# Extreme Risk, Value-At-Risk And Expected Shortfall In The Gold Market

Knowledge Chinhamu, University of KwaZulu-Natal, South Africa
Chun-Kai Huang, University of Cape Town, South Africa
Chun-Sung Huang, University of Cape Town, South Africa
and Delson Chikobvu, University of Free State, South Africa

**ABSTRACT**

*Extreme value theory (EVT) has been widely applied in fields such as hydrology and insurance. It is a tool used to reflect on probabilities associated with extreme, and thus rare, events. EVT is useful in modeling the impact of crashes or situations of extreme stress on investor portfolios. It describes the behavior of maxima or minima in a time series, i.e., tails of a distribution. In this paper, we propose the use of generalised Pareto distribution (GPD) to model extreme returns in the gold market. This method provides effective means of estimating tail risk measures such as Value-at-Risk (VaR) and Expected Shortfall (ES). This is confirmed by various backtesting procedures. In particular, we utilize the Kupiec unconditional coverage test and the Christoffersen conditional coverage test for VaR backtesting, while the Bootstrap test is used for ES backtesting. The results indicate that GPD is superior to the traditional Gaussian and Student's t models for VaR and ES estimations.*

**Keywords:** Gold Prices; Generalised Pareto Distribution; Value-At-Risk; Expected Shortfall; Kupiec; Christoffersen

## INTRODUCTION

Risk measures are used primarily to safeguard a financial position against severe losses. To successfully model such tail-related risks, we need to find suitable techniques to measure and capture these extreme events. Despite certain drawbacks, Value-at-Risk (VaR) and Expected Shortfall (ES) remain popular measures of financial risk among practitioners. Hence, there is a further need for the development of more robust methods in estimating VaR and ES. In particular, this paper aims to improve current assumptions of appropriate underlying distributions to capture extreme tails, and as a result, improve the estimation of VaR and ES.

The implementation of VaR to identify appropriate regulatory capital requirement suffers from a number of setbacks, such as its inability to capture "tail loss". Such drawbacks have recently been highlighted by the Basel Committee on Banking Supervision. The committee has also recommended a shift of focus to the alternative ES measure to address the drawbacks of VaR. Although a number of operational challenges has been identified by the committee for the move to ES, it is believed that the benefit outweighs the disadvantages. The use of ES has been proposed for the internal model-based approach, and to be utilised in determining risk weights for the standardised approach (Basel, 2012). In this paper, both VaR and ES are implemented under the assumption of Extreme value theory (EVT). In addition, backtesting procedures are also conducted to analyse the model adequacy.

Formally, 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. ES measures the mean of losses that are equal to, or greater than, a corresponding VaR value. In order to capture the effect of market behavior under extreme events, EVT has been widely adopted in VaR estimation in recent years. Since EVT are derived from sound statistical theory and provides a parametric form for the tails of a distribution, its methodologies are attractive for risk assessments.

There is a large literature that studies EVT for risk measures in areas where extreme observations are of interest, such as finance, insurance, hydrology, climatology and engineering. Specifically, numerous studies in finance and commodity markets have been conducted using EVT, including Embrechts *et al.* (1997), Gençay & Selçuk (2004) and Gilli & Këllezi (2006). Byström (2005) applied EVT to the case of extreme large electricity prices and declared a good fit with the generalized Pareto distribution (GPD). Bali (2003) determined the type of asymptotic distribution for modeling the extreme changes in US treasury yields. He found that the thin-tailed Gumbel and exponential distributions perform worse than the fat-tailed Frechet and Pareto distributions. Marohn (2005) studied the tail index in the case of generalised order statistics and determined the asymptotic properties of the Frechet distribution. However, to the best of our knowledge, there are limited discussions on the application of EVT to the gold market, which represents a crucial commodity to the world economy. Large losses are the main concern in the field of financial risk management. For example, it may signify the situation of stock market crash. While a lion's share of the literature focuses on extreme losses, extreme gains in financial returns, on the other hand, are also of importance for financial leverage in the likes of security options and hedge funds.

In this paper, we extend some of the work by Jang (2007) and Chaithep *et al.* (2012). Specifically, we look at modeling both gains (i.e., positive returns) and losses (i.e., negative returns) for short and long position of trade, respectively, in the gold market, while utilizing graphical analyses (such as excess distribution plots, plot of the tail of underlying distribution and scatter plot of residuals) and various backtesting procedures (i.e., Kupiec test, Chritoffersen test and Bootstrap test) to draw robust conclusions on the adequacy of GPD models for VaR and ES estimates.

The remainder of the paper is organised as follows. In Section 2, we present a short literature review on the gold market. Section 3 describes the GPD and the peaks-over-threshold method for establishing extremes. Section 4 introduces the risk measures, and their corresponding backtesting procedures, that are utilised for this study. Empirical results obtained in GPD estimation are discussed in Section 5. Finally, Section 6 concludes the study.

## GOLD

Gold, as a financial indicator, is one of the most important commodities in the world and it is largely held by central banks. Central banks must maintain a proportion of their foreign exchange reserves in gold, as a store of value and as an assurance to redeem promises to pay depositors, note holders, or trading peers, or to secure a currency. Gold is also used by jewelers and investors as a hedging instrument (Sari *et al.*, 2010). When currencies devaluate, investors move to the gold market and when currencies revaluate investors move away from the gold market (Capie *et al.*, 2005).

Gold has an influence on other precious metals. Sari *et al.* (2010) states: "*Among the major precious metal class, an increase in the gold price seem to lead to parallel movements in the prices of the other precious metals which are also considered investment assets as well as industrial commodities*". The statement suggests that a model adequately explaining the gold prices could also contributes to models used in predicting the prices of other precious metals. Hence, many economists consider gold as a leading indicator in the precious metal pack.

The Bretton Woods system, for which the US dollar was expressed in terms of a fixed gold price, collapsed in 1971 (Capie *et al.*, 2005). Accordingly, it seems appropriate to start our investigation around this period. High inflation, uncertain international politics and low confidence in the US dollar are some of the main reasons advanced for the rapid increase in gold prices between September 1976 and January 1980. A combination of worries pushed investors to diversify their holdings of paper currencies into more tangible gold (Cheung & Lai, 1993). The swift increase in gold prices during 1980 was caused by technically driven trading in the futures market. The gold price reached US$700 for eleven days during 1980 but then returned to around US$300 by middle 1982. Between mid-1982 and June 2002, gold was seen trading in the US$250-US$500 range (Mills, 2004).

A number of authors have reported on the role gold plays as an inflation hedge and the role inflation plays on the gold price. However, according to Lawrence (2003), no significant correlations exist between returns on gold and changes in certain macroeconomic variables such as inflation, GDP and interest rates. Sjaastad & Scacciavillani (1996) reported that gold is a store value against inflation. Baker & van Tassel (1985) documented that the price of

gold depends on the future inflation rate. Sherman (1983) noted the log of the gold price is positively related to the anticipated inflation.

According to Kaufmann & Winters (1989), the price of gold is based on changes in the US rate of inflation, as well as other variables. Traditionally, gold has played a significant role during times of political and economic crises and during equity market crashes, whereby gold has responded with higher prices. According to Smith (2002), "*when the economic environment becomes more uncertain, attention turns to investigating in gold as a safe haven.*" The author also noted that following the September 11th, 2001, attack, the FTSE All share Index decreased by 9% while the London gold afternoon fixing price increased by 7.45%. Lawrence (2003) reported that gold returns are less correlated with returns on equity and bond indices than returns of other commodities. In line with gold's role as an asset last resort, Koutsoyiannis (1983) stated that the price of gold is strongly related to the state of the US economy and geopolitical factors.

The above motivations demonstrate the importance in measuring and capturing the stylised facts exhibits in the gold market prices. In particular, such prices display fatter tails and excess kurtosis (shown in Section 5), which cannot be fully captured by the widely exhausted Gaussian and Student's $t$ distributions.

## GPD AND EVT

The two-parameter GPD (with scale parameter $\beta$ and shape parameter $\xi$) has the following representation (Tsay, 2013):

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

where $x > 0$ when $\xi \geq 0$, $0 \leq x \leq -\beta/\xi$ when $\xi < 0$, and $\beta > 0$.

### Excess Distribution

For a random variable $X$, the excess distribution function $F_u$ above a certain threshold $u$ is defined as

$$F_u(x) = P(X - u \leq x \mid X > u) \quad (2)$$

where $x$ represent the size of exceedances over $u$. Furthermore, if we denote $F$ as the distribution function for $X$, then we may write

$$F_u(x) = \frac{F(x+u) - F(u)}{1 - F(u)} \quad (3)$$

A fundamental theorem in EVT, by Balkema & de Haan (1974) and Pickands (1975), identifies the asymptotic behavior of these exceedances with GPD. Hence, the excess distribution function $F_u$ can be well approximated by GPD for large enough $u$.

### Peaks Over Threshold (POT)

To fit a GPD to our data set, we adopt the peak over threshold (POT) method that focuses on the distribution of exceedances above some high threshold. For $x - u \geq 0$, we can rewrite the excess distribution function (3) as

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

and, hence, deduce the following reverse expression

$$F(x) = \big(1 - F(u)\big)F_u(x-u) + F(u) \tag{5}$$

which allows us to apply the POT method.

There are two steps in applying the POT method. Firstly, we need to choose an appropriate threshold. Secondly, fit the GPD function to data. Given the choice of a sufficiently high threshold, we may estimate $F(u)$ by $(1 - N_u/n)$, where *n* is the total sample size and $N_u$ is the amount of observations above the chosen threshold. And, $F_u(x-u)$ can be estimated by a GPD using maximum likelihood estimation (Embrechts *et al.*, 1997). We then obtain the following tail estimator (Ren & Giles, 2010)

$$\hat{F}(x) = 1 - \frac{N_u}{n}\left(1 + \frac{\hat{\xi}}{\hat{\beta}}(x-u)\right)^{-1/\hat{\xi}} \tag{6}$$

**Threshold Selection**

In this paper, we utilise the empirical mean excess plot for threshold selections. For a random variable *X*, the mean excess function is defined as

$$e(u) = E(X - u | X > u) \tag{7}$$

i.e., the mean of exceedances over a threshold *u*. If the underlying distribution of $X > u$ follows a GPD, then the corresponding mean excess is

$$e(u) = \frac{\beta + \xi u}{1 - \xi},\ \beta + \xi u > 0 \tag{8}$$

provided $\xi < 1$. From equation (8), we can clearly see that the mean excess function must be linear in *u*. More precisely, $X > u$ follows a GPD if, and only if, the mean excess function is linear in *u* (Coles, 2001). This gives us a way of selecting an appropriate threshold.

Given the data, we define the empirical mean excess function as

$$e_n(u) = \frac{\sum_i^n (X_i - u) I_{\{X_i > u\}}}{\sum_i^n I_{\{X_i > u\}}} \tag{9}$$

where *n* is the sample size. The empirical excess plot is a graphical representation of the locus of $(u, e_n(u))$ and we can examine this plot to choose the threshold *u* such that $e_n(u)$ is approximately linear for $X > u$.

**Parameter Estimation**

There are various techniques for estimating the parameters of the GPD, such as maximum likelihood estimation (MLE), method of moments and the method of probability-weighted moments. We adopt the MLE method in this paper because the maximum likelihood estimator is asymptotically normal and allows simple approximations for standard errors and confidence intervals (Azzalini, 1996).

Given that we have a sufficiently high threshold $u$ and, assuming there are $m$ observations with $x_i - u \geq 0$, the subsample $\{x_1 - u, \ldots, x_m - u\}$ has an underlying distribution of GPD, where $x_i - u \geq 0$ for $\xi \geq 0$, $0 \leq x_i - u \leq -\beta/\xi$ for $\xi < 0$, then the logarithm of the probability density function of $x_i$ can be derived from Equation (1) as

$$\ln f(x_i - u) = \begin{cases} -\ln(\beta) - \frac{1+\xi}{\xi}\ln\left(1 + \xi\left(\frac{x_i - u}{\beta}\right)\right) & \text{for } \xi \neq 0 \\ -\ln(\beta) - \frac{1}{\beta}(x_i - u) & \text{for } \xi = 0 \end{cases} \quad (10)$$

Hence, the log-likelihood function $L(\xi, \beta \,|x_i - u)$ for the GPD is the logarithm of the joint density of the $m$ observations, i.e.,

$$L(\xi, \beta \,|x_i - u) = \begin{cases} -n\ln(\beta) - \frac{1+\xi}{\xi}\sum_{i=1}^{m}\ln\left(1 + \xi\left(\frac{x_i - u}{\beta}\right)\right) & \text{for } \xi \neq 0 \\ -n\ln(\beta) - \frac{1}{\beta}\sum_{i=1}^{m}(x_i - u) & \text{for } \xi = 0 \end{cases} \quad (11)$$

Therefore, we can obtain the estimates for $\xi$ and $\beta$ by maximizing the log-likelihood function of the subsample under a suitable threshold $u$.

**Model Validation**

We can use quantile plots to assess the quality of a fitted generalised Pareto model (Coles, 2001). Assume we have a chosen threshold $u$, the ordered threshold excesses $y_{(1)} \leq \cdots \leq y_{(m)}$ and an estimated model $\hat{G}$ with $\hat{\xi} \neq 0$.

The quantile (Q-Q) plot consists of the pairs

$$\left\{\left(\hat{G}^{-1}\left(\frac{i}{m+1}\right), y_{(i)}\right) ;\ i = 1, \ldots, m\right\} \quad (12)$$

where

$$\hat{G}^{-1}(y) = u + \frac{\hat{\beta}}{\hat{\xi}}\left[y^{-\hat{\xi}} - 1\right] \quad (13)$$

If GPD is a reasonable fit for the exceedances above $u$, then the Q-Q plot should depict points that are approximately linear. Furthermore, we may confirm the goodness-of-fit of GPD by utilizing the excess distribution plot and plot of the tail of underlying distribution (McNeil *et al.*, 2005). For a good fit, the exceedances should lie close to the theoretical curves. Lastly, a scatter plot of residuals should not depict any visible pattern to indicate independence of the exceedances.

**RISK MEASURES**

The amount of asset risk capital, reserved by financial institutes as per Basel accords, is directly associated to the portfolio risk level and two of the most common benchmark measure for evaluating such risk are VaR and ES. VaR is intended to assess the maximum possible loss of a portfolio over a given time period, and its calculations focus on the tails of a distribution, whereas ES evaluates the expected value of losses (or gains) that exceed a corresponding VaR level. Hence, the accuracies of VaR and ES estimation are dependent on how well a selected model portrays the extreme data observations (McNeil *et al.*, 2005).

**VaR**

For a random variable $X$ (usually the return in some risky financial instrument) with distribution function $F$ over a specified time period, the VaR, for a given probability $p$, can be defined as the $p$-th quantile of $F$, i.e.,

$$\mathrm{VaR}_p = F^{-1}(1-p) \tag{14}$$

where $F^{-1}$ is the quantile function. VaR is a common measure of extreme risks and we use GPD to approximate this measure. In particular, using Equation (6) we obtain

$$\widehat{\mathrm{VaR}}_p = \begin{cases} u + \frac{\hat{\beta}}{\hat{\xi}}\left\{\left(\frac{n}{N_u}p\right)^{-\hat{\xi}} - 1\right\} & \hat{\xi} \neq 0 \\ u - \hat{\beta}\log\left(\frac{n}{N_u}(1-p)\right) & \hat{\xi} = 0 \end{cases} \tag{15}$$

where $\hat{\beta}$ and $\hat{\xi}$ are the maximum likelihood estimates of the GPD parameters (Tsay, 2010).

**ES**

Although VaR is often considered as an adequate risk measure, it does not capture all aspects of market risks, such as subadditivity. Hence, Artzner *et al.* (1999) proposed ES as a better measure of risk, which is subadditive and also informs us about the likely magnitude of exceedances. In contrast to VaR, ES measures the riskiness of an instrument by considering both the size and likelihood of losses above a particular threshold (Basel, 2012). ES gives the expected size of return that exceeds VaR, i.e., for a probability level $p$,

$$\mathrm{ES}_p = E\left(X \mid X > \mathrm{VaR}_p\right) \tag{16}$$

And, equivalently,

$$\mathrm{ES}_p = \mathrm{VaR}_p + E\left(X - \mathrm{VaR}_p \mid X > \mathrm{VaR}_p\right) \tag{17}$$

where the second term above represent the mean of the excess distribution $F_{\mathrm{VaR}_p}(x)$ (treating $\mathrm{VaR}_p$ as the threshold). Proceeding as before, if the threshold $\mathrm{VaR}_p$ is sufficiently large then $F_{\mathrm{VaR}_p}(x)$ is a GPD, i.e.,

$$F_{\mathrm{VaR}_p}(x) = G_{\xi,\beta+\xi(\mathrm{VaR}_p-u)}(x) \tag{18}$$

Thus, the mean of the excess distribution $F_{\mathrm{VaR}_p}(x)$ can be calculated as

$$\frac{\beta + \xi\left(\mathrm{VaR}_p - u\right)}{1-\xi} \tag{19}$$

where $\xi < 1$, and substituting into Equation (17) will yield

$$\widehat{\mathrm{ES}}_p = \frac{\widehat{\mathrm{VaR}}_p}{1-\hat{\xi}} + \frac{\hat{\beta} - \hat{\xi}u}{1-\hat{\xi}} \tag{20}$$

**Backtesting**

To examine the adequacy and effectiveness of VaR and ES estimates, we utilise various backtesting procedures. In particular, VaR backtesting is performed using the Kupiec likelihood ratio unconditional coverage test (Kupiec, 1995) and Christoffersen conditional coverage test (Christoffersen, 1998). While for ES, we follow the backtesting procedure in McNeil and Frey (2000), with and without bootstrapping.

The Kupiec test exploits the fact that an adequate model ought to have its proportion of violations of VaR estimates close to the corresponding tail probability level. The method consists of calculating the number of times $x^{\alpha}$ the observed returns fall below (for long positions) or above (for short positions) the VaR estimate at level $\alpha$, i.e., $r_t < VaR^{\alpha}$ or $r_t > VaR^{\alpha}$, and compare the corresponding failure rates to $\alpha$. The null hypothesis is that the expected proportion of violations is equal to $\alpha$. Under this null hypothesis, the Kupiec statistic, given by

$$LR_{UC} = 2\ln\left(\left(\frac{x^{\alpha}}{N}\right)^{x^{\alpha}} \cdot \left(1-\frac{x^{\alpha}}{N}\right)^{N-x^{\alpha}}\right) - 2\ln\left(\alpha^{x^{\alpha}}(1-\alpha)^{N-x^{\alpha}}\right) \quad (21)$$

is asymptotically distributed according to a chi-square distribution with one degree of freedom. The Christoffersen test extends the Kupiec test to account for serial independence of violations (i.e., clustering of extremes). The Christoffersen test statistic can be represented by

$$LR_{CC} = LR_{UC} + 2\ln\frac{\lfloor(1-\pi_0)^{\varphi_{00}}\pi_0{}^{\varphi_{01}}(1-\pi_1)^{\varphi_{10}}\pi_1{}^{\varphi_{11}}\rfloor}{\ln\lfloor(1-\pi)^{(\varphi_{00}+\varphi_{10})}\pi^{(\varphi_{01}+\varphi_{11})}\rfloor} \quad (22)$$

where $\varphi_{ij}$ is defined as the number of returns in state *i* while they have been in state *j* previously (state 1 indicates the VaR estimate is violated and state 0 indicates it is not) and $\pi_i$ is defined as the probability of having an exception that is conditional on state *i* the previous day. This statistic is asymptotically chi-square distributed with two degrees of freedom.

The null hypothesis of the ES backtest is that the excess conditional shortfalls (excess of the actual data series when VaR is violated), are i.i.d. and has zero mean. The test is a one sided t-test against the alternative that the excess shortfall has mean greater than zero and thus that the conditional shortfall is systematically underestimated. The test statistics is given by

$$T = \frac{\bar{r}-\mu_0}{\bar{\sigma}/\sqrt{m}} \quad (23)$$

where $\bar{r}$ and $\bar{\sigma}$ are the mean and standard deviation of "exceedance residuals" $\{r_1, r_2, \ldots, r_m\}$. The bootstrap techniques can also be utilised to alleviate any bias with respect to assumptions about the underlying distribution of the excess shortfall. For the bootstrap test, we sample $\{\tilde{r}_1^*, \tilde{r}_2^*, \ldots, \tilde{r}_m^*\}$ without replacement from the shifted residuals $\tilde{r}_i = r_i - \bar{r} + \mu_0$ and compute the test statistic

$$T_j^* = \frac{\bar{\tilde{r}}^* - \mu_0}{\bar{\tilde{\sigma}}/\sqrt{m}} \quad (24)$$

for each bootstrap sample *j* (McNeil & Frey, 2000).

## EMPIRICAL RESULTS AND DISCUSSIONS

### Stylised Facts

A number of facts about the volatility of financial assets have emerged over the years and have been confirmed in numerous studies. Hence, a good volatility model must be able to capture and reflect these stylized facts. These features are commonly found in financial and commodity markets. Financial returns are almost unpredictable; they have surprisingly large number of extreme values where both extremes and quiet periods are clustered in time. These features are often described as unpredictability, fat tails and volatility clustering (Engle, 2003).

**The Data**

The data used in this study are the monthly gold prices, quoted in US dollars, and is taken from the following website: www.gold.org/investment/statistics/gold_price_chart/. The data cover 515 observations from January 1969 to October 2012. The time series exhibits a number of price shocks, e.g., during the period around September 11, 2001, the beginning of the Iraq war in 2003 as well as the global crisis in 2008. The data are transformed into monthly log-returns by taking first backward differences in the logarithm of prices. For the observed gold prices $P_t$, the monthly log-returns are calculated using

$$r_t = \log P_t - \log P_{t-1} \tag{25}$$

Table 1 provides a summary of descriptive statistics for the considered return series.

**Table 1:** Descriptive Summary Statistics of Gold Returns

| Minimum | SD | Skewness | Kurtosis | JB statistic (*p*-value) | Maximum | Mean | N |
|---|---|---|---|---|---|---|---|
| -0.1100 | 0.0251 | 0.3769 | 6.7290 | 309.0970 (<0.0001) | 0.1099 | 0.0033 | 514 |

We observe that the mean of monthly returns is positive, indicating that the overall gold prices were increasing during the considered time period. The magnitude of the average return is very small compared to the standard deviation. Further, the large kurtosis of 6.7290 indicates the leptokurtic characteristics of returns. The series has a distribution with tails that are significantly fatter than those of the normal distribution. This indication of non-normality is also supported by the Jacque-Bera test, which rejects the null hypothesis of a normal distribution at all levels of significance.

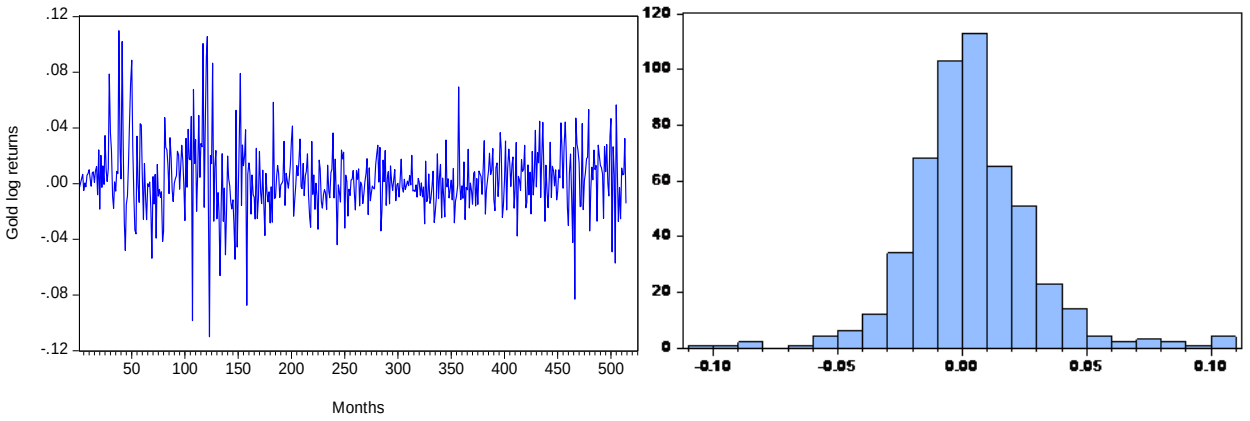


**Figure 1:** Time Series Plot And Histogram Of Gold Returns

Figure 1 provides a plot of the monthly log returns as well as a histogram of the returns distribution. The figures indicate heteroscedasticity and volatility clustering for the return series that also exhibits a number of isolated extreme occurrences caused by unforeseen events or shocks to the gold market.

The unpredictability of returns and volatility clustering can also be shown by observing the autocorrelations. Autocorrelations are correlations calculated between the value of a random variable today and its value in the past. Significant autocorrelations in returns indicate predictability, and volatility clustering is evidenced through the significance of squared or absolute returns. Figure 2 provides the autocorrelation function plot (ACF) of the returns and the ACF of squared returns. Clearly, the return autocorrelations are almost all insignificant while the squared returns have significant autocorrelations. Furthermore, the squared returns autocorrelations are all positive which is highly unlikely to occur by chance. The figures give significant evidence for both the unpredictability of returns and volatility clustering.

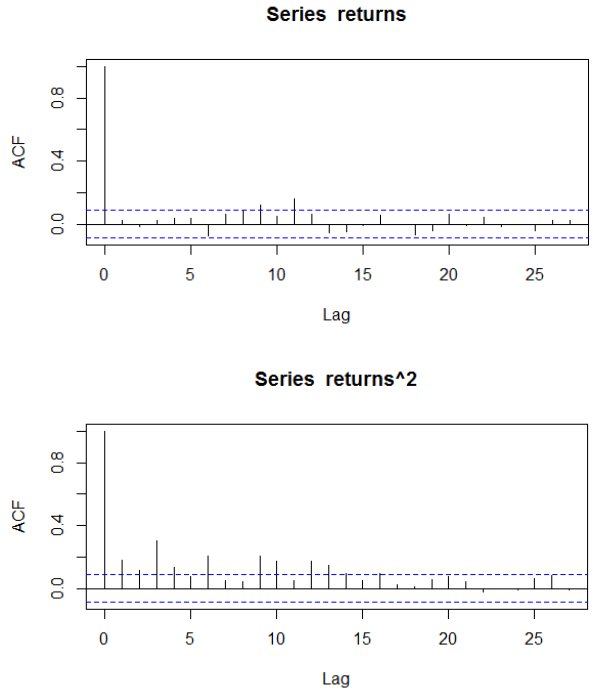


**Figure 2:** ACF Of Gold Returns (Left) And Squared Returns (Right)

The kurtosis is substantial at 6.7290; this is strong evidence that extremes are more substantial than one would expect from a normal random variable. Similar evidence is seen graphically in the Q-Q plot for gold returns (see Figure 2).

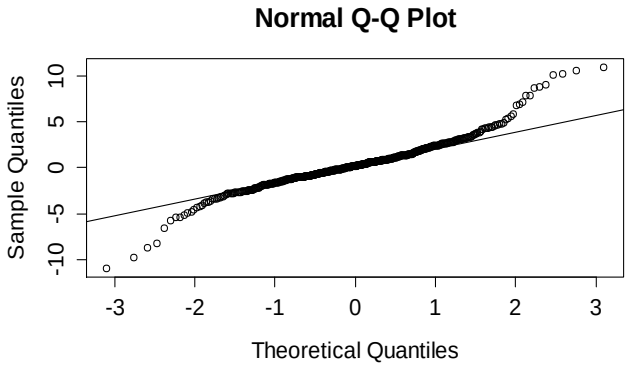


**Figure 3:** Normal Q-Q Plot For Gold Returns

We further test for stationarity of the return series using the Augmented Dickey-Fuller (ADF) and Philips Perron (PP) unit root tests. The ADF test is set to lag 0 using the Schwartz Information Criterion (SIC) and the PP test is conducted using the Bartlett Kernel spectral estimation method. Results are reported in Table 2 and indicate that the null hypothesis of unit root is rejected for both tests. Hence, the return series of gold prices can be considered to be stationary.

**Table 2:** Results For ADF And PP Unit Root Tests For Gold Return Series

| Unit root test | Test statistic | p-value |
|---|---|---|
| ADF test | -22.0778 | 0.0000 |
| Philips-Perron test | -22.0876 | 0.0000 |

For convenience of presentation, the data are now re-scaled as $r_t \to 100r_t$. Further, we produce all analogous results for negative returns by taking into account the relation $\min\{X_1, \ldots \ldots, X_n\} = \max\{-X_1, \ldots, -X_n\}$.

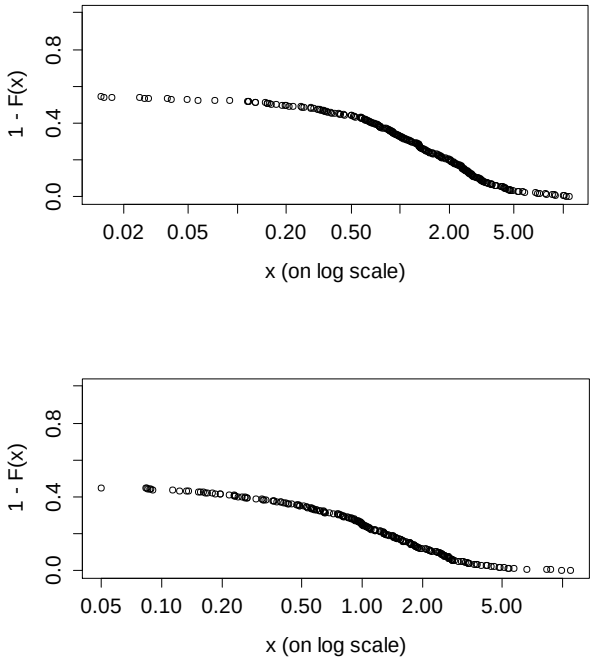


**Figure 4:** Empirical distribution for positive returns (left) and negative returns (right)

Figure 4 shows empirical distribution for both positive and negative returns. For both positive and negative returns, the tails are approximately linear, implying the Pareto behavior, hence we are justified in fitting GPD to the tails.

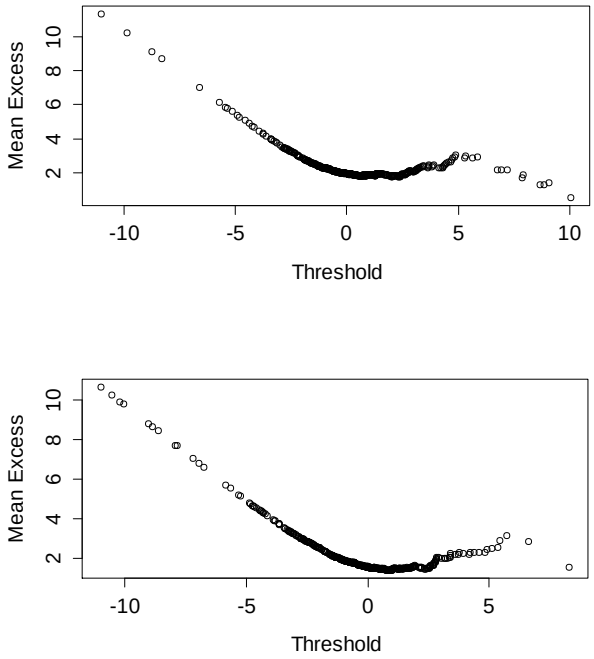


**Figure 5:** Mean Excess Function For Positive Returns (Left) And Negative Returns (Right)

Figure 5 shows the mean excess plot (mean residual life plot) of the positive monthly gold log-returns and negative monthly gold log-returns. A threshold of 2.5% seems to be reasonable for both positive (monthly gain) and negative (monthly loss) returns, i.e., 0.025 for positive log returns and -0.025 for negative log returns.
Table 3 shows the results of fitting a GPD to exceedances of positive and negative returns using thresholds of 0.025 (2.5%) and -0.025 (-2.5%), respectively.

**Table 3:** Results From Fitted GPD

| | **Positive Returns** | **Negative Returns** |
|---|---|---|
| Threshold | 2.5% | -2.5% |
| Percentile of the Threshold | 0.8560 | 0.9105 |
| Number of Points Exceeding Threshold | 74 | 46 |
| Estimates for $\xi$ | 0.2238 | 0.4347 |
| Standard Error of Estimates for $\xi$ | 0.1699 | 0.2503 |
| Estimates for $\beta$ | 1.4911 | 0.9392 |
| Standard Error of Estimates for $\beta$ | 0.3045 | 0.2661 |
| Variance-Covariance Matrix of Estimates | $\begin{bmatrix} 0.0289 & -0.0375 \\ -0.0375 & 0.0927 \end{bmatrix}$ | $\begin{bmatrix} 0.0627 & -0.0461 \\ -0.0461 & 0.0708 \end{bmatrix}$ |

Diagnostic plots for threshold excess model (GPD) fitted to monthly positive log returns and negative log returns are shown in Figures 6 and 7.

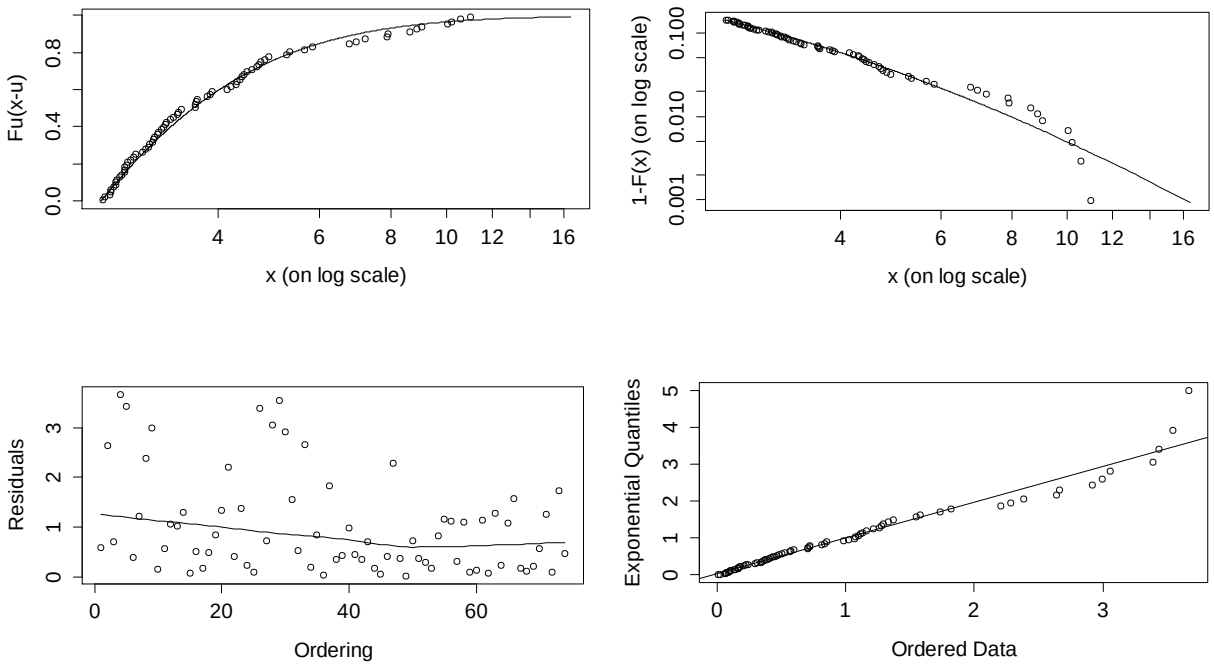


**Figure 6:** Excess Distribution (Top Left), Tail Of Underlying Distribution (Top Right), Scatterplot Of Residuals (Bottom Left) And Q-Q Plot (Bottom Right) For GDP With 74 Exceedances For *Positive* Returns

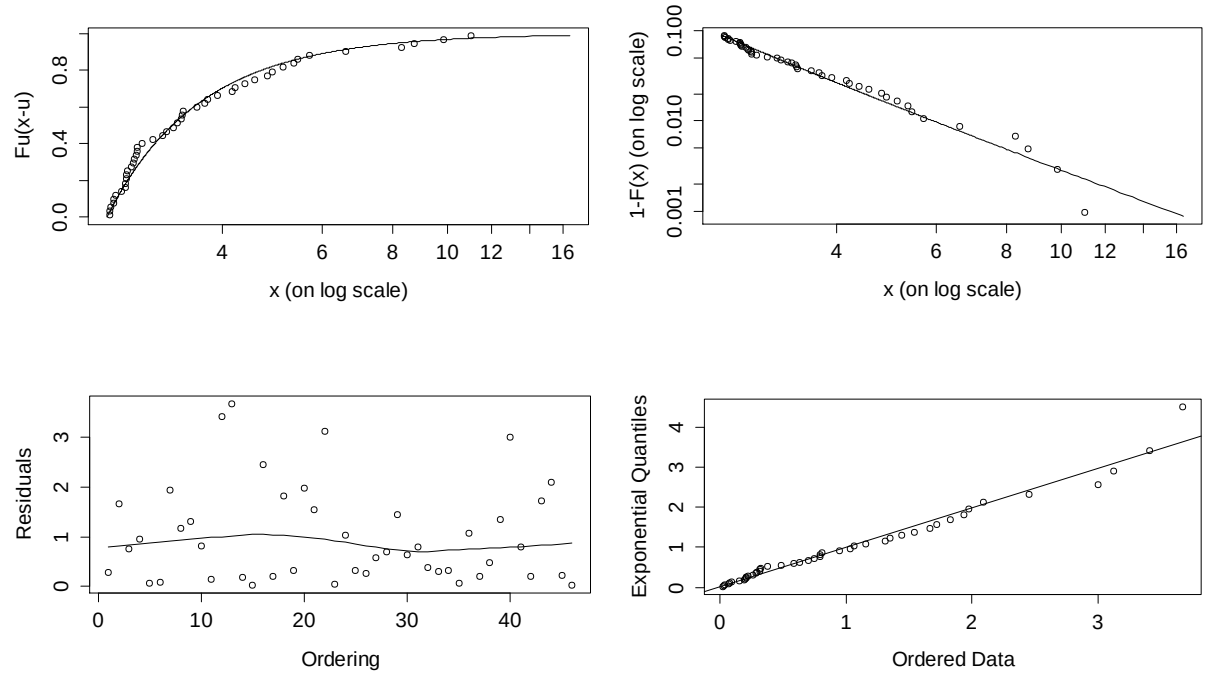


**Figure 7:** Excess Distribution (Top Left), Tail Of Underlying Distribution (Top Right), Scatterplot Of Residuals (Bottom Left) And Q-Q Plot (Bottom Right) For GPD With 46 Exceedances For *Negative* Returns

For positive returns, the graphs of the excess distribution and tail of underlying distribution follow the traces of the corresponding GPD, implying that the GPD model provides a good fit to exceedances in the upper tail of our data. This is further confirmed by the approximate linearity in the Q-Q plot. Thus, the positive extreme values (beyond 0.025) can be modeled by

$$G_x = 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi}$$

where $\xi$ =0.1699 and $\beta$=1.4911.

In Figure 7, the shape of the excess distribution graph corresponds closely to the shape of a GPD (Embrechts *et al.*, 1997) and the Q-Q plot is approximately linear. Hence, we again confirm that the GPD is a good fit for exceedances in the lower tail of our data. This implies that the negative extreme values (beyond -0.025) can be modeled by:

$$G_x = 1 - \left(1 + \frac{\xi x}{\beta}\right)^{-1/\xi}$$

where $\xi$ = 0.4347 and $\beta$= 0.9392.

Table 4 provides the estimates of VaR and ES for both positive and negative returns, at various quantiles levels. The table presents the estimates constructed from the fitted GPD model and these are contrasted against the estimates drawn from the traditional Gaussian model and the Student's t model.

**Table 4:** Estimates For Var And ES For Positive And Negative Returns

| | | **Positive returns** | | **Negative returns** | |
|---|---|---|---|---|---|
| Model | *p*-values | Estimate of VaR | Estimate of ES | Estimate of VaR | Estimate of ES |
| Normal | 0.9 | 3.5402 | 5.8512 | 2.8831 | 4.8983 |
| | 0.95 | 4.4507 | 6.9612 | 3.7936 | 6.0645 |
| | 0.99 | 6.1586 | 8.7573 | 5.5015 | 8.3640 |
| Student's t | 0.9 | 3.7972 | 6.1173 | 3.1401 | 5.1250 |
| | 0.95 | 4.9257 | 7.9494 | 4.2686 | 6.5182 |
| | 0.99 | 7.4025 | 9.3525 | 6.7454 | 9.4670 |
| GPD | 0.9 | 3.0662 | 5.1506 | 2.3982 | 3.9814 |
| | 0.95 | 4.2793 | 6.7135 | 3.1222 | 5.2620 |
| | 0.99 | 7.9399 | 11.4298 | 5.9411 | 10.2482 |

At a quantile level of 90%, the estimated VaR from GPD is 3.0662 for gains and 2.3982 for losses. This is, with the GPD model, we are 90% confidence that the expected market value of gold would not gain by more than 3.0662% for the best case scenario or lose more than 2.3982% for the worst case scenario, within one-month durations. For GPD, VaR is estimated as 7.9399% at the 99$^{th}$ percentile for the right tail, i.e., we expect the monthly changes in the value of gold would not increase by more than 7.9399%. Given the quantile levels, the corresponding VaR estimates in the right tail are larger than those in left tail. The GPD estimates of ES under different quantile levels exhibits analogous characteristics as observed from VaR. Similar interpretations can be made for the Gaussian model and the Student's t model.

In comparison of different models, it is also interesting to note that GPD produced lower VaR and ES estimates than the two other models, at both 90% and 95% quantile levels. However, the GPD estimates at the 99% level are higher than those from the normal and Student's t models, except for the VaR estimate of the negative tail.

**Table 5:** Backtesting Of Var For Gold Returns

| | | ***p*-values for Kupiec test** | | | ***p*-values for Chritoffersen test** | | |
|---|---|---|---|---|---|---|---|
| Model | Level | 0.9 | 0.95 | 0.99 | 0.9 | 0.95 | 0.99 |
| Normal | Gains | 0.0266 | 0.5783 | 0.0095 | 0.0131 | 0.2069 | 0.0029 |
| | Losses | <0.001 | 0.0354 | 0.7103 | <0.001 | 0.0653 | 0.8694 |
| Student's t | Gains | 0.0040 | 0.0354 | 0.1219 | 0.0073 | 0.0262 | 0.1019 |
| | Losses | <0.001 | 0.0047 | 0.5990 | <0.001 | 0.0130 | 0.8439 |
| GPD | Gains | 0.9530 | 0.6461 | 0.4345 | 0.0870 | 0.4758 | 0.1535 |
| | Losses | 0.9530 | 0.9517 | 0.9503 | 0.9977 | 0.8407 | 0.9501 |

**Table 6:** Backtesting Of ES For Gold Returns

| | | **Boot *p*-value** | | | ***p*-value** | | |
|---|---|---|---|---|---|---|---|
| | Level | 0.9 | 0.95 | 0.99 | 0.9 | 0.95 | 0.99 |
| Normal | Gains | 0.5213 | 0.5185 | 0.5110 | 0.5000 | 0.5000 | 0.5000 |
| | Losses | 0.5126 | 0.5174 | 0.4786 | 0.5000 | 0.5000 | 0.5000 |
| Student's t | Gains | 0.5084 | 0.4962 | 0.5013 | 0.5000 | 0.5000 | 0.5000 |
| | Losses | 0.5192 | 0.5285 | 0.5177 | 0.5000 | 0.5000 | 0.5000 |
| GPD | Gains | 0.5420 | 0.6679 | >0.999 | 0.5213 | 0.6965 | >0.999 |
| | Losses | 0.6406 | 0.6498 | 0.9255 | 0.6486 | 0.6935 | 0.9663 |

Tables 5 and 6 provide the results of backtesting of VaR and ES estimates from different models. Both the Kupiec test and the Christoffersen test suggest that the VaR estimates from GPD cannot be rejected. In particular, our model seems to produce very suitable VaR estimate for long positions (as indicated by high *p*-values for losses at all levels and for both tests). As for backtesting of ES, results from both tests, with and without bootstrapping, are presented. Again, the high *p*-values indicate very suitable ES estimates from GPD, at all levels. Furthermore, at all quantile levels and for both tails, GPD produced the highest *p*-value for all tests. These are strong evidence that GPD is a more adequate model for VaR and ES estimations, as compared to the traditional Gaussian model and the Student's t model.

## CONCLUSION

In this paper, we have illustrated the use of EVT to model tail-related risk measures, such as VaR and ES, for the gold market**.** In particular, GPD was found to be an appropriate model to describe the conditional excess distributions of a heteroscedastic gold log return series and provides adequate estimations for VaR and ES. These were confirmed by various statistical graphical analyses and backtesing procedures. Moreover, the superior performances of GPD were contrasted against the normal distribution and the Student's t distribution.

Further work may include comparative analyses with other heavy-tail distributions, that are suitable for the depiction of financial returns, and incorporation of GPD in the framework of the well-known GARCH-based VaR models. For example, comparisons may be drawn with the generalised logistic distribution (Tolikas & Brown, 2006) and the class of generalised hyperbolic distributions (Huang *et al.*, 2014), with the inclusion of backtesting results on the VaR and ES estimates.

*R* and *EViews* were used in this paper to produce figures and results from various tests.

## AUHTOR INFORMATION

**Knowledge Chinhamu** is a lecturer in Statistics at the University of KwaZulu-Natal and a member of the South African Statistical Association. His research interests lie in financial time series and econometric modeling.
Email: chinhamu@ukzn.ac.za

**Chun-Kai Huang** is a lecturer in Statistics at the University of Cape Town. He is a member of the South African Statistical Association and a fellow of the Cambridge Commonwealth Society. His research interests lie in probabilistic exchangeability, moment problems, extreme value theory and statistical applications to financial data.
Email: chun-kai.huang@uct.ac.za (Contact author)

**Chun-Sung Huang** is a lecturer in Finance and an associate of the African Collaboration for Quantitative Finance and Risk Research (ACQuFRR) at the University of Cape Town. His research interests lie in volatility modeling and forecasting, Value-at-Risk (VaR) modeling and derivatives pricing in incomplete markets.
Email: chun-sung.huang@uct.ac.za

**Delson Chikobvu** is a senior lecturer in Mathematical Statistics and Actuarial Science at the University of Free State. He is a member of the South African Statistical Association. His research interests lie in econometrics, mathematical finance, decision sciences and energy forecasting.
Email: chikobvu@ufs.ac.za

**REFERENCES**

1. Artzner, P., Delbaen, F., Eber, J-M., & Heath, D. (1999). Coherent measures of risk. *Mathematical Finance*, 9(3), 203-228.
2. Azzalini, A. (1996). *Statistical Inference Based On the Likelihood*. London: Chapman & Hall.
3. Baker, S. A., & van Tassel, R. C. (1985). Forecasting the price of gold: A fundamentalist approach. *Atlantic Economic Journal*, 13(4), 43-51.
4. Bali, T. G. (2003). An extreme value approach to estimating volatility and value at risk. *The Journal of Business*, 76(1), 83-108.
5. Balkema, A., & de Haan, L. (1974). Residual life time at great age. *Annals of Probability*, 2, 792-804.
6. Basel (2012). Basel Committee on Banking Supervision - Fundamental review of the trading book. Bank for International Settlements. Retrieved from http://www.bis.org/publ/bcbs219.pdf
7. Byström, H. N. E. (2005). Extreme value theory and extremely large electricity price changes. *International Review of Economics & Finance*, 14(1), 41-55.
8. Capie, F., Mills, T. C., & Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money*, 15(4), 343-352.
9. Chaithep, K., Sriboonchitta, S., Chaiboonsri, C., & Pastpipatkul, P. (2012). Value at risk analysis of gold price returns using extreme value theory. *The Empirical Econometrics and Quantitative Economics Letters*, 1(4), 151-168.
10. Cheung, Y. W., & Lai, K. S. (1993). Do gold market returns have long memory? *Financial Review*, 28(2), 181-202.
11. Christoffersen, P. F. (1998). Evaluating interval forecasts. *International Economic Review*, *39*(4), 841–862.
12. Coles, S (2001). *An Introduction to Statistical Modeling of Extreme Values.* London: Springer-Verlag.
13. Embrechts, P., Klüppelberg, C., & Mikosh, T. (1997). *Modelling Extremal Events: For Insurance and Finance.* Berlin: Springer-Verlag.
14. Engle, R. F. (2003). *Risk and Volatility: Econometric Models and Financial Practice*. Nobel Lecture, New York University.
15. Gençay, R., & Selçuk, F. (2004). Extreme value theory and value-at-risk: Relative performance in emerging markets. *International Journal of Forecasting*, 20(2), 287-303.
16. Gilli, M., & Këllezi, E. (2006). An application of extreme value theory for measuring financial risk. *Computational Economics*, 27(2-3), 207-228.
17. Huang, C-K., Chinhamu, K., Huang, C-S., & Hammujuddy, J. (2014). Generalized hyperbolic distributions and Value-at-Risk estimation for the South African Mining Index. *International Business & Economics Research Journal*, 13(2), 319-328.
18. Jang, J-B. (2007). An extreme value theory approach for analyzing the extreme risk of the gold prices. *Journal of Financial Review*, 6, 97-109.
19. Kaufmann, T. D., & Winters, R. A. (1989). The price of gold: a simple model. *Resources Policy*, 15(4), 309-313.
20. Koutsoyiannis, A. (1983). A short-run pricing model for a speculative asset, tested with data from the gold bullion market. *Applied economics*, 15(5), 563-581.
21. Kupiec, P. H. (1995). Techniques for verifying the accuracy of risk management models. *Journal of Derivatives*, 3(2), 73–84.
22. Lawrence, C. (2003). *Why is gold different from other assets? An empirical investigation*. World Gold Council. London, UK.

23. Marohn, F. (2005). Tail index estimation in models of generalized order statistics. *Communications in Statistics: Theory & Methods*, 34(5), 1057-1064.
24. McNeil, A. J., & Frey, R. (2000). Estimation of tail-related risk measures for heteroscedastic financial time series: an extreme value approach. *Journal of Empirical Finance*, 7(3-4), 271–300.
25. McNeil, A. J., Frey, R., & Embrechts, P. (2005). *Quantitative Risk Management: Concepts, Techniques, and Tools.* Princeton: Princeton University Press.
26. Mills, T. C. (2004) Statistical analysis of daily gold price data. *Physica A: Statistical Mechanics and its Applications*, 338(3-4), 559-566.
27. Pickands, J. (1975). Statistical inference using extreme order statistics. *Annals of Statistics*, 3(1), 119-131.
28. Ren, F., & Giles, D. E. (2010). Extreme value analysis of daily Canadian crude oil prices. *Applied Financial Economics*, 20(12), 941-954.
29. Sari, R., Hammoudeh, S., & Soytas, U. (2010). Dynamics of oil prices, precious metal prices, and exchange rate. *Energy Economics*, 32(2), 351-362.
30. Sherman, E. J. (1983). A gold pricing model. *The Journal of Portfolio Management*, 9(3), 68-70.
31. Sjaastad, L. A., & Scacciavillani, F. (1996). The price of gold and the exchange rate. *Journal of International Money and Finance*, 15(6), 879-897.
32. Smith, G. (2002). Tests of the random walk hypothesis for London gold prices. *Applied Economics Letters*, 9(10), 671-674.
33. Tolikas, K., & Brown, R. A. (2006). The distribution of the extreme daily share returns in the Athens Stock Exchange. *The European Journal of Finance*, 12(1), 1-22.
34. Tsay, R. S. (2010). *Analysis of Financial Time Series* (3rd ed). New Jersey: Wiley & Sons.
35. Tsay, R.S. (2013). *An Introduction to Analysis of Financial Data in R.* New Jersey: Wiley & Sons.