**Forecasting the Nominal Broad US Dollar Index**

**Using Time-Series Models**

**Sreeram Santhinilayam Ramankutty**

**INTRODUCTION**

The financial markets are complex and unpredictable. It is challenging to come up with a model that is reliable and accurate has always been a challenge for investors as well as analysts. The primary motive of this project is to predict the Nominal Broad US Dollar Index (NBUSDX). This project also aims to gain potential investment insights based on forecast insights. Therefore, the main goal of this project can be stated as predicting the Nominal Broad US Dollar Index for 16th March 2024.

I’m a short-term trader investing mainly in international stocks. I use the forecast to gain valuable insights for my investments abroad to gain profits. The Nominal Broad US Dollar Index and investment in stocks abroad are highly related. Buying international stocks when expecting the dollar is going to strengthen can be a good investment strategy. If the dollar strengthens after I purchase the foreign stocks and if I decide to sell these stocks after the change in value, I receive more dollars back in the exchange than the dollars I used to purchase these stocks. I’m currently interested in Switzerland stocks and I’m ready to spend 1000 CHF every day if I predict the Nominal Broad US Dollar Index is going to increase tomorrow.

According to my forecast, the Nominal Broad US Dollar Index on 16th March 2024 will be between 120.7636 and 120.821. The exchange rate on 15th March 2024 is 120.7884. Thus, I would like to purchase 1000 CHF worth of shares because of the higher probability of the exchange rate index going above 120.7884.

**Loss Function**

My investment strategy involves daily transactions based on a forecast of the Nominal Broad US Dollar Index. Each day I face a decision which is either to buy 1000 CHF worth of shares or not to buy. I make this decision solely by predicting the NBUSDX in a particular way. If I predict the NBUSDX will increase tomorrow, I proceed with buying the shares, and if I see that the NBUSDX will reduce tomorrow, I do not buy the shares. The decision and forecast horizon align perfectly since the shares are purchased after the forecast and are made to sell the very next day.

Here, we are trying to see the amount of profit/loss according to the forecast accuracy/mistakes. Consider the latest NBUSDX as of 16th March 2024 is 122. If I predict the rate to be 123 tomorrow, I will buy the 1000 CHF worth of shares. If the actual NBUSDX is 124 tomorrow, I gain a $2000 profit. Similarly, I can also lose money if my prediction is wrong. If I predict 123 and the actual is 120 tomorrow, I lose $2000. I also lose the opportunity to gain profit if I make mistakes in the forecast. If I predict the NBUSDX to fall from 122 to 121 tomorrow, I do not buy. But, if the actual NBUSDX has increased to 123, I lost the opportunity to gain $1000 profit. Similarly, if the actual NBUSDX has reduced tomorrow (like to 120), my loss is 0.

It is important to see here that the loss increases linearly with the size of the mistakes. It is illustrated as an example below:

*If the exchange rate falls to 120:*

*If the exchange rate falls to 118:*

Thus, we can see that the loss increases linearly as the size of the mistakes increases.

The analytical form of the loss can be described below:

The loss function is a function of three variables: Current exchange rate rt, my forecast of tomorrow’s exchange rate rft+1, and tomorrow’s exchange rate rt+1.

L(rt, rt+1, rft+1) = D1(rt+1- rt)qt + (1-D1)D2(rt+1- rt)qt

D1 = 1 when (rft+1 > rt) i.e., when I predict the rate will increase

D2 = 1 when (rt+1 > rt) i.e., when the rate for tomorrow increases.

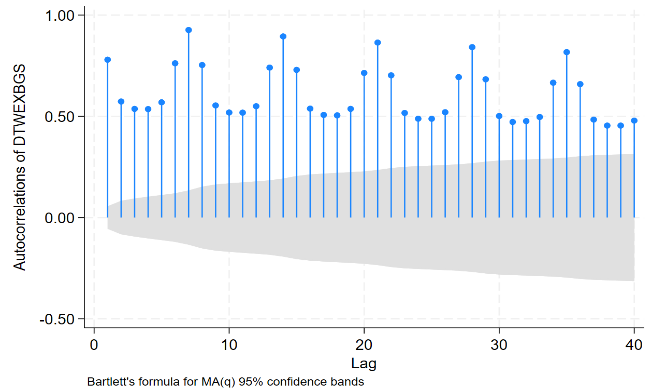
Checking Examples

1. L(rt = 122, rt+1 = 124, rft+1 = 123) = 1\*(124- 122)\*1000 + (1-1)\*1(124- 122)\*1000 = 2000
2. L(rt = 122, rt+1 = 120, rft+1 = 123) = 1\*(120- 122)\*1000 + (1-1)\*0\*(120- 122)1000 = -2000
3. L(rt = 122, rt+1 = 123, rft+1 = 121) = 0\*(123- 122)\*1000 + (1-0)\*1\*(123- 122)1000 = 1000
4. L(rt = 122, rt+1 = 120, rft+1 = 121) = 0\*(120- 122)\*1000 + (1-0)\*0\*(118- 122)1000 = 0

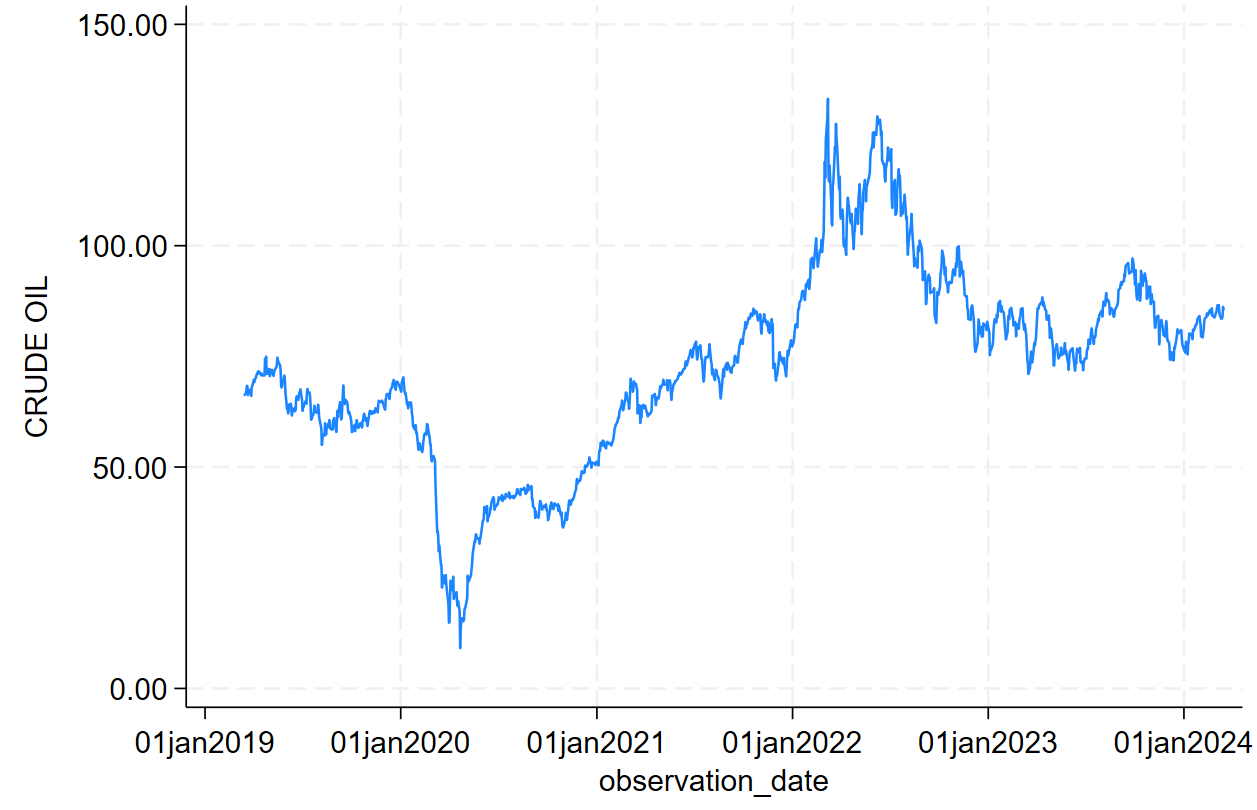
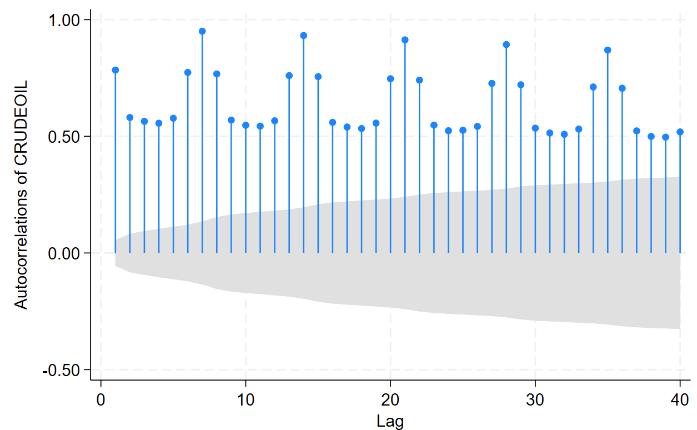
The analytical form gives the same answers that we predicted from intuition.

**Data & Stationarity & Exogeneity**

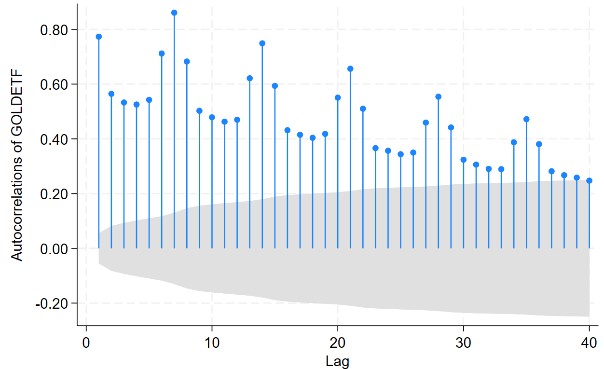
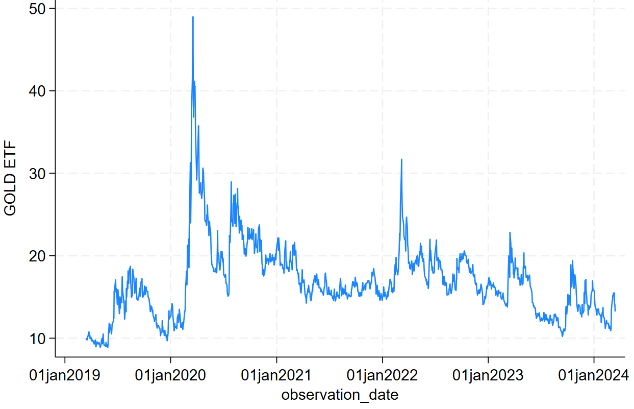
The main variable in interest is the Nominal Broad US Dollar Index. We have data from 15th March 2019 to 15th March 2024. The data is generated daily. When we plot the NBUSDX, we get:



It is important to note here that we don’t see a trend in the plot. We can observe a seasonal pattern but it is not clear. There is evidence of non-stationarity in the model. When we carefully look at the autocorrelation graph, we see that the highest spikes reduce very slowly. This is an indication of non-stationarity. However, formal testing is required to check for stationarity.

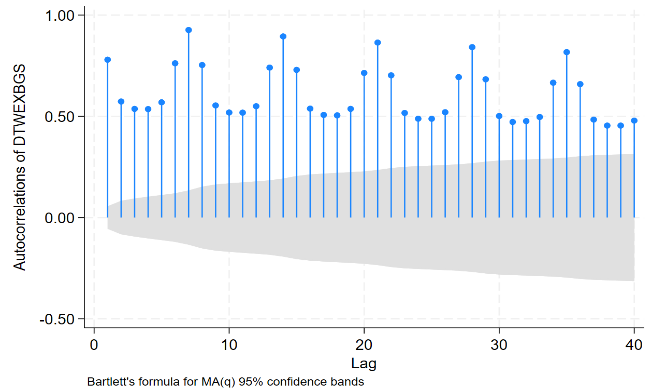


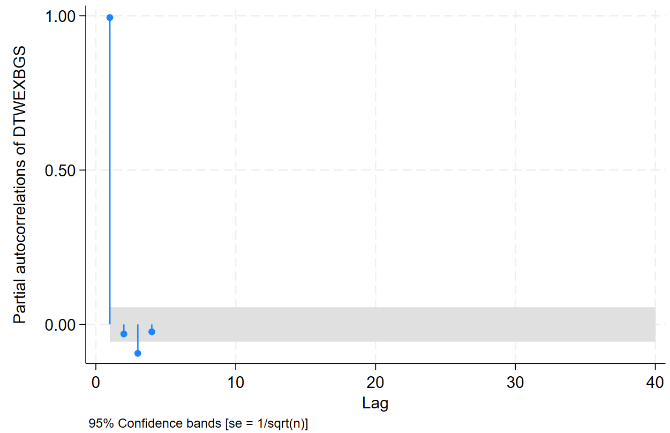
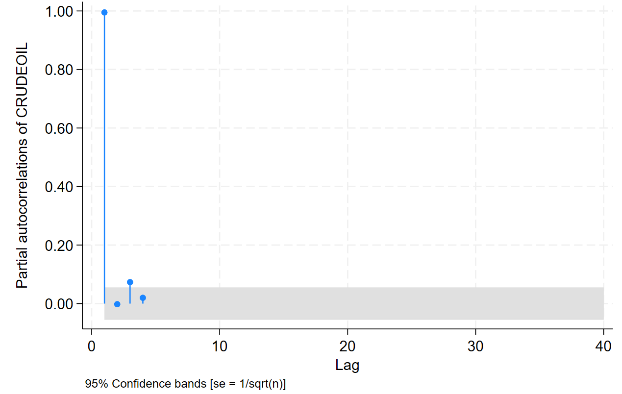
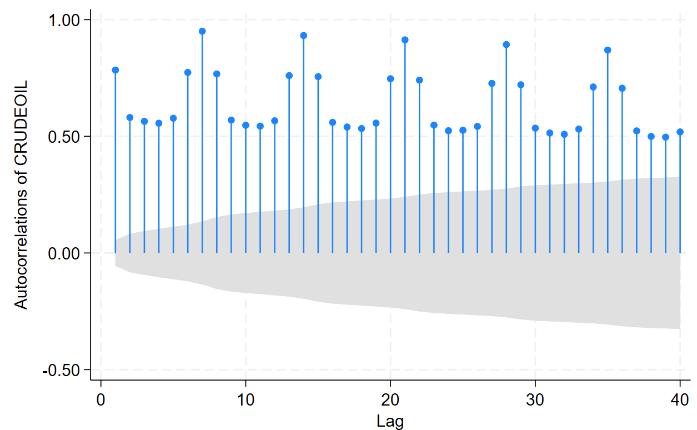
The crude oil prices data ranges from 15th March 2019 to 15th March 2024. It is a daily generated variable. From the graph of the series as well as from the autocorrelation graph, we can gain valuable insights. Even for the crude oil prices, we do not see a trend in the plot. We observe a seasonal pattern in the autocorrelation graph. There is clear evidence of non-stationarity in the model from the graph. We can’t imagine the crude oil diagram in a box as well as the autocorrelation function is showing a slower decline (looking at the highest spikes), which is an indication of non-stationarity. However, formal testing of non-stationarity is required.

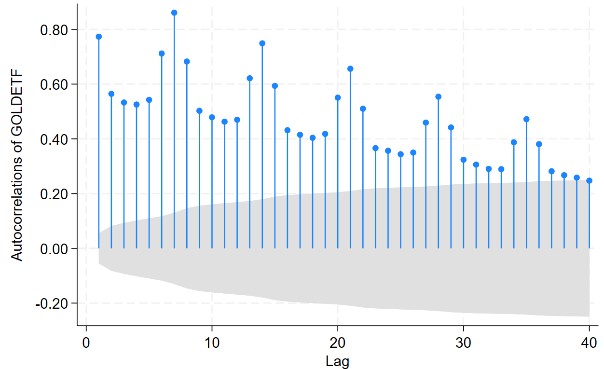
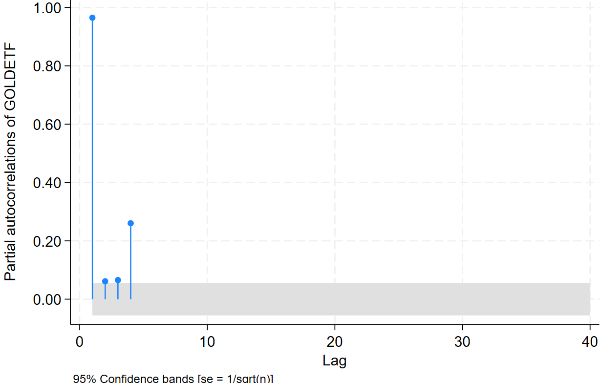
****

Similar insights can be procured from the CBOE Gold ETF Volatility Index. The CGETV is generated daily. I have data from 15th March 2019 to 15th March 2024. The graph of CGETV cannot be imagined in a box which is an indication of non-stationarity. The autocorrelation also shows very slow decline which is again an indication of non-stationarity. However formal testing is required to determine whether the series is stationary or not.

I do not take logs or do similar transformations because I do not see a trend in any of the series.

Plotting autocorrelation partial-autocorrelation of Nominal Broad US Dollar Index



****

From the autocorrelations and partial autocorrelation graphs, it is clear that all the series are showing signs of non-stationarity. (Write about partial autocorrelation).

Performing the unit-root test is crucial for any series to check whether it is stationary or not. First, I conduct the unit-root test for NBUSDX. Since, there is no trend, but a particular shift of mean, which is a drift, I use the tau-mu test. The number of lags selected is 6 because lag 6 is significant and I removed all the lags above that. (The selection of lags is given in APPENDIX). The null hypothesis of the test is that the series has a unit root or the data is non-stationary. The alternate hypothesis is that the data is stationary.

(Full output in the APPENDIX)

H0: Random walk with drift, d = 0

t-distribution

Test -------- critical value ---------

statistic 1% 5% 10%

--------------------------------------------------------------

Z(t) -1.658 -2.331 -1.647 -1.283

--------------------------------------------------------------

p-value for Z(t) = 0.0489

The output shows that the test statistic is less negative at a 5% significance level. Therefore, I cannot reject the null hypothesis at the 5% level. Thus, we cannot reject that the NBUSDX series is non-stationary.

The unit-root test for crude oil series also involves similar processes. Since there is no trend but a significant shift in the mean, I use the tau-mu unit-root test. The number of lags selected is 5 because lag 5 is significant. I remove all the lags above that. (The selection of lags given in the APPENDIX).

Augmented Dickey–Fuller test for unit root

Variable: CRUDEOIL Number of obs = 1,102

Number of lags = 5

H0: Random walk with drift, d = 0

t-distribution

Test -------- critical value ---------

statistic 1% 5% 10%

--------------------------------------------------------------

Z(t) -1.286 -2.330 -1.646 -1.282

Since the test statistic is not more negative than any of the critical values, we do not have evidence to reject the null hypothesis. Therefore, we do not have evidence to say that the series is stationary.

The unit-root test of GOLDETF is also following similar processes. Since there is no trend and evidence of a shift of mean, I performed the tau-mu unit-root test. The number of lags selected is 3 because lag 3 was significant. I removed all the lags above 3 for the test.

Augmented Dickey–Fuller test for unit root

Variable: GOLDETF Number of obs = 1,082

Number of lags = 3

H0: Random walk with drift, d = 0

t-distribution

Test -------- critical value ---------

statistic 1% 5% 10%

--------------------------------------------------------------

Z(t) -3.153 -2.330 -1.646 -1.282

--------------------------------------------------------------

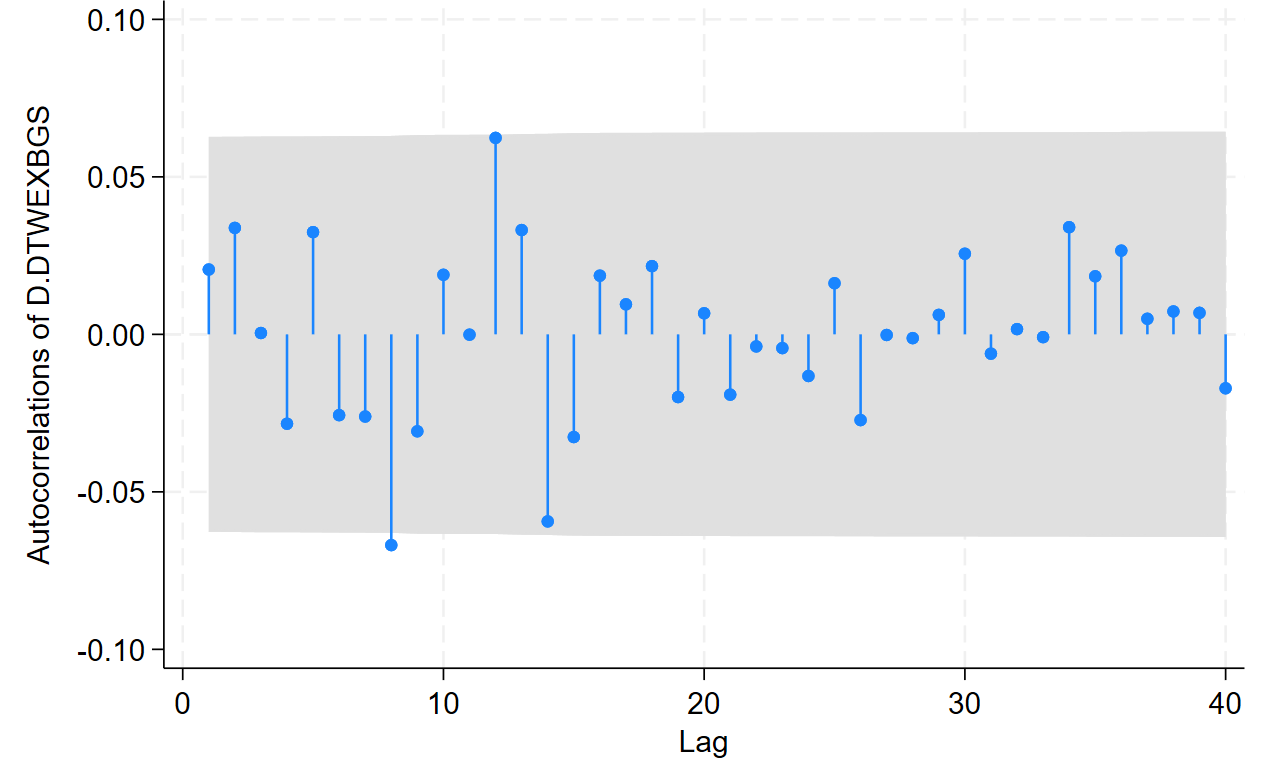
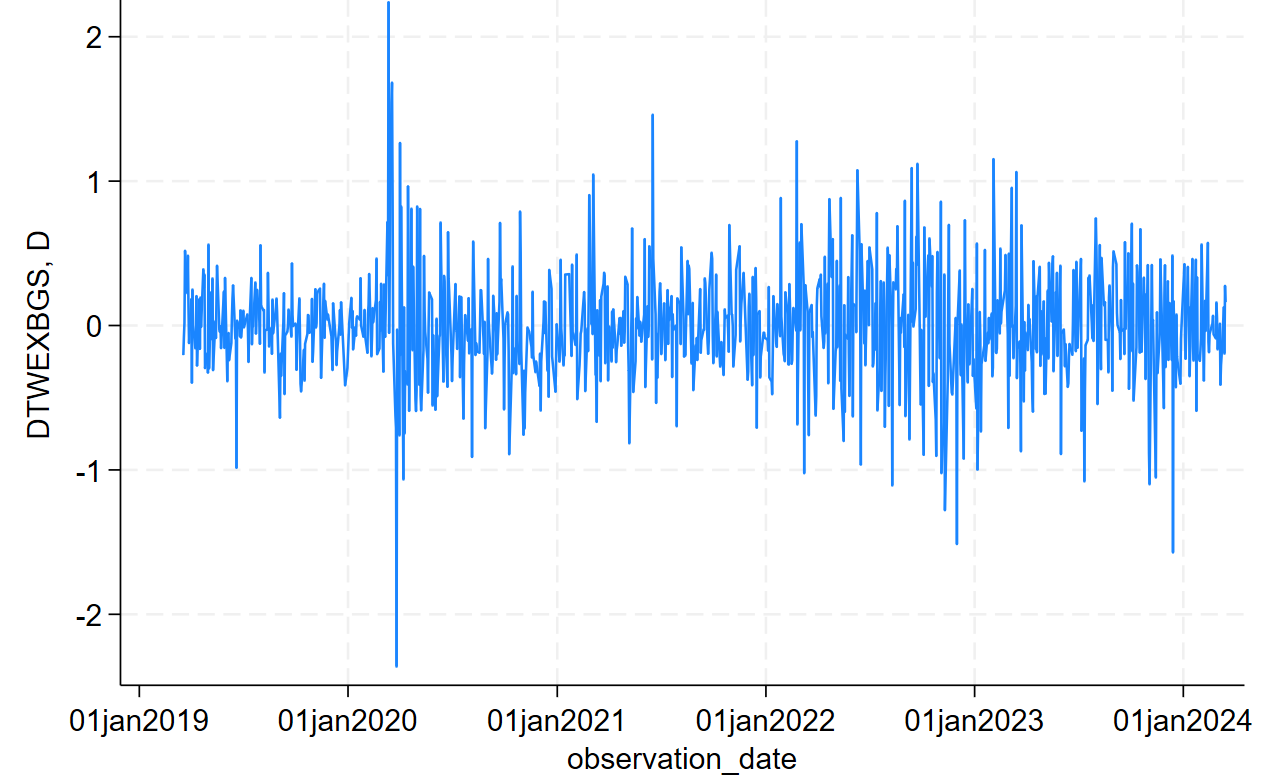
p-value for Z(t) = 0.0008

The unit-root test output for GOLDETF shows evidence of rejecting the null since the test statistic is more negative than the critical value. However, since the autocorrelation function not showing an exponential decline, which is an indicator of non-stationarity, we have to transform the series to stationary either way.

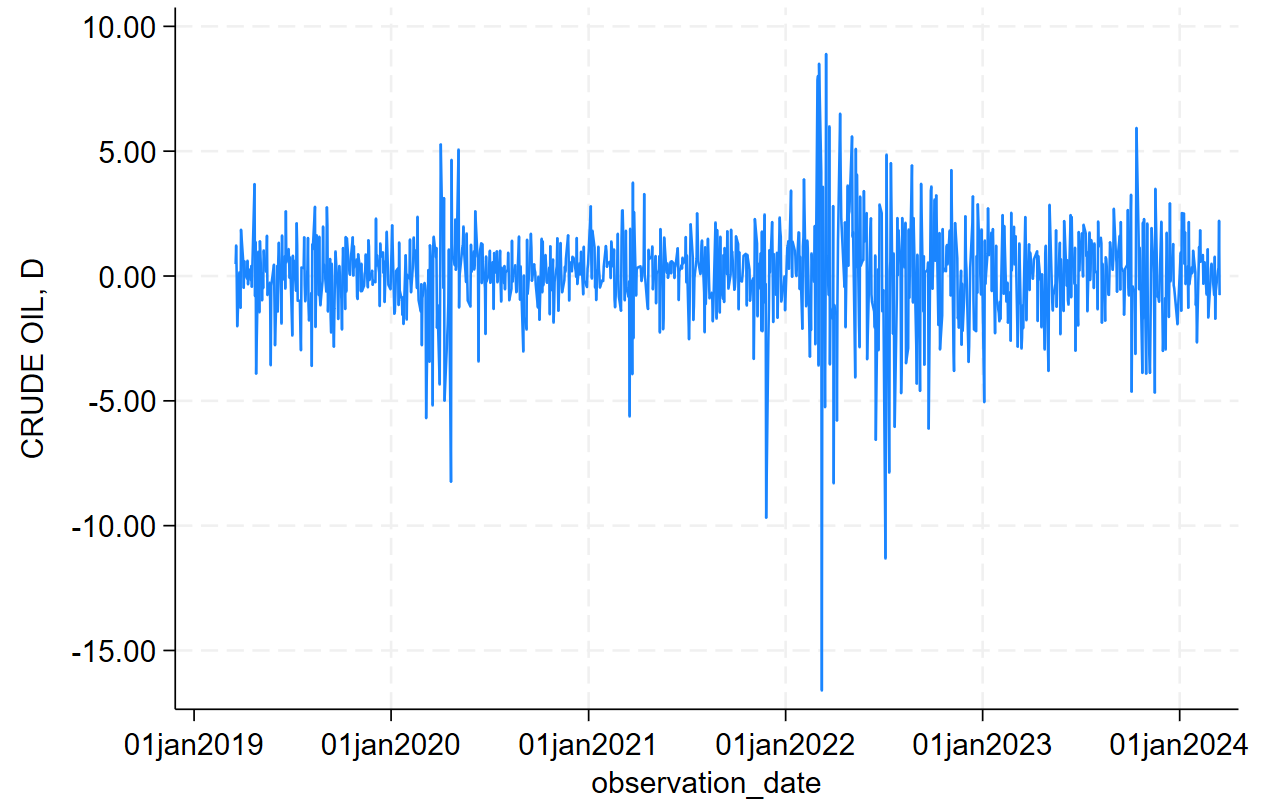
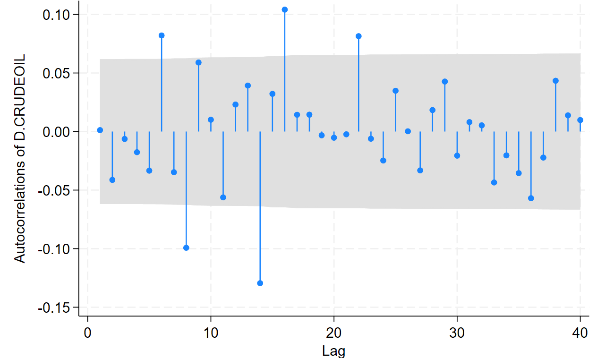
To transform all the series into stationary, I take the first difference of all the series. Let’s plot and see if the transformed variables show an indication of stationarity.

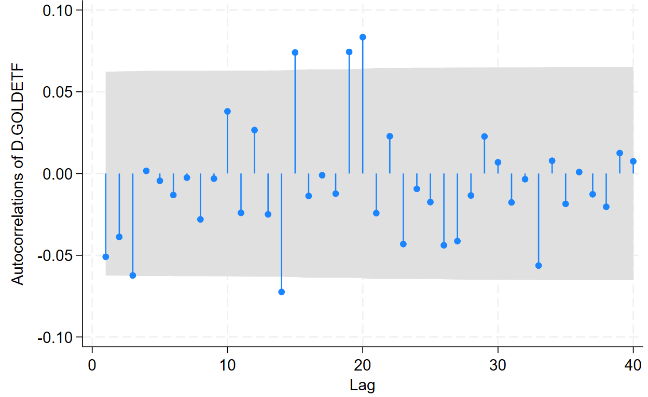
**After transformation figures:**

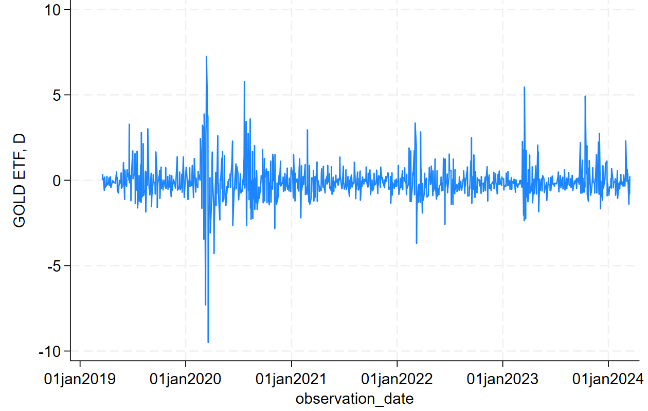
**Nominal Broad US Dollar Index:**



**Crude Oil Prices**



**Gold ETF:**



From observing all the transformed series and its autocorrelations, we find that the variables are out of indications of non-stationarity. All the series can be imagined in a box and the autocorrelations of all the variables are inside the confidence bounds which is good when looking for stationarity.

After differencing all the series, we do not see a trend or drift in any of the series. Hence, the tau test is the most suitable unit-root test. The number of lags selected for NBUSDX was 6 since lag 6 is significant. The number of lags selected for CRUDEOIL is 4 since lag 4 is significant. The number of lags selected for GOLDETF is 3 since lag 3 is significant. We reject the null hypothesis for unit-root tests for all three variables since the obtained test statistics for all three were more negative than the critical values at a 5% level. Hence, we conclude that all three variables have been transformed to stationary by taking the first difference of the series.

**How do crude oil prices help in forecasting the nominal broad US dollar index?**

The US exports tons of oil. Thus, higher prices of crude oil lead to more revenue from these exports. Exports lead to more inflow of foreign currency, exchanged for dollars increasing the demand of dollar, which potentially make the dollar strong. An oil price increase can lead to higher inflation. To control this US would increase the interest rate, which attract foreign investors, making the dollar stronger. Similarly, lower interest rates can showcase a slowing down economy. It can lead to reduced investments from abroad. Thus, weakening the dollar.

**How Gold ETF Volatility Index help in forecasting the nominal broad US dollar index?**

Investors are uncertain about the future price of gold when the index is high because of various economic concerns. It directs the investors to invest in dollar-denominated assets, potentially strengthening the dollar. Similarly, when the index is low, it indicates stability in gold prices. Investors move to gold assets rather than dollar-denominated assets. Thus, weakening the dollar.

**Granger causality test**

Output:

vargranger

Granger causality Wald tests

+------------------------------------------------------------------------+

| Equation Excluded | F df df\_r Prob > F |

|--------------------------------------+---------------------------------|

| D\_CRUDEOIL D.GOLDETF | 2.9649 3 759 0.0314 |

| D\_CRUDEOIL D.DTWEXBGS | .59816 3 759 0.6163 |

| D\_CRUDEOIL ALL | 1.6173 6 759 0.1394 |

|--------------------------------------+---------------------------------|

| D\_GOLDETF D.CRUDEOIL | 3.0186 3 759 0.0292 |

| D\_GOLDETF D.DTWEXBGS | 4.8502 3 759 0.0024 |

| D\_GOLDETF ALL | 3.6306 6 759 0.0015 |

|--------------------------------------+---------------------------------|

| D\_DTWEXBGS D.CRUDEOIL | 1.3253 3 759 0.2649 |

| D\_DTWEXBGS D.GOLDETF | 7.4963 3 759 0.0001 |

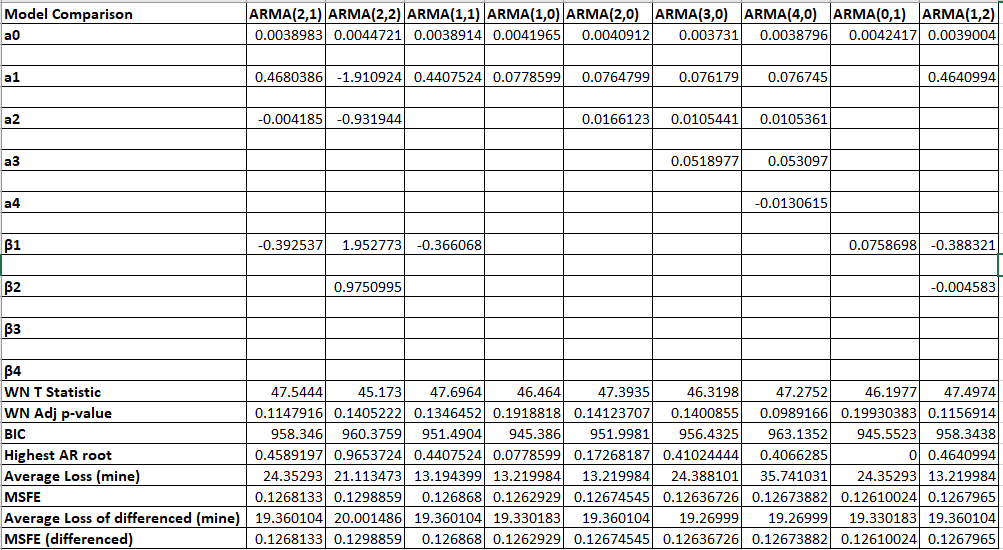
| D\_DTWEXBGS ALL | 4.4447 6 759 0.0002 |

|  |
| --- |
|  |

From the output, it is clear that Granger causality was found between the GOLD ETF and the Nominal Broad US Dollar Index. It implies that historical changes in Gold ETF contain predictive information about future changes in the US dollar index. There was no significant Granger causality seen between crude oil prices and the Nominal Broad US Dollar Index. Given this interdependence between variables, we can say that both US Nominal Broad Dollar Index and Gold ETF can be considered as endogenous variables. Since there was no reported Granger causality between Crude oil and NBUSDX, we can say that the crude oil prices is an exogenous variable. Given these findings, VAR would be the most suitable approach compared to ARIMAX.

**Univariate Estimation**

9 ARMA models are compared according to various criteria. It is interesting to see that all the models compared here have residuals which showcases white noise. The highest AR root among the models lies within the unit-root circle. The summary of the test criteria among all 9 models is given below. ‘a0’ represents the model constant while ‘a1’, ‘a2’, ‘a3’ and ‘a4’ represent AR coefficients. Similarly, the Beta values represent the MA coefficients.



\*\*Average Loss (mine): The computed loss function in the original series

\*\*Avg loss differenced (mine): The computed loss function in the differenced series

\*\*MSFE (differenced): The mean squared forecast error in the differenced series

We chose ARMA (1,1) to be the best model among all because of the lowest average loss according to the loss function we computed before. The residuals are white noise and the BIC value is also lower compared to other ARMA models. Even in the different form, ARMA (1,1) gives a lower loss (not the lowest) which is a good indication. This proves that ARMA (1,1) is a good model that helps in the prediction of the Nominal Broad US Dollar Index.

**Multivariate Estimation**

I would like to forecast the Nominal Broad US Dollar Index with the help of both Crude Oil prices and the Gold ETF Index. As mentioned before, both the variables were non-stationary and I took the first difference to transform it into stationarity. For the multivariate estimation I need to select a VAR model with a suitable lag. For that I conducted a separate test for lag selection and the output is given below:

. varsoc d.CRUDEOIL d.GOLDETF d.DTWEXBGS if t<1045

Lag-order selection criteria

Sample: 6 thru 1044, but with gaps Number of obs = 723

+---------------------------------------------------------------------------+

| Lag | LL LR df p FPE AIC HQIC SBIC |

|-----+---------------------------------------------------------------------|

| 0 | -3081.08 1.01786 8.53133 8.53867 8.55035\* |

| 1 | -3062.08 37.991 9 0.000 .990101 8.50368 8.53305\* 8.57976 |

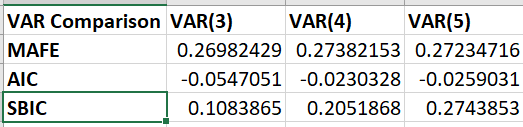
| 2 | -3058.27 7.6137 9 0.573 1.00443 8.51805 8.56943 8.65118 |

| 3 | -3038.2 40.146\* 9 0.000 .974133\* 8.48742\* 8.56082 8.6776 |

| 4 | -3033.35 9.7006 9 0.375 .985385 8.4989 8.59433 8.74614 |

The HQIC is pointing to lag 2. However, since the AIC and FPE (Final Prediction Error) are pointing to VAR with lag 3, we choose the VAR model with lag 3 for the multivariate estimation.

I also compared VAR with 3 lags to VAR with lags 4 and 5.



The VAR with the least MAFE, AIC, and SBIC values are VAR with 3 lags. Hence it is optimal to choose VAR with 3 lags for multivariate estimation.

The VAR model with 3 lags output is shown below.

**| Coefficient Std. err. t P>|t| [95% conf. interval]**

**-------------+----------------------------------------------------------------**

**D\_CRUDEOIL |**

**CRUDEOIL |**

**LD. | .0688916 .0368529 1.87 0.062 -.0034541 .1412373**

**L2D. | -.039525 .0372894 -1.06 0.290 -.1127276 .0336777**

**L3D. | -.0363207 .0372497 -0.98 0.330 -.1094454 .036804**

**|**

**GOLDETF |**

**LD. | -.040513 .0652853 -0.62 0.535 -.1686742 .0876482**

**L2D. | .0067883 .0667658 0.10 0.919 -.1242793 .1378559**

**L3D. | -.1958665 .0668647 -2.93 0.003 -.3271283 -.0646047**

**|**

**DTWEXBGS |**

**LD. | -.0137505 .1989321 -0.07 0.945 -.4042729 .3767719**

**L2D. | .2277818 .1979128 1.15 0.250 -.1607396 .6163033**

**L3D. | .1194443 .1971604 0.61 0.545 -.2676002 .5064888**

**|**

**\_cons | .0347471 .0766259 0.45 0.650 -.1156768 .1851709**

**-------------+----------------------------------------------------------------**

**D\_GOLDETF |**

**CRUDEOIL |**

**LD. | .0578008 .0206196 2.80 0.005 .0173226 .0982789**

**L2D. | .0187322 .0208638 0.90 0.370 -.0222254 .0596897**

**L3D. | .0032127 .0208416 0.15 0.878 -.0377013 .0441267**

**|**

**GOLDETF |**

**LD. | -.012914 .0365277 -0.35 0.724 -.0846214 .0587934**

**L2D. | -.0232812 .0373561 -0.62 0.533 -.0966148 .0500524**

**L3D. | -.1246064 .0374115 -3.33 0.001 -.1980487 -.0511642**

**|**

**DTWEXBGS |**

**LD. | .2263581 .1113044 2.03 0.042 .0078571 .4448592**

**L2D. | .1759652 .1107341 1.59 0.112 -.0414163 .3933467**

**L3D. | .2868528 .1103132 2.60 0.009 .0702977 .5034079**

**|**

**\_cons | .0080992 .0428729 0.19 0.850 -.0760644 .0922629**

**-------------+----------------------------------------------------------------**

**D\_DTWEXBGS |**

**CRUDEOIL |**

**LD. | -.0074816 .0068529 -1.09 0.275 -.0209346 .0059714**

**L2D. | -.0060662 .0069341 -0.87 0.382 -.0196785 .0075461**

**L3D. | -.0097172 .0069267 -1.40 0.161 -.0233151 .0038806**

**|**

**GOLDETF |**

**LD. | .0443562 .0121401 3.65 0.000 .0205241 .0681883**

**L2D. | -.0136654 .0124154 -1.10 0.271 -.0380379 .0107072**

**L3D. | .0356685 .0124338 2.87 0.004 .0112598 .0600771**

**|**

**DTWEXBGS |**

**LD. | .0848956 .0369922 2.29 0.022 .0122764 .1575148**

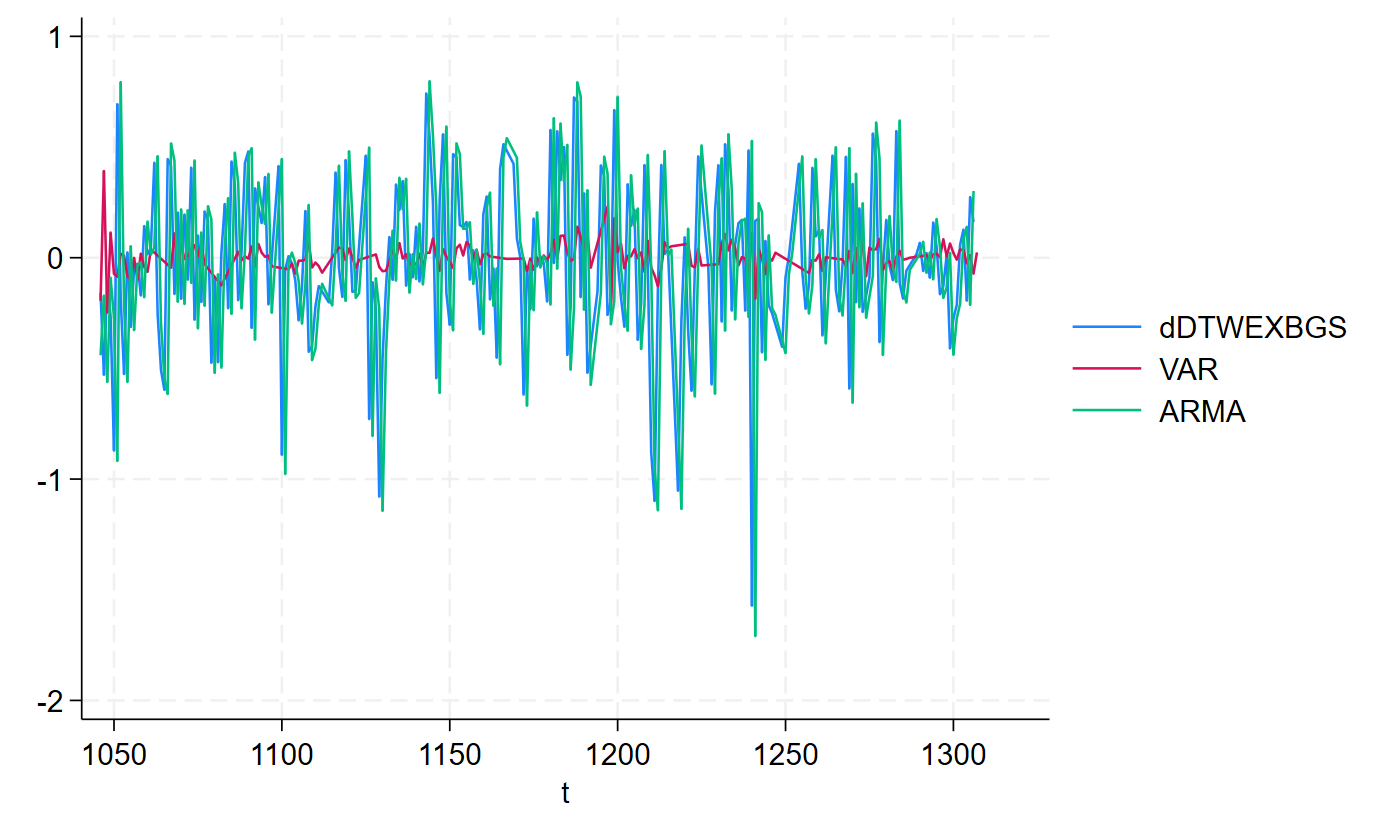
**L2D. | -.0183059 .0368027 -0.50 0.619 -.090553 .0539412**

**L3D. | .0067465 .0366628 0.18 0.854 -.065226 .078719**

**|**

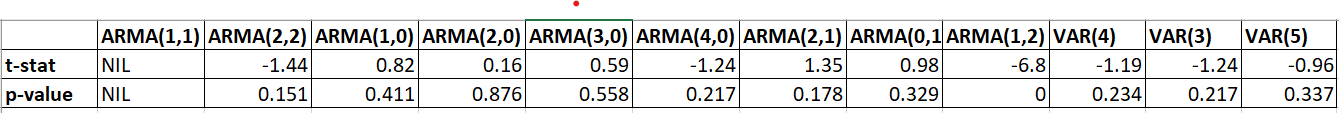
**\_cons | .0060281 .0142489 0.42 0.672 -.0219438 .034**

When we observe the coefficients of all three equations, we see that most of the coefficients are not statistically significant. But I’m satisfied with the fact that all the roots lie within the unit circle which is an indication of stationarity. From the Granger Causality test, it is clear that the Gold ETF and Nominal Broad US Dollar Index are interdependent. Let’s compare the predictability of the univariate and multivariate models.



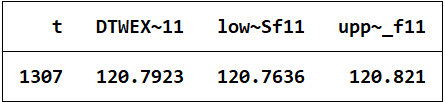
Here, the dDTWEXBGS represents the original differenced series, and the red line and the green line represent the VAR (lag 3) and ARMA (1, 1) respectively. From the graph, it is clear that the VAR model does not have much predictable power. The VAR model is not even close to the original differenced series. Thus, it is good to conclude that the multivariate model is not good at forecasting the Nominal Broad US Dollar Index. The ARMA (1,1) model is clearly showing nearly accurate predictions. Thus, the univariate model proves to outperform the multivariate model. Therefore, we choose ARMA (1,1) as the best model.

We check whether the ARMA (1,1) is statistically better than all other models. I did the Diebold Mariano test to check whether ARMA (1,1) is statistically better than all other models. For that, I take the difference between the absolute forecast errors of ARMA (1,1) and all other models. I plotted the autocorrelation of the differences. We did not see a serial correlation in any of the model comparisons. We obtain the t-statistic and associated p-value from the regression output. We see that only ARMA (1,2) has a statistically significant p-value. This says that ARMA (1,2) may outperform the ARMA (1,1). But when considering the obtained loss function and other statistics, we go ahead with the ARMA (1,1) model.



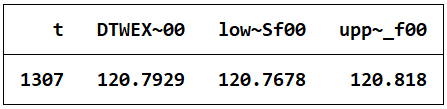
\*\*VAR(3), VAR(5), and VAR(4) represent VAR models with lags 3, 5, and 4 respectively.

I have my data till observation 1306 which is 15th March 2024. When I compute the forecast of observation 1307 using the model ARMA (1,1), I get:



Thus, 19 out of 20 times the Nominal Broad US Dollar Index for 16th March 2024 will be between 120.7636 and 120.821.

Computing the out-of-sample forecast using ARMA (0,0)



We do not see much difference in the forecast interval of ARMA (1,1) and ARMA (0,0). Only a small shrunk of 1% was observed in the 95% probability interval. In other words, compared to ARMA (0,0), the best model’s probability interval is smaller by 1%.

The actual Nominal Broad Exchange rate of 16th March 2024 is 120.9572. The obtained forecast is 120.7923 with a probability interval of 120.7636 and 120.821. The last recorded nominal interest rate is 120.7884. So, we predicted an increase in the exchange rate index and the actual rate index has also increased. According to the loss function:

L(rt, rt+1, rft+1) = D1(rt+1- rt)qt + (1-D1)D2(rt+1- rt)qt

D1 = 1 when (rft+1 > rt) i.e., when I predict the rate will increase

D2 = 1 when (rt+1 > rt) i.e., when the rate for tomorrow increases.

‘rt’ = 120.7884

‘rt+1’ = 120.9572

‘rft+1’ = 120.7923

When we compute the loss, we get the loss as 168.8.

The loss interval can be computed by:

Lower bound: (120.7636 – 120.7884)\*1000 = -24.8

Upper bound: (120.821 – 120.7884)\*1000 = 32.6

Therefore, the lower and upper bounds are -24.8 and 32.6.

The out-of-sample forecast on 15th March 2024 is that the Nominal Broad US Dollar Index would increase the next day. The decision would be to purchase 1000 CHF worth of shares on 15th March 2024 and on 16th March, when the Nominal Broad US Dollar Index has increased, sell the 1000 CHF worth of shares.

**CONCLUSION**

The attempt was to forecast the Nominal Broad US Dollar index to decide whether or not to buy 1000 CHF worth of shares to sell it at an increased exchange rate. I converted the series into stationary by taking the first difference. I also included two more variables for trying multivariable forecasts. The two variables that are highly connected to the Nominal Broad US Dollar Index are the Crude Oil Prices and the Gold ETF Volatility Index. I compared different models for the forecast which included both univariate and multivariate models. ARMA(1,1) was selected as the best model and I computed of sample forecast for 16th March 2024. We predicted that the index would increase. Thus, making the decision of purchasing 1000 CHF worth of shares on 15th of March and selling it on the next day.