td><strong>Right Tail</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\xi}$</td>
<td>0.136</td>
<td>0.212</td>
<td>0.304</td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.655</td>
<td>0.728</td>
<td>0.809</td>
</tr>
<tr>
<td>$VaR_{1%}$</td>
<td>3.238</td>
<td>3.377</td>
<td>3.534</td>
</tr>
<tr>
<td>$ES_{1%}$</td>
<td>4.334</td>
<td>4.676</td>
<td>5.204</td>
</tr>
</tbody>
</table>

Table 6: POT- Point and interval estimates for ASX-All Ordinaries and S&P-500.

Looking at the VaR and ES from table-3.6, with a 1% confidence level we can predict tomorrow’s loss (left tail) for the ASX-All Ordinaries to exceed 3.693% and if this happens the corresponding expected loss will be 5.635%. The same inferences can be drawn for the S&P-500 and for both right and left tails.

Next we look at the results from the POT method fitted to excesses over 90% of the return data, i.e. $u = 90\%$ quantile. Figure-12 gives the GPD fit plots for ASX-All Ordinaries (left tail) the plots show that the estimated model fit the exceedances.
Table 7 gives the point and interval estimates of GPD fits along with 1% VaR and ES.

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lower Bound</td>
<td>Point Estimate</td>
</tr>
<tr>
<td><strong>Left Tail</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\xi}$</td>
<td>0.117</td>
<td>0.182</td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.597</td>
<td>0.656</td>
</tr>
<tr>
<td>$\hat{VaR}_{1%}$</td>
<td>2.923</td>
<td>3.001</td>
</tr>
<tr>
<td>$\hat{ES}_{1%}$</td>
<td>3.923</td>
<td>4.221</td>
</tr>
<tr>
<td><strong>Right Tail</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{\xi}$</td>
<td>0.081</td>
<td>0.151</td>
</tr>
<tr>
<td>$\hat{\sigma}$</td>
<td>0.585</td>
<td>0.645</td>
</tr>
<tr>
<td>$\hat{VaR}_{1%}$</td>
<td>2.818</td>
<td>2.938</td>
</tr>
<tr>
<td>$\hat{ES}_{1%}$</td>
<td>3.747</td>
<td>4.014</td>
</tr>
</tbody>
</table>

Table 7: POT- Point and interval estimates for ASX-All Ordinaries and S&P-500 ($u=90\%$ quantile)

Figure-13 gives the tail plot of ASX-All Ordinaries with 1% VaR and 1% ES with their 95% confidence interval estimates.
5.3 Dynamic-EVT VaR

EVT can not only be used in a static approach to predict VaR as seen in the results of previous subsection, it can also be used in a dynamic model to predict time varying VaR estimates (section-3.4). Here we use a moving window of the last 1000 days log returns for ASX-All Ordinaries and S&P-500 indices to forecast one day ahead 1% and 5% VaR estimates. The total data period is approximately 10 years (Jan-2000 to Dec-2010) containing 2850 daily log returns for both the indices, which gives us a total of 1850 predictions.

The method uses a two step approach in which we predict the next day volatility ($\sigma$) and mean expected return ($\mu$) using a GARCH (1,1) model in first step and in the second we fit the residuals of the step-1 to GPD to get quantile values for final VaR calculations (equation-28). We chose a 90% quantile level as threshold, $u$ to fit the residuals from the GARCH(1,1) model to GPD. The forecasts from this method are compared with the forecasts from normal a GARCH (1,1) where residuals are assumed to belong to normal distribution and to the RiskMetrics$^{TM}$ model (Morgan, 1996).

We use a violation based backtesting method (McNeil and Frey, 2000) for the forecasted 1% and 5% VaR estimates. If we have a next day predicted quantile $\hat{r}_q^t$ and the actual return $r_{t+1}$, a violation is said to occur if $r_{t+1} > \hat{r}_q^t$, i.e. the actual loss is greater than the forecasted VaR. A binomial test for the success of these VaR forecasting models can be developed based on the number of violations. The test based on violations counts only two possible (binomial) outcomes of a violation or no violation. If $q$ is the quantile for VaR (95% and 99%) the estimated number of violations are given by $(1 - q)Total\ Predictions\ (Trials)$. We will calculate a two sided binomial test of the null hypothesis against the alternative that the method has prediction errors and it underestimates (too many violations) or overestimates (too few violations) the
conditional quantile.

Figure-14 gives the plot of 1% VaR estimates of ASX-All Ordinaries from the three models plotted with the actual return series for the prediction period. It is evident from the figure that the dynamic-EVT method closely follows the changing return dynamics of the market. Also in figure-15 we plot dynamic-EVT based on 1% VaR with the original return series for the ASX-All Ordinaries, the estimates here are changing closely with the changing market dynamics and they estimate the extreme risk better in the extreme market conditions.

Figure 14: VaR Forecasts-ASX-All Ordinaries

Figure 15: Dynamic-EVT VaR forecasts.

Table-8 gives the backtest statistics for the models along with the two-sided p-value, a p-value greater than 0.05 shows the rejection of alternate hypothesis and hence is significant. The
results show that apart from on one occasion (5% VaR for S&P-500) the dynamic-EVT method works better than all the other methods, in this case even when the p-value does not approve the method the method still has the least number of violations. Other significant result is that the other two models i.e. GARCH(1,1) and RiskMetrics\textsuperscript{TM} fail for both quantile levels except RiskMetrics\textsuperscript{TM} for the ASX-All Ordinaries (q=0.95).

<table>
<thead>
<tr>
<th></th>
<th>ASX-All Ordinaries</th>
<th>S&amp;P-500</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Predictions</td>
<td>1850</td>
<td>1850</td>
</tr>
<tr>
<td>q=0.99</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Dynamic-EVT</td>
<td>23(0.29)</td>
<td>34(0.00)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>42(0.00)</td>
<td>49(0.00)</td>
</tr>
<tr>
<td>RiskMetrics\textsuperscript{TM}</td>
<td>43(0.00)</td>
<td>45(0.00)</td>
</tr>
<tr>
<td>q=0.95</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Expected</td>
<td>92</td>
<td>92</td>
</tr>
<tr>
<td>Dynamic-EVT</td>
<td>81(0.24)</td>
<td>104(0.23)</td>
</tr>
<tr>
<td>GARCH(1,1)</td>
<td>123(0.00)</td>
<td>115(0.02)</td>
</tr>
<tr>
<td>RiskMetrics\textsuperscript{TM}</td>
<td>107(0.12)</td>
<td>117(0.01)</td>
</tr>
</tbody>
</table>

Table 8: Results-Backtesting VaR

The forecasted period here includes the period of the GFC and it can be seen from the forecasted VaR that the method works well in the crisis period as well, which shows the capabilities of the EVT approach for modelling extreme market events. The dynamic model changes itself with changing market dynamics and hence the forecasted VaR values represent more closely the extreme risk of the market.

## 6 Conclusion

In this paper we focused on the extreme market risk of financial markets and illustrated how extreme value theory can be used to model extreme events and their associated risk levels. EVT as demonstrated by the empirical exercise here, can be used to model financial risk measures like VaR, ES and Return Level to asses extreme tail events.

EVT can be used to quantify the size of extreme events, with the help of the two major approaches to application of EVT; BMM and POT. The applicability of both the methods depend on the the availability of data, the desired time horizon and the kind of risk measures we want to forecast. For fairly large data sets with non overlapping block periods BMM method can be a useful technique as it is simple to implement and provides Return Level and Return Period forecasts which are useful for stress testing purposes. The POT method has its advantages in modelling the available data more efficiently than BMM as it uses excesses over a threshold and can be more effective if we have limited data sets. The techniques give point as well as interval estimates of the risk measures which are useful in risk assessment in financial risk management.

We also demonstrate how we can use a GARCH based dynamic-EVT approach to model VaR for short term forecasting. The dynamic-EVT method has the advantage of dynamically reacting to changing market conditions which is useful in getting better VaR forecasts. We
show with our analysis that this method performs better than the other widely used methods of normal GARCH(1,1) and RiskMetrics™, not only in normal market conditions but also in extreme market conditions such as the recent GFC.

To summarize we followed both unconditional and conditional EVT based VaR estimation models to forecast VaR for Australian and USA stock markets. We also used EVT to quantify ES in an unconditional model (which can also be used in a conditional model as VaR). Our results show that EVT can be quite useful in financial risk modelling and given the occurrence of extreme market events it can be used for efficient extreme market risk modelling.

7 References


Nystrom, K., & Skoglund, J. (2002). Univariate Extreme Value Theory, GARCH and Measures


