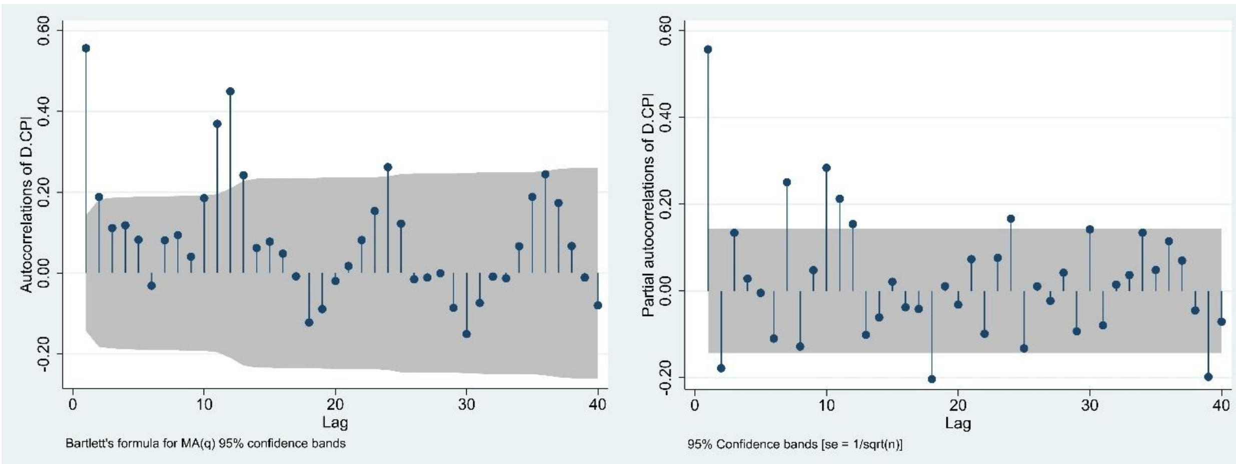
Assignment 5

Dương Hạnh Trang - FAECIU23029

Question 1:

i. Import your data and check for autocorrelation at the first-differenced ΔCPI_t



There is significant autocorrelation remaining in the first-differenced CPI. The ACF plots show spikes outside the confidence bands, particularly at lags 1, 2, 8, 10, and 11, indicating that the series is not white noise.

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).

Because ACF spikes at 1, 2, 8, and 10, I test the significant estimated coefficients of AR(1), AR(2), AR(8), AR(10) and AR(11).

	AR(1) dCPI	AR(2) dCPI	AR(8) dCPI	AR(10) dCPI	AR(11) dCPI
L.dCPI	0.557*** (9.10)	0.658*** (9.04)	0.740*** (9.71)	0.731*** (9.81)	0.673*** (8.82)
L2.dCPI		-0.178* (-2.45)	-0.316*** (-3.45)	-0.287** (-3.09)	-0.254** (-2.75)
L3.dCPI			0.151 (1.64)	0.0676 (0.71)	0.0817 (0.87)
L4.dCPI			-0.0426 (-0.46)	0.0378 (0.41)	-0.0252 (-0.27)
L5.dCPI			0.157 (1.70)	0.115 (1.28)	0.179 (1.95)
L6.dCPI			-0.319*** (-3.46)	-0.314*** (-3.49)	-0.335*** (-3.76)
L7.dCPI			0.341*** (3.71)	0.310** (3.33)	0.300** (3.26)
L8.dCPI			-0.128 (-1.68)	-0.0697 (-0.72)	-0.0814 (-0.85)
L9.dCPI				-0.167 (-1.77)	-0.110 (-1.15)
L10.dCPI				0.284*** (3.76)	0.129 (1.37)
L11.dCPI					0.212** (2.75)
_cons	0.258*** (3.99)	0.298*** (4.47)	0.249** (3.08)	0.182* (2.20)	0.148 (1.79)
N	186	185	179	177	176

t statistics in parentheses

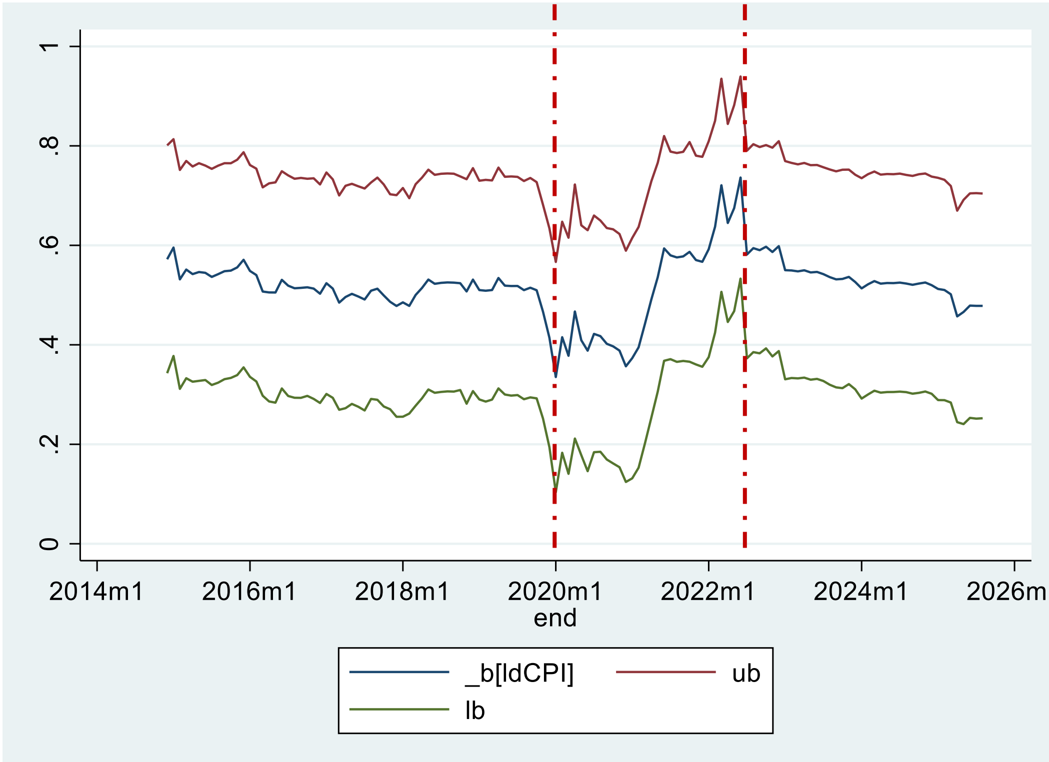
*p<0.05 **p<0.01 *** p<0.001

AR(1) and AR(2): all estimated coefficients are significant at 5% significance level.

Since estimated coefficients from AR1 are significant at 0.1% significance level, and have fewest lags, therefore AR(1) is the best model in the most parsimonious form.

iii. Based on your best AR(p) model for for $\Delta CPIt$, 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.



The estimated coefficient for ldCPI showed significant volatility, deviating from its historical range of 0.5 to 0.6 by first decreasing below 0.4 in early 2020m1 and then rising above 0.7 by mid-2022. This non-constant parameter behavior indicates a structural change in the underlying price dynamics during the 2020-2022 period, suggesting a single, fixed coefficient model may be inappropriate.

iv. Write brief comments (3–4 sentences) summarizing your findings with a connection to relevant macroeconomic events in the U.S. economy.

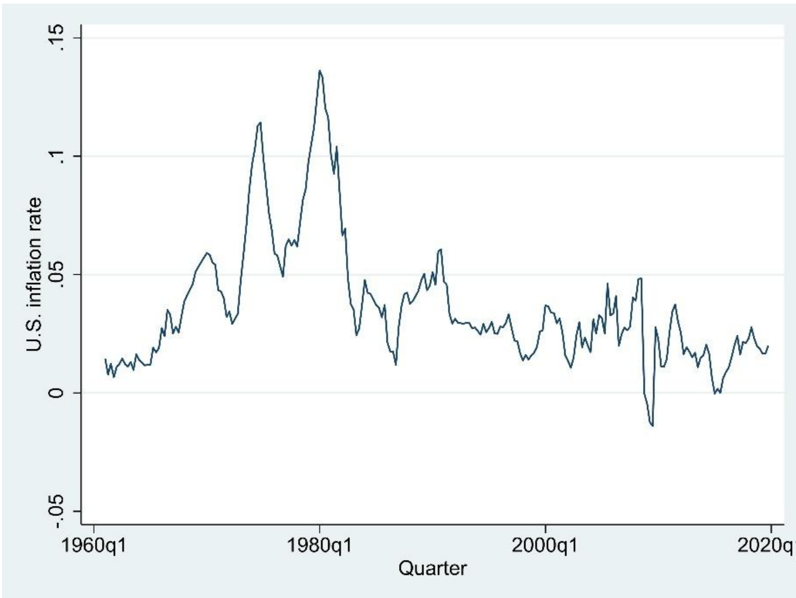
The initial drop below 0.4 aligns with the COVID-19 demand shock in early 2020, where lockdowns caused sharp, temporary disinflation and erratic price behavior. The subsequent rise above 0.7 corresponds to the post-pandemic inflation surge driven by massive fiscal stimulus, supply chain disruptions, and tight labor markets in 2021-2022, which fundamentally changed the persistence and responsiveness of inflation.

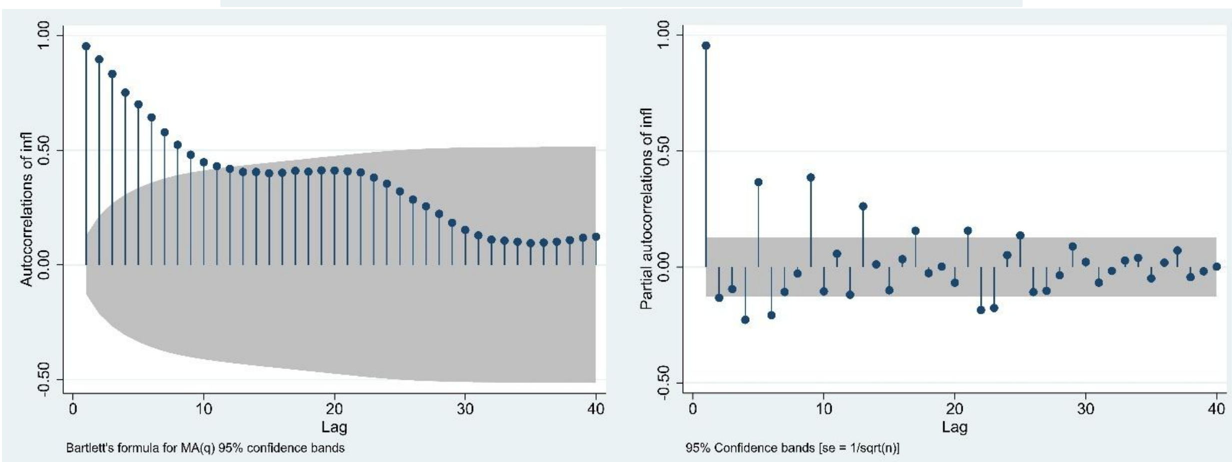
Question 2:

i.

The Phillips Curve illustrates a inverse relationship where inflation rise rapidly when unemployment is low and slowly when unemployment is high. However, analyses by Friedman and Phelps distinguish between the "short-run", and the "long-run". The trade-off between inflation and unemployment holds in short run. But the unemployment returns to its natural rate in the long run (represented as a vertical line) because workers' inflation expectations adjust.

ii.





The ACF and PACF plots show that the inflation series is likely nonstationary due to the very slow decay of the ACF.

```
. dfuller infl, lags(10)

Augmented Dickey-Fuller test for unit root

Variable: infl                          Number of obs  = 225
                                        Number of lags =  10

H0: Random walk without drift, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	-2.105	-3.468	-2.882	-2.572

MacKinnon approximate *p*-value for Z(t) = **0.2427**.

Fail to reject null hypothesis => infl nonstationary even at 10% significance level

```
. dfuller unem, lags(10)

Augmented Dickey-Fuller test for unit root

Variable: unem                          Number of obs  = 229
                                        Number of lags =  10

H0: Random walk without drift, d = 0
```

	Test statistic	Dickey-Fuller critical value 1%	5%	10%
Z(t)	-2.974	-3.467	-2.881	-2.571

MacKinnon approximate *p*-value for Z(t) = **0.0374**.

Since p = 0.0374, we fail to reject null hypothesis at 1% significance level => unem nonstationary

iii.

Static model

```
. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
```

lags(*p*)	chi2	df	Prob > chi2
1	214.767	1	0.0000

H0: no serial correlation

=> reject the null hypothesis => serial correlation between independent variables and residuals

=> correct by robust

```
. estat dwatson

Durbin-Watson d-statistic(  2,   236) =  .0918878
```

Since the calculated Durbin-Watson d-statistic of 0.0918878 is extremely close to 0

=> serial correlation between independent variables and residuals

iv.

Because the data is quarter, I choose q = 4

$$infl_t = \beta_0 + \sum_{i=1}^{4} \beta_i unem_{t-i} + u_t$$

Breusch-Godfrey LM test for autocorrelation

lags(*p*)	chi2	df	Prob > chi2
1	**215.668**	**1**	**0.0000**

H0: no serial correlation

Since p-value = 0.0000, I reject the null hypothesis => serial correlation between independent variables and residuals

=> Solution: robust

```
. estat dwatson

Durbin-Watson d-statistic(  5,   236) =  .1258888
```

Since the calculated Durbin-Watson d-statistic of 0.1259 is close to 0

=> serial correlation between independent variables and residuals

I try to avoid autocorrelation by add for more lags, now I choose q = 6

```
. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
```

lags(p)	chi2	df	Prob > chi2
1	**213.036**	**1**	**0.0000**

H0: no serial correlation

Since p-value = 0.0000, I reject the null hypothesis => serial correlation between independent variables and residuals

v.

ADL(p,q)

```
. est stat ADL11 ADL21 ADL31 ADL41 ADL51 ADL61 ADL71 ADL81 ADL12 ADL22 ADL32 ADL42 ADL52 ADL62 ADL72 ADL82

Akaike's information criterion and Bayesian information criterion
```

Model	N	ll(null)	ll(model)	df	AIC	BIC
ADL11	235	-565.5596	-275.5902	3	557.1805	567.5592
ADL21	234	-563.0706	-273.33	4	554.6601	568.4813
ADL31	233	-560.7508	-271.9744	5	553.9488	571.204
ADL41	232	-558.2092	-265.7014	6	543.4029	564.0833
ADL51	231	-555.8313	-245.8264	7	505.6527	529.7497
ADL61	230	-553.4964	-241.4252	8	498.8504	526.355
ADL71	229	-551.232	-239.7109	9	497.4217	528.3252
ADL81	228	-548.8874	-238.5113	10	497.0226	531.3161
ADL12	235	-565.5596	-270.1087	4	548.2175	562.0558
ADL22	234	-563.0706	-268.9156	5	547.8312	565.1078
ADL32	233	-560.7508	-267.8051	6	547.6101	568.3164
ADL42	232	-558.2092	-261.3286	7	536.6571	560.7843
ADL52	231	-555.8313	-242.0198	8	500.0396	527.5789
ADL62	230	-553.4964	-237.0244	9	492.0488	522.9915
ADL72	229	-551.232	-234.7156	10	489.4312	523.7684
ADL82	228	-548.8874	-233.3285	11	$\boxed{488.657}$	526.3798

AIC lowest at ADL(8,2) => ADL(8,2) is the optimal model

Check autocorrelation for ADL(8,2)

```
. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
```

lags(p)	chi2	df	Prob > chi2
1	**2.033**	**1**	**0.1539**

H0: no serial correlation

p-value = 0.1539 => fail to reject the null hypothesis => no serial correlation between independent variables and residuals

vi.

```
. estat hettest

Breusch-Pagan/Cook-Weisberg test for heteroskedasticity
Assumption: Normal error terms
Variable: Fitted values of infl

H0: Constant variance

    chi2(1) =   2.16
Prob > chi2 = 0.1418
```

Since p-value = 0.1418, we fail to reject the null hypothesis at even 10% significant level => error terms is homoskedasticity

=> It implies that the error terms are white noise

To varify my implication, I do white noise test for residuals

```
. wntestq ress

Portmanteau test for white noise

 Portmanteau (Q) statistic =  1686.8735
 Prob > chi2(40)           =     0.0000
```

Since p-value = 0.000 => We reject the null hypothesis => residual is white noise

vii.

```
. est stat ADL11a ADL21a ADL31a ADL41a ADL51a ADL61a ADL71a ADL81a ADL12a ADL22a ADL32a ADL42a ADL52a ADL62a ADL72a ADL82a
```

Akaike's information criterion and Bayesian information criterion

Model	N	ll(null)	ll(model)	df	AIC	BIC
ADL11a	148	-246.7309	-170.5202	3	347.0405	356.0321
ADL21a	148	-246.7309	-170.2936	4	348.5872	360.5761
ADL31a	148	-246.7309	-170.259	5	350.5179	365.504
ADL41a	148	-246.7309	-168.1174	6	348.2347	366.218
ADL51a	148	-246.7309	-149.8568	7	313.7136	334.6941
ADL61a	148	-246.7309	-148.3519	8	312.7039	336.6816
ADL71a	148	-246.7309	-147.8036	9	313.6071	340.582
ADL81a	148	-246.7309	-147.5832	10	315.1663	345.1385
ADL12a	148	-246.7309	-168.0582	4	344.1163	356.1052
ADL22a	148	-246.7309	-168.0346	5	346.0692	361.0552
ADL32a	148	-246.7309	-167.99	6	347.9799	365.9632
ADL42a	148	-246.7309	-166.2717	7	346.5434	367.5239
ADL52a	148	-246.7309	-147.4759	8	310.9519	334.9296
ADL62a	148	-246.7309	-146.1231	9	310.2463	337.2212
ADL72a	148	-246.7309	-145.4433	10	310.8865	340.8586
ADL82a	148	-246.7309	-144.6494	11	311.2987	344.2681

Note: BIC uses N = number of observations. See [R] BIC note.

AIC lowest at ADL(6,2) => ADL(6,2) is the optimal model for the sample beginning in 1983q1

Check autocorrelation for ADL(6,2)

```
. estat bgodfrey

Breusch-Godfrey LM test for autocorrelation
```

lags(p)	chi2	df	Prob > chi2
1	**2.033**	**1**	**0.1539**

```
H0: no serial correlation
```

p-value = 0.1539 => fail to reject the null hypothesis => no serial correlation between independent variables and residuals

Heteroskedasticity test

```
. estat hettest

Breusch-Pagan/Cook-Weisberg test for heteroskedasticity
Assumption: Normal error terms
Variable: Fitted values of infl

H0: Constant variance

    chi2(1) =   2.16
Prob > chi2 = 0.1418
```

Since p-value = 0.1418, we fail to reject the null hypothesis at even 10% significant level => error terms is homoskedasticity

viii.

Since ADL(8,2) and ADL(6,2) satisfy the assumptions of no serial autocorrelation and homoskedasticity, therefore, we use these two model to explain the relationship between inflation and unemployment.

- Short-Run: At lag 1, there is a negative relationship consistent with the traditional trade-off between inflation and unemployment. This result is statistically significant at the 5% level, with coefficients of -0.496 for ADL(8,2) and -0.409 for ADL(6,2).
- Long-Run: The trade-off disappears as the relationship becomes positive. This suggests that once unemployment returns to its natural level and workers' inflation expectations adjust, the inverse correlation no longer holds. This long-run effect is also significant at the 5% level, indicated by coefficients of 0.467 for ADL(8,2) and 0.380 for ADL(6,2).

	static infl	staticrobust infl	DL4 infl	DL4robust infl	ADL82 infl	ADL62a infl
L.unem	0.124 (1.14)	0.124 (1.35)	1.326* (2.25)	1.326 (1.64)	-0.496*** (-3.35)	-0.409* (-2.16)
L2.unem			-1.003 (-1.00)	-1.003 (-0.78)	0.467** (3.18)	0.380* (2.06)
L3.unem			-0.418 (-0.41)	-0.418 (-0.34)		
L4.unem			0.183 (0.31)	0.183 (0.26)		
L.infl					1.148*** (17.02)	0.952*** (11.32)
L2.infl					-0.145 (-1.40)	-0.170 (-1.62)
L3.infl					0.193 (1.87)	0.169 (1.80)
L4.infl					-0.646*** (-6.73)	-0.543*** (-5.84)
L5.infl					0.591*** (6.18)	0.566*** (5.50)
L6.infl					-0.0798 (-0.78)	-0.124 (-1.60)
L7.infl					-0.0429 (-0.42)	
L8.infl					-0.0589 (-0.87)	
_cons	2.915*** (4.35)	2.915*** (5.55)	3.135*** (4.46)	3.135*** (6.42)	0.317 (1.62)	0.533* (2.10)
N	236	236	236	236	228	148

t statistics in parentheses

="*p<0.05 **p<0.01 *** p<0.001"