ASSIGNMENT 5

Questions 1 (20 points): In this question, you will examine the presence of structural breaks in US CPI monthly data for the period of 2010-2025. The data are in the sheet '`US CPI`'. To ensure compatibility with available packages in STATA, you can model the data using ONLY autoregressive AR(p) models with the command '*regress*'. Your task is to look for evidence of structural breaks by following these steps:

i. Import your data and check for autocorrelation at the first-differenced ΔCPI_t.
ii. Use the first-differenced ΔCPI_t and identify the AR(p) orders with significant estimated coefficients. Then, select the best model in the most parsimonious form (i.e., with the fewest AR terms necessary).
iii. Based on your best AR(p) model for for ΔCPI_t, look for the evidence of structural break. You may use tools such as: dummy variables, Chow test, QLR statistics, or rolling/recursive regressions.
Note: limit the number of observations to 60 in your rolling regressions to make them meaningful in the case you use this method.
iv. Write brief comments (3–4 sentences) summarizing your findings with a connection to relevant macroeconomic events in the U.S. economy.

Question 2 (80 points): In this exercise, we study the relationship between inflation and unemployment—the Phillips Curve—using 1960-2019 U.S. data. The dataset is quarterly and provided in the file "Assignment 5.xlsx", sheet '`PC`'. The data has been seasonally adjusted. Each sheet contains three columns: `date` (day–month–year), `unem` (unemployment rate, in percent), and `CPI` (consumer price index, with 2015Q1 = 100). Data was collected from FRED. Perform these tasks:

i. Review the concept of the Phillips Curve and write a short explanation (3–4 sentences) describing this relationship[1].
ii. Import the data from sheet PC. Construct a year-on-year inflation rate based on the CPI variable (hint: use log transformation and an appropriate lag). Name this variable `infl`. Then plot the time series, plot the ACF, and conduct unit root tests to assess the stationarity of `unem` and `infl` at level.
iii. Now, ignoring stationarity conditions for the moment, estimate the static regression version of the Phillips Curve:

$$\text{infl}_t = \beta_0 + \beta_1 \text{unem}_{t-1} + u_t$$

Test for autocorrelation in the residuals. If it is present, choose a solution and report the results.

[1]You can find a short and concise article summarizing the Phillips Curve from https://www.econlib.org/library/Enc1/PhillipsCurve.html

iv. Next, estimate the distributed lag (DL) model:

$$\text{infl}_t = \beta_0 + \sum_{i=1}^{q} \beta_i \, \text{unem}_{t-i} + u_t$$

Test for autocorrelation. If it exists, propose a remedy and present the corrected results.

v. Then estimate an ADL(p, q) model with lagged inflation and unemployment:

$$\text{infl}_t = \beta^0 + \sum_{i=1}^{p} \beta_i \, \text{infl}_{t-i} + \sum_{i=1}^{q} \gamma_i \, \text{unem}_{t-i} + u_t$$

Use the AIC information criterion to select the optimal model. Test for autocorrelation. If present, apply an appropriate method and report results. Suggested maximum lags: $p = 8$ for inflation, $q = 2$ for unemployment.

vi. Test for heteroskedasticity in the optimal model from Step (v). Report on the findings and explain what they imply.

vii. Restrict the sample to begin in 1983Q1 and re-estimate the ADL(p, q) model with p up to 8 and $q = 2$. Use AIC for model selection. Test for serial correlation and heteroskedasticity. If either exists, correct it and report results.

viii. Based on all estimated models, briefly discuss whether U.S. data from the sheet provide evidence supporting the Phillips curve.

We need to use this table to present the estimation outputs:

Variables	**Model 1 Static**	**Model 2 DL(p)**	**Model 3 DL(p) with robust s.e.**	**Model 4 ADL(p,q)**	**Model 5 ….**
constant	x.xxx (y.yyy)*				
$\boldsymbol{infl_{t-1}}$	-				
$\boldsymbol{infl_{t-2}}$	-				
…	-				
$\boldsymbol{infl_{t-p}}$	-				
$\boldsymbol{unem_t}$	x.xxx (y.yyy)				
$\boldsymbol{unem_{t-1}}$	-				
…					

Note: (**) and (*) respectively present significance level at 1% và 5%.