Pricing and Volatility Relationships for the Largest Oil Producer: Saudi Arabia

Saad Alshahrani and Ikhlaas Gurrib

***Abstract*—Using OLS, ARIMA and Change Point Analysis, this paper empirically analyzes the time series relation between oil price and production and explore various kinds of model that can be utilized to overcome statistical problems and to understand the relationship in Saudi Arabia, one of the fastest growing economies in the Middle East and North Africa and the largest oil producing country in the world, with a large current account surplus. This paper methodically analyzed a times series data and built up models to overcome statistical problems such as non linearity, autocorrelation, and change point. Our findings indicate Saudi Arabia had been consistently producing oil in linear fashion before 2001. However, any international or Middle East dispute can change the trend, as is evident from our change point analysis. Nonetheless, it is likely that Saudi Arabia may increase their oil production during such times of conflict to meet the energy requirements of the world.**

***Index Terms*—Oil prices, production, change point, Saudi Arabia.**

I. Introduction

The Kingdom of Saudi Arabia is the largest country on the Arabian Peninsula and primarily borders the Persian Gulf to the east, the Red Sea on the west, Yemen to the south and Iraq to the north. Its lesser borders include the U.A.E. and Oman to the southeast and Kuwait and Jordan to the northwest. Saudi Arabia is about one-fifth the size of the United States, with a total area of over 2.5 million square kilometers. The climate is mostly harsh, dry desert with extreme temperatures and mostly uninhabited sandy desert. Saudi Arabia's primary natural resource is oil; hence, this country has been significantly and highly dependent upon oil exports. Oil export revenues make up approximately 90% to 95% of total Saudi export earnings, 70% to 80% of state revenues, and around 40% of the country's gross domestic product (GDP) [1]. Saudi Arabia's economy remains, despite attempts at diversification, heavily dependent on oil (although investments in petrochemicals have increased the relative importance of the downstream petroleum sector in recent years).

The combination of relatively high oil prices and exports has led to a revenue windfall for Saudi Arabia during 2004 and early 2006 [2]. Increasing oil prices for various years can be observed in Fig. 1 below.

Manuscript received November 16, 2012; revised January 15, 2013.

S. Alshahrani is with the International Monetary Fund (IMF), Fiscal Affairs Department, 700 19th St, Washington, DC 20431, USA (e-mail: salshahrani@imf.org).

I. Gurrib is with the Canadian University of Dubai, Sheikh Zayed Road, PO Box 117781, Dubai, United Arab Emirates (e-mail: ikhlaas@cud.ac.ae).

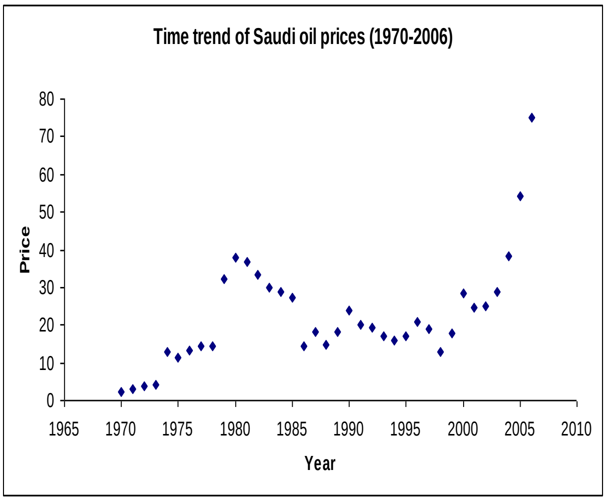

Fig. 1. Oil price trend from 1970-2010.

As the graph suggests, oil prices show an increasing trend over the 30+ year span. In addition, prices doubled four times between 1970 and 1979, reflecting a 1600 percentage increase in only 9 years. The price approximately doubled again between 1989 and 2000, which is only another 11 years. And, even though the price declined gradually between 1981 and 1988, a very sharp increase in prices can be seen from 2001 until current years, after which the price doubled again, but this time in only 5 years between the years of 2002 to 2006.

Because Saudi Arabia is the world's leading oil producer and exporter, and its location is in the politically volatile Gulf region, there is an added element of concern for its major customers, including the United States. However, because Saudi Arabia has one-fourth of the worlds proven oil reserves and some of the lowest production costs, it will still likely remain the world's largest net oil exporter for the foreseeable future [3]. This prediction is based, in part, on the fact that although Saudi Arabia maintains a crude oil production capacity of around 10.5-11.0 million barrels/day, there are claims that it is "easily capable" of producing up to 15 million barrels/day in the future and can maintain that production level for 50 years. In June 2005, the Saudi Aram Senior Vice President of Gas Operations, The Oil Drum report stated that Saudi Arabia would raise production capacity to more than 12 million barrels/day by 2009, and then possibly to 15 million barrels/day "if the market situation justifies it"[4]. The report added that by 2006, Saudi Arabia would have 90 drilling rigs in the Kingdom, more than double the number of rigs operating in 2004 [4].

Little attention has been devoted to inquiring about the effects of fluctuations on the price of crude oil on employment and output for other medium-sized economies

DOI: 10.7763/JOEBM.2013.V1.14

that are represented in the Organization for Economic Co-operation and Development (OECD). The lack of unequivocal evidence on the relationship between oil price changes and production creates an obvious deficiency that may affect applied research and any analysis of policy in finance and economic development. Crude oil prices behave much as any other commodity, with wide price swings in times of shortage or oversupply. The crude oil price cycle may extend over several years, responding to changes in demand in OPEC and non-OPEC supply [5].

II. Data

The data source for this study is drawn from the Saudi Arabian Monetary Agency's Research and Statistics Department's Forty First Annual Report [5]. This agency has recorded the data for oil production and prices since 1970. Also, this data set includes share price indices for major industrial sectors for the same years. This report contains annual data on oil prices and production information since 1970. The variables that were chosen for this analysis are *OP (Oil prices for each year, SCOP (Saudi crude oil production), SCOP*^{2} *(Square of SCOP variable), and* X_2 *(Dummy variable for oil prices: 1 if year less than equal to 2001; 0 otherwise).* We created variable SCOP^2 to explore the non linear relationship between price and production. We also created a dummy variable X_2, as we suspect a change point from year 2001**.** The data spans from 1971 to 2006.

III. Research Methodology

A. Exploratory Data Analysis (EDA)

The correlation between OP and SCOP was found to be 0.4489, which is moderately high. Also, correlation coefficients between SCOP and other industries were calculated, which is shown in Table I as follows.

TABLE I: Spearman Correlation Coefficients for Saudi Crude Oil Production (SCOP) Against Share Prices Index by Other Sectors

Share Price Index by Sector of Banking	0.71257
Share Price Index by Sector of Industry	0.71647
Share Price Index by Sector of Cement	0.73076
Share Price Index by Sector of Services	0.51380
Share Price Index by Sector of Agriculture	0.20201

It can be inferred from the above table that there is a high correlation between share price indices of sectors like banking, industry, cement, services, and SCOP. Also, agriculture and SCOP had a correlation of 0.20201. As we can see a high correlation amongst SCOP and other industrial sectors, then it makes sense to use SCOP as a variable that will affect OP. Additionally, a scatter plot between OP and SCOP suggests the presence of no heteroskedasticity. However, it can be observed that all the observations are clumped into two distinctive groups. One is before year 2001 and one is after year 2001. As part time series modeling, as these are time series data, we need to also check for possible autocorrelation. The Durbin-Watson (DW) test is useful for finding repeating patterns which could be buried under noise, or to identify a potentially missing frequency. Specifically, the DW test is traditionally used *to check the* presence of first-order autocorrelation. The DW results indicate that there is a first order autocorrelation - *AR (1*). To account for this issue, an ARIMA model is constructed [6].

Further, although the normal probability plot (available on request) seems to be fairly straight, suggesting errors to be normally distributed, we performed a normality test as part of robust testing. Looking at the result from the normality test of Shapiro-Wilk, we rejected the null hypothesis that the data is normally distributed, and thus we need to account for this. One of the ways is to introduce some polynomial function in our regression; creating the SCOP^2 variable and using it in our regression can be justified to capture non linear trend. If normality is present, the model is well-behaved and many assumptions and analyses can be performed, based on the test of whether the data coincide with a normal distribution. In contrast, if the data conform to a normal distribution, then certain tests results should not be considered in the analysis. Some of these tests that would probably not be statistically significant are the Z tests, t-tests, F-tests and chi-square tests because they are all derived from a normal distribution [7].

Based on the EDA in the above section, we clearly have three problems: non normal data, autocorrelation, and change point. To deal with non normal data, we fit a normal regression (OLS) with quadratic parameter. For the autocorrelation problem we ran an ARMA model, and then finally a change point test [7]. The Ordinary Least Squares (OLS) model used is as follows:

$$OP_t = \beta_0 + \beta_1 SCOP + \beta_2 SCOP^2 + \beta_3 SCOP * X2 + \beta_4 SCOP^2 * X2 + \varepsilon_t$$

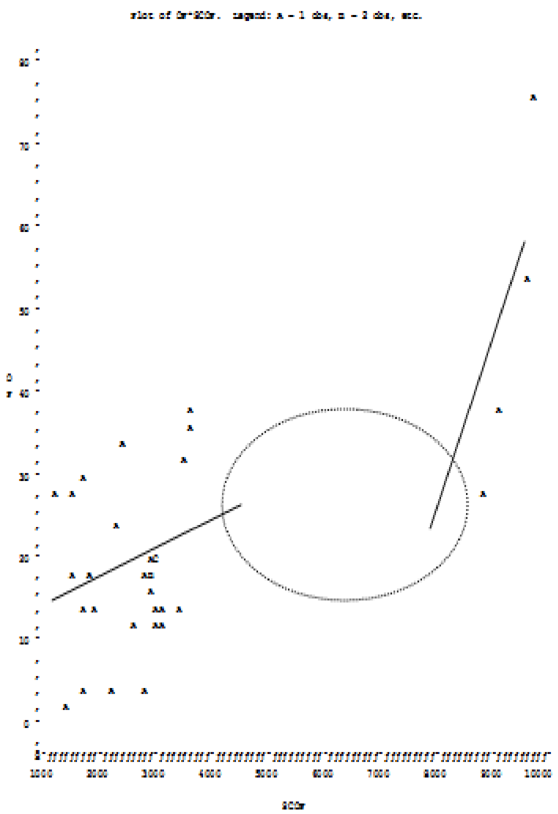

Fig. 2. Least square fit between OP and SCOP.

The reason to include $SCOP^2$ is because of EDA, as we suspect some non linear relation. Also, we interacted before year 2001 and after year 2001 with SCOP with $SCOP^2$ to see if there is an interaction effect. Therefore, if X_2 (the dummy variable) = 0, $OP_t = \beta_0 + \beta_1 SCOP + \beta_2 SCOP^2$. Otherwise,

$$OP_t = \beta_0 + \beta_1 SCOP + \beta_2 SCOP^2 + \beta_3 SCOP + \beta_4 SCOP^2 + \varepsilon_t$$

$$OP_t = (\beta_0 + \beta_5) + (\beta_1 + \beta_3) SCOP + (\beta_2 + \beta_4) SCOP^2 + \varepsilon_t$$

A plot of the least square fit between OP and SCOP as shown in Fig. 2 suggests there are two distinct groups. It appears that there is some kind of discontinuity (the circled area).

B. Change Point Analysis

Change-point analysis is a powerful new tool for determining whether a change has taken place. Change-point analysis is capable of detecting subtle changes missed by control charts [8]. Further, this type of analysis better characterizes the changes detected by providing confidence levels and confidence intervals, which, by definition, suggests a range of values where likely estimates are given. This would result in a collection of likely estimates that could be statistically significantly and, therefore creates an increased reliability of the estimates. Nevertheless, when analyzing historical data, especially when dealing with large data sets, change-point analysis can be preferable to control charting [8]. A change-point analysis is more powerful, better characterizes the changes, controls the overall error rate, is robust to outliers, is more flexible and is simpler to use [6]. As we suspect a change point due to discontinuity earlier, we decided to use a non parametric estimation technique rather than a parametric estimation. One of the methods that can be used is non- parametric regression (PROC LOESS in SAS) where a particular smoothening parameter can be selected and connect the gap between these two groups (circle area in Fig. 2). Thus, we need to test for a change point at year 2001; if there is a change point, then, as we not know a suitable parametric for the regression surface, we will need to implement a non parametric regression such as local regression (PROC LOESS). Another reason to use this process is that we may suspect couple of outliers [7].

IV. Research Findings

By using the SAS program to form an ANOVA table with five degrees of freedom based on the five variables, we found that the overall model was significant, as shown in Table II. Then, by estimating the β parameters for the model, significant parameters are clear from Table II. This table reflects significant parameters for the X2 and interaction of X2 with SCOP and $SCOP^2$. As a result, the X2 parameter for the interactive term suggests that there is significance between the Saudi Crude Oil Production (SCOP) and the dummy variable of whether the oil was produced before 2001, which was before the sharp price increase, or whether the oil was produced more currently. Also, this is significant with the square of SCOP that is $SCOP^2$, which means we are capturing some non linear effect, and selection of this model is justified.

TABLE II: Least Square Estimates of Model

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
Intercept	1	13.04357	8.31112	1.57	0.1267
SCOP	1	0.00230	0.00439	0.52	0.6041
$SCOP^2$	1	-6.71113E-8	4.469588E-7	-0.15	0.8816
SCOP*X2	1	1466.70890	695.11052	2.11	*0.0430
$SCOP^2$*X2	1	-0.35997	0.16197	-2.22	*0.0337
X2	1	0.00002196	0.00000937	2.34	*0.0257

Significant at 95% Level.

Based upon these significant parameters, it is clear that the estimated Oil Price (OP_t) is significantly based upon whether we are estimating before the price increase or after. We would expect this from a logical perspective. However, that being said, more analysis must be performed to account for autocorrelation. ARMA results suggest that all the variables that were significant in OLS analysis are still significant, and we have taken care of autocorrelation. This suggests a better estimate than a model based just on OLS. The prediction model from ARIMA is shown in Fig. 3 below.

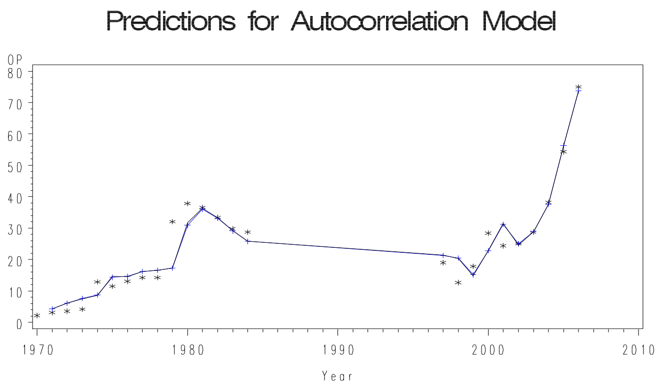

Fig. 3. Prediction model from ARIMA.

Results from the change point analysis, with a p-value is almost zero, provide some indication of the overall significance of model. The values of T_n, have a maximum value is 188.984, which corresponds to year 2001, indicating that it is reasonable to have a change point at year 2001. Studying the two cases we present in our study, stability in fluctuations in prices between 1970 to 2001 are clear, and then there is a strongly significant increase in oil prices and production between 2001-2006, for various reasons. In observing the historical events of the Saudi oil market, the war on Iraq has significantly affected the demand on Saudi oil. Specifically, Iraq was in oil market competition with Saudi Arabia. During the war on Iraq, Iraqi oil demand decreased, whereas, in contrast, the demand for Saudi oil increased, which effectively created a monopoly, or at least an oil market that reflects an oligopoly. In addition, with oil prices having doubled five times from 1970 until 2001, which includes the decline in price during the years of 1981 to 1988, the world is in the grip of an energy crunch that promises to become even more difficult for oil consuming countries. A "perfect storm" of rapidly rising demand, political instability, environmental restrictions, signs of diminishing resources, and increasingly violent weather patterns are all contributors to pushing oil prices steadily higher. Finally, but not least, we ran PROC LOESS for a local regression estimation. To do this, we have to first select an appropriate smoothening parameter. The smoothening parameter that was selected on the basis of lowest Akaike Information Criterion (AIC) criteria was 0.4. After selecting this parameter, a local regression was fitted, as shown below in Fig. 4. Based on the 37 observations, which represent data for the last 37 years, the predicted oil prices line to move almost in a linear fashion with the oil prices historical data.

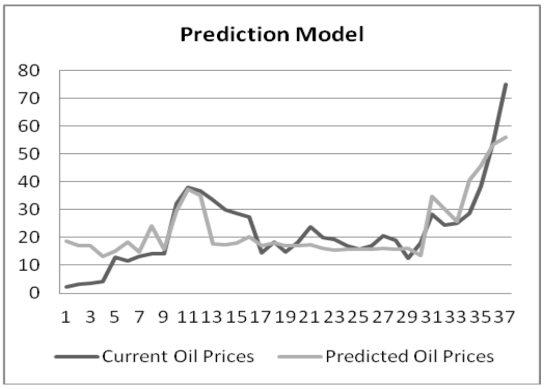

Fig. 4. Prediction model for oil prices.

V. Conclusive Remarks

It is evident that Saudi Arabia has been consistently producing oil in a linear fashion before 2001. However, any international or Middle East dispute can change the trend, as is evident from our change point analysis. Thus, it becomes difficult to predict future oil production, as this is very much subject to the geo-political situation of other oil producing countries. It is likely that Saudi Arabia may increase their oil production during such times of conflict to meet the energy requirements of the world, and especially of industrialized countries like the U.S. This may increase oil prices, as demand for Saudi oil will increase. This project systematically analyzed a times series data and built up models to overcome statistical problems such as non linearity, autocorrelation, and change point. The end result was a prediction model for oil prices using Saudi Arabia's oil production as the primary independent variable. Future research should encompass cross checking the validity of such a model for oil prices in other parts of the world, and also possibly validating some policy implications in terms of whether the largest oil producer has some impact on the global oil prices and economy.

Appendix

Formulation for the Change Point Test; Data: $X_1,\ldots\ldots\ldots\ldots,X_n$ (n=37) is given below:

$$\hat{\sigma}_t^2 = \frac{1}{n}\left[\sum_{i=1}^{t}\left(X_i - \overline{X}_{1:t}\right)^2 + \sum_{i=t+1}^{n}\left(X_i - \overline{X}_{t+1:n}\right)^2\right]$$

$$\frac{L_1(t)}{L_0} = \frac{\prod_{i=1}^{t} f\left(x_i,\hat{\mu}_{1:t},\hat{\sigma}_t\right).\prod_{i=t+1}^{n} f\left(x_i,\hat{\mu}_{t+1:n},\hat{\sigma}_t\right)}{\prod_{i=1}^{n} f\left(x_i,\hat{\mu}_{1:n},\hat{\sigma}_n\right)}$$

$$\log\frac{L_1(t)}{L_0} = \sum_{i=1}^{t}\log f\left(x_i,\hat{\mu}_{1:t},\hat{\sigma}_t\right) + \sum_{i=t+1}^{n}\log f\left(x_i,\hat{\mu}_{t+1:n},\hat{\sigma}_t\right)$$

$$-\sum_{i=1}^{n}\log f\left(x_i,\hat{\mu}_{1:n},\hat{\sigma}_n\right)$$

$$= -t.\log\sqrt{2\pi} - \frac{t}{2}\log\hat{\sigma}_t^2 -$$

$$\frac{1}{2\hat{\sigma}_t^2}\sum_{i=1}^{t}\left(x_i - \hat{\mu}_{1:t}\right)^2 - (n-t)\log\sqrt{2\pi} - \frac{(n-t)}{2}\log\hat{\sigma}_t^2 -$$

$$\frac{1}{2\hat{\sigma}_t^2}\sum_{i=t+1}^{n}\left(x_i - \hat{\mu}_{t+1:n}\right)^2$$

$$-n\log\sqrt{2\pi} - \frac{n}{2}\log\hat{\sigma}_n^2 + \frac{1}{2\hat{\sigma}_n^2}\sum_{i=1}^{n}\left(x_i - \hat{\mu}_{1:n}\right)^2$$

$$= -\frac{n}{2}\log\frac{\hat{\sigma}_t^2}{\hat{\sigma}_n^2} - \frac{1}{2\hat{\sigma}_t^2}\left[\sum_{i=1}^{t}\left(x_i - \hat{\mu}_{1:t}\right)^2 + \sum_{i=t+1}^{n}\left(x_i - \hat{\mu}_{t+1}\right)^2\right] +$$

$$\frac{1}{2\hat{\sigma}_n^2}\sum_{i=1}^{n}\left(x_i - \hat{\mu}_{1:n}\right)^2$$

$$= 2\log\frac{L_1(t)}{L_0} = n\log\left(\frac{1}{n}\sum_{i=1}^{t}\left(X_i - \overline{X}_{1:n}\right)^2\right) -$$

$$n\log\left(\frac{1}{n}\left(\sum_{i=1}^{t}\left(X_i - \overline{X}_{1:t}\right)^2 + \sum_{i=t+1}^{n}\left(X_i - \overline{X}_{t+1:n}\right)^2\right)\right)$$

$$T_n = \max_{2\le t\le n-1} 2\log\frac{L_1(t)}{L_0}$$

$$\overline{X}_{1:t} = \frac{1}{t}\sum_{i=1}^{t} X_i$$

$$\overline{X}_{t+1:n} = \frac{1}{n-t}\sum_{i=t+1}^{t} X_i$$

$$t = \left(2\log\log nT_n\right)^{\frac{1}{2}} - 2\log\log n - \frac{1}{2}\log\log\log n + \frac{1}{2}\log\pi$$

Acknowledgment

The authors would like to thank both the Canadian University of Dubai and IMF for providing support towards this research paper. The views expressed in this paper are those of the authors and do not necessarily represent those of the Canadian University of Dubai, IMF or IMF policy.

REFERENCES

[1] Saudi Arabian Monetary Agency (SAMA). (2006). The Forty First Annual Report. Research and Statistics Department. [Online]. Available: http://www.samaksa.org/newreports/annual/en/index.htm.

[2] J. Duarte. (October 2007). Today's Analysis: Report: Saudi Arabia Oil Production Tapped Out. [Online]. Available: http://www.rigzone.com/news/article.asp?a_id=26351

[3] Energy Information Administration (EIA). (August 2005). *Saudi Arabia Country Analysis Brief (SACAB)* [Online]. Available http://www.eia.doe.gov/emeu/cabs/saudi.html.

[4] The Oil Drum. (2005). Updating Saudi Oil Production Plans. [Online]. Available: http://www.theoildrum.com/classic/2005/08/updating-saudi-oil-production-plans.html.

[5] J. L. Williams. (2006). Oil Price History and Analysis, WTRG Economics. [Online]. Available: http://www.wtrg.com/prices.htm.

[6] H. Lutkepohl, *Introduction to Multiple Time Series Analysis*, Springer, Berlin, 1991.

[7] W. H. Greene, *Econometric Analysis*, New York: New York University, 2003.

[8] W. A. Taylor. (2006). Change-Point Analysis: A Powerful New Tool for Detecting Changes. [Online]. Available: http://www.variation.com/cpa/tech/changepoint.html

S. Alshahrani is currently an Economist, with the Fiscal Affairs Department of the International Monetary Fund (IMF) and is based in Washington, U.S. He previously held position at the Central Bank of Saudi Arabia (SAMA). He held some academic positions in USA & GCC universities. He holds a PhD (Economics) from Washington State University (USA) and 2 graduate degrees in applied economics and statistics.

I. Gurrib is currently the MBA Finance Program Coordinator at the Canadian University of Dubai, which is ranked no. 1 in Dubai for its MBA program. He is a holder of a PhD (Economics and Finance) from Curtin University, together with a MPA (Master of Professional Accounting) and MFin (Master of Finance) from Victoria State University, Australia. He held various academic positions in Australia, Saudi Arabia and UAE. He was also a finance manager at Westpac Bank, the 2nd largest bank in Australia.