Working Paper 2000-20 / Document de travail 2000-20

Steps in Applying Extreme Value Theory to Finance: A Review

by

Younes Bensalah

Bank of Canada Banque du Canada

ISSN 1192-5434

Printed in Canada on recycled paper

Bank of Canada Working Paper 2000-20

November 2000

Steps in Applying Extreme Value Theory to Finance: A Review

by

Younes Bensalah

Research and Risk Management Section
Financial Markets Department
Bank of Canada
Ottawa, Ontario, Canada K1A 0G9
beny@bankofcanada.ca

The views expressed in this paper are those of the author. No responsibility for them should be attributed to the Bank of Canada.

Contents

Acknowledgments. iv
Abstract/Résumé v

1. Introduction. 1

2. EVT: Some Theoretical Results 2

2.1 The GEV distribution (Fisher-Tippett, Gnedenko result) 2

2.2 The excess beyond a threshold (Picklands, Balkema-in Haan result) 3

3. Steps in Applying EVT. 4

3.1 Exploratory data analysis. 4

3.1.1 Q-Q plots 4

3.1.2 The mean excess function. 4

3.2 Sampling the maxima and some data problems. 5

3.3 The choice of the threshold 6

3.3.1 The graph of mean excess 6

3.3.2 Hill graph. 6

3.4 Parameters and quantiles estimation 7

3.4.1 Estimate the distribution of extremes. 7

3.4.2 Fitting excesses over a threshold. 9

4. Extreme VaR. 10

5. The Limitations of the EVT 11

6. Applying EVT Techniques to a Series of Exchange Rates of Canadian/U.S. Dollars 12

7. Conclusions. 13

Bibliography 15

Figures. 16

Acknowledgments

I am grateful to Shafiq Ebrahim, John Kiff, Samita Sareen, Uri Ron, and Peter Thurlow for helpful comments, suggestions, and documentation. Also, I would like to thank André Bernier for providing me with data.

Abstract

Extreme value theory (EVT) has been applied in fields such as hydrology and insurance. It is a tool used to consider probabilities associated with extreme and thus rare events. EVT is useful in modelling the impact of crashes or situations of extreme stress on investor portfolios. Contrary to value-at-risk approaches, EVT is used to model the behaviour of maxima or minima in a series (the tail of the distribution). However, implementation of EVT faces many challenges, including the scarcity of extreme data, determining whether the series is “fat-tailed,” choosing the threshold or beginning of the tail, and choosing the methods of estimating the parameters. This paper focuses on the univariate case; the approach is not easily extended to the multivariate case, because there is no concept of order in a multidimensional space and it is difficult to define the extremes in the multivariate case. Following a review of the theoretical literature, univariate EVT techniques are applied to a series of daily exchange rates of Canadian/U.S. dollars over a 5-year period (1995–2000).

JEL classification: C0, C4, C5, G1
Bank classification: Extreme value theory; Measure of risk; Risk management; Value at risk

Résumé

Appliquée dans des domaines aussi divers que l’hydrologie et l’assurance, la théorie des valeurs extrêmes permet d’estimer la probabilité associée à des événements extrêmes, donc rares. Aussi cet outil peut-il servir à modéliser l’incidence de krachs boursiers ou de tensions extrêmes sur les portefeuilles des investisseurs. Contrairement aux méthodes fondées sur la valeur exposée au risque, la théorie des valeurs extrêmes permet de reproduire le comportement des maxima ou des minima d’une série (les queues de la distribution). Son application soulève cependant de nombreuses difficultés : les occurrences extrêmes sont par définition rares, et il faut établir dès le départ si les queues de la distribution sont épaisses ainsi que choisir le seuil (ou début de la queue) et les méthodes d’estimation des paramètres. L’auteur se limite au cas univarié, car la théorie en question se prête mal à l’étude du cas multivarié en raison de l’absence de notion d’ordre dans un espace multidimensionnel et du problème que pose la définition des extrêmes lorsqu’il y a plusieurs variables. Après avoir effectué une revue de la littérature, l’auteur applique des techniques univariées s’inspirant de cette théorie à la série des taux de change quotidiens Canada/États-Unis pour les années 1995-2000.

Classification JEL : C0, C4, C5, G1
Classification de la Banque : Théorie des valeurs extrêmes; Mesure du risque; Gestion du risque; Valeur exposée au risque

1. Introduction

The study of extreme events, like the crash of October 1987 and the Asian crisis, is at the centre of investor interest. Until recently, the value-at-risk (VaR) approach was the standard for the risk-management industry. VaR measures the worst anticipated loss over a period for a given probability and under normal market conditions. It can also be said to measure the minimal anticipated loss over a period with a given probability and under exceptional market conditions (Longuin 1999). The former perspective focuses on the centre of the distribution, whereas the latter models the tail of the distribution.

The VaR approach (see Jorion 1996) has been the subject of several criticisms. The most significant is that the majority of the parametric methods use a normal distribution approximation. Using this approximation, the risk of the high quantiles is underestimated, especially for the fat-tailed series, which are common in financial data. Some studies have tried to solve this problem using more appropriate distributions (such as the Student-t or mixture of normals), but all the VaR methods focus on the central observations or, in other words, on returns under normal market conditions. Non-parametric methods make no assumptions concerning the nature of the empirical distribution, but they suffer from several problems. For example, they cannot be used to solve for out-of-sample quantiles; also, the problem of putting the same weight on all the observations remains unsolved.

Investors and risk managers have become more concerned with events occurring under extreme market conditions. This paper argues that extreme value theory (EVT) is a useful supplementary risk measure because it provides more appropriate distributions to fit extreme events. Unlike VaR methods, no assumptions are made about the nature of the original distribution of all the observations. Some EVT techniques can be used to solve for very high quantiles, which is very useful for predicting crashes and extreme-loss situations.

This paper is organized as follows. Section 2 introduces some theoretical results concerning the estimation of the asymptotic distribution of the extreme observations. Section 3 describes some data sampling problems, the choice of the threshold (or beginning of the tail), and parameter and quantile estimation. Section 4 estimates an extreme VaR and Section 5 describes the limitations of the theory. Section 6 provides a complete example of EVT techniques applied to a series of daily exchange rates of Canadian/U.S. dollars over a 5-year period (1995–2000). Section 7 concludes.

2. EVT: Some Theoretical Results

EVT has two significant results. First, the asymptotic distribution of a series of maxima (minima) is modelled and under certain conditions the distribution of the standardized maximum of the series is shown to converge to the Gumbel, Frechet, or Weibull distributions. A standard form of these three distributions is called the generalized extreme value (GEV) distribution. The second significant result concerns the distribution of excess over a given threshold, where one is interested in modelling the behaviour of the excess loss once a high threshold (loss) is reached. This result is used to estimate the very high quantiles (0.999 and higher). EVT shows that the limiting distribution is a generalized Pareto distribution (GPD).

2.1 The GEV distribution (Fisher-Tippett, Gnedenko result)

Let $\{X_1.....,X_n\}$ be a sequence of independent and identically distributed (iid) random variables. The maximum $Xn=Max(X_1,....,X_n)$ converges in law (weakly) to the following distribution:

$$H_{(\xi,\mu,\sigma)} = \begin{pmatrix} \exp\left(-[1+\xi(x-\mu)/\sigma]^{-1/\xi}\right) \\ \exp(-e^{-(x-\mu)/\sigma}) \end{pmatrix} if \left| \begin{matrix} \xi \neq 0 \\ \xi = 0 \end{matrix} \right.$$

$$while \qquad 1+\zeta(x-\mu)/\sigma > 0$$

The parameters μ and σ correspond, respectively, to a scalar and a tendency; the third parameter, ξ, called the tail index, indicates the thickness of the tail of the distribution. The larger the tail index, the thicker the tail. When the index is equal to zero, the distribution *H* corresponds to a Gumbel type. When the index is negative, it corresponds to a Weibull; when the index is positive, it corresponds to a Frechet distribution. The Frechet distribution corresponds to fat-tailed distributions and has been found to be the most appropriate for fat-tailed financial data. This result is very significant, since the asymptotic distribution of the maximum always belongs to one of these three distributions, whatever the original distribution. The asymptotic distribution of the maximum can be estimated without making any assumptions about the nature of the original distribution of the observations (unlike with parametric VaR methods), that distribution being generally unknown.

2.2 The excess beyond a threshold (Picklands, Balkema-in Haan result)

After estimating the maximum loss (in terms of VaR or another methodology), it would be interesting to consider the residual risk beyond this maximum. The second result of the EVT involves estimating the conditional distribution of the excess beyond a very high threshold.

Let *X* be a random variable with a distribution *F* and a threshold given x_F, for *U* fixes < x_F. F_u is the distribution of excesses of *X* over the threshold *U*

$$F_u(x) = P(X - u \le x | X > u), \quad x \ge 0$$

Once the threshold is estimated (as a result of a VaR calculation, for example), the conditional distribution F_u is approximated by a GPD.

We can write:

$$F_u(x) \approx G_{\xi,\beta(u)}(x), u \to \infty, \quad x \ge 0$$

where:

$$G_{\xi,\beta(u)}(x) = \begin{pmatrix} 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi} \\ 1 - e^{-x/\beta} \end{pmatrix} if \Big| \begin{matrix} \xi \ne 0 \\ \xi = 0 \end{matrix}$$

Distributions of the type *{H}* are used to model the behaviour of the maximum of a series. The distributions *{G}* of the second result model excess beyond a given threshold, where this threshold is supposed to be sufficiently large to satisfy the condition $u \to \infty$ (for a more technical discussion, see Castillo and Hadi 1997).

The application of EVT involves a number of challenges. The early stage of data analysis is very important in determining whether the series has the fat tail needed to apply the EVT results. Also, the parameter estimates of the limit distributions *H* and *G* depend on the number of extreme observations used. The choice of a threshold should be large enough to satisfy the conditions to permit its application (*u* tends towards infinity), while at the same time leaving sufficient observations for the estimation. Different methods of making this choice will be examined in Section 3.3. Finally, until now, it was assumed that the extreme observations are iid. The choice of the method for extracting maxima can be crucial in making this assumption viable. However, there are some extensions to the theory for estimating the various parameters for dependent observations; see Embrechts, Kluppelberg, and Mikosch (1997).

3. Steps in Applying EVT

3.1 Exploratory data analysis

The first step is to explore the data; see Bassi, Embrechts, and Kafetzaki (1997). Q-Q plots and the distribution of mean excess are used. They are described below.

3.1.1 Q-Q plots

Usually, one starts by studying a histogram of the data. In most of the VaR methods, the approximation by a normal distribution remains a basic assumption. However, most financial series are fat-tailed. The graph of the quantiles makes it possible to assess the goodness of fit of the series to the parametric model.

Let $X_1, \ldots, X_n$ be a succession of random variables iid, and $X_{n,n} < \ldots < X_{1,n}$ the order statistics, F_n being the empirical distribution. Note that $F_n(X_{k,n}) = (n-k+1)/n$ and F the estimated parametric distribution of the data.

The graph of quantiles (Q-Q plots) is defined by the set of the points,

$$\left\{X_{k,\, n}, F^{-1}\left(\frac{n-k+1}{n}\right), k = 1, \ldots, n\right\}$$

If the parametric model fits the data well, this graph must have a linear form. Thus, the graph makes it possible to compare various estimated models and choose the best. The more linear the Q-Q plots, the more appropriate the model in terms of goodness of fit. Also, if the original distribution of the data is more or less known, the Q-Q plots can help to detect outliers; see Embrechts, Kluppelberg, and Mikosch (1997).

Finally, this tool makes it possible to assess how well the selected model fits the tail of the empirical distribution. For example, if the series is approximated by a normal distribution and if the empirical data are fat-tailed, the graph will show a curve to the top at the right end or to the bottom at the left end.

3.1.2 The mean excess function

Definition: Let X be a random variable and given threshold x_F; then

$$e(u) = E(X-u|X>u), 0 \le u < x_F$$

e(.) is called the mean excess function. *e(u)* is called the mean excess over the threshold *u*. See Embrechts, Kluppelberg, and Mikosch (1997) for a discussion of the properties of this function.

If X follows an exponential distribution with a parameter λ, the function is equal to $e(u) = \lambda^{-1}$ for any $u > 0$. For the GPD,

$$e(u) = \frac{\beta + \xi u}{1 - \xi}, \beta + \xi u > 0$$

The mean excess function for a fat-tailed series is located between the constant mean excess function of an exponential distribution $e(u) = \lambda^{-1}$ and the GPD, which is linear and tends towards infinity for high thresholds as u tends towards infinity; see Embrechts, Kluppelberg, and Mikosch (1997).

A graphical test to establish the behaviour of the tail can be performed based on the form of the distribution of mean excess. Let $X_1 ..., X_n$ be iid with F_n the corresponding empirical distribution and $\Delta_n(u) = \{i, i = 1, \ldots, n, X_i > u\}$. Then,

$$e_n(u) = \frac{1}{card(\Delta_n(u))} \sum_{i \in \Delta_n(u)} (X_i - u), u \geq 0$$

where *card* refers to the number of points in the set $\Delta_n(u)$.

The set $\{(X_{k,n}, e_n(X_{k,n})), k = 1, ,,,, n\}$ forms the mean excess graph.

Fat-tailed distributions yield a function *e(u)* that tends towards infinity for high-threshold *u* (linear shape with positive slope).

3.2 Sampling the maxima and some data problems

Two approaches can be considered in building the series of maxima or minima. The first consists of dividing the series into non-overlapping blocks of the same length and choosing the maximum from each block. The assumption that the extreme observations are iid is viable in this case. Indeed, as financial data contain periods of high volatility followed by periods of low volatility (clustering), sampling the maxima using this technique reduces this phenomenon as soon as the size of the block is increased. However, the risk of losing extreme observations within the same block is still present, making the choice of the size of the block problematic.

The second approach consists of choosing a given threshold (high enough) and considering the extreme observations exceeding this threshold. The choice of the threshold 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 index of tail is more precise (less variance) but biased. On the other hand, choosing a high threshold reduces the bias but makes the estimator more volatile (fewer observations). The problem of dependent observations is also present. Some studies, such as by Resnick and Starica (1996), suggest using standardized observations to fit the various parameters to deal with this problem.

3.3 The choice of the threshold

3.3.1 The graph of mean excess

An immediate result of Section 3.2 enables us to provide a graphical tool to choose the threshold. The mean excess function for the GPD is linear (tends towards infinity). According to the result of Picklands, Balkema-in Haan, for a high threshold, the excess over a threshold for a given series converges to a GPD. It is possible to choose the threshold where an approximation by the GPD is reasonable by detecting an area with a linear shape on the graph.

Another graphical tool used to choose the threshold is the Hill graph.

3.3.2 Hill graph

Let $X_1 > ... > X_n$ be the ordered statistics of random variables iid. The Hill estimator of the tail index ξ using $k+1$ ordered statistics is defined by:

$$H_{k,n} = \frac{1}{k}\sum_{i=1}^{k} ln\left(\frac{Hi}{H_{k+1}}\right) = \hat{\xi}$$

The Hill graph is defined by the set of points

$$\{(k, H_{k,n}^{-1}), 1 \le k \le n-1\}$$

The threshold u is selected from this graph for the stable areas of the tail index. However, this choice is not always clear. In fact, this method applies well for a GPD or close to GPD type distribution. The Hill estimator is the maximum likelihood estimator for a GPD and since the extreme distribution converges to a GPD over a high threshold u (see the result of Picklands, Balkema-in Haan), its use is justified. For other distributions, some researchers suggest alternatives to the Hill graph to make this choice easier. Dress, de Haan, and Resnick (1998) propose using the graph defined by the set of points:

$$\{(\theta, H^{-1}_{[n^\theta], n}), 0 \le \theta < 1\}$$

In fact, they use a logarithmic scale for the axis of *k* *([*n^θ*]* integer less than or equal to n^θ*)*. This gives more space on the graph for the small values of *K*. Dress, de Haan, and Resnick (1998) found that the Hill graph is superior for the GPD type distribution, while the alternate graph adapts better for a large variety of distributions. They quantify the superiority in terms of “time occupation” around the true value of the tail index. The percentage *PERHILL* of “time” that the points *H* pass in a neighbourhood ε of the true value of index ξ is:

$$PERHILL(\varepsilon, n, l) = \frac{1}{l}\sum_{i=1}^{l} 1_{\{|H_{(i,n)} - \xi| \le \varepsilon\}}$$

The value of *PERHILL* quantifies stability around a selected value of the index, and can be used to choose among various values of the index (various stable areas of the graph).

3.4 Parameters and quantiles estimation

The parameters of the extreme distribution can be estimated under different assumptions. First, we can assume that the extreme observations follow exactly the GEV distribution. Second, and possibly more realistically, we can assume that the observations are roughly distributed like the GEV distribution. More specifically, the distribution of the observations belongs to the maximum domain of attraction (MDA) of H_ξ; for more details, see Embrechts, Kluppelberg, and Mikosch (1997). From there, the quantiles estimators can differ. Last, the parameters and quantiles are estimated for the distribution of excess over a threshold.

3.4.1 Estimate the distribution of extremes

Parametric methods. Assuming that the extreme observations follow exactly the GEV distribution, maximum likelihood estimation (MLE) can be used. Unfortunately, there is no closed form for the parameters, but numerical methods provide good estimates. For other estimation methods, see Embrechts, Kluppelberg, and Mikosch (1997). The p-quantile is defined as $\hat{x}_p = H_{\xi, \hat{\mu}, \hat{\sigma}}^{-1}(p)$. Then,

$$ln(p) = \left(-[1+\hat{\xi}(\hat{x}_p-\hat{\mu})/\hat{\sigma}]^{-1/\hat{\xi}}\right)$$

$$(-ln(p)) = [1+\hat{\xi}(\hat{x}_p-\hat{\mu})/\hat{\sigma}]^{-1/\hat{\xi}}$$

$$\hat{x}_p = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}[(-ln(p))^{-\xi}-1]$$

$$p = \exp\left(-[1+\hat{\xi}(\hat{x}_p-\hat{\mu})/\hat{\sigma}]^{-1/\hat{\xi}}\right)$$

MLE offers the advantage of simultaneous estimation of the three parameters, and it applies well to the series of maxima per block (see Section 3.2). Also, MLE seems to give good estimates for the case ξ>-1/2. As the majority of financial series have a positive tail index ξ>0, it offers a good tool for estimation in our field of interest.

Semi-parametric methods. The assumption that the extreme observations converge or are distributed exactly as H_ξ seems very strong. By relaxing this assumption, the observations are roughly distributed like the GEV distribution; then *F*, the distribution of the observations, belongs to the MDA of H_ξ. Thus, the parameters and quantiles estimation differs from that of the *Parametric methods* section.

Let $X_1 ... ,X_n$ be random variables that are iid, with a distribution $F \in MDA(H_\xi)$. This is equivalent to

$$\lim_{n\to\infty} \bar{F}(\sigma_n x + \mu_n) = -ln(H_\xi(x))$$

In this case (> *0*):

$$\bar{F}(x) = x^{-1/\xi}L(x), x > 0$$

with *L* a slow variation function, or more exactly,

$$\forall\lambda > 0, \lim_{x\to\infty}\frac{L(\lambda x)}{L(x)} = 1$$

Therefore, for a high *u* such as $u = \sigma_n x + \mu_n$, we have

$$n\bar{F}(u) \approx [1 + \hat{\xi}(u - \hat{\mu}_n)/\hat{\sigma}_n]^{-1/\hat{\xi}}$$

The p-quantile is defined as $\hat{x}_p = H^{-1}_{\xi, \hat{\mu}_n, \hat{\sigma}_n}(p)$; then,

$$\hat{x}_p = \hat{\mu}_n + \frac{\hat{\sigma}_n}{\hat{\xi}}[(n(1-p))^{-\xi} - 1]$$

Usually we are more interested in very high quantiles beyond the series or, in other words, out-of-sample estimation. For this purpose, let $u=X_k$, where u is a very high threshold, k/n its associated probability (probability from the empirical distribution of a series of length N and u is the K order statistic $X_1>.. >X_n$), and p is the associated probability of the p-quantile x_p; then,

$$\bar{F}(X_{k,n}) = X_{k,n}^{-1/\xi}L(X_{k,n}) = \frac{k}{n} \qquad (1)$$

$$\bar{F}(\hat{x}_p) = \hat{x}_p^{-1/\xi}L(\hat{x}_p) = 1 - p \qquad (2)$$

Dividing (2) by (1), and $x_p>X_k$, we have $L(\hat{x}_p)/L(X_{k,n}) \approx 1$; then,

$$\hat{x}_p = \left(\frac{n}{k}(1-p)\right)^{-\hat{\xi}} X_{k,n}$$

The tail index ξ is chosen from a stable area on the Hill graph.

3.4.2 Fitting excesses over a threshold

The GPD estimation involves two steps:

1. The choice of the threshold u. The mean excess graph can be used where u is chosen such that $e(x)$ is approximately linear for $x > u$ ($e(u)$ is linear for a GPD).

2. The parameter estimations for ξ and β can be done using MLE.

Once the distribution of excesses over a threshold is estimated, an approximation of the unknown original distribution (that generates the extreme observations) and an estimation of the p-quantile from it can be used to estimate the extreme VaR.

Thus, let u be a threshold, $X_1...,X_n$ the random variables exceeding this threshold (following a distribution $F \in MDA(H_\xi)$), and $Y_1...,Y_n$ the series of exceedances $(Y_i=X_i-u)$. The distribution of excesses beyond u is given by:

$$F_u(y) = P(X - u \le y | X > u) = P(Y \ge y | X > u), y \ge 0$$

and the distribution, *F*, of the extreme observations, X_i, is given by:

$$\bar{F}(u + y) = P(X \ge u + y) = P((X \ge u + y | X > u) \cdot P(X > u))$$

$$F(u + y) = P(X - u \ge y | X > u) \cdot P(X > u)$$

$$F(u + y) = \bar{F}_u(y) \cdot \bar{F}(u)$$

This result makes it possible to estimate the tail of the original distribution, by separately estimating *F* and F_u. According to the result of Picklands, Balkema-in Haan, and for a high threshold *u*:

$$(\hat{\bar{F}_u(y)}) \approx \bar{G}_{\hat{\xi}, \hat{\beta}(u)}(y)$$

F(u) can be estimated from the empirical distribution of the observations:

$$(\hat{F(u)}) = \frac{1}{n} \sum_{i=1}^{n} I_{\{X_i > u\}} = \frac{N_u}{n}$$

and

$$(\hat{\bar{F}(u + y)}) = \frac{N_u}{n}\left(1 + \hat{\xi}\frac{y}{\hat{\beta}}\right)^{-1/\hat{\xi}}$$

The estimation of the p-quantile for a given threshold, *u*, using this distribution is straightforward:

$$\hat{x}_p = u + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\frac{n}{N_u}(1 - p)\right)^{-\hat{\xi}} - 1\right]$$

4. Extreme VaR

By definition, VaR is the p-quantile of the distribution of the log change in price. EVT makes it possible to model the empirical distribution of the extreme observations. Extreme VaR is defined as the p-quantile estimated from the extreme distribution. Various estimators are available, depending on the estimation method and the assumptions (as discussed in Section 3.4).

VaR is estimated assuming that the extreme observations follow exactly the GEV distribution (see *Parametric methods*, in Section 3.4.1):

$$VaR_{extreme} = \hat{\mu} + \frac{\hat{\sigma}}{\hat{\xi}}[(-ln(p))^{-\xi} - 1]$$

VaR is estimated assuming that the observations follow approximately the GEV distribution (see *Semi-parametric methods*, in Section 3.4.1).

VaR in-sample:

$$VaR_{extreme(in-sample)} = \hat{\mu}_n + \frac{\hat{\sigma}_n}{\hat{\xi}}[(n(1-p))^{-\xi} - 1]$$

and VaR out-of-sample:

$$VaR_{extreme(out-of-sample)} = \left(\frac{n}{k}(1-p)\right)^{-\hat{\xi}} X_{k,n}$$

The approximation of excesses over a threshold by a GPD leads to the following estimator:

$$VaR_{extremeGPD} = u + \frac{\hat{\beta}}{\hat{\xi}}\left[\left(\frac{n}{N_u}(1-p)\right)^{-\hat{\xi}} - 1\right]$$

Usually, the estimation using a GPD is repeated several times to have a graph of quantiles. Various results might be presented: quantiles for a stable area of the graph and a quantile corresponding to the pessimistic version (peak). This method is called peak over threshold (POT); McNeil and Saladin (1997) discuss this method in detail.

5. The Limitations of the EVT

As stated earlier, the approaches described here relate to aggregate positions. Most of the articles in the literature use the same approach. The application of the EVT results in a multivariate case faces a basic problem. There is no standard definition of order in a vectorial space with dimensions greater than 1 and thus it is difficult to define the extreme observations for n-dimension vectors *(n>1)*; see Embrechts, de Haan, and Huang (1999).

To solve this problem, Longuin (1999) proposes estimation of the extreme marginal distribution for each asset (use the maxima for the short positions w_i and the minima for the long positions), solving for the extreme p-quantiles VaR_i, computing the correlations r_{ij} between the series of maxima and minima, and calculating the extreme VaR of a portfolio of N assets:

$$VaRextreme = \sqrt{\sum_{i=1}^{n} \sum_{j=1}^{n} \rho_{ij} \cdot w_i \cdot w_j \cdot VaR_i \cdot VaR_j}$$

Unfortunately, the joint distribution of the extreme marginal distributions is not necessarily the distribution of the extremes for the aggregate position. In other words, extreme movements of the log change in prices for the different assets do not necessarily result in extreme movements for the whole portfolio. This will depend on the composition of the portfolio (position on each instrument) and on the relations (dependencies or correlations) between the various assets. For an interesting discussion of the potential and limitations of EVT, see Embrechts (2000).

6. Applying EVT Techniques to a Series of Exchange Rates of Canadian/U.S. Dollars

In this section, EVT techniques are applied to a series of daily exchange rates of Canadian/U.S. dollars over a 5-year period (1995–2000) using EVIS software (www.math.ethz.ch/~mcneil).

Table 1 lists VaR results for different distributions. The VaR listed is the maximum loss that can not be exceeded over a one-day horizon with a given confidence level (probability). A dollar VaR is given for US$10 billion. The VaR is in Canadian dollars.

Table 1: VaR in Can$ million assuming different distributions

Confidence level	VaR_Normal[a]	VaR_HS[b]	VaR_GEV[c]	VaR_GPD[d]
95% (in-sample)	22	20	72	NA
97.5% (in-sample)	26	25	89	NA
99% (out-of-sample)	31	37	116	70
99.9% (out-of-sample)	42	61	NA	136
99.99% (out-of-sample)	50	70	NA	257
99.999% (out-of-sample)	584	71	NA	479

a. VaR assuming a normal distribution.
b. VaR assuming an empirical distribution (historical simulation).
c. VaR assuming a GEV distribution: the maxima are sampled from non-overlapping blocks of 90 observations. The GEV distribution was estimated using MLE; the parameters are $\mu(0.002865304)$, $\sigma(0.000961009)$, and $\xi(0.274717)$.
d. VaR using the GPD. Usually, for very high confidence levels, the approximation by the GPD is used. The GPD was estimated using 90 observations over the threshold $u(0.001836343)$. The GPD parameters are $\xi(0.2643934)$ and $\beta(0.00068409)$.

Figures 1 to 13 illustrate the graphical techniques and describe the goodness of fit of the different models.

7. Conclusions

EVT can be used to supplement VaR methods, which have become a standard for measuring market risk. VaR methods present two problems. First, most VaR methods use the normal distribution approximation, which underestimates the risk of the high quantiles because of the fat-tail phenomenon. Some studies have tried to solve this problem by using more appropriate distributions (like the Student-t). Second, VaR methods use all the data of the series for the estimation. However, because most of the observations are central, the estimated distribution tends to fit central observations, while falling short on fitting the extreme observations because of their scarcity. However, it is these extreme observations that are of greater interest for investors and risk managers.

EVT techniques make it possible to concentrate on the behaviour of these extreme observations. Also, the loss over a very large threshold can be estimated. While these results apply well to the univariate case, the multivariate one seems to define the limits of this theory. Indeed, the joint distribution of the marginal extreme distributions is not necessarily an extreme distribution. However, the joint distribution can be used to estimate the probability associated with an extreme scenario, but incorporating this information in the market risk framework remains an open question (see Berkowitz 1999). The majority of the results were presented assuming that the observations are independent and identically distributed. The sampling of maxima can make this assumption viable, but it is also possible to estimate an index of dependency to incorporate it in the calculation of the extreme VaR.

Asset allocation is usually conducted by using the variance (Markowitz) or VaR (parametric normal) as measures of risk, implicitly giving the same weights to the positive and negative returns (see Huisman, Koedijk, and Pownall 1999). Some studies try to concentrate on the downside risk by using the variance of negative returns as a measure of risk, but focusing on the central observations in their model. The allocation of capital using an extreme VaR as a risk measure will lead to a more conservative allocation, because the resulting portfolio is designed to hedge against worst-case market conditions.

Bibliography

Bassi, F., P. Embrechts, and M. Kafetzaki. 1997. "A Survival Kit on Quantile Estimation." ETH preprint (www.math.ethz.ch/~embrechts).

Berkowitz, J. 1999. *A Coherent Framework for Stress-Testing.* Federal Reserve Board.

Castillo, E. and S.A. Hadi. 1997. "Fitting the Generalized Pareto Distribution to Data." *Journal of the American Statistical Association* 92(440): 1609–20.

de Vries, G.C. and J. Danielsson. 1997. "Value at Risk and Extreme Returns." Timbergen Institute, Erasmus University, Discussion paper.

Dress, H., L. de Haan, and S. Resnick. 1998. *How to Make a Hill Plot.* Timbergen Institute, Erasmus University, Rotterdam.

Embrechts, P. 2000. "Extreme Value Theory: Potential and Limitations as an Integrated Risk Management Tool." ETH preprint (www.math.ethz.ch/~embrechts).

Embrechts, P., L. de Haan, and X. Huang. 1999. "Modelling Multivariate Extremes." ETH preprint (www.math.ethz.ch/~embrechts).

Embrechts, P., S. Resnick, and G. Samorodnitsky. 1998. "Extreme Value Theory as a Risk Management Tool." ETH preprint (www.math.ethz.ch/~embrechts).

Embrechts, P., C. Kluppelberg, and T. Mikosch. 1997. *Modelling Extremal Events for Insurance and Finance.* Berlin: Springer.

Huisman, R., G.K. Koedijk, and A.J.R. Pownall. 1999. *Asset Allocation in a Value at Risk Framework.* Erasmus University, Rotterdam, Faculty of Business Administration.

Jorion, P. 1996. *Value at Risk—The New Benchmark for Controlling Derivatives Risk*. Chicago: Irwin.

Kluppelberg, C. and A. May. 1998. "The Dependence Function for Bivariate Extreme Value Distributions—A Systematic Approach." Center of Mathematical Sciences, Munich University of Technology (http://www-m4.mathematik.tu-muenchen.de/m4).

Longuin, M.F. 1999. "From Value at Risk to Stress Testing: The Extreme Value Approach." Center for Economic Policy Research, Discussion Paper No. 2161.

McNeil, A.J. 1999. "Extreme Value Theory for Risk Managers." *Risk* special volume, to be published. ETH preprint (www.math.ethz.ch/~mcneil/pub_list.html).

McNeil, A.J. and T. Saladin. 1997. "The Peak Over Thresholds Method for Estimating High Quantiles of Loss Distributions." ETH preprint (www.math.ethz.ch/~mcneil/pub_list.html).

Resnick, S. and C. Starica. 1996. "Tail Index for Dependent Data." School of ORIE, Cornell University, Discussion paper.

Smith, L.R. 1987. "Estimating Tails of Probability Distributions." *The Annals of Statistics* 15(3): 1174–1207.

Figure 1: Histogram of the tail of the empirical distribution.

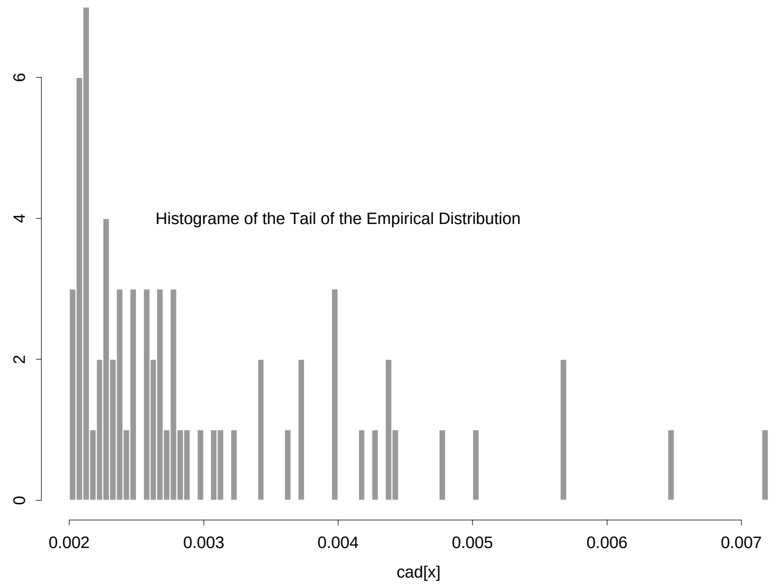


Figure 2: The Hill plot; the tail index is stable around values from 3 to 4, which is typical for the financial series.

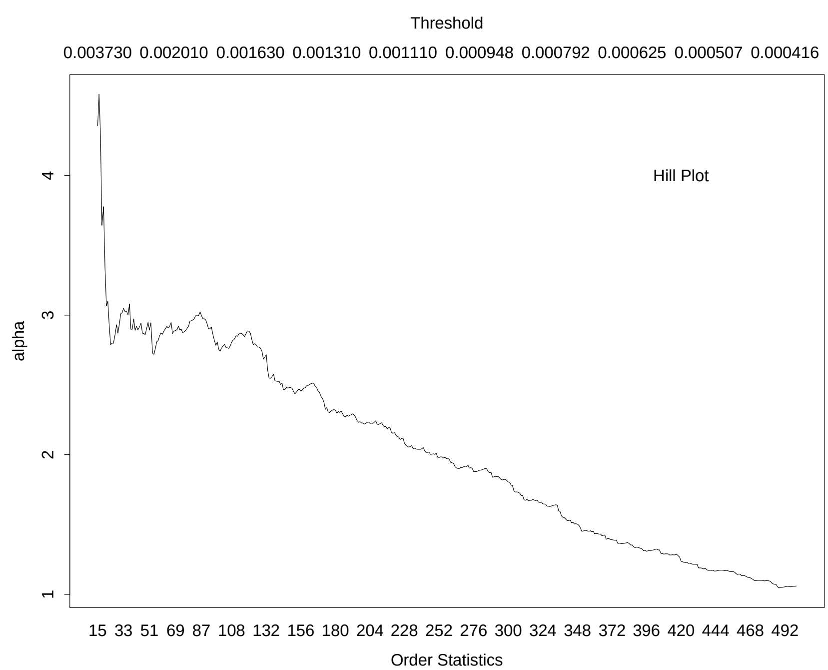


Figure 3: The average function of excess: for an adequate graph, certain extreme observations at the end of the tail were omitted (the value of the function for these points is the average of all the extreme observations). Note the linear shape of the graph around certain thresholds (0.0020, 0.0026), and that the function tends to infinity like a GPD.

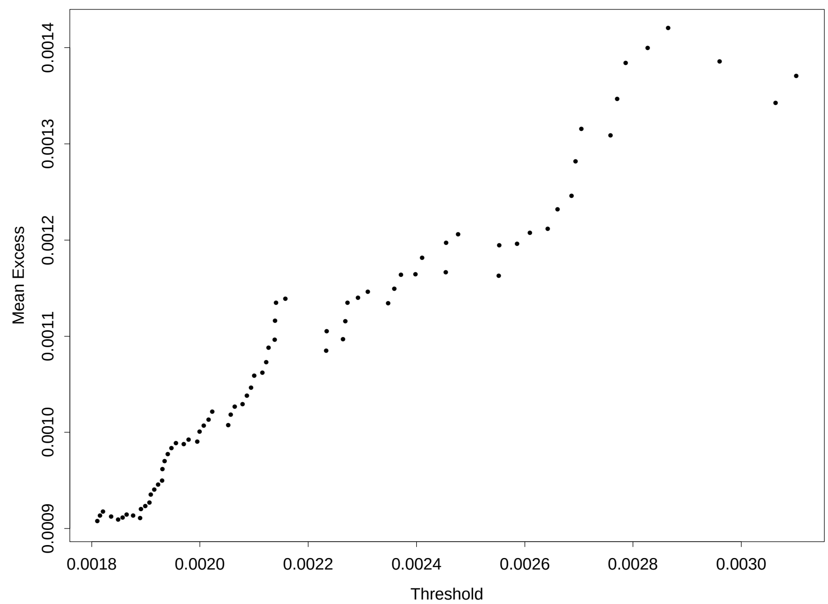


Figure 4: Q-Q plot for EVDistribution, maxima from blocks of 28 observations.

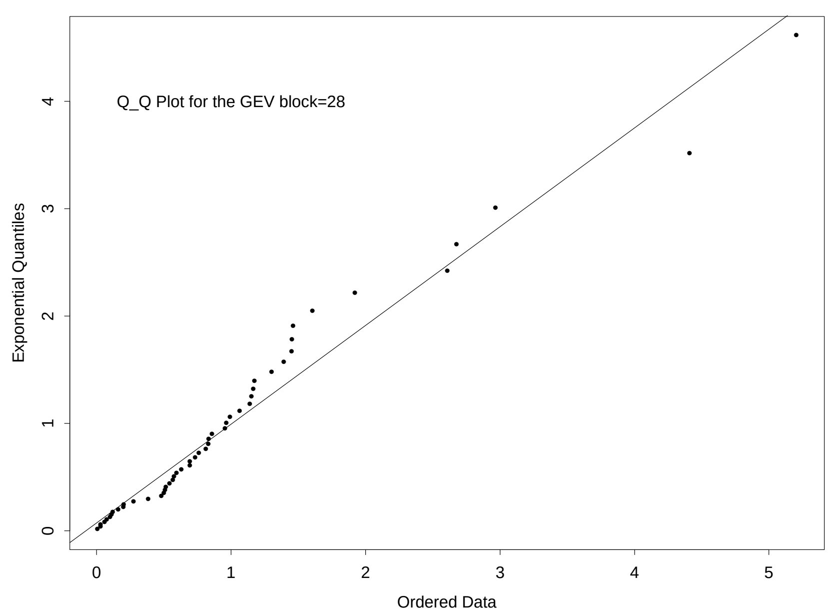


Figure 5: Q-Q plot for EVDistribution, maxima from blocks of 60 observations.

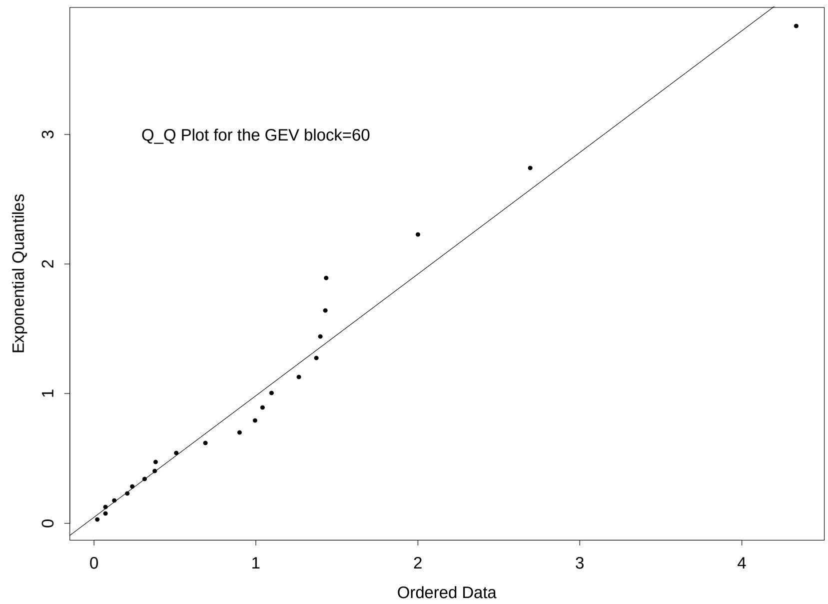


Figure 6: Q-Q plot for EVDistribution, maxima from blocks of 90 observations.

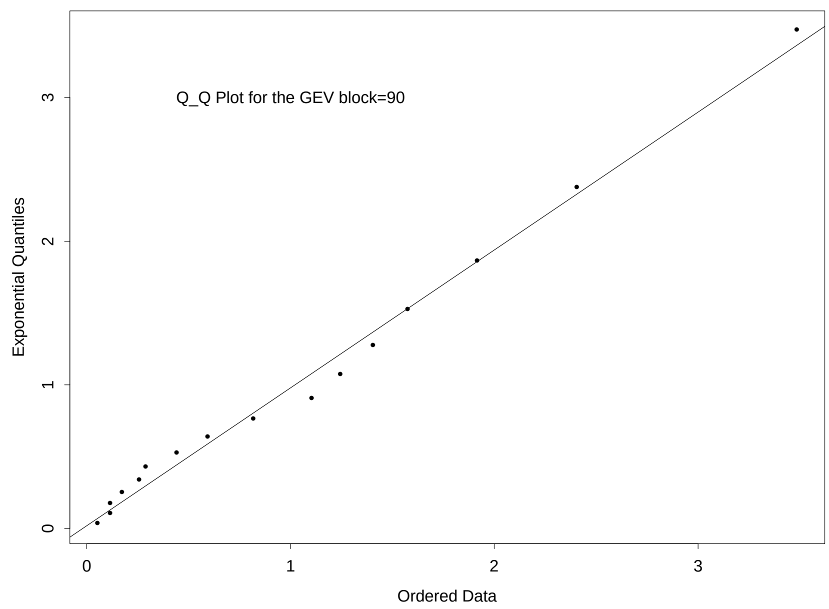


Figure 7: To estimate the GPD, start with the distribution of excesses. The shape of the graph corresponds exactly to the shape of a GPD (see Embrechts, Kluppelberg, and Mikosch 1997, p. 164).

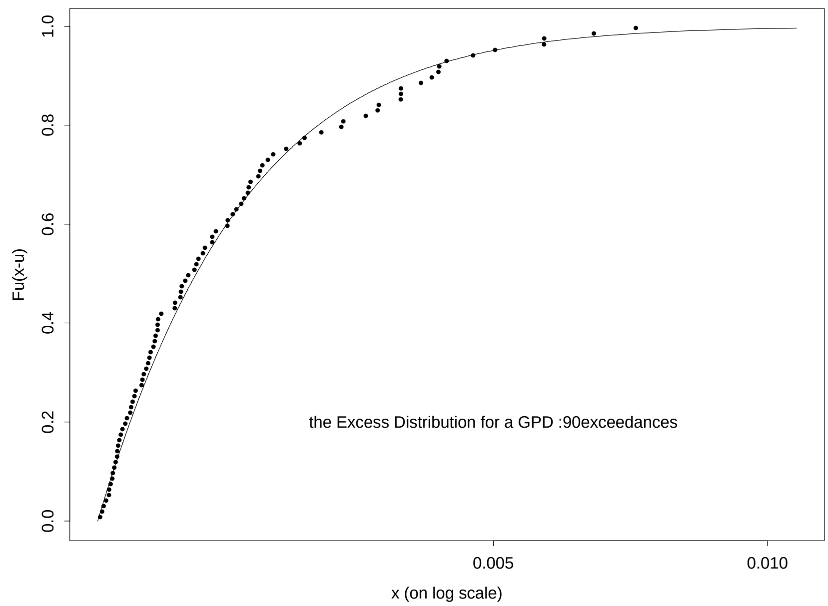


Figure 8: Q-Q plot GPD with 90 exceedances.

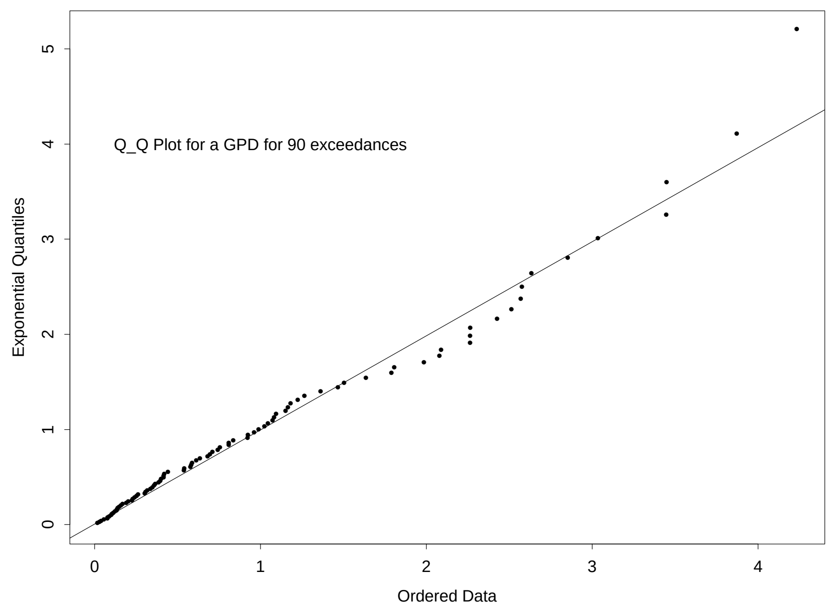


Figure 9: Tail of the underlying distribution using a GPD.

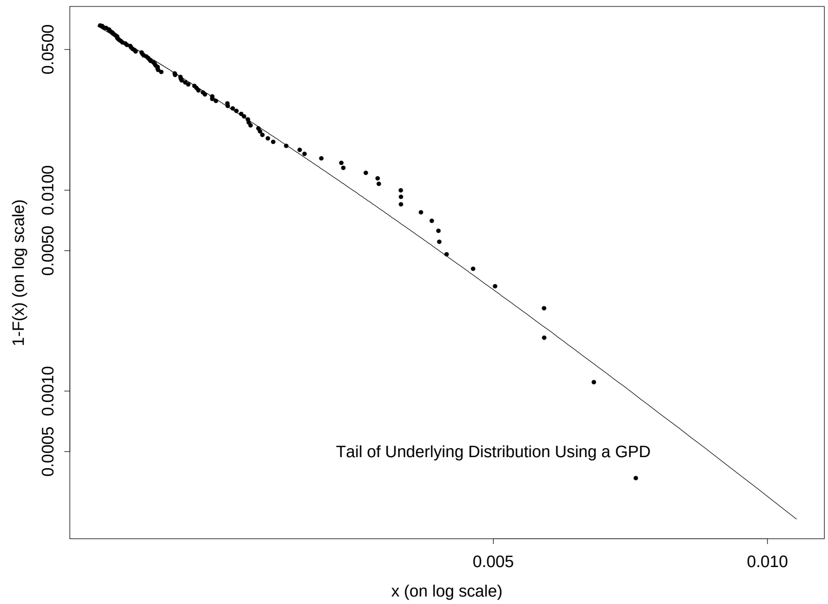


Figure 10: Q-Q plot from a normal distribution. While comparing this plot with a Q-Q plot from a normal distribution, one must remember that the empirical distribution presents a fat-tail.

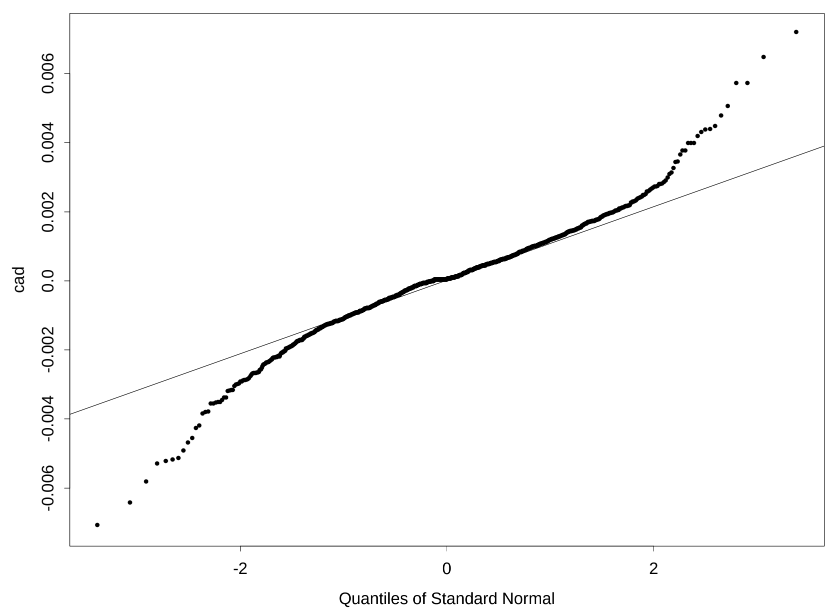


Figure 11: The 0.999(0.001%) quantile for different thresholds. The graph of the quantiles is presented for various thresholds. Usually, it is not satisfactory to have only one result (one extreme VaR); one can choose the threshold from a stable area of the graph of the quantiles and present a pessimistic version for an area of peak.

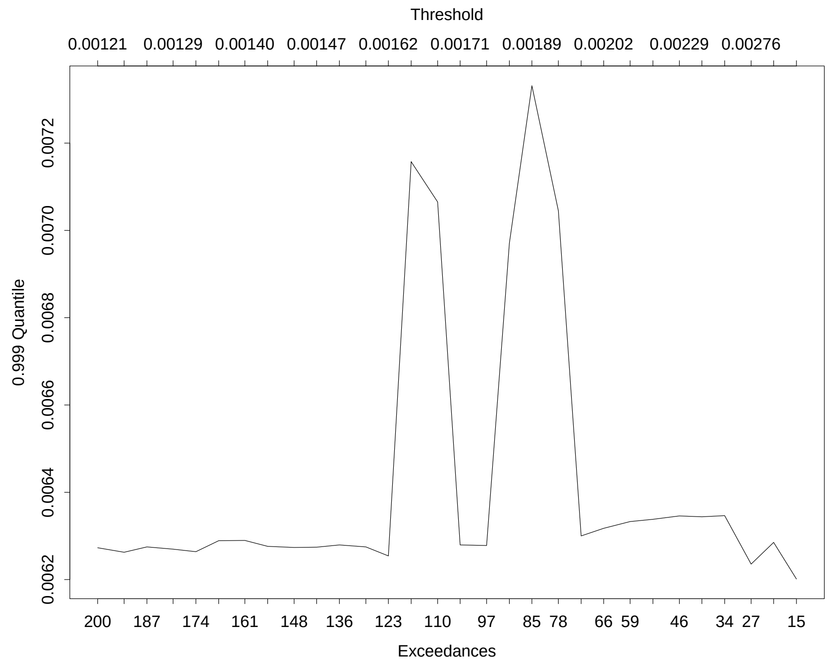


Figure 12: The 0.9999(0.01%) quantile for different thresholds (GPD).

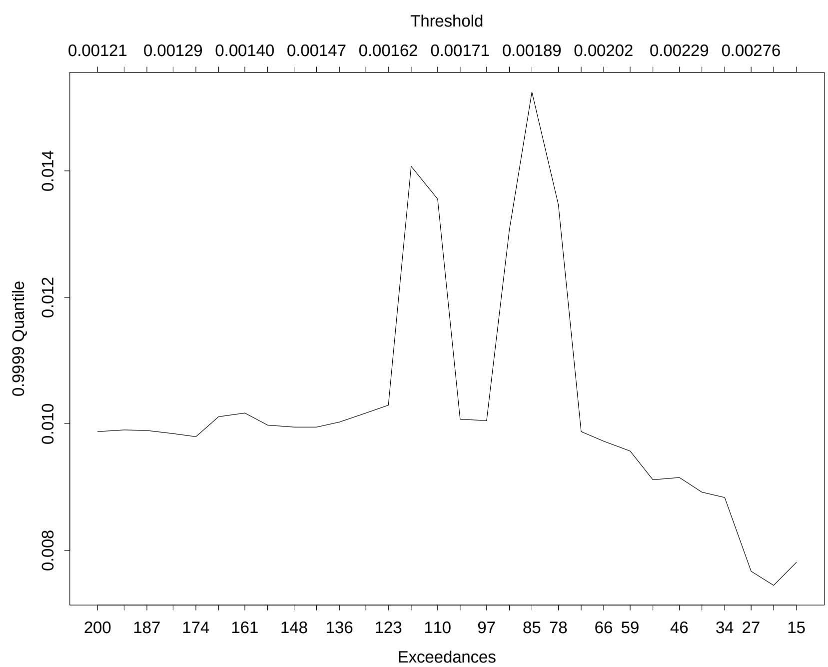
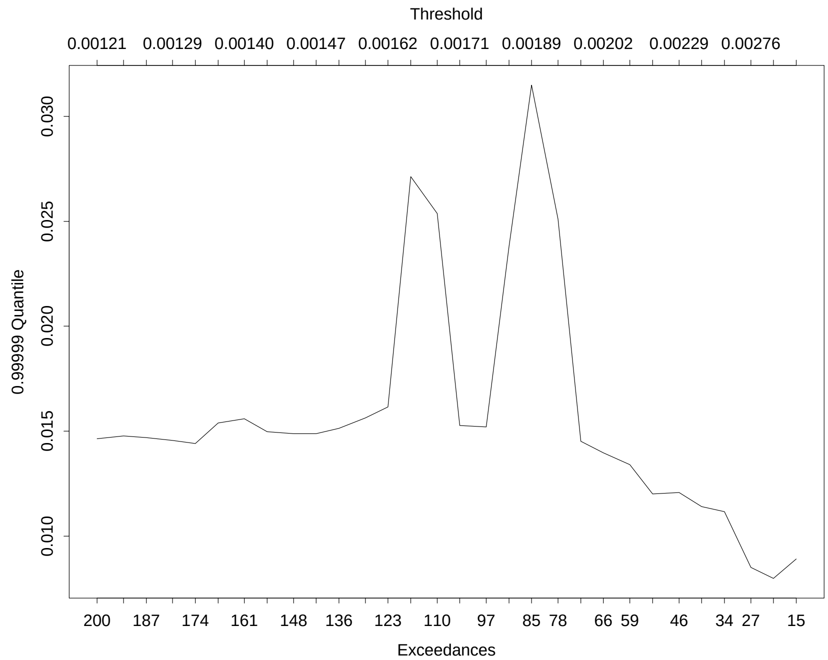


Figure 13: The 0.99999(0.001%) quantile for different thresholds (GPD).



Bank of Canada Working Papers
Documents de travail de la Banque du Canada

Working papers are generally published in the language of the author, with an abstract in both official languages. ***Les documents de travail sont publiés généralement dans la langue utilisée par les auteurs; ils sont cependant précédés d'un résumé bilingue*.**

2000

2000-19	Le modèle USM d'analyse et de projection de l'économie américaine	R. Lalonde
2000-18	Inflation and the Tax System in Canada: An Exploratory Partial-Equilibrium Analysis	B. O'Reilly and M. Levac
2000-17	A Practical Guide to Swap Curve Construction	U. Ron
2000-16	Volatility Transmission Between Foreign Exchange and Money Markets	S.K. Ebrahim
2000-15	Private Capital Flows, Financial Development, and Economic Growth in Developing Countries	J.N. Bailliu
2000-14	Employment Effects of Nominal-Wage Rigidity: An Examination Using Wage-Settlements Data	U.A. Faruqui
2000-13	Fractional Cointegration and the Demand for M1	G. Tkacz
2000-12	Price Stickiness, Inflation, and Output Dynamics: A Cross-Country Analysis	H. Khan
2000-11	Identifying Policy-makers' Objectives: An Application to the Bank of Canada	N. Rowe and J. Yetman
2000-10	Probing Potential Output: Monetary Policy, Credibility, and Optimal Learning under Uncertainty	J. Yetman
2000-9	Modelling Risk Premiums in Equity and Foreign Exchange Markets	R. Garcia and M. Kichian
2000-8	Testing the Pricing-to-Market Hypothesis: Case of the Transportation Equipment Industry	L. Khalaf and M. Kichian
2000-7	Non-Parametric and Neural Network Models of Inflation Changes	G. Tkacz
2000-6	Some Explorations, Using Canadian Data, of the *S*-Variable in Akerlof, Dickens, and Perry (1996)	S. Hogan and L. Pichette
2000-5	Estimating the Fractional Order of Integration of Interest Rates Using a Wavelet OLS Estimator	G. Tkacz
2000-4	Quelques résultats empiriques relatifs à l'évolution du taux de change Canada/États-Unis	R. Djoudad and D. Tessier

Copies and a complete list of working papers are available from:
Pour obtenir des exemplaires et une liste complète des documents de travail, prière de s'adresser à:

Publications Distribution, Bank of Canada
234 Wellington Street Ottawa, Ontario K1A 0G9

Diffusion des publications, Banque du Canada
234, rue Wellington, Ottawa (Ontario) K1A 0G9

E-mail / Adresse électronique: publications@bankofcanada.ca
WWW: http://www.bankofcanada.ca/