HW#3

Name: **Jingyi Wu**

Due Date: Feb 1, 2019

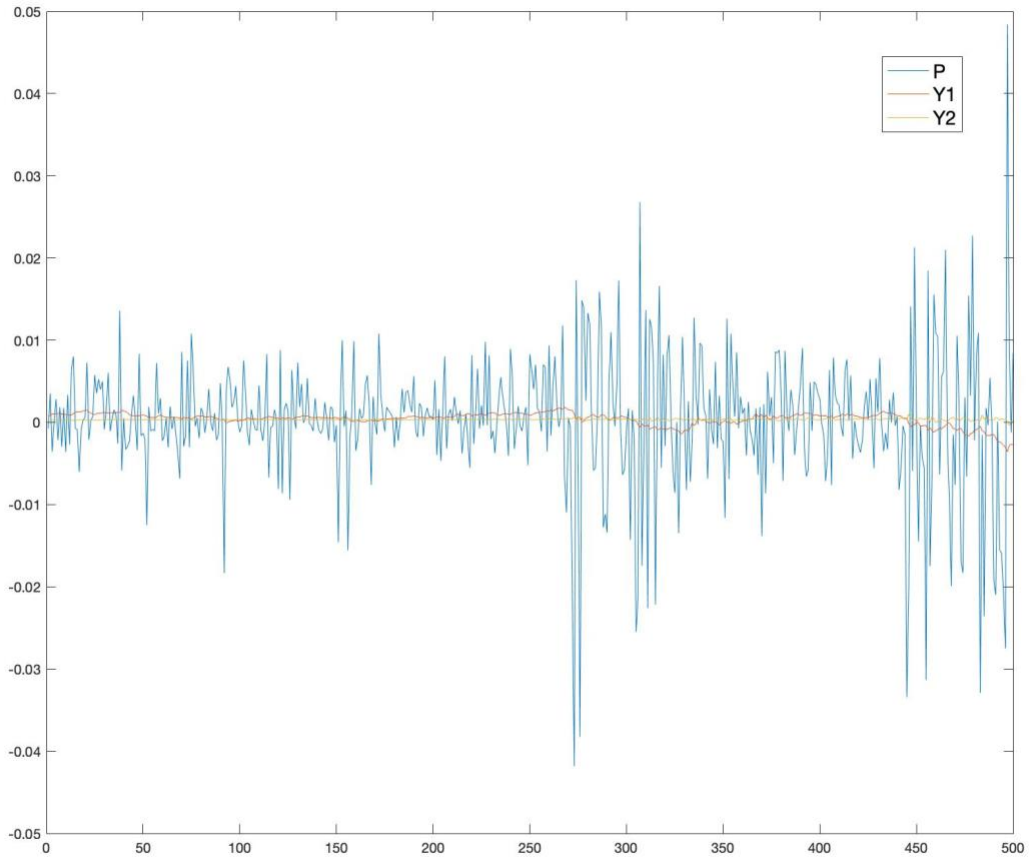
Course Code: ECON 623 Forecasting Financial Markets

Instructor: Professor Andrew Patton

Exercise 1

Data: S&P 500 from %S&P500 from 01/01/2014-12/31/2018
Finance.yahoo.com

(b)



(c)

	Y1_Forecast	Y2_Forecast
b0	"0.00018237"	"0.0035192"
b1	"0.03724"	"-9.7304"
se0	"0.00059599"	"0.0015973"
se1	"0.84205"	"4.8947"
chi2-critical value	"5.9915"	"5.9915"
chi2-statistics	"2.7674"	"4.8803"
conclusion	"fail to reject"	"fail to reject"

Because in both cases, we fail to reject the null, so both Forecast 1 and Forecast 2 are not optimal.

(d)

mse1	mse2	DM	CriticalValue
6.754e-05	6.718e-05	0.33542	1.96

DM test statistics is less than critical value. So we can't reject the null. Therefore, there is not evidence to show that forecast 1 is better than 2 or vice versa.

Code:

```
%% Exercise 1
SP = csvread('^GSPC.csv', 1, 0); %S&P500 from 01/01/2014-12/31/2018 on
Finance.yahoo.com
P_SP = SP(:, 6);
ret_SP = log(P_SP(2:end)./P_SP(1:end-1)); %calculate log return

T = length(ret_SP);
R = ret_SP(1:T-500);
P = ret_SP(T-500+1:T);

%(a)
Y1 = zeros(500, 1); %create a default column for Y1
for i = 1:500
    Y1(i) = mean(ret_SP(T-500+i-60:T-500+i-1)); %calculate forecast Y1
end

%(b)
Y2 = zeros(500, 1); %create a default column for Y1
for i = 1:500
    Intercept = ones(size(ret_SP(1:T-500+i-1-1), 1), 1); %create intercept
    X = [Intercept ret_SP(1:T-500+i-1-1)]; %create independent variable
    Y = ret_SP(2:T-500+i-1); %create dependent variable
    results = ols(Y,X);
    Beta = results.beta'; %extract estimated parameters
    Ylag = [1 ret_SP(T-500+i-1)]';
    Y2(i) = Beta*Ylag; %calculate predict value
end

plot(P)
hold on
plot(Y1)
hold on
plot(Y2)
hold off

%(c)
% MZ for Y1
Intercept = ones(size(Y1, 1), 1);
X = [Intercept Y1];
Y = P;
results = nwest(Y,X);
beta1 = results.beta;
se1 = results.se;
vcv1 = results.vcv;
```

```
R1 = [1 0
      0 1];
%Calculate test statistics
TestValue1 = (R1*beta1 - [0; 1])'*((R1'*vcv1*R1)\(R1*beta1 - [0; 1]))
%Calculate Critical Value for 5% significance level
CriticalValue1 = chi2inv(1-0.05,2)
%Test if test statistics is less than the Critical Value
TestValue1 < CriticalValue1
Y1_Forecast = [beta1; se1; CriticalValue1; TestValue1; "fail to reject"];

% MZ for Y2
Intercept = ones(size(Y2, 1), 1);
X = [Intercept Y2];
Y = P;
results = nwest(Y,X);
beta2 = results.beta;
se2 = results.se;
vcv2 = results.vcv;
R2 = [1 0
      0 1];
%Calculate test statistics
TestValue2 = (R2*beta2 - [0; 1])'*((R2'*vcv2*R2)\(R2*beta2 - [0; 1]))
%Calculate Critical Value for 5% significance level
CriticalValue2 = chi2inv(1-0.05,2)
%Test if test statistics is less than the Critical Value
TestValue2 < CriticalValue2
Y2_Forecast = [beta2; se2; CriticalValue2; TestValue2; "fail to reject"];

Summary = table(Y1_Forecast,Y2_Forecast);
Summary.Properties.RowNames = {'b0', 'b1', 'se0', 'se1', 'chi2-critical
value', 'chi2-statistics', 'conclusion'};
Summary

%(d)
e1 = P - Y1; %error for Y1
squarederror1 = e1.^2; %squared error
mse1 = mean(squarederror1); %MSE for Y1
e2 = P - Y2;
squarederror2 = e2.^2;
mse2 = mean(squarederror2);
d = squarederror1 - squarederror2;
dbar = mean(d);
dvar = var(d);
dbarvar = dvar/500;
DM = dbar/sqrt(dbarvar);
CriticalValue = 1.96;
Summary2 = table(mse1, mse2, DM, CriticalValue)
```

Exercise 2

(a)

	Beta	Tstatistics
Base case	−0.026276	−2.6062

(b)

	Beta	Tstatistics
Base case	−0.026276	−2.6062
Dropping Oct 19, 1987	−0.035981	−3.5696
Dropping Oct, 1987	−0.025285	−2.505
Dropping 1987	−0.0329	−3.2215
Dropping 1980–1989	−0.052854	−4.5229

In all case, there is significant first-order autocorrelation. But it seems that by dropping more, the significance goes up. So, this outlier (Oct 19, 1987) may not contribute that much to the autocorrelation.

	Beta	Tstatistics
Base case	−0.026276	−2.6062
Dropping May 6, 2010	−0.027769	−2.7542
Dropping May, 2010	−0.025476	−2.5242
Dropping 2010	−0.02654	−2.5984
Dropping 2010–2018	−0.022714	−1.9763

In all case, there is significant first-order autocorrelation. And it seems that by dropping more, the significance goes down. So, this outlier (May 6, 2010) may contribute some extent to the autocorrelation.

(c)

	Beta	Tstatistics
Base case	−0.026276	−2.6062
1980–1989	0.051327	2.582
1990–1999	0.017518	0.88023
2000–2009	−0.087278	−4.3968
2010–2018	−0.043299	−2.0595

There is positive autocorrelation in subsample 1 and 2 (1980-1989, 1990-1999), and negative autocorrelation in subsample 3 and 4 (2000-2009, 2010-2018). The base case seems to be an average of these four subsamples. The second subsample (1990-1989) does not have a significant autocorrelation, so it may not contribute too much to the overall autocorrelation.

(d)

	Beta	Tstatistics
Base case (OLS sd)	-0.026276	-2.6062
Base case (Newey-West sd)	-0.026276	-1.8195

When Newey-West standard error is applied, there is no evidence to show that there is significant autocorrelation. So, this model is kind of sensitive to the method we use to calculate standard error.

Code:

```
%% Exercise 2
%(a)
SP2 = csvread('^GSPC-2.csv', 1, 0); %S&P500 from 01/01/1980-12/31/2009 on
Finance.yahoo.com
P_SP2 = SP2(:, 6);
ret_SP2 = log(P_SP2(2:end)./P_SP2(1:end-1)); %calculate log return

Intercept = ones(size(ret_SP2(1:end-1), 1), 1);
X = [Intercept ret_SP2(1:end-1)];
Y = ret_SP2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics = results.tstat(2);
Original = table(Beta, Tstatistics);
Original.Properties.RowNames = {'Base case'};
Original

%(b)
%investigate Oct 19, 1987
date1 = SP2(:, 6);
date1(1972) = [];
ret_date1 = log(date1(2:end)./date1(1:end-1)); %calculate log return
Intercept = ones(size(ret_date1(1:end-1), 1), 1);
X = [Intercept ret_date1(1:end-1)];
Y = ret_date1(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Date1 = table(Beta, Tstatistics);
Date1.Properties.RowNames = {'Dropping Oct 19, 1987'};
Date1;

month1 = SP2(:, 6);
month1(1960:1981) = [];
```

```
ret_month1 = log(month1(2:end)./month1(1:end-1)); %calculate log return
Intercept = ones(size(ret_month1(1:end-1), 1), 1);
X = [Intercept ret_month1(1:end-1)];
Y = ret_month1(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Month1 = table(Beta, Tstatistics);
Month1.Properties.RowNames = {'Dropping Oct, 1987'};
Month1;

year1 = SP2(:, 6);
year1(1771:2023) = [];
ret_year1 = log(year1(2:end)./year1(1:end-1)); %calculate log return
Intercept = ones(size(ret_year1(1:end-1), 1), 1);
X = [Intercept ret_year1(1:end-1)];
Y = ret_year1(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Year1 = table(Beta, Tstatistics);
Year1.Properties.RowNames = {'Dropping 1987'};
Year1;

decade1 = SP2(:, 6);
decade1(1:2528) = [];
ret_decade1 = log(decade1(2:end)./decade1(1:end-1)); %calculate log return
Intercept = ones(size(ret_decade1(1:end-1), 1), 1);
X = [Intercept ret_decade1(1:end-1)];
Y = ret_decade1(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Decade1 = table(Beta, Tstatistics);
Decade1.Properties.RowNames = {'Dropping 1980-1989'};
Decade1;

outlier1 = [Original; Date1; Month1; Year1; Decade1]

%investigate May 6, 2010
date2 = SP2(:, 6);
date2(7657) = [];
ret_date2 = log(date2(2:end)./date2(1:end-1)); %calculate log return
Intercept = ones(size(ret_date2(1:end-1), 1), 1);
X = [Intercept ret_date2(1:end-1)];
Y = ret_date2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Date2 = table(Beta, Tstatistics);
Date2.Properties.RowNames = {'Dropping May 6, 2010'};
Date2;

month2 = SP2(:, 6);
month2(7654:7673) = [];
ret_month2 = log(month2(2:end)./month2(1:end-1)); %calculate log return
Intercept = ones(size(ret_month2(1:end-1), 1), 1);
X = [Intercept ret_month2(1:end-1)];
```

```
Y = ret_month2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Month2 = table(Beta, Tstatistics);
Month2.Properties.RowNames = {'Dropping May, 2010'};
Month2;

year2 = SP2(:, 6);
year2(7572:7823) = [];
ret_year2 = log(year2(2:end)./year2(1:end-1)); %calculate log return
Intercept = ones(size(ret_year2(1:end-1), 1), 1);
X = [Intercept ret_year2(1:end-1)];
Y = ret_year2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Year2 = table(Beta, Tstatistics);
Year2.Properties.RowNames = {'Dropping 2010'};
Year2;

decade2 = SP2(:, 6);
decade2(7572:9835) = [];
ret_decade2 = log(decade2(2:end)./decade2(1:end-1)); %calculate log return
Intercept = ones(size(ret_decade2(1:end-1), 1), 1);
X = [Intercept ret_decade2(1:end-1)];
Y = ret_decade2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Decade2 = table(Beta, Tstatistics);
Decade2.Properties.RowNames = {'Dropping 2010-2018'};
Decade2;

outlier2 = [Original; Date2; Month2; Year2; Decade2]

%(c)
sub1 = SP2(1:2528, 6);
ret_sub1 = log(sub1(2:end)./sub1(1:end-1)); %calculate log return
Intercept = ones(size(ret_sub1(1:end-1), 1), 1);
X = [Intercept ret_sub1(1:end-1)];
Y = ret_sub1(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Sub1 = table(Beta, Tstatistics);
Sub1.Properties.RowNames = {'1980-1989'};
Sub1;

sub2 = SP2(2529:5056, 6);
ret_sub2 = log(sub2(2:end)./sub2(1:end-1)); %calculate log return
Intercept = ones(size(ret_sub2(1:end-1), 1), 1);
X = [Intercept ret_sub2(1:end-1)];
Y = ret_sub2(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Sub2 = table(Beta, Tstatistics);
```

```
Sub2.Properties.RowNames = {'1990-1999'};
Sub2;

sub3 = SP2(5057:7572, 6);
ret_sub3 = log(sub3(2:end)./sub3(1:end-1)); %calculate log return
Intercept = ones(size(ret_sub3(1:end-1), 1), 1);
X = [Intercept ret_sub3(1:end-1)];
Y = ret_sub3(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Sub3 = table(Beta, Tstatistics);
Sub3.Properties.RowNames = {'2000-2009'};
Sub3;

sub4 = SP2(7573:9835, 6);
ret_sub4 = log(sub4(2:end)./sub4(1:end-1)); %calculate log return
Intercept = ones(size(ret_sub4(1:end-1), 1), 1);
X = [Intercept ret_sub4(1:end-1)];
Y = ret_sub4(2:end);
results = ols(Y,X);
Beta = results.beta(2);
Tstatistics= results.tstat(2);
Sub4 = table(Beta, Tstatistics);
Sub4.Properties.RowNames = {'2010-2018'};
Sub4;

Subsample = [Original; Sub1; Sub2; Sub3; Sub4]

%(d)
Intercept = ones(size(ret_SP2(1:end-1), 1), 1);
X = [Intercept ret_SP2(1:end-1)];
Y = ret_SP2(2:end);
results = nwest(Y,X);
Beta = results.beta(2);
Tstatistics = results.tstat(2);
NW = table(Beta, Tstatistics);
NW.Properties.RowNames = {'Base case (Newey-West sd)'};
Original.Properties.RowNames = {'Base case (OLS sd)'};
SD = [Original; NW]
```