The Volatility of the Price of Gold:

An Application of Extreme Value Theory

By

Qinlu(Louisa) Chen

An Honours Thesis Submitted in Partial Fulfillment of the
Requirements for the Degree of

BACHELOR OF SCIENCE, HONOURS
in the Department of Economics

We accept this thesis as conforming
to the required standard

______________________________
Dr. David Giles, Supervisor (Department of Economics)

______________________________
Dr. Herbert Schuetze, Honours Co-Supervisor (Department of Economics)

______________________________
Dr. Pascal Courty, Honours Co-Supervisor (Department of Economics)

© Qinlu(Louisa) Chen, 2014
University of Victoria
All rights reserved. This extended essay may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.
Acknowledgement

I would like to express my deep gratitude to my thesis supervisor, Dr. David E. A. Giles, professor in the Department of Economics, University of Victoria. His invaluable technical guidance, encouragement, and constant support have greatly helped me in presenting this extended essay. I would also like to sincerely thank the Honours Program supervisors, Dr. Pascal Courty and Dr. Herbert Schuetze, for their constructive suggestions and useful critiques; and thanks to my classmates for their friendship and support. Without their assistance and dedicated involvement in every step throughout the process, this paper would have never been accomplished.
Abstract

Gold is one of the oldest and most widely held commodities used as a hedge against the risk of disruptions in financial markets. The price of gold fluctuates substantially over time and day-to-day, and this introduces a risk of its own. Surprisingly, there is very little formal analysis of the risk of holding gold. This paper’s goal is to analyze the risk of gold investment by employing the Extreme Value Theory (EVT) to historical daily data for extreme daily losses and gains in the price of gold. The risk measures adopted in this paper are Value at Risk (VaR) and Expected shortfall (ES). Point estimates of VaR and ES are obtained by fitting the Generalized Pareto Distribution (GPD)—the Peak Over Threshold (POT) model, to the daily gold price changes in the London Bullion Market from 1968 to 2014. The interval estimates of VaR and ES are calculated using the delta method, and a time period sensitivity check is conducted by applying the EVT to daily gold returns from 1982 to 2014. Also, a comparison of the risks associated with owning precious metals, including gold, silver and platinum, is presented at the end of the paper.
Table of Contents

1. Introduction ......................................................................................................................... 5
2. Models and Methods ............................................................................................................. 9
   2.1 Extreme Value Theory .................................................................................................... 9
   2.2 Generalized Pareto Distribution .................................................................................... 10
   2.3 Peak Over Threshold method ....................................................................................... 12
3. Risk Measures ....................................................................................................................... 15
   3.1 Value at Risk (VaR) and Expected Shortfall (ES) ......................................................... 15
   3.2 Interval estimation of VaR and ES ................................................................................ 18
4. Data Characteristics and Application .................................................................................. 20
   4.1 Data Characteristics ..................................................................................................... 20
   4.2 Determination of threshold ......................................................................................... 23
   4.3 Parameter Estimation .................................................................................................. 25
   4.4 Risk Measures Estimation ............................................................................................ 28
5. Data time sensitivity check .................................................................................................. 30
6. Comparison of Risks of Different Precious Metals ............................................................. 32
7. Conclusion ............................................................................................................................ 35
References: ............................................................................................................................. 37
Appendix: ................................................................................................................................. 39
1. Introduction

The financial crisis in 2008 has led to a worldwide economic recession. The crash in the subprime mortgage market was the start of a nightmare for many investors. One of the lessons we have learnt from this financial disaster, which caused trillions of dollars in losses, is the importance of risk management. Gold is one of the oldest and most widely-held commodities used as a hedge against the risk of disruptions in financial markets, but the price of gold fluctuates substantially over time and day-to-day, and this introduces a risk of its own. While the US economy was improving in 2013, gold prices kept decreasing. On April 15th, the gold price dropped 9.60 percent in one day as a result of the decreasing inflation rate and the increasing real interest rate in the US. The fact that many investors sold off gold to exit the market accounts for this steep drop in gold prices. In reality, gold is not as safe as people think; last year’s considerable loss in gold’s value was almost the historical largest one-day decline in recent decades. According to data from the London bullion market between 1968 and 2014, the daily percentage change in the gold price was as high as 12.50 percent (1980/1/3) and as low as -14.20 percent (1980/1/22). 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 economic agents. Surprisingly, there has been very little formal analysis of the risk of holding gold as an asset either by itself, or as part of a portfolio. Exceptions include Jang (2007), and Trück and Liang (2012). Jang used extreme value theory to analyze the risk of holding gold, which is quite similar to the goal of my paper. More discussion of Jang’s paper is presented in section 5 when we check the time sensitivity of the VaR and ES estimates. Trück and Liang applied a threshold ARCH model to gold data from the London bullion market from 1999-2008, which is quite different from the method 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. (more on http://www.bis.org). 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 calculated at 99 per cent) for a certain time horizon. Generally speaking, for a given confidence interval and a given portfolio, VaR is the estimated potential loss/gain, the quantile of the loss/gain distribution. For example, if a trader whose portfolio has a 1 percent VaR (at the 99
percentile) of $1 million in one day, it means there is a 1 percent probability that the trader will lose more than $1 million 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 the financial market after the Basel II Accord adopted VaR as a preferred risk measure in the late 1990s. There is much literature about this financial risk measure, for example, the book Value at Risk: the New Benchmark for Controlling Market Risk by Jorion (1996), the book Beyond Value at Risk: the New Science of Risk Management by Dowd (1998) and the article "An Overview of Value at Risk" By Duffie and Pan (1997). Another common risk measure 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), (Artzner et al., 1999). This paper analyzes the effects of both VaR and Expected Shortfall on daily gold returns.

In order to measure extreme risk and to be prepared for irregular losses, we are interested in the behavior of the tail of a distribution. 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.
Extreme Value Theory 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). 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, Resnick and Samorodnitsky (1999), Gilli (2006), and McNeil (1999). EVT is also applied to many other markets, like the crude oil and agricultural markets (see Giles and Ren (2007), and Odening and Hinrichs (2003)). Also the pitfalls of Extreme Value Theory are discussed in Diebold, Schuermann and Stroughair (2000). When studying extreme events, Extreme Value Theory plays the same role as the Central Limit Theorem plays in the studies of sums of a random variable. We use the Peak Over Threshold (POT) method to analyze the distribution (GPD) of exceedances above a specific high threshold.

In this paper, we employ the Peak Over Threshold- Generalized Pareto Distribution method to model the extreme risk associated with the daily returns of gold. Section 2 presents the framework of Extreme Value Theory and the Peak Over Threshold model. Section 3 introduces the standard risk measures—Value at Risk and Expected shortfall. Then we present the gold data, the process of the tail estimation, and the computation of estimates of risk measures in section 4. The time sensitivity of risk measure estimates is discussed in section 5. In section 6, we compare the risk of holding gold to the risk of
holding silver and platinum, by modeling daily silver returns and daily platinum returns with EVT. The interval estimates of VaR and ES of gold, silver and platinum are calculated by using the delta method. The last section summarizes the empirical analysis results and briefly discusses the limitations of this paper.

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 tail of the financial time series. Those rare but plausible events could cause significant damage to society and have a strong negative impact on the livelihood of economic agents. Many conventional models failed to model those irregular events properly. Past literature has discussed the superiority of EVT to other approaches, such as the GARCH model, Historical Simulation, Variance and Covariance method, and Monte Carlo Simulation. The EVT based VaR is more robust than other model based VaR. (See Paul and Barnes (2010), Gençay, Selçuk and Ulugülyağcı (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 (BMM) which involves the Generalized Extreme Value (GEV) distribution and another one is the Peak Over Threshold method which involves the Generalized Pareto
Distribution (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 normalized maxima of BMM follow one of the Fréchet, Weibull and Gumbel distributions according to the Fisher–Tippet theorem. The block maxima method requires the variable to be i.i.d. and must converge to a non-degenerate distribution function. The Peak Over Threshold method models the behavior of exceedances over a given high threshold. EVT implies that the limiting distribution of the exceedances is a Generalized Pareto Distribution (Pickands (1975), Balkema and de Haan (1974)).

Previous research indicated that GPD uses data more efficiently and effectively than BMM. (For detailed discussion of these two methods please refer to Jang (2007), Gilli (2006), Allen, Singh and Powell (2011)). Dividing the data into artificial blocks ignores the fact that each block may have a different situation. In some blocks, all 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’s inefficient to artificially “block” the data. This paper will mainly focus on the application of the POT method.

2.2 Generalized Pareto Distribution

EVT identifies that the maxima of a random variable above certain high threshold should be generated from GPD. The distribution of the exceedances is presented by a Conditional Excess Distribution $F_u$. 
Conditional Excess Distribution Function:

\[ F_u(y) = P(X - u \leq y | X > u) \quad \text{for} \quad 0 \leq y \leq x_F - u \]  \hspace{1cm} (1)

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

Equation (1) could be rewritten in terms of \( F \):

\[ F_u(y) = \frac{F(u+y) - F(u)}{1 - F(u)} = \frac{F(x) - F(u)}{1 - F(u)} \]  \hspace{1cm} (2)

The related EVT reveals the asymptotic behavior of the scaled exceedances.

**Pickands-Balkema-de Haan Theorem** (Balkands & de Haan (1974); Pickands (1975)):

For large enough \( u \), the Conditional Excess Distribution \( F_u(y) \) is well approximated by the **Generalized Pareto Distribution**.

That being said:

\[ F_u(y) \approx G_{\xi,\sigma}(y), \text{ as } u \rightarrow \infty \]

Two-parameter GPD in terms of \( y \):

\[ G_{\xi,\sigma}(y) = \begin{cases} 
1 - (1 + \frac{\varepsilon}{\sigma} y)^{-1/\xi} & \text{if } \xi \neq 0 \\
1 - e^{-y/\sigma} & \text{if } \xi = 0
\end{cases} \]  \hspace{1cm} (3)
for $\in [0,(x_F - u)]$, if $\xi \geq 0$; and $y \in [0,-\frac{\xi}{\sigma}]$, if $\xi < 0$

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

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

(4)

$u$ is the threshold, $\xi$ is the shape parameter, $\sigma$ is the scale parameter.

The shape parameter $\xi$ is related to the heaviness of the tail. The larger the shape parameter is, the heavier the tail will be. The shape parameter $\xi$ could be positive, negative or equal to zero. If $\xi < 0$, it means the Conditional Excess Distribution $F_u(y)$ has an upper bound. If $\xi \geq 0$, it means that the corresponding distribution has a fat tail. Since nobody can set an upper bound for financial losses and a financial time series is usually fat-tailed, the case where $\xi \geq 0$ is of our interest. We see later in the data characteristics section that our data is fat-tailed.

2.3 Peak Over Threshold method

The POT model is a preferable approach to analyze extreme risks because it’s 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 maximizing the likelihood
function. 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 (ME plot)**

   It is a plot of the mean of the excesses over a group of thresholds. We need to choose the threshold $u$ after which the ME plot is approximately linear and upward-sloping for some $x \geq u$. For a random variable $X$ with right end point $x_F$, its mean excess function is defined as:

   $$e(u) = E(X - u|X > u)$$

   (5)

   *for* $u < x_F$. If the underlying distribution is a generalized Pareto distribution, then the corresponding mean excess function is:

   $$e(u) = \frac{\sigma + \xi u}{1 - \xi}$$

   (6)

   *for* $\sigma + \xi u > 0$, *and* $\xi < 1$.

   As we can see from equation (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.
The sample mean excess plot is the locus of

\[(u, e_n(u)), \quad x_1 < u < x_n\]

The empirical mean excess function is

\[e_n(u) = \frac{\sum_{i=k}^{n}(x_i^n - u)}{n-k+1}, \quad k = \min\{i \mid x_i^n < u\} \quad (7)\]

\(n - k + 1\) is the number of observations over the threshold \(u\).

We use R to graph the plot. More discussion will be presented in section 4.2, threshold selection for our gold data.

2. Parameter Plot

This is the plot of the maximum likelihood estimates of the shape and scale parameters under various thresholds. We chose the threshold before which the estimates of parameters are relatively stable. Above a certain threshold, the exceedances should approach GPD, thus the estimated parameters should be roughly constant.

The log-likelihood function for GDP is

\[L(\xi, \sigma \mid y) = \begin{cases} 
-\log \sigma - \left(\frac{1}{\xi} + 1\right) \sum_{i=1}^{n} \log \left(1 + \frac{\xi}{\sigma} y_i\right) & \text{if } \xi \neq 0 \\
-\log \sigma - \frac{1}{\sigma} \sum_{i=1}^{n} y_i & \text{if } \xi = 0 
\end{cases} \quad (8)\]
Both of these plots (especially the ME Plot) are of limited help in practice since they typically do not offer clear outcomes. There is a trade-off between minimizing bias and variance in the process of threshold selection. We need a large $u$ for the EVT to hold. But as the threshold gets larger, there will be fewer observations in the tail. If the threshold is too low, the conditional excess distribution function will not approach the GPD. The basic rule is to choose a threshold as low as possible, as long as the GPD holds. We need to investigate the robustness of the results to choice of threshold $u$ by comparing the confidence intervals and standard errors. For detailed discussion about threshold selection see also Matthys and Beirlant (2000), Embrechts, Klüppelberg, and Mikosch (1997), and McNeil, Frey, and Embrechts (2010).

3. Risk Measures

3.1 Value at Risk (VaR) and Expected Shortfall (ES)

We employ two standard choices of risk measurement in this paper. One is Value at Risk and the other one is Expected Shortfall. Those two risk measures involve extreme quantile estimation.

1. “Value at Risk” (VaR):

VaR quantifies the maximum loss /gain occurring over a given time-period, with a specified (low) probability. It is the capital sufficient to cover the losses from a portfolio over a given holding period. VaR is usually calculated at 99 percentile over a one-day or
ten day-period. In this paper, we compute the VaR in one day. To better address the
relevant and unexpected events happening around the world every day, investors would
like to know the newest VaR on a daily basis for decision making purpose. Because, EVT
only analyzes extreme values, the larger the sample size is the better. Usually the
number of observations above the threshold will be up to 20% of the sample size. Given
the sample data of gold prices from 1968 to 2014, if we compute the VaR on a daily basis
by using gold daily returns, we will have more observations.

Let $X$ be a random variable with continuous distribution function $F$, $VaR_p$ is $(1 - p)^{th}$
percentile of the distribution $F$: $VaR_p = F^{-1}(1 - p)$

The VaR can be defined as a function of GPD parameters:

Reorganize equation (2), we will get

$$F(x) = (1 - F(u))F_u(y) + F(u) \quad (9)$$

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

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

After simplifying and inverting the equation (9), for a given percentile $p$, we will have
\[ \hat{\xi}_p = u + \frac{\hat{\alpha}}{\hat{\xi}} \left( \frac{n}{N_u} \right)^{-\hat{\xi}} - 1 \]  

(11)

Since VaR is the quantile of the density distribution, we can get the formula for the VaR estimator:

\[ \hat{VaR}_p = u + \frac{\hat{\alpha}}{\hat{\xi}} \left( \frac{n}{N_u} \right)^{-\hat{\xi}} - 1 \]  

(12)

2. “Expected shortfall” (ES):

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

The function of ES estimator is:

\[ ES_p = \hat{VaR}_p + E(X - \hat{VaR}_p | X > \hat{VaR}_p) \]  

(13)

The second term of the equation (12) is the expected value of exceedances above the threshold \( VaR_p \), which is the mean excess function of \( F_{VaR_p}(x) \) (see equation (5)). As long as the threshold \( VaR_p \) is large enough, the mean excess function \( F_{VaR_p}(x) \) should follow GPD. The mean excess function for GPD is equation (6) in section 2.3. Thus we get:

\[ F_{VaR_p}(x) = \frac{\hat{\alpha} + \hat{\xi}(\hat{VaR}_p - u)}{1 - \hat{\xi}} \]  

(14)
Plug equation (13) into equation (12) we get the formula for ES estimator:

\[
\hat{ES}_p = \sqrt{\text{VAR}_p} + \frac{\hat{\xi}(\text{VaR}_p - u)}{1 - \hat{\xi}} = \sqrt{\text{VAR}_p} + \frac{\hat{\sigma} - \hat{\xi} u}{1 - \hat{\xi}}
\]

(15)

After we estimate the tail distribution using GPD, we can calculate the risk measures by plugging in the GPD parameter estimates in the above estimator formulas.

3.2 Interval estimation of VaR and ES

It would be interesting to compute the interval estimates of VaR and ES. The interval estimates could help check the robustness of our results. Since both the VaR and ES functions are non-linear, we use the delta method to calculate the confidence interval. The delta method is used to obtain the limiting variance and an approximate probability distribution of an estimator that is asymptotically normal. Because our shape parameter \( \hat{\xi} \) and scale parameter \( \hat{\sigma} \) are maximum likelihood estimators, they are asymptotic normal distributed. We only present the steps and functions used to calculate the confidence interval in this section. Thanks to Ren and Giles (2007) for making their formulas available. For more information about the delta method, please refer to Oehlert (1994).

From the Theorem of Limiting Normal Distribution of a Set of Functions (Greene 2003):

the variance of VaR could be calculated as follows:

\[
\text{var}(\text{VaR}_p) \approx (\hat{d}_1', \hat{c}, \hat{d}_2) = \begin{pmatrix} d_1' \\ d_2' \end{pmatrix} \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \end{pmatrix}
\]
Where \( \hat{d} = \begin{pmatrix} d_1 \\ d_2 \end{pmatrix} = \begin{pmatrix} \frac{\partial \text{VaR}_p}{\partial \xi} \\ \frac{\partial \text{VaR}_p}{\partial \hat{\beta}} \end{pmatrix} \), \quad \hat{C} = \begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{pmatrix} = \text{cov}(\hat{\xi}, \hat{\beta}) \quad c_{12} = c_{21}

\hat{C} is the estimated variance-covariance matrix of the shape and scale parameters. It is available in the R results of maximum likelihood estimation of the GPD parameters.

Whereas \( \hat{d} \) is calculated as following in Excel:

\[
\frac{\partial \text{VaR}_p}{\partial \xi} = -\frac{\hat{\beta}}{\xi^2}\left(\frac{N_u}{n_p}\xi - 1\right) + \frac{\hat{\beta}}{\xi} \left(\frac{N_u}{n_p}\right) \ln \left(\frac{N_u}{n_p}\right)
\]

\[
\frac{\partial \text{VaR}_p}{\partial \hat{\beta}} = \frac{1}{\xi}\left(\frac{N_u}{n_p}\right) - 1
\]

We get the asymptotic variance of VaR after we plug in all the results in equation (15).

The associated standard error is

\[
s.e.(\text{VaR}_p) = \sqrt{\text{var}(\text{VaR}_p)} \quad (17)
\]

As the VaR and ES estimators are normal, the 95% confidence interval is constructed as below:

\[
(\text{VaR}_p - 1.96 \times s.e.(\text{VaR}_p), \text{VaR}_p + 1.96 \times s.e.(\text{VaR}_p)) \quad (18)
\]

We compute the standard error for ES in the same way as VaR by substituting all the
$VaR_p$ with $ES_p$. The 95% confidence interval is constructed as below:

\[
(ES_p - 1.96 \times s.e.(ES_p), ES_p + 1.96 \times s.e.(ES_p))
\]  

(19)

4. Data Characteristics and Application

4.1 Data Characteristics

The data we use is London Fixings P.M. Gold Price from the London Bullion Market Association (LBMA)\(^1\). Fixing levels are set per troy ounce. The data time period is from April 1, 1968 to Jan 8, 2014. There are 11,496 prices in total. All prices are in USD. The daily gold returns are calculated by taking log differences of the daily gold prices.

\[
Daily\ Return_t = \log P_t - \log P_{t-1}
\]

(20)

The total number of observations is 11,496, so there are 11,495 gold returns, among which 5447 are negative and 6048 are positive. We separated the positive returns and the negative returns, and then took the absolute values of the negative returns because EVT is only defined for non-negative values. Since we did not want to impose the assumption that the risk is symmetric, the positive and negative returns are modeled separately. But actually, as we will see later in section 4.4, the risks of positive and negative gold daily returns are symmetric; the VaR and ES we get for both sides are quite

---

\(^1\)The data were retrieved on Jan 9, 2014 from: http://www.quandl.com/OFDP-Open-Financial-Data-Project/GOLD_2-LBMA-Gold-Price-London-Fixings-P-M
close to each other.

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

Figure 2: Daily gold returns. London Bullion Market, April 1, 1968 to Jan 8, 2014.

As we can see in figure 2, in the early 1980s, the daily return of gold fluctuated significantly. The choice of the data time horizon may affect the risk measurement result.

We check the time sensitivity of the risk measures in section 5.
Figure 3: Descriptive statistics of daily Gold returns.

Figure 3 presents the descriptive statistics. As we can see in the figure, the average daily gold return is 0.03%, with a maximum of 12.5% and a minimum of -14.2%. This figure shows that daily gold returns have a positive skew (0.1018>0) and large kurtosis (14.38>3). It suggests the distribution of returns has a fat tail. The null hypothesis for the Jarque-Bera statistic has a probability of zero, which means we should reject the null hypothesis in favor of the alternative hypothesis that this distribution is not normal. For long time horizons, an unconditional approach is better suited. Thus, we prefer the GPD-POT approach to BMM.

Below we use the quantile-quantile (QQ) plot to further analyze the data (Figure 4). The QQ plot is a graphical technique used to assess the effectiveness of fitting a data series to a known distribution. The quantile function \( Q \) is the inverse function of the cumulative distribution function \( F: Q(p) = F^{-1}(p) \).
Figure 4: QQ plots of daily gold returns applied to normal distribution (left panel) and student t distribution (right panel).

The blue points are the empirical data, the daily gold returns. For the student t distribution, the degree of freedom is estimated within the sample. The graphs show that the data do not fit well with normal distribution or student t distribution. In fact, many researchers found that the normal distribution will underestimate the VaR. This result suggests that, instead of considering those tail observations as outliers, we need a proper method to address the fat tail and model those extreme events. EVT is designed to deal with this situation. The application of EVT to daily gold returns will be presented in the next section. The descriptive statistics and QQ plots are obtained with the help of Eview 7.

4.2 Determination of threshold

As we have discussed in section 2.3, there are two plots that could help with the
determination of the threshold. Figure 5 presents the ME plot and parameter plot for positive and negative returns. The two plots are obtained with the help of R.

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)

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

The parameter plot is a plot of maximum likelihood estimates of the shape parameter $\xi$ and scale parameter $\sigma$ of the GPD. The maximum likelihood estimator of the shape parameter is asymptotically normal. We check the robustness of our choice of threshold by comparing the standard errors of the parameter estimates in Section 4.3. According to the graph, for positive returns, we choose threshold $u=0.03$; for negative returns, we choose $u=0.022$. Before those two thresholds, the plots are relatively flat. The selection of threshold is a bit arbitrary. However, as we can see in Section 4.4, the estimated risk measures are not very sensitive 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 will be quite stable.

### 4.3 Parameter Estimation

Given the thresholds selected in the previous section, we could estimate the shape and scale parameters of the corresponding GPD. Please refer to Section 2.3, equation (8) for the maximum likelihood function. The software used for modeling the extreme events is R.

Table 1 summarizes all the parameter estimates and their standard errors under
different thresholds. The results in the table confirmed my choice of the threshold: threshold \( u=0.03 \) for positive returns and threshold \( u=0.022 \) for negative returns. For positive returns, there are 210 observations above the threshold \( u=0.03 \). Compared to the sample size 6048, the tail only accounts for 3.47% of the total distribution. As mentioned in previous sections, EVT only analyzes the extreme events, thus even if we started with a large sample, we end up with few observations. So compared to using weekly returns, daily returns provide a bigger sample size. This in turn reduces the variance by increasing the number of observations in the tail. It is the same situation for negative returns.

Since the maximum likelihood estimator is asymptotically normal, the associated maximum likelihood estimates of the parameters under selected thresholds are statistically significant at the 5% significance level. The second and fourth columns with different thresholds are presented to check the sensitivity of the parameter estimates to the choice of threshold. The AIC, Akaike’s Information Criteria, is used to compare competing models. Because we use the maximum likelihood estimation, the higher the likelihood, the smaller the AIC will be. Again, this confirms that our choice of thresholds is better than comparable thresholds, since the associated AIC for \( u=0.03 \) (positive) and \( u=0.022 \) (negative) is smaller than the AIC for \( u=0.038 \) (positive) and \( u=0.042 \) (negative).
Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Threshold</th>
<th>No. Exceedances</th>
<th>Positive Returns</th>
<th>Negative Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>u = 0.03</td>
<td>u = 0.038</td>
</tr>
<tr>
<td></td>
<td></td>
<td>ξ</td>
<td>(s.e.)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.1848</td>
<td>(0.0910)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>0.0123</td>
<td>(0.0014)</td>
</tr>
<tr>
<td>AIC</td>
<td></td>
<td>-1347.096</td>
<td>-684.8385</td>
</tr>
</tbody>
</table>

Table 1: Maximum likelihood estimates under different thresholds for positive and negative returns

Figure 6 is a comparison of the empirical Cumulative Distribution Function (CDF) against the theoretical CDF under the threshold we chose. The black points are the empirical data and the red line is the theoretical CDF. The left panel is a plot of GPD for positive returns with the threshold \( u = 0.03 \), shape parameter 0.1848 and scale parameter 0.0123. The right panel is a plot of GPD for negative returns with the threshold \( u = 0.022 \), shape parameter 0.1689 and scale parameter 0.0105. Both of the graphs show that the GPD models the tail of the corresponding distribution quite well. The negative daily return series fits a GPD slightly better than the positive return series does.
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 goal of risk measures, such as VaR, is to help set risk adjusted minimum capital requirement to protect financial institutions from irregular, large losses. VaR and ES analyze the worst case scenario: if things go wrong, how wrong could it go? Point estimates of VaR and ES are calculated at the conventional 99th percentile and 95th percentile following the formula in section 3.1. Firms use the 99th percentile more frequently than the 95th percentile in evaluating market risk. Our discussion and interpretation focus on the 99th percentile. The calculation is done in Excel. The results are presented in table 2.

For Positive returns, with 1% probability (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 $1 million 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.

<table>
<thead>
<tr>
<th>VaR - Positive Returns</th>
<th>VaR - Negative Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile</td>
<td>Threshold</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.03</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.038</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.03</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.038</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expected Shortfall - Positive Returns</th>
<th>Expected Shortfall - Negative Returns</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentile</td>
<td>Threshold</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.03</td>
</tr>
<tr>
<td>99&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.038</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.03</td>
</tr>
<tr>
<td>95&lt;sup&gt;th&lt;/sup&gt;</td>
<td>u = 0.038</td>
</tr>
</tbody>
</table>

*Table 2: Point estimates of VaR and Expected Shortfall for gold daily positive and negative returns.*

As we can see in table 2, at the 99th percentile, the VaR for positive returns is 0.0472 under threshold u=0.03 and is 0.0477 under threshold u=0.038. The VaR estimates are quite stable under different thresholds. For ES, the estimates under different thresholds are quite stable at the same percentile as well. This shows that the estimates of VaR and ES are not that sensitive to the choice of threshold u, which confirms our argument that as long as the threshold is within a proper range, the estimates of risk measures are robust. Also, the table shows that the risk is symmetric as the VaR or ES for positive
returns (VaR_0.01=0.0472, u=0.03) is quite close to that for negative returns (VaR_0.01=0.0468, u=0.022) at the same percentile.

5. Data time sensitivity check

In this section, we check the sensitivity of our results against the choice of the data time horizon. We model the gold daily returns from Jan 4th, 1982 to Jan 8, 2014 with GPD. The whole process is the same as presented in the previous section. There are 8025 gold daily returns during this period and among which 4178 are positive and 3847 are negative. The highest daily return is 10.48% and the lowest is -12.9%. We intentionally omit the data from 1980 to 1982 since that’s the most volatile period in my sample. We wanted to see how this omission would affect our risk measurement results.

Again, the positive returns and negative returns are modeled separately. We first use ME plot and parameter plot to find a proper threshold and then estimate the GPD parameters with maximum likelihood estimator. The threshold chosen for positive returns is 0.032 and for negative ones it is 0.028. The ME plot and parameter plot are presented in appendix 1. The following table summarizes the estimates of risk measures for gold daily returns over different time periods.

<table>
<thead>
<tr>
<th>Time Sensitivity Check</th>
</tr>
</thead>
</table>


<table>
<thead>
<tr>
<th></th>
<th>Positive Returns</th>
<th></th>
<th>Negative Returns</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1% VaR</td>
<td>ES</td>
<td>1% VaR</td>
<td>ES</td>
</tr>
<tr>
<td>Gold 1968</td>
<td>0.0472</td>
<td>0.0661</td>
<td>-0.0468</td>
<td>-0.0645</td>
</tr>
<tr>
<td>Gold 1982</td>
<td>0.0383</td>
<td>0.0516</td>
<td>-0.0403</td>
<td>-0.0553</td>
</tr>
</tbody>
</table>

Table 3: Point estimates of VaR and ES for gold daily returns with different time horizon.

As we can see in Table 3, the risk measures get smaller as we omit the data from the early 1980s. The changes are not very significant since we include another volatile period 1982-1985 in our modeling. The choice of time horizon does affect the estimation of risk. According to this feature, we could infer that if we model data from late 1980s to 2014, the associated VaR and ES will be even smaller. The results presented in the paper “An Extreme Value Theory Approach for Analyzing the Extreme Risk of the Gold Prices” written by Jang (2007) support our argument. Jang only analyzed negative returns of daily gold return, and he got 2.4% for 1% VaR and 3.13% for associated ES, which is much smaller than our results (0.0468 for 1% VaR and -0.0645 for ES). His data is from 1985 to 2006, excluded the volatile period from 1980-1985 and the volatile period after 2008 financial crisis. The difference in time period could be one of the reasons why his results are much smaller than ours. Regulators could use this information when monitoring company’s market risks. The time horizon could 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.
6. Comparison of Risks of Different Precious Metals

1. EVT Application to silver and platinum

In this section, we apply EVT to silver and platinum to compare the risks of different precious metals. We repeat the process discussed in section 2 and 3 for silver and platinum to compute the VaR and ES for each of them. For silver, we use *LBMA Silver Price: London Fixings.* from London Bullion Market Association (LBMA). Fixing levels are set per troy ounce. The prices are from Jan 2, 1968 to March 14, 2014, and are all in USD. We have 11,680 gold daily returns in total, among which 5968 are positive returns and 5712 are negative returns. The highest daily return is 31.2% and the lowest is -25.7% over the given period. In our sample, the data for some dates are missing. In order to make sure the return is calculated on a daily basis (over one trading day), we convert available data in GBP to USD for those missing dates. Again all returns are computed by taking log differences of the prices and positive returns and negative returns are modeled separately. We first used ME plot and parameter plot to find the proper threshold. The descriptive statistics and two plots are provided in appendix. We choose threshold $u=0.018$ for positive returns and $u=0.038$ for negative returns. The point and interval estimates calculated at 99th percentile are presented in Table 4. The confidence interval is 95% and the formulas for the standard error are provided in section 3.2. For

---

2 Data retrieved on March 15 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
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 choose the Johnson Matthey Base Prices, London 8:00 am., which are the company’s quoted selling prices. The price is for metal in sponge form, with minimum purities of 99.95% for platinum. All prices are in USD per troy oz. The time horizon is from July 1, 1992 to March 24, 2014. This time horizon is shorter than the time horizon for gold and silver. In order to have a better understanding of the risk of holding platinum, we used all the available data. However, when comparing the risk of various commodities, we should keep it in mind that the risk measurement could be affected by the differences in data time period. For some dates, there is no data available in London market. Because all the prices provided by Johnson Matthey are in USD, we simply use the platinum prices from New York and Hong Kong market to make sure the returns we get are daily returns. Again, we model the positive and negative returns separately and select the thresholds with visual aid. The ME plot and parameter lot along with the descriptive statistics for platinum are provided in the Appendix 2. The highest daily platinum return is 13.9% and the lowest is 15.5%. According to the two plots, we choose threshold \( u=0.03 \) for positive returns and \( u=0.025 \) for negative returns.

\[ \text{Data and data description were from http://www.quandl.com/JOHNMMATT/PLAT-Platinum-Prices} \]
The point and interval estimates of corresponding VaR and ES are presented in Table 4. 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%.

3. Comparison of the Risks of Gold, Silver, Platinum

After applying EVT to silver and platinum, we compute the point and interval estimates of VaR and ES for silver and platinum (formulas are provided in section 3.1 and 3.2). The following table summarizes all the estimates of risk measure at 99\textsuperscript{th} percentile. The 95% confidence intervals are obtained using delta method. 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 care more about the downside risk, we conclude that gold is the least risky rare metal compared to silver and platinum. In addition, the narrow confidence intervals for the estimates indicate that EVT works well in modeling extreme events and our risk measure estimates are quite robust.

<table>
<thead>
<tr>
<th></th>
<th>Percentile</th>
<th>VaR</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gold</strong></td>
<td>99\textsuperscript{th}</td>
<td>0.0472</td>
<td>0.0444</td>
<td>0.05</td>
</tr>
<tr>
<td><strong>Silver</strong></td>
<td>99\textsuperscript{th}</td>
<td>0.0774</td>
<td>0.0726</td>
<td>0.0821</td>
</tr>
<tr>
<td><strong>Platinum</strong></td>
<td>99\textsuperscript{th}</td>
<td>0.0453</td>
<td>0.0422</td>
<td>0.0484</td>
</tr>
<tr>
<td>Percentile</td>
<td>ES</td>
<td>95% CI Lower</td>
<td>95% CI Upper</td>
<td></td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>--------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>Gold</td>
<td>99th</td>
<td>0.0661</td>
<td>0.0587</td>
<td>0.0736</td>
</tr>
<tr>
<td>Silver</td>
<td>99th</td>
<td>0.1082</td>
<td>0.0972</td>
<td>0.1191</td>
</tr>
<tr>
<td>Platinum</td>
<td>99th</td>
<td>0.0606</td>
<td>0.0527</td>
<td>0.0684</td>
</tr>
</tbody>
</table>

**Point and Interval Estimates-Negative Returns**

<table>
<thead>
<tr>
<th>Percentile</th>
<th>VaR</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td>99th</td>
<td>0.0468</td>
<td>0.044</td>
</tr>
<tr>
<td>Silver</td>
<td>99th</td>
<td>0.0838</td>
<td>0.0778</td>
</tr>
<tr>
<td>Platinum</td>
<td>99th</td>
<td>0.0511</td>
<td>0.0462</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentile</th>
<th>ES</th>
<th>95% CI Lower</th>
<th>95% CI Upper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gold</td>
<td>99th</td>
<td>0.0645</td>
<td>0.0571</td>
</tr>
<tr>
<td>Silver</td>
<td>99th</td>
<td>0.1301</td>
<td>0.1078</td>
</tr>
<tr>
<td>Platinum</td>
<td>99th</td>
<td>0.0752</td>
<td>0.0583</td>
</tr>
</tbody>
</table>

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

7. Conclusion

This paper shows that EVT is a reliable method for predicting future potential extreme losses/gains. As we have seen in section 5, the choice of a data time horizon will affect the estimation of risk measures. This result may be interesting to regulators when they evaluate the reported market risk of firms. The price of gold is volatile, but compared to other precious metals, it’s still the safest. Also, since the time horizons of the gold, silver, and platinum daily returns are different, the comparison has some limitations. We use all the available data to estimate the risks of the three precious metals in order to improve the accuracy of the estimation. Future analysis could compare the risks of different commodities by analyzing data that have the same time horizon. Furthermore, our interval estimates of VaR and ES show that our results are quite robust. We have
demonstrated how to apply EVT to model fat tail distributions, like daily gold returns. The results provide useful information for investors or portfolio managers who hold gold as an asset or a hedge against financial market risks. Future research could analyze the suitability of applying EVT to other markets. The main limitation of this model is the choice of the threshold. There is no mathematical way to get a “most” efficient threshold. Obtaining the threshold only with visual aids makes this process a bit arbitrary. Future improvement could be achieved in this area.
References:


Appendix:

1. ME plots and parameter plots of gold daily returns from 1982-2014

2. Descriptive statistics of silver
2.1 ME plots and parameter plots of silver daily returns
3. Descriptive statistics of platinum

3.1 ME plots and parameter plots of platinum daily returns