Question 1

Data Cleaning

Period	GDP
Q1 1961	343577
Q2 1961	352261
Q3 1961	360748
Q4 1961	365775
Q1 1962	375481
Q2 1962	378238

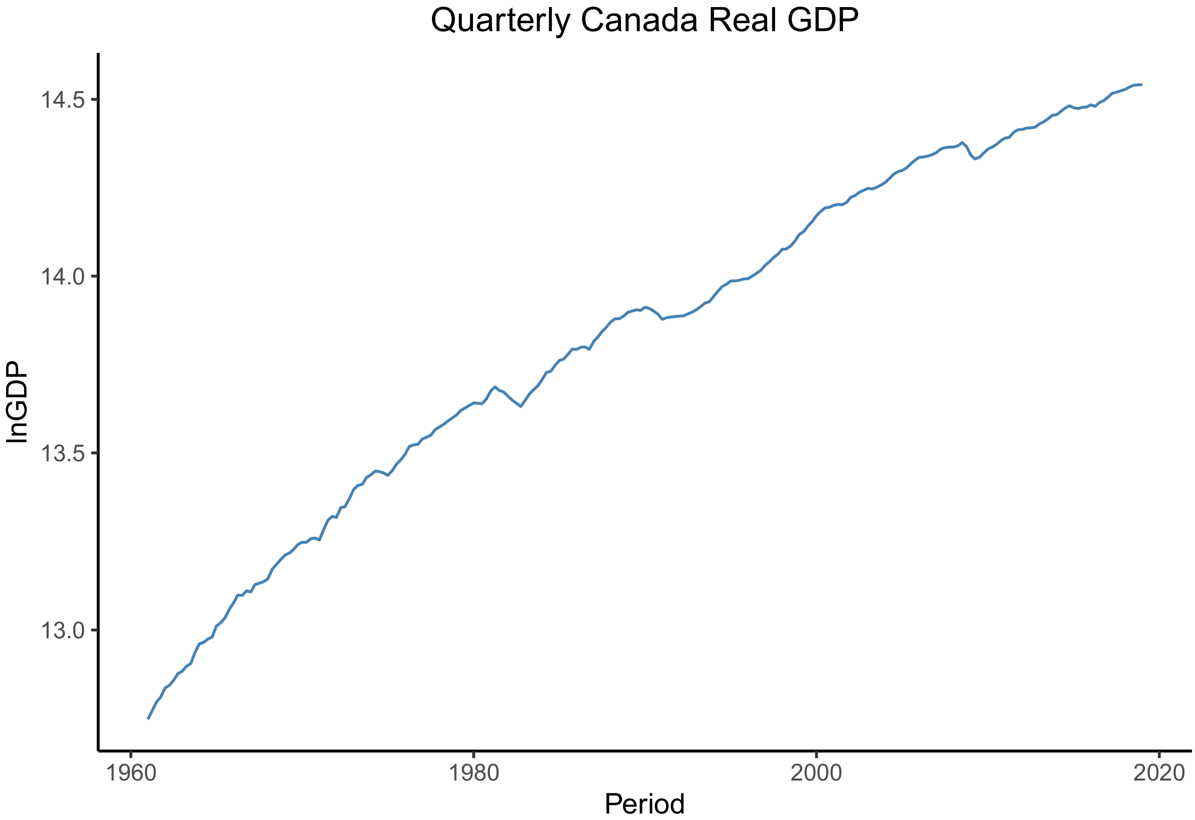
First, we add a column called `lnGDP` and convert `Period` to date-time

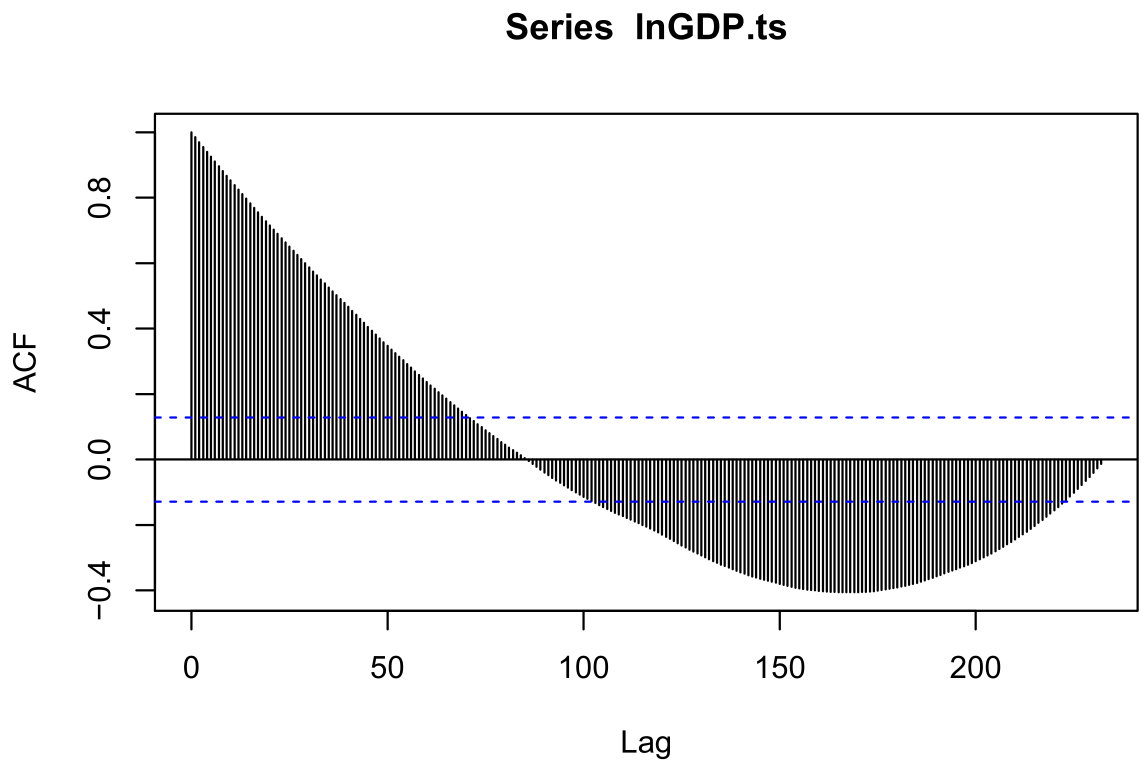
Convert `Period` to a workable date-time format and add the lag. Now, the dataset looks like this.

Period	GDP	lnGDP	laggedlnGDP
1961 Q1	343577	12.74717	NA
1961 Q2	352261	12.77213	12.74717
1961 Q3	360748	12.79593	12.77213
1961 Q4	365775	12.80977	12.79593
1962 Q1	375481	12.83596	12.80977
1962 Q2	378238	12.84328	12.83596

(a)

Now, graph it. We observe that the trend is between that of a positive linear function and a log function. There's no clear structural breaks and that it appears to be stationary.





The log level of real GDP is clearly upward trended and a time trend should therefore be included. Per correlgram, it appears that there is a positive auto-correlation until a certain point between 50 and 100 lags, indicating the shock will last for years. Then, it changes to a negative auto-correlation, where the mean reversion takes place. It is also noticeable that auto correlations for the change in the logarithm of real GDP will also be positive.

(b)

First, we estimate the autoregressive function. The estimated regression equation we get is $\hat{RGDP}_t = 0.007715 + 0.9938 RGDP_{t-1}$.

```
##
## Call:
## ar.ols(x = lnGDP.ts, order.max = 1, dmean = F, Intercept = T)
##
## Coefficients:
##      1
## 0.9938
##
## Intercept: 0.007715 (0.0005197)
##
## Order selected 1  sigma^2 estimated as  6.267e-05
```

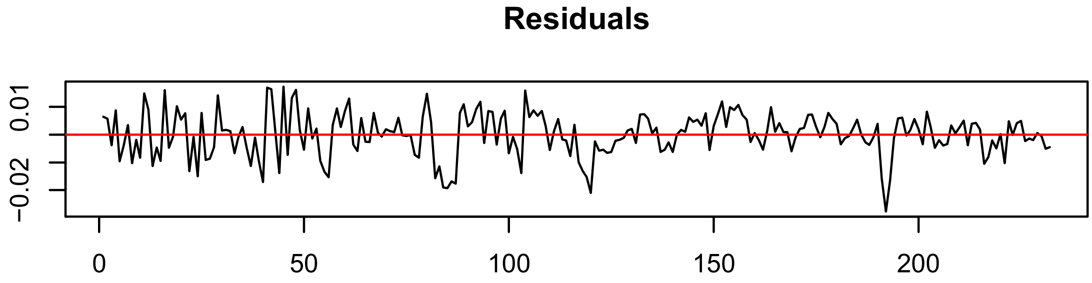
Now, we estimated the mean-adjusted auto-regressive function. The estimated regression equation we get is $R\hat{GDP}_t = 0.007715 + 0.9938RGDP_{t-1}$. As the estimated regression function is the exact same, the model is stable.

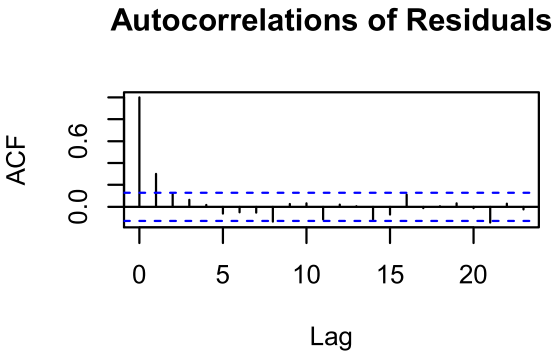
```
##
## Call:
## ar.ols(x = MAlnGDP, order.max = 1, dmean = F, Intercept = T)
##
## Coefficients:
##      1
## 0.9938
##
## Intercept: 0.007715 (0.0005197)
##
## Order selected 1  sigma^2 estimated as  6.267e-05
```

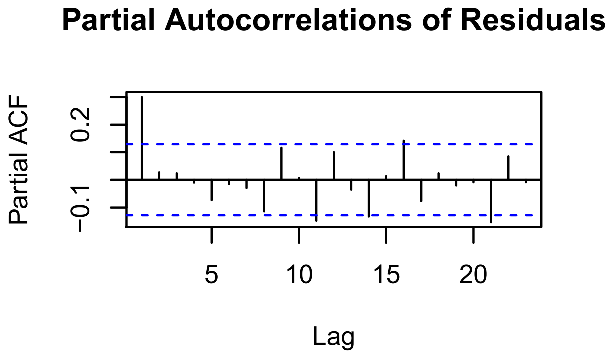
As the coefficients of both models are close to 1, it is believed that the variable, log level of real GDP, appears non-stationary, indicating that error terms will accumulate and Yt will blow up either.

(c)

We fit a model with time trend. The output below is for the model $\Delta GDP_t = \beta_0 + \delta GDP_{t-1} + \mu t + u_t$, where the coefficient for `z.lag.1` corresponds with δ and the coefficient for `tt` corresponds with μ. There is definitely a trend here as δ can be rejected at 5% significance when compared to the DF test statistic.



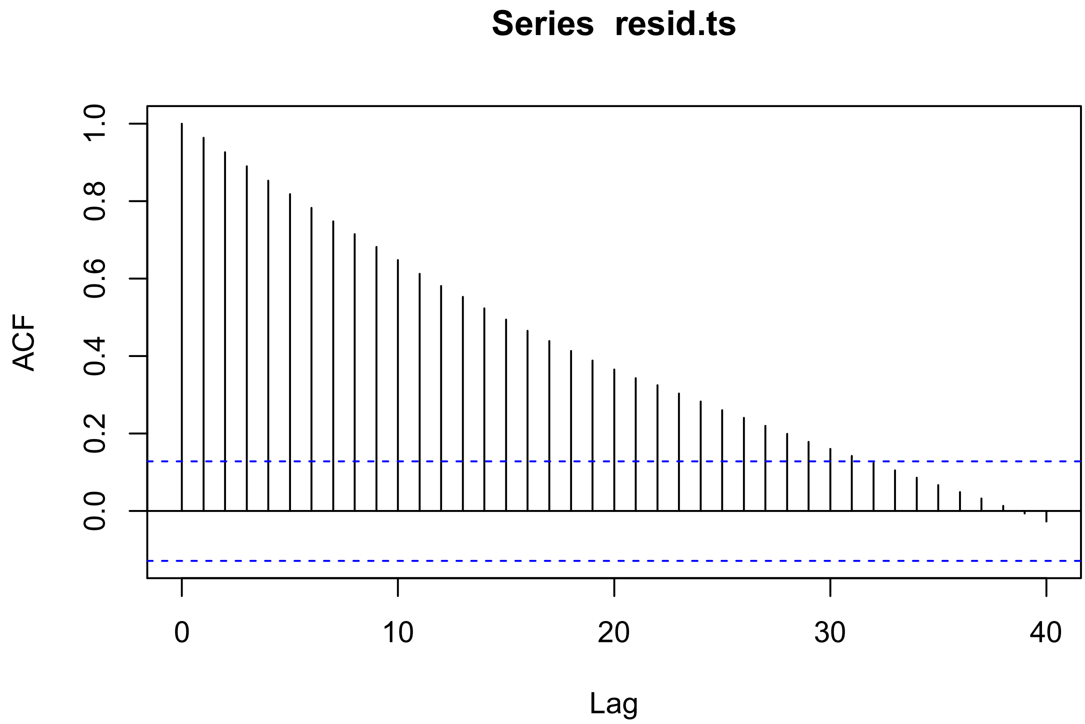


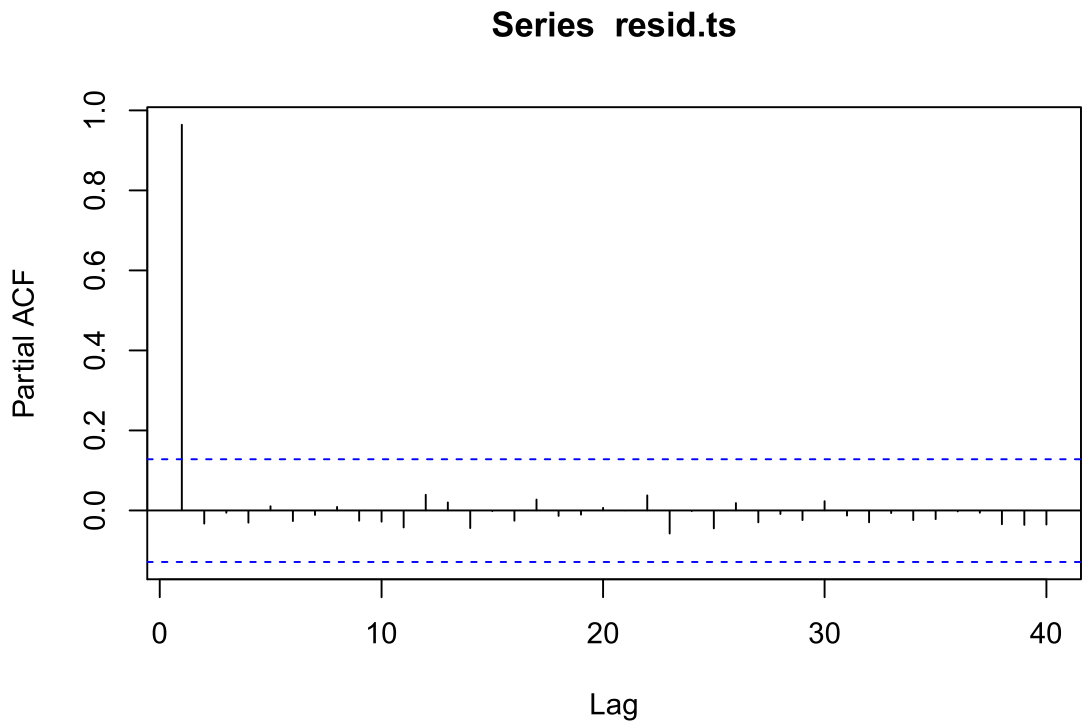


The blue bands in the correlogram below represent values beyond which the autocorrelations are significantly different from zero at 5% significance. For most lags beyond the second lag, the autocorrelation does not exceed the band or only marginally exceed the band. We probably should take a look at the lags of the first and two.The graphs show that there is a positive auto-correlation yet negative partial auto-correlation.

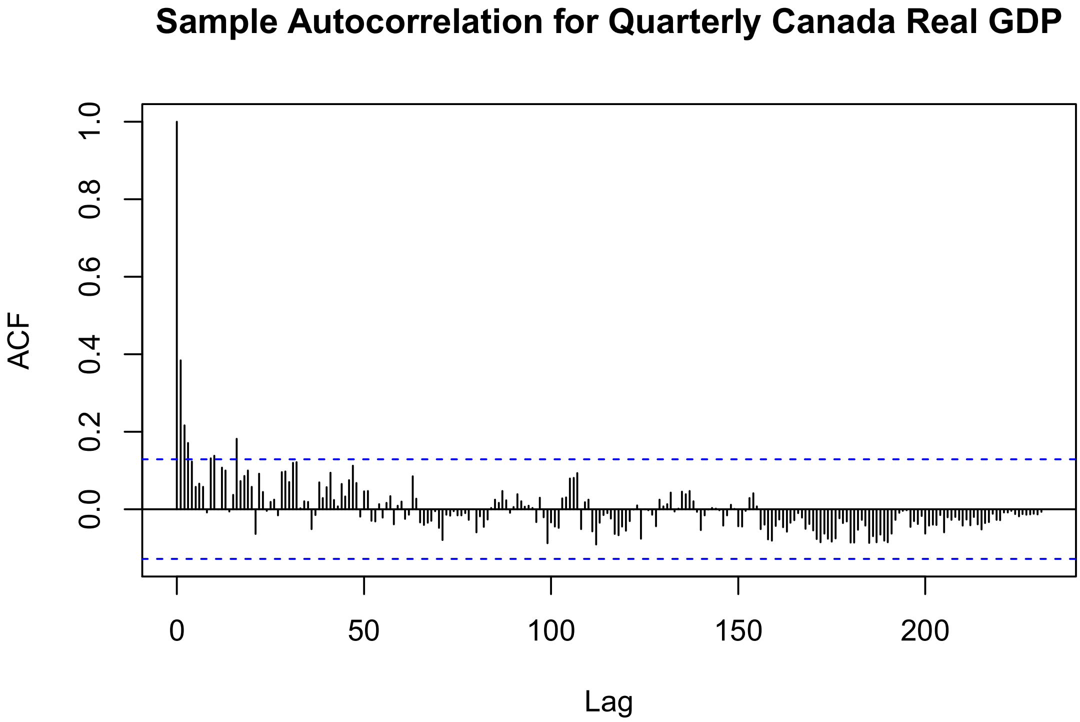
Per ACF correlogram, autocorrelation is probably not significant after the 1st or 2nd lag. Thus, AR(1) or AR(2) is more appropriate.

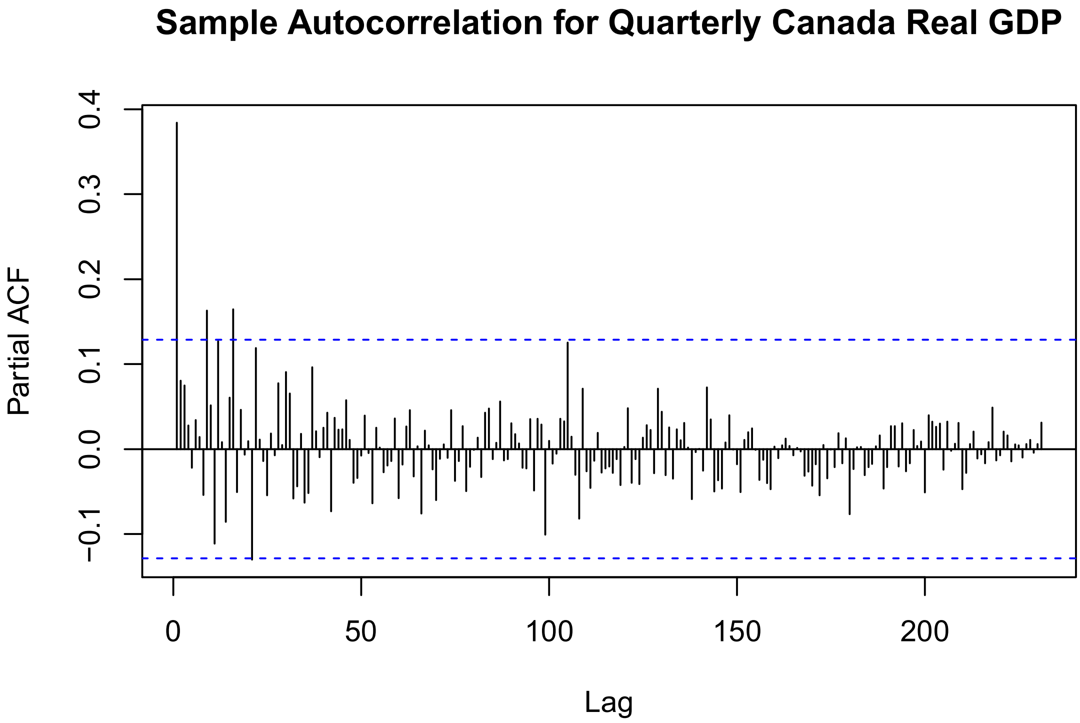
Now, we de-trend using residuals and show the ACF and PACF for the de-trended series.


Series resid.ts
ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Lag


Series resid.ts
Partial ACF
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
Lag

(d)

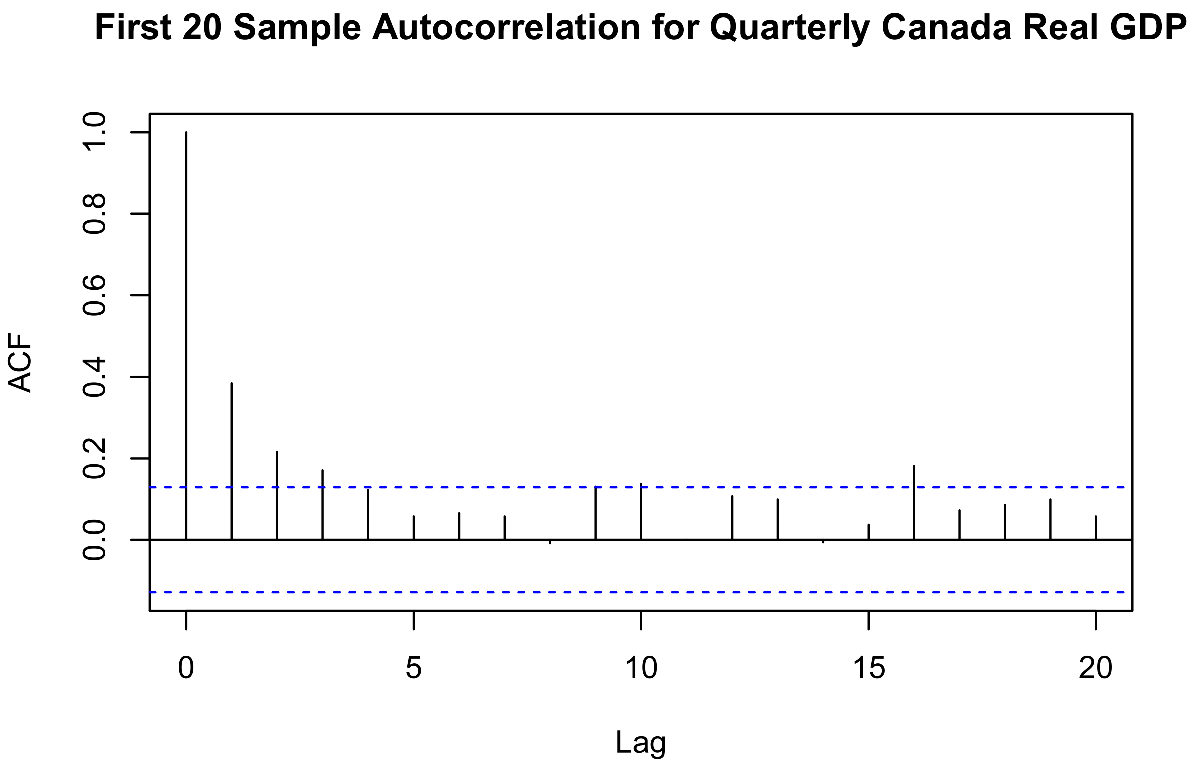

Sample Autocorrelation for Quarterly Canada Real GDP
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0
50
100
150
200
Lag


Sample Autocorrelation for Quarterly Canada Real GDP
Partial ACF
−0.1
0.0
0.1
0.2
0.3
0.4
0
50
100
150
200
Lag

The

blue bands in the correlogram below represent values beyond which the autocorrelations are significantly different from zero at 5% significance. For most lags beyond the third lag, the autocorrelation does not exceed the band or only marginally exceed the band. We probably should take a look at the lags of the first few.

The graphs show that there is a positive auto-correlation: the shock will last for a few quarters.Yet, the PACF correlogram indicates a non-significant negative partial auto-correlation, which indicate that the mean-reversion effects from intermediate days are not that significant.



Autocorrelation probably not significant after the 3rd or 4th lag. Thus, AR(3) or AR(4) is more appropriate.

(e)

We compute the AIC and BIC for p = 0, 1, 2, 3, 4, based on the model in part C. We need to write the model in another way so that it is in a regression class. First we write our own function for a custom output.

```
## Write a criteria function
IC <- function(model) {

  ssr = sum(model$residuals^2)
  t = length(model$residuals)
  npar = length(model$coef)
```

```r
  return(
    round(c("p" = npar - 1,
          "AIC" = log(ssr/t) + npar * 2/t,
          "BIC" = log(ssr/t) + npar * log(t)/t,
          "R2" = summary(model)$r.squared), 4)
  )
}
```

```r
# apply the IC() to an intercept-only model of GDP growth
IC(dynlm((dlnGDP.ts) ~ 1))
```

```
##       p     AIC     BIC      R2
##  0.0000 -9.5320 -9.5172  0.0000
```

```r
# loop IC over models of different orders
order <- 1:4

ICs <- sapply(order, function(x)
        "AR" = IC(dynlm(dlnGDP.ts ~ L(lnGDP.ts, 1:x))))
ICs
```

```
##        [,1]    [,2]    [,3]    [,4]
## p    1.0000  2.0000  3.0000  4.0000
## AIC -9.6605 -9.7541 -9.7454 -9.7337
## BIC -9.6307 -9.7094 -9.6856 -9.6587
## R2   0.1281  0.2021  0.1927  0.1919
```

Then we show the output. Model selection criteria suggests 2 lags.

Question 2

(a)

The graph shows a strong positive autocorrelation for the inflation rate level, indicating level of the inflation rate will persist for quite a while.However, we suspect that the Canadian inflation rate might have a stochastic trend. It also appears a break taking place around 1982.

Taking a look at the data prior to 1982, the level of inflation rate had never returned to the original point, which supports our suspicion of stochastic trend. To further demonstrate it, the Dickey-Fuller test of trend and intercept can be conducted.

(b)

T-statistic is -2 = -0.1/0.05, which is smaller in absolute terms than DF critical values of 2.86 at 5% of significance. Thus, we fail to reject the null of stationary variable and conclude that there is a random walk for inflation.

Robust standard errors are not required. When adding the lag of variable, we already eliminate the effect of autocorrelation in error terms. It doesn't mean we should use 4 lags because the data is quarterly.Yet, we should determine the number of lags to use by computing the BIC and/or AIC for each p, and then selecting the p that minimizes the absolute value of the information criteria. As BIC is more consistent compared to AIC, it is more likely for us to apply BIC in terms of choosing the numbers of lags.

(c)

We are given the following data

Quarter	UrateCt	Inft	Inft.1	cInft
1999:I	7.7	0.8	0.8	0.0
1999:II	7.9	4.3	0.8	3.5
1999:III	7.7	2.9	4.3	-1.4
1999:IV	7.0	1.3	2.9	-1.5
2000:I	6.8	2.1	1.3	0.8

The calculations are as follow:

```
## AR(1)
pred.ar1 = 0.002-0.31*(-1.5)
inf.ar1 = 1.3 + pred.ar1
error.ar1 = inf.ar1 - 2.1

## AR(4)
pred.ar4 = 0.02-0.46*(-1.5)-0.39*(-1.4)-0.25*3.5+0.03*0
inf.ar4 = 1.3 + pred.ar4
error.ar4 = inf.ar4 - 2.1

## ADL(4, 1)
pred.adl = 1.279-0.51*(-1.5)-0.44*(-1.4)-0.30*3.5-0.02*0-0.16*7
inf.adl = 1.3 + pred.adl
error.adl = inf.adl - 2.1
```

The results are as follow:

Metric	AR.1.	AR.4.	ADL.4..1.
Predicted Change in Inflation	0.467	0.381	0.49
Predicted Inflation	1.767	1.681	1.79
Forecast Error	-0.333	-0.419	-0.31

(d)

Perform a test on whether or not Canadian unemployment rates Granger-cause the Canadian inflation rate. Per ADL(4,1) model, t-statistic of 2.29 (ABS(-0.16/0.07)) on the coefficient of Canadian unemployment rate is the only restriction. We can apply t^2 as F-statistics = 5.22, which is greater than the critical value. Therefore, we can reject the null hypothesis that the coefficient on Canadian unemployment rates is zero; namely that Canadian unemployment rates does Granger cause the Canadian inflation rate and simultaneously serves the purpose of predictive content.

Question 3-1

(a)

The average percent of accidents resulting in at least a fatality is 0.8856%.

```
mean(Driving$prcfat)
```

```
## [1] 0.8856363
```

(b)

The regression of prcfat on a linear time trend, 11 monthly dummies, wkends, unem, spdlaw, and beltlaw:

```
##
## Call:
## glm(formula = prcfat ~ t + wkends + unem + spdlaw + beltlaw +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + nov +
##     dec, data = Driving)
##
## Deviance Residuals:
##       Min         1Q     Median         3Q        Max
## -0.132085  -0.032574   0.000621   0.038226   0.132080
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.030e+00  1.030e-01  10.003 2.52e-16 ***
## t           -2.235e-03  4.208e-04  -5.312 7.66e-07 ***
## wkends       6.259e-04  6.162e-03   0.102 0.919329
## unem        -1.543e-02  5.544e-03  -2.782 0.006563 **
## spdlaw       6.709e-02  2.057e-02   3.262 0.001560 **
## beltlaw     -2.951e-02  2.323e-02  -1.270 0.207287
## feb          8.607e-04  2.900e-02   0.030 0.976384
## mar          9.226e-05  2.741e-02   0.003 0.997321
## apr          5.822e-02  2.782e-02   2.093 0.039152 *
## may          7.164e-02  2.764e-02   2.592 0.011129 *
```

```
## jun          1.013e-01  2.809e-02   3.604 0.000510 ***
## jul          1.766e-01  2.726e-02   6.479 4.64e-09 ***
## aug          1.926e-01  2.744e-02   7.018 3.91e-10 ***
## sep          1.600e-01  2.820e-02   5.674 1.64e-07 ***
## oct          1.010e-01  2.767e-02   3.651 0.000435 ***
## nov          1.395e-02  2.814e-02   0.496 0.621346
## dec          9.200e-03  2.786e-02   0.330 0.741960
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 0.00330791)
##
##     Null deviance: 1.06525  on 107  degrees of freedom
## Residual deviance: 0.30102  on  91  degrees of freedom
## AIC: -292.84
##
## Number of Fisher Scoring iterations: 2
```

Based on the results, the higher state unemployment rate decreases the percent of fatal accidents by 0.01543. With the p-value of 0.0066,this is a statistically significant effect. This may be because the higher state unemployment rate will reduce the economics activity, and then probably decrease the chance of an accident results in a fatality.

The higher speed limits, the higher the percent of fatal accidents, which increased by 0.06709 percentage points. This is also a statistically significant effect because of the very small p-value(0.00156).

Lastly, the implementation of seat belt law is estimated to decrease the percent of fatal accidents by 0.02951 with the p-value of 0.2073.

(c)

(c) To test the errors for AR(1) serial correlation, we conduct the Durbin-Waston statistic: H0: $\rho = 0$ (no serial correlation) H1: $\rho \neq 0$ (there is serial correlation)

```
##
##  Durbin-Watson test
##
## data:  prcfat_ltr
## DW = 1.43, p-value = 0.001057
## alternative hypothesis: true autocorrelation is greater than 0
```

The result above shows that the Durbin-Waston d-statistic is 1.43,which is between 0 and d-lower (1.522 with n=100 & k=1). As a result, we reject the null and conclude that there is autocorrelation.

(d)

Compute the first order autocorrelations for `unem` and `prcfat`.

```
##
## Partial autocorrelations of series 'na.omit(Driving$unem)', by lag
##
##     1     2     3     4     5
## 0.941 0.065 0.118 0.090 0.057
```

```
##
## Partial autocorrelations of series 'na.omit(Driving$prcfat)', by lag
##
##      1      2      3      4      5
##  0.708 -0.114 -0.181 -0.166 -0.115
```

The first order autocorrelation for unem and prcfat are 0.941 and 0.708 respectively. For `unem` it is pretty high, so its very possible that there's a unit root. It is not as high for `prcfat`, so concerns about a unit root there are not as significant.

(e)

Estimate the model in (b) using first differences for unem and prcfat

```
##
## Call:
## lm(formula = disprcfat ~ t + wkends + disunem + spdlaw + beltlaw +
##     feb + mar + apr + may + jun + jul + aug + sep + oct + nov +
##     dec, data = new_driving)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.130088 -0.044411  0.002195  0.032828  0.166416
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.1268680  0.1048114  -1.210  0.22928
## t            0.0001433  0.0004849   0.296  0.76823
## wkends       0.0068097  0.0072276   0.942  0.34862
## disunem      0.0125343  0.0161094   0.778  0.43857
## spdlaw      -0.0071825  0.0237979  -0.302  0.76349
## beltlaw      0.0008251  0.0265048   0.031  0.97523
## feb          0.0346229  0.0370460   0.935  0.35250
## mar          0.0419347  0.0389248   1.077  0.28421
## apr          0.0985704  0.0382989   2.574  0.01170 *
## may          0.0568103  0.0374417   1.517  0.13270
## jun          0.0540340  0.0347738   1.554  0.12373
```

```
## jul          0.0878394  0.0331103   2.653  0.00943 **
## aug          0.0589255  0.0396687   1.485  0.14092
## sep          0.0065434  0.0379741   0.172  0.86358
## oct         -0.0323897  0.0352025  -0.920  0.35998
## nov         -0.0591083  0.0354151  -1.669  0.09859 .
## dec          0.0272795  0.0363246   0.751  0.45462
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06724 on 90 degrees of freedom
## Multiple R-squared:  0.3436, Adjusted R-squared:  0.2269
## F-statistic: 2.945 on 16 and 90 DF,  p-value: 0.0006167
```

After replacing the orginal unem and prcfat with the first difference, we found that the sign of coefficient for spdlaw and beltlaw become negatvie and positive respectively. This conflicts the result in part(b) and it does no make sense. Moreover, the multiple R-squared drops to 0.3436 from 0.7174. Therefore, we conclude that the change in prcfat cannot be explained by the change in unem or any of the policy variables.

Question 3 - 2

(a)

To test if ltotacc has unit root, we conduct the hypothesis test as follows: H0: $\delta = 0$ H1: $\delta < 0$

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.207558 -0.033619  0.003362  0.039209  0.181057
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.0514     0.6192   3.313  0.00127 **
## z.lag.1      -0.1923     0.0581  -3.311  0.00128 **
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06488 on 105 degrees of freedom
## Multiple R-squared:  0.09451,    Adjusted R-squared:  0.08589
## F-statistic: 10.96 on 1 and 105 DF,  p-value: 0.001277
##
##
## Value of test-statistic is: -3.3105 5.506
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

The t value is -3.3105,which is smaller than the t = -3.136 at 2.5%. Therefore, we reject the null and conclude that ltotacc does not have a unit root at 2.5%.

(b)

Repeat (a) with two lagged changes and conduct the ADF test as follows: H0: $\delta = 0$ H1: $\delta < 0$

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.179847 -0.022024  0.008181  0.036650  0.130035
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.90975    0.60361   1.507  0.13488
## z.lag.1     -0.08503    0.05663  -1.501  0.13636
## z.diff.lag1 -0.47275    0.09956  -4.748 6.78e-06 ***
## z.diff.lag2 -0.28700    0.09372  -3.062  0.00282 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.05802 on 101 degrees of freedom
## Multiple R-squared:  0.2635, Adjusted R-squared:  0.2417
## F-statistic: 12.05 on 3 and 101 DF,  p-value: 8.318e-07
##
##
## Value of test-statistic is: -1.5015 1.3178
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

The t value with two lagged change is -1.5015,which becomes bigger than the t = -3.136 at 2.5%. Therefore, we cannot reject the null, and we conclude that ltotacc has a unit root at 2.5% with two lagged changes.

(c)

Add a linear trend and repeat (b), we conduct the ADF test again as follows: H0: $\delta = 0$ H1: $\delta < 0$

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.162429 -0.026676  0.004885  0.034603  0.127044
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  5.89132    1.11293   5.294 7.12e-07 ***
## z.lag.1     -0.42527    0.11601  -3.666 0.000397 ***
## tt           0.42654    0.11635   3.666 0.000397 ***
## z.diff.lag1 -0.25256    0.11591  -2.179 0.031677 *
## z.diff.lag2 -0.17249    0.09585  -1.799 0.074964 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
```

```
## Residual standard error: 0.05535 on 100 degrees of freedom
## Multiple R-squared:  0.3364, Adjusted R-squared:  0.3099
## F-statistic: 12.67 on 4 and 100 DF,  p-value: 2.215e-08
##
##
## Value of test-statistic is: -3.6658 56.4384 6.7302
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

We look at the first t-statistic at the bottm, which is -3.6658. This is slightly bigger than the critical value at 2.5% (-3.683). Therefore, we cannot reject the null , and we conclude that ltoacc still has a unit root after adding a linear trend.

(d)

Compared the results from parts a-c, ltotacct has unit root with either two lagged changes or even a linear time trend. Therefore, the best characterization seems to be suggested by the regression in part(c).

(e) Repeat (a-d) for prcfat

a.

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.141962 -0.036649 -0.004246  0.038963  0.207079
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.25690    0.06146   4.180 6.06e-05 ***
## z.lag.1     -0.29061    0.06893  -4.216 5.28e-05 ***
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07106 on 105 degrees of freedom
## Multiple R-squared:  0.1448, Adjusted R-squared:  0.1366
## F-statistic: 17.77 on 1 and 105 DF,  p-value: 5.282e-05
##
##
## Value of test-statistic is: -4.216 8.8914
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

The t value is -4.216, which is smaller than the t = -3.136 at 2.5%. Therefore, we reject the null and conclude that prcfat does not have a unit root at 2.5%.

b.

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.152015 -0.040662 -0.003668  0.036441  0.206577
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.33928    0.07201   4.712 7.87e-06 ***
## z.lag.1     -0.38370    0.08087  -4.744 6.89e-06 ***
## z.diff.lag1  0.15163    0.09953   1.523   0.1308
## z.diff.lag2  0.18345    0.09800   1.872   0.0641 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.07075 on 101 degrees of freedom
```

```
## Multiple R-squared:  0.1843, Adjusted R-squared:  0.1601
## F-statistic: 7.607 on 3 and 101 DF,  p-value: 0.0001229
##
##
## Value of test-statistic is: -4.7445 11.2618
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.46 -2.88 -2.57
## phi1  6.52  4.63  3.81
```

The t value with two lagged change is -4.7445, which is still smaller than the t = -3.136 at 2.5%. Therefore, we reject the null again, and conclude that prcfat does not have a unit root at 2.5% with two lagged changes.

c.

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.139622 -0.046311  0.002222  0.036136  0.226407
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  1.02631    0.12900   7.956 2.83e-12 ***
## z.lag.1     -0.45812    0.08663  -5.288 7.28e-07 ***
## tt           0.45759    0.08653   5.288 7.29e-07 ***
## z.diff.lag1  0.18925    0.09933   1.905   0.0596 .
## z.diff.lag2  0.21860    0.09765   2.239   0.0274 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.06951 on 100 degrees of freedom
## Multiple R-squared:  0.2205, Adjusted R-squared:  0.1893
## F-statistic: 7.073 on 4 and 100 DF,  p-value: 4.681e-05
```

```
##
##
## Value of test-statistic is: -5.2882 31.6121 13.9839
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.99 -3.43 -3.13
## phi2  6.22  4.75  4.07
## phi3  8.43  6.49  5.47
```

The first t-statistic at the bottm is -5.2882, which is smaller than the critical value at 2.5% (-3.683). Therefore, we reject the null and conclude that prcfat still does not have a unit root after adding a linear trend.

d.

Unlike ltoacc, prcfat does not have a unit root in any cases. The evidence is strong no matter we included two lagged change/ a linear time trend in our regression or not.

Question 4

Dataset is as follow:

year	unem	inf	inf_1	unem_1	cinf	cunem
1948	3.8	8.1	NA	NA	NA	NA
1949	5.9	-1.2	8.1	3.8	-9.3	2.1000000
1950	5.3	1.3	-1.2	5.9	2.5	-0.5999999
1951	3.3	7.9	1.3	5.3	6.6	-2.0000000
1952	3.0	1.9	7.9	3.3	-6.0	-0.3000000
1953	2.9	0.8	1.9	3.0	-1.1	-0.0999999

(a)

We got to estimate the level of unemployment on the levels of inflation. That is, run OLS with `inf` as response variable and `unem` as explanatory variable. The interpretation directly from this is that 1% increase in unemployment rate is related to 0.5024% increase in inflation rate. This positive relationship is weird and against economic theories. That is, we estimated the equation $inf_t = 1.0536 + 0.5024 unem_t$. Then, the natural rate of unemployment, μ_0 cannot be estimated using this regression.

```
##
## Call:
## glm(formula = inf ~ unem, data = PCurve)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
```

```
## -5.2176  -1.7812  -0.6659   1.1473   8.8795
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   1.0536     1.5480   0.681   0.4990
## unem          0.5024     0.2656   1.892   0.0639 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 8.82992)
##
##     Null deviance: 508.42  on 55  degrees of freedom
## Residual deviance: 476.82  on 54  degrees of freedom
## AIC: 284.86
##
## Number of Fisher Scoring iterations: 2
```

(b)

Run the Durbin-Watson test on the model from (a). We reject the null, meaning that autocorrelation is indeed present in the OLS model from part (a).

```
##
##  Durbin-Watson test
##
## data:  PCurve.a
## DW = 0.80148, p-value = 1.486e-07
## alternative hypothesis: true autocorrelation is greater than 0
```

(c)

Now, we estimate the adaptive expectations model using `cinf` and `unem`. Obtain the regression equation $\Delta inf_t = 2.8282 - 0.5176 unem_t$

```
##
## Call:
## glm(formula = cinf ~ unem, data = PCurve)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -9.0741  -0.9241   0.0189   0.8606   5.4800
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   2.8282     1.2249   2.309   0.0249 *
## unem         -0.5176     0.2090  -2.476   0.0165 *
```

```
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 5.321809)
##
##     Null deviance: 314.69  on 54  degrees of freedom
## Residual deviance: 282.06  on 53  degrees of freedom
##   (1 observation deleted due to missingness)
## AIC: 252
##
## Number of Fisher Scoring iterations: 2
```

Then, check for autocorrelation using Durbin-Watson. We cannot reject the null that autocorrelation exists, so there's no evidence that there's autocorrelation in these residuals.

```
##
##  Durbin-Watson test
##
## data:  PCurve.c
## DW = 1.771, p-value = 0.1673
## alternative hypothesis: true autocorrelation is greater than 0
```

For reference, the implied natural rate of unemployment is 5.464%.

```
2.8282/0.5176
```

```
## [1] 5.464065
```

(d)

Now, we estimate the adaptive expectations model using `cinf` and `cunem` based on the alternative model.Obtain the regression equation $\Delta inf_t = -0.07214 - 0.83281 \Delta unem_t$. In a, both intercept and slope are positive while in d they are both negative. The model in part a suggests that unemployment rate is positively related to inflation (at 10% significance), which seems to be against the theoretical framework that higher unemployment leads to lower inflation (deflationary pressure). d is essentially a regression on the first derivative, which suggests that increases in unemployment is related to decreased inflation. This would fit the theoretical framework.

```
##
## Call:
## glm(formula = cinf ~ cunem, data = PCurve)
##
## Deviance Residuals:
##     Min       1Q   Median       3Q      Max
## -7.4790  -0.9441   0.1384   1.0889   5.4551
##
## Coefficients:
```

```
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -0.07214    0.30584  -0.236  0.81443
## cunem       -0.83281    0.28984  -2.873  0.00583 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## (Dispersion parameter for gaussian family taken to be 5.137264)
##
##     Null deviance: 314.69  on 54  degrees of freedom
## Residual deviance: 272.28  on 53  degrees of freedom
##   (1 observation deleted due to missingness)
## AIC: 250.05
##
## Number of Fisher Scoring iterations: 2
```

Is there autocorrelation? No.

```
##
##  Durbin-Watson test
##
## data:  PCurve.d
## DW = 1.8414, p-value = 0.2747
## alternative hypothesis: true autocorrelation is greater than 0
```

(e)

Summary of the result as follow:

```
##
## Regression Results
## ===============================================
##                        Dependent variable:
##                   -----------------------------
##                      inf           cinf
##                      (1)       (2)       (3)
## -----------------------------------------------
## unem               0.502*   -0.518**
##                   (0.266)   (0.209)
##
## cunem                                -0.833***
##                                       (0.290)
##
## Constant            1.054    2.828**   -0.072
##                   (1.548)   (1.225)   (0.306)
##
## -----------------------------------------------
```

```
## Observations          56        55        55
## Log Likelihood    -140.430  -123.998  -123.027
## Akaike Inf. Crit.  284.861   251.996   250.055
## ===============================================
## Note:               *p<0.1; **p<0.05; ***p<0.01
```

A is obviously not accurate, because the residuals showed autocorrelation. The residuals in C & D have no such issues. Choosing between C or D require us to test whether or not a unit root is present in `unem`. If a unit root is present, then D is better. We run a Dickey-Fuller test with no lags to see whether we need a unit root.

Question 5

US Income and Australian exports: Exogeneity is unlikely to hold perfectly here because Australia and the United States trade with each other, so there's simultaneous causality. A decline in Australian exports could lead to Austrlians having less income, that reduces their imports from America, which in turn reduces the US income. US Income can be exogenous if we decide that this effect is not strong enough to significantly affect US income.

Oil prices and inflation: Exogeneity is most unlikely to hold here, since oil prices is set based on a bunch of economic indicators, including the inflation rate of oil-consuming countries as it relates to real purchasing power. It is most likely endogenous.

Monetary policy and inflation: They are most likely endogenous too. One of the things used to determine interest rate/target rate is inflation and expected inflation. Often, changes in monetary policy is based upon expectations of inflationary or deflationary pressure in the market. Therefore, they are endogenous.

Phillips Curve: The Phillips curve is a regression of the change in inflation on lagged changes in inflation and lagged unemployment rates. It is not exogenous because past unemployment is determined by past inflation, so there's again simultaneous causality.

In general, it is very difficult to find examples of distributed lag regressions in economics because economic data are often observed data from societal behaviour and not randomized data from a controlled experiment. As society is interconnected, it is very difficult to not have simultaneous causality.

Question 6

(a)

To test if lnSP500 = ln(SP500) and lnip = ln(ip) have unit root, we conduct the hypothesis test as follows. We conduct the test with both drift and trend to encompass all possibilities: H0: $\delta = 0$ H1: $\delta < 0$

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
```

```
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.120331 -0.018027  0.001964  0.020409  0.110882
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.010416   0.007296   1.428    0.154
## z.lag.1     -0.001284   0.001618  -0.794    0.428
## z.diff.lag1  0.264461   0.042682   6.196 1.14e-09 ***
## z.diff.lag2 -0.083192   0.044133  -1.885    0.060 .
## z.diff.lag3  0.016918   0.044073   0.384    0.701
## z.diff.lag4  0.043858   0.042547   1.031    0.303
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03265 on 547 degrees of freedom
## Multiple R-squared:  0.06896,    Adjusted R-squared:  0.06045
## F-statistic: 8.103 on 5 and 547 DF,  p-value: 2.173e-07
##
##
## Value of test-statistic is: -0.7936 5.6273
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.43 -2.86 -2.57
## phi1  6.43  4.59  3.78

##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
```

```
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.118961 -0.018376  0.002415  0.020044  0.110475
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.254e-02  1.721e-02   2.472   0.0137 *
## z.lag.1     -1.259e-02  5.717e-03  -2.201   0.0281 *
## tt           6.331e-05  3.073e-05   2.060   0.0398 *
## z.diff.lag1  2.684e-01  4.260e-02   6.302 6.08e-10 ***
## z.diff.lag2 -7.699e-02  4.411e-02  -1.746   0.0815 .
## z.diff.lag3  2.209e-02  4.401e-02   0.502   0.6160
## z.diff.lag4  5.093e-02  4.256e-02   1.197   0.2320
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03255 on 546 degrees of freedom
## Multiple R-squared:  0.07614,    Adjusted R-squared:  0.06599
## F-statistic:   7.5 on 6 and 546 DF,  p-value: 9.722e-08
##
##
## Value of test-statistic is: -2.2012 5.1887 2.4392
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.96 -3.41 -3.12
## phi2  6.09  4.68  4.03
## phi3  8.27  6.25  5.34

##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.042715 -0.005229 -0.000132  0.004683  0.067796
##
```

```
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0063813  0.0036230   1.761   0.0787 .
## z.lag.1     -0.0012191  0.0008888  -1.372   0.1707
## z.diff.lag1  0.3487046  0.0427129   8.164 2.25e-15 ***
## z.diff.lag2  0.0708172  0.0451230   1.569   0.1171
## z.diff.lag3  0.0643923  0.0451187   1.427   0.1541
## z.diff.lag4  0.0026028  0.0426664   0.061   0.9514
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01008 on 547 degrees of freedom
## Multiple R-squared:  0.1705, Adjusted R-squared:  0.1629
## F-statistic: 22.49 on 5 and 547 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -1.3717 5.9587
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.43 -2.86 -2.57
## phi1  6.43  4.59  3.78

##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression trend
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.044484 -0.005189 -0.000138  0.004644  0.067419
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  4.292e-02  1.620e-02   2.650  0.00829 **
## z.lag.1     -1.279e-02  5.077e-03  -2.519  0.01207 *
## tt           3.565e-05  1.541e-05   2.314  0.02104 *
## z.diff.lag1  3.505e-01  4.255e-02   8.236 1.32e-15 ***
## z.diff.lag2  7.684e-02  4.502e-02   1.707  0.08845 .
```

```
## z.diff.lag3  7.204e-02  4.506e-02   1.599  0.11046
## z.diff.lag4  1.531e-02  4.285e-02   0.357  0.72108
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.01004 on 546 degrees of freedom
## Multiple R-squared:  0.1786, Adjusted R-squared:  0.1695
## F-statistic: 19.78 on 6 and 546 DF,  p-value: < 2.2e-16
##
##
## Value of test-statistic is: -2.5186 5.7888 3.6253
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau3 -3.96 -3.41 -3.12
## phi2  6.09  4.68  4.03
## phi3  8.27  6.25  5.34
```

Based on the results, the ADF statistic for lnSP500 without trend is -0.7936 and t = -2.2012 with a trend.These all above the 5% critical value, which is -2.86 and -3.41 for without and with trend respectively. As for lnip, the ADF statistic without trend and wtih trend are -1.3717 and -2.5186 respectively. These are not closing to rejcting at 5% critical value. As a result, we fail to reject the null for both lnSP500 and lnip, and conclude that they both have unit root.

(b)

(b) Run a regression of lnSP500 on lnip.

```
##
## Call:
## lm(formula = l_sp500 ~ l_ip)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.6615 -0.2012 -0.0231  0.2236  0.5565
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) -2.40173    0.09549  -25.15   <2e-16 ***
## l_ip         1.69407    0.02354   71.97   <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2714 on 556 degrees of freedom
## Multiple R-squared:  0.9031, Adjusted R-squared:  0.9029
```

```
## F-statistic:  5179 on 1 and 556 DF,  p-value: < 2.2e-16
```

The t-statistic for lnip is 71.97. With such a small p-value, we can concludes that lnip has significant effect on lnSP500. In addition, the multiple and adjusted R-squared are both above 0.9, means that the goodness-of-fit is high.

(c)

To test whether the residual of the regression in part (b) is unit root. we conduct the hypothesis test as follows: H0: $\delta = 0$ H1: $\delta < 0$

```r
re <- residuals(regre)
cbdf <- adf.test(re, k=0)
summary(ur.df(re, type = "drift", lags = 2, selectlags = "Fixed"))
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.137993 -0.021211  0.000479  0.023226  0.134036
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.001012   0.001556   0.651    0.516
## z.lag.1     -0.009133   0.005807  -1.573    0.116
## z.diff.lag1  0.282832   0.042388   6.672 6.14e-11 ***
## z.diff.lag2 -0.064498   0.042535  -1.516    0.130
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03662 on 551 degrees of freedom
## Multiple R-squared:  0.07682,    Adjusted R-squared:  0.07179
## F-statistic: 15.28 on 3 and 551 DF,  p-value: 1.452e-09
##
##
## Value of test-statistic is: -1.5726 1.456
##
```

```
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.43 -2.86 -2.57
## phi1  6.43  4.59  3.78
```

```
cbdf
```

```
##
##  Augmented Dickey-Fuller Test
##
## data:  re
## Dickey-Fuller = -1.1512, Lag order = 0, p-value = 0.9143
## alternative hypothesis: stationary
```

The t-statistic for DF and ADF are -1.1512 and -1.5726 respectively, which is all above the 5% critical value (-2.86). Therefore, we fail to reject the null and conclude that the residual of the regression in part(b) is unit root.

(d)

We add a time trend to the regression and redo part (c) based on part(b).

```
##
## Call:
## lm(formula = l_sp500 ~ t + l_ip, data = S_P)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -0.5159 -0.2180  0.0561  0.2309  0.3767
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 1.8737656  0.3835509   4.885 1.35e-06 ***
## t           0.0041557  0.0003633  11.438  < 2e-16 ***
## l_ip        0.3438902  0.1199295   2.867   0.0043 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2443 on 555 degrees of freedom
## Multiple R-squared:  0.9215, Adjusted R-squared:  0.9213
## F-statistic:  3260 on 2 and 555 DF,  p-value: < 2.2e-16
```

```
##
## ###############################################
## # Augmented Dickey-Fuller Test Unit Root Test #
## ###############################################
##
```

```
## Test regression drift
##
##
## Call:
## lm(formula = z.diff ~ z.lag.1 + 1 + z.diff.lag)
##
## Residuals:
##       Min        1Q    Median        3Q       Max
## -0.121321 -0.019124  0.002408  0.020734  0.114141
##
## Coefficients:
##               Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.0008073  0.0013858   0.583   0.5604
## z.lag.1     -0.0107410  0.0057237  -1.877   0.0611 .
## z.diff.lag1  0.2681779  0.0422459   6.348 4.56e-10 ***
## z.diff.lag2 -0.0806177  0.0423367  -1.904   0.0574 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.03263 on 551 degrees of freedom
## Multiple R-squared:  0.07251,    Adjusted R-squared:  0.06746
## F-statistic: 14.36 on 3 and 551 DF,  p-value: 5.094e-09
##
##
## Value of test-statistic is: -1.8766 1.9338
##
## Critical values for test statistics:
##       1pct  5pct 10pct
## tau2 -3.43 -2.86 -2.57
## phi1  6.43  4.59  3.78

##
##  Augmented Dickey-Fuller Test
##
## data:  dre
## Dickey-Fuller = -1.5409, Lag order = 0, p-value = 0.7726
## alternative hypothesis: stationary
```

After adding the time trend to regression, the t-statistic for DF and ADF become -1.5409 and -1.8766 respectively, which is still all above the 5% critical value (-2.86). Therefore, we fail to reject the null and conclude that the residual of the regression in part(b) is unit root. In other words, lnSP500 and lnip are not cointegrated even with a time trend.

(e)

The results above shows that there is no cointegration of lnSP500 and lnip even with a linear time trend. Therefore, we conclude that there is no long time equilibrium relationship between stock market and economic activity.

Question 7

(a)

If money and government expenditures are exogenous, then we can use a distributed lag model to estimate the multipliers. The coefficients on the money variables represent the cumulative dynamic monetary multiplier for that period, and the coefficients on the fiscal variables represent the cumulative dynamic fiscal multipliers for that period. They are the cumulative multipliers, and not the dynamic multipliers because the regression was conducted on first differences.

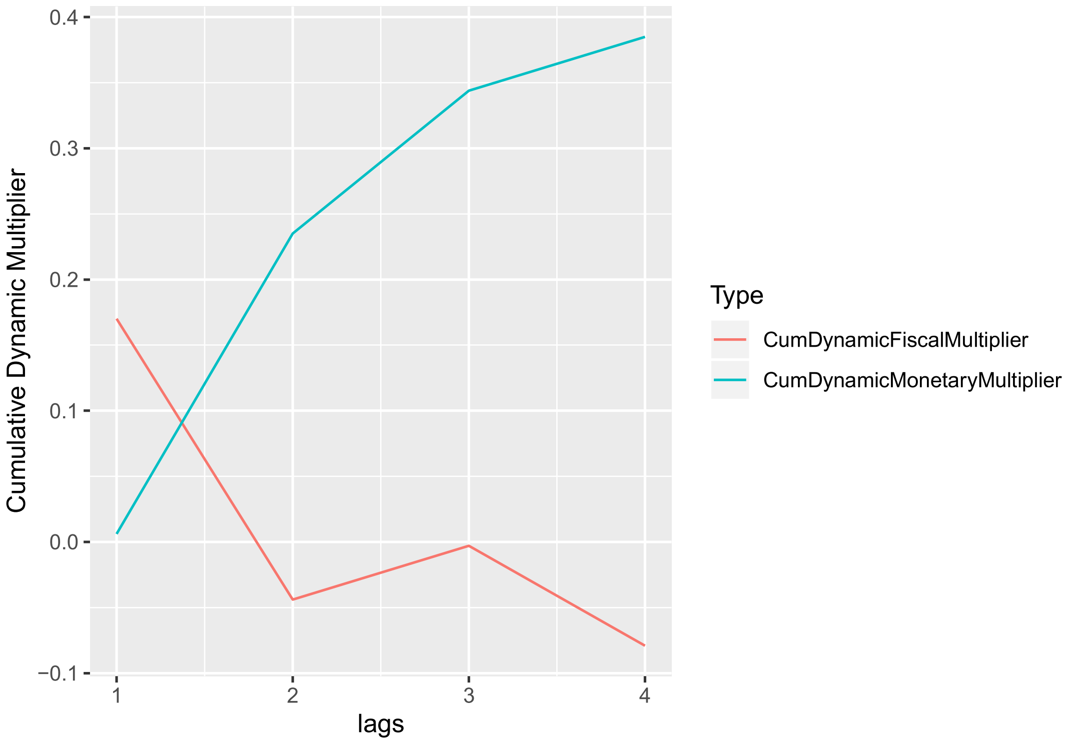
The long-run cumulative dynamic monetary multiplier is 0.425, and 0.018 is the long-run cumulative dynamic fiscal multiplier.

lags	CumDynamicMonetaryMultiplier	CumDynamicFiscalMultiplier	Total
1	0.006	0.170	0.176
2	0.235	-0.044	0.191
3	0.344	-0.003	0.341
4	0.385	-0.079	0.306

We can use a t-test to test for the significance of coefficients of these multipliers using HAC standard errors. Rejection criteria is the same as usual.

(b)

We plot from the data earlier.



(c)

It is unlikely to be the case because both monetary and fiscal policy decisions take into account a number of the same factors such as current and future expected output growth, therefore making them endogenous.