University of Victoria
Econometrics Working Paper EWP0708

ISSN 1485-6441

Department of Economics

Extreme Value Analysis of Daily Canadian Crude Oil Prices

Feng Ren and David E. Giles

Department of Economics, University of Victoria
Victoria. B.C.. Canada V8W 2Y2

October 2007

Abstract

Crude oil markets are highly volatile and risky. Extreme value theory (EVT), an approach to modelling and measuring risks under rare events, has seen a more prominent role in risk management in recent years. This paper presents an application of EVT to the daily returns of crude oil prices in the Canadian spot market between 1998 and 2006. We focus on the peak over threshold method by analyzing the generalized Pareto-distributed exceedances over some high threshold. This method provides an effective means for estimating tail risk measures such as Value-at-Risk and Expected Shortfall. The estimates of risk measures computed under different high quantile levels exhibit strong stability across a range of the selected thresholds. At the 99th quantile, the estimates of VaR are approximately 6.3% and 6.8% for daily positive and negative returns, respectively.

Keywords: Crude oil, daily returns, market volatility, extreme value analysis
JEL Classifications: C46, G32, Q40

Contact Author: David Giles, Dept. of Economics, University of Victoria, P.O. Box 1700, STN CSC, Victoria, B.C., Canada; e-mail: dgiles@uvic.ca; Voice: (250) 721-8540; FAX: (250) 721-6214

1. Introduction

Crude oil markets can be volatile and risky. The world crude oil prices have risen dramatically during the past decade. However, oil prices did not sustain a constant rise – rather, they showed high volatility, reflecting market conditions such as political turmoil, supply disruptions, unexpected high demand and speculation. Research conducted by Acadian Asset Management Inc. shows that the daily returns on crude oil (West Texas Intermediate) have a wider range than that of gold, copper and major U.S. stock market indices through the period of 1990-2006. In the Canadian crude oil market, the oil price is represented by the Edmonton par crude oil price. The percentage change in price over one trade day was as high as 17% and as low as -24% between 1998 and 2006. For instance, following the September 11, 2001 attacks, the price of oil plummeted as oil traders believed that weakened economies in the U.S. and elsewhere would use less oil. In particular, on September 24, the Edmonton oil price fell by 24%, its biggest one-day drop through that period. In contrast, on March 23, 1998, the crude oil price increased sharply by 17% over one day because of the news that three of the world's biggest oil producers agreed to cut supply. Relative to the average positive daily return of 1.74% and the average negative daily return of -1.79% during that period, these cases provide examples of extreme events.

Moreover, volatile oil prices may lead to price variability of other energy commodities and can have wide-spread impacts on the international economy. Canada's (known) reserves of oil are among the largest in the world. The bulk of Canadian crude oil is exported to the U.S.A., Canada being the third largest supplier of oil to that country after Saudi Arabia and Mexico. So, volatility in the price of Canadian crude oil impacts significantly on the global economy. Specific examples of impacts include the obvious example of gasoline, and hence transportation costs. Other examples include the stock market and exchange rates, which can be affected substantially by the price of oil (see Nandha and Faff, 2007), as well as the chemical industry. If relevant risk management organizations and investors in these markets can not predict and capture the risks appropriately, their losses could be huge. The highly volatile behaviour of crude oil prices and the substantial impacts of this volatility motivate us to undertake research on modelling oil price fluctuations and providing an effective instrument to measure energy price risks.

In order to better disclose the nature of the risks under extreme situations, and finally avoid the risks in the most degree, we need certain risk measures. Extreme value theory (EVT), a theory for assessing the asymptotic probability distribution of extreme values, models the tail part of the distribution where the risk exists. This theory is playing an increasingly important role in dealing with modelling rare events. While application of EVT is not foolproof, it provides a relatively safe method for extrapolating beyond what has been observed (Embrechts et al., 1997). On successfully modelling tail-related risks, we then need to find suitable instruments to measure these risks. Two popular measures are Value-at-Risk (VaR) and Expected Shortfall (ES). VaR is the maximum loss of a portfolio such that the likelihood of experiencing a loss exceeding that amount over a specified risk horizon is equal to a pre-specified tolerance level. Expected Shortfall measures the mean of the losses that are equal to or greater than VaR. In particular, VaR has become one of the most used techniques in risk management. In order to capture the effect of market behaviour under extreme events, EVT has been widely adopted in VaR estimation in recent years.

Because extreme value methods are derived from a sound statistical theory, and provide a parametric form for the tail of a distribution, these methods are attractive when dealing with measuring risks. There is a large literature that studies extreme value theory for risk measures in areas where extreme observations may appear, such as finance, insurance, hydrology, climatology, engineering and modern science. Specifically, numerous studies in finance and insurance have been conducted, including Embrechts et al. (1997), Reiss and Thomas (1997), Danielsson and de Vires (2000), McNeil and Frey (2000), Gencay et al. (2003) and Gilli and Këllezi (2006). However, to the best of our knowledge, there is only limited discussion of the application of EVT to markets for crude oil, which is a crucial commodity to the world economy. Among the studies on Value-at-Risk estimation on energy markets with extreme value approaches, Krehbiel and Adkins (2005) examine the price risk in the NYMEX energy complex. This study constructs risk statistics for unconditional distributions of daily price changes and applies the conditional extreme value method for estimating VaR and related risk statistics. Another research undertaken by Marimoutou et al. (2006) explores the daily spot Brent oil prices and compares the performances of unconditional and conditional EVT models with that of conventional models such as GARCH and historical simulation.

The remainder of this paper proceeds as follows: section 2 presents an overview of the theoretical framework of extreme value theory and the statistical approach of fitting the generalized Pareto distribution – the peak over threshold model. Section 3 describes the measures of extreme risks – VaR and ES. We discuss the tail modelling of the price time series and assess the outcomes in section 4. Point and interval estimates of risk measures are provided in section 5, and concluding remarks are given in the final section.

2. Extreme Value Theory and Statistical Approaches

Extreme value theory relates to the asymptotic behaviour of extreme observations of a random variable. It provides the fundamentals for the statistical modelling of rare events, and is used to compute tail risk measures. Researchers have contributed abundant theoretical discussion on EVT such as Embrechts et al. (1997), Reiss and Thomas (1997), and Coles (2001). Two different but related methods may be applied in modelling extremes: block maxima models and threshold models. The first way concentrates on the distribution of block maxima, which can be modelled by the generalized extreme value (GEV) distribution. The second one identifies realized large values over some high threshold, which can be simulated by the generalized Pareto distribution (GPD).

The Fisher-Tippet Theorem (Fisher and Tippett, 1928; Gnedenko, 1943) implies that, if maxima are suitably normalized and converge in distribution to a non-degenerate limit, then the limiting distribution must be one of the Fréchet, Gumbel and Weibull families, indicating that these three classes are the only possible forms of extreme value distributions. Furthermore, these three types of distributions can be nested into a single class of continuous probability distributions – the GEV distribution (Embrechts et al., 1997). Inspired by the theorem, the distribution of block maxima (the maxima for n observation periods of fixed and equal length) can be modelled by fitting the GEV to the set of block maxima. In practice, this approach is subject to pitfalls: one key issue in implementing this model is the choice of block size, and another difficulty rises with the use of relatively attractive likelihood-based methods for the GEV. More importantly, modelling block maxima only uses partial information if other data on extreme values are available. This especially may be an issue when one block contains more extremes than another. In contrast, threshold models provide a more efficient means to dealing with extremes, as the information from the entire time series above some threshold will be used.

2.1. The generalized Pareto distribution

By introducing the shape parameter ξ, location parameter μ and scale parameter σ, a three-parameter generalized Pareto distribution has the following representation[1]

---

[1] The one-parameter standard GPD is defined as $G_{\xi}(x) = \begin{cases} 1-(1+\xi x)^{-1/\xi} & \text{if } \xi \neq 0 \\ 1-e^{-x} & \text{if } \xi = 0 \end{cases}$

$$G_{\xi,u,\beta}(x)=\begin{cases}1-\left(1+\xi\left(\dfrac{x-u}{\beta}\right)\right)^{-1/\xi} & \text{if } \xi\neq 0\\ 1-e^{-\left(\frac{x-u}{\beta}\right)} & \text{if } \xi=0\end{cases} \tag{1}$$

where (i) $x-u\geq 0$ when $\xi\geq 0$, and $0\leq x-u\leq -\beta/\xi$ when $\xi<0$, and (ii) $\beta>0$.

The excess distribution above a certain threshold u for a random variable X is referred to as the excess distribution function F_u and is defined as

$$F_u(x)=P(X-u\leq x\mid X>u)$$

for $0\leq x<x_F-u$; where $x_F\leq\infty$ is the right endpoint of F; x represents exceedances over u; and the above expression can be written as

$$F_u(x)=\frac{F(x+u)-F(u)}{1-F(u)}\ . \tag{2}$$

Pickands-Balkema-de Haan Theorem: (Pickands, 1975; Balkema and de Haan, 1974) It is possible to find a positive measurable function β, where β is a function of u, such that

$$\lim_{u\to x_F}\ \sup_{0\leq x\leq x_F-u}|F_u(x)-G_{\xi,\beta}(x)|=0$$

if and only if[2] $F\in MDA\left(H_\xi(x)\right)$.

That is, provided the underlying distribution belongs to the maximum domain of attraction of the GEV, as the threshold u becomes large, the excess distribution function of exceedances over the threshold has approximately a generalized Pareto distribution. This theorem suggests that we use the following approximation conditional on a sufficiently high threshold u and an appropriately positive β:[3]

$$F_u(x-u)\approx G_{\xi,u,\beta}(x)\qquad u\to\infty \tag{3}$$

where the exceedances $x-u\geq 0$ and $G_{\xi,\beta}(x-u)=G_{\xi,u,\beta}(x)$.

---

[2] MDA represents the maximum domain of attraction; $H_\xi(x)$ represents the GEV.

[3] The two-parameter form is written as $F_u(x)\approx G_{\xi,\beta}(x)\qquad u\to\infty$

The generalized extreme value distribution measures the limiting distribution of normalized maxima for a series of i.i.d. random variables, while the generalized Pareto distribution describes the limiting distribution of scaled exceedances above certain high thresholds (residuals beyond the maxima). The key point from the Pickands-Balkema-de Haan Theorem is that, if block maxima have the GEV distribution, then exceedances over some high threshold will have an associated GPD. Furthermore, the parameters of the GPD of exceedances over a certain threshold are determined by that of the associated GEV of block maxima. The shape parameter ξ in the GPD exactly equals that of the corresponding GEV distribution (Coles, 2001), and acts as the dominant factor in determining the tail properties of the GPD and thus measures the fatness of the tail. As for the GEV family, the distribution of exceedances is unbounded if $\xi > 0$ or $\xi = 0$, while the excess distribution has an upper bound if $\xi < 0$. The condition $\xi > 0$ associates with a re-parameterized ordinary Pareto distribution; if $\xi < 0$ then the distribution is Pareto type II; and if $\xi = 0$ then the distribution is exponential. The case of $\xi > 0$ is of our interest, as the corresponding GPD is fat-tailed with a positive tail index, and the larger the shape parameter, the heavier is the tail of the distribution. The choice of the threshold and the estimation of the parameters conditional on that threshold will be discussed in the next section.

2.2. The peak over threshold model: a GPD approach

Based on the Pickands-Balkema-de Haan Theorem, the peak over threshold (POT) model focuses on the distribution of exceedances above some high threshold.
For $x - u \geq 0$, we rewrite the excess distribution function (2) as

$$F_u(x-u) = \frac{F(x) - F(u)}{1 - F(u)} \ , \tag{4}$$

and then we derive the following expression by rearranging (4):

$$F(x) = \left(1 - F(u)\right) F_u(x-u) + F(u) \ .$$

In practice, the application of the POT method involves two steps: choose an appropriate threshold, and then fit the GPD function. Provided that the threshold u is sufficiently high, we can use the method of historical simulation to estimate $F(u)$ by $(1 - N_u/n)$, where n is the total number of sample observations and N_u is the number of observations above the threshold

u . By applying maximum likelihood estimation, $F_u(x-u)$ can be estimated by a GPD approximation (Embrechts et al., 1997). We then obtain the tail estimator

$$\hat{F}(x) = 1 - \frac{N_u}{n}\left(1 + \hat{\xi}\frac{x-u}{\hat{\beta}}\right)^{-1/\hat{\xi}} \tag{5}$$

The choice of threshold is an issue of balancing bias and variance. Coles (2001) points out that if a threshold is too low, it is more likely to violate the asymptotic property of the model and cause bias; if a threshold is too high, it will generate few exceedances with which the model can be estimated and result in high variance. A basic strategy is to select a threshold as low as possible, as long as the limiting approximation of the model can provide a reasonable result. McNeil (1997), Danielsson and de Vires (1997) and others have exploited this data-analytic issue, but there is still not a clear rule or treatment that we can follow in dealing with such a question. In this paper, we apply two approaches: one is to use exploratory tools prior to model estimation; the other is to assess the stability of the estimates of parameters, based on fitting the model across a range of different thresholds.

The first approach for threshold selection utilizes the empirical mean excess (ME) plot. A mean excess function is the mean of the exceedances over a certain threshold *u.* For a random variable *X* with right endpoint x_F , its mean excess function is defined as

$$e(u) = E(X - u \mid X > u)$$

for $u < x_F$. If the underlying distribution of $X > u$ has a generalized Pareto distribution, then the corresponding mean excess is

$$e(u) = \frac{\beta + \xi u}{1 - \xi} \tag{6}$$

where $\xi < 1$ so that $e(u)$ exists. According to equation (6), the mean excess function is linear in the threshold *u*. More precisely, the mean excess function is linear if, and only if, $X > u$ has a generalized Pareto distribution. This important property can help with the selection of the threshold. Provided with data, we define the empirical mean excess function as

$$e_n(u) = \frac{\sum_i^n (X_i - u)_{X_i > u}}{\sum_i^n 1_{X_i > u}} \tag{7}$$

where n represents the number of sample observations that exceed threshold u. The locus of $\left(u, e_n\left(u\right)\right)$ is termed the empirical mean excess plot, and we can choose the threshold u in such a way that $e_n\left(u\right)$ is approximately linear for $x \geq u$.

A complementary method is to fit the GPD by maximizing the log-likelihood to acquire estimates of the shape parameter ξ and the scale parameter β corresponding to different thresholds, and then determine thresholds for which the estimates of the shape parameter are relatively stable. This technique is based on a favourable property of the GPD: above the level of a certain threshold, the asymptotic characteristics of the GPD are valid, and thus estimates of the shape parameter should be roughly constant.

Various techniques have been applied in estimating the parameters of the GPD, including maximum likelihood estimation (MLE), method of moments and the method of probability-weighted moments. The method of MLE is adopted in this study. A major reason is that the maximum likelihood estimator is asymptotically normal, allowing approximations for standard errors and confidence intervals. A detailed introduction to likelihood-based statistical inference can be found in Azzalini (1996). Provided a sufficiently high threshold u and the sample $\{x_1 - u, \ldots, x_n - u\}$ with a underlying distribution of GPD, where $x_i - u \geq 0$ for $\xi > 0$ and $0 \leq x_i - u \leq -\beta / \xi$, for $\xi < 0$, the individual probability density function in logarithmic form is derived from equation (1):

$$\log f\left(x_i\right) = -\log\left(\beta\right) - \frac{1+\xi}{\xi} \log\left(1 + \xi\left(\frac{x_i - u}{\beta}\right)\right).$$

In the case $\xi = 0$, the individual logarithmic density function is obtained as

$$\log f\left(x_i\right) = -\log\left(\beta\right) - \frac{1}{\beta}\left(x_i - u\right).$$

The log-likelihood function $L\left(\xi, \beta \mid x_i - u\right)$ for the GPD is the logarithm of the joint density of the n observations

$$L\left(\xi, \beta \mid x_i - u\right) = \begin{cases} -n\log\left(\beta\right) - \dfrac{1+\xi}{\xi} \displaystyle\sum_{i=1}^{n} \log\left(1 + \xi\left(\frac{x_i - u}{\beta}\right)\right) & \text{if } \xi \neq 0 \\ -n\log\left(\beta\right) - \dfrac{1}{\beta} \displaystyle\sum_{i=1}^{n}\left(x_i - u\right) & \text{if } \xi = 0 \end{cases} \tag{8}$$

Consequently, we may calculate the estimates of tail index ξ and scale parameter β by maximizing the log-likelihood function for the sample corresponding to a suitable threshold u.

3. Measures of Extreme Risks

Risk management in energy markets, especially the market for crude oil, manages pricing risks associated with changing market conditions. One of the most frequent questions concerning risk management in volatile markets is about extreme quantile estimation. Typical examples of such tail-related risk measures are Value-at-Risk (VaR) and Expected Shortfall (ES). This section describes how extreme value theory can be applied to model and compute these measures of extreme risks.

3.1. Value-at-Risk and Expected Shortfall

The origin of the concepts of Value-at-Risk or VaR-like measures can be traced back as far as 1922 to capital requirements the New York Stock Exchange imposed on member firms. VaR was widely adopted for measuring market risks starting in the early 1990s (Holton, 2002). In particular, the Basel Committee on Banking Supervision (1996) at the Bank for International Settlements imposed on financial institutions such as banks and investment firms to meet capital requirements based upon VaR measures. Value-at-Risk measures the maximum potential losses in the market value of, say, a financial portfolio with a given level of confidence over a specified period. Essentially, VaR is a high quantile on the distribution of returns.

Mathematically, let X be a random variable with continuous cumulative distribution function F_x that models the return distribution of a risky financial instrument over the specified time horizon. For a given probability p, VaR can be defined as the p-th quantile of the distribution F_x:

$$\text{VaR}_p = F^{-1}(1-p) \tag{9}$$

where F^{-1}, the inverse of the distribution function F, is the quantile function. [4]

VaR is a good risk measure, but it does not capture all the aspects of market risks. A better methodology for risk measurement would be to combine VaR and other complement measure vehicles. Artzner et al. (1997, 1999) argue that ES, as opposed to VaR, is a coherent risk measure, and thus suggest the use of Expected Shortfall. ES describes the expected size of the return that exceeds VaR. This risk measure exhibits some information on the size of the potential returns

[4] In terms of the quantile q, the equivalent expression is $\text{VaR}_q = F^{-1}(q)$, where $q = 1 - p$

given that the return is bigger than VaR. The expression of ES for risk *X* at given probability level *p* is defined as

$$ES_p = E(X \mid X > VaR_p) \tag{10}$$

3.2. Calculation of VaR and ES

Traditional approaches to compute VaR include parametric and non-parametric approaches. The parametric approach applies historical data on the underlying variable, and assumes that the relevant variable owns a normal distribution. The limitation of this approach is obvious: the assumption of normality for the underlying distribution is not realistic. For example, the financial and commodity data often exhibit the properties of asymmetry and heavy tails. Thus the inferences drawn about the future level of the variable is lack of reliability and accuracy. An alternative way is known as non-parametric approach, represented by Monte Carlo simulation method and historical simulation method. The principle is that specific values can be forecasted based on the estimation of a range for the future values, and then VaR can be calculated using these values. This approach of measure is still restricted under limitations. On the other hand, the extreme value approach based VaR estimation is superior to the traditional parametric and non-parametric methods in identifying extreme risks (for example, Aragonés et al., 2000).

For given level of probability *p* (correspondingly the quantile level is $q = 1 - p$), the tail quantile can be obtained by inverting the tail estimator formula provided in equation (5)

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

where $\hat{\xi}$ and $\hat{\beta}$ are the maximum likelihood estimates of the GPD parameters. As VaR is exactly the extreme quantile, it can be estimated by

$$V\hat{a}R_p = u + \frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}p\right)^{-\tilde{\xi}} - 1\right). \tag{11}$$

An equivalent expression of formula (10) is

$$ES_p = VaR_p + E(X - VaR_p \mid X > VaR_p) \tag{12}$$

where the second term represents the mean of the excess distribution $F_{VaR_p}(x)$ above the threshold VaR_p. Applying the Pickands-Balkema-de Haan Theorem, if the threshold VaR_p for

$1-p>F(u)$ is sufficiently high, then the excess distribution above this threshold is a GPD, implying the following relation

$$F_{VaR_p}(x)=G_{\xi,\beta+\xi(VaR_p-u)}(x)\ .$$

Thus, the mean of the excess distribution $F_{VaR_p}(x)$ can be easily calculated as

$$\left(\beta+\xi\left(VaR_p-u\right)\right)/\left(1-\xi\right)$$

where $\xi<1$. Substituting this result into equation (12), we obtain the expression of ES:

$$E\hat{S}_p=\frac{V\hat{a}R_p}{1-\hat{\xi}}+\frac{\hat{\beta}-\hat{\xi}u}{1-\hat{\xi}}\ . \tag{13}$$

Correspondingly, the explicit form of the ES estimator is

$$E\hat{S}_p=\frac{1}{1-\hat{\xi}}\left(u+\frac{\hat{\beta}}{\hat{\xi}}\left(\left(\frac{n}{N_u}p\right)^{-\hat{\xi}}-1\right)\right)+\frac{\hat{\beta}-\hat{\xi}u}{1-\hat{\xi}}\ . \tag{14}$$

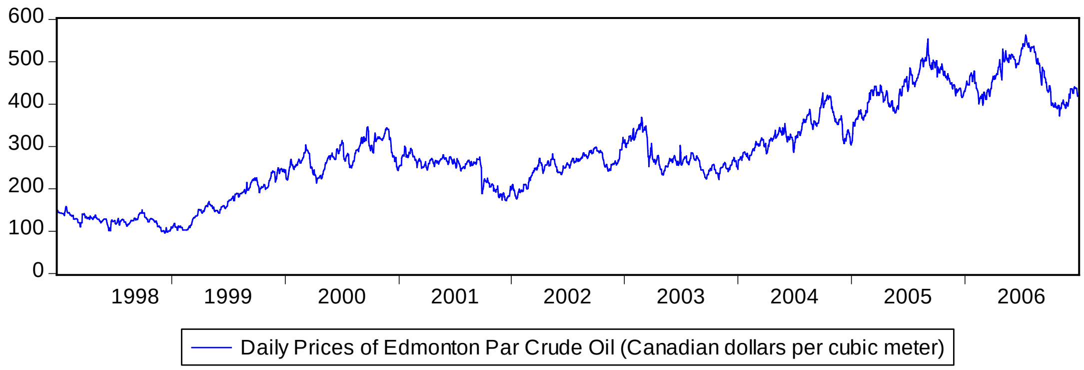
McNeil (1999) stresses that formula (13) indicates the importance of the shape parameter ξ in tail estimation, and essentially ξ determines how VaR and ES differ as risk measures.

4. Modelling the Tails of Return Distributions Using EVT

Let us now return to our primary objective – analyzing the volatility of daily crude oil prices by applying the extreme value theory to model the tails of the distribution for daily returns. In this section we describe the historical data for spot oil prices on the Canadian market, the preliminary tests undertaken on the data and exploratory techniques, the determination of thresholds, the fitting of the GPD, and the examination of tail modelling. The empirical analysis has been undertaken by writing program code that was executed using the EViews 5 econometrics package.

4.1. Data description

There are four key players in the Canadian oil market: Imperial Oil, Petro-Canada, Shell Canada and Suncor Energy. The Edmonton par crude oil prices refer to the average prices of par crude postings at Edmonton from these integrated oil companies. Our raw data represents per-cubic-meter oil prices in Canadian dollars over the period January 4, 1998 through December 31, 2006[5]. After preliminary treatment[6] to the raw data, we focus on the daily returns of Canadian crude oil prices, measured as differences in the natural logarithms of the price.



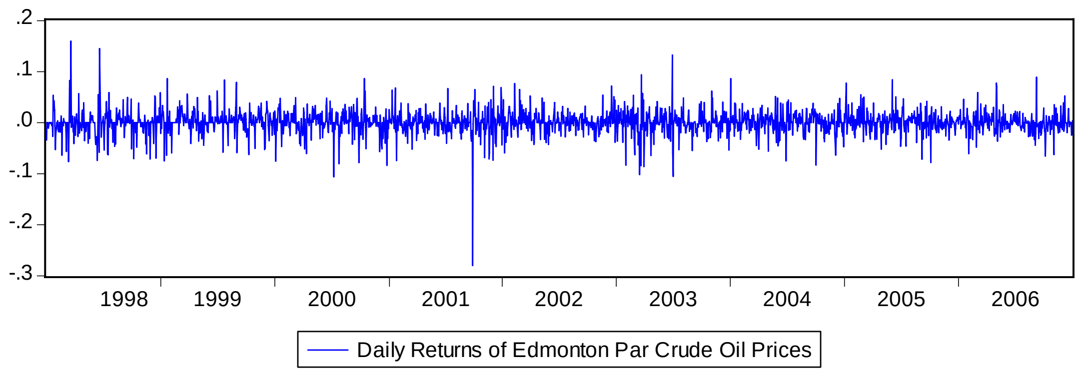


Fig. 1. Daily oil prices (top) and daily logarithmic returns (bottom).

There are 2,319 adjusted observations, including 1,018 observations of gains and 897 observations of losses. The plot of the daily crude prices (Fig. 1, top) shows a substantial increase since 1998 with lots of fluctuations, and the graph of daily returns (Fig. 1, bottom) confirms the volatility of the Canadian crude oil market.

4.2. Preliminary tests and exploratory analysis

The Jarque-Bera (1980) test uses sample skewness and kurtosis to measure the deviation of a distribution from normality. Under the null hypothesis, both the skewness and excess kurtosis

[5] Price data were retrieved from the "Summary of Par Crude Postings at Edmonton – Daily, Crude Oil Data", Natural Resources Canada.

[6] Treatment of non-trade-day data: duplicated data on Saturday and Sunday are removed in order to better describe the price movements in trade days. For same reason, data on fixed-date Canadian public holidays including January 1, July 1 and December 25 and 26 are also eliminated. The elimination of the duplicated data on weekends and holidays does not affect data accuracy.

are zero. In our sample, the Jarque-Bera statistic is very large and the associated p-value essentially equals zero, suggesting that the normality of the return distribution is unrealistic. The larger kurtosis value (kurtosis = 16.2) relative to that under the normal distribution implies the existence of fat tails, and the statistic also shows the return distribution skews to the left. Krehbiel and Adkins (2005) examine different spot market commodities of the NYMEX energy complex, and conclude that the commodity returns are not normally distributed and exhibit the asymmetry and heavy tails; in particular, all empirical distributions demonstrate significant negative skewness. Our finding is in line with their study. Because of the existence of asymmetry of the return distribution in crude oil market, it is necessary to model left and right tails separately in order to capture their distinct characteristics. For the negative one, we follow the rule that the absolute values of losses are examined.

McNeil and Frey (2000) propose a conditional extreme value method: first a GARCH-type model is fitted to the return data using quasi-maximum likelihood, and then a generalized Pareto distribution approximation is fitted to the tails of the innovations. Following this thought, we fit various forms of GARCH-type models. The results show that there is no evidence indicating the existence of conditional heteroskedasticity in our sample data. Our finding from this preliminary analysis coincides with the conclusion of Danielsson and de Vires (2000): for long time horizons an unconditional approach is better suited. For this reason, the GARCH-GPD specification is not adopted in our study.

A popular tool in conducting exploratory data analysis is the quantile-quantile (QQ) plot. The QQ plot is a graphical technique to check whether our sample data is consistent with some known distribution and thus can be used to assess goodness of fitting. The quantile function Q is the generalized inverse function of the cumulative distribution function F:

$$Q(p) = F^{-1}(p) \quad \text{where } 0 < p < 1$$

The quantity $x_p = F^{-1}(p)$ defines the $p-th$ quantile of the distribution function F. A QQ plot compares the quantiles of the empirical distribution function with the quantiles of the reference distribution model. If the empirical data comes from the family of reference distribution, the plot will be approximately linear. If the plotted values deviate far from a straight line, the sample likely comes from a different distribution. For a normal QQ plot, the points on the QQ plot should have an S shape if the sample data has heavy tails relative to a normal distribution.

Fig. 2 illustrates the QQ plots of daily returns against the normal and the Student's t distribution respectively. The plots suggest that the underlying distribution of daily returns does not fit the normal or Student's t distribution. The plot curves down to the right, implying that the data has a heavier right tail compared with the normal distribution; on the contrary, the plot curves up to the left, which indicates that the data has a fatter left tail than the normal distribution. This property confirms the heavy-tailed behaviour of the crude oil daily price change series. Compared to the Student's t distribution, the underlying distribution of daily returns fits to some degree. There is still evidence showing that the distribution of our sample data is heavier in the tail against the Student's t distribution. In addition, we examine the QQ plots of daily price changes against quantiles from various standard statistical distributions. Essentially, no distribution matches the sample data exactly.

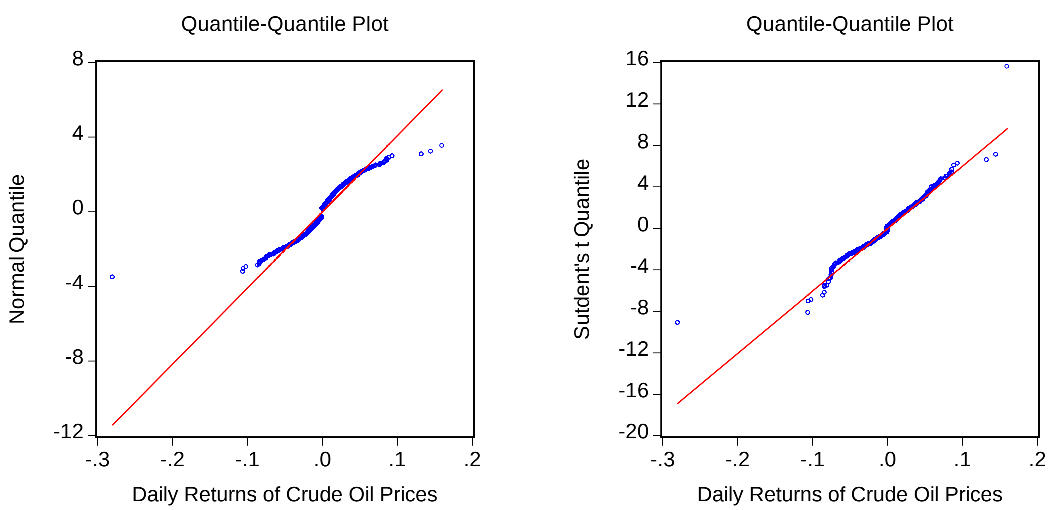


Fig. 2. QQ plots for daily returns against the normal and the Student's t distributions.

4.3. Determination of thresholds

Choosing some suitable threshold is critical in order to adopt the POT method to model the tails of the distribution of daily returns. Two techniques can contribute to the determination of an appropriate threshold level: one is the mean excess plot and the other is the maximum likelihood estimation to a parametric GPD.

The ME plot is helpful in detecting graphically the quantile above which the Pareto's relationship is valid. Section 2.2 details that the empirical mean excess plot is approximately linear in the threshold u given that the underlying distribution of sample data is a GPD. More specifically, the ME plot of the data can be used to distinguish between light- and heavy-tailed models: the plot of a heavy-tailed distribution shows an upward trend, a medium tail shows a horizontal line, and the plot is downward-sloped for light-tailed data. A common ground in our sample data is that both the ME plots of positive and negative returns have an upward-trend part followed by an irregular portion in the far end. The initial and small part of the gain plot is downward-sloping until $u \approx 0.012$, followed by a roughly upward-sloping straight line until $u \approx 0.073$, where upon it varies sharply. The case of losses shows an approximate linearity with slightly upward trend in the threshold from $u = 0$ to $u \approx 0.064$. Therefore, there is some evidence to choose thresholds from 0.012 to 0.073 for the right tail and from 0.000 to 0.064 for the left tail based on the criterion of linearity in the ME plots shown in Fig. 3.

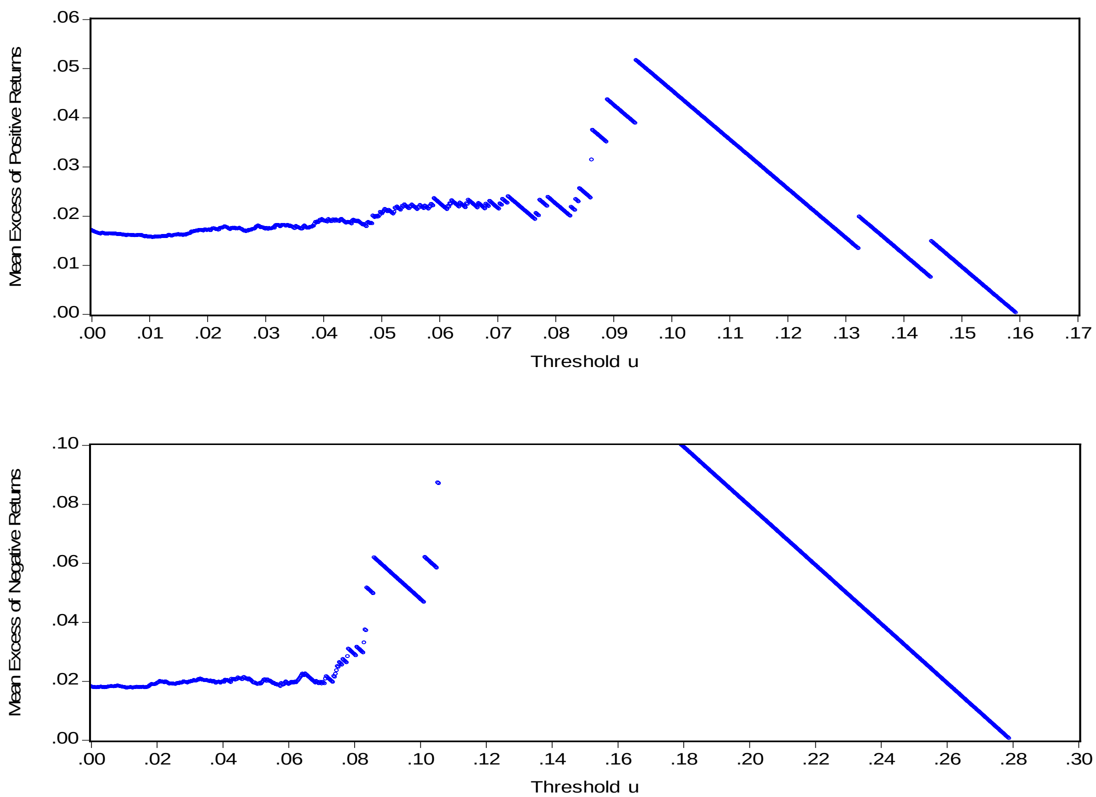


Fig. 3. The ME plots for positive and negative returns.

As a means of threshold selection, the ME plots may be difficult to interpret, and the results can be treated as preliminary conclusions. A further step is to apply the GPD fitting and look for stability of shape parameter estimates. We fit the exceedances of daily returns beyond the associated threshold in each tail to the GPD. Because the maximum likelihood estimator of the shape parameter is asymptotically normal, we can calculate associated approximate standard errors and construct confidence intervals for this parameter. The plots of the shape parameter estimates against different threshold levels are shown in Fig. 4. The upper and lower dashed lines constitute confidence intervals at an approximate 95% level. There exists some scope of the threshold that the estimated shape parameter is relatively steady: from 0.006 to 0.046 for the tail of gains, and from 0.000 to 0.048 for the tail of losses. The estimated shape parameter $\hat{\xi}$ and scale parameter $\hat{\beta}$ as well as their associated standard errors under different thresholds for both tails are listed in Table 1.

In order to apply extreme value theory, the threshold should be sufficiently large so that only the tail of the distribution is being analyzed. When the threshold is close to zero, there are too many observations included. Practical experience suggests it is reasonable including observations up to roughly one-fifth of the total number of observations for both positive returns and negative returns. This is somewhat arbitrary, but provides a reasonable compromise. Combining this restriction and results from the ME plots and the shape parameter plots, we choose the range of the threshold from 0.027 to 0.046 for positive returns and from 0.028 to 0.048 for negative returns.

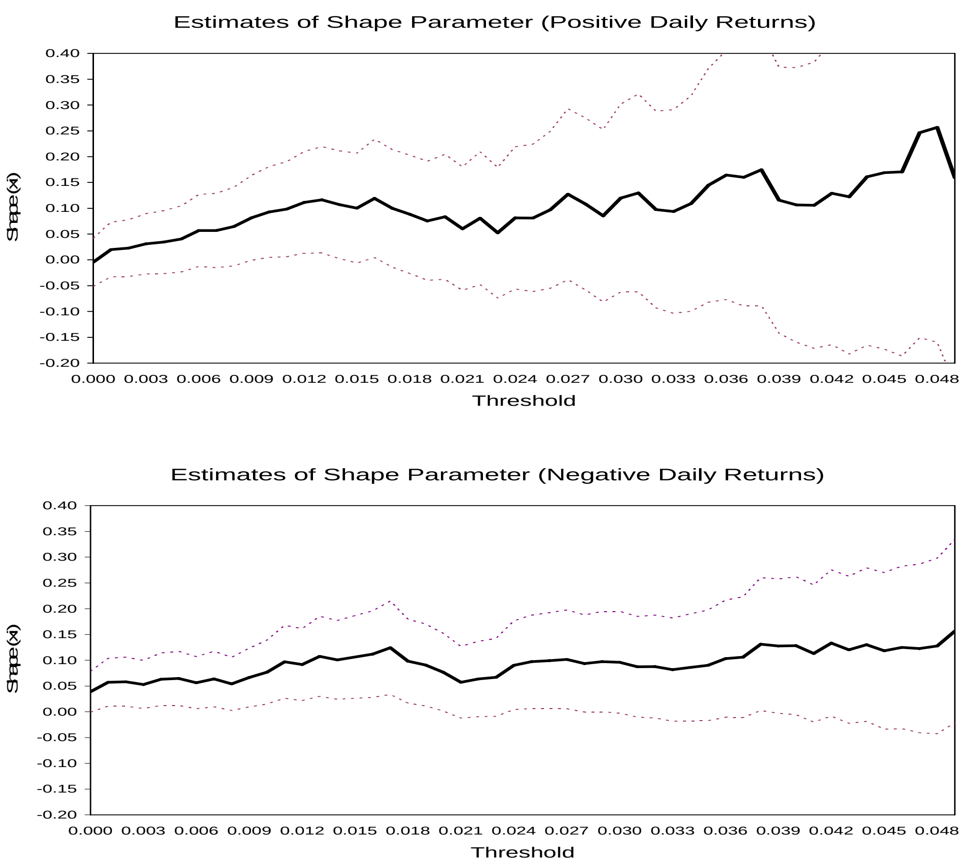


Fig. 4. Estimates of the shape parameter ξ of the GPD for both returns.

Table 1
Maximum likelihood parameter estimation under different thresholds for both returns

	Parameter estimates for positive returns			Parameter estimates for negative returns		
	u =0.006	u =0.027	u =0.046	u =0.000	u =0.028	u =0.048
$\hat{\xi}$ (s.e.)	0.0569 (0.0348)	0.1272 (0.0828)	0.1704 (0.1784)	0.0388 (0.0199)	0.0935 (0.0471)	0.1277 (0.0851)
$\hat{\beta}$ (s.e.)	15.1490 (0.8049)	14.7867 (1.5860)	15.8048 (3.2247)	17.5690 (0.7628)	17.6787 (1.7990)	18.0102 (3.8846)

4.4. Examination of tail modelling

Extreme value theory suggests that the excess distribution above a suitable threshold of daily returns should follow a generalized Pareto distribution. To determine how the GPD fits the tails of the return distribution, we plot the empirical distribution of exceedances along with the cumulative distribution simulated by a GPD and compare the results visually. Fig. 5 provides the plots corresponding to different threshold levels for gains and losses, respectively. For gains, the plots show the GPD fitting to 201 exceedances at the threshold $u = 0.027$ with the shape parameter estimate $\hat{\xi}$ = 0.127 (top, left), and the GPD fitting to 63 exceedances at the threshold $u = 0.046$ with the shape parameter estimate $\hat{\xi}$ = 0.170 (top, right). For losses, the plots give the GPD fitting to 184 exceedances at the threshold $u = 0.028$ with the shape parameter estimate of $\hat{\xi}$ = 0.094 (bottom, left), and the GPD fitting to 64 exceedances at the threshold $u = 0.048$ with the shape parameter estimate $\hat{\xi}$ = 0.128 (bottom, right). For both positive and negative returns, the graph of the empirical excess distribution function follows closely the trace of a corresponding GPD, implying that the GPD models the exceedances very well. Two points are noticeable in the plots: as the increase of the threshold, the fit is getting less precise for both gains and losses; at an either lower or higher threshold level, positive daily return series fits a GPD slightly better than negative returns does.

The survival function is the probability of observing a value from the series x at least as large as some specified value u. It equals one minus the cumulative distribution function[7]. The survival functions plotted in Fig. 6 under different thresholds for both daily gains and losses depict similar facts as to what has happened in the excess distribution functions.

[7] $S_u(x) = 1 - F(x)$

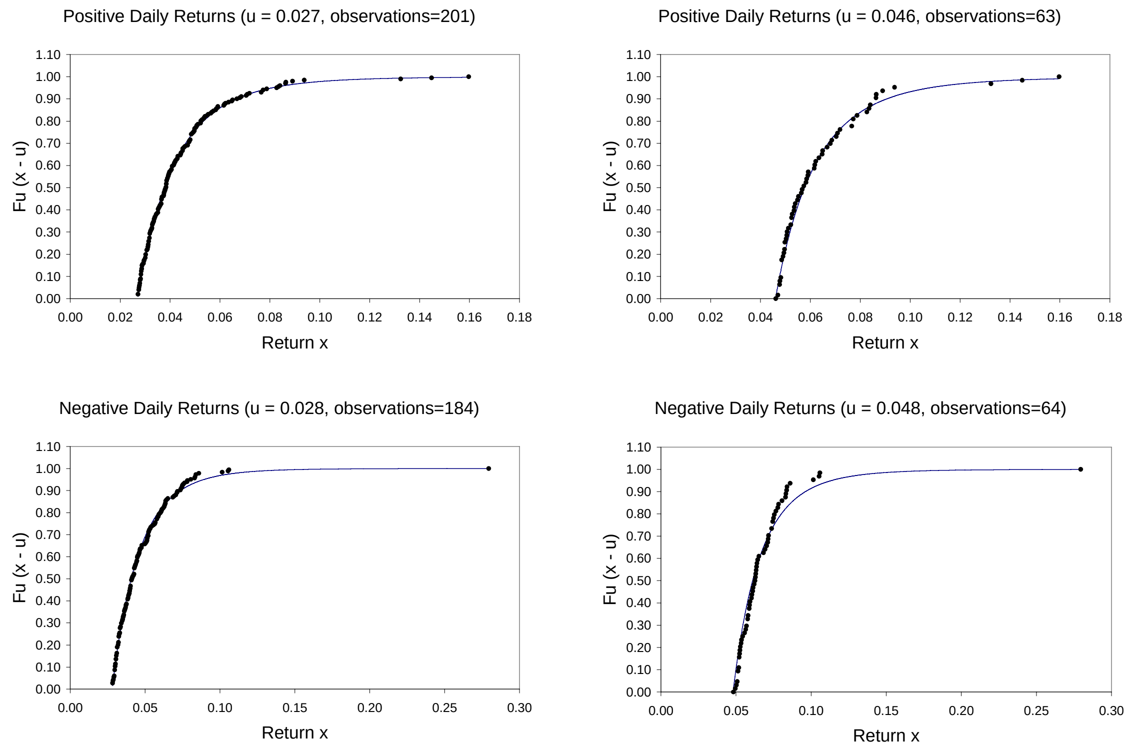


Fig. 5. Excess distribution functions for positive and negative returns.

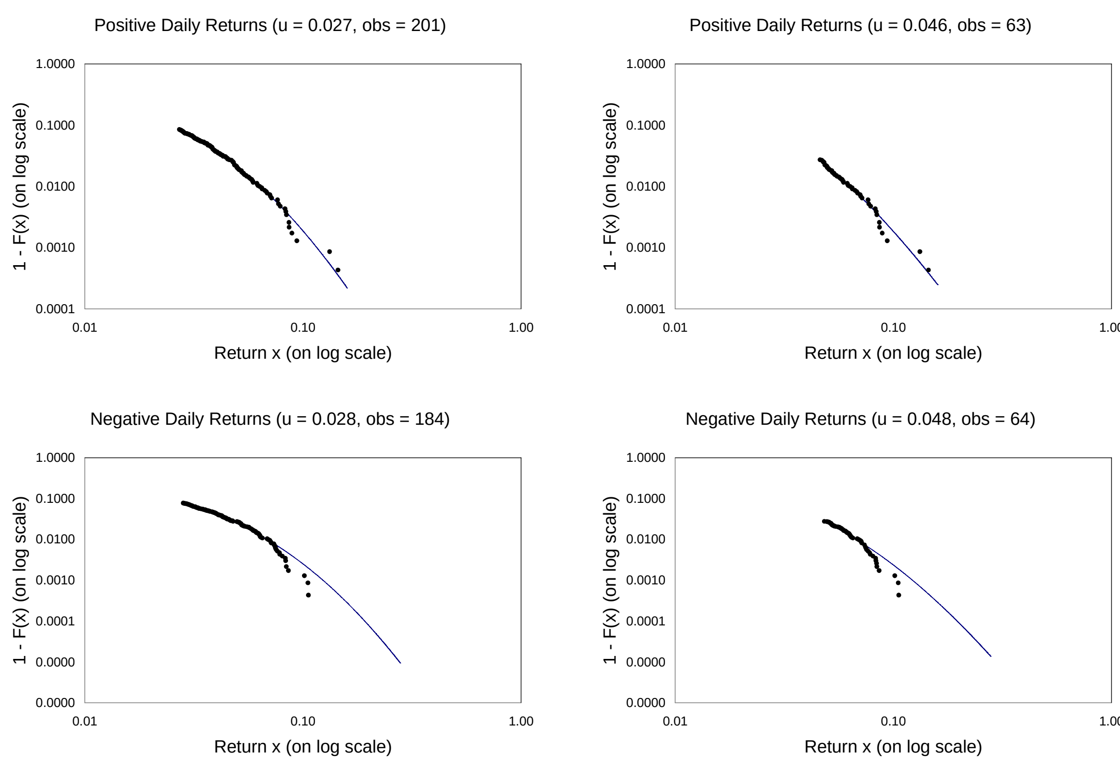


Fig. 6. Survival functions for positive and negative returns.

Overall, based on different fittings with different values of the threshold and associated parameter estimates, the GPD models the tail behaviour of our daily returns very well and the fits exhibit reasonable robustness to the choice of thresholds.

5. Computation of Extreme Risk Measures

We conduct one-period ahead return estimation in both tails of the return distribution at two tail quantiles. Point estimation of VaR and ES follow the formula in equations (11) and (14). Interval estimates are generated using the delta method (see Oehlert, 1992). Table 2 provides the estimates of VaR and ES and their respective confidence intervals for both positive and negative returns, given different levels of the threshold and different tail quantiles.

Table 2
Point and interval estimates of VaR and ES

Value-at-Risk for positive returns				
Quantile	Threshold	Estimate	CI Lower	CI Upper
99-level	$u = 0.027$	0.0637	0.0579	0.0696
99-level	$u = 0.046$	0.0632	0.0580	0.0684
95-level	$u = 0.027$	0.0354	0.0339	0.0370
95-level	$u = 0.046$	0.0368	0.0325	0.0412
Value-at-Risk for negative returns				
Quantile	Threshold	Estimate	CI Lower	CI Upper
99-level	U = 0.028	0.0684	0.0619	0.0750
99-level	U = 0.048	0.0675	0.0604	0.0747
95-level	U = 0.028	0.0363	0.0348	0.0379
95-level	U = 0.048	0.0377	0.0330	0.0424
Expected Shortfall for positive returns				
Quantile	Threshold	Estimate	CI Lower	CI Upper
99-level	U = 0.027	0.0860	0.0723	0.0998
99-level	U = 0.046	0.0858	0.0712	0.1005
95-level	U = 0.027	0.0536	0.0492	0.0580
95-level	U = 0.046	0.0540	0.0508	0.0572
Expected Shortfall for negative returns				
Quantile	Threshold	Estimate	CI Lower	CI Upper
99-level	U = 0.028	0.0921	0.0812	0.1029
99-level	U = 0.048	0.0910	0.0780	0.1040
95-level	U = 0.028	0.0567	0.0521	0.0613
95-level	U = 0.048	0.0568	0.0542	0.0595

In this study, Value-at-Risk measures the best/worst case scenario on the market value of the Edmonton par crude oil over one trade day and given a specified degree of confidence. We first consider the cases of point estimates under the lower threshold for both tails (0.027 for the right tail and 0.028 for the left tail) with statistics shown in Table 2. For example, we calculate VaR as 0.0637 at the 99th percentile for the right tail. That is, given usual conditions, we expect a daily change in the value of par crude oil in the Edmonton market would not increase by more than 6.37%. In other words, the market value, with a probability of 1%, would be expected to gain by $63,700 or more if we have an investment of $1 million in that market. On the other hand, VaR is estimated as 0.0684 at the 1st percentile for the left tail. This implies that, for the lowest 1% negative daily returns, the worst daily loss in the market value may exceed 6.84% in expectation. Put differently, if we invest $1 million in Canadian par crude oil, we are 99% confident that our daily loss at worst will not exceed $68,400 during one trade day. Similarly, at a lower quantile of 95-level, the estimated VaR is 0.0354 for gains and 0.0363 for losses. We can state that, with 95% confidence, the expected market value of par crude oil would not gain by more than 3.54% for the best case scenario or lose more than 3.63% for the worst case scenario within one-day duration. Under the higher threshold for both tails (0.046 for the right tail and 0.048 for the left tail), the estimates of VaR are very close to their corresponding values under the lower threshold, and the estimates may or may not be larger than that under a lower threshold.

Some characteristics of the estimation can be summarized as follows: (1) under different thresholds, the estimates of VaR exhibit strong stability; (2) given the quantile level, under either the lower or higher threshold, the corresponding VaR estimate in the left tail is larger than that in the right tail, but the difference is small, implying that the behaviour difference in both tails is likely to be small.

Krehbiel and Adkins (2005) present risk statistics results for the 4,102 conditional-EVT backtest trials in the West Texas Intermediate (WTI) crude oil spot market between December 1986 and April 2003. The average of VaR calculated at the 99-level quantile is 0.0528 for the right tail and 0.0677 for the left tail. Compared with their results, we have similar estimates of the left-tail VaR, while their estimate in the right tail is significantly smaller than ours.

The usefulness of VaR estimates crucially depends on their precision. The most straightforward way to examine this is to construct confidence intervals. For instance, we discuss the right tail here. An approximate 95% confidence interval constructed for the 99-level quantile VaR is [0.0579, 0.0696] under $u = 0.027$ and [0.0580, 0.0684] under $u = 0.046$; for the 95-level quantile VaR the associated 95% confidence intervals are [0.0339, 0.0370] and [0.0325, 0.0412]. Obviously the two confidence intervals under different thresholds overlap considerably for both quantile levels. In the case of the left tail, the interval estimates tell a similar story.

Table 2 also provides the estimates and confidence intervals of ES under various combinations of different quantile levels and thresholds. The estimates and intervals exhibit similar characteristics to those observed from the VaR. In contrast with the results provided by Krehbiel and Adkins (2005), our estimates of ES are a little larger in the left tail and more than 30 percent larger in the right tail. Noticeably, in our study, the left tail ES exceeds VaR by a slightly greater margin than the right rail ES exceeds VaR.

Researchers have conducted sound studies on the tail distribution modelling by applying some methods of univariate extreme value theory, especially in the financial field. An important argument is that the EVT approach well captures the features of the innovation distribution and can provide more accurate estimates of risk measures compared with other approaches (for example, McNeil, 1997; Gencay *et al*., 2003; Fernandez, 2005), and one can obtain better estimates with the application of the GPD fitting of the excess distribution based on threshold models (for example, Coles, 2001; Gilli and Këllezi, 2006; Marimoutou et al., 2006). This confirms our belief of choosing the POT method to apply the extreme value theory. Overall, the assessment of our results shows that the point and interval estimations are stable and reliable, implying that this approach of modelling extreme values can be used to further application of extreme events. Some studies, including Krehbiel and Adkins (2005), also claim that the upper and lower tails behave differently, and thus should be treated separately while estimating risk measures. Evidences from our empirical study show the small difference in risk statistics on both tails, implying that the thickness of two tails is likely to be similar.

6. Conclusions

The high volatility of prices in oil markets requires the implementation of effective risk management. Extreme value theory is a powerful tool to estimate the effects of extreme events in risky markets based on sound statistical methodology. This study exhibits how EVT can be used to model tail-related risk measures such as Value-at-Risk and Expected Shortfall by applying it to the daily returns of crude oil prices in the Canadian spot market. Our application captures the heavy-tailed behaviour in daily returns and the asymmetric characteristics in distributions, suggesting us to treat positive and negative returns separately. An unconditional approach is favoured as no evidence indicates the existence of conditional heteroskedasticity in our sample data. In the context of applying EVT, the peak over threshold method provides a simple and effective means to choose thresholds and estimate parameters. By assessing empirical excess distribution functions and survival functions with associated theoretical distribution simulations, we find the goodness of fit in tail modelling. Furthermore, as the increase of the threshold, the fit is getting less precise for both gains and losses; at an either lower or higher threshold level, positive daily return series fits a GPD slightly better than negative one does. The point and interval estimates of VaR and ES computed under different high quantile levels exhibit strong stability through a range of the selected thresholds, implying the accuracy and reliability of the estimated quantile-based risk measures.

Our EVT-based Value-at-Risk approach provides quantitative information for analyzing the extent of potential extreme risks in energy markets, particularly the crude oil markets. Interested organizations and corporations could employ this technique as one of the means of risk management. For those who invest in the Canadian crude oil market, our estimates of VaR and ES provide quantitative indicators for their investment decisions.

Acknowledgement

We are grateful to Judith Clarke for her helpful comments and suggestions which led to several improvements to this paper.

References

Acadian Asset Management Inc., 2006. Assessing extreme risk. Retrieved from http://www.acadian-asset.com/Documents/167c7006-fe50-4e5b-a913-772634b6e236.pdf

Aragonés, J.R., Dowd, K., Blanco, C., 2000. Extreme value VaR. Derivatives Week 7-8.

Artzner, P., Delbaen, F., Eber, J., Heath, D., 1997. Thinking coherently. Risk 10, 68-71.

Artzner, P., Delbaen, F., Eber, J., Heath, D., 1999. Coherent measures of risk. Mathematical Finance 9, 203-228.

Azzalini, A., 1996. Statistical Inference Based on the Likelihood. Chapman & Hall, London.

Balkema, A. A., de Haan, L., 1974. Residual lifetime at great age. Annals of Probability 2, 792-804.

Coles, S., 2001. An Introduction to Statistical Modeling of Extreme Values. Springer-Verlag, London.

Danielsson, J., de Vries, C., 1997. Tail index and quantile estimation with very high frequency data. Journal of Empirical Finance 4, 241-257.

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

Embrechts, P., Klüppelberg, C., Mikosch, T., 1997. Modelling Extremal Events for Insurance and Finance. Springer-Verlag, Berlin.

Fernandez, V., 2005. Risk management under extreme events. International Review of Financial Analysis 14, 113-148.

Fisher, R. A., Tippett, L. H. C., 1928. Limiting forms of the frequency distribution of the largest or smallest member of a sample. Proceedings of the Cambridge Philosophical Society 24, 180-190.

Gencay, R., Selcuk, F., Ulugülyagci, A., 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. Insurance: Mathematics and Economics 33, 337-356.

Gilli, M., Këllezi, E., 2006. An application of extreme value theory for measuring financial risk. Computational Economics 27, 207-228.

Gnedenko, B. V., 1943. Sur la distribution limite du terme d'une série aléatoire. Annals of Mathematics 44, 423-453.

Holton, G. A., 2002. History of value-at-risk: 1922-1988. Contingency Analysis Working Paper.

Jarque, C. M., Bera, A. K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. Economics Letters 6, 255–259.

Krehbiel, T., Adkins, L.C., 2005. Price risk in the NYMEX energy complex: an extreme value approach. Journal of Futures Markets 25, 309-337.

Marimoutou, V., Raggad, B., Trabelsi, A., 2006. Extreme value theory and value at risk: application to oil market. GREQAM Working Paper 2006-38.

McNeil, A. J., 1997. Estimating the tails of loss severity distributions using extreme value theory. ASTIN Bulletin 27, 117-137.

McNeil, A. J., 1999. Extreme value theory for risk managers. British Bankers' Association, Internal Modelling and CAD II: Qualifying and Quantifying Risk within a Financial Institution 93-113. RISK Books, London.

McNeil, A. J., Frey, R., 2000. Estimation of tail-related risk measures for heteroskedastic financial time series: an extreme value approach. Journal of Empirical Finance 7, 271-300.

Nandha, M., Faff, R., 2007. Does oil move equity prices? A global view. Energy Economics.

Oehlert, G. W., 1992. A note on the delta method. The American Statistician 46, 27-29.

Pickands, J., 1975. Statistical inference using extreme order statistics. Annals of Statistics 3, 119–131.

Reiss, R.D., Thomas, M., 1997. Statistical Analysis of Extreme Values with Applications to Insurance, Finance, Hydrology and Other Fields. Birkhäuser Verlag, Basel.