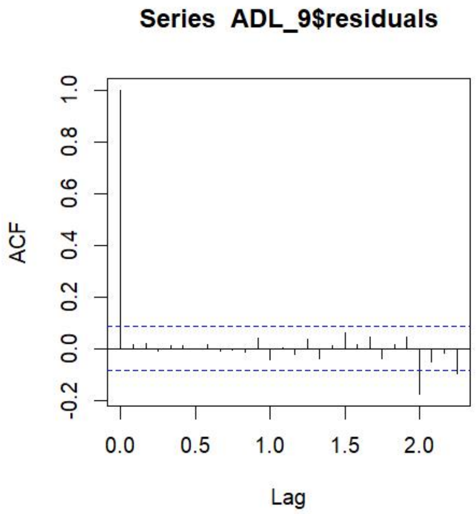
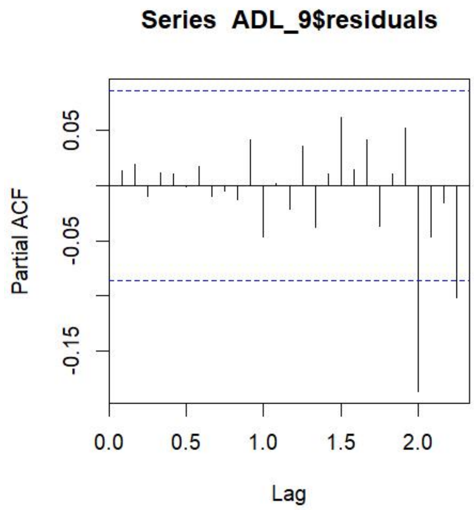
1.

```
> aic_diffADL <- numeric()
> for (i in 1:15) {
+   adl_model <- dynlm(diff_Yt ~ lag(diff_Yt, i)
+                      + diff_OIL_P + lag(diff_OIL_P, i)
+                      + diff_OIL_PROD + lag(diff_OIL_PROD, i)
+                      + diff_OIL_STOCKS + lag(diff_OIL_STOCKS,
 i)
+                      + diff_WORLD_IP + lag(diff_WORLD_IP, i)
+                      + diff_US_CPI+ lag(diff_US_CPI, i)
+                      , data = data)
+   aic_diffADL[i] <- AIC(adl_model)
+ }

> best_AICorder <- which.min(aic_diffADL)
> print(best_AICorder)

[1] 9
```

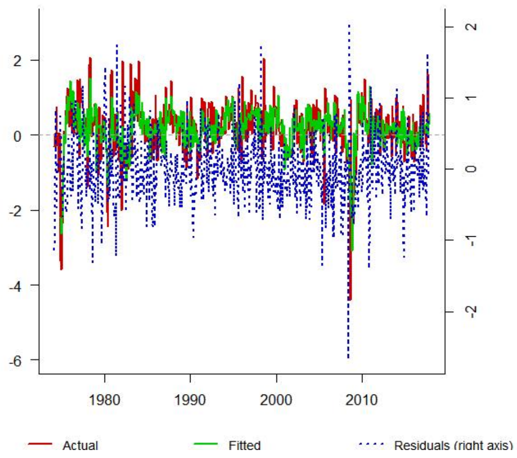
```
> Bic_diffADL <- numeric()
> for (i in 1:15) {
+   adl_model <- dynlm(diff_Yt ~ lag(diff_Yt, i)
+                      + diff_OIL_P + lag(diff_OIL_P, i)
+                      + diff_OIL_PROD + lag(diff_OIL_PROD, i)
+                      + diff_OIL_STOCKS + lag(diff_OIL_STOCKS,
 i)
+                      + diff_WORLD_IP + lag(diff_WORLD_IP, i)
+                      + diff_US_CPI+ lag(diff_US_CPI, i)
+                      , data = data)
+   Bic_diffADL[i] <- BIC(adl_model)
+ }

> best_BICorder <- which.min(Bic_diffADL)
> print(best_BICorder)

[1] 9
```

```
stargazer(ADL_9,type = "text")

====================================================
                            Dependent variable:
                        ----------------------------
                                  diff_Yt
----------------------------------------------------
.....

----------------------------------------------------
Observations                        518
R2                                 0.532
Adjusted R2                        0.472
Residual Std. Error          0.526 (df = 458)
F Statistic               8.823*** (df = 59; 458)
====================================================
Note:                    *p<0.1; **p<0.05; ***p<0.01
```






```
> LjungBox(ADL_9$residuals,90,59)

1    1    0.0931    NA
2    2    0.2966    NA
3    3    0.3427    NA
4    4    0.4115    NA
5    5    0.4730    NA
6    6    0.4731    NA
7    7    0.6307    NA
8    8    0.6791    NA
9    9    0.6891    NA
10   10   0.7780    NA
11   11   1.7067    NA
12   12   2.7902    NA
13   13   2.7928    NA
14   14   3.1070    NA
15   15   3.8005    NA
16   16   4.5570    NA
17   17   4.6039    NA
18   18   6.6305    NA
19   19   6.7634    NA
20   20   7.8589    NA
21   21   8.6177    NA
22   22   8.7558    NA
23   23   9.9298    NA
24   24   26.8169   NA
25   25   28.2201   NA
26   26   28.4317   NA
27   27   33.8216   NA
28   28   35.7241   NA
29   29   41.3947   NA
30   30   42.8718   NA
31   31   42.9850   NA
32   32   43.8020   NA
33   33   43.8723   NA
34   34   43.9069   NA
35   35   43.9092   NA
36   36   44.6362   NA
37   37   44.7052   NA
38   38   45.0020   NA
39   39   45.3004   NA
40   40   46.1999   NA
41   41   46.2028   NA
42   42   47.0674   NA
43   43   47.2669   NA
44   44   47.2743   NA
45   45   49.3025   NA
46   46   50.1197   NA
47   47   50.2510   NA
48   48   51.1866   NA
49   49   52.1228   NA
50   50   52.2342   NA
51   51   53.0309   NA
52   52   53.5732   NA
53   53   53.6836   NA
54   54   58.8572   NA
55   55   58.9393   NA
56   56   60.3026   NA
57   57   61.6636   NA
58   58   63.1226   NA
59   59   64.0594   NA
60   60   64.3398   0.3272
61   61   64.3539   0.3600
62   62   64.9216   0.3752
63   63   65.1769   0.4009
64   64   66.1696   0.4019
65   65   66.6154   0.4212
66   66   66.9905   0.4428
67   67   67.4130   0.4629
68   68   67.6358   0.4896
69   69   68.5380   0.4931
70   70   68.6327   0.5239
71   71   68.8658   0.5497
72   72   73.7267   0.4214
73   73   74.0893   0.4424
74   74   74.1404   0.4735
75   75   74.1825   0.5050
76   76   75.9113   0.4813
77   77   75.9221   0.5133
78   78   77.2344   0.5032
79   79   78.5154   0.4942
80   80   78.5307   0.5255
```



Can you do standard inference based on the estimation results?

No I cannot do standard inference, because now the model is ADL in first difference and I didnt test if there are cointegration relationshps. The model would be misspecification if there is cointegration relationship which will make the estimation inconsistent and not asymtotically normal, thus standard inference is not possible.

Compute the dynamic response of the growth rate of U.S. industrial production to a 10% increase in the oil price.

From the IRF_calc, it can calculate two months dynamic impact of the 10% increase in the oil price. Pay attention, the 10% is permanent increase in oil price and transitory in oil price in first difference (and it is 10 unit shock).

Outline possible disadvantages of the ADL model you have estimated.

First, it used all the variables in first differences without testing for the cointegrations. if there is cointegrations, the model will be misspecified and inconsistent, and the estimations is not asymptotically normal, standard inference is not possible,

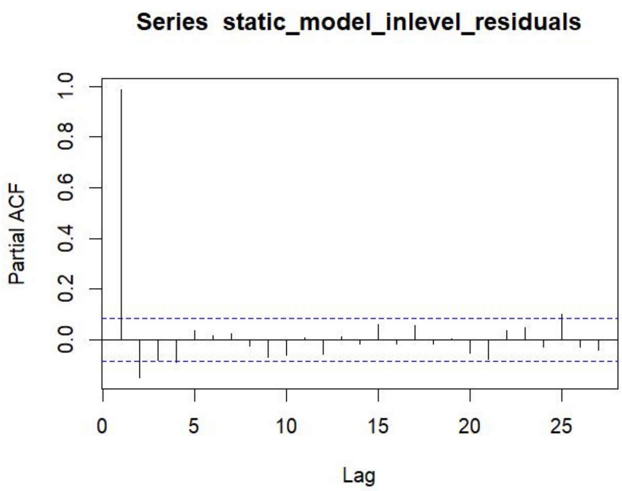
Second, it ignore the endogenity, which will cause the endogenous problem in the model, the error term and the explanatory variables will be correlated and the model is inconsistent. Also the endogenous problem will make the casual inference incorrectly identified.

In addition, all the variables choose the same lag, in fact this is not possible and there will be overfitting problem in the specification.

2.

```
> static_model_inlevel <- lm(Yt ~ OIL_P + OIL_PROD + OIL_STOCK
S
+                           +  WORLD_IP + US_CPI)

> stargazer(static_model_inlevel, type = "text")
```

```
===============================================
                        Dependent variable:
                    ---------------------------
                                Yt
-----------------------------------------------
OIL_P                       -0.069***
                             (0.005)

OIL_PROD                      0.094
                             (0.070)

OIL_STOCKS                   -0.023
                             (0.032)

WORLD_IP                    0.689***
                             (0.054)

US_CPI                        0.093
                             (0.057)

Constant                     22.971
                            (70.752)

-----------------------------------------------
Observations                   528
R2                            0.972
Adjusted R2                   0.972
Residual Std. Error      4.856 (df = 522)
F Statistic         3,669.979*** (df = 5; 522)
===============================================
Note:               *p<0.1; **p<0.05; ***p<0.01
```



```
> AR2_residuals_static = arima(static_model_inlevel_residuals,
+                             order=c(2,0,0), include.mean = TRUE)

> AR4_residuals_static = arima(static_model_inlevel_residuals,
+                             order=c(4,0,0), include.mean = TRUE)

> LR_test_residuals_AR2_AR4 <- lrtest(AR2_residuals_static, AR
4_residuals_static)> print(LR_test_residuals_AR2_AR4)

                       Likelihood ratio test

Model 1: arima(x = static_model_inlevel_residuals, order = c(2,
 0, 0),
    include.mean = TRUE)
Model 2: arima(x = static_model_inlevel_residuals, order = c(4,
 0, 0),
    include.mean = TRUE)
  #Df  LogLik Df  Chisq Pr(>Chisq)
1   4 -592.88
2   6 -589.93  2 5.8936    0.05251 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1



> LjungBox(AR2_residuals_static$residuals, 25, 2)

lag Ljung-Box p-value
1    1    0.0622      NA
2    2    2.9459      NA
3    3    6.1464  0.1047
4    4    6.6499  0.1556
5    5    6.7265  0.2418
6    6    7.2505  0.2983
7    7    7.2570  0.4026
8    8    9.8638  0.2747
9    9   10.5997  0.3042
10  10   10.5997  0.3895
11  11   12.9352  0.2976
12  12   12.9416  0.3733
13  13   13.3746  0.4193
14  14   14.0105  0.4489
15  15   14.0460  0.5220
16  16   16.9694  0.3876
17  17   16.9738  0.4561
18  18   17.0937  0.5167
19  19   17.4898  0.5567
20  20   20.9262  0.4015
21  21   21.8423  0.4086
22  22   23.6081  0.3681
23  23   24.8282  0.3592
24  24   32.5530  0.1139
25  25   32.5537  0.1426



> ADF1_cointegration = ur.df(static_model_inlevel_residuals, l
ags = 1, type = "none")

> summary(ADF1_cointegration)
```

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

Test regression none


Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min      1Q  Median      3Q     Max
-4.0737 -0.4240  0.0336  0.4606  2.8265

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1    -0.012811   0.006743  -1.900 0.058001 .
z.diff.lag  0.154756   0.043249   3.578 0.000378 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.7439 on 524 degrees of freedom
Multiple R-squared:  0.02841,     Adjusted R-squared:  0.0247
F-statistic:  7.66 on 2 and 524 DF,  p-value: 0.0005258


Value of test-statistic is: -1.8999

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.58 -1.95 -1.62
```

Test for cointegration (taking the order of integration of the variables into account).

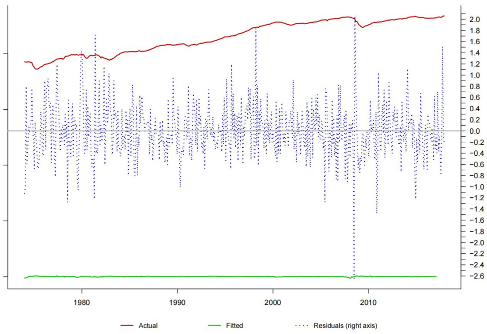
From the ADF test on residuals, the ADF test value is -1.8999 and the critical value by MacKi nnon is -3.7429 - 8.352/528 - 13.41/528 = -3.784116 (variables=3, no trend, T=528). so we ac cept the H0 of the error term is not I(0) and there is no cointegration. and this may because th e 6 variables may have different orders of integration.

Can you do standard inference using the estimation r esults ?

No, there is no cointegration, there will be spurious regression in the model which bring incon sistent result, and it is not asymptotically normal, the standard inference is not possible.

3.

```
> staticResid  = ts(static_model_inlevel$residuals, start = 19
74, frequency = 12)

> ECMmodel   = dynlm(diff_Yt ~ L(diff_Yt,1:9)
+                    + diff_OIL_P  + L(diff_OIL_P,1:9)
+                    + diff_OIL_PROD   + L(diff_OIL_PROD,1:9)
+                    + diff_OIL_STOCKS + L(diff_OIL_STOCKS,1:9)
+                    + diff_WORLD_IP   + L(diff_WORLD_IP,1:9)
+                    + diff_US_CPI     + L(diff_US_CPI,1:9)
+                    + L(staticResid,1) )
```




```
> LjungBox(ECMresid, 40,12)
```

1	1	0.0773	NA	21	21	8.9312	0.9897
2	2	0.2580	NA	22	22	9.1004	0.9928
3	3	0.3135	NA	23	23	10.3545	0.9889
4	4	0.3548	NA	24	24	26.9931	0.3048
5	5	0.3941	NA	25	25	28.3217	0.2933
6	6	0.3967	NA	26	26	28.5413	0.3324
7	7	0.5002	NA	27	27	33.9127	0.1686
8	8	0.5969	NA	28	28	35.8945	0.1453
9	9	0.6613	NA	29	29	41.8512	0.0579
10	10	0.7327	NA	30	30	43.5425	0.0524
11	11	1.6658	NA	31	31	43.6407	0.0655
12	12	2.7088	NA	32	32	44.5164	0.0697
13	13	2.7169	0.9988	33	33	44.6104	0.0855
14	14	2.9556	0.9991	34	34	44.6591	0.1044
15	15	3.7468	0.9985	35	35	44.6593	0.1270
16	16	4.4214	0.9980	36	36	45.3428	0.1367
17	17	4.4850	0.9989	37	37	45.4711	0.1600
18	18	6.8264	0.9915	38	38	45.8681	0.1782
19	19	7.0386	0.9940	39	39	46.2124	0.1989
20	20	8.2491	0.9901	40	40	47.0401	0.2065

```
> stargazer(ECMmodel, type = "text")

====================================================
                            Dependent variable:
                         ---------------------------
                                  diff_Yt
----------------------------------------------------
.....

L(staticResiduals, 1)            -0.014***
                                  (0.005)

Constant                          -0.030
                                  (0.050)

----------------------------------------------------
Observations                        518
R2                                 0.539
Adjusted R2                        0.478
Residual Std. Error          0.522 (df = 457)
F Statistic               8.898*** (df = 60; 457)
====================================================
Note:                    *p<0.1; **p<0.05; ***p<0.01
```

Test for cointegration in this ECM and explain in what respect this test differs from the one under question 2 ?

From the cointegration test in ECM, we accept of H0 of a=0, there is no cointegration. (we can not use standard critical value because this is not a normal distribution **under H0**), so we use Mackinnon critical value as a proxy.

The differences is, by this way, we **test the significance of the coefficient of the error correction term to see if it exists.** And if it is significant, then there is cointegration relationship. And **the Engle Granger two step approach is conducted by the ADF test on the residuals to see if it is I(0).** if it is I(0), there is cointegration.

And by this way, we have more power than ADF test to reject there is no cointegration. And we can also avoid the common factor restrictions on the ECM.

4.

```
> ADL_10 = dynlm(Yt ~ L(Yt,1:10)
+                + OIL_P + L(OIL_P,1:10)
+                + OIL_PROD + L(OIL_PROD,1:10)
+                + OIL_STOCKS +L(OIL_STOCKS, 1:10)
+                + WORLD_IP+ L(WORLD_IP,1:10)
+                + US_CPI + L(US_CPI,1:10)+ )
```

```
> stargazer(ADL_10,type = "text")

==================================================
                          Dependent variable:
                      ----------------------------
                                   Yt
--------------------------------------------------
...
--------------------------------------------------
Observations                      518
R2                               1.000
Adjusted R2                      1.000
Residual Std. Error        0.514 (df = 452)
F Statistic           24,694.970*** (df = 65; 452)
==================================================
Note:                  *p<0.1; **p<0.05; ***p<0.01



> LjungBox(ADL_10$residuals, 100,65)


1     1    0.0153      NA
...

65   65   67.4197      NA
66   66   67.4837  0.4262
67   67   68.1374  0.4383
68   68   68.3513  0.4652
69   69   69.5975  0.4572
70   70   69.6344  0.4898
71   71   69.9234  0.5139
72   72   73.8817  0.4165
73   73   73.9802  0.4459
74   74   73.9804  0.4788
75   75   74.1088  0.5074
76   76   75.6316  0.4903
77   77   75.6873  0.5210
78   78   76.7854  0.5177
79   79   77.5475  0.5252
80   80   77.5687  0.5562
81   81   78.9699  0.5431
82   82   79.4866  0.5581
83   83   80.6818  0.5516
84   84   82.4574  0.5272
85   85   82.5150  0.5561
86   86   82.5477  0.5854
87   87   82.5979  0.6136
88   88   89.3336  0.4403
89   89   91.2048  0.4153
90   90   91.2052  0.4447
91   91   92.6792  0.4313
92   92   92.7865  0.4574
93   93   94.2175  0.4452
94   94   94.9859  0.4521
95   95   94.9893  0.4810
96   96   95.9079  0.4835
97   97   95.9093  0.5122
98   98   98.6600  0.4623
99   99   98.6628  0.4907
100 100   98.7161  0.5175
```

Is this a valid approach ?

Yes, it is valid. In principle, we should use the ADL model in first difference if the variables is not stationary and there is no cointegration. However, when conducting cointegration test. use the residuals from static model will ignore the dynamics which will lead a biased and not even asymptotically normal distribution result, also there may be less power to reject the H0 of no cointegration. **so, there may be type II error. And if incorrectly accept the H0 of no cointegration and use the model in first difference, the model will be misspecified and inconsistent.**

and by estimating the ADL model in level, the result will be biased but consistent, because the real DGP will be nested in our model. so this is also a valid approach.

Can you do standard inference based on the estimation results ?

No, because there may be type II error in our cointegration test, we are not sure if there is error correction term in **real data generating process,** so in general,the standard inference is not possible, unless you know about whether there is cointegration.

What is the link between this ADL model in levels, the ADL in first differences (question 1) and the error correction model (question 3) ?

The ADL in levels can be reparameterization to the error correction model which equals the ADL in first differences plus error correction term.

5.

```
> ECMcoint = nls(formula, data = ECMdata, start = list(theta1
= theta1init,
+ delta0 = delta0init, delta1 = delta1init, kappa0 = kappa0ini
t, kappa1 = kappa1init,
+ Oilstock0 = Oilstock0init, Oilstock1 = Oilstock1init,
+ worldip0 = worldip0init, worldip1 = worldip1init, uscpi0 = u
scpi0init, uscpi1 = uscpi1init,
+ alpha = alphainit, + beta0 = beta0init, beta1 = beta1init, b
eta2 = beta2init, beta3 = beta3init,+ beta4 = beta4init, beta5
 = beta5init))

Error in nls(formula, data = ECMdata, start = list(theta1 = theta1init, :
  Singular gradient
```

```
> formula  = diff_Yt ~ theta1*diff_Yt_1 + delta0*diff_OIL_P +
delta1*diff_OIL_P_1 ++   kappa0*diff_OIL_PROD + kappa1*diff_OI
L_PROD_1 + +   alpha*( Yt_1 - beta0 - beta1*OIL_P_1 - beta2*OI
L_PROD_1 )>

> ECMcoint = nls(formula, data = ECMdata, start = list(theta1
= theta1init, + delta0 = delta0init, delta1 = delta1init, kapp
a0 = kappa0init, kappa1 = kappa1init,+ alpha = alphainit, + be
ta0 = beta0init, beta1 = beta1init, beta2 = beta2init))

> summary(ECMcoint)

Formula: diff_Yt ~ theta1 * diff_Yt_1 + delta0 * diff_OIL_P +
delta1 *
    diff_OIL_P_1 + kappa0 * diff_OIL_PROD + kappa1 * diff_OIL_
PROD_1 +
    alpha * (Yt_1 - beta0 - beta1 * OIL_P_1 - beta2 * OIL_PROD
_1)

Parameters:
         Estimate Std. Error t value Pr(>|t|)
theta1  3.287e-01  4.126e-02   7.966 1.05e-14 ***
delta0  1.667e-03  3.963e-03   0.421   0.6742
delta1  6.731e-03  3.966e-03   1.697   0.0903 .
kappa0  8.395e-02  1.922e-02   4.368 1.51e-05 ***
kappa1  3.759e-03  1.959e-02   0.192   0.8479
alpha  -3.184e-03  2.750e-03  -1.158   0.2474
beta0  -1.317e+03  9.494e+02  -1.387   0.1660
beta1  -5.194e-01  4.197e-01  -1.238   0.2165
beta2   1.750e+00  8.028e-01   2.179   0.0298 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.6622 on 517 degrees of freedom

Number of iterations to convergence: 3
Achieved convergence tolerance: 4.749e-08
  (Three observations have been deleted)
```

Test for cointegration in this ECM.

The t-statistic value of alpha is -1.158, compared to the MacKinnon critical value as a proxy o f the real distribution. it accept H0 of there is no cointegration.

Of course there still can be type II error, because **when firstly using the nonlinear least squa re method to set the result of the ECM from the static OLS regression as the initial value to estimate this error correction model of order(9) with full variables,** there is singular gr adient in the coefficient matrix. And this is very likely because of the multicolinearity in such a complex model. So I have to reduce the lag and even the variables to avoid multicolinearity. But by doing that, I omitted some dynamics and the error term will be autocorrelated.

What are the advantages of this approach ?

By this one step approach, first, it have more powers than the ADF test to reject the H0.

Second, it doesn't ignore the dynamics so because used the cointegrating vector within the E CM, **so if there is cointegration, it can do standard inference on the coefficients of the lev el variables (beta) because now it is unbiased and normal distributed.**

Third, by testing within the ECM, it does not impose common factor restrictions on the ECM.

Can you do standard inference based on the estimatio n results ?

Yes, because **by ECM it will give me biased but consistent result, standard inference on t he coefficient of the differenced variables is appropriate for it is stationary and asymptot ically normal.** And now the alpha is not significant compared to its critical value, and there i s **no cointegration, so the error correction term can be dropped out, standard inference on the level variables within the error correction term is not appropriate but is relevant now. So the standard inference is appropriate.**