ECON8040: APPLIED ECONOMETRICS ASSIGNMENT

Name: Mai Vy (Vivian) Nguyen

Student ID: 47554029

TABLE OF CONTENTS

QUESTION 1:

Plot the CPI inflation rate and comment on its time series behaviour. Does it look stationary? Would you expect it to be stationary?

This is output for the times series of CPI by using View Times Series Plot function on Gretl

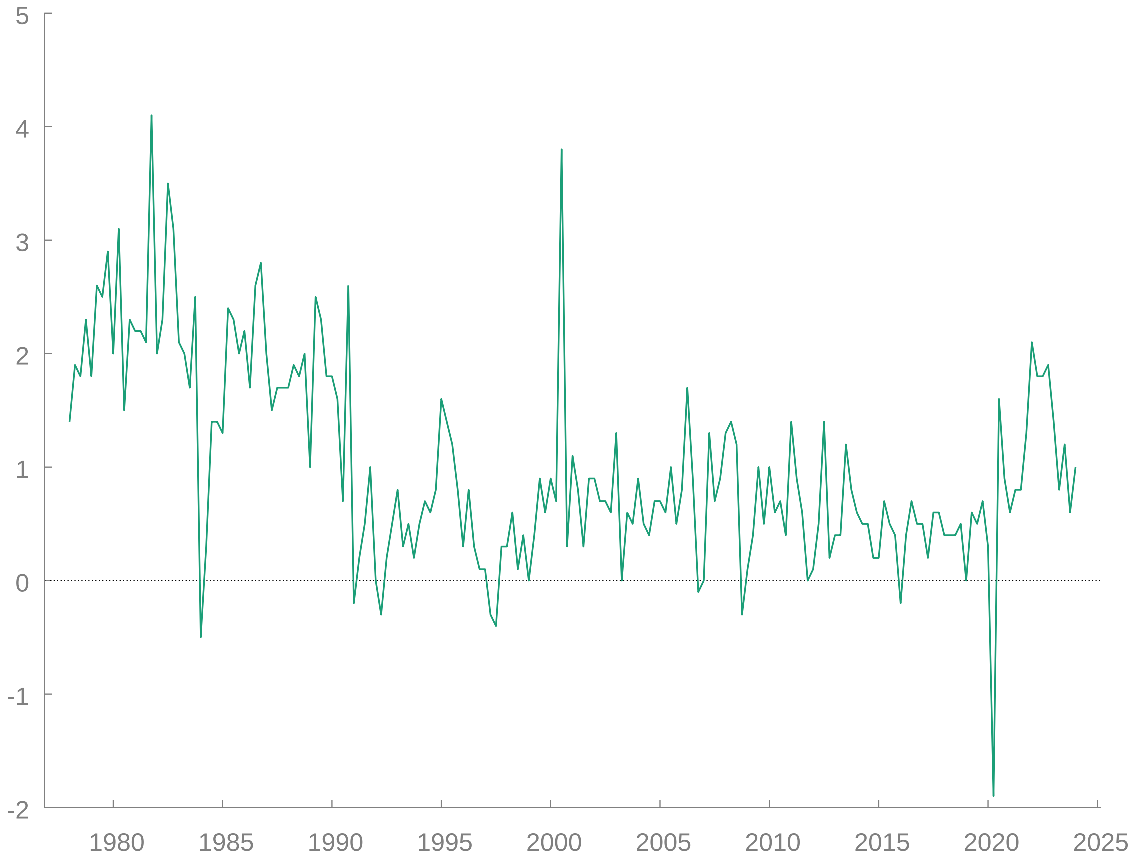


Figure 1: Time series plot of CPI inflation rate

The CPI inflation rate time series plot demonstrates significant fluctuations from the first quarter of 1978 to the first quarter of 2024. High volatility is most noticeable before 1985, throughout the early 2000s, and in 2020. From 1978 to 1985, the rate witnessed significant declines, reaching a low of less than 0% in 1984, and reaching its maximum point of 4% in 1981, reflecting the economic instability of that time. Before 2000, the rate steadied in a range of 0 and 2%. However, the early 2000s experienced another spike, to hit the highest point of nearly 4%. Recently, the CPI inflation rate witnessed a lower rate compared to the previous years in the time series. And the inflation rate hit the bottom of nearly -2% in the earlier period of the COVID-19 pandemic in 2020.

Based on Figure 1, the data clearly shows inflation trends over time. On average, the inflation rate from 1978 to 1990 is higher than in the period following 1990. It also showed higher inflation rates in the earlier years and more stability in the mid-period. Throughout the whole period, there are no obvious long-term trends whether upward or downward. Therefore, the variability indicates that the time series plot for the CPI inflation rate is not stationary

According to the macroeconomic theory, I expect the inflation rate to be stationary. That also is the aim of central banks to stabilise inflation around the target rate. A stable inflation rate is crucial for economic stability and growth. However, the stationary inflation rate has not remained because of the policy and economic changes.

Note: From question 2, I use a common significance level of 5% to test the null hypothesis if the question does not mention any specific significance level.

QUESTION 2:

Estimate an autoregression for the level of the CPI inflation rate. You should estimate the order of the autoregression using the AIC and a maximum possible lag of 6 quarters. In your answer, you should provide a table that contains each order considered and the corresponding value of the AIC. You should place an asterisk next to the smallest value of the AIC.

To estimate an autoregression for the level of CPI inflation rate, I use the Gretl script to create models with lag orders ranging from 1 to 6 quarters.

In this case, I use AIC to select the optimal lag length.

The Gretl Script is below:

```
pmax = 6

besttp =0

bestic = 99999999999999

smpl +(pmax+1)#Ensure the same observation value in any model

loop p = 1..pmax

     ols CPI const CPI(-1 to -p)
     ic = $aic
     if ic < bestic
          besttp = p
          bestic = ic
     endif
endloop
```

The Gretl output below shows 6 models with AIC that are compatible with using each lag order from 1 to 6:

Model 1: AIC model for CPI lag order from 1 to 1

```
OLS, using observations 1979:4-2024:1 (T = 178)
Dependent variable: CPI

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.417294     0.0810541     5.148    6.98e-07 ***
  CPI_1       0.570905     0.0611203     9.341    4.23e-17 ***

Mean dependent var   0.983708   S.D. dependent var   0.875088
Sum squared resid    90.61988   S.E. of regression   0.717555
R-squared            0.331430   Adjusted R-squared   0.327631
F(1, 176)            87.24824   P-value(F)           4.23e-17
Log-likelihood      -192.4863   Akaike criterion     388.9726
Schwarz criterion    395.3361   Hannan-Quinn         391.5531
rho                 -0.200261   Durbin's h          -4.615853
```

Model 2: AIC model for CPI lag order from 1 to 2

```
OLS, using observations 1979:4-2024:1 (T = 178)
Dependent variable: CPI

             coefficient   std. error   t-ratio   p-value
  ---------------------------------------------------------
  const       0.278248     0.0821029     3.389    0.0009   ***
  CPI_1       0.375830     0.0711136     5.285    3.71e-07 ***
  CPI_2       0.331468     0.0705001     4.702    5.21e-06 ***

Mean dependent var   0.983708   S.D. dependent var   0.875088
Sum squared resid    80.45672   S.E. of regression   0.678051
R-squared            0.406411   Adjusted R-squared   0.399627
F(2, 175)            59.90833   P-value(F)           1.51e-20
Log-likelihood      -181.8994   Akaike criterion     369.7987
Schwarz criterion    379.3441   Hannan-Quinn         373.6696
rho                 -0.078080   Durbin's h          -3.297015
```

Model 3: AIC model for CPI lag order from 1 to 3

OLS, using observations 1979:4-2024:1 (T = 178)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.219382	0.0829959	2.643	0.0090	***
CPI_1	0.305037	0.0738802	4.129	5.65e-05	***
CPI_2	0.248211	0.0748745	3.315	0.0011	***
CPI_3	0.211216	0.0733072	2.881	0.0045	***

Mean dependent var	0.983708	S.D. dependent var	0.875088
Sum squared resid	76.79290	S.E. of regression	0.664333
R-squared	0.433441	Adjusted R-squared	0.423673
F(3, 174)	44.37248	P-value(F)	2.39e-21
Log-likelihood	−177.7513	Akaike criterion	363.5026
Schwarz criterion	376.2297	Hannan-Quinn	368.6638
rho	−0.045902	Durbin's h	−3.632265

Model 4: AIC model for CPI lag order from 1 to 4

OLS, using observations 1979:4-2024:1 (T = 178)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.179602	0.0833692	2.154	0.0326	**
CPI_1	0.265155	0.0745825	3.555	0.0005	***
CPI_2	0.201434	0.0761837	2.644	0.0089	***
CPI_3	0.152314	0.0760783	2.002	0.0468	**
CPI_4	0.182809	0.0739296	2.473	0.0144	**

Mean dependent var	0.983708	S.D. dependent var	0.875088
Sum squared resid	74.17139	S.E. of regression	0.654780
R-squared	0.452782	Adjusted R-squared	0.440130
F(4, 173)	35.78619	P-value(F)	9.01e-22
Log-likelihood	−174.6600	Akaike criterion	359.3200
Schwarz criterion	375.2289	Hannan-Quinn	365.7715
rho	−0.006190	Durbin's h	−0.831446

Model 5: AIC model for CPI lag order from 1 to 5

OLS, using observations 1979:4-2024:1 (T = 178)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.180050	0.0846592	2.127	0.0349	**
CPI_1	0.265623	0.0760770	3.492	0.0006	***
CPI_2	0.201845	0.0773696	2.609	0.0099	***
CPI_3	0.152845	0.0779105	1.962	0.0514	*
CPI_4	0.183529	0.0771562	2.379	0.0185	**
CPI_5	−0.00254530	0.0755097	−0.03371	0.9731	

Mean dependent var	0.983708	S.D. dependent var	0.875088
Sum squared resid	74.17090	S.E. of regression	0.656678
R-squared	0.452786	Adjusted R-squared	0.436879
F(5, 172)	28.46388	P-value(F)	5.78e-21
Log-likelihood	−174.6594	Akaike criterion	361.3188
Schwarz criterion	380.4095	Hannan-Quinn	369.0606
rho	−0.006669	Durbin's h	NA

Excluding the constant, p-value was highest for variable 6 (CPI_5)

Model 6: AIC model for CPI lag order from 1 to 6

OLS, using observations 1979:4-2024:1 (T = 178)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.180792	0.0859059	2.105	0.0368	**
CPI_1	0.265609	0.0762988	3.481	0.0006	***
CPI_2	0.202666	0.0789291	2.568	0.0111	**
CPI_3	0.153532	0.0790667	1.942	0.0538	*
CPI_4	0.184448	0.0790517	2.333	0.0208	**
CPI_5	−0.00134326	0.0786298	−0.01708	0.9864	
CPI_6	−0.00432348	0.0761045	−0.05681	0.9548	

Mean dependent var	0.983708	S.D. dependent var	0.875088
Sum squared resid	74.16950	S.E. of regression	0.658589

```
R-squared             0.452796   Adjusted R-squared   0.433596
F(6, 171)             23.58298   P-value(F)           3.26e-20
Log-likelihood       -174.6577   Akaike criterion     363.3155
Schwarz criterion     385.5880   Hannan-Quinn         372.3476
rho                  -0.006921   Durbin's h                 NA

Excluding the constant, p-value was highest for variable 6 (CPI_5)
```

Based on 6 models above, I have the table to summarise the value of AIC for each model with each lag order.

Model	**Lag Order**	**AIC**
1	1	388.9726
2	2	369.7987
3	3	363.5026
4	4	**359.3200***
5	5	361.3188
6	6	363.3155

Model 4 with a lag of 4 quarters has the smallest AIC value of 359.3200 that provides the best fit for the data.

QUESTION 3:

Does your autoregression have a heteroscedasticity problem? How do you know? If it does, propose and implement a remedy.

Based on question 2, I found the best autoregressive model with a lag order is 4. Because I updated the data to make sure the same observation sample for all models with lag orders from 1 to 6 for question 2. Before doing the model for question 3, I restored full range of the data.

I use the Breusch-Pagan-Koenker to check for heteroscedasticity by examining whether the variance of the residuals is constant or not. In this case, using a 5% significance level, conduct a Breusch-Pagan-Koenker test for heteroscedasticity.

There are hypotheses:

$\mathbf{H_0}$: The residual is homoscedastic

$\mathbf{H_1}$: The residual is heteroscedasticity

First, I estimated the linear regression model for CPI with the lag order from 1 to 4.

Model 1: OLS AR(4) model of CPI

OLS, using observations 1979:1-2024:1 (T = 181)

Dependent variable: CPI

	Coefficient	Std. Error	t-ratio	p-value	
const	0.172252	0.0830741	2.073	0.0396	**
CPI_1	0.272487	0.0738478	3.690	0.0003	***
CPI_2	0.206789	0.0758397	2.727	0.0070	***
CPI_3	0.155708	0.0758299	2.053	0.0415	**
CPI_4	0.182379	0.0737230	2.474	0.0143	**

Mean dependent var	1.005525	S.D. dependent var	0.885169
Sum squared resid	75.22618	S.E. of regression	0.653775
R-squared	0.466611	Adjusted R-squared	0.454489
F(4, 176)	38.49145	P-value(F)	4.02e-23
Log-likelihood	−177.3691	Akaike criterion	364.7382
Schwarz criterion	380.7306	Hannan-Quinn	371.2219

```
rho                      0.000561     Durbin's h            0.066396
```

Then I choose the Koenker test on Gretl for testing heteroscedasticity. The Gretl output for Breusch-Pagan-Koenker model is:

```
Breusch-Pagan test for heteroskedasticity
OLS, using observations 1979:1-2024:1 (T = 181)
Dependent variable: scaled uhat^2 (Koenker robust variant)

             coefficient   std. error   t-ratio    p-value
  --------------------------------------------------------
  const      -0.165415      0.132921    -1.244     0.2150
  CPI_1      -0.00252058    0.118159    -0.02133   0.9830
  CPI_2      -0.00970944    0.121346    -0.08001   0.9363
  CPI_3       0.0743033     0.121330     0.6124    0.5411
  CPI_4       0.0992066     0.117959     0.8410    0.4015

  Explained sum of squares = 3.14003

Test statistic: LM = 2.903772,
with p-value = P(Chi-square(4) > 2.903772) = 0.574056
```

The null hypothesis for the Breusch-Pagan-Koenker test is that the residual is homoscedastic. It can be seen that the p_value is 0.574056 which is higher than 0.05. **The null hypothesis is not rejected at the 5% significance level** and it is concluded that there is no evidence of heteroscedasticity in the residual of the AR(4) model.

The Breusch-Pagan-Koenker test results demonstrate that there is no heteroscedasticity problem in the AR(4) model for the CPI inflation rate. That means the variance of residual is constant, which contributes to have the reliable time series modelling and forecasting.

QUESTION 4:

Test the null hypothesis that your autoregression does not have autocorrelation against the alternative hypothesis that it has sixth-order autocorrelation. What do you conclude? If you find evidence of autocorrelation, propose and implement a remedy.

Based on the previous questions, I choose to estimate the model with the best lag = 4. In order to test for autocorrelation, I use the Breusch-Godfrey test with a significance level of 5% to test the null hypothesis that autoregression does not have autocorrelation against the alternative hypothesis that it has sixth-order autocorrelation.

There are hypotheses:

$\mathbf{H_0}$**:** The autoregression does not have autocorrelation

$\mathbf{H_1}$**:** The autoregression has sixth-order autocorrelation

I estimate the OLS model for CPI with lag order =4

```
Model 1: OLS AR (4) model of CPI
OLS, using observations 1979:1-2024:1 (T = 181)
Dependent variable: CPI
```

	Coefficient	Std. Error	t-ratio	p-value	
const	0.172252	0.0830741	2.073	0.0396	**
CPI_1	0.272487	0.0738478	3.690	0.0003	***
CPI_2	0.206789	0.0758397	2.727	0.0070	***
CPI_3	0.155708	0.0758299	2.053	0.0415	**
CPI_4	0.182379	0.0737230	2.474	0.0143	**

Mean dependent var	1.005525	S.D. dependent var	0.885169
Sum squared resid	75.22618	S.E. of regression	0.653775
R-squared	0.466611	Adjusted R-squared	0.454489
F(4, 176)	38.49145	P-value(F)	4.02e-23
Log-likelihood	-177.3691	Akaike criterion	364.7382
Schwarz criterion	380.7306	Hannan-Quinn	371.2219
rho	0.000561	Durbin's h	0.066396

Then I test the autocorrelation of the AR(4) model with lag order test = 6

```
Breusch-Godfrey test for autocorrelation up to order 6
OLS, using observations 1979:1-2024:1 (T = 181)
Dependent variable: uhat

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      -0.117108      0.185435    -0.6315    0.5285
  CPI_1      -0.0850480     0.987247    -0.08615   0.9315
  CPI_2       1.01594       1.04625      0.9710    0.3329
  CPI_3      -0.993600      0.684680    -1.451     0.1486
  CPI_4       0.183109      0.563982     0.3247    0.7458
  uhat_1      0.0888493     0.989839     0.08976   0.9286
  uhat_2     -1.00217       0.922934    -1.086     0.2791
  uhat_3      0.747674      0.536166     1.394     0.1650
  uhat_4     -0.191965      0.396273    -0.4844    0.6287
  uhat_5     -0.0469890     0.138881    -0.3383    0.7355
  uhat_6     -0.150851      0.138359    -1.090     0.2771

  Unadjusted R-squared = 0.020013

Test statistic: LMF = 0.578623,
with p-value = P(F(6,170) > 0.578623) = 0.747

Alternative statistic: TR^2 = 3.622404,
with p-value = P(Chi-square(6) > 3.6224) = 0.728

Ljung-Box Q' = 0.377251,
with p-value = P(Chi-square(6) > 0.377251) = 0.999
```

It can be seen that p_value is 0.728 which is higher than 0.05. **The null hypothesis is not rejected at the 5% significance level** and it is concluded that there is no evidence of sixth-order autocorrelation in the AR(4) model of the CPI inflation rate.

QUESTION 5:

Use your estimated autoregression to forecast the June Quarter 2024 CPI inflation rate. Also, compute an 80% forecast interval for the June Quarter CPI Inflation rate.

I estimate the OLS model for CPI with lag order =4.

Model 1: OLS AR (4) model of CPI

OLS, using observations 1979:1-2024:1 (T = 181)

Dependent variable: CPI

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.172252	0.0664795	2.591	0.0104	**
CPI_1	0.272487	0.0815448	3.342	0.0010	***
CPI_2	0.206789	0.0576057	3.590	0.0004	***
CPI_3	0.155708	0.0545624	2.854	0.0048	***
CPI_4	0.182379	0.0610160	2.989	0.0032	***

Mean dependent var	1.005525	S.D. dependent var	0.885169
Sum squared resid	75.22618	S.E. of regression	0.653775
R-squared	0.466611	Adjusted R-squared	0.454489
F(4, 176)	35.41599	P-value(F)	1.09e-21
Log-likelihood	-177.3691	Akaike criterion	364.7382
Schwarz criterion	380.7306	Hannan-Quinn	371.2219
rho	0.000561	Durbin's h	NA

Then I compute an 80% forecast interval for the June Quarter CPI Inflation rate

The Gretl output for the forecast:

For 80% confidence intervals, t(176, 0.1) = 1.286

Obs	CPI	prediction	std. error	80% interval
2024:2	undefined	0.901564	0.653775	(0.0605613, 1.74257)

The AR(4) model predicts a CPI inflation rate of 0.901564 for the June Quarter 2024. The 80% confidence interval for this forecast ranges from 0.0605613 to 1.74257.

QUESTION 6:

Test the null hypothesis that the autoregression does not have any structural breaks over the time span covered by the sample. Use the critical value method with a 5% significance level. If you find evidence of a structural break, when do you estimate that it occurs?

I use Quandt Likelihood Ratio (QLR) test with the significance level of 5% and 15% trimming to test whether a structural break exists for the AR(4) model

Hypothesis:

$\mathbf{H_0}$: The autoregression does not have any structural breaks over the time span covered by the sample

$\mathbf{H_1}$: The autoregression has a structural break over the time span covered by the sample

First, I estimate the OLS AR(4) model for CPI inflation rate

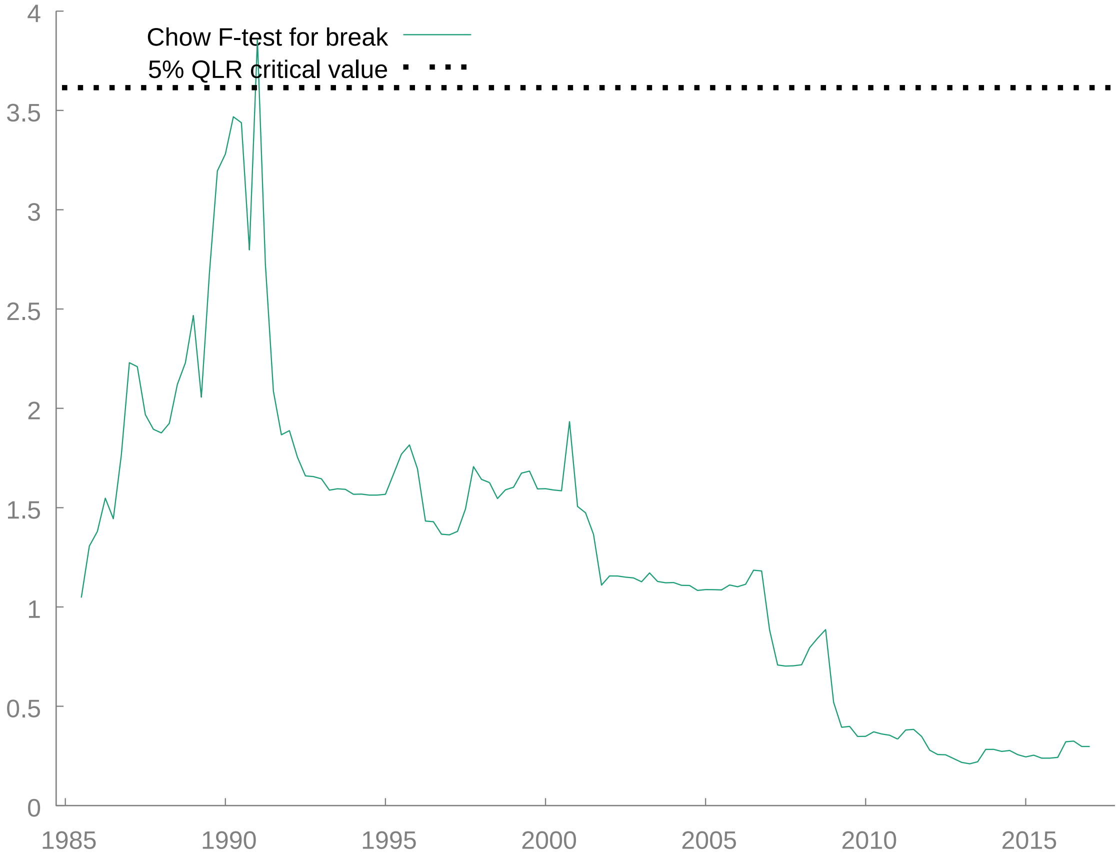
Model 1: OLS AR (4) model of CPI

```
OLS, using observations 1979:1-2024:1 (T = 181)
Dependent variable: CPI
```

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.172252	0.0830741	2.073	0.0396	**
CPI_1	0.272487	0.0738478	3.690	0.0003	***
CPI_2	0.206789	0.0758397	2.727	0.0070	***
CPI_3	0.155708	0.0758299	2.053	0.0415	**
CPI_4	0.182379	0.0737230	2.474	0.0143	**

Mean dependent var	1.005525	S.D. dependent var	0.885169
Sum squared resid	75.22618	S.E. of regression	0.653775
R-squared	0.466611	Adjusted R-squared	0.454489
F(4, 176)	38.49145	P-value(F)	4.02e-23
Log-likelihood	-177.3691	Akaike criterion	364.7382
Schwarz criterion	380.7306	Hannan-Quinn	371.2219
rho	0.000561	Durbin's h	0.066396

Then I compute QLR test for the model and get the Gretl output:




```
Quandt likelihood ratio test for structural break at an unknown point,
with 15 percent trimming:
The maximum F(5, 171) = 3.85592 occurs at observation 1991:1
Asymptotic p-value = 0.0329128 for chi-square(5) = 19.2796
```

Based on the QLR test results, I find evidence of a structural break at quarter one 1991 (1991:1) in the AR(4) model for the CPI inflation rate. The test results show that the p_value of 0.0329128 which is lower than 0.05. Besides, the critical value of QLR test for 5 restrictions with 15% trimming and 5% significance level is 3.66 is smaller than F-statistic (3.85592). **Therefore, the null hypothesis of no structural breaks over the time span covered by the sample is rejected.**

To confirm the period that the autoregression has a structural break, I use Chow test to check whether a structural break occurs in the first quarter of 1991 (1991:1)

There are hypotheses for the Chow test:

$\mathbf{H_0}$: The autoregression does not have a structural break in the first quarter of 1991

$\mathbf{H_1}$: The autoregression has a structural break in the first quarter of 1991

This is the Gretl output after testing Chow test for AR(4) model

```
 Augmented regression for Chow test
OLS, using observations 1979:1-2024:1 (T = 181)
Dependent variable: CPI

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const       1.01283       0.390754     2.592     0.0104  **
  CPI_1       0.182500      0.129116     1.413     0.1593
  CPI_2       0.108888      0.135175     0.8055    0.4216
  CPI_3       0.143774      0.135187     1.064     0.2890
  CPI_4       0.0612379     0.132445     0.4624    0.6444
  splitdum   −0.616404      0.405257    −1.521     0.1301
  sd_CPI_1   −0.0178375     0.158557    −0.1125    0.9106
  sd_CPI_2   −0.00613653    0.164051    −0.03741   0.9702
  sd_CPI_3   −0.127626      0.163932    −0.7785    0.4373
  sd_CPI_4    0.0261694     0.160477     0.1631    0.8707

Mean dependent var   1.005525   S.D. dependent var   0.885169
Sum squared resid    67.60407   S.E. of regression   0.628765
R-squared            0.520656   Adjusted R-squared   0.495427
F(9, 171)            20.63748   P-value(F)           2.85e-23
Log-likelihood      −167.7009   Akaike criterion     355.4017
Schwarz criterion    387.3867   Hannan-Quinn         368.3691
rho                  0.022813   Durbin-Watson        1.952999

Chow test for structural break at observation 1991:1
  F(5, 171) = 3.85592 with p-value 0.0025
```

Based on the Chow test, the p_value is 0.0025 is significantly lower than the 5% significant level. Besides, the F-statistic is 3.85592 is higher than the critical value F(5, 171, 0.05) is 2.26698. Therefore**, the null hypothesis is rejected that means AR(4) model for CPI inflation rate has a structural break at the first of 1991(1991:1).**

Note: Question 7, I use the data for CPI ranges from 1992:1 to 2024:1

QUESTION 7:

Adjust the sample so that it includes data only from the first quarter of 1992 to the first quarter of 2024. Repeat your analysis from questions 2 to 6 with this smaller data set.

For question 7, I use the range of value for the data CPI from 1992:1 to 2024:1

Repeating the analysis from questions 2 to 6:

7.1. Question 2:

To estimate an autoregression for the level of CPI inflation rate, I use the Gretl script to create models with lag orders ranging from 1 to 6 quarters.

In this case, I use AIC to select the optimal lag length.

The Gretl Script is below:

```
smpl 1992:1 2024:1
pmax = 6
bestp = 0
bestic = 9999999999999

loop p =1..pmax
    ols CPI const CPI(-1 to -p)
    ic = $aic
    if ic < bestic
        bestp = p
        bestic = ic
    endif
endloop
```

The Gretl output below shows 6 models with AIC that are compatible with using each lag order from 1 to 6:

Model 1: AIC model for CPI lag order from 1 to 1

```
OLS, using observations 1992:1-2024:1 (T = 129)
Dependent variable: CPI

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const       0.486305     0.0754339     6.447    2.17e-09 ***
  CPI_1       0.248702     0.0859476     2.894    0.0045   ***

Mean dependent var   0.647287   S.D. dependent var   0.595042
Sum squared resid    42.51830   S.E. of regression   0.578610
R-squared            0.061853   Adjusted R-squared   0.054466
F(1, 127)            8.373180   P-value(F)           0.004483
Log-likelihood      -111.4559   Akaike criterion     226.9119
Schwarz criterion    232.6315   Hannan-Quinn         229.2359
rho                 -0.028230   Durbin's h          -1.477723
```

Model 2: AIC model for CPI lag order from 1 to 2

```
OLS, using observations 1992:1-2024:1 (T = 129)
Dependent variable: CPI

             coefficient   std. error   t-ratio   p-value
  --------------------------------------------------------
  const       0.418214     0.0863665     4.842    3.69e-06 ***
  CPI_1       0.213969     0.0881883     2.426    0.0167   **
  CPI_2       0.140096     0.0881695     1.589    0.1146

Mean dependent var   0.647287   S.D. dependent var   0.595042
Sum squared resid    41.68307   S.E. of regression   0.575168
R-squared            0.080281   Adjusted R-squared   0.065683
F(2, 126)            5.499211   P-value(F)           0.005132
Log-likelihood      -110.1763   Akaike criterion     226.3526
Schwarz criterion    234.9321   Hannan-Quinn         229.8386
rho                 -0.000637   Durbin's h                 NA
```

Model 3: AIC model for CPI lag order from 1 to 3

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.394129	0.0941013	4.188	5.26e-05	***
CPI_1	0.206299	0.0891676	2.314	0.0223	**
CPI_2	0.127433	0.0904757	1.408	0.1615	
CPI_3	0.0582943	0.0893154	0.6527	0.5152	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	41.54150	S.E. of regression	0.576482
R-squared	0.083405	Adjusted R-squared	0.061407
F(3, 125)	3.791435	P-value(F)	0.012114
Log-likelihood	−109.9569	Akaike criterion	227.9137
Schwarz criterion	239.3530	Hannan-Quinn	232.5617
rho	0.001501	Durbin's h	NA

Excluding the constant, p-value was highest for variable 4 (CPI_3)

Model 4: AIC model for CPI lag order from 1 to 4

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.352419	0.0998145	3.531	0.0006	***
CPI_1	0.200817	0.0890937	2.254	0.0260	**
CPI_2	0.114133	0.0909303	1.255	0.2118	
CPI_3	0.0345843	0.0911814	0.3793	0.7051	
CPI_4	0.109353	0.0886969	1.233	0.2200	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	41.03845	S.E. of regression	0.575287
R-squared	0.094505	Adjusted R-squared	0.065295
F(4, 124)	3.235408	P-value(F)	0.014562
Log-likelihood	−109.1710	Akaike criterion	228.3421
Schwarz criterion	242.6411	Hannan-Quinn	234.1521
rho	−0.001466	Durbin's h	NA

Excluding the constant, p-value was highest for variable 4 (CPI_3)

Model 5: AIC model for CPI lag order from 1 to 5

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.336897	0.105101	3.205	0.0017	***
CPI_1	0.195381	0.0900676	2.169	0.0320	**
CPI_2	0.112933	0.0912452	1.238	0.2182	
CPI_3	0.0310891	0.0917463	0.3389	0.7353	
CPI_4	0.102218	0.0901765	1.134	0.2592	
CPI_5	0.0414668	0.0853824	0.4857	0.6281	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	40.95990	S.E. of regression	0.577068
R-squared	0.096238	Adjusted R-squared	0.059499
F(5, 123)	2.619549	P-value(F)	0.027473
Log-likelihood	−109.0475	Akaike criterion	230.0949
Schwarz criterion	247.2538	Hannan-Quinn	237.0669
rho	0.008885	Durbin's h	NA

Excluding the constant, p-value was highest for variable 4 (CPI_3)

Model 6: AIC model for CPI lag order from 1 to 6

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	coefficient	std. error	t-ratio	p-value	
const	0.374655	0.109571	3.419	0.0009	***
CPI_1	0.199477	0.0899764	2.217	0.0285	**
CPI_2	0.126292	0.0917697	1.376	0.1713	
CPI_3	0.0324764	0.0915942	0.3546	0.7235	
CPI_4	0.111434	0.0903494	1.233	0.2198	
CPI_5	0.0562602	0.0861278	0.6532	0.5148	
CPI_6	−0.103360	0.0864732	−1.195	0.2343	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	40.48579	S.E. of regression	0.576065
R-squared	0.106699	Adjusted R-squared	0.062766

```
F(6, 122)             2.428684   P-value(F)           0.029773
Log-likelihood       −108.2965   Akaike criterion     230.5930
Schwarz criterion     250.6117   Hannan-Quinn         238.7270
rho                  −0.003616   Durbin's h                 NA

Excluding the constant, p-value was highest for variable 4 (CPI_3)
```

Based on 6 models above, I have the table to summarise value of AIC for each model with each lag orders

Model	**Lag Order**	**AIC**
1	1	226.9119
2	2	**226.3526***
3	3	227.9137
4	4	228.3421
5	5	230.0949
6	6	230.5930

Model 2 with a lag of 2 quarters has the smallest AIC value of 226.3526 that provides the best fit for the data.

7.2. Question 3:

Based on the question 2, I found the best autoregressive model with lag order is 2. After that, the Breusch-Pagan-Koenker is the test the AR(2) model I chose to check for heteroscedasticity by examining whether the variance of the residuals is constant or not. In this case, using a 5% significance level, conduct a Breusch-Pagan-Koenker test for heteroscedasticity.

There are hypotheses:

$\mathbf{H_0}$: the residual is homoscedastic

$\mathbf{H_1}$: the residual is heteroscedasticity

First, I estimated the linear regression model for CPI with the lag order from 1 to 2.

```
Model 1: AIC model for CPI lag order from 1 to 2
OLS, using observations 1992:1-2024:1 (T = 129)
Dependent variable: CPI

                Coefficient Std. Error  t-ratio     p-value
const           0.418214    0.0863665   4.842       <0.0001     ***
CPI_1           0.213969    0.0881883   2.426       0.0167      **
CPI_2           0.140096    0.0881695   1.589       0.1146

Mean dependent var     0.647287     S.D. dependent var    0.595042
Sum squared resid      41.68307     S.E. of regression    0.575168
R-squared              0.080281     Adjusted R-squared    0.065683
F(2, 126)              5.499211     P-value(F)            0.005132
Log-likelihood        -110.1763     Akaike criterion      226.3526
Schwarz criterion      234.9321     Hannan-Quinn          229.8386
rho                   -0.000637     Durbin's h                  NA
```

Then I choose Koenker test for testing heteroscedasticity. The Gretl ouput for Breusch-Pagan-Koenker model is:

```
Breusch-Pagan test for heteroskedasticity
OLS, using observations 1992:1-2024:1 (T = 129)
Dependent variable: scaled uhat^2 (Koenker robust variant)

             coefficient   std. error   t-ratio   p-value
  -------------------------------------------------------
  const      -0.0344840     0.157153    -0.2194   0.8267
  CPI_1      -0.0565765     0.160468    -0.3526   0.7250
  CPI_2       0.109983      0.160434     0.6855   0.4943

  Explained sum of squares = 0.55369

Test statistic: LM = 0.515471,
with p-value = P(Chi-square(2) > 0.515471) = 0.772800
```

The null hypothesis for the Breusch-Pagan-Koenker test is that the residual is homoscedastic. The p-value is 0.772800 is higher than the 5% significance level. Therefore, **the null hypothesis is not rejected at the 5% significance level** and it is concluded that there is no evidence of heteroscedasticity in the residual of the AR(2) model.

The Breusch-Pagan-Koenker test results demonstrate that there is no heteroscedasticity problem in the AR(2) model for the CPI inflation rate. That means the variance of residual is constant, which contributes to have the reliable time series modelling and forecasting.

7.3. Question 4 :

Based on the previous questions, I choose to estimate the model with the best lag = 2.

To test for autocorrelation, I used the Breusch-Godfrey test with a significance level of 5% to test the null hypothesis that autoregression does not have autocorrelation against the alternative hypothesis that it has sixth-order autocorrelation

There are hypotheses:

$\mathbf{H_0}$: The autoregression does not have autocorrelation

$\mathbf{H_1}$: The autoregression has sixth-order autocorrelation

I estimate the OLS model for CPI with lag order = 2

Model 1: AIC model for CPI lag order from 1 to 2

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.418214	0.0863665	4.842	<0.0001	***
CPI_1	0.213969	0.0881883	2.426	0.0167	**
CPI_2	0.140096	0.0881695	1.589	0.1146	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	41.68307	S.E. of regression	0.575168
R-squared	0.080281	Adjusted R-squared	0.065683
F(2, 126)	5.499211	P-value(F)	0.005132
Log-likelihood	-110.1763	Akaike criterion	226.3526
Schwarz criterion	234.9321	Hannan-Quinn	229.8386
rho	-0.000637	Durbin's h	NA

Then I test the autocorrelation of the AR(2) model with lag order test = 6

Breusch-Godfrey test for autocorrelation up to order 6

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: uhat

	coefficient	std. error	t-ratio	p-value	
const	3.03993	1.47436	2.062	0.0414	**
CPI_1	-2.76402	1.58091	-1.748	0.0830	*
CPI_2	-1.90950	1.41773	-1.347	0.1806	
uhat_1	2.73131	1.57720	1.732	0.0859	*

```
  uhat_2       2.47945      1.50351      1.649    0.1017
  uhat_3       0.942575     0.460400     2.047    0.0428  **
  uhat_4       0.636461     0.313898     2.028    0.0448  **
  uhat_5       0.351411     0.154874     2.269    0.0251  **
  uhat_6       0.105896     0.114166     0.9276   0.3555

  Unadjusted R-squared = 0.053219

Test statistic: LMF = 1.124201,
with p-value = P(F(6,120) > 1.1242) = 0.352

Alternative statistic: TR^2 = 6.865201,
with p-value = P(Chi-square(6) > 6.8652) = 0.333

Ljung-Box Q' = 2.6166,
with p-value = P(Chi-square(6) > 2.6166) = 0.855
```

It can be seen that p_value is 0.333 which is higher than 0.05. **The null hypothesis is not rejected at the 5% significance level** and it is concluded that there is no evidence of sixth-order autocorrelation in AR(2) model for the CPI inflation rate.

7.4. Question 5:

I estimate the OLS model for CPI with lag order =2.

Model 1: AIC model for CPI lag order from 1 to 2

OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.418214	0.0863665	4.842	<0.0001	***
CPI_1	0.213969	0.0881883	2.426	0.0167	**
CPI_2	0.140096	0.0881695	1.589	0.1146	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	41.68307	S.E. of regression	0.575168
R-squared	0.080281	Adjusted R-squared	0.065683
F(2, 126)	5.499211	P-value(F)	0.005132
Log-likelihood	-110.1763	Akaike criterion	226.3526
Schwarz criterion	234.9321	Hannan-Quinn	229.8386
rho	-0.000637	Durbin's h	NA

Then I compute an 80% forecast interval for the June Quarter CPI Inflation rate

The Gretl output for forecast:

For 80% confidence intervals, t(126, 0.1) = 1.288

Obs	CPI	prediction	std. error	80% interval
2024:2	undefined	0.716240	0.575168	(-0.0247522, 1.45723)

The AR(2) model predicts a CPI inflation rate of 0.716240 for the June Quarter 2024. The 80% confidence interval for this forecast ranges from -0.0247522 to 1.45723.

7.5. Question 6:

I use Quandt Likelihood Ratio (QLR) test with the significant level 5% and 15% trimming to test whether a structural break exists for AR(2) model.

There are hypothesises:

$\mathbf{H_0}$: The autoregression does not have any structural breaks over the time span covered by the sample

$\mathbf{H_1}$: The autoregression has a structural break over the time span covered by the sample

First, I estimate the OLS AR(2) model for CPI

Model 1: AIC model for CPI lag order from 1 to 2

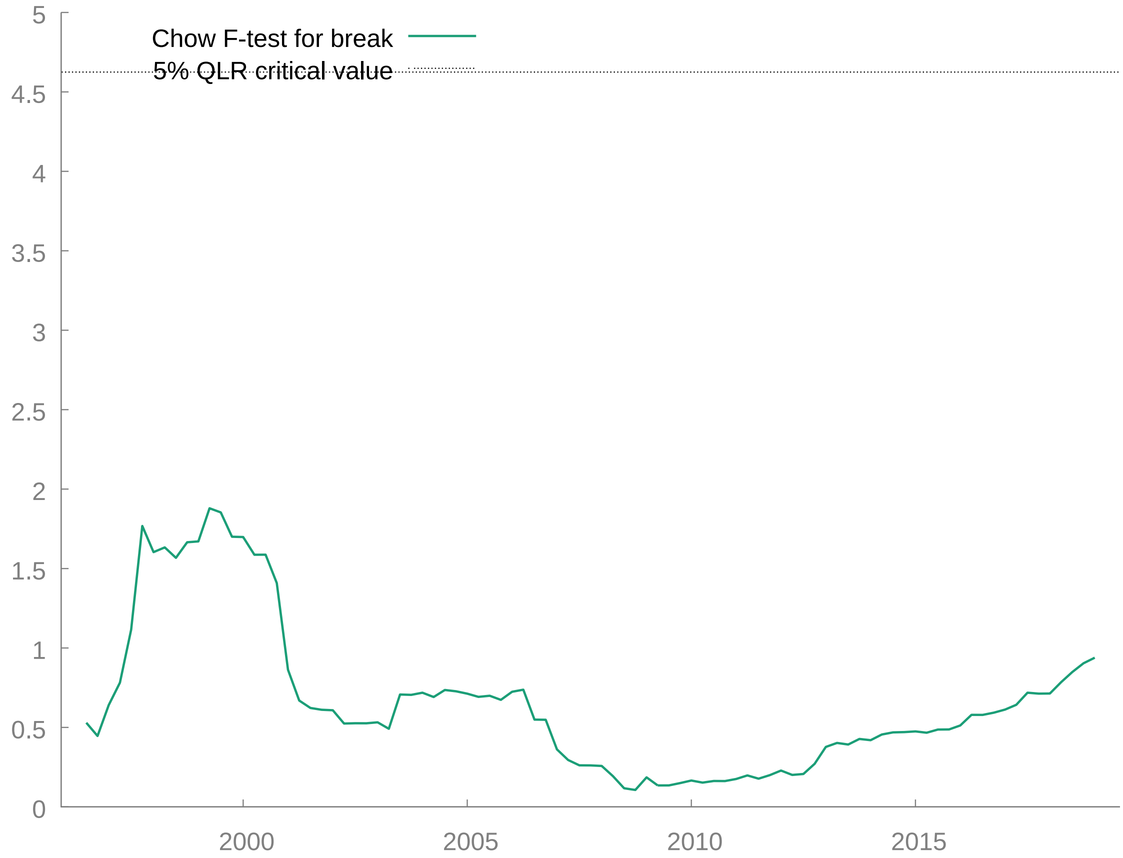
OLS, using observations 1992:1-2024:1 (T = 129)

Dependent variable: CPI

	Coefficient	*Std. Error*	*t-ratio*	*p-value*	
const	0.418214	0.0863665	4.842	<0.0001	***
CPI_1	0.213969	0.0881883	2.426	0.0167	**
CPI_2	0.140096	0.0881695	1.589	0.1146	

Mean dependent var	0.647287	S.D. dependent var	0.595042
Sum squared resid	41.68307	S.E. of regression	0.575168
R-squared	0.080281	Adjusted R-squared	0.065683
F(2, 126)	5.499211	P-value(F)	0.005132
Log-likelihood	-110.1763	Akaike criterion	226.3526
Schwarz criterion	234.9321	Hannan-Quinn	229.8386
rho	-0.000637	Durbin's h	NA

Then I compute QLR test for the model then have the Gretl output:




```
Quandt likelihood ratio test for structural break at an unknown point,
with 15 percent trimming:
The maximum F(3, 123) = 1.87941 occurs at observation 1999:2
Asymptotic p-value = 0.712295 for chi-square(3) = 5.63824
```

Based on the QLR test results, I find evidence of a structural break in second quarter 1999 (1999:2) in the AR(2) model for the CPI inflation rate. The test results show that the p_value of 0.712295 is higher than the 5% significance level. Furthermore, the critical value of QLR test for 3 restrictions with 15% trimming and a 5% significance level is 4.71, which is higher than F-statistic (1.87941). Therefore, **the null hypothesis of no structural breaks over the time span covered by the sample is not rejected**. That means there is no structural breaks over the time span covered by the sample for AR(2) model for CPI inflation rate.

QUESTION 8:

Is your forecast interval computed in Question 7 narrower or wider than the forecast interval that you computed in Question 5? Why do you think that this is the case?

Forecast in Question 5 (Data from 1978:1 to 2024:1 using AR(4) model):

- Forecasted CPI Inflation Rate in the second quarter of 2024: 0.901564
- 80% Confidence Interval: [0.0605613, 1.74257]
- Width of Interval = 1.74257 – 0.0605613 = 1.6820087

Forecast in Question 7 (Data from 1992:1 to 2024:1 using AR(2) model):

- Forecasted CPI Inflation Rate in the second quarter of 2024: 0.716240
- 80% Confidence Interval: [- 0.0247522, 1.45723]
- Width of Interval = 1.45723 – (- 0.0247522) = 1.4819822

⇨ The forecast interval computed in Question 7 is narrower than the forecast interval computed in Question 5

There are several key reasons:

- The sample period for Question 7 (1992:1 to 2024:1) is shorter than the sample used in Question 5 (1978:1 to 2024:1). In Question 7, the shorter period can reduce standard errors that lead to have resulted in a narrower interval.
- The AR(2) model used in Question 7 fits the recent data better, leading to more accurate predictions with lower forecast uncertainty.