ECONOMETRICS - FIN41660

MATLAB ASSIGNMENT

Lecturer name(s): [redacted]

Submission Date: 21/12/2021

Student Name: Ross Kearney

Student Number: [redacted]

A SIGNED COPY OF THIS FORM MUST ACCOMPANY ALL SUBMISSIONS FOR ASSESSMENT.

STUDENTS SHOULD KEEP A COPY OF ALL WORK SUBMITTED.

Procedures for Submission and Late Submission

Ensure that you have checked the School's procedures for the submission of assessments.

Note: There are penalties for the late submission of assessments. For further information please see the University's Policy on Late Submission of Coursework, (http://www.ucd.ie/registrar/)

Plagiarism: the unacknowledged inclusion of another person's writings or ideas or works, in any formally presented work (including essays, examinations, projects, laboratory reports or presentations). The penalties associated with plagiarism designed to impose sanctions that reflect the seriousness of University's commitment to academic integrity. Ensure that you have read the University's Briefing for Students on Academic Integrity and Plagiarism and the UCD Plagiarism Statement, Plagiarism Policy and Procedures, (http://www.ucd.ie/registrar/)

Declaration of Authorship

I declare that all material in this assessment is my own work except where there is clear acknowledgement and appropriate reference to the work of others.

Signed: Ross Kearney

TABULATED SUMMARY OF RESULTS

	BETA COEFFICIENTS	STD ERRORS	T VALUES	P VALUES	CONFIDENCE INTERVALS	
CONSTANT	0.0048963	0.022018	0.22238	0.82406	-0.038301	0.048093
ln GOLD	-0.089617	0.02644	-3.3894	0.000723	-0.14149	-0.037743
ln SP500	0.93052	0.018448	50.44	0	0.89433	0.96672
ln VOLUME	0.0013244	0.00060909	2.1744	0.029872	0.00012939	0.0025194

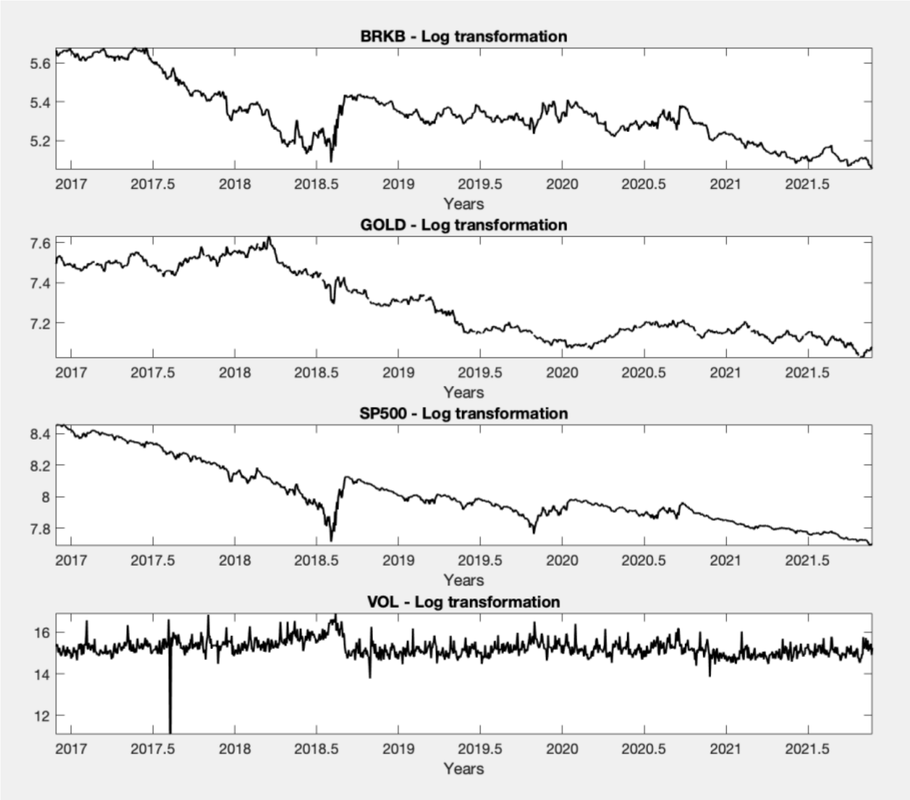
R SQUARED	0.68133		
ADJUSTED R SQUARED	0.68054		
F-STATISTIC	859.49	**F-STAT P-VALUE**	0

JARQUE BERA TEST	3.1206	**JB P-VALUE**	0.19918	**JB CRITICAL VALUE**	5.9423
BREUSCH PAGAN TEST	119.95	**BP P-VALUE**	0	**BP CRITICAL VALUE**	7.8147
RAMSEY RESET F-STAT	17.344	**RESET P-VALUE**	3.732e-08		
T-STATISTIC CRITICAL VALUE	1.9619	**DURBIN WATSON**	0.059825		

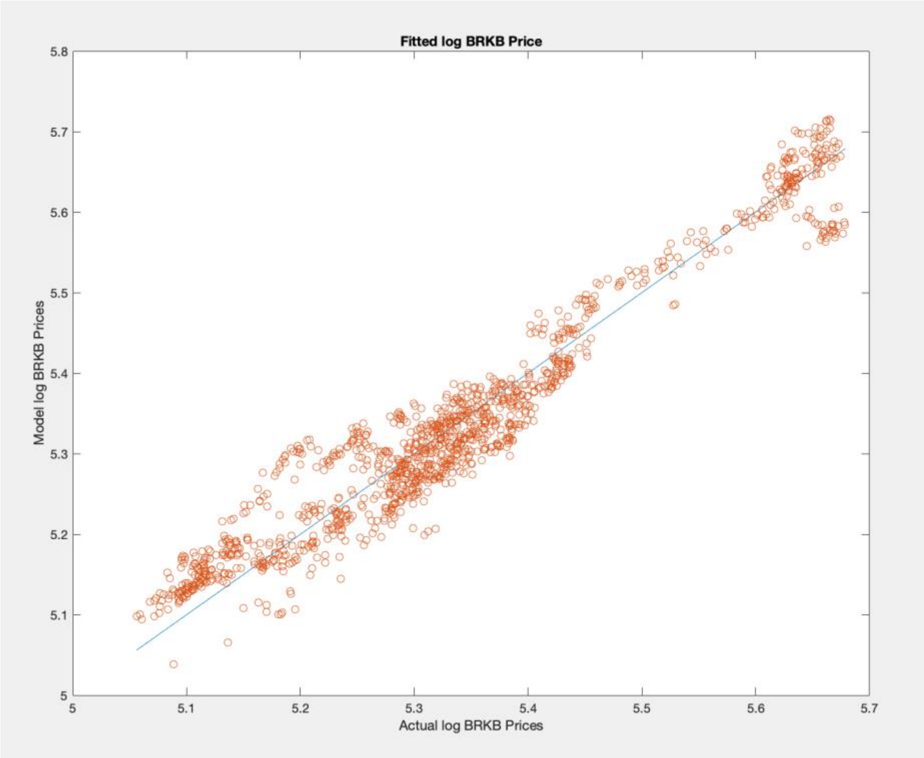
Residual Summary	
0% Quartile	-0.11521
25% Quartile	-0.27232
50% Quartile	0.00069703
75% Quartile	0.027909
100% Quartile	0.11275

Table 1: Summary of Regression Analysis Results

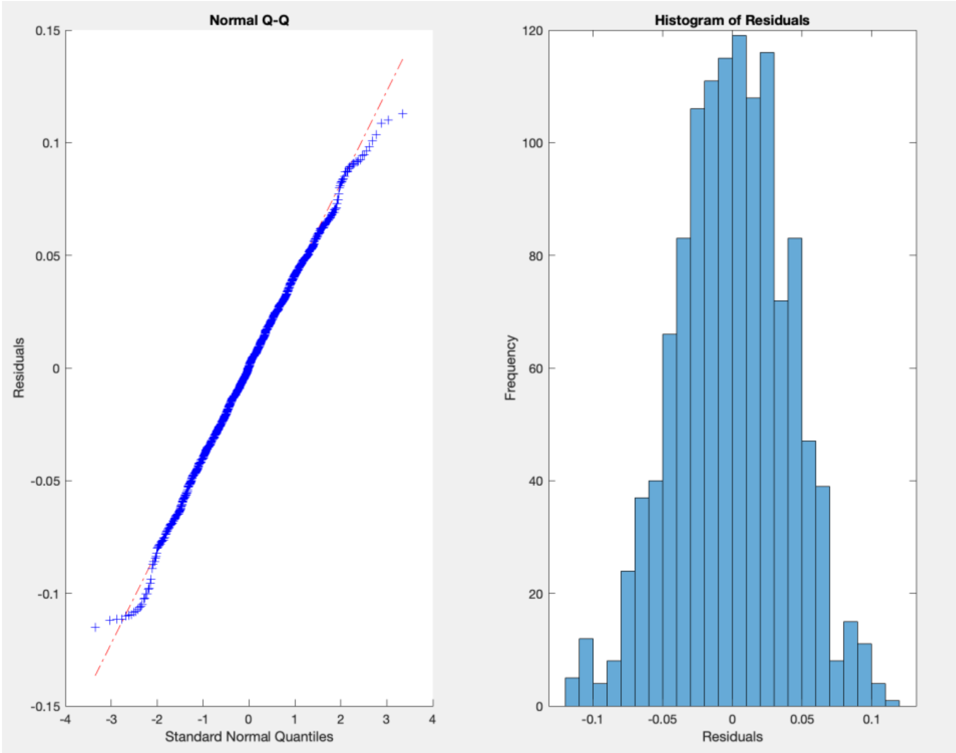
GRAPHS



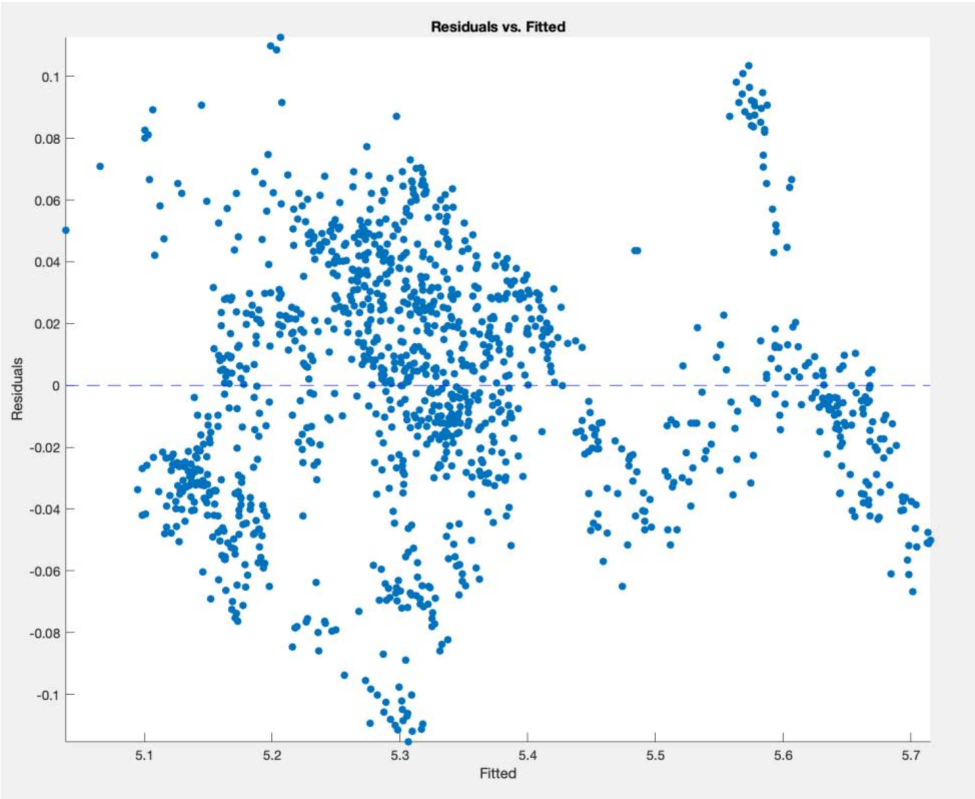
Graph 1: Log Transformation of Data over Sample Time Period



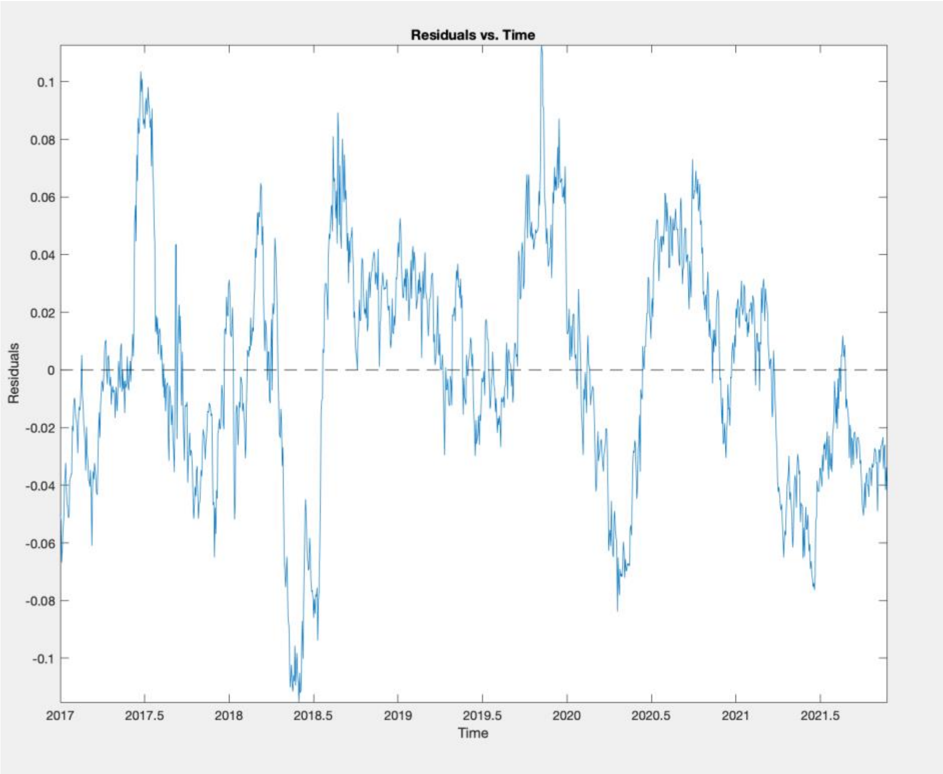
Graph 2: Linear Regression of The Fitted Model



Graph 3: Scatter Plot and Histogram for Normality



Graph 4: Residuals Plotted Against Fitted Values



Graph 5: Plot of Residuals over Sample Time

DISCUSSION

Given all of the above graphs and the tabulated results of the regression analysis, a few observations and comments can be made:

Beta Coefficients

The OLS estimations give beta (β) coefficients for the regression as seen in Table 1. The first beta coefficient suggests that for every 1% increase in the price of gold, the stock price of BRKB (Berkshire Hathaway) *decreases* by ~ 0.09%. While, as may have been expected, the S&P 500 coefficient almost follows a one-to-one relationship with BRKB, as a 1% increase in the S&P 500 would result in a 0.93% increase in BRKB. The volume of the S&P 500 traded each day has a low correlation to the stock price of BRKB, where every 1% increase in the volume traded, the price of BRKB is expected to increase 0.00061%. An almost negligible amount.

P-values

The p-values of these beta coefficients are all statistically significant to a 5% level of confidence, except for the intercept coefficient. This has a significantly high p-level of 0.824. This indicates that the intercept of the regression is not significantly different from 0. When thought about logically, this would seem to make sense. It is essentially saying, if all regressors reduced to zero, i.e. if the price of the S&P 500 and the price of gold, along with the volume of S&P 500 shares traded reduced to zero, the price of BRKB would also reduce to zero. A very reasonable assumption to make.

Confidence Interval
The confidence interval for this regression analysis was set to 95%. The results of the tabulated confidence interval in Table 1 can therefore be determined as saying: with 95% confidence, a 1% rise in the S&P 500 price will cause the price of BRKB to rise at least 0.894% and at most 0.967%, holding all else constant.

Standard Errors
The standard error of the regression or estimate, is the average distance that the observed values deviate from the regression line. That is to say, on average, the price of gold deviates from the regression line by 0.0264%. Obviously a lower standard error translates to a 'tighter' spread of the estimated values around the regression line.

R^2 and Adjusted R^2
The R^2 of the regression analysis is 0.68133, with an Adjusted R^2 of 0.68054. an R^2 value of 0.6813 means that ~68.13% of the variation in the stock price of BRKB is explained by the variation in the three explanatory variables. Clearly a higher R^2 indicates a better regression analysis. An R^2 of 0.68133 is not a bad coefficient of determination, but it is not a great one either. Adjusted R^2 takes into account the fact that an increased number of regressors causes R^2 to increase. Adjusted R^2 adjusts for the degrees of freedom in the model.

F-Statistic
The F-statistic gives the overall significance of the regression. It tests the hypothesis all the slope coefficients are simultaneously equal to zero. In the case of this regression, the p-value of the F-statistic equals zero. This means the null hypothesis can be rejected and it is concluded that at least one regressor is statistically significant.

Graph 2: Fitted Model
The linear trend model in Graph 2 shows the fitted model of the regression analysis. The graph looks as though it confirms our previous R^2 value, the scatter plot is reasonably tight around the trend line of the data, with a strong positive correlation between the regressors and the dependant variable.

Normality
The left hand side plot in Graph 3 shows the residuals of the regressions plotted against the standard normal quantiles. The more the × marks deviate from the red line, the more indicative the graph is of the presence of no-normality in the residuals. That is to say, the more closely the × marks line up with the line, the more normally distributed the residuals of the regression are.

As can be seen from both plots in Graph 2, the residuals look to be normally distributed. This can be confirmed in Table 1, the Jarque-Bera test reveals a value of 3.126 with a p-value of 0.199. The null hypothesis for the Jarque-Bera test is that the data is normally distributed. The above results indicate that we *fail to reject* the null hypothesis just below a 20% confidence level. In other words, the data is in fact normally distributed.

Heteroskedasticity
Graph 4 is a scatter plot of the residuals against the fitted values. This plot is very ‘noisy’ and it is difficult to say directly from the plot whether there is heteroskedasticity present in the data. The Breusch-Pagan test allows a more exact answer to be extracted from the data. As seen in Table 1, the Breusch-Pagan test results in a Breusch-Pagan statistic of 119.95, and a critical value of ~7.815, the result also includes a p-value of zero. These results – the large statistic relative to the critical value – indicate that the null hypothesis of homoskedasticity can be rejected. That is to say, the regression suffers from heteroskedasticity, i.e. the variance of the dependent variable is *not* constant over time.

Serial Correlation
As seen in Table 1, the Durbin-Watson result are 0.0598 which is very close to 2, indicating there is probably evidence of serial correlation.

Collinearity
To test for misspecification, a Ramsey RESET Test is used. A Ramsey RESET Test consists of an F-test and a corresponding p-value. As seen in Table 1, the F-statistic equates to 17.334 and the p-value is ~ 0. This is a highly statistically significant p-value, indicating that we can reject the null hypothesis of the Ramsey RESET Test, which states that the correct specification is linear.

REFERENCES

"plotx1.m" Courtesy of Author: Fabio Parla, Dublin, November 2020

```
%-------------------------------------------------------------------------
% plotx1.m
%-------------------------------------------------------------------------
% This function plots the transformed time series from FRED-MD
%
% Input:
% Y = T x N matrix containing time series
% TCODE = vector of tcode (see Appendix FRED-MD)
% NAMEVAR = vector containing the names of the time series to be plotted
% DATES = vector containing the dates
%
% Output:
% The function plots the figure as programmed
%
% Author: Fabio Parla,
% Dublin, November 2020
%
% Disclaimer: If you find any errors, please email me at
% fabioparla123@gmail.com
%--------------------------------------------------------------------------

function plotx1(Y,TCODE,NAMEVAR,DATES)

n=size(Y,2);  % number of time series to be plotted
labtransform=["no transformation" "Delta (FOD) transformation" "Delta^2 transformation",...
            "Log transformation" "Delta (FOD) log transformation",...
            "Delta^2 Log transformation", "Delta growth rate transformation"];

figure;
for i=1:n  % alternatively cols(DATASELECT)
   subplot(size(Y,2),1,i)
   plot(DATES,Y(:,i),'LineWidth',1.3,'Color','k')
   axis tight
```

```
    hold on;
    xlabel("Years")
    title(strcat(NAMEVAR(i), " - ", labtransform(TCODE(i))))
end

end
```

" getdatatransform.m" Courtesy of Author: Fabio Parla, Dublin, November 2020

```
%-----------------------------------------------------------------------
% getdatatransform.m
%-----------------------------------------------------------------------
% This function compute transformations for series as suggested by
% the Appendix of the FRED-MD database
%
% Input:
% Y = T x N matrix containing time series
% TCODE = vector of tcode (see Appendix FRED-MD)
%
% Output:
% X = T x N matrix of transformed time series
%
% Author: Fabio Parla,
% Dublin, November 2020
%
% Disclaimer: If you find any errors, please email me at
% fabioparla123@gmail.com
%----------------------------------------------------------------------

function X=getdatatransform(Y,TCODE)

% Check if the number of columns in Y is equal to the number of elements
% in TCODE. This function has been prepared to work only in case of number
% of columns in Y equal to the number of elements in TCODE.
% If the numbers are the different, the function prints an error message

[T,N]=size(Y);

if N~=size(TCODE,2)
  error('Error! The lenght of tcode must be equal to the number of time series.')
end

% Apply transformation to the time series
X=NaN(T,N);   % Create matrix of NaN where we store the transformed time
```

```
              % series
for j=1:N
    x=Y(:,j); % select the time series

    if TCODE(j)==1 % no transformation
       X(:,j)=x;
    elseif TCODE(j)==2 % FOD transformation
       X(2:end,j)=diff(x,1);
    elseif TCODE(j)==3 % Second order diff. transformation
       X(3:end,j)=diff(x,2);
    elseif TCODE(j)==4 % Log transformation
       X(:,j)=log(x);
    elseif TCODE(j)==5 % Delta (FOD) log transformation
       X(2:end,j)=diff(log(x),1);
    elseif TCODE(j)==6 % Delta^2 (SOD) log transformation
       X(3:end,j)=diff(log(x),2);
    elseif TCODE(j)==7 % Delta (FOD) of the growth rate transformation
       X(3:end,j)=diff((x(2:end,:)./x(1:end-1))-1,1);
    end
end
end
```

" removenan.m" Courtesy of Author: Fabio Parla, Dublin, November 2020

```
%----------------------------------------------------------------------
% removenan.m
%----------------------------------------------------------------------
% This function remove NaNs from a matrix of observations
%
% Input:
% X = T x N matrix containing time series
%
% Output:
% out = T x (N-NaNS) matrix containing time series
%
% Author: Fabio Parla,
```

```
% Dublin, November 2020
%
% Disclaimer: If you find any errors, please email me at
% fabioparla123@gmail.com
%------------------------------------------------------------------------
function out=removenan(X)

id=sum(isnan(X),2)==0;
out=X(id,:);

end
```