Econ 3342
Assignment #5

Instructions. Please type your answer key before turning it in. Don't forget to include your name and to include every graph/output that is requested in this activity.

Go to https://fred.stlouisfed.org/ and search for Real Personal Consumption Expenditures (series *pcecca*). This series is available since 1929. Download the data in annual frequency as an EXCEL file. Import your files into Eviews. File should show the data to be available from 1929 to 2019.

a. Let y_t be the variable *pcecca* in period t. Create the variable gc (annual consumption growth rate) with the formula $gc = 100 * \frac{y_t - y_{t-1}}{y_{t-1}}$. When creating the variable, make sure that you have not made a mistake in the formula. Show the time series for gc for the period 1929 - 2019 (paste the graph in your homework).

b. For the period 1980 - 2017 (NOTICE THE PERIOD) create the correlogram for the variable gc, use up to 12 lags.

c. The requested correlogram of the consumption growth (variable gc) could be consistent with processes in the AR family, as well as in the MA family (or ARMA processes). For the period 1980-2017, estimate an *AR(1)* process of the form $y_t = a + \phi y_{t-1} + \varepsilon_t$. Use Maximum Likelihood estimation (under menu **Estimation**, after inputting your equation **gc c ar(1)**, select **Options**, then under **Method** select **ML**).[1] Show your output.

d. For the same period as above, now estimate an *MA(2)* process of the form $y_t = \mu + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2}$. Use Maximum Likelihood estimation. The particular specification for your model is **gc c ma(1) ma(2).** Show your output**.**

e. Are your *AR(1)* and *MA(2)* models covariance-stationary and invertible? Explain each case separately and back up your answer. For each, explain why the answer is yes or no and why.

f. Are your residuals white noise? From your estimation exercises, paste your residuals of both of your estimation exercises here and explain why they appear to be white noise or not.

g. Report both the **AIC** and **SIC** of your models. Based on each of these criterions (AIC and SIC), which model should be preferred ?

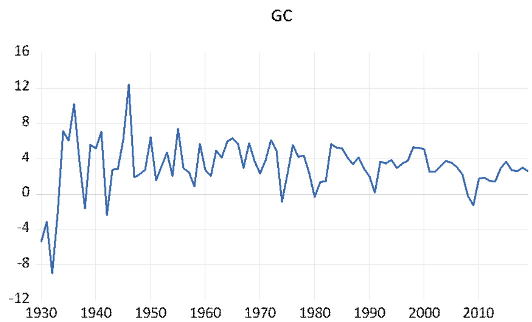
h. From parts (e), (f) and (g), there should be a model that should be statistically preferred to the other one. Under the assumption of quadratic loss (means you can use the conditional mean as the optimal forecast) use the best model to forecast gc for the years 2018 and 2019. Select a **dynamic forecast** (as done in the videos). Compare your forecast to the actual growth rates in consumption that were observed in 2018 and 2019. Paste your forecast diagram here and then type the forecasted values.

i. Refer to your forecast for 2018 (imagine you are in 2017 providing the forecast for year 2018). Let Y be the random variable consumption growth in year 2018 (gc_{2018}). Given the conditional mean and standard deviation of your forecast, find the level of the growth rate y_0 such that the following holds $P[Y \leq y_0] = 10\%$ and then explain the meaning of such probabilistic statement. Assume that the innovation ε_t is distributed normally.

[1] A deep explanation of Maximum Likelihood is beyond the scope of this course. Maximum likelihood estimators start from knowing (or assuming) a particular probability distribution for the random variables. Once these distributions are known, one can produce estimators that maximize the probability of observing the particular sample of data points that we are observing. In our exercise, the ML estimation is implicitly assuming that the innovations ε_t are distributed $N\left(0, \sigma_\varepsilon^2\right)$. Maximum likelihood is not synonimous with assuming normality; rather, it is a method for obtaining estimators (estimators depend on the particular p.d.f.'s assumed for the random variables). Conditional least squares (CLS in Eviews) are an approximation to maximum likelihood. Since computers have a lot of power nowadays, estimating a model via ML is not a problem anymore. As an example, under the classical regression model $Y_i = \beta_0 + \beta_1 X_i + u_i$, if the error is assumed to be $u_i \tilde{} N\left(0, \sigma^2\right)$, then it is possible to show that the ML estimators for β_0 and β_1 are the same estimators that one obtains from OLS. ML estimators are typically consistent and asymptotically efficient (minimum variance under large samples). Although OLS in classical regression yields the same estimators as ML, that's not the case in ARMA models.

Answer Key

a.



b.

Date: 04/01/20 Time: 21:00
Sample: 1980 2017
Included observations: 38

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	0.582	0.582	13.935	0.000
		2	0.252	-0.133	16.606	0.000
		3	-0.052	-0.217	16.723	0.001
		4	-0.000	0.252	16.723	0.002
		5	-0.043	-0.164	16.809	0.005
		6	-0.095	-0.140	17.237	0.008
		7	-0.138	0.080	18.176	0.011
		8	-0.189	-0.199	19.982	0.010
		9	-0.130	0.054	20.870	0.013
		10	-0.052	0.095	21.016	0.021
		11	0.022	-0.101	21.044	0.033
		12	-0.069	-0.141	21.325	0.046

c.

Dependent Variable: GC
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/01/20 Time: 23:42
Sample: 1980 2017
Included observations: 38
Convergence achieved after 17 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.757058	0.578614	4.764932	0.0000
AR(1)	0.632939	0.134571	4.703389	0.0000
SIGMASQ	1.629527	0.321415	5.069848	0.0000

R-squared	0.379673	Mean dependent var	2.903072
Adjusted R-squared	0.344226	S.D. dependent var	1.642523
S.E. of regression	1.330113	Akaike info criterion	3.497531
Sum squared resid	61.92204	Schwarz criterion	3.626815
Log likelihood	-63.45310	Hannan-Quinn criter.	3.543529
F-statistic	10.71094	Durbin-Watson stat	1.808130
Prob(F-statistic)	0.000235		

Inverted AR Roots	.63

d.

Dependent Variable: GC
Method: ARMA Maximum Likelihood (OPG - BHHH)
Date: 04/01/20 Time: 23:43
Sample: 1980 2017
Included observations: 38
Convergence achieved after 24 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.847422	0.492913	5.776725	0.0000
MA(1)	0.826809	0.182682	4.525959	0.0001
MA(2)	0.613612	0.167501	3.663331	0.0008
SIGMASQ	1.361193	0.291056	4.676731	0.0000

R-squared	0.481823	Mean dependent var	2.903072
Adjusted R-squared	0.436101	S.D. dependent var	1.642523
S.E. of regression	1.233423	Akaike info criterion	3.389645
Sum squared resid	51.72533	Schwarz criterion	3.562022
Log likelihood	-60.40325	Hannan-Quinn criter.	3.450975
F-statistic	10.53820	Durbin-Watson stat	2.064034
Prob(F-statistic)	0.000048		

Inverted MA Roots	-.41+.67i	-.41-.67i