Oil Price Dynamics: Economic Linkages, Price Wars and Forecasting Models During COVID-19

Review of Integrative Business & Economics Research

Ayaan Sohhel*
Hindu College, University of Delhi

Prashant Rao
Hindu College, University of Delhi

ABSTRACT

The COVID-19 outbreak placed millions of workers in quarantine and dissipated factories, affecting small business owners and disrupting global supply chains that are dependent on China for manufacturing. An oil price war erupted in March 2020 after the unexpected faltering of a cooperative agreement between OPEC and Russia, a pact that was crucial in forming the basis for stability in global oil markets for the preceding three years. In the month following this crisis, crude oil plummeted by more than 50% and its future contracts reached negative levels in the following month, sending shockwaves throughout international commodity markets that were already in a rut. This extreme volatility raises the question of unpredictability of oil prices due to high dependence on political turbulences, when compared to its linkages with macroeconomic factors. Studies have shown that the accuracy of univariate models tends to perform better than linear regression models, yet the latter is relevant for knowing the relation of each variable to crude oil prices. Faced with this situation, we analyze the fragilities of linear regression and the potential of the univariate ARIMA for predicting oil prices, evaluating it through the lens of the Russian Oil Crisis, OPEC+ price war and COVID-19.

Keywords: Transmission mechanism; Price war; Linear regression; Partial autocorrelation.

Received 19 March 2021 | Revised 1 July 2021 | Accepted 8 August 2021.

1. INTRODUCTION

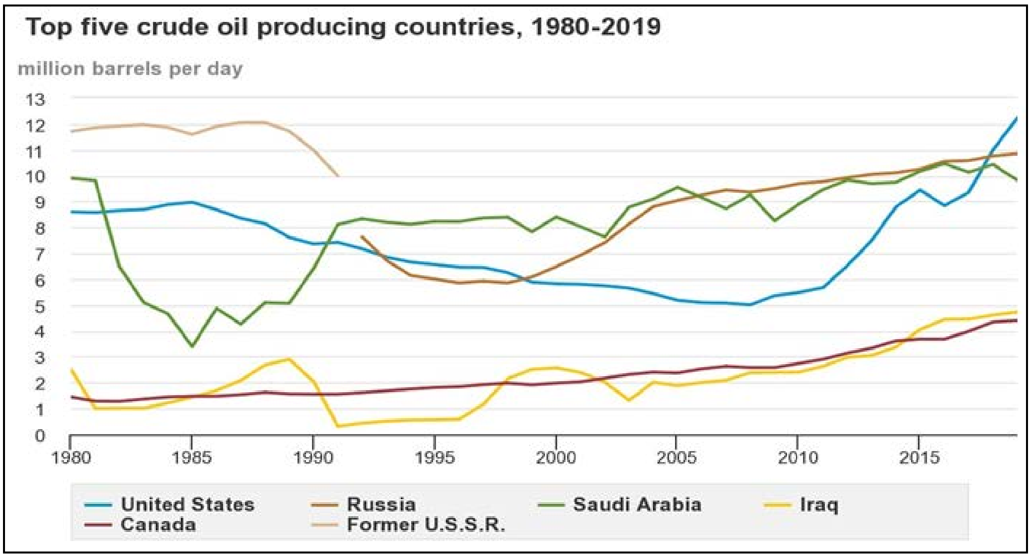
It is beneficial to have detailed oil market forecasts in many sectors of the economy. For example, both central banks and the private sector use these predictions in many scenarios to formulate monetary policy and assessing risk in the future. Industries like transportation, manufacturing and utility companies, for example, are heavily dependent on these forecasts for decision-making. Oil prices can drastically change within a short period of time. The demand and supply of supply of oil are fairly inelastic in the short run, which makes the price either plummet or skyrocket, when one exceeds the other. It is also highly influenced by geopolitical disruptions, which makes it even more difficult to predict oil prices.

The crude oil market was relatively stable before 2014. After 2014, the price has been on a declining trend overall and has been affected by various economic and political activities.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

The price of Brent, the benchmark crude oil contract, declined from a high of $112 in June 2014 to a low of $30 in January 2016, reflecting the acute volatility of prices. The price gradually recovered and spiked in October 2018 at $80 only to drop steeply from that point onwards. Similarly, the West Texas Intermediate benchmark fell from $102 in June 2014 to $31 in January 2016.

Figure 1: Oil production (by country)



Source: EIA

Two factors have been playing a major role in price movements. First, the ‘Shale Revolution’ in the United States which transformed it from one of the largest importers in the world to the largest producer in the world (Figure 1). The United States production has increased its daily production from 5.48 million barrels per day (bpd) in 2010 to 9.44 million bpd in 2015, and surpassed Russia and Saudi Arabia in 2018 with 10.99 million bpd. In 2019, the United States constituted a share of 15% in the world crude oil production. The country now finds itself in a new position and unsure how to respond to oil price fluctuations. It was simple and straightforward in the past: being a large importer, lower prices were seen as conducive to the growth of the US economy. With the US now the world’s largest producer and consumer, political leaders have been grappling with the difficulty of devising policies that balance the trade-off between the well-being of consumers and producers.

Second, Russia and nine other countries had joined forces with the 13-member OPEC in 2016. The extended group, dubbed “OPEC+” controlled almost half of the world’s oil production and led to a resurgence of the cartel. On November 30, 2016, OPEC decided to cut output by 1.2 million bpd, with top exporter and OPEC Leader Saudi Arabia cutting as much as 486,000 bpd to balance excess supply in the market, primarily by the United States. Producers from outside the 13-country group agreed to reduce output by 558,000 bpd, with Russia contributing a reduction of 300,000 bpd. It was the largest production cut since 2008 and caused a rise in the price of Brent to $50 per barrel. It was extended till December 2017 when Russia and OPEC increased the productions cuts to 1.8 million bpd until December

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

2018. Over the past 6 years, OPEC+ has increasingly played the role of a swing producer in the market, absorbing the demand that remains after non-OPEC supply is covered. In this scenario, OPEC along with its non-member collaborators are essentially 'price takers' in the market, with the United States significantly influencing the price making process.

2. LITERATURE REVIEW

Hamilton (1983) wrote the most influential paper in the field of oil price linkages with the macroeconomy, arguing that the price increases were at least partially responsible for every US recession after World War 2, except the one in 1960. In addition to this, studies by Santini (1985), Burbidge and Harrison (1984) identified the relationship between economic growth and oil price movements in the US economy. The existing literature examines the relationship by modelling the transmission channels responsible for it. The Russia Economic Report by World Bank in 2016 identifies the characteristics of its internal production structure and composition that led to the oil crisis in 2014. The lack of diversification from commodities in the economy was a major drawback and inevitably led to an economic recession when the ruble depreciated. The Russia Economic Report for 2020, highlights the events and the economic context leading up to the price war and its aftermath. The first half of this paper revolves around the interdependence of oil with other key metrics, analyzing it through the perspective of economic crises, technological innovations (Shale) and geopolitical developments. The second half of the paper is centered around the uncertainty and reliability of the short/medium-term projections of future oil prices, with a focus on WTI Crude. Traditionally, structural models have performed poorly for forecasting oil prices and univariate time series models have shown better accuracy in predicting prices (D. Lam). F. Bosler evaluated the time series approach, which included analysis of both linear and non-linear time series. Comparing the ARIMA model and a neural network autoregressive model for nonlinear time series, he concluded that the price forecasts by non-linear models showed lower error. In a separate paper, D. Lam used the Box-Jenkins method to model oil prices on a univariate time series. Using the Autocorrelation and Partial Autocorrelation functions (ACF, PACF), an ARIMA model with the appropriate parameters was chosen. In addition, Lam developed a regression model to contrast it with the results of the nonlinear model. Eight macroeconomic explanatory variables were chosen: production, consumption, net import, ending stock, utilization rate of refineries, interest rates (U.S.), an oil futures contract from NYMEX and the S&P 500 index. In this paper, we construct a structural model and the univariate Box-Jenkins model using Python's statsmodels library.

3. MACROECONOMIC LINKAGES, OPEC+ PRICE WAR AND COVID-19

3.1 Economic Linkages

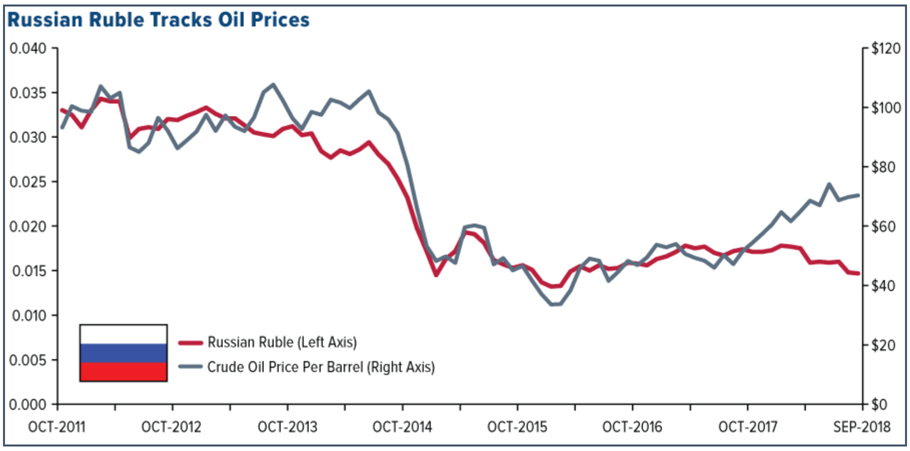
The Russian Oil Crisis of 2014 revealed that oil prices could have significant linkages with a number of macroeconomic aggregates. Russia is one of the biggest players in the oil market,

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

with both oil and natural gas constituting a major part of its GDP growth. As a result, its revenues are highly dependent on oil exports, which constitutes 44% of the national budget. Due to the dependence on commodities, the economy has a relatively higher exposure to the risk of crude oil prices and exchange rate fluctuations. Russia's export revenue has a high susceptibility to the volatility of the RUB/USD exchange rate, but the exchange rate tends to have a positive correlation with oil prices, while oil itself is a source of large money inflows (Figure 2 shows the positive relation between the RUB/USD exchange rate and oil prices). This trade off reflects the macroeconomic risks associated with the Russia's economic policy, with a rise in oil prices resulting in appreciation (driving down export demand) of the currency but also increasing the prices of exports simultaneously. To ameliorate these policy risks, in 2015 Russia started accepting payments in the form of rubles and yuan for its export sales, a policy known as "de-dollarization", effectively reducing its dependence on the US dollar, Engdahl et al. (2017). It was combined with an effort to spread investment risk across non-commodity sectors to reduce dependence on natural resources, which is highlighted by the divergence towards the end of 2017 (Figure 2).

The oil price shock of 2014 caused the GDP growth rate to fall to -0.5% in Q2, 2014 and the UN sanctions in 2015 further affected the growth of GDP negatively, contracting by 3.7%. The low oil prices resulted in steep depreciation of the ruble, which led to a sharp rise in inflation. Due to falling real wages and increasingly poor quality of consumer debt, consumption demand was negatively affected (Verma and Gupta, 2018). The increase in doubtful account write-offs by stressed PSBs led to a rise in interest rates, consequently resulting in the decline of private investment and therefore, capital formation (Table 1). The free-float exchange rate mechanism also allowed the imports to adjust quickly to depreciation, with import quantities falling by 25.7%. The contraction of GDP was accompanied by the rise of unemployment and rise in countercyclical fiscal expenditure by the government (Figure 3).

Figure 2: Ruble vs Crude Oil



Source: Bloomberg, L.P.

But on the other hand, oil price increases have also been shown to cause inflation in other countries, with oil prices directly influencing the prices of products using petroleum inputs.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

They indirectly have an effect on costs such as transportation, manufacturing, and heating, as sellers and manufacturers pass the production costs on to consumers, lowering the purchasing power of the domestic currency. This case was especially highlighted with the United States prior to the Shale Revolution, when it was a net importer of oil. A rise in oil prices exerts a downward pressure on aggregate supply because higher energy prices indicate that firms relying on energy inputs will buy a lower volume of energy. Consequently, the productivity of a fixed amount of capital and labour will fall, and productive capacity contracts. The decline in factor productivity results in lower real wages, negatively affecting aggregate demand, which will trigger a recession. This transmission mechanism results in a clear negative relationship between economic growth and oil prices, with the dynamics being more pronounced in the medium-term than the short-term. If an aggregate demand-side shock due to swinging oil prices occurs during a business cycle fluctuation, or a period of economic turmoil, it is found to have exhibited a heightened correlation between oil prices and economic growth (Ftiti, Guesmi et al., 2016).

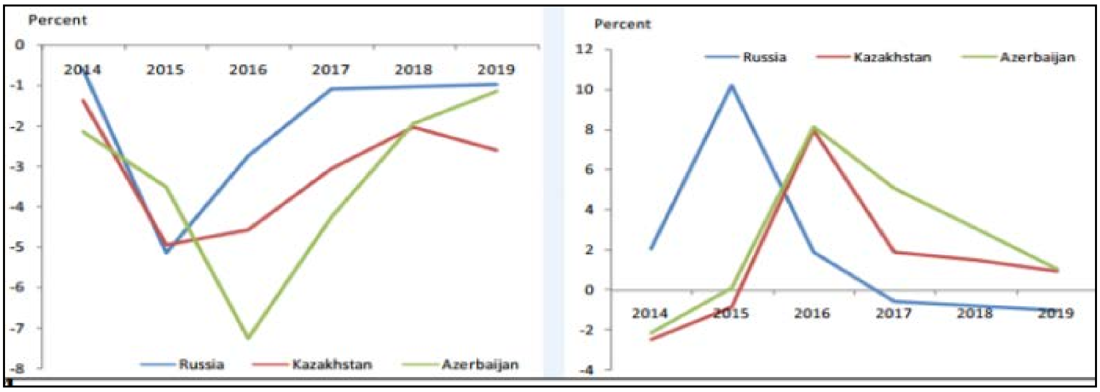
Another crucial part of oil prices declining in 2014 originated with the expansionary monetary policy of the Federal Reserve and other central banks prior to and following the subprime mortgage crisis of 2008. It resulted in an oil price bubble due to an excessive amount of cheap credit and leverage, raising the speculative demand for oil before and while the global economy was in recession (Tagizhadeh-Hesary, 2017). From this we can infer that the interplay of oil prices with the exchange rate, GDP, investment, consumption and imports are highly influenced by the degree of diversification in an economy or its internal production structure, leverage in financial markets and its export-to-import ratio for energy products.

Table 1: Oil prices vs Macroeconomic aggregates in Russia

	2015	2016	2017	2018
Oil price (US$ per barrel, WB average)	51.9	43.3	55.2	59.9
GDP growth, percent	-3.7	-0.6	1.5	1.7
Consumption growth, percent	-7.5	-2.5	2.0	1.6
Gross capital formation growth, percent	-18.7	1.9	6.0	4.9
General government balance, percent of GDP	-3.5	-4.2	-2.5	-0.5
Current account (US$ billions)	69.0	27.6	26.5	25.4
Current account, percent of GDP	5.2	2.2	1.8	1.6
Capital and financial account (US$ billions)	-86.1	-27.4	-26.4	-25.4
Capital and financial account, percent of GDP	-5.3	-2.2	-1.8	-1.6
CPI inflation (average)	15.5	7.1	4.5	4.0

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Figure 3: i) Oil price effect on GDP, ii) inflation for OPEC+ Members

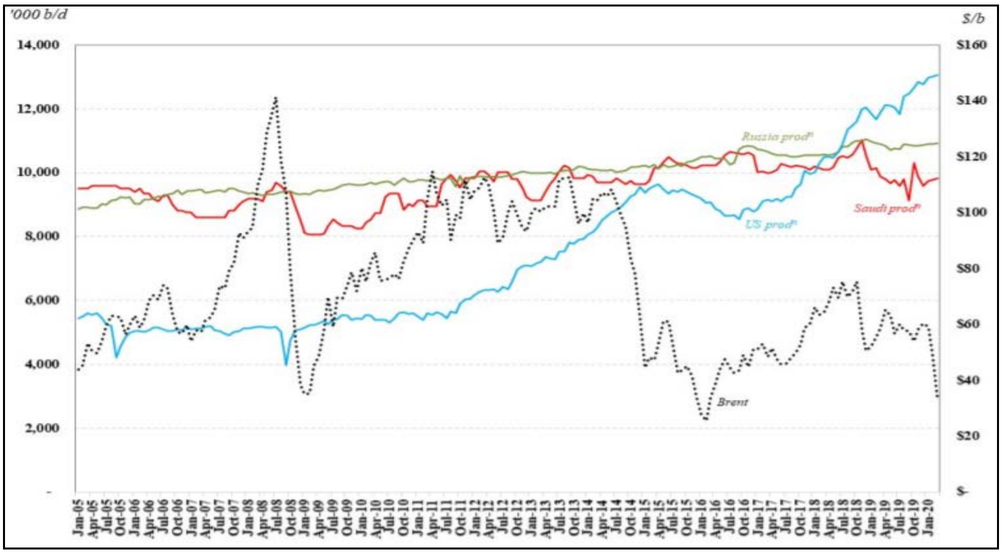


Source: cbr.ru (Table 1), World Bank

3.2 Dynamics of OPEC+ Price War During COVID-19

During the OPEC+ meetings in Vienna in early March, Saudi Arabia had proposed for a collective production cut of 1.5 Mb/d, over the existing cuts of the past 2 years. The Kremlin was apprehensive about the additional cuts to its oil production, believing that its three-year pact with OPEC had further propelled America's shale oil industry. An attempt to combat the competitiveness of shale had been attempted before as well. When the new technology (shale) was advancing in 2014, Saudi Arabia adopted a strategy to ramp up production and flood the market, expecting that a plunge in prices would scotch the new competition. But as shale drillers and refiners found more cost-effective ways to operate, a global supply surplus dragged on and OPEC inevitably resorted to its signature tool of constraining output. But the Saudi (and OPEC+) decline was accompanied by a persistent rise in U.S shale production by 4.4 million bpd, which has been subverting Saudi's position as the dominant oil supplier of the world (as evident in Figure 4), resulting in oil prices remaining relatively stagnant.

Figure 4: Production vs Brent (blue-USA, red-Saudi, grey-Russia)



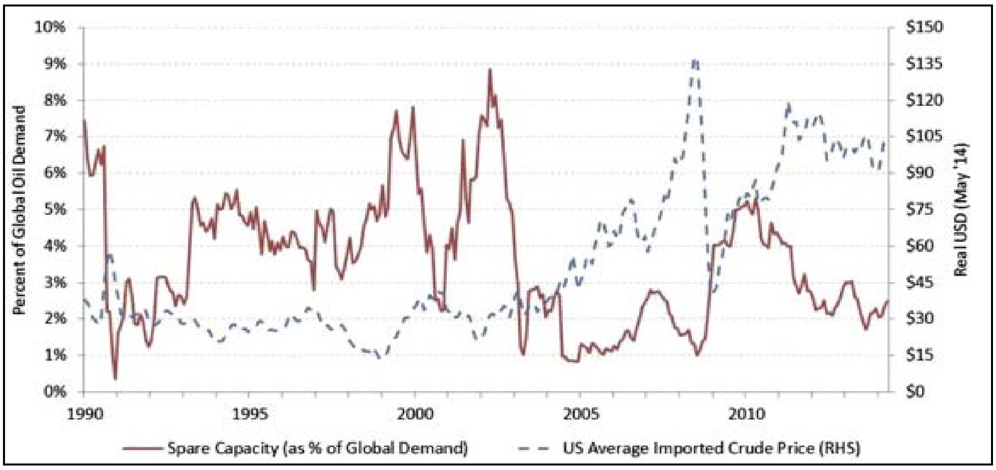
Source: IEA Oil Market Report

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

After Russia refused to follow the production cuts, Saudi Arabia, in an unanticipated move, abruptly responded by sanctioning a record of 12.3 million bpd in April and raising sustainable production capacity to 13 million bpd. The Oil Minister also commissioned $30 billion in capital investment, effectively launching a price war with non-OPEC countries (Ratcliffe, Di Paola et al., 2020).

So why was a price war launched? The dynamics of a price war can be analyzed through the lens of a simple game theory construct, a dominant producer like Saudi Arabia can extract monopolistic rents from residual demand i.e demand that cannot be supplied at a specific price by other producers, as a result of being a low-cost producer. This strategy indicated that maintaining "spare capacity" and withholding production is an efficient outcome for the dominant player, all other things being equal. But in the current day's scenario with exogenous factors like shale production expanding productive capacity significantly, a production cut by Saudi Arabia can backfire due to the inelasticity of shale oil supply. Saudi Arabia can then employ an increase in production as a deterrent to punish its competitors, with its cost advantage remaining intact. Spare capacity is the volume of production that can be brought on within 30 days and lasts for at least 90 days and indicates the degree to which production is below full capacity.

Figure 5: Spare Capacity vs. Crude Price of US Imports



Source: EIA

Usually, Saudi Arabia, utilises its spare capacity (which is the largest in the world) to mitigate large swings in oil market prices and to prevent the global economy from being negatively affected by a volatile, cyclical market. These fluctuations in spare capacity are transmitted to the price of crude imports in non-OPEC economies, like the United States (Figure 5). Lower spare capacity tends to accompany an upswing in the price levels, and a sharp rise in prices are followed by an expansion of spare capacity. It helps Saudi Arabia in attaining the goal of stabilising oil demand in the long-run. But in the case of the price war, the intent was to amplify volatility to capture a higher market share in the oil industry. The Saudi and Russian utilisation of spare capacity resulted in increases of around 3 million bpd of low-cost oil to

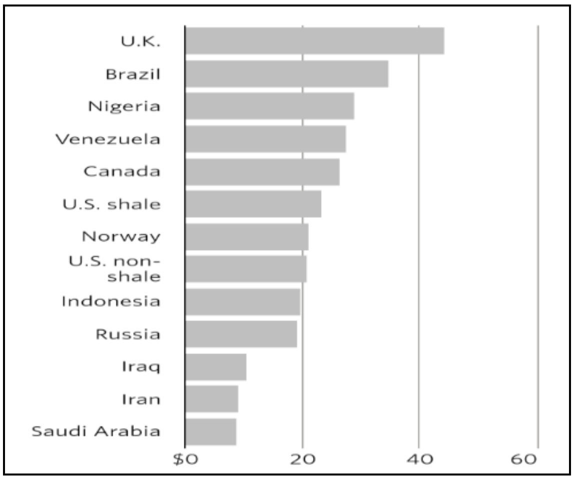
ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

global supply, waging a downward pressure on crude prices (Pierru et al., 2018). Waging a price war is endemic and sometimes an involuntary part of deploying the dominant producer strategy as it gestures a threat to higher cost competitors, and other low-cost producers as well.

The end objective is that the sacrifice of profit margins in the short-term (lower prices) for the dominant player is followed by a long-term increase in market share, but at a lower price. The prices will stabilise in the long-run with the use of alternate fuels being discouraged by the threat of another price war.

The Kingdom's average production cost per barrel of crude oil is 5-10 USD, while the average production cost for the shale oil drillers and refiners in the US is approximately within the range of 25 USD and 35 USD (Figure 6). A price war would have left the US shale oil producers with no other choice but to halt production by around two to three million barrels, forcing at least a few companies to either comply and face losses or file for bankruptcy. This is especially effective when a competitor producer's supply becomes more elastic - as shale supply has recently, becoming more responsive to global price movements. This would push the dominant producer, Saudi Arabia, to further punish their competitors.

Figure 6: Oil Production Costs

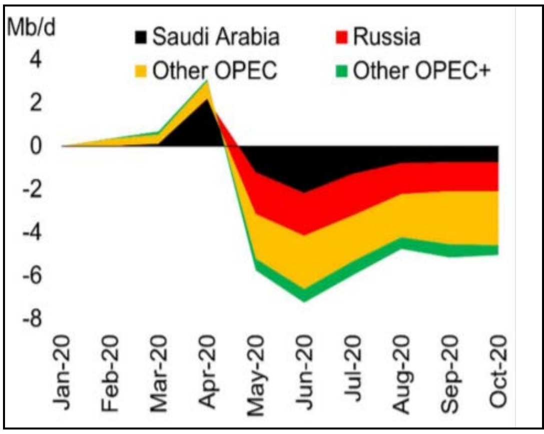


Source: World Bank

Although they had accounted for the effects on worldwide demand from Covid-19, in March Saudi forecasted only a loss of around 1 million bpd of global oil demand growth in 2020. At that point in time, the degree of demand destruction from the pandemic had not been fully realized. Shortly thereafter, the panic of Covid-19 began to kindle more substantial action in the countries that were affected, including the United States, the largest oil consumer. As more countries directed its citizens to quarantine, an unexpectedly large amount of global transportation was halted. It was estimated by EIA that the global oil demand had reduced by 33%, or 35 million bpd **[7]**.

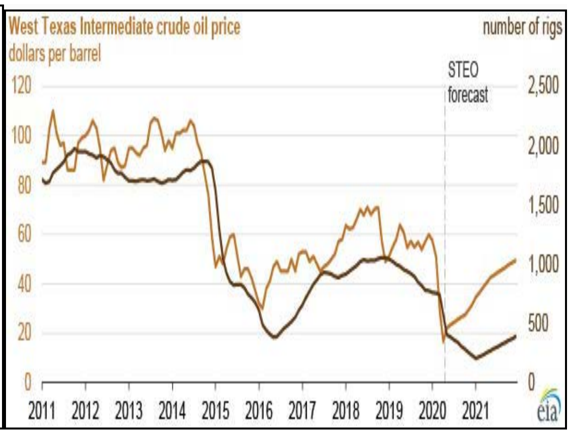
ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Figure 7a: OPEC+ Production



Source: World Bank; OPEC

Figure 7b: WTI vs US Oil Rig Count



Source: EIA, STEO

In April, future contract prices plummeted to negative levels, production fell significantly, with OPEC+ eventually agreeing to cut production by a historically large amount of 9.7 m/bd from May (Figure 7a). In Russia, production of crude oil fell from a mean of 11.3 mb/d in Q1 to 9.4 mb/d from May to July. Shortly after (August-December), production was raised by 2 mb/d and further by 0.5 mb/d (Finley et al., 2020). In retrospect, the launching of the price war was self-sabotaging and a major lapse in foresight.

Countries outside of OPEC+ cut production as well, declining by around 20% in the United States between April and May (Figure 7b). It has mildly recovered since then, with a moderate rise in the number of oil rigs in November.

4. ECONOMETRIC MODELS

4.1. LINEAR REGRESSION

We have discussed the effects of crude price movements on the economy and the transmission channels responsible for influencing a sequence of economic variables. But causation in one direction does not imply that it is true in the opposite direction. In this section, we attempt to build a model that predicts the price of crude oil, using a structural model. Structural modelling uses explanatory variables (independent) and computes its relationship with the response variable (dependent) to make forecasts. The model we create in this paper, uses the multivariate linear regression model, to find the best fitting regression equation. It is implicit that the spot price of oil has a linear correlation with the historical values of each explanatory variable.

There are a few essential conditions for the model that needs to be tested. First, there should be an absence of multicollinearity among the explanatory variables. This indicates that the parameters used for predicting oil prices should not be highly correlated to each other (<0.75). The explanatory variables for the WTI spot price are Production, Sales (Consumption), Ending Stocks (inventory of oil), Imports and economic performance across various sectors,

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

which is reflected by the S&P 500 index (Lam, 2013). The second condition is that the variables should also have constant error variance across its values, a property known as homoscedasticity. The errors of the explanatory variables should not be correlated with each other and should be normally distributed, which can be checked by QQ plots.

Figure 8: Correlation matrix of WTI Crude Oil

Source: Generated from Python's seaborn library

Model Specification: We plot a correlation matrix using the seaborn library in Python (Figure 8). It suggests that the variables taken into account for predicting crude oil prices have an issue of multicollinearity (when two explanatory variables have a correlation above the threshold value 0.75). The variables are Production and the S&P 500. Multicollinearity can be corrected by using regularization techniques. Only two of the covariates are highly correlated, while another pair, Sales and Imports, have a medium correlation of 0.62. The Elastic net regularization would be suitable for such a regression, which is a combination of L1 and L2 regularization with both terms being multiplied by a weight.

$$0.5 * RSS/n + alpha * ((1 - L1\_wt) * |params|_2^2/2 + L1\_wt * |params|_1)$$

where RSS is the residual sum of squares, alpha is the penalty weight for L2, n is the sample size and |params| indicates the L1 and L2 norms. (using 0.5 as L1_wt).

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Table 2: Regression Analysis Summary (WTI Price for thousand barrels)

	Coefficients	Std. error	t-Statistic	P-value
Intercept	3.813e+05	3.03e+04	12.594	0.000
Production	-0.1060	0.041	-2.616	0.011
Consumption	0.0547	0.062	0.878	0.382
Imports	-0.0071	0.090	-0.079	0.937
Ending Stock	-0.1660	0.012	-13.879	0.000
Regression Statistics				
R-squared	0.770			
Adj. R-squared	0.759			
F-statistic	68.64			
Observations	87			
Skew	-0.272			
AIC	1863			

Source: Generated from Python's statsmodels library

Note:

1. The metrics above were generated using data since 2013 [10], with the 2014 Russian Oil Crisis and Shale supply changing the nature of macro-relationships
2. The S&P 500 Index has been omitted from the model because it has a high correlation with Production, one of the explanatory variables. This will prevent the issue of endogeneity within the model.

Analysis of Results: There is a negative relationship between crude prices and production. In a business cycle expansion, the supply (Production in Table 2) will catch up to and eventually outstrip global demand once economic growth starts to decelerate, exerting a downward pressure on prices. But if the cycle is accompanied by an accomodative monetary policy, the formation of a speculative bubble is possible, postponing the price decline to a later date (2008 global financial crisis). There is a positive relationship between crude prices and consumption. This is likely due to the supply lag that tends to persist in oil markets, since supply elasticity is relatively lower. The oil producers' expectations of higher demand in the future requires significant capital investment by drillers and refineries for them to expand. It leads to the production having a slower growth rate than demand for a period of time, resulting in a temporary shortage that drives prices upward. The p-value of 0.382 indicates that it does not have a significant effect on crude prices.

The regression model indicates a slight negative relationship between crude prices and imports. This can be explained by international competitiveness with foreign exporters, an increase in imports implies a potential increase in market share for foreign companies. Domestic refiners will respond by lowering their market price for a higher market share. The p-value of 0.937 indicates that imports does not have any effect on oil price movements. The

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

negative coefficient for Production is likely due to the reversal of the transmission mechanism through aggregate demand, as stated in the previous section.

The accuracy of the model can be tested by applying it on the dataset with all the explanatory variables. The forecasted data from the fitted model is compared to the true values of the WTI prices and the root mean square error (RMSE) is calculated. The predictions were carried out using the beta coefficients and intercept, with the following computation:

Root Mean Square Error: 21.5192

When compared to the average price range of crude oil, the RMSE is too high to make efficient predictions using the regression model. The model's inefficiency is also reflected by the p-values of explanatory variables. With the exception of Sales, all the variable's p-values breach a significance level of 5%. The reasons for these shortcomings can be attributed to heavy dependence of crude prices on geopolitical disruptions. This could include anything from an oil tanker attack on the Persian Gulf to a country's political leader refusing to cooperate with OPEC. Therefore, it can only be applied in periods with low volatility in the data trends.

4.2. AUTOREGRESSIVE INTEGRATED MOVING AVERAGE (ARIMA, BOX-JENKINS APPROACH)

Now we test a univariate time series model for forecasting, since they are more effective at dealing with higher levels of volatility, particularly for volatility clustered at specific points in financial time series. These models use the methods of autoregressive functions and moving averages of historical prices and returns, which are applied to forecast future prices. ARIMA is also referred to as the Box Jenkins approach. It is a linear model and can predict non-stationary time series, by differencing the data and converting to a stationary series.

The ARIMA (p, d, q) parameter model for variable Yt is shown in the general equation below. The parameters are:

- p: number of autoregressive terms
- d: number of orders of differencing required for stationarity
- q: number of moving average terms;

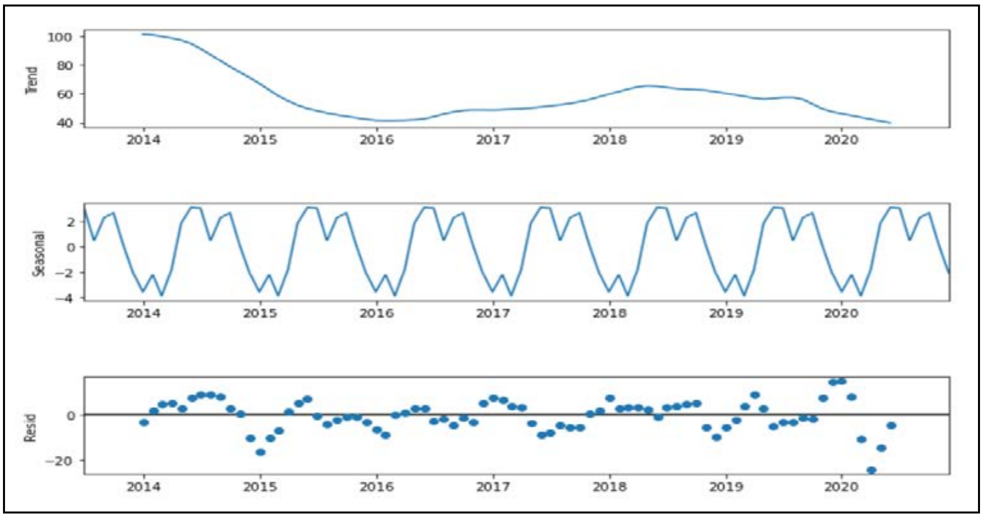
$$Y_t = \phi_0 + \sum_{i=1}^{p} \phi_i Y_{t-i} + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

Here φi's are the autoregressive "AR" terms and represent lags of stationary series and θj's are moving average "MA" terms and represent lags of the forecast errors. Here εs represents white noise, with zero mean and zero correlation of its values across time. A stationary series is one with constant mean and variance across time. This implies that a stationary series does not depend on the time at which the observation is recorded. A non-stationary series can become stationary by differencing it to the nth order.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

We generate Figure 9using the seasonal_decompose function from the statsmodels library in Python to check the 3 components of the WTI Spot Price data: **trend, seasonality, and randomness**.

Figure 9: Decomposition of Data (Source: Python statsmodels library)



(Source: Python statsmodels library)

There is an element of seasonality in the data, which means it affects the value of time series at different times, indicating that it is probably not stationary. The residual component of the data tends to revert to a mean of zero.

Pre-processing Data: In order to validate the data, we apply a test known as the **Augmented Dickey-Fuller test** on the differenced oil prices to determine whether it is stationary (Enders, 2015). The following was the output from Python's adfuller function:

Augmented Dickey-Fuller Test:
ADF Test Statistic : -6.0081228234168 79
p-value : 0.08579536295493445
#Lags Used : 1
Number of Observations Used : 87
weak evidence against the null hypothesis. Data has a unit root, indicating it is non-stationary.

Since the p-value is higher than 0.05, there is a high probability that the time series is non-stationary and we cannot use ARIMA with d=0. We difference the data further for it to become more stationary. This means that for every data point at time 't' the value of the data at 't+1' will be deducted from it. After differencing again, this is the output returned by the adfuller function:

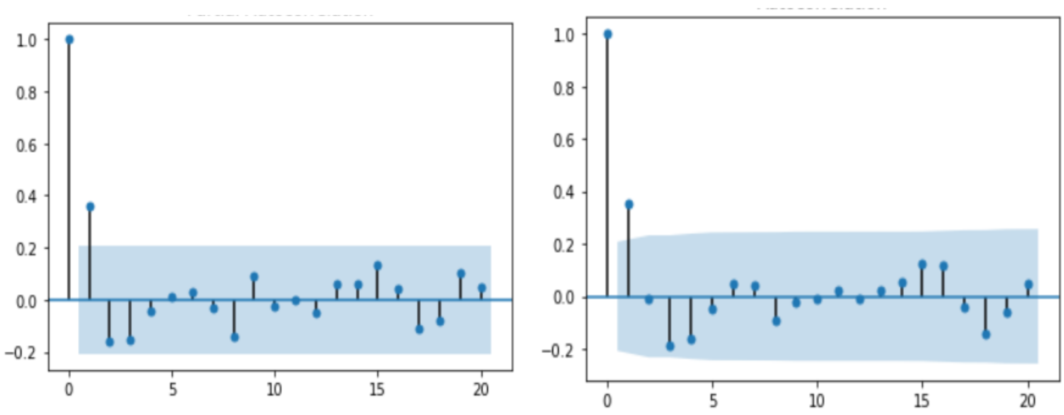
Augmented Dickey-Fuller Test:
ADF Test Statistic : -7.141622300681548

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

p-value : 3.3127603370557743e-07
#Lags Used : 4
Number of Observations Used : 83
strong evidence against the null hypothesis, reject the null hypothesis. Data has no unit root and is stationary.

Since the data is now stationary, we can expect an order of differencing d=1. A partial autocorrelation function (PACF) is also used to verify the order of differencing, giving the partial correlation of a stationary time series with its own lagged values. Unlike an Autocorrelation function (ACF), it controls for correlation between shorter lagged values. The ACF and PACF plots of the first differenced data with different values of t (difference in time period), also suggests that the differenced data is stationary, since there is only one significant spike at x=1 (Figure 10).

Figure 10: i) PACF After Differencing, ii) ACF After Differencing



Source: statsmodels library (Python)

The parameters of the ARIMA model are decided by executing the all variations of the model and finding the combination with the highest log likelihood. Using the statsmodels tool SARIMAX, we find that the optimal model is found to be ARIMA (0,1,2) [0,1,1,12], with seasonal parameters in the square brackets. It is a second-order moving average and with no autoregressive term, with a seasonal 'MA' term. The (0,1,2) variation is a damped Holt's Model or damped exponential smoothing (Holt's Linear Trend Method). The exponential smoothing of the first moving average term is flattened out by the second moving average term, by extrapolating local trends towards the end of the dataset and adjusting it for a longer time horizon. The forecasts are made using a level equation to estimate the level of the series at time 't', and a trend equation to estimate trend (slope) at time 't'. Smoothing parameters (from 0 to 1) are used in both equations. The metrics for the fitted ARIMA model are given below (Table 3).

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Table 3: ARIMA Statistics Summary

	Coefficients	Std. Error	t-Statistic	p-value
ma.L1	0.438	0.128	3.424	0.001
ma.L2	0.1273	0.098	1.295	1.95E-01
ma.S.L12	-0.8779	0.432	-2.03	0.042
sigma2	30.5983	12.067	2.536	0.011
ARIMA Statistics				
Log Likelihood	-248.95			
Heteroscedasticity	2			
Observations	90			
Skew	-0.27			
AIC	505.90			

Source: Generated from Python's statsmodels library

Note: These metrics are not generalizable across all countries, the model above uses WTI Spot Price data. The results will be different for Brent Crude Oil.

Analysis of Results: It can be inferred that for the prediction of crude prices, exponential smoothing is much more significant than the autoregressive component. The p-values of the components are all under significance level 5% except for ma.L2, but it generates a higher log likelihood than the other variations. The largest coefficient is attached to sigma2, which indicates the white noise in the data. It reflects that price volatility is largely explained by the exogenous effects due to geopolitical turbulences. The accuracy of the model can be tested by applying it on the historical dataset, and the forecasted data from the fitted model is compared to the true values of WTI prices. The root mean square error of the values is computed:

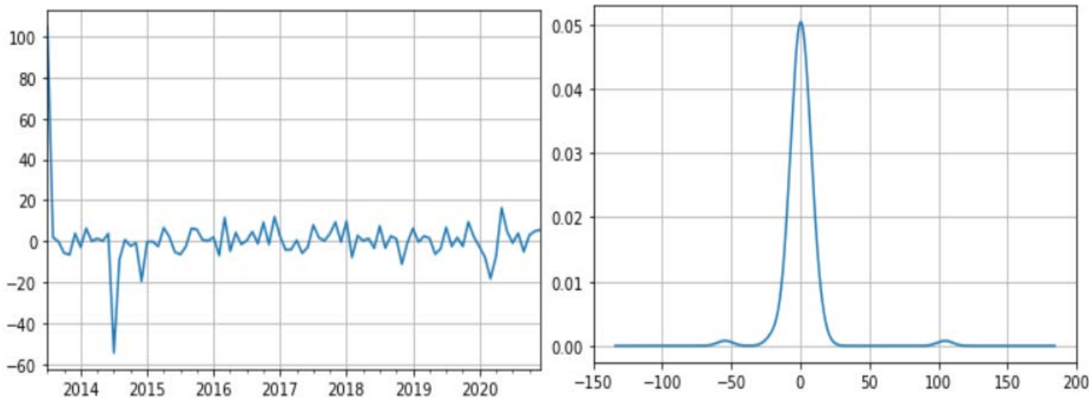
Root Mean Square Error: 8.42203

The RMSE and the AIC (Alkaike Information Criterion) are much lower, and we can conclude that the ARIMA method's predictive accuracy is better than the structural model. We can verify this by analysing the distribution of the fitted model's residual error terms (Figure 11).

Overall, the model is more suitable for forecasting than the OLS regression model because it addresses the drawback of volatility clustering of linear regression. We use the fitted ARIMA model to predict crude oil prices in the future after the slump it experienced this year (Figure 12).

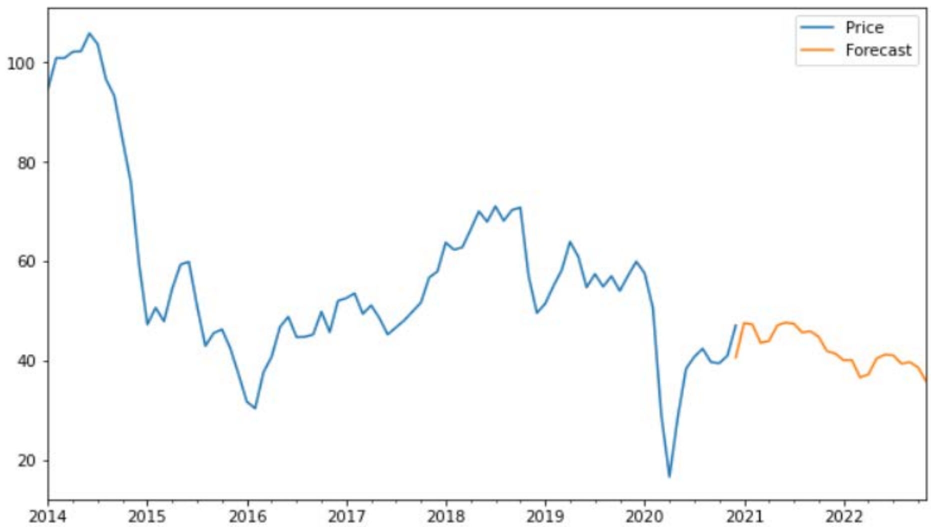
ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

Figure 11: i) Plot of Fitted Residuals, ii) Probability Density of Residual Terms



Source: Python's Seaborn Library

Figure 12: ARIMA Forecast (till November, 2022)



Source: Python's seaborn library

The model suggests that stability in US crude oil markets will regain its footing and prices will be relatively stable in the next two years, barring any exogenous shocks. Crude prices will follow a mild downward trend from mid-2021 to the beginning of 2022, at which point average monthly prices are forecasted to take a dip to around 35 USD, rebounding to 40 USD shortly after that. The prediction's accuracy depends on the correctness of assuming minimal geopolitical disruptions between the United states and OPEC+, or OPEC's internal conflicts with Russia.

5. CONCLUSION

In Section 2, the structural OLS regression model was explored with 5 explanatory variables being validated for the model. Due to the correlations between covariates, it was regularized

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

using a combination of L1 and L2 regularization (elastic net). The fitted model was applied on a dataset, finding the RMSE. The error was too large to make meaningful forecasts, due to clustering of volatility. To address this issue, a univariate linear ARIMA model was constructed. The best fitting parameters was found to be (0,1,2) [0,1,1,12] seasonal-damped exponential smoothing, with the root mean square error of the fitted model being less than half that of the OLS model. However, due to the complex cost dynamics of oil, even the most fine-tuned models can have very large errors in predicting future prices of oil, particularly when it is forecasting for the long-term.

In the absence of a major Middle Eastern conflict or price war, crude oil prices should stabilize at least partially by the first quarter of 2021. The ARIMA model in this paper extrapolates a rise from the trough of the mid-April 2020 level to remain within the range of a monthly average price of $35 and $50 per barrel throughout 2021, forecasting a low in the beginning of 2022. However, the forecast is heavily dependent on future production decisions by OPEC+, the momentum of oil demand growth and the elasticity of U.S. oil production with respect to prices. The market forces from the effects of COVID-19 as well as the conflicts among OPEC, Russia and the United States are still likely to linger, which could trigger fluctuations outside of the forecasted range. A bilateral effort must be made between the power centres to ameliorate the risks of a potential oil price shock in the future. A shock would exacerbate the recession in oil importing countries, affecting aggregate supply, consequently pushing real wages, and hence aggregate demand further down.

ACKNOWLEDGEMENT

We are indebted to Dr. Puja Saxena and Mrs. Nidhi Dhamija (Department of Economics, Hindu College) for their invaluable insights and whose guidance allowed us to think laterally and tightened our grasp over the process of writing a research paper. We would like to thank them for making this memorable experience.

REFERENCES

[1] Derek Lam (2013), Time Series Modelling of Monthly WTI Crude Oil Returns, University of Oxford; https://core.ac.uk/download/pdf/13501939.pdf

[2] Engdahl, F. W. (2017). *Russia and China Challenge US Dollar Domination*.

[3] Finley et al., IEA Oil Market Report, February 2020 and Baker Institute; from https://www.bakerinstitute.org/media/files/files/b24814f4/ces-pub-oilpricecrash-040620.pdf

[4] Ftiti, Guesmi et al. (2016), Journal of Applied Business Research, *Relationship Between Crude Oil Prices and Economic Growth in OPEC Countries*; https://core.ac.uk/download/pdf/268105506.pdf

[5] Pierru, Smith and Zamrik (2018), *OPEC's Impact on Oil Price Volatility: The Role of Spare Capacity*, The Energy Journal 39.

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)

[6] Ratcliffe, Di Paola et al. (2020), Saudi Arabia Pledges to Expand Oil Output Capacity, Bloomberg News; from https://www.bloomberg.com/news/articles/2020-03-11/aramco-will-boost-oil-output-capacity-to-13-million-barrels-day

[7] Robert Nau, Arima Models For Time Series Forecasting, Duke University; https://people.duke.edu/~rnau/seasarim.htm

[8] Tagizhadeh-Hesary (2017), *Trade Linkages and Transmission of Oil Price Fluctuations in a Model*; https://www.adb.org/sites/default/files/publication/363561/adbi-wp777.pdf

[9] Verma, Gupta (2018), *Russia's Policy Framework in a Multifaceted Macro Risk Environment*, Review of Integrative Business and Economics Research, Vol. 7 (s2), 200-212.

[10] World Bank, Russia Economic Report, 2020; http://pubdocs.worldbank.org/en/520231608062784328/Russia-Economic-Report-44-in-English.pdf

[11] Holt's Linear Trend Method, from https://otexts.com/fpp2/holt.html

ISSN: 2304-1013 (Online); 2304-1269 (CDROM); 2414-6722 (Print)