Introductory Econometrics
Tutorial 11 Solutions

<u>PART A:</u> To be done before you attend the tutorial. The solutions will be made available at the end of the week.

1. **Answer**: Agree. Mostly. Seasonality usually refers to inter-annual variations, but really, the phenomenon is simply one of repeated deterministic patterns, which could cover any period of time. So since the most obvious deterministic periodic patterns that affect our behaviour are seasons or months or days of the week, the best answer is 'agree', seasonality is not an issue when using annual time series observations. However, technically deterministic patterns could also exist that occur deterministically every x-years, for example leap years happen deterministically once every four years. If you answered 'disagree' and gave a compelling argument with a good example, your answer is correct also.

2. **Answer**:

   (a) A covariance stationary process $\{y_t,\ t = 1, 2, ...\}$ is a sequence of random variables with finite mean, variance that do not depend on t and covariances that only depend on time distance between the variables and not on time itself. That is

$$\begin{aligned} E(y_t) &= \mu \text{ for all } t \\ Var(y_t) &= \sigma^2 \text{ for all } t, \text{ and} \\ Cov(y_t, y_{t+j}) &= \gamma_j \text{ for all } t \text{ and all } j. \end{aligned}$$

   The correlogram of a time series sample from this process will have autocorrelations that decay to zero exponentially.

   (b) A white noise process is a covariance stationary process that is uncorrelated over time, i.e. $\gamma_j = 0$ for $j = 1, 2, ...$ The correlogram of a time series sample from a white noise process will have autocorrelations that are all insignificant.

   (c) A mean reverting process is commonly used as another name for a covariance stationary process (although technically they need not be the same, but we do not get into those technicalities here).

   (d) A trend stationary process is a process whose mean depends linearly on time, and after removing this time trend, the remainder is covariance stationary. i.e. $\{y_t,\ t = 1, 2, ...\}$ is trend stationary if for some β $\{y_t - \beta t,\ t = 1, 2, ...\}$ is covariance stationary. The correlogram of a time series sample from a trend stationary process shows very high persistence, but after removing its trend by a regression on time, the correlogram of the residuals will look like the correlogram of a time series sample from a stationary process.

   (e) A random walk is an AR(1) process with intercept equal to zero and variance equal to 1, i.e.

$$y_t = y_{t-1} + e_t$$

   where e_t is white noise. The correlogram of a time series sample from a random walk will show first order autocorrelation very close to 1 and autocorrelations that decay very slowly.

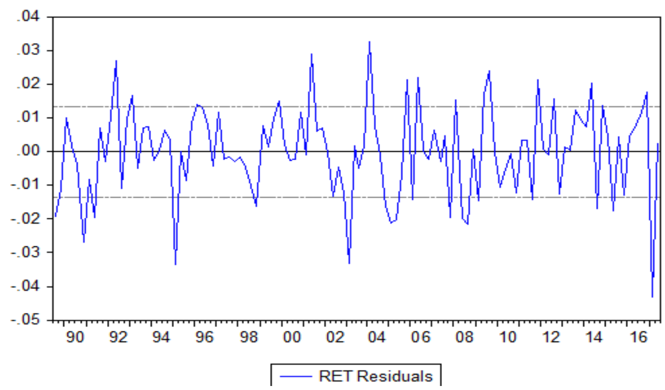
3. **Answer**:Eviews output:

Dependent Variable: RET
Method: Least Squares
Date: 10/01/17 Time: 19:31
Sample (adjusted): 1989Q3 2017Q2
Included observations: 112 after adjustments

Variable	Coefficien...	Std. Error	t-Statistic	Prob.
C	-0.000100	0.001278	-0.078434	0.9376
RET(-1)	-0.190765	0.093856	-2.032533	0.0445

R-squared	0.036197	Mean dependent var	-6.10E-05
Adjusted R-squared	0.027435	S.D. dependent var	0.013716
S.E. of regression	0.013526	Akaike info criterion	-5.750697
Sum squared resid	0.020125	Schwarz criterion	-5.702152
Log likelihood	324.0390	Hannan-Quinn criter.	-5.731001
F-statistic	4.131188	Durbin-Watson stat	1.962740
Prob(F-statistic)	0.044509		

The plot or residual is given below:



It is (sort of) evident that there is a pattern in the graph of the residuals and that they are not an i.i.d. sequence. Then, from the estimation output table, select: View–> Residual Diagnostics –> Serial Correlation LM Test –> 4 lags. The Eviews outcome for the Breusch-Godfrey test is given below:

Breusch-Godfrey Serial Correlation LM Test:

F-statistic	5.091057	Prob. F(4,106)	0.0009
Obs*R-squared	18.04936	Prob. Chi-Square(4)	0.0012

It is clear from both statistics that the null of no serial correlation in the residuals is rejected. This supports our observations from the plot of the residuals. You might want to investigate running model with larger lags and find out the best AR model for this time series.

Do not forget to bring your answers to PART A and a copy of the tutorial questions to your tutorial.

<u>Part B:</u> This part will be covered in the tutorial. It is still a good idea to attempt these questions before the tutorial.

1. Let
$$y_t = c + \varphi_1 y_{t-1} + u_t, \text{ with } |\varphi_1| < 1 \text{ and} \tag{1}$$
$$u_t = \rho u_{t-1} + e_t, \text{ with } |\rho| < 1 \text{ and} \tag{2}$$
$$e_t \sim i.i.d(0, \sigma^2).$$

Answer:

(a) Substitute for u_t in (1) from (2) to obtain:
$$y_t = c + \varphi_1 y_{t-1} + \rho u_{t-1} + e_t, \tag{3}$$
and note that by lagging (1) we see that
$$u_{t-1} = y_{t-1} - c - \varphi_1 y_{t-2}. \tag{4}$$
So, we use (4) to substitute for u_{t-1} in (3) and simplify:
$$\begin{aligned} y_t &= c + \varphi_1 y_{t-1} + \rho\left(y_{t-1} - c - \varphi_1 y_{t-2}\right) + e_t \\ &= (1-\rho)\, c + (\varphi_1 + \rho)\, y_{t-1} - \rho\varphi_1 y_{t-2} + e_t. \end{aligned}$$
or
$$y_t = \alpha_0 + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + e_t,$$
where
$$\alpha_0 = (1-\rho)c, \ \alpha_1 = (\varphi_1 + \rho), \ \alpha_2 = -\rho\varphi_1.$$
as required.

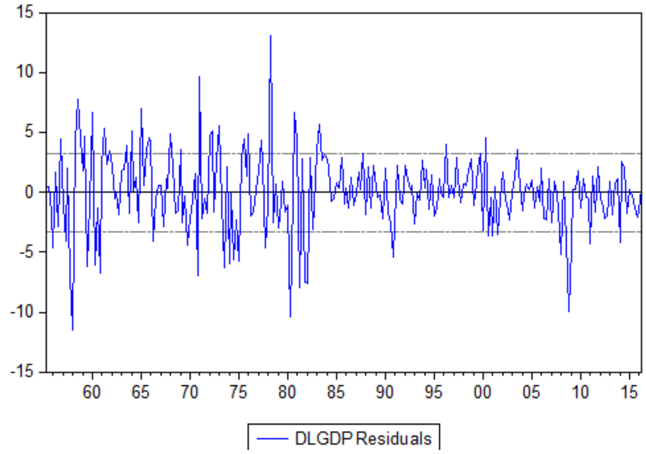
(b) In an AR model, if we find evidence of serial correlation in errors, it means that we need a higher order AR model.

2. **Answer**: The graph of the residuals is given below (the scale may be different depending on whether you used lag differences only or you multiplied them by 100 (to convert them to percentage points) or by 400 (to convert them to annual rates).

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.742741	0.300922	5.791343	0.0000
DLGDP(-1)	0.306892	0.063042	4.868049	0.0000
DLGDP(-2)	0.108769	0.063053	1.725045	0.0858

R-squared	0.130079	Mean dependent var	2.999617
Adjusted R-squared	0.123035	S.D. dependent var	3.499670
S.E. of regression	3.277315	Akaike info criterion	5.223854
Sum squared resid	2652.976	Schwarz criterion	5.266111



We can detect a slight pattern in the evolution of residuals but it does not necessarily persist over time. It is not clear from the visual inspection if the residuals are serially correlated. The correlogram gives us better information, and shows no evidence of serial correlation in errors.

Sample: 1955Q1 2017Q2
Included observations: 250
Q-statistic probabilities adjusted for 2 dynamic regressors

Autocorrelation	Partial Correlation		AC	PAC
		1	-0.008	-0.008
		2	0.006	0.006
		3	-0.009	-0.009
		4	0.008	0.008
		5	-0.094	-0.094
		6	0.033	0.032
		7	-0.012	-0.011
		8	-0.094	-0.097
		9	0.041	0.043
		10	0.083	0.076
		11	0.028	0.032
		12	-0.072	-0.076

A formal test for no serial correlation in errors against an AR(8) alternative starts with specifying an AR equation for errors:

$$
\begin{aligned}
dlgdp_t &= \beta_0 + \beta_1 dldgp_{t-1} + \beta_2 dlgdp_{t-2} + u_t \\
u_t &= \rho_1 u_{t-1} + \rho_2 u_{t-2} + \cdots + \rho_8 u_{t-8} + e_t \\
H_0 &: \rho_j = 0 \text{ for } j = 1, 2, \ldots, 8 \\
H_1 &: \text{at least one } \rho \text{ is not zero} \\
BG &= (n-8)\, R^2_{\hat{u}} \sim \chi^2_8 \text{ under } H_0
\end{aligned}
$$

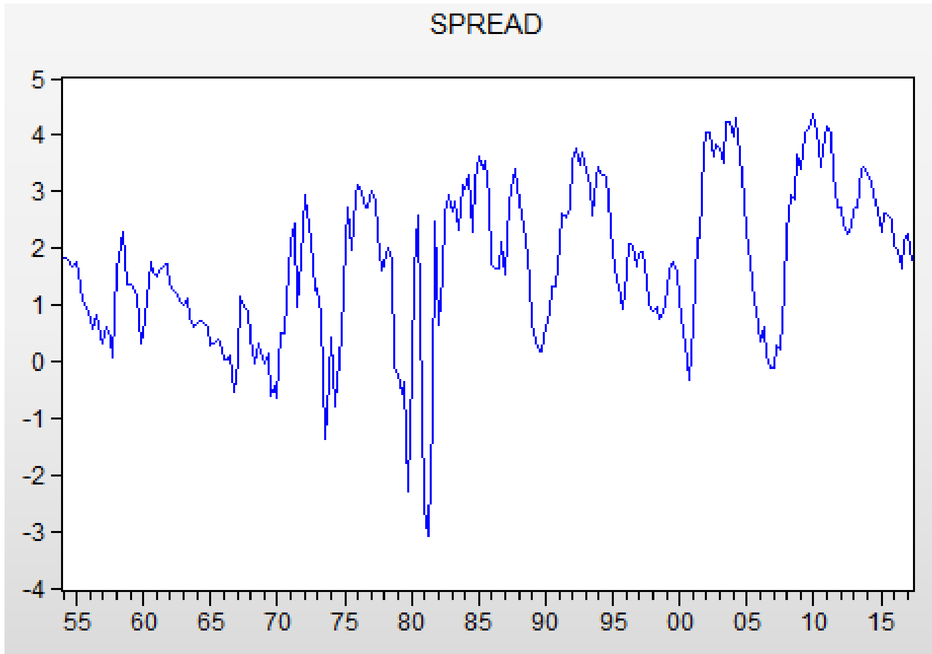
where $R^2_{\hat{u}}$ is the R^2 of the regression of residuals on a constant, $dldgp_{t-1}$, and $dlgdp_{t-2}$ and 8 lags of residuals. From the Eviews output we get

$$BG_{calc} = 8.972$$

with p-value of 0.345, which is much larger than 0.05. This means that we cannot reject the null at the 5% level of significance. Hence the AR(2) model seems adequate. We can then proceed to use t-test to see if both lags are needed, or an AR(1) would be sufficient.

3. **Answer**:

(a) No, *spread* is not white noise. Its plot shows long swings. And its correlogram shows significant autocorrelation. It does return to its mean regularly, although it has noticeable persistence. The correlogram shows that autocorrelations decay exponentially, so it looks like a covariance stationary process.



Sample: 1954Q1 2017Q2
Included observations: 254

Autocorrelation	Partial Correlation		AC	PAC
		1	0.875	0.875
		2	0.723	-0.183
		3	0.606	0.074
		4	0.522	0.044
		5	0.419	-0.160
		6	0.316	-0.013
		7	0.236	0.008
		8	0.184	0.025
		9	0.117	-0.119
		10	0.036	-0.074
		11	-0.028	0.006
		12	-0.059	0.026

(b) The unrestricted model:

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.103416	0.394741	2.795293	0.0056
DLGDP(-1)	0.277960	0.063335	4.388718	0.0000
DLGDP(-2)	0.107821	0.062999	1.711453	0.0883
SPREAD(-1)	-0.122029	0.315459	-0.386831	0.6992
SPREAD(-2)	0.547081	0.317431	1.723466	0.0861

R-squared	0.160183	Mean dependent var	2.999617
Adjusted R-squared	0.146472	S.D. dependent var	3.499670
S.E. of regression	3.233226	Akaike info criterion	5.204635
Sum squared resid	2561.168	Schwarz criterion	5.275064

$$
\begin{aligned}
H_0 &: \beta_3 = \beta_4 = 0 \\
H_1 &: \text{ at least one of the above is not zero} \\
F &= \frac{(SSR_r - SSR_{ur})/2}{SSR_{ur}/(250-5)} \sim F_{2,245} \text{ under } H_0 \\
F_{calc} &= \frac{(2652.976 - 2561.168)/2}{2561.168/245} = 4.391 \\
F_{crit} &= 3.03 \text{ (from Eviews)} \approx 3.07 \text{ from the table for } F_{2,120} \\
F_{calc} &> F_{crit} \Rightarrow \text{ we reject the null.} \\
& \text{Conclusion is that at least one of the spread lags is significant}
\end{aligned}
$$

(c) The first lag of spread is insignificant judging by its t-statistic. We drop that and re-estimate:

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.058540	0.376656	2.810361	0.0053
DLGDP(-1)	0.281108	0.062701	4.483279	0.0000
DLGDP(-2)	0.111661	0.062105	1.797950	0.0734
SPREAD(-2)	0.438723	0.149062	2.943227	0.0036

R-squared	0.159670	Mean dependent var	2.999617
Adjusted R-squared	0.149422	S.D. dependent var	3.499670
S.E. of regression	3.227633	Akaike info criterion	5.197246
Sum squared resid	2562.733	Schwarz criterion	5.253589

According to these estimates, a one percentage point decrease in the spread does not change the growth rate immediately. It takes two quarters before this will start to affect the GDP growth, when it is expected to decrease by 0.44 percentage points. In the long-run the GDP growth will decline by

$$\frac{0.438723}{(1 - 0.281108 - 0.111661)} = 0.723 \text{ percentage points.}$$

Note that I have generated growth rate using $400 * \triangle \log(GDP)$ to get the annualised rate. If you multiply by 100 only point estimates will be different but the test results should all be the same as here. Some may notice that the second lag of GDP growth is not significant at the 5% level either. Dropping that does not cause any serial correlations in the errors either. With that decision, we get

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.286932	0.356188	3.613068	0.0004
DLGDP(-1)	0.320048	0.059107	5.414679	0.0000
SPREAD(-2)	0.434482	0.149715	2.902056	0.0040

R-squared	0.148628	Mean dependent var	2.999617
Adjusted R-squared	0.141734	S.D. dependent var	3.499670
S.E. of regression	3.242187	Akaike info criterion	5.202301
Sum squared resid	2596.409	Schwarz criterion	5.244558

which tells us that it takes two periods before a decline in the spread affect GDP growth by decreasing it by 0.43 percentage points initially and by

$$\frac{0.434482}{1 - 0.320048} = 0.639$$

in the long-run.

(d) We generate a dummy variable called $pre86 = @year < 1986$. This uses the EViews function @year that extracts the year component of the date for each observation. This dummy variable is zero for all observations before 1986 and is 1 for all observation from 1986 onward. Since the hypothesis is only about the effect of *spread* on GDP growth, we interact this dummy with *spread* only (they may want to consider more elaborate test of structural break, that is OK if they want to explore, but not necessary for answering this specific question). Using this with either ARDL(2,2) or ARDL(1,2) confirms that the leading indicator power of *spread* has deteriorated after 1986.

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.211209	0.362702	3.339408	0.0010
DLGDP(-1)	0.203281	0.062346	3.260521	0.0013
DLGDP(-2)	0.078950	0.059966	1.316575	0.1892
PRE86*SPREAD(-2)	1.213964	0.217398	5.584054	0.0000
(1-PRE86)*SPREAD(-2)	0.228862	0.149687	1.528941	0.1276

R-squared	0.230086	Mean dependent var	2.999617
Adjusted R-squared	0.217516	S.D. dependent var	3.499670
S.E. of regression	3.095743	Akaike info criterion	5.117730
Sum squared resid	2347.988	Schwarz criterion	5.188159

Dependent Variable: DLGDP
Method: Least Squares
Sample: 1955Q1 2017Q2
Included observations: 250

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	1.375443	0.341083	4.032578	0.0001
DLGDP(-1)	0.227953	0.059552	3.827831	0.0002
PRE86*SPREAD(-2)	1.235853	0.217085	5.692954	0.0000
(1-PRE86)*SPREAD(-2)	0.219178	0.149729	1.463832	0.1445

R-squared	0.224639	Mean dependent var	2.999617
Adjusted R-squared	0.215183	S.D. dependent var	3.499670
S.E. of regression	3.100354	Akaike info criterion	5.116780
Sum squared resid	2364.600	Schwarz criterion	5.173124