University of Victoria | Economics

Econometrics Working Paper EWP1704

ISSN 1485-6441

RISK ANALYSIS FOR THREE PRECIOUS METALS: AN APPLICATION OF EXTREME VALUE THEORY

Qinlu Chen
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2
&
David E. Giles
Department of Economics, University of Victoria
Victoria, B.C., Canada V8W 2Y2

Revised, August 2017

Abstract

Gold, and other precious metals, are among the oldest and most widely held commodities used as a hedge against the risk of disruptions in financial markets. The prices of such metals fluctuate substantially, introducing risks of their own. This paper's goal is to analyze the risk of investment in gold, silver, and platinum by applying Extreme Value Theory to historical daily data for changes in their prices. The risk measures adopted in this paper are Value at Risk and Expected Shortfall. Estimates of these measures are obtained by fitting the Generalized Pareto Distribution, using the Peaks-Over-Threshold method, to the extreme daily price changes. The robustness of the results to changes in the sample period, threshold choice, and distributional assumptions, are discussed. Our results show that silver is the most risky metal among the three considered. For negative daily returns, platinum is riskier than gold; while the converse is true for positive returns.

Keywords: Precious metals; extreme values; portfolio risk; value-at-risk; generalized Pareto distribution

JEL Classifications: C46 ; C58 ; G10; G32

Author Contact:
David E. Giles 58 Rock Lake Court, L'Amable, ON, Canada K0L 2L0; E-mail: dgiles@uvic.ca; Voice: (613) 332-6833

1. Introduction

Gold, and other precious metals, have long been held by investors as a hedge against the instability of international financial markets. Typically, the price of precious metals increases in times of actual, or perceived, international political unrest. Often, such price movements will run counter to movements in equity prices, allowing investors to spread the risk associated with their portfolios.

The prices of gold, silver, and other precious metals fluctuate substantially over time and day-to-day, and this introduces risks of their own. For example, while the US economy was recovering during 2013, gold prices kept decreasing. On April 15th of that year, the price of gold dropped 9.6 percent in one day as a result of the decreasing inflation rate, and the increasing real interest rates in the US. According to data from the London Bullion Market between 1968 and 2014, the daily percentage change in the gold price was as high as 12.5% (on 3 January 1980) and as low as -14.2% (on 22 January of that year). The Soviet invasion of Afghanistan and the Iranian revolution of late 1979 and early 1980 motivated people to buy gold, which drove up gold prices.

Conversely, the Hunt brothers' failed attempt to corner the silver market, and the US Federal Reserve's new policy that raised interest rates dramatically to about 20 percent, pushed gold prices downward because of panic selling. These fluctuations had a direct impact on international financial markets and therefore on the business cycles of the major economies, and the livelihood of their populations. Surprisingly, there has been very little formal empirical analysis of the risk of holding gold and other precious metals as assets in their own right, or as part of a portfolio. Exceptions include Jang (2007), Dunies *et al*. (2010), and Trück and Liang (2012). As in the present paper, Jang used extreme value theory to analyze the risk of holding gold. In contrast, Trück and Liang applied a threshold ARCH model to gold data from the London bullion market for the period 1999 - 2008, and Dunis *et al*. analyzed data from 2002 - 2008 using a neural network model. These methods are quite different from that adopted in this paper.

Regulators and supervisors of financial institutions have been struggling for years to raise public awareness of risk control in investment activities. The Basel II Accord, an international agreement on banking regulation, sets a minimum capital requirement for banks according to the risk forecasts calculated by the banks on a daily basis. The Basel II Accord was widely implemented in many countries including the US, Canada, Australia, and the EU. The risk measure adopted by the Basel Committee to quantify the market or operational risks is Value at Risk (VaR). Value at Risk is the maximum loss/gain at a low probability (usually one per cent) for a certain time horizon. For example, if a trader whose

portfolio has a 1 percent VaR of $1 million in one day, it means there is a 1 percent probability that the trader will lose $1 million or more "overnight". VaR was first brought to the public's attention by JP Morgan as its internal risk measure in its publication, Riskmetrics (1996), and it became widely accepted as a basic risk measure in financial markets after the Basel II Accord adopted VaR as a preferred risk measure in the late 1990s. There is a considerable literature relating to this financial risk measure. For example, see Jorion (1996), Dowd (1998), and Duffie and Pan (1997). Another common risk measure that is often used as an alternative to VaR is Expected Shortfall (ES). Expected Shortfall is the average loss/gain given that VaR already has been exceeded. ES is considered as a coherent alternative to VaR since VaR is non-subadditive and may have a misleading effect on portfolio diversification (Artzner *et al*., 1997, 1999). This paper estimates the values of both VaR and ES for daily changes in the prices of gold, silver, and platinum.

In order to measure extreme risk and to be prepared for irregular losses, we are interested in the behavior of the tail of the distribution of price changes. Most conventional models take the normality assumption for granted and consider the tail part of a distribution as outlier events. In this paper, we employ a well-developed statistical method that models low frequency, high severity events, the Extreme value theory (EVT). EVT provides a firm theoretical foundation for analyzing those rare events that have serious consequences. EVT identifies the limiting distribution of the maxima of a random variable. Those exceedances, the values above a specified high threshold, must be generated by the Generalized Pareto Distribution (GPD) (Balkema and de Haan, 1974; Pickands, 1975). When studying extreme events, EVT plays a role that is analogous to that played by the Central Limit Theorem in the study of sums of a random variable. There has been much research related to the application of Extreme Value Theory to risk management and financial series, for example, the work of Embrechts *et al*. (1999), Gilli (2006), and McNeil (1999). EVT has been applied to many other markets, such as those for crude oil and agricultural products (*e.g*., Giles and Ren, 2007; and Odening and Hinrichs, 2003). Some of the pitfalls of EVT are discussed in Diebold, Schuermann and Stroughair (2000).

In this paper, we use the Peaks-Over-Threshold (POT) method, and the GPD, to model the extreme risks associated with the daily price returns for gold, silver, and platinum. Section 2 presents the framework of EVT and the POT methodology. Section 3 introduces the standard risk measures - VaR and ES. Then we discuss our data, the process of the tail estimation, and the computation of estimates of our risk measures in section 4. The robustness of our risk estimates to the choice of sample, and to the choice of threshold in the POT analysis, is considered in section 5. The last section summarizes the empirical results and briefly discusses some directions for future research.

2. Models and Methods

2.1 Extreme Value Theory

In order to avoid systematic risk, regulators and supervisors of large financial institutions are concerned about the heavy tails of the time series for returns on financial assets. Many conventional models fail to model those irregular events properly. The past literature has discussed the superiority of EVT to other approaches, such as the GARCH model, historical simulation, the variance-covariance method, and Monte Carlo simulation. The EVT-based VaR is more robust than other model-based VaR - see Paul and Barnes (2010), and Gençay *et al*. (2003) for more details. Avdulaj (2011) found that the historical simulation method tends to overestimate the VaR, while the variance-covariance method tends to underestimate it. Two different methods are used to model extreme events. One is the block maxima method which involves the Generalized Extreme Value distribution; and another one is the POT method which involves the GPD. The block maxima method chooses the maximum values of a variable during successive periods to constitute the extreme events, and it is based on the Fisher–Tippet theorem (Fisher and Tippett, 1928; Gnedenko, 1943). The latter result ensures that the normalized maxima of the blocks of data follow one of the Fréchet, Weibull and Gumbel distributions. The block maxima method requires the data to be i.i.d. and they must converge to a non-degenerate distribution function. The POT method models the behavior of exceedances over a given high threshold, and requires that individual data-points be availabe. EVT implies that the limiting distribution of the exceedances (not the original data) is the GPD (Pickands, 1975; Balkema and de Haan, 1974). Some recent contributions to the theory of the GPD include those of Oakes and Dasu (1990), Asadi *et al*. (2001), and Tavangar and Asadi (2008, 2012).

Previous research has indicated that the GPD uses the data more efficiently and effectively than does the block maxima method (*e.g*., Jang; 2007, Gilli, 2006; and Allen *et al*., 2013). Unnecessarily dividing the data into artificial blocks ignores the fact that each block may have a different characteristics. In some blocks, all of the values could be much smaller than in most blocks, and in other blocks, all the values might be quite large compared to the whole sample. It is inefficient to artificially "block" the data. Accordingly, in this paper we focus on the POT method, as the original daily data are readily available.

2.2 Generalized Pareto Distribution

An implication of EVT is that the maxima of a random variable above certain high threshold should be generated from a GPD. The distribution of the exceedances is presented by a Conditional Excess Distribution F_u , defined as

$$F_u(y) = P(X - u \leq y | X > u) \; ; \; for \; 0 \leq y \leq x_F - u \quad . \tag{1}$$

X is a random variable, u is a given threshold, $y = x - u$ are the exceedances, and $x_F < \infty$ is the right endpoint of the support of the unknown population distribution, *F*, of X.

Equation (1) can be rewritten in terms of *F*:

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

For large enough u, the Conditional Excess Distribution $F_u(y)$ is approximated increasingly well by the GPD (Balkands and de Haan, 1974; Pickands, 1975):

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

where the two-parameter GPD in terms of y takes the form:

$$G_{\xi,\sigma}(y) = \begin{cases} 1 - (1 + \frac{\varepsilon}{\sigma} y)^{-1/\xi} & if \; \xi \neq 0 \\ 1 - e^{-y/\sigma} & if \; \xi = 0 \end{cases} \tag{3}$$

for $y \in [0, (x_F - u)], \; if \;\; \xi \geq 0$; and $y \in \left[0, -\frac{\xi}{\sigma}\right], \; if \;\; \xi < 0$.

Let $x = u + y$, after reorganizing the equations we can get a three-parameter GPD in terms of x:

$$G_{u,\xi,\sigma}(x) = \begin{cases} 1 - (1 + \frac{\varepsilon}{\sigma}(x - u))^{-1/\xi} & if \; \xi \neq 0 \\ 1 - e^{-(x-u)/\sigma} & if \; \xi = 0 \end{cases} \quad . \tag{4}$$

Here, *u* is the threshold, ξ is the shape parameter, σ is the scale parameter.

The shape parameter, ξ, determines the heaviness of the tail of the distribution. The larger the shape parameter is, the heavier the tail will be. The shape parameter ξ can be positive, negative or equal to zero. If $\xi < 0$, the Conditional Excess Distribution $F_u(y)$ has an upper bound. If $\xi \geq 0$, the corresponding distribution has unbounded support and a fat tail. The latter case is usually the one of interest when modelling financial data, and indeed this is the case in our study.

2.3 Peaks Over Threshold Method

The POT methodology is a desirable approach to analyze extreme risks because it is based on a sound statistical theory, and it offers a parametric form for the tail of a distribution. The POT method focuses on

the exceedances above a specified high threshold. First, we need to determine the proper threshold. Then, with the given threshold we can fit the GPD to our data. The parameter estimates are computed by the method of maximum likelihood. There are two plots that help with the selection of thresholds. One is the Sample Mean Excess Plot and another one is the Shape Parameter Plot. So far, there is no algorithm based method available for the selection of the threshold u. Many researchers have analyzed this issue, but none have provided a convincing solution.

1. Sample Mean Excess Plot

For a random variable X with right end point x_F, its mean excess function is defined as:

$$e(u) = E(X - u|X > u) \tag{5}$$

for $u < x_F$. If the underlying distribution is a GPD, then the corresponding (population) mean excess function is:

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

for $\sigma + \xi u > 0$, and $\xi < 1$. (*e.g.*, Cole, 2001, p.79.)

As we can see from (6), the mean excess function is *linear* in the threshold u when the exceedances follow GPD. This important property can help with the selection of the threshold value in practice.

The *empirical* mean excess function is

$$e_n(u) = \frac{\sum_{i=k}^{n}(x_i^n - u)}{n-k+1}, \;\; k = min\{\, i \mid x_i^n < u\} \tag{7}$$

where n - k + 1 is the number of observations over the threshold u. The sample mean excess plot is the locus of $(u, e_n(u))$, $x_1 < u < x_n$, and an inspection of this plot facilitates the choice of threshold value. Specifically, we seek a threshold beyond which the empirical mean excess function is roughly upwards, and linearly, sloped.

2. Parameter Plot

We use the method of maximum likelihood to estimate the parameters. The log-likelihood function, based on a sample of n observations, is

$$L(\xi,\sigma|y) = \begin{cases} -nlog\sigma - (\frac{1}{\xi}+1)\sum_{i=1}^{n} log(1+\frac{\xi}{\sigma}y_i) & if\ \xi \neq 0 \\ -nlog\sigma - \frac{1}{\sigma}\sum_{i=1}^{n} y_i & if\ \xi = 0 \end{cases} \quad . \tag{8}$$

When choosing the value of the threshold, there is a trade-off between the bias and variance of the maximum likelihood estimates of the shape and scale parameters. We need a large value of u for the EVT to hold, and to minimize bias. But as the threshold gets larger, there will be fewer observations in the tail, and this reduces the efficiency of estimation. If the threshold is too low, the conditional excess distribution function will not converge to that of the GPD. It is important to investigate the robustness of the results to the choice of the threshold, u. For a detailed discussion relating to threshold selection see Matthys and Beirlant (2000), Embrechts *et al*. (1997), and McNeil *et al*. (2010).

Above a certain threshold, the exceedances should approach the GPD, so the estimated values of the parameters should be roughly constant. The so-called parameter plot is a graph of the maximum likelihood estimates of the shape and scale parameters for various choices of the thresholds. We choose the threshold beyond which the estimates of parameters become unstable.

3. Risk Measures

3.1 Value at Risk and Expected Shortfall

We consider two standard choices of risk measurement in this paper. One is Value at Risk (VaR), and the other one is Expected Shortfall (ES). These two risk measures involve the estimation of extreme quantiles of the underlying distribution.

VaR quantifies the maximum loss /gain occurring over a given time-period, with a specified (low) probability. VaR is often calculated at the 99th percentile over a one-day or ten day-period. In this paper, we compute the VaR associated with one day returns. Let X be a random variable with continuous distribution function F, VaR_p is $(1-p)^{\text{th}}$ percentile of the distribution F : $VaR_p = F^{-1}(1-p)$

In our analysis, the VaR can be defined as a function of the parameters of the GPD. Re-organizing (2), we get:

$$F(x) = (1-F(u))F_u(y) + F(u) \quad . \tag{9}$$

Replacing F_u with the GPD, and $F(u)$ with $(\frac{n-N_u}{n})$, where n is the sample size and N_u is the number of observations above the threshold, we have:

$$F(x) = \frac{N_u}{n}\left(1 - \left(1 + \frac{\hat{\xi}}{\hat{\sigma}}(x-u)\right)^{-1/\hat{\xi}}\right) + \left(1 - \frac{N_u}{n}\right) \quad . \tag{10}$$

After simplifying and inverting (10), for a given percentile, *p*, we have

$$\hat{x}_p = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{N_u}p\right)^{-\hat{\xi}} - 1\right) \quad . \tag{11}$$

Since VaR is a quantile of the distribution, we can obtain the formula for the VaR estimator:

$$\widehat{VaR_p} = u + \frac{\hat{\sigma}}{\hat{\xi}}\left(\left(\frac{n}{N_u}p\right)^{-\hat{\xi}} - 1\right) \quad . \tag{12}$$

ES describes the expected size of the return exceeding VaR. It is a *conditional* mean, given that VaR is exceeded, and is defined as $ES_p = E(X|\, X > VaR_p)$.

By invariance, the maximum likelihood estimator of ES is:

$$\widehat{\mathrm{ES}}_{\mathrm{p}} = \widehat{VaR_p} + E\left(X - \widehat{VaR_p} | X > \widehat{VaR_p}\right) \quad . \tag{13}$$

The second term in equation (13) is the expected value of the exceedances above the VaR_p, which is the mean excess function of $F_{VaR_p}(x)$ (see equation (5)). The mean excess function for the GPD is given in (6) in section 2.3. Thus, we get:

$$F_{VaR_p}(x) = \frac{\hat{\sigma} + \hat{\xi}(\widehat{VaR_p} - u)}{1 - \hat{\xi}} \quad . \tag{14}$$

Substituting (12) into (13) yields the formula for the ES estimator:

$$\widehat{ES_p} = \widehat{VaR_p} + \frac{\hat{\sigma} + \hat{\xi}(\widehat{VaR_p} - u)}{1 - \hat{\xi}} = \frac{\widehat{VaR_p}}{1 - \hat{\xi}} + \frac{\hat{\sigma} - \hat{\xi}u}{1 - \hat{\xi}} \quad . \tag{15}$$

After we estimate the tail distribution using the GPD, we can calculate these risk measures by inserting the GPD parameter estimates into the above formulae.

3.2 Interval Estimation of VaR and ES

As the expressions for both the VaR and ES are non-linear in the parameters, we use the delta method (*e.g.*, Oehlert, 1992) to calculate asymptotic standard errors for these measures. The asymptotic normality and invariance of maximum likelihood estimators imply that the estimates of VaR and ES are

also asymptotically normal in distribution. So, the construction of asymptotically valid confidence intervals is straightforward. Although most authors report only *point* estimates of VaR and ES, Ren and Giles (2007) also use the delta method in a similar context to obtain confidence intervals.

Using the delta method, the estimated asymptotic variance of VaR can be calculated as:

$$\widehat{var}\left(\widehat{VaR}_p\right) \approx \left(\hat{d}'\hat{C}\hat{d}\right) = \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix}' \begin{pmatrix} \hat{c}_{11} & \hat{c}_{12} \\ \hat{c}_{21} & \hat{c}_{22} \end{pmatrix} \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix}$$

$$= \hat{d}_1{}^2 \hat{c}_{11} + \hat{d}_2{}^2 \hat{c}_{22} + \hat{d}_1 \hat{d}_2 (\hat{c}_{12} + \hat{c}_{21}) \quad , \qquad (16)$$

where $\hat{d} = \begin{pmatrix} \hat{d}_1 \\ \hat{d}_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial \widehat{VaR}_p}{\partial \xi} \\ \frac{\partial VaR_p}{\partial \beta} \end{pmatrix}$, $\hat{C} = \begin{pmatrix} \hat{c}_{11} & \hat{c}_{12} \\ \hat{c}_{21} & \hat{c}_{22} \end{pmatrix} = \widehat{cov}\left(\hat{\xi}, \hat{\beta}\right)$, $\hat{c}_{12} = \hat{c}_{21}$

$\hat{C}$ is the estimated variance-covariance matrix of the (estimated) shape and scale parameters, and $\hat{d}$ is obtained by noting that:

$$\hat{d}_1 = -\frac{\hat{\beta}}{\hat{\xi}^2}\left(\left(\frac{N_u}{np}\right)^{\hat{\xi}} - 1\right) + \frac{\hat{\beta}}{\hat{\xi}}\left(\frac{N_u}{np}\right)^{\hat{\xi}} ln\left(\frac{N_u}{np}\right)$$

$$\hat{d}_2 = \frac{1}{\hat{\xi}}\left(\left(\frac{N_u}{np}\right)^{\hat{\xi}} - 1\right).$$

The asymptotic standard error for the estimated VaR is

$$s.e.\left(\widehat{VaR}_p\right) = \sqrt{\widehat{var}\left(\widehat{VaR}_p\right)} \quad . \qquad (17)$$

As the VaR and ES maximum likelihood estimators are asymptotically normal, a 95% confidence interval is constructed as

$$\left(\widehat{VaR}_p - 1.96 * s.e.\left(\widehat{VaR}_p\right) \ , \ \widehat{VaR}_p + 1.96 * \ s.e.\left(\widehat{VaR}_p\right)\right) \quad . \qquad (18)$$

We compute the asymptotic standard error and 95% confidence interval for ES in the same way.

4. Data Characteristics and Application

4.1 Data Characteristics

The data that we use for gold is *London Fixings p.m. Gold Price* from the London Bullion Market Association (LBMA)[1]. Fixing levels are in USD per troy ounce. The sample period is from 1 April, 1968 to 8 January, 2014. Of the 11,495 gold price returns, 5,447 are negative and 6,048 are positive.

For silver, we use *LBMA Silver Price: London Fixings*, from London Bullion Market Association (LBMA)[2]. Again, fixing levels are set in troy ounces. The prices are from 2 January, 1968 to 14 March, 2014, and are in USD. We have 11,680 silver daily returns in total, of which 5,968 are positive and 5,712 are negative. In our sample, the data for some dates are missing. In these cases we converted data that are available in GBP to USD[3].

For platinum, we choose the *Johnson Matthey Base Prices*, London 8:00 a.m., which are the company's quoted selling prices. The price is for metal in sponge form, with minimum purities of 99.95% for platinum[4]. Again, prices are in USD per troy ounce. The time horizon is the longest available – namely, from 1 July, 1992 to 24 March, 2014 (N = 5,578). There are 3,079 positive returns and 2,499 negative returns. However, when comparing the risks of holding various commodities, we should keep in mind that these risks could be affected by the choice of time period. For a small number of dates, data are unavailable for the London market. Because all of the prices provided by Johnson Matthey are in USD, we simply use the platinum prices from the New York and Hong Kong markets in these cases.

Daily price data for each precious metal are shown in Figures 1 to 3, and these are converted to daily returns by taking log-differences. We have modelled the positive returns and the (absolute values of the) negative returns for each metal separately, as there is no prior justification for assuming that the risk is symmetric, and there is recent evidence to suggest asymmetry (Blose and Gondhalekar, 2014).

[1]The data were retrieved on 9 January, 2014 from: http://www.quandl.com/OFDP-Open-Financial-Data-Project/GOLD_2-LBMA-Gold-Price-London-Fixings-P-M
[2] Data retrieved on 15 March 2014 from http://www.quandl.com/OFDP/SILVER_5-LBMA-Silver-Price-London-Fixings
[3] The exchange rate was obtained from http://fxtop.com/en/historical-exchange-rates.php
[4] The data and data description are from http://www.quandl.com/JOHNMATT/PLAT-Platinum-Prices

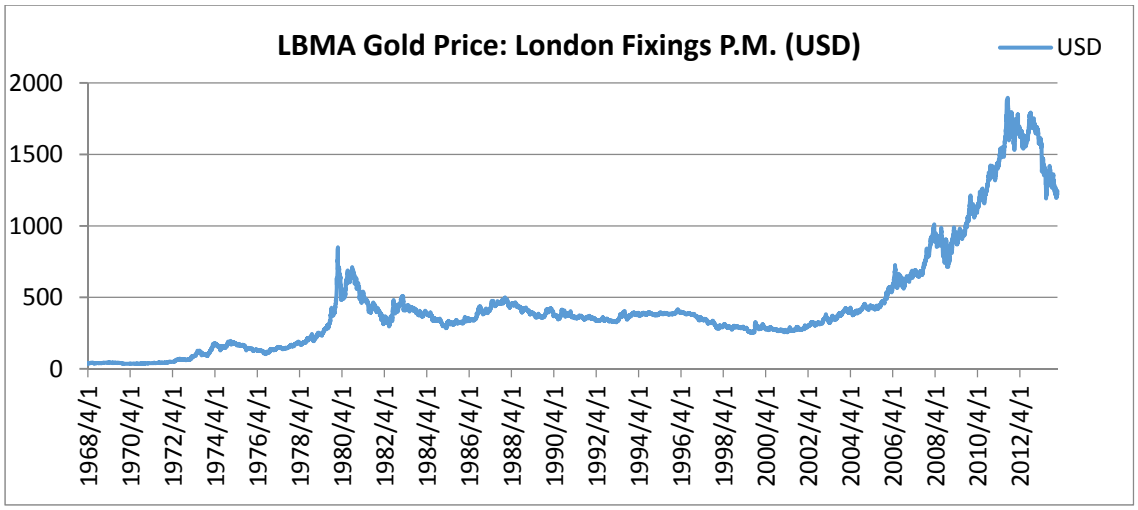


Figure 1: Daily Gold Prices from the London Bullion Market for the period 1 April, 1968 to 8 January, 2014.

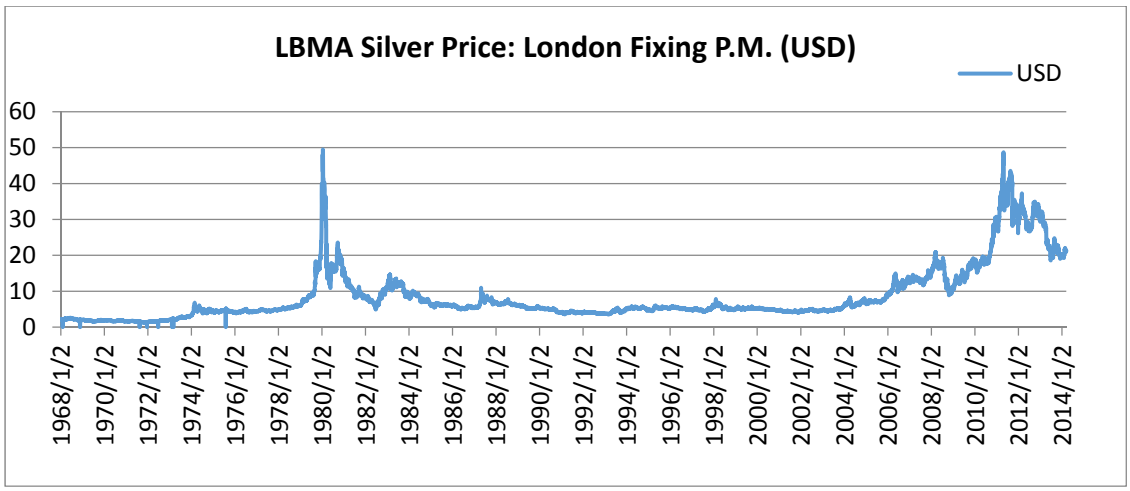


Figure 2: Daily Silver Prices from the London Bullion Market for the period 2 January, 1968 to 14 March, 2014.

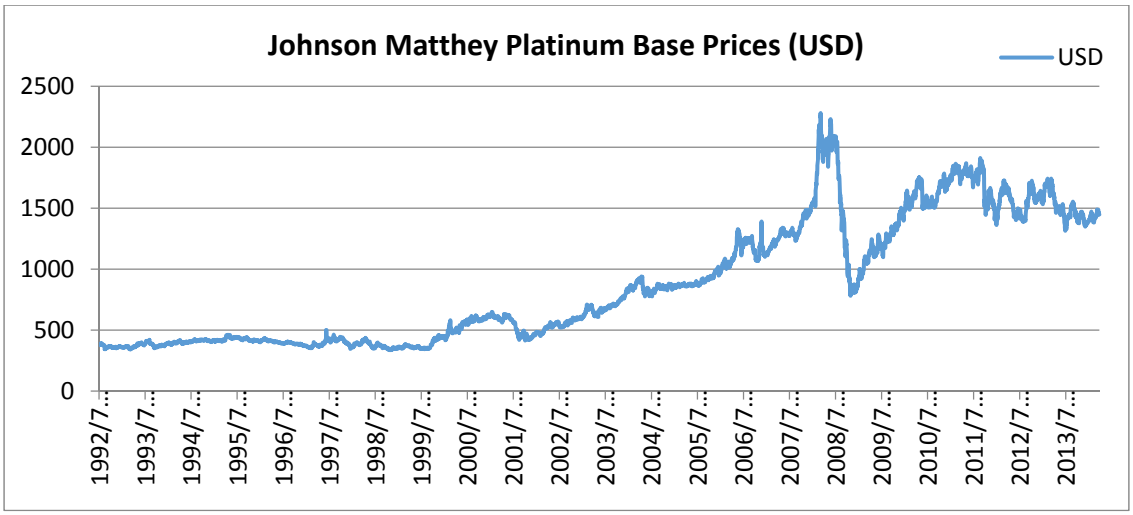


Figure 3: Daily Platinum Prices are the Johnson Matthey Base Prices, London 8:00 a.m. for the period 1 July, 1992 to 24 March, 2014.

	Mean	Max.	Min.	t-stat.	Skewness	Kurtosis	J-B Test	ADF
Gold	0.03%	12.50%	-14.20%	0.02	0.10	14.38	62048.84	-109.93
Silver	0.02%	31.18%	-25.75%	0.01	-0.06	19.23	128271.90	-14.86
Platinum	0.02%	13.93%	-15.54%	0.02	-0.31	13.79	27145.25	-74.33

Table 1: Descriptive statistics and basic tests of daily gold, silver, and platinum returns.

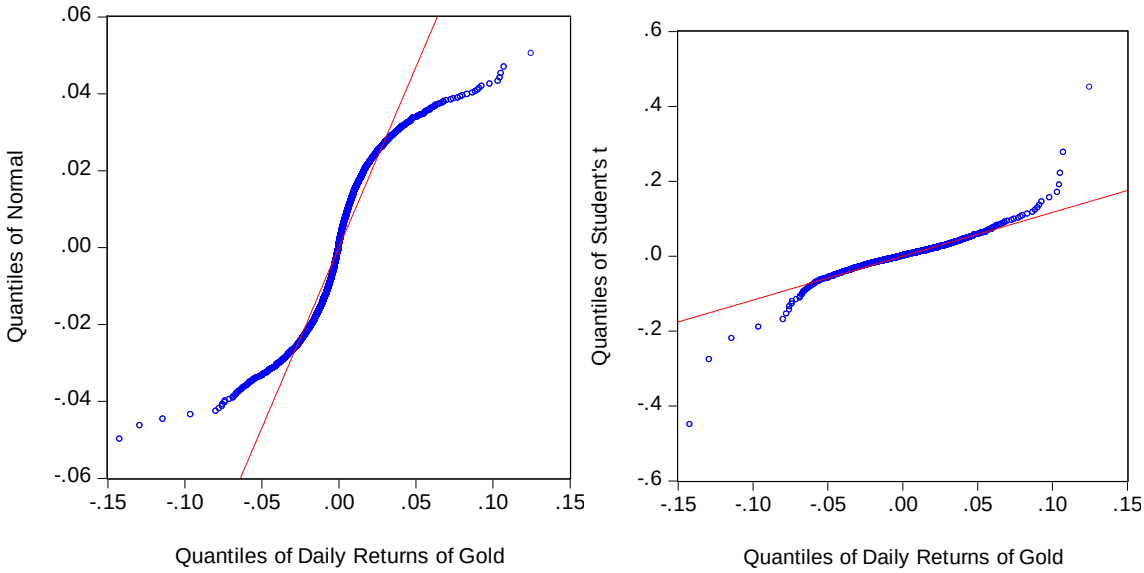


Figure 4: QQ plots of daily gold returns applied to normal distribution (left panel) and Student-t distribution (right panel).

The stationarity of the returns time-series is assessed by means of the augmented Dickey-Fuller (ADF) test (with no drift or trend), where the augmentation level is chosen by minimizing the Schwartz information criterion. For our large sample sizes the (one-sided negative) 1% critical value is -2.566. The results in Table 1 indicate a very clear *rejection* of the hypothesis that the returns series are non-stationary.

Also in Table 1, we see that the mean daily returns for each metal are essentially zero, and the t-statistics for testing this null hypothesis are so small that zero means cannot be rejected at any reasonable significance level. Further, the daily returns are slightly skewed, and they have high kurtosis. The Jarque-Bera (J-B) test statistics, which are asymptotically chi-square with two degrees of freedom under the null hypothesis of normality, clearly indicate that the daily returns are non-normal for each metal. Indeed, the quantile-quantile plots in Figure 4 show that the gold price returns data are not well explained by either a normal distribution or a Student-t distribution (with estimated degrees of freedom). The same is true for the other two metals under study. In fact, many researchers have found that an analysis based on the normal distribution will underestimate the VaR – a point to which we return in section 5.1. This suggests that, instead of considering those tail observations as outliers, we need a proper method to address the fat tails and model those extreme events. EVT is designed to deal with this situation.

The application of this theory requires that the extremes data are independent. In particular it is important to check that the "exceedances" exhibit these properties before estimating the parameters of the GPD distribution in the context of the Peaks-Over-Threshold analysis. The independence of the various returns exceedances series has been verified by inspection of the associated correlograms and Ljung-Box Q statistics. We have also used the more general independence test of Brock *et al*. (1997). This test is applicable for a range of *nonlinear* multi-dimensional alternatives, and we have bootstrapped the p-values for the test statistic to allow for the modest sample sizes associated with the exceedances series. These results, for two and three embedding dimensions, appear in Table 2. The results for higher dimensions are very similar, and in all cases clearly support the null hypothesis of independence.

We discuss the time sensitivity of our estimated risk measures in section 5. In some parts of the following discussion we use gold as an illustrative example, to conserve space.

	Positive Returns			Negative Returns		
	u	**BDS-2**	**BDS-3**	**u**	**BDS-2**	**BDS-3**
	(N)	**(p)**	**(p)**	**(N)**	**(p)**	**(p)**
Gold	0.030	2.080	1.959	0.022	-0.237	-0.333
	(210)	(0.225)	(0.239)	(398)	(0.648)	(0.578)
Silver	0.018	-0.084	(0.946)	0.038	-0.238	-0.280
	(1554)	-0.121	(0.882)	(391)	(0.872)	(0.894)
Platinum	0.030	-0.206	-0.251	0.025	-0.105	-0.107
	(119)	(0.951)	(0.991)	(180)	(0.638)	(0.590)

Table 2: Tests of independence of the excess returns series using the BDS test with 2 and 3 embedding dimensions. u = threshold value, determined in section 4. Sample sizes and bootstrapped p-values appear in parentheses.

4.2 Determination of the Threshold

As was discussed in section 2.3, there are two plots that can help us to determine the threshold level that is central to the Peaks-Over-Threshold analysis. Figure 5 presents the mean excess (ME) plots and parameter plots for positive and negative gold daily returns. The corresponding plots for the silver and platinum daily returns are available upon request. These results, and the associated maximum likelihood estimation results below, were obtained using the POT package (Ribabet, 2006) for the R statistical environment.

For the ME plot, the upper and lower dashed lines constitute confidence bands at an approximate 95% level. The ME plots are not very helpful in our case. We are looking for a point where the plot starts to be linear and upward sloping, but both positive and negative returns have positive slopes under all thresholds. Therefore, we focus primarily on the parameter plots to determine the thresholds.

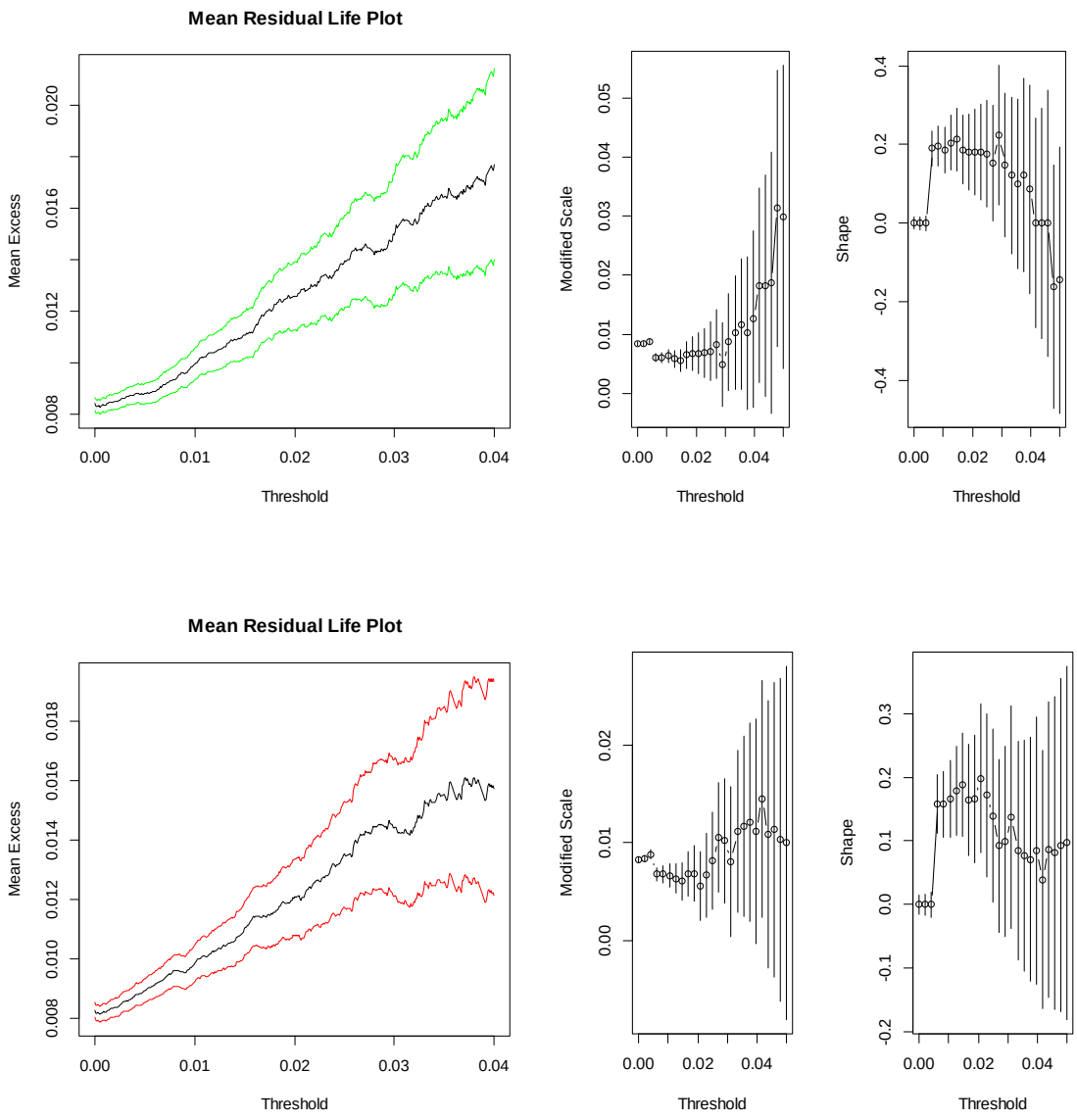


Figure 5: ME plots of positive returns (top left) and negative returns (bottom left). Parameter plots for positive returns (top right) and negative returns (bottom right).

On the basis of the plots in Figure 5, for positive returns, we choose a threshold u = 0.030; for negative returns, we choose u = 0.022. As we will see in Section 4.4, the estimated risk measures are quite robust to the choice of threshold, u. As long as the threshold is within a proper range so that the exceedances above the threshold follow GPD, the estimates of VaR and ES are quite stable. We use corresponding plots to determine the thresholds for the silver and platinum daily returns. For silver, we

choose thresholds u = 0.018 for positive returns and u = 0.038 for negative returns. For platinum, based on the two plots, we choose thresholds u = 0.030 for positive returns and u = 0.025 for negative returns.

4.3 Parameter Estimation

Given the thresholds selected in the previous section, we can estimate the shape and scale parameters of the corresponding GPD. Again, the POT package for R is used to implement maximum likelihood estimation, based on (8). Table 3 summarizes the parameter estimates and their asymptotic standard errors. In the case of gold, we illustrate the sensitivity of the parameter estimates to different threshold choices. The preferred thresholds, based on Akaike's information criterion (AIC), are indicated with asterisks. These results are also statistically more significant than the others shown.

Figure 6 provides a comparison of the empirical Cumulative Distribution Function (CDF) against the theoretical CDF for the GPD, for each of the metals' returns. The theoretical CDFs are computed using the threshold values and parameter estimates in Table 3. All of the graphs show that the GPD models the tails of the corresponding empirical distributions extremely well.

Maximum Likelihood Estimates: Gold

	Parameter Estimates For Positive Returns		Parameter Estimates For Negative Returns	
Threshold	u = 0.030*	u = 0.038	u = 0.022*	u = 0.042
No. Exceedances	210	112	398	78
$\hat{\xi}$	0.1848	0.1054	0.1689	0.0563
(s.e.)	(0.0910)	(0.1267)	(0.0616)	(0.1078)
$\hat{\sigma}$	0.0123	0.0153	0.0105	0.0156
(s.e.)	(0.0014)	(0.0024)	(0.0008)	(0.0024)
AIC	-1347.096	-684.839	-2695.506	-480.997

(*continued*)

Maximum Likelihood Estimates: Silver

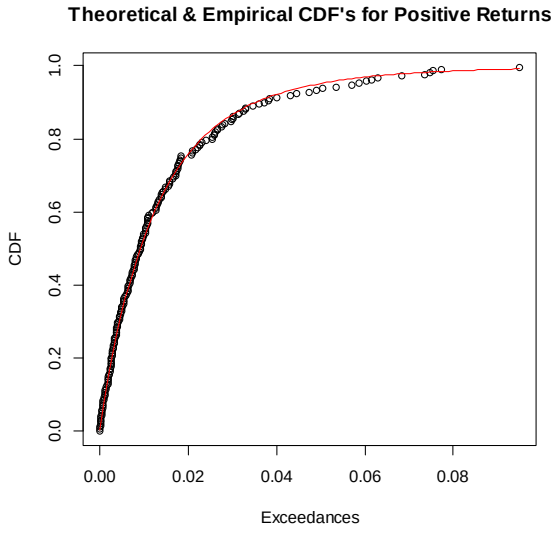
	Parameter Estimates For Positive Returns	Parameter Estimates For Negative Returns
Threshold	u = 0.018	u = 0.038
No. Exceedances	1554	391
$\hat{\xi}$	0.1980	0.3140
(s.e.)	(0.0282)	(0.0740)
$\hat{\sigma}$	0.0130	0.0174
(s.e.)	(0.0005)	(0.0015)
AIC	-9776.956	-2140.080

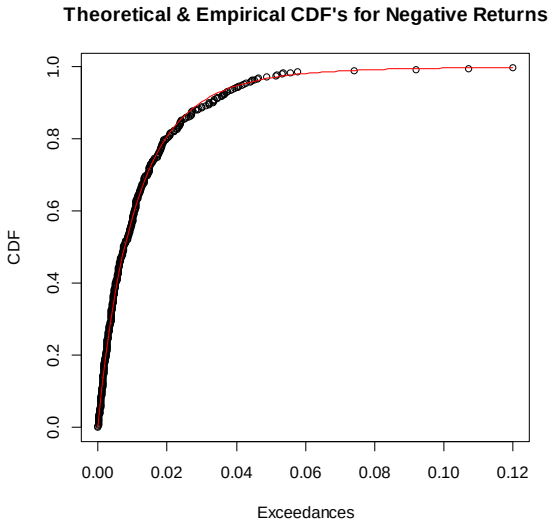
Maximum Likelihood Estimates: Platinum

	Parameter Estimates For Positive Returns	Parameter Estimates For Negative Returns
Threshold	u = 0.030	u = 0.025
No. Exceedances	119	180
$\hat{\xi}$	0.1692	0.2829
(s.e.)	(0.0965)	(0.1058)
$\hat{\sigma}$	0.0101	0.0099
(s.e.)	(0.0013)	(0.0012)
AIC	-812.765	-1197.287

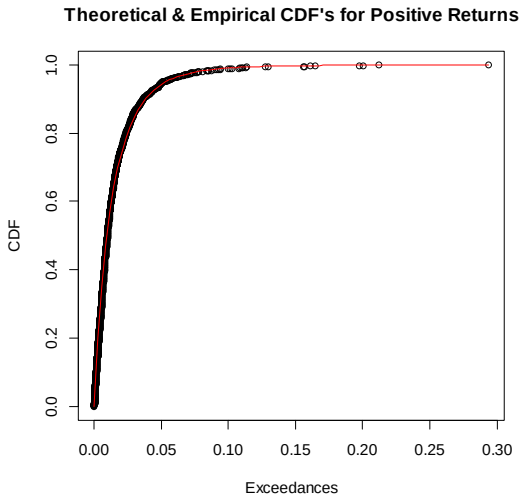
Table 3: Maximum likelihood parameter estimates for gold, silver, and platinum.

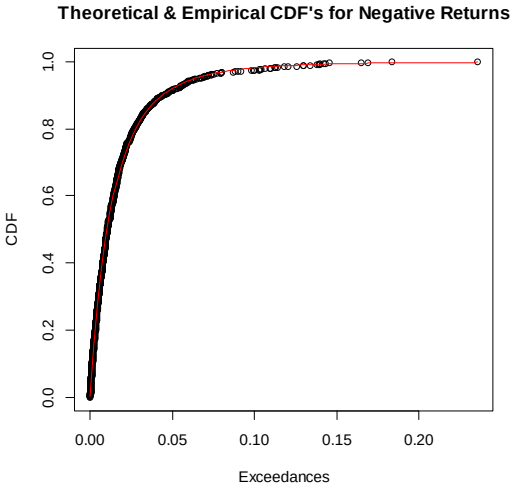
CDF for Gold Daily Returns





CDF for Silver Daily Returns





CDF for Platinum Daily Returns

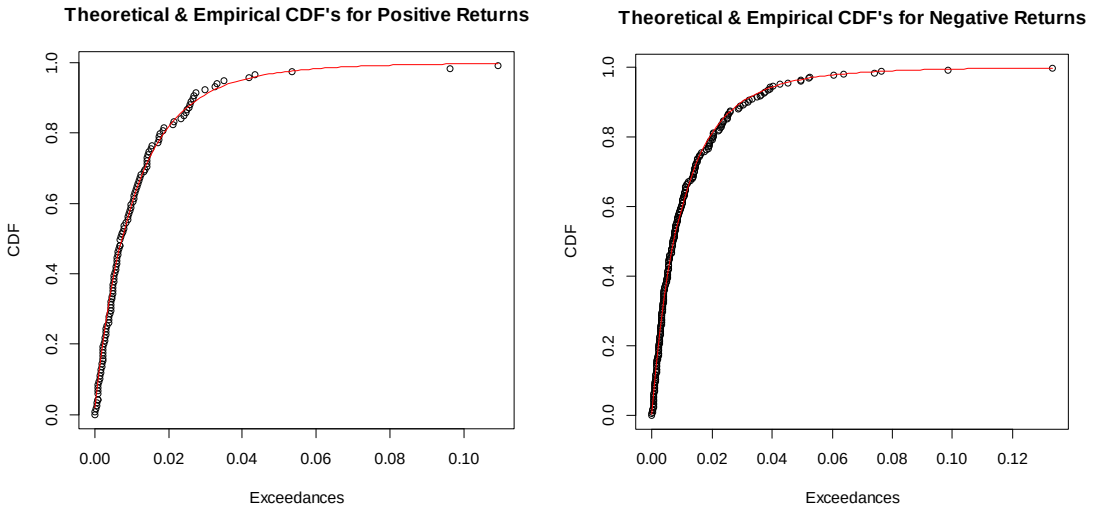


Figure 6: Empirical Cumulative Distribution Function (CDF) against the theoretical CDF for Positive Returns (left panel) and negative returns (right panel).

4.4 Risk Measures Estimation

The ultimate use of risk measures, such as VaR, is to help set risk adjusted minimum capital requirements to protect financial institutions from irregular, large losses. VaR and ES analyze the worst case scenario: if things go wrong, how wrong could they go? In this section we present both point and interval estimation results for gold, silver and platinum. Point estimates of VaR and ES are calculated at the conventional 99th percentile (*i.e.*, 1% VaR and 1% ES), using the formulae in section 3.1. The results are presented in Table 4.

For positive gold returns, with 1% probability (at the 99th percentile), the daily return for the gold price could exceed 4.72%, and the average return above this level will be 6.61%. For negative returns, with 1% probability, the daily return for the gold price could fall below -4.68%, and the average return below this level will be -6.45%. That means a trader holding a $1million position in gold faces a 1% chance of losing $46,800 or more "overnight". If such an event occurred, the expected loss would be $64,500.

For silver, the point and interval estimates of 1% VaR and 1% ES are also presented in Table 4. For positive returns, with 1% probability, the daily return for silver price could exceed 7.74%, and the average return above this level will be 10.82%. For negative returns, with 1% probability, the daily return for silver price could fall below -8.38%, and the average return below this level will be -13.01%.

For platinum, we see that for positive returns, there is 1% probability that the daily return for platinum price could exceed 4.53%, and the average return above this level will be 6.06%. For negative returns, there is 1% probability that the daily return for platinum price could fall below -5.11%, and the average return below this level will be -7.52%.

Point and Interval Estimates - Positive Returns

	Percentile	**VaR**	**95% CI Lower**	**95% CI Upper**
Gold	99th	0.0472	0.0444	0.0500
Silver	99th	0.0774	0.0726	0.0821
Platinum	99th	0.0453	0.0422	0.0484
	Percentile	**ES**	**95% CI Lower**	**95% CI Upper**
Gold	99th	0.0661	0.0587	0.0736
Silver	99th	0.1082	0.0972	0.1191
Platinum	99th	0.0606	0.0527	0.0684

Point and Interval Estimates - Negative Returns

	Percentile	**VaR**	**95% CI Lower**	**95% CI Upper**
Gold	1st	-0.0468	-0.0440	-0.0496
Silver	1st	-0.0838	-0.0778	-0.0899
Platinum	1st	-0.0511	-0.0462	-0.0561
	Percentile	**ES**	**95% CI Lower**	**95% CI Upper**
Gold	1st	-0.0645	-0.0571	-0.0718
Silver	1st	-0.1301	-0.1078	-0.1524
Platinum	1st	-0.0752	-0.0583	-0.0921

Table 4: Point and interval estimates of VaR and ES for gold, silver and platinum.

Table 4 summarizes all of the estimates of risk measure at the 99th percentile. Some evidence relating to the 95th percentile (*i.e.*, 5% VaR and 5% ES) is given in section 5.3. The 95% confidence intervals are obtained using the delta method, described earlier. The results show that silver is the most

risky metal among the three. For negative returns, platinum is riskier than gold. For positive returns, gold is riskier than platinum. As investors and financial regulators generally care more about the downside risk, we conclude that gold is the least risky of these three precious metals. In addition, the narrow confidence intervals for the estimates indicate that EVT works well in modeling these extreme events and our risk measure estimates are quite precise.

5. Some Robustness Checks

5.1 Sensitivity to Distributional Assumption

As we noted in section 1, many conventional approaches to risk analysis make the naïve assumption that the data are normally distributed. As we showed in section 4.1, this assumption is not tenable for our data. However, it is interesting to investigate the extent to which are results differ from those that would be obtained under the assumption of normally distributed returns for gold, silver, and platinum.

For each metal, the returns data were standardized, and normal quantiles were computed. This yielded the 99th and 1st percentile values – *i.e.*, the 1% VaR estimates – shown in Table 5. Then, the ES values in Table 5 were estimated by computing the mean of the sub-sample of returns in the tail beyond the VaR. Several results emerge.

First, none of the VaR or ES values computed assuming normality are covered by the corresponding GPD-based 95% confidence intervals in Table 4. Loosely speaking, the results based on the normality assumption are significantly different from those based on the GPD. Second, in all cases, the VaR and ES estimates that are obtained when normality is assumed are smaller (in absolute value) than those based on the more appropriate GPD assumption. That is, the naïve methodology based on the assumptions that daily returns are normally distributed leads to a systematic under-estimation of the "risks" associated with holding these precious metals. We foreshadowed this result in our discussion in section 4.1. In many instances, this under-estimation is substantial in value, and this underscores the importance of using the appropriate EVT.

Risk Measures Based on Assumption that Returns are Normally Distributed

	Positive Returns		Negative Returns	
	VaR	**ES**	**VaR**	**ES**
Gold	0.0299	0.0449	-0.0295	-0.0441
Silver	0.0450	0.0658	-0.0469	-0.0812
Platinum	0.0324	0.0449	-0.0323	-0.0490

Table 5: 1% estimates of VaR and ES for gold, silver and platinum, assuming normally distributed returns.

5.2 Sensitivity to Choice of Sample

In this section, using the gold returns by way of illustration, we check the sensitivity of some of our results to the choice of sample period. We repeat our analysis for the gold daily returns using data only from 4 January 1982 to 8 January 2014. This omits the highly volatile period from 1980 to 1982. There are 8,025 daily returns in the shortened sample, of which 4,178 are positive and 3,847 are negative. The highest daily return is 10.48% and the lowest is -12.9%. Again, the positive returns and negative returns are modeled separately. We first use the ME and parameter plots to find appropriate thresholds and then estimate the GPD parameters by maximum likelihood. The threshold chosen for positive returns is 0.032, and for negative ones it is 0.028.

Comparing Tables 6 and 3 allows us to assess the robustness of the parameter estimates to the choice of sample period. Although the point estimates of the shape and scale parameters are very similar in each case, the precision of estimation is generally greater when the full sample period is used.

In Table 7, we see that the values of the risk measures are reduced when we omit the data from the early 1980s. The changes are not very significant as we include another volatile period 1982-1985 in our modeling. The choice of the time horizon does affect the estimation of risk, as expected. Accordingly, we might infer that if we modelled data from the late 1980s to 2014, the associated VaR and ES would be even smaller. The results of Jang (2007) support our argument. Jang analyzed only negative daily returns for gold and the estimated the 1% daily VaR to be 2.4%, and the associated ES to be 3.13%, both of which values are (absolutely) smaller than ours. His data were from 1985 to 2006, and so he excluded the volatile periods from 1980-1985 and after the 2008 financial crisis. The time horizon can have a

significant impact on the estimation of market risks. When using VaR and ES to compare the risk of different portfolios or assets, financial regulators and supervisors should take this factor into account.

Maximum Likelihood Estimates: Gold

	Parameter Estimates For Positive Returns	Parameter Estimates For Negative Returns
Threshold	$u = 0.032$	$u = 0.028$
No. Exceedances	75	112
$\hat{\xi}$	0.1607	0.1615
(s.e.)	(0.1370)	(0.1016)
$\hat{\sigma}$	0.0102	0.01056
(s.e.)	(0.0018)	(0.0014)
AIC	-510.071	-755.899

Table 6: Maximum likelihood parameter estimates for gold daily returns from 4 January 1982 to 8 January 2014.

Time Sensitivity Check

	Positive Returns		**Negative Returns**	
	1% VaR	1% ES	1% VaR	1% ES
Gold 1968	0.0472	0.0661	-0.0468	-0.0645
Gold 1982	0.0383	0.0516	-0.0403	-0.0553

Table 7: Point estimates of VaR and ES for gold daily returns with different time horizons.

5.3 Threshold Selection Sensitivity

As we can see in Table 8, the 1% daily VaR for positive gold price returns is 4.72% when the threshold is $u = 0.030$, and it is 4.77% for a threshold $u = 0.038$. This VaR estimate is very stable under different thresholds. This is true for both positive and negative returns. The estimates of 1% ES are also quite stable, consistent with our earlier assertion that as long as the threshold is within an appropriate range, the estimates of these risk measures will be robust.

VaR - Positive Returns		
Percentile	Threshold	Estimate
99th	u = 0.030	0.0472
99th	u = 0.038	0.0477
95th	u = 0.030	0.0257
95th	u = 0.038	0.0236

VaR - Negative Returns		
Percentile	Threshold	Estimate
1st	u = 0.022	-0.0468
1st	u = 0.042	-0.0476
1st	u = 0.022	-0.0261
1st	u = 0.042	-0.0232

ES - Positive Returns		
Percentile	Threshold	Estimate
99th	u = 0.030	0.0661
99th	u = 0.038	0.0666
95th	u = 0.030	0.0398
95th	u = 0.038	0.0390

ES - Negative Returns		
Percentile	Threshold	Estimate
1st	u = 0.022	-0.0645
1st	u = 0.042	-0.0654
1st	u = 0.022	-0.0396
1st	u = 0.042	-0.0386

Table 8: Point estimates of VaR and ES for gold daily positive and negative returns.

Table 8 also includes 5% VaR and 5% ES estimates. Of course, these are lower than their 1% counterparts, but they too are very robust to the choice of threshold values, over sensible ranges. Although gold prices are used again here for illustrative purposes, our conclusions apply equally in the cases of silver and platinum.

6. Conclusions and Policy Implications

In this paper we have used extreme value theory to estimate potential extreme losses and gains in the markets for three key precious metals – gold, silver, and platinum. The Peaks-Over-Threshold method, in conjunction with maximum likelihood estimation of the parameters of the Generalized Pareto Distribution, is used. We report estimates of Value at Risk and Expected Shortfall associated with the daily returns on the prices of each of these metals, using a long time-span of data. One novel aspect of our results is that interval estimates of these risk measures are provided, rather than just point estimates. Of the three precious metals considered in this study, we find that gold has the least downside risk. The difference between the downside risk for gold, and that for platinum, is not significant at the 5% level.

However, the downside risk associated with holding silver is significantly greater than that of holding either platinum or gold, again at the 5% significance level.

There are several important policy implications arising from this study. First, naively constructing risk measures on the assumption that the underlying data are normally distributed is extremely dangerous. Specifically, this can lead to estimates of Value at Risk that significantly under-state the risk involved. Portfolio managers are advised to base these measures, instead, on appropriately applied extreme value theory.

Second, the choice of the time horizon for the data can affect the estimation of risk measures, and we have explored this issue using daily gold returns as an illustration. Even when the appropriate extreme value theory is applied, the choice of the sample period can affect the conclusions significantly, and can again lead to under-statements of Value at Risk. This is a point that must be borne in mind by regulators when they evaluate the reported market risk of firms, or assets.

Third, portfolio managers should recognize that the downside risk associated with holding precious metals can vary substantially depending on the asset that is chosen. Among the three widely traded precious metals that we have considered, gold is the preferred choice in this respect. However, because the time horizons used for our gold, silver, and platinum daily returns data are slightly different, the comparisons that we make have some limitations. This warrants further study.

Future work could also consider formalizing the choice of the threshold value when applying the Peaks-Over-Threshold methodology. Although our results appear to be very robust to this choice, this is an issue that is relevant for all studies that employ this particular methodology. Finally, our extreme value analysis has been based on univariate theory, with each of the three precious metals treated separately. An analysis of *multivariate* extremes is possible. However, this needs to be explored with care as it is well known that many dependent multivariate processes are almost independent in the tails of the distributions. In such cases a univariate analysis remains more appropriate.

References:

Singh, A., D. Allen, and R. Powell, 2013. Extreme market risk and extreme value theory. *Mathematics and Computers in Simulation*, 94, 310-328.

Artzner, P., F. Delbaen, J-M. Eber, and D. Heath, 1997. Thinking coherently: Generalised scenarios rather than VaR should be used when calculating regulatory capital. *Risk-London-Risk Magazine Limited*, 10, 68-71.

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

Asadi, M., C. R. Rao, and D. N. Shanbhag, 2001. Some unified characterization results on the generalized Pareto distribution. *Journal of Statistical Planning and Inference*, 93, 29-50.

Avdulaj, K., 2011. The extreme value theory as a tool to measure market risk. IES Working Paper No. 26/2011.

Basel Committee membership. Bank for International Settlements. 29 March 2014 http://www.bis.org/bcbs/membership.htm

Basel Committee on Banking Supervision (1996). Amendment to the Capital Accord to Incorporate Market Risks, Technical report, Bank for International Settlements. http://www.bis.org

Balkema, A. A. and L. De Haan, 1974. Residual life time at great age. *The Annals of Probability*, 2, 792-804.

Blose, L. E. and V. Gondhalekar, 2014. Overnight gold returns. *Applied Economics Letters*, 21, 1269-1272.

Brock, W. A., W. D. Dechert, J. A. Scheinkman and B. LeBaron, 1996. A test for independence based on the correlation dimension. *Econometric Reviews*, 15, 197-235.

Cole, S., 2001. *An Introduction to Statistical Modeling of Extreme Values*. Springer, London.

Dowd, K., 1998. *Beyond Value at Risk: The New Science of Risk Management*, *Vol. 3*. Wiley, Chichester.

Duffie, D. and J. Pan, 1997. An overview of value at risk. *Journal of Derivatives*, 4, 7-49.

Diebold, F. X., T. Schuermann, and J. D. Stroughair, 2000. Pitfalls and opportunities in the use of extreme value theory in risk management. *Journal of Risk Finance,* 1, 30-35.

Dunis, C. L., J. Laws, and G. Sermpinis, 2010. Modelling commodity value at risk with higher order neural networks. *Applied Financial Economics*, 20, 585-600.

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

Embrechts, P., S. I. Resnick, and G. Samorodnitsky, 1999. Extreme value theory as a risk management tool. *North American Actuarial Journal* 3, 30-41.

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

Gençay, R., F. Selçuk, and A. Ulugülyağci, 2003. High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics* 33.2 (2003): 337-356.

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

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

Jorion, P., 1997. *Value at Risk: The New Benchmark for Controlling Market Risk Vol. 2.* McGraw-Hill, New York.

Jang, J-B., 2007. An extreme value theory approach for analyzing the extreme risk of the gold prices. *Finance and Economics Journal*, 6, 97-109.

Matthys, G., and J. Beirlant, 2000. Adaptive threshold selection in tail index estimation. In P. Embrechts (ed.), *Extremes and Integrated Risk Management*. Risk Books, London, 37-49.

McNeil, A. J., 1999. Extreme value theory for risk managers. *Departement Mathematik ETH Zentrum*.

McNeil, A. J., R. Frey, and P. Embrechts, 2010. *Quantitative Risk Management: Concepts, Techniques, and Tools*. Princeton University Press, Princeton, NJ.

Oakes, D. and T. Dasu, 1990. A note on residual life. *Biometrika*, 77, 409-410

Odening, M. and J. Hinrichs, 2003. Using extreme value theory to estimate value-at-risk. *Agricultural Finance Review*, 63, 55-73.

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

Paul, S. and A. Barnes, 2010. Application of extreme value theory to economic capital estimation. In *Advances in Data Analysis*. Birkhäuser, Boston, 329-336.

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

Ribabet, M., 2006. A user's guide to the POT package, version 1.4. http://pot.r-forge.r-project.org/docs/guide.pdf

Riskmetrics: Technical document, 1996. Morgan Guaranty Trust Company of New York.

Ren, F. and D. E. Giles, 2010. Extreme value analysis of daily Canadian crude oil prices. *Applied Financial Economics*, 20, 941-954.

Tavangar, M. and M. Asadi, 2008. On a characterization of generalized Pareto distribution based on generalized order statistics. *Communications in Statistics: Theory and Methods*, 37, 1347-1352.

Tavangar, M. and M. Asadi, 2012. Some unified characterization results on the generalized Pareto distributions based on generalized order statistics. *Metrika*, 75, 997-1007.

Trück, S. and K. Liang, 2012. Modelling and forecasting volatility in the gold market. *International Journal of Banking and Finance*, 9, Article 3.