Analysis of the Relationship between Brent Crude Oil and Gold Prices in R

Hilala Jafarova*, Rovshan Aliyev**, Sudaba Salihova***

* Department of Stochastic control and Applied statistics, Institute of Control a Systems of ANAS, Baku, Azerbaijan, Department of Information technology in public administration, State Management Academy under the President Republic of Azerbaijan, Baku
** Department of Stochastic control and Applied statistics, Institute of Control a Systems of ANAS, Baku, Azerbaijan, Department of Operation Research and Probability Theory, Baku State University, Baku, Azerbaijan
*** Department of Business Economics and Management, Azerbaijan State University of Economics, Baku, Azerbaijan

E-mail: hilalajafarova@gmail.com

Copyright © 2018 Hilala Jafarova, Rovshan Aliyev, Sudaba Salihova. This is an open access article distributed under the Eurasian Academy of Sciences License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT
This paper focuses on the relationship between Brent Crude Oil and gold prices from 2010.01.01 to 2016.03.01. The aim of this article is to analyze and determine the character of the co-movement between price levels. For statistical computing R language and environment be used. This article also presents the basic characteristic and determinants of current price trends.

Keywords: oil, gold, linear regression model, R code, plots.

JEL-Clasification: B41, B40, C20

1. Introduction

Gold and oil prices are definitely the most important economic indicators, and this article is devoted to the relationship between the price of foreign currencies, the dollar and gold and oil prices. However, on the relationship of currencies and commodity markets has long been known.

Dollar and gold or a dollar and oil - 90% of traders are using these correlations in their trade. But gold and oil are not the only products that influence the currency market quotes. There is a correlation between the market much more subtle and less well-known (see, References).

Bildirici (2015) the relationship between oil prices and the price of gold and silver was analyzed by BDS test, non-linear ARDL approach and two non-linear Granger causality methods for 1973:1 – 2012:11 period in Turkey. Šimáková (2011) is included a quantitative
analysis of the variables, such as Granger causality test, Johansen cointegration test and Vector Error Correction model. This paper reveals

*Corresponding author's email: hilalajafarova@gmail.com

the existence of a long-term relationship between analyzed relationship between oil and gold prices from January 1970 to December 2010. In study [8] optimal two-month lag in Johansen cointegration test was determined by lag structure in EViews5. Cointegration test confirmed the existence of one cointegration equation. The equation is following:

\[
\text{LNGOLD} = 0.64 \times \text{LNOIL} + 3.73C
\]

which means that percentage increase in oil price affect 0.64 % increase in gold price. The value 0.64 is consequently a long-term relationship.

Unlike of above mentions studies in present paper analyzed relationship between Brent Crude Oil and Gold prices from 2010.01.01 to 2016.03.01. For statistical computing R language and environment be used. R is a language and environment for statistical computing and graphics. It provides a wide variety of statistical (linear and nonlinear modelling, classical statistical tests, time-series analysis, classification, clustering and so on) and graphical techniques, and is highly extensible.

2. PRIMARILY DISCUSSIONS AND FORMULATION OF THE PROBLEM

Oil or as it is also called "black gold" is the raw material of which is also an integral part of economic stability and competitiveness of the economies of many countries in the world. It so happened that the dollar is tied to oil, since the US is one of the largest earners of this raw material, and one of the largest consumers of oil. Following the descriptive statistics of oil prices and graphic was used in the article from 2010.01.01 to 2016.03.01 are given:

\[
\begin{array}{cccccc}
\text{Min.} & \text{1st Qu.} & \text{Median} & \text{Mean} & \text{3rd Qu.} & \text{Max.} \\
33.18 & 71.02 & 106.80 & 93.82 & 111.20 & 124.70 \\
\end{array}
\]
Analysis of the Relationship between Brent Crude Oil and Gold Prices in R

Fig.1. Graph of Brent Crude Oil Price at monthly intervals with indicator RSI (Relative Strength Index).

Gold is still very popular in the world because of its "conversion" in any commodity or currency. This is important for the gold in the world-historical fact. The volume of gold reserves of a particular state is not the last feature of the domestic economy and the stability of the country. Similarly as the oil price of gold on the world market or supports the American currency or vice versa exerts pressure on it. Following the descriptive statistics of Gold prices and graphic was used in the article from 2010.01.01 to 2016.03.01 are given:

```r
> summary(gold)
Min. 1st Qu. Median Mean 3rd Qu. Max.   
1061  1231  1342  1401  1614  1781
```

Fig.2. Graph of Gold price at monthly intervals with indicator RSI (Relative Strength Index).
The expected model for the data is
\[ \text{oil} = \beta_0 + \beta_1 \times \text{gold} \]
where \( \beta_0 \) is the theoretical y-intercept and \( \beta_1 \) is the theoretical slope. The goal of a linear regression is to find the best estimates for \( \beta_0 \) and \( \beta_1 \) by minimizing the residual error between the experimental and predicted oil. The final model is
\[ \text{oil} = b_0 + b_1 \times \text{gold} + e \]
where \( b_0 \) and \( b_1 \) are the estimates for \( \beta_0 \) and \( \beta_1 \) and \( e \) is the residual error.

### 3. MAIN RESULTS

Syntax for a regression analysis in R is
\[ \text{lm( } Y \sim \text{ model} \)\]
where \( Y \) is the object containing the dependent variable to be predicted and \( \text{model} \) is the formula for the chosen mathematical model. The command \( \text{lm( )} \) provides the model’s coefficients but no further statistical information; thus

```r
> lm(log(gold)~log(oil))
```

**Call:**
\[ \text{lm(formula = log(gold) ~ log(oil))} \]

**Coefficients:**

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 5.8799    |   0.3017 |          |
| log(oil)   |   0.3017  |         |          |

Some functions allow to display the resulting R model including \textbf{summary ( )} - displays a certain set statistical parameters (statistical tests ...), \textbf{residuals ( )} - indicates the remains of a regression, \textbf{predict ( )} - predicted values, and \textbf{coef ( )} - the vector parameter estimates. In the system R for calculating the sampling of the correlation coefficient use function \textbf{cor ( )}.

```r
> summary(lm.r)
```

**Call:**
\[ \text{lm(formula = log(gold) ~ log(oil))} \]

**Residuals:**

<table>
<thead>
<tr>
<th>Min</th>
<th>1Q</th>
<th>Median</th>
<th>3Q</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.17988</td>
<td>-0.09576</td>
<td>-0.01968</td>
<td>0.10075</td>
<td>0.17111</td>
</tr>
</tbody>
</table>

**Coefficients:**

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 5.8799    |   0.3017 |          |
| log(oil)   |   0.3017  |         |          |
The section of output labeled ‘Residuals’ gives the difference between the experimental and predicted gold price. Estimates for the model’s coefficients are provided along with their standard deviations (‘Std Error’), and a t-value and probability for a null hypothesis that the coefficients have values of zero. In this case, for example, we see that there is no evidence that the intercept (β₀) is different from zero and strong evidence that the slope (β₁) is significantly different than zero. At the bottom of the table we find the standard deviation about the regression (sᵣ or residual standard error), the correlation coefficient and an F-test result on the null hypothesis that the MSreg/MSres is 1.

From the analysis of the results we obtain the following:

Residual standard error (0.1075) - estimate of the standard deviation of residuals. It is assumed that the residuals are normally distributed with mean a value of 0 and a standard deviation σ. The analysis results of this line is exactly assesses the value of σ.

Multiple R-squared (0.5096) and Adjusted R-squared (0.5015) - the coefficient of determination and the coefficient of determination adjusted for the number of model parameters respectively.

F-statistic (63.38) - the value of Fisher's criterion, by which is verified the null hypothesis that all the coefficients of the true model (in this case, β₁ and β₂) are equal to 0.

The value of F-test for linear models is calculated as the ratio of the variance in the data, "explained" the parameters of the model, to the residue of the dispersion (i.e, of the total variance, that model "Doesn't explain"). It is easy to see, using object lm.r as a parameter of function anova ()

> anova(lm.r)
Analysis of Variance Table

Response: log(gold)
  Df  Sum Sq Mean Sq  F value    Pr(>F)
log(oil)  1 0.73219 0.73219  63.378 5.133e-11 ***
Residuals 61 0.70471 0.01155

...
Obtained regression equation is

\[ \ln(\text{gold}) = 5.8799 + 0.3017 \times \ln(\text{oil}) \]

which means that percentage increase in oil price affect 0.30% increase in gold price. The value 0.30 is consequently from 2010.01.01 to 2016.03.01 relationship.

**Fig.3.** The residual errors plotted versus their fitted values.

In Fig.3 the residuals should be randomly distributed around the horizontal line representing a residual error of zero; that is, there should not be a distinct trend in the distribution of points.

**Fig.4.** The standard Q-Q plot.
In Fig. 4 plot which should suggest that the residual errors are normally distributed. The scale-location plot in the upper right shows the square root of the standardized residuals (sort of a square root of relative error) as a function of the fitted values.

![Scale-Location Plot](image)

**Fig. 5.** The scale-location plot.

In Fig. 5 shows the square root of the standardized residuals (sort of a square root of relative error) as a function of the fitted values. Again, there should be no obvious trend in this plot.

![Residuals vs Leverage Plot](image)

**Fig. 6.** The residual errors plotted versus their fitted values.
In Fig. 6 shows the each points leverage, which is a measure of its importance in determining the regression result.

Superimposed on the plot are contour lines for the Cook’s distance, which is another measure of the importance of each observation to the regression. Smaller distances means that removing the observation has little affect on the regression results.

4. CONCLUSIONS

In this paper analyzed relationship between Brent Crude Oil and Gold prices from 2010.01.01 to 2016.03.01 in R codes. Obtained regression equation shows that percentage increase in oil price affect 0.30 % increase in gold price. From the analysis above it can be concluded that Brent Crude Oil and gold prices have a positive correlation (0.7). Obtained plots show that the residual errors plotted versus their fitted values, the residual errors are normally distributed (a standard Q-Q), the square root of the standardized residuals, the each points leverage, which is a measure of its importance in determining the regression result.

REFERENCES


