The Actuarial Profession
making financial sense of the future

NETWORK OF CONSULTING ACTUARIES

James Sharpe
Based on a presentation from the Extreme Events Working Party 2011

Extreme Value Theory
For a 1-in-200 event

January 2021

Contents

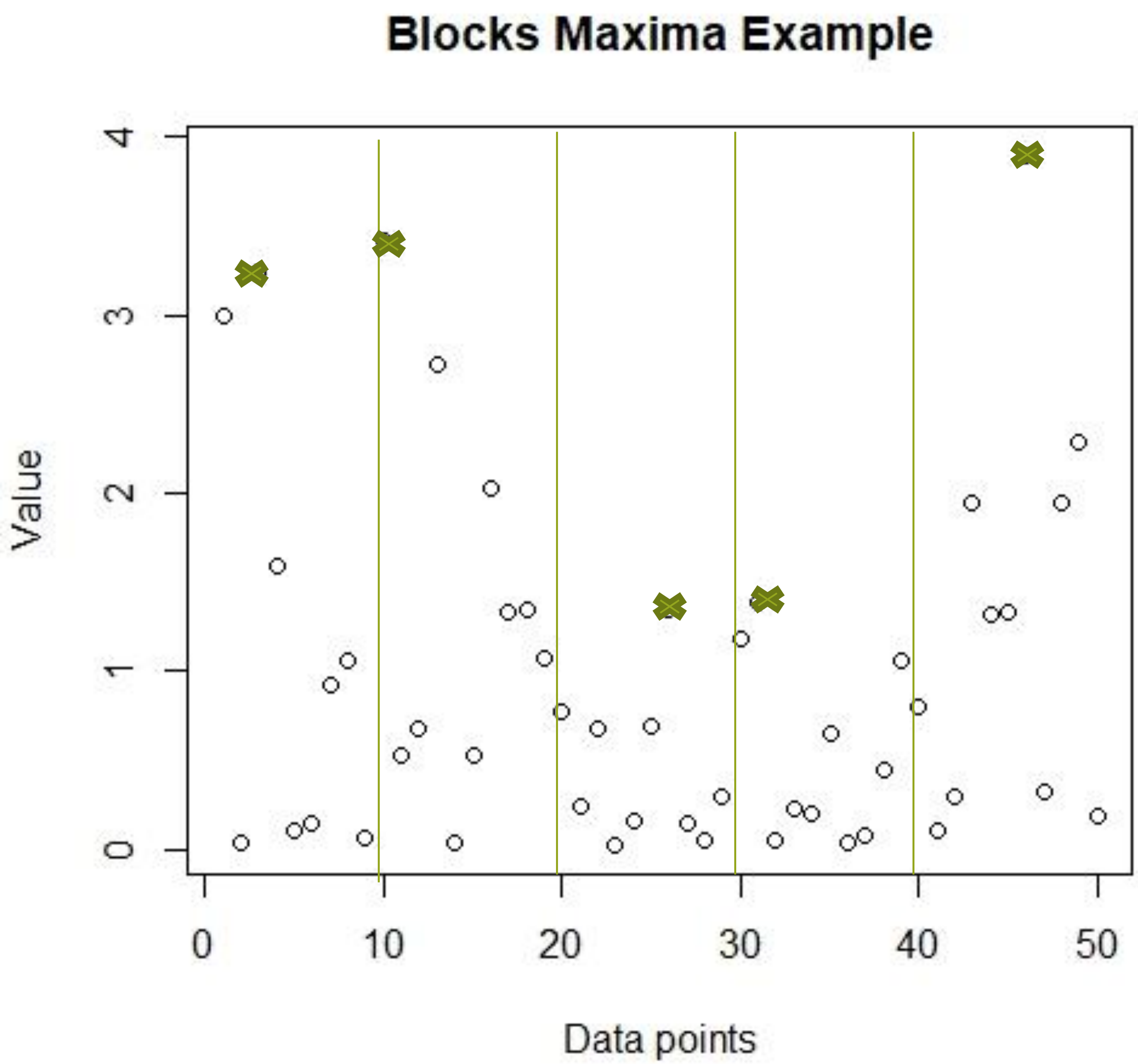
Introduction to EVT

- Extreme Value Theory(EVT) is a statistical approach that allows a practitioner to model the occurrence of extreme events with relatively small amounts of extreme data.
- A key difference between EVT and other statistical approaches is that, in EVT we fit a distribution to a subset of the available data, while in other statistical approaches, we fit a chosen distribution to the entire set of data.
- This allows us to focus on just the tail of the distribution and what this tells us about extreme percentiles. It contrasts with the approach of fitting probability distributions to all the data which fits to the body of the distribution and extrapolates out to extreme percentiles.

Introduction to EVT

- There are two EVT approaches:
  - Block Maxima Models (BMM): this is the traditional approach and it involves the following:
    - Grouping the data into samples/blocks;
    - Calculating the maximum observation in each block;
    - Fitting the Generalised Extreme Value (GEV) distribution to the maxima of the blocks
    - Estimating the risk measure we are interested in from the fitted GEV distribution.
  - The Peak over Threshold (POT): this is a more recent technique which involves the following:
    - Selecting a threshold that defines which observation are included in modelling;
    - Calculate the exceedances (this is the excess of the observations over the threshold);
    - Fit the Generalised Pareto Distribution (GPD) to the exceedances; and
    - Compute the measure of risk that is desired (e.g. 99.5^{th} percentile)
- The POT is typically preferred over the BMM because data is used more efficiently.

Block Maxima Models - Brief description



Data set is divided into blocks and only the maximum data point in each block is used.

The maximums in each block follow a Generalised Extreme Value distribution. This is regardless of the underlying probability distribution of the data.

Only the maxima data points are used (five in this example) and all the other data points are not used.

Block Maxima Models – Historic Example

- An historic example of the BMM approach is to consider the October 17th 1987 Black Monday S&P 500 one day fall of 20.4%
- Many models based on fitting the Normal distribution and extrapolating to the tails said this event was 1 in a “Trillion+” year event
- To apply the BMM approach we take a data set of daily returns from 1960 to October 16th 1987 and divide it into blocks of 1 year. This includes 28 annual maxima. The largest fall in the data set is just 6.7%.
- We fit a Generalised Extreme Value distribution to these 28 data points. This gives a 1 in 50 year event of a 24% fall, which is in excess of the 20.4% seen on October 17th 1987
- This shows the BMM gives far more plausible model. It avoid the problem of only preparing for the worst event seen in history (i.e. thinking the tallest mountain is the tallest one personally seen)

[Example from McNeil 1998 On extremes and crashes, Risk Magazine Jan 1999]

Peaks over Threshold - Introduction to GPD

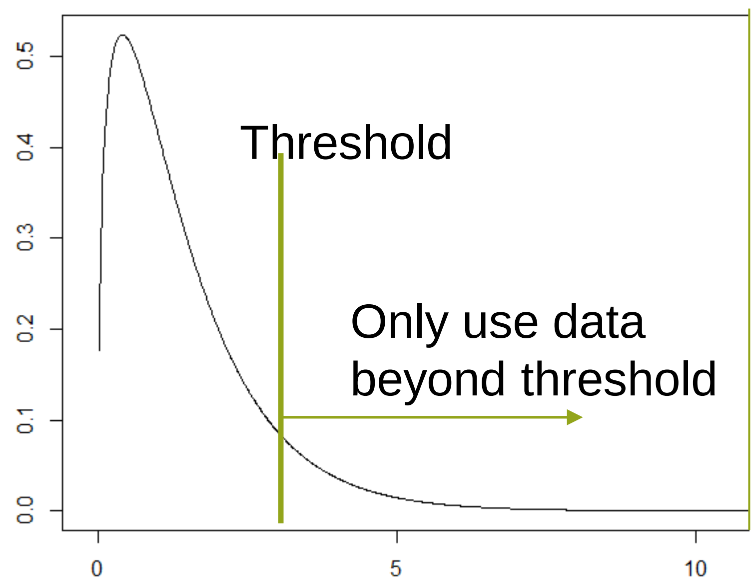
- The GPD distribution is used in the POT approach to model data above a certain threshold
- We use the Pickands Balkema de haan theorem to say that data in the tail of a data set follows a GPD distribution.
- Thus GPD is commonly used to model the tail of other distributions. It is defined by three parameters and its cumulative distribution function is:

$$F_{(\xi,\mu,\sigma)}(x) = \begin{cases} 1 - \left(1 + \frac{\xi(x-\mu)}{\sigma}\right)^{-1/\xi} & \text{for } \xi \neq 0, \\ 1 - \exp\left(-\frac{x-\mu}{\sigma}\right) & \text{for } \xi = 0. \end{cases}$$

μ is the location parameter;

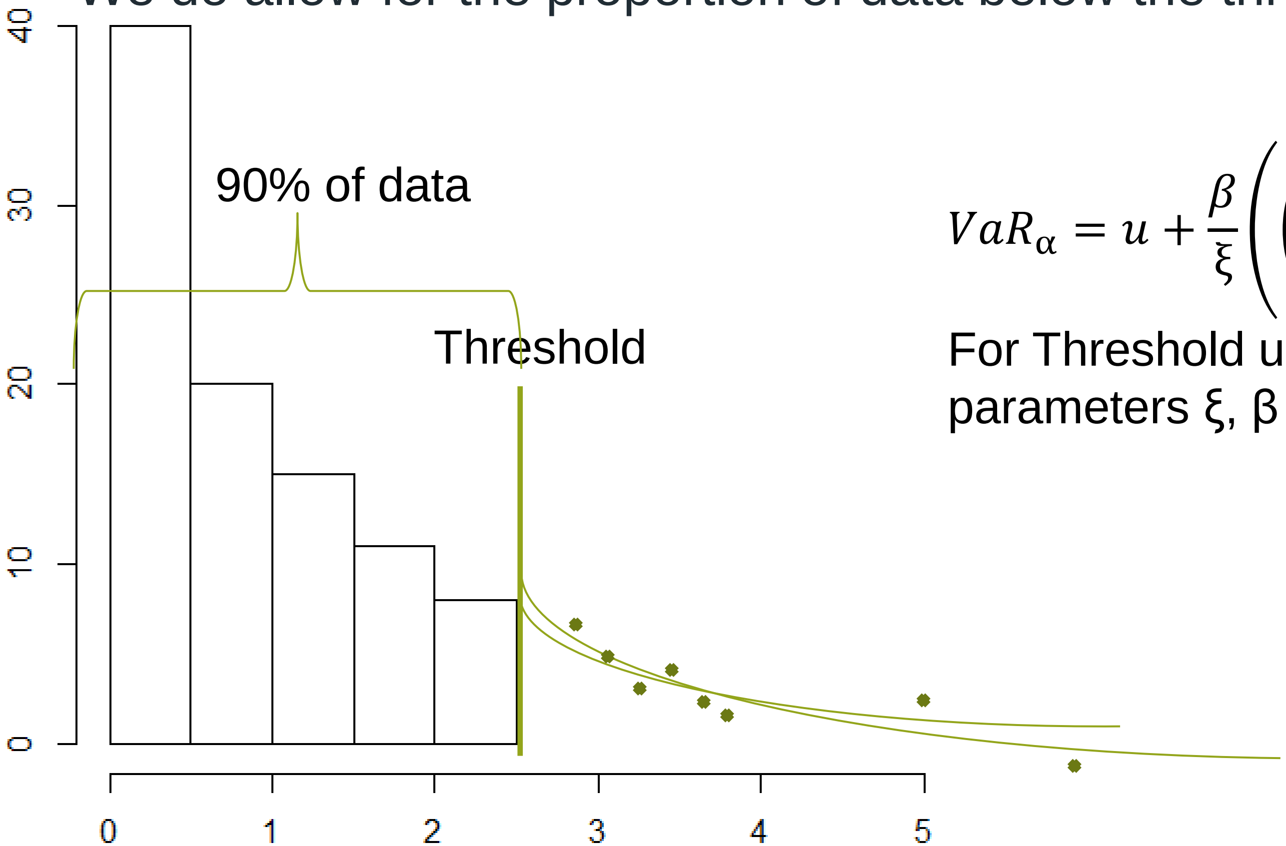
σ > 0 the scale parameter; and

ξ is the shape parameter.



Peaks over Threshold - Introduction to GPD

- The GPD tail is fit based on the values above the threshold; ignoring the exact values below the threshold
- We do allow for the proportion of data below the threshold in VaR estimates



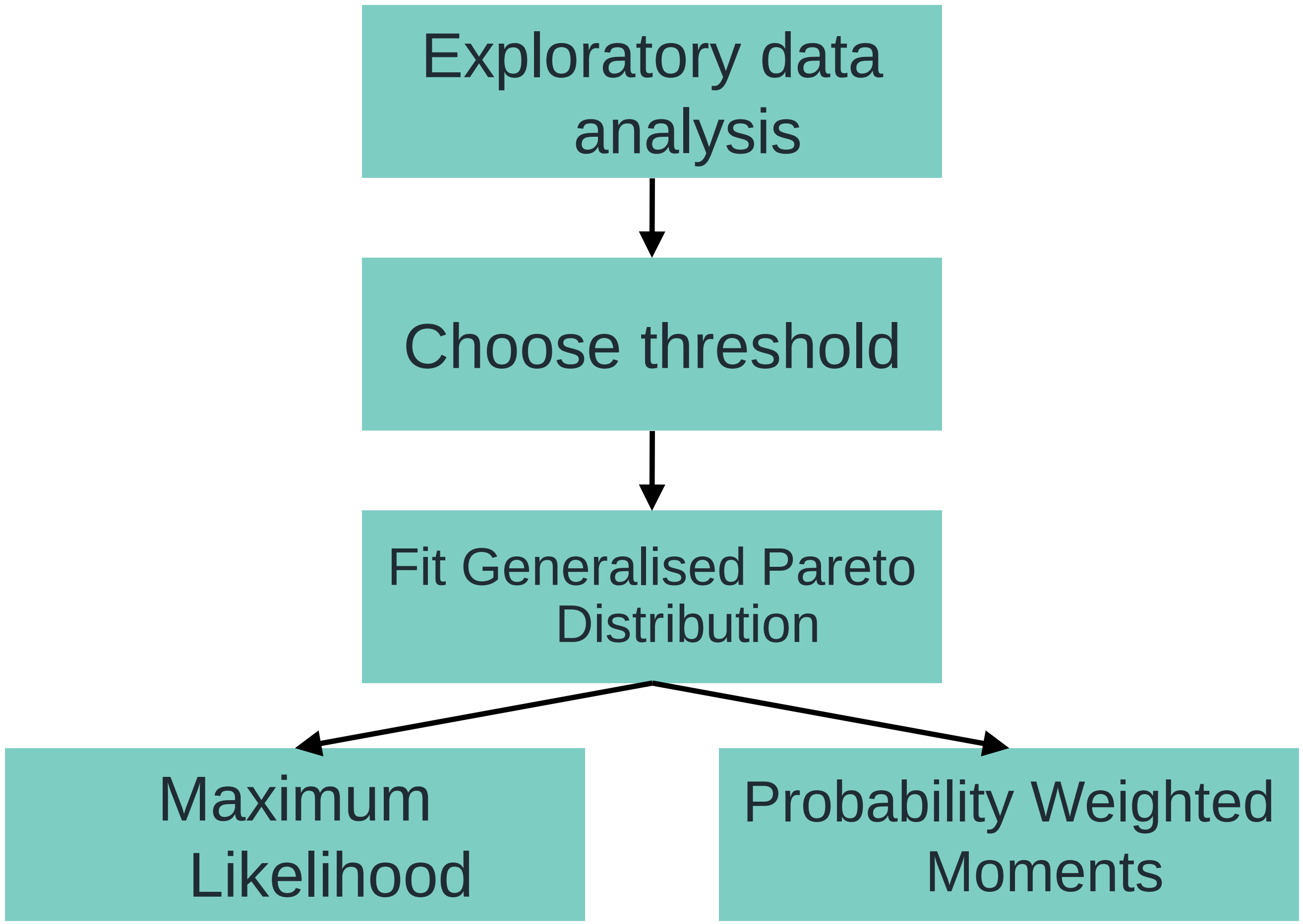
$$VaR_{\alpha} = u + \frac{\beta}{\xi}\left(\left(\frac{1-\alpha}{P(X>u)}\right)^{-\xi} - 1\right)$$

For Threshold u, percentile α, GPD parameters ξ, β

Peaks over Threshold - Introduction to GPD

- The following are examples of tail behaviours where the GPD can be used as an approximation:
  - Decreasing Exponential: such as the gamma, normal and log-normal. Achieved by setting the shape parameter of the GPD to zero;
  - Decreasing Polynomial: such as Student's t and Cauchy. Achieved by setting the shape parameter of the GPD to a positive number;
  - Finite: such as the beta distribution. Achieved by setting the shape parameter of the GPD to a negative number.
- As the name suggests, the GPD simplifies to other distributions under specific conditions. For example:
  - When $\xi = 0$, the tail simplifies to an exponential distribution;
  - When $\xi > 0$, the tail simplifies to an ordinary Pareto distribution (fat tailed); and
  - When $\xi < 0$, the tail simplifies to a Pareto type II distribution (thin tailed with finite right end point).

How to apply Peaks over Threshold to a data set



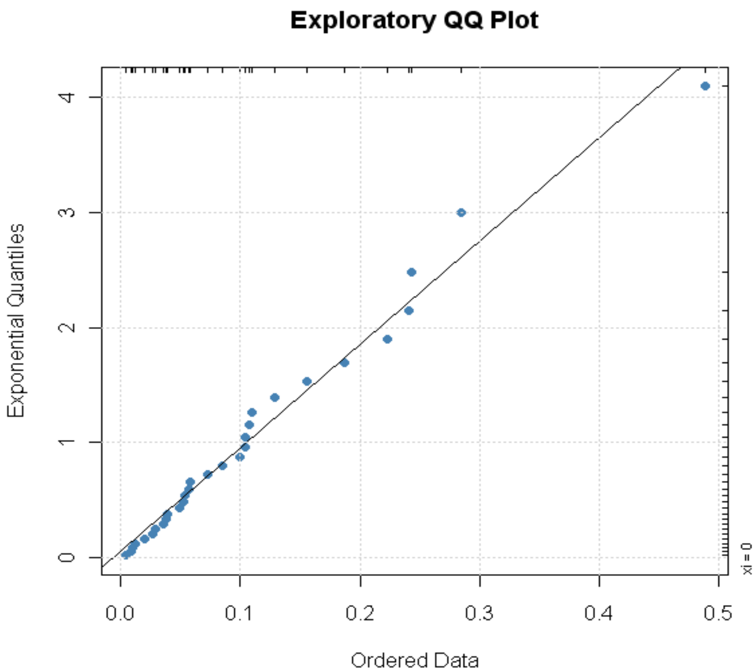
How to apply Peaks over Threshold to a data set

- The first step is to analyse the available data to assess whether the data, is fat tailed enough for us to apply EVT to it. The key tools used in this analysis are the Q-Q plots and the mean excess plots.
  - The Q-Q plot is a graphical means of assessing whether a sample data follows a given probability distribution.
  - The mean excess function express the mean of the excess over a threshold as a function of the threshold. The mean excess function for the GPD is a linear function of the threshold u.
- The next step in applying EVT is selecting an appropriate threshold.
  - A lower threshold reduces the variance of the estimates of the GPD model. However, a lower threshold can introduce bias in the data.
  - A higher threshold reduces bias but increase the volatility of the estimate of the GPD distribution.
  - The mean excess analysis may be used to select an optimum threshold.
  - An alternative approach is to fit the GPD to the data using different thresholds.
- Finally the parameters of the GPD can be estimated by the following approaches:
  - Maximum likelihood;
  - Probability-weighted moments.

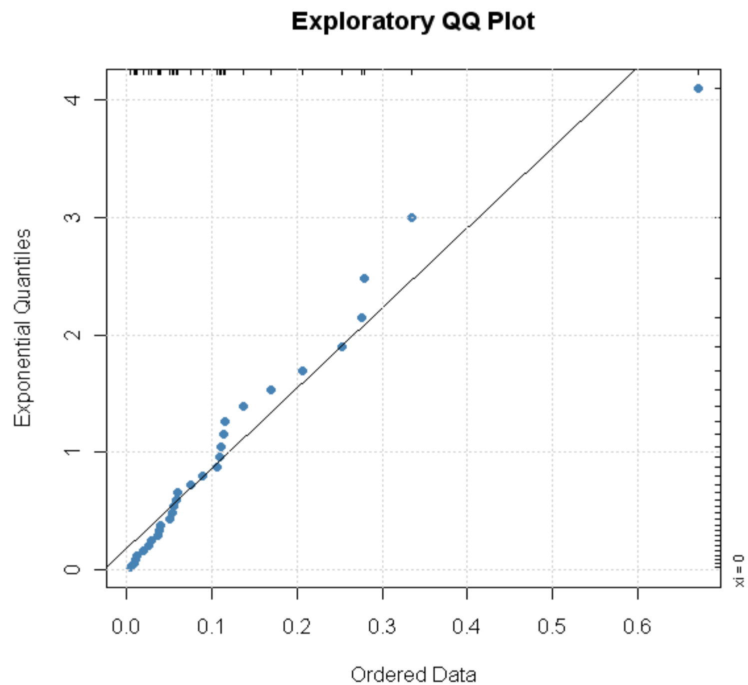
Case Study - Fitting GPD to Equity returns

- This section of the presentation looks at applying the Generalised Pareto distribution (GPD) to equity returns
- The data set used is the DMS data with Annual equity returns 1900-2008

Exploratory data analysis – QQ plots of negative annual returns vs. Exponential (GPD ξ=0) UK



QQ plot of UK simple returns



QQ plot of UK log returns

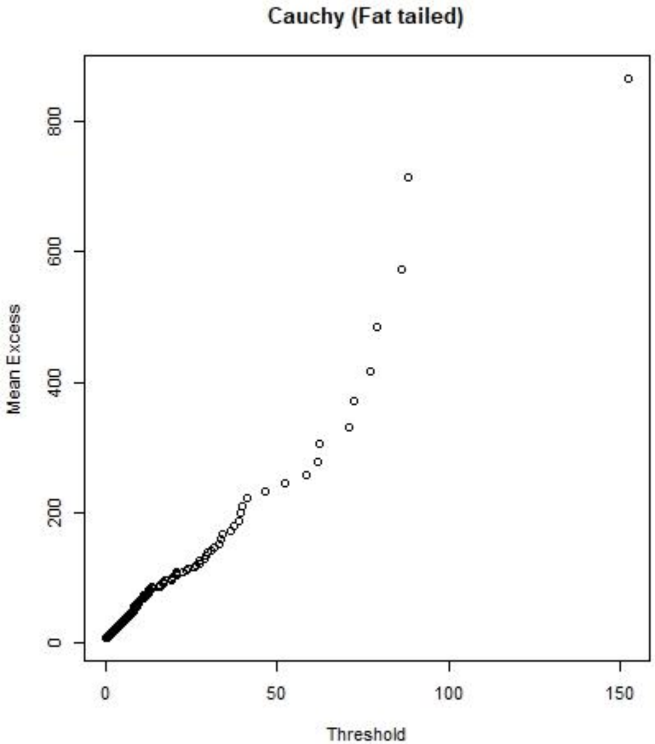
QQ plots - Interpretation

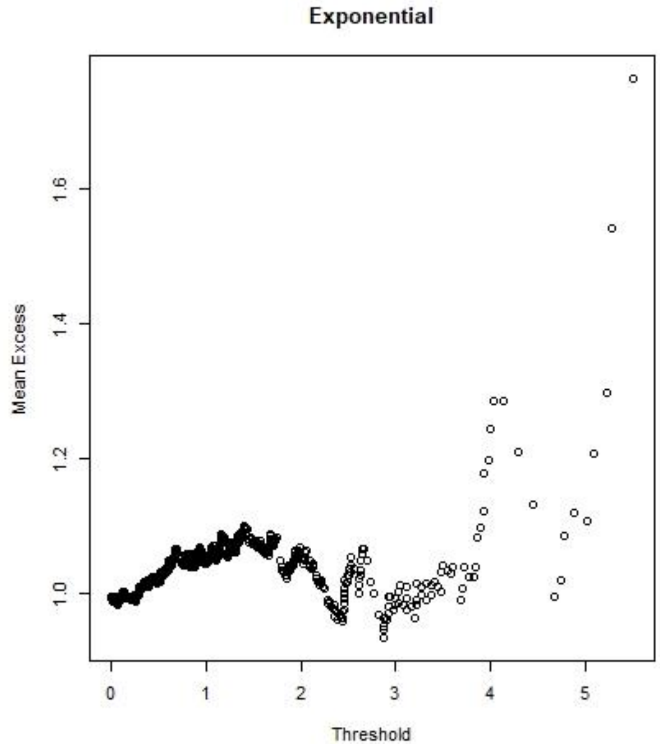
- QQ plot with a straight diagonal line of data points from the bottom left of the chart to the top right indicates an exponential distribution is a relatively good fit to the tail of this data
- A concave (i.e. starting bottom left and curving round in the top half of C shape to a horizontal line) shape to the QQ plot indicates fatter tail than the exponential distribution; so would suggest fitting a GPD with a $\xi > 0$
- A convex shape (i.e. starting bottom left and curving round in the right half of a U shape to a vertical line) to the QQ plot indicates a thinner tail than the exponential distribution with $\xi < 0$
- From the QQ plots, the exponential distributions appears a relatively good fit to the negative annual returns
- This indicates that annual returns in the UK from 1900 -2008 have not been particularly fat tailed

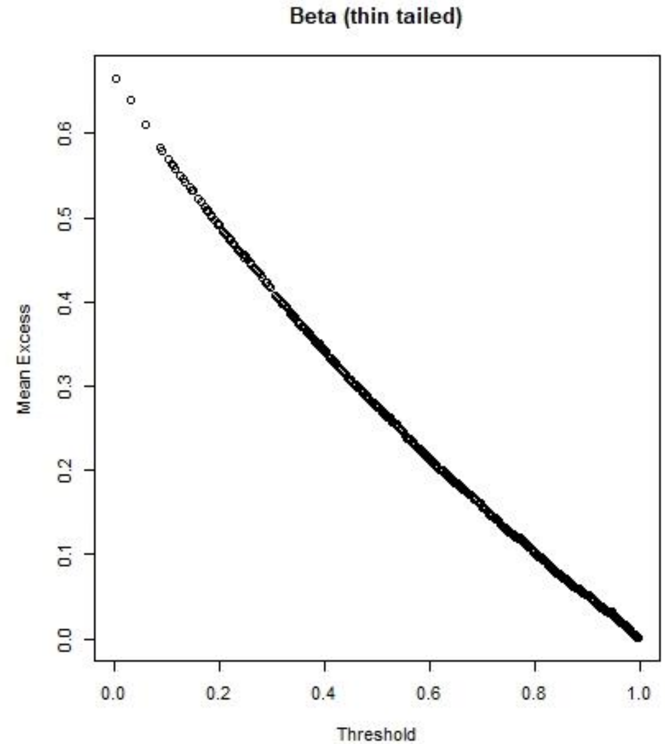
Exploratory Data Analysis - Mean Excess Plots

- The mean excess function is the mean of the values above a threshold. This can be plotted for varying thresholds as the mean excess plot.
- The mean excess plot is a tool used to aid the choice of threshold and also to determine the adequacy of the GPD model in practice.
- A characteristic of a fat tailed GPD type distribution with positive shape parameter is a straight line from bottom left to top right of the mean excess plot
- A mean excess plot with a downwards sloping line from top left to bottom right indicates thin tailed behaviour. A straight horizontal line indicates exponential type behaviour

Exploratory data analysis – Mean Excess plots of negative annual returns





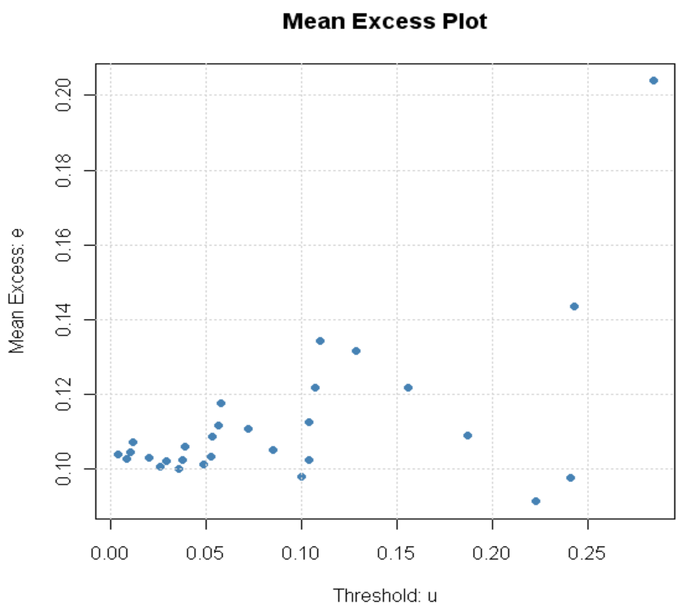


Mean excess plots are:

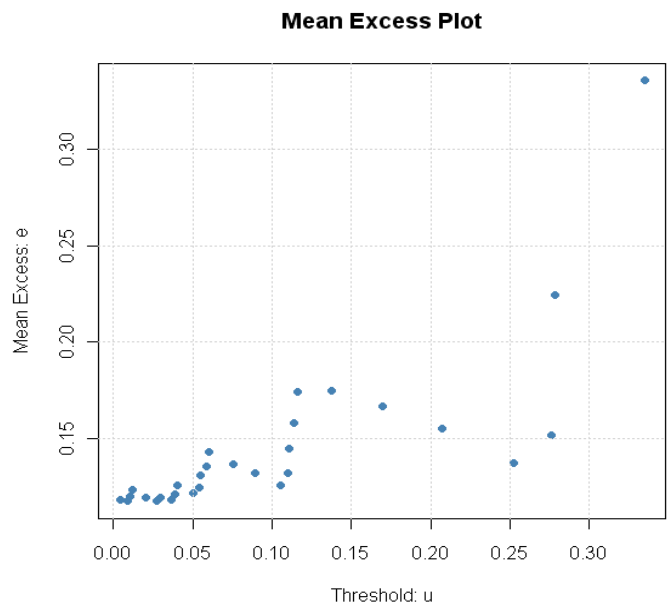
- Upwards sloping for the fat tailed Cauchy
- Broadly flat for the exponential and
- Downwards sloping for the thin tailed Beta distribution

This is useful for comparing to the plots from the equity data

Exploratory data analysis – Mean Excess plots of negative annual returns



Mean Excess UK simple returns



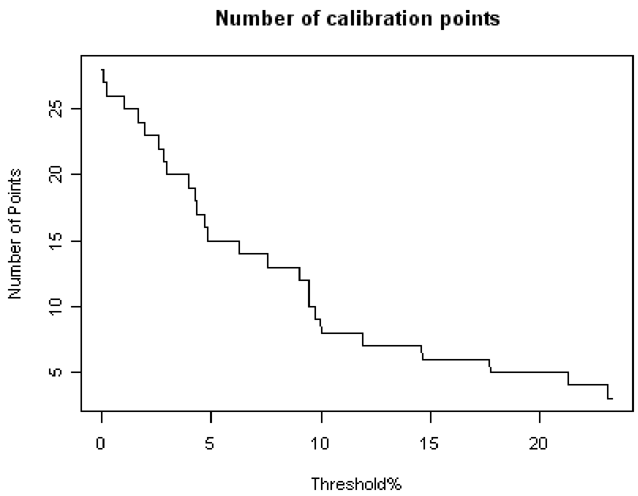
Mean Excess UK log returns

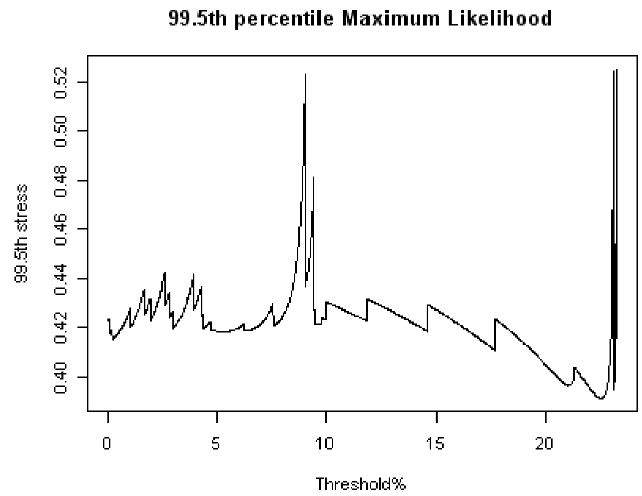
There is some significant up and down behaviour in the UK mean excess plots, with a broad upwards trend

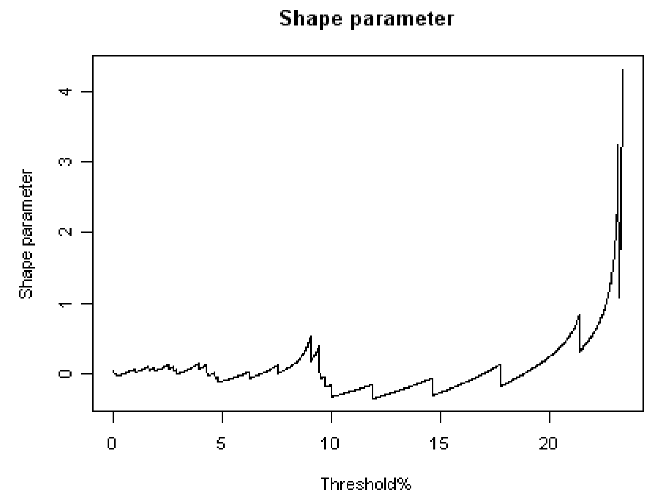
Generalised Pareto Distribution – Calibrated to UK annual returns

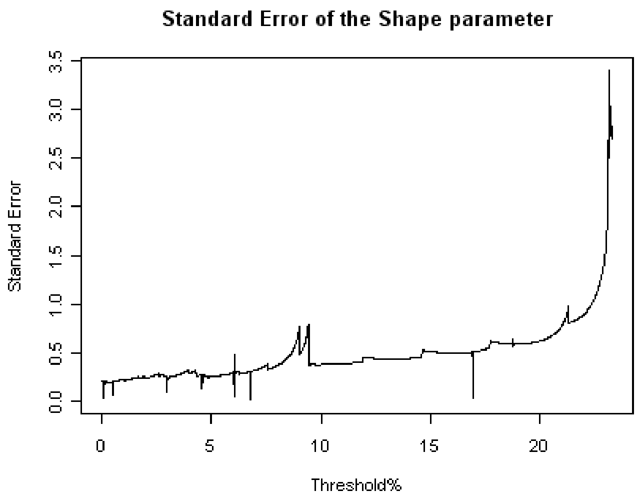
- The GPD is calibrated to the annual returns based on the points over a certain threshold
- Calibrations are made at each threshold
- Five plots are presented:
  1. The number of points above each threshold
  2. The 99.5th percentile for each calibration at each threshold
  3. The shape parameter ξ for each calibration at each threshold
  4. The standard error for the shape parameter for each calibration at each threshold
  5. The maximum log likelihood for each calibration at each threshold

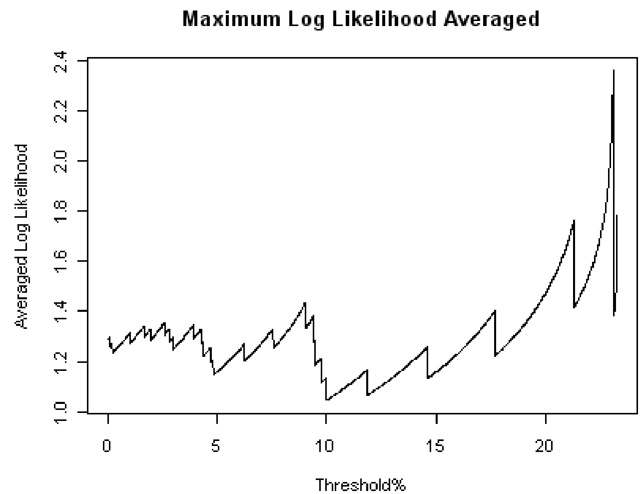
Generalised Pareto Distribution – Calibrated to UK annual returns – Maximum Likelihood Estimate







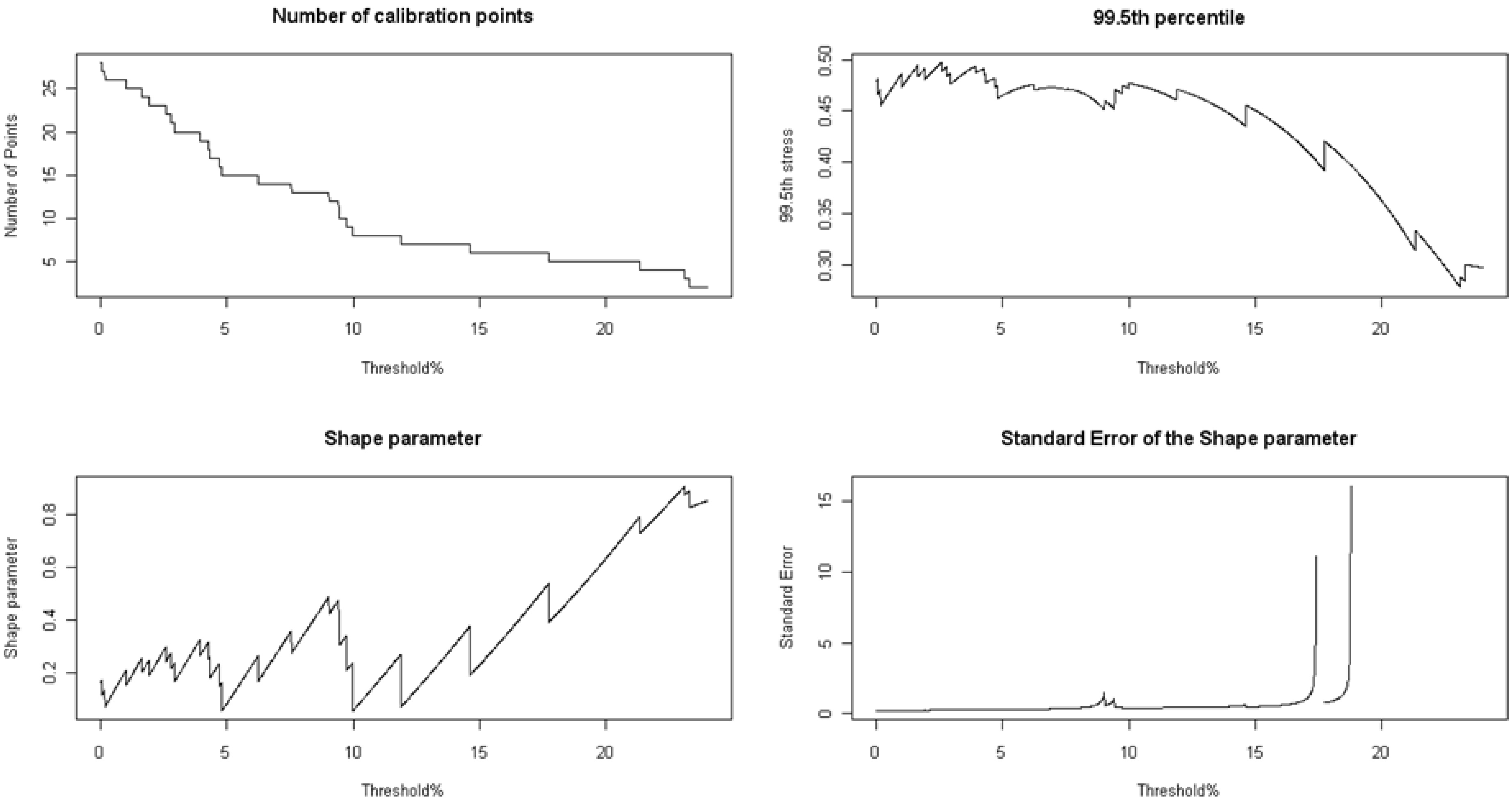




Generalised Pareto Distribution – Calibrated to UK annual returns – Maximum Likelihood Estimate

- Starting in the top left, the first plot shows how the number of points in the tail decreases as the calibration moves further into the tail
- The second plot shows the calibration of the 99.5th percentile at each of the thresholds. We can see how this varies as the calibration moves into the tail. There is a significant spike at the -10% threshold where a number of points are clustered. Otherwise the calibration of the 99.5th percentile is fairly stable with changes in threshold
- The top right plot shows the shape parameter for the calibration at each threshold. This stays around the 0 level for most of the thresholds, spiking upwards as the number of points in the calibration falls
- The bottom left plot shows the standard error of the shape parameter estimate. The standard error rises as the number of points in the calibration falls
- The final plot shows the maximum log likelihood, which may indicate better calibration for higher values

Generalised Pareto Distribution – Calibrated to UK annual returns – Probability Weighted Moments



Generalised Pareto Distribution – Calibrated to UK annual returns – Probability Weighted Moments

- Starting in the top left, the first plot shows how the number of points in the tail decreases as the calibration moves further into the tail
- The second plot shows the calibration of the 99.5th percentile at each of the thresholds. We can see how this falls from around 45% to around 30% as the threshold increases
- The bottom left plot shows the shape parameter for the calibration at each threshold. This is generally over 0 for most of the thresholds.
- The bottom right plot shows the standard error of the shape parameter estimate. The standard error rises as the number of points in the calibration falls

Confidence intervals around estimated 99.5th percentiles – UK annual simple returns

- Using three different thresholds with the maximum likelihood calibration method, confidence intervals have been estimated for the 99.5th percentile stress level

	Thresholds		
UK 99.5th percentiles	**-1%**	**-5%**	**-10%**
Lower CI	-29.2%	-28.6%	-27.2%
Best Estimate	-42.2%	-42.9%	-51.3%
Upper CI	-97.7%	-97.7%	-97.7%

- Confidence intervals are very wide reflecting significant uncertainties in calibration of the 99.5th percentile stress
- Confidence intervals created using profile likelihood method [McNeil et al ., 2005]

Uses of Extreme Value Theory for Actuaries

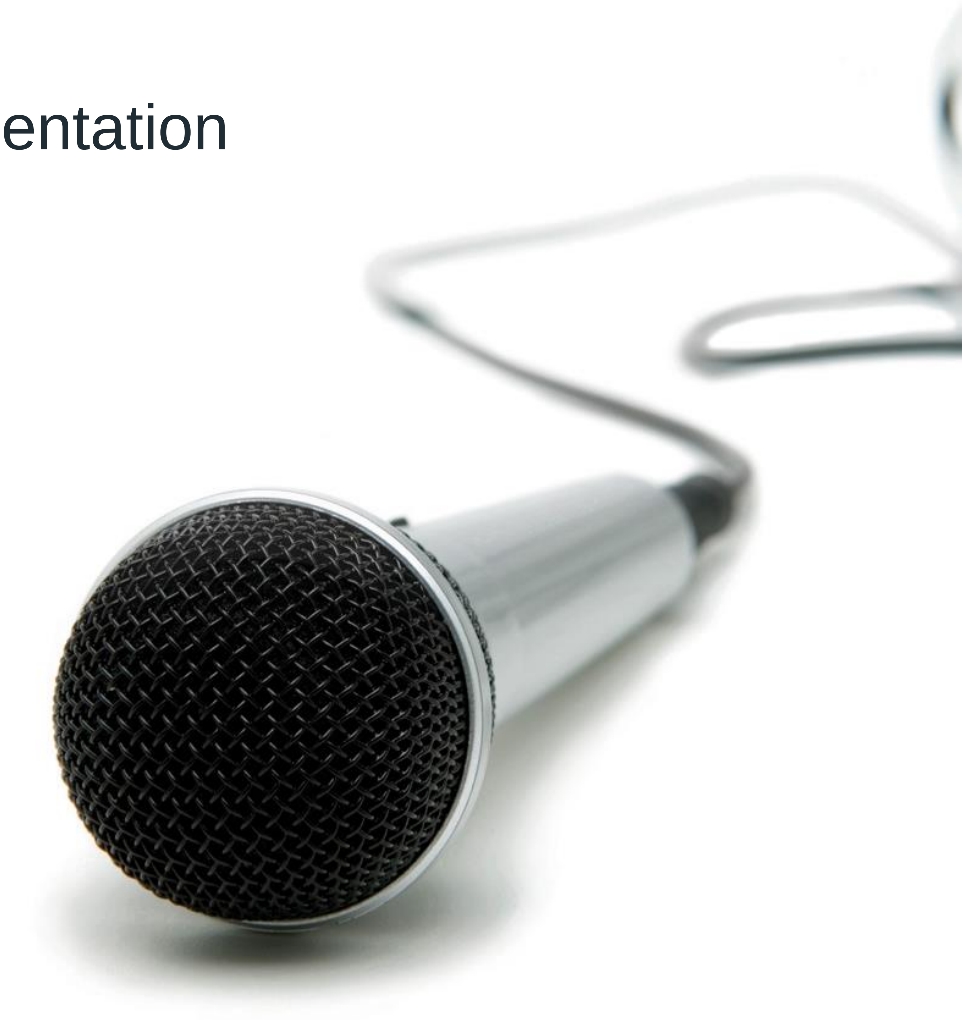
- EVT is widely used by Actuaries in General Insurance for pricing, reserving and capital
- For Solvency II capital modelling the full risk distribution is required across all percentiles. EVT is only calibrating in the tail so cannot be used for the full risk distribution
- However, it has a clear application in validating other probability distributions in the tails of the distribution. This is because when we fit probability distributions we are extrapolating from the body of the data into the tail. This extrapolation may not be accurate and EVT can be used to validate the tail fit

Software used for calibration

- Statistics package R has been used using the QRM library
- Other R libraries such as the fExtremes library were also investigated
- The work was partly re-created in excel as a check on the results
- The R libraries are freely available online and the code used to produce the results in this presentation and excel checking tool will be made available

Questions or comments?

The views expressed in this presentation are those of the presenter.



References

McNeil, A., Frey, R., and Embrechts, P. (2005).
Quantitative Risk Management: Concepts, Techniques and Tools.
Princeton University Press, Princeton.