Macro-econometrics: Notes on Week 4 Exercise

Julia Giese
JULIA.GIESE@NUFFIELD.OX.AC.UK

- These notes will be posted on:
  `http://www.nuff.ox.ac.uk/Users/giese/Teaching.htm`

Notes on suggested solutions for Question 2:[1]

a) Consider the graphical properties of $cons$, inc and $inflat$.

The data are plotted in levels ($cons$, inc and $inflat$) in Figure 1, and in first differences ($\Delta cons$, Δinc and $\Delta inflat$) in Figure 2. Further plots are the ACF and PACF in Figure 3, and the densities and scatter plots in Figure 4. Given the lower case data descriptions, it seems safe to assume these data are already in log form.

Looking firstly at the levels in Figure 1, one can see that $cons$ and inc could not be described as possessing a constant mean throughout the sample; both have a much higher mean before 1974, and a lower mean since 1976. There appears to be some kind of structural break between 1974 and 1976. The bottom panel, $inflat$ gives some hint; around this same period is a large increase and then fall in the inflation rate. It appears also in this series there is a different mean after 1976 than before 1974; before it is about 0.0125, but after it is around 0.025.

So it might be possible to conclude that the first two series are stationary around a structural break in 1974-76, with the exception of $inflat$, which appears to be stationary albeit with slow mean reversion.

Considering now the first differences, from Figure 2 both Δinc and $\Delta cons$ look stationary; their variances look reasonably constant, while their mean appears also constant over the series. There is perhaps less certainly over the stationarity of the $\Delta inflat$ series; the 1970s sees a number of long and large deviations from mean. However, before and after the 1970s, the series appears stationary.

A further graphical inspection that will help determine appropriate lag length of any models of these variables is looking at the ACF plots and PACF plots. These are given in Figure 3. They suggest that all three series have an autoregressive structure, and that $cons$ and inc would require one lag, but $inflat$ three. Furthermore, the ACF shows very high persistence for $cons$ and inc, but not quite as much for $inflat$.

The densities plotted in Figure 4 show that both $cons$ and inc display bi-modality. This suggests that there is a structural break, as the process gathers around the two different means seen in Figure 1. This problem is not seen in the $inflat$ series. Often, taking a transformation of the data renders it more ‘normal’, and indeed this

---

[1] Many thanks to James Reade for providing notes from last year.

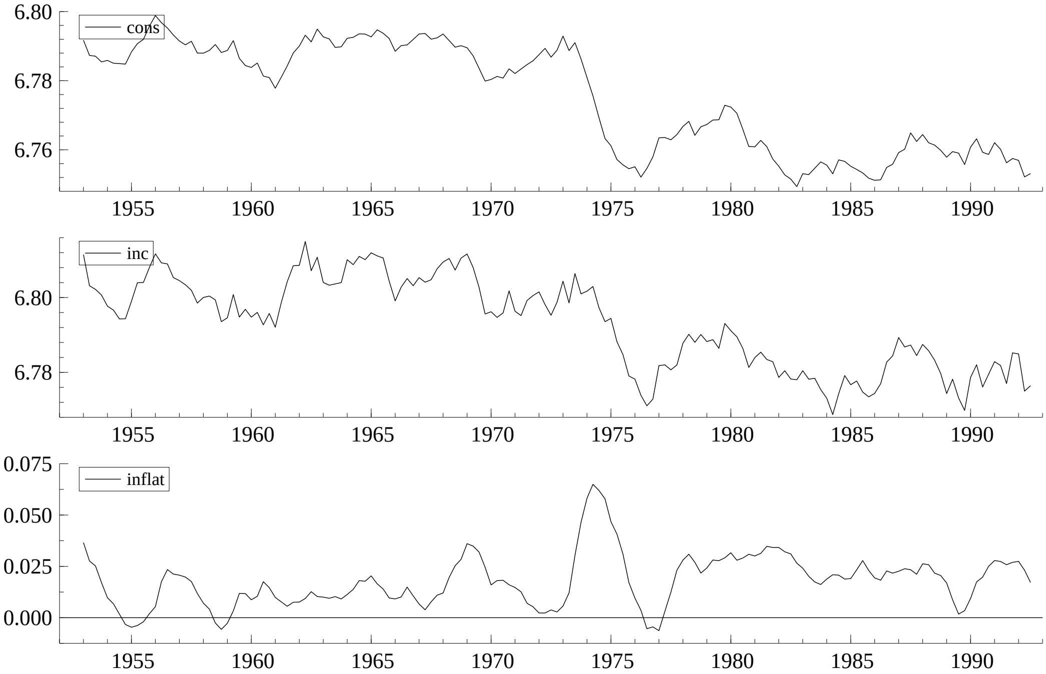

Figure 1: Plot of the data in levels

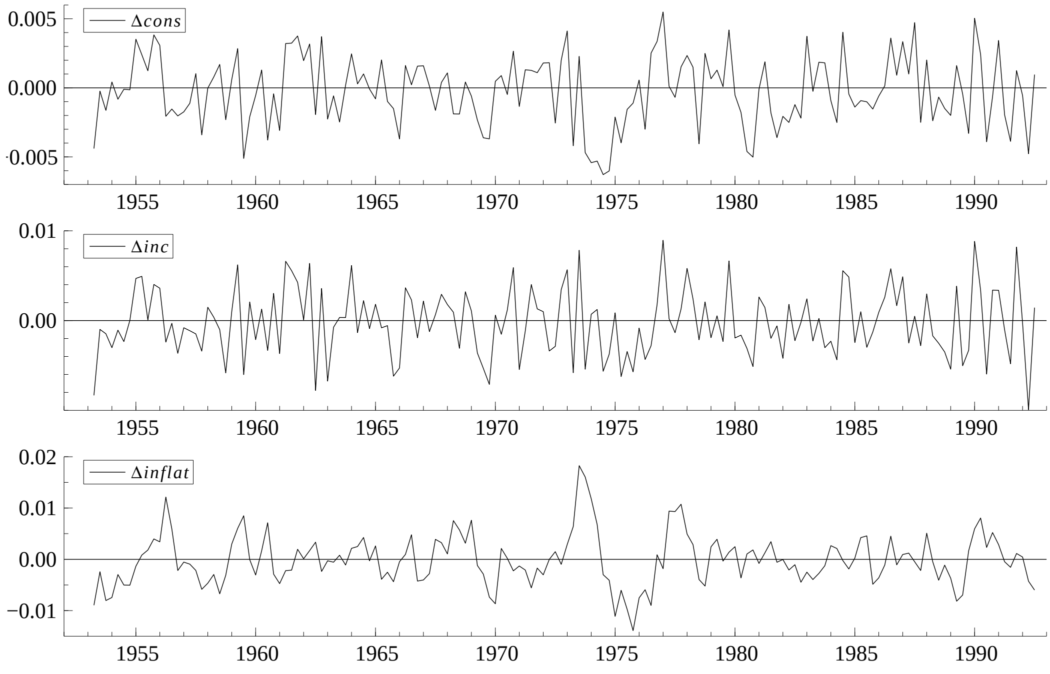

Figure 2: Plot of the data in first differences

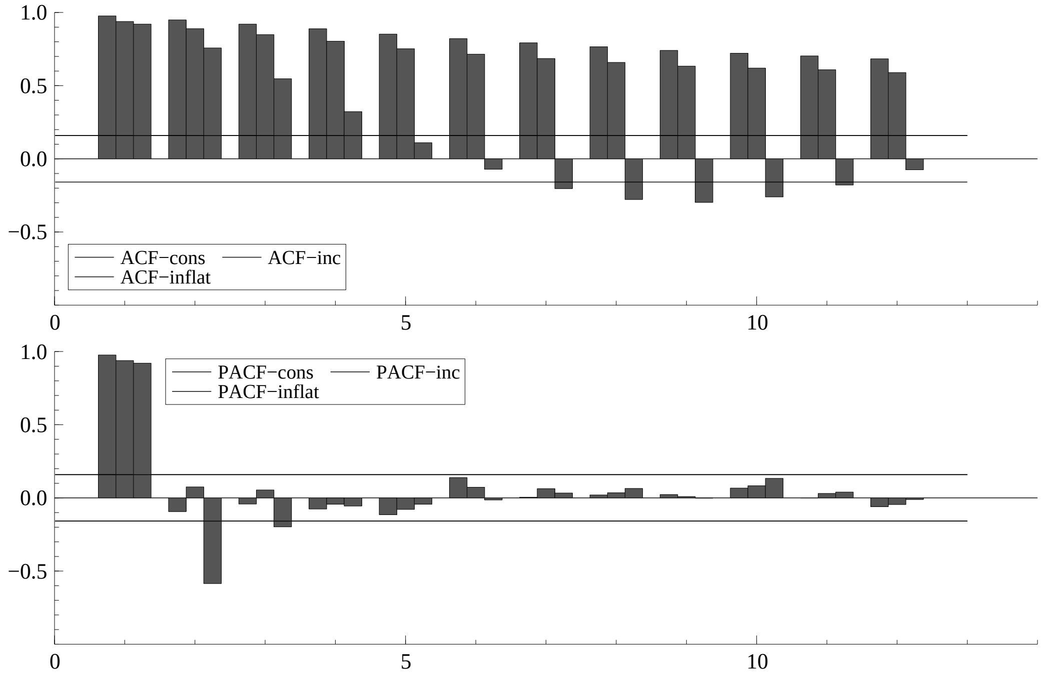

Figure 3: Plots of the autocorrelation function (ACF) and partial autocorrelation function (PACF) of the series in levels.

is the case here: taking first differences, as seen again from Figure 4 in the bottom three panels on the left hand side, renders *cons* and *inc* looking normal.

Scatter plots of the three series against each other shows that there is strong correlation, easily visible in Figure 1, between *cons* and *inc*, but not between either of these series and $inflat$.

b) Undertake unit root tests on the variables.

Testing on $cons$, inc and $inflat$, given the potential description of the data as unit root with drift, it seems sensible to carry out a Dickey-Fuller unit root test with

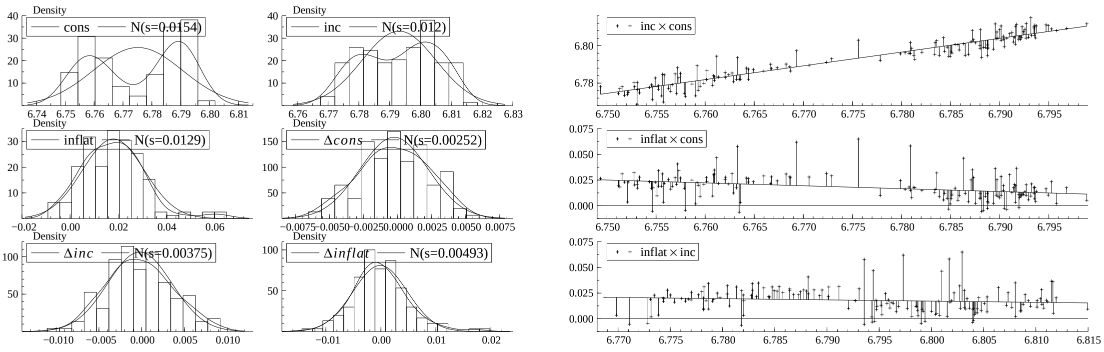

Figure 4: Plots of the densities of the three data series, and the cross correlations between the series (all in levels).

constant in PcGive. The models will be:

$$\begin{aligned}
\Delta cons_t &= \mu_{cons} + \gamma_{cons} cons_{t-1} + \sum_{i=0}^{D} \phi_{cons,i} \Delta cons_{t-i} + \epsilon_{cons,t}, \\
\Delta inc_t &= \mu_{inc} + \gamma_{inc} inc_{t-1} + \sum_{i=0}^{D} \phi_{inc,i} \Delta inc_{t-i} + \epsilon_{inc,t}, \\
\Delta inflat_t &= \mu_{inflat} + \gamma_{inflat} inflat_{t-1} + \sum_{i=0}^{D} \phi_{inflat,i} \Delta inflat_{t-i} + \epsilon_{inflat,t}.
\end{aligned}$$

The two tests will be $H_0 : \gamma_{cons} = 0$, $H_0 : \gamma_{inc} = 0$, and $H_0 : \gamma_{inflat} = 0$. The PcGive output for this test is given below, where up to three lags are used, since we noted in the previous section that $inflat$ requires three lags. While three lags aren't necessary for the other two series, nevertheless for simplicity this output is given, since with 157 observations, losing two is not likely to be harmful:

```
Unit-root tests (using DATA.in7) The sample is 1954 (1) - 1992 (3)

cons: ADF tests (T=155, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf       beta Y_1     sigma   t-DY_lag  t-prob       AIC  F-prob
  3     -1.266        0.98341 0.002466       1.404  0.1623    -11.98
  2     -1.110        0.98550 0.002474       1.694  0.0924    -11.98  0.1623
  1    -0.9238        0.98793 0.002490       2.481  0.0142    -11.97  0.0916
  0    -0.6348        0.99163 0.002531                        -11.95  0.0130

inc: ADF tests (T=155, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf       beta Y_1     sigma   t-DY_lag  t-prob       AIC  F-prob
  3     -1.682        0.95642 0.003737      0.9126  0.3629    -11.15
  2     -1.581        0.95939 0.003734    0.006880  0.9945    -11.15  0.3629
  1     -1.601        0.95942 0.003722     -0.7453  0.4573    -11.17  0.6601
  0     -1.759        0.95615 0.003717                        -11.18  0.7097

inflat: ADF tests (T=155, Constant; 5%=-2.88 1%=-3.47)
D-lag    t-adf       beta Y_1     sigma   t-DY_lag  t-prob       AIC  F-prob
  3     -5.400**      0.85702 0.003516       1.177  0.2410    -11.27
  2     -5.364**      0.86932 0.003521       1.062  0.2900    -11.27  0.2410
  1     -5.389**      0.87969 0.003522       11.61  0.0000    -11.28  0.2872
  0     -2.381        0.92852 0.004822                        -10.66  0.0000
```

The PACF plots in Figure 3 suggested an AR(1) model `D-lag=0` for $cons$ and inc, and an AR(3) model for $inflat$, and we can assess the significance of each lag from the output given here. `t-DY_lag` is the t-statistic on the longest lag, where `D-lag` is the number of lagged differences (one less than the number of lagged levels), and `t-prob` is the probability of incorrectly rejecting the null hypothesis (that this lag is insignificantly different from zero). So for both inc and $cons$ the second and third lagged differences can be omitted as insignificant, but the first lag of inc can be omitted, it cannot for $cons$, as the t-statistic is 2.481. The first lag in differences is significant for $inflat$, which is important, as without this lag, the test outcome is significantly different: without a lag, the null hypothesis of $\gamma_{inflat} = 0$ would not be rejected as the t-statistic is -2.381, inside the critical value of -2.88 at the 5% level, while with the lagged difference, the statistic is -5.389, which is rejected conclusively; there is no random walk with drift in inflation, instead it is found to be stationary around a non-zero mean. For *inc* and *cons*, the additional lags make little difference, in all cases the t-statistics are well inside the critical values and the null hypotheses of $\gamma_{inc} = 0$ and $\gamma_{cons} = 0$ cannot be rejected. So the data support the conclusion that both *cons* and *inc* are random walk with drift.

c) Estimate the model:

$$cons_t = \beta_0 + \beta_1 cons_{t-1} + \beta_2 inc_t + \beta_3 inc_{t-1} + \beta_4 inflat_t + \epsilon_t.$$

Running this regression provides us with:

```
EQ( 1) Modelling cons by OLS (using DATA.in7)
       The estimation sample is: 1953 (2) to 1992 (3)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
cons_1               0.812547    0.02518     32.3   0.000   0.8719
Constant            -0.167215    0.06696    -2.50   0.014   0.0392

inc                  0.515763    0.02835     18.2   0.000   0.6838

inc_1               -0.303907    0.03472    -8.75   0.000   0.3337

inflat              -0.113270    0.009765  -11.6    0.000   0.4679

sigma               0.0012391  RSS            0.000234911463
R^2                   0.99375  F(4,153) =     6082 [0.000]**
log-likelihood 835.901
DW                       1.95
no. of observations       158 no. of parameters          5
mean(cons)            6.77507 var(cons) 0.000237902

AR 1-5 test:      F(5,148) =   1.1195 [0.3525]
ARCH 1-4 test:    F(4,145) =  0.73098 [0.5722]
Normality test:   Chi^2(2) =   1.7383 [0.4193]
hetero test:      F(8,144) =   1.3224 [0.2369]
hetero-X test:    F(14,138)=   1.1776 [0.2991]
RESET test:       F(1,152) =   2.0634 [0.1529]
```

The output shows that all the diagnostic tests are not rejected, suggesting that the model is well specified. The coefficients are significant, but we need to be careful with regard to interpretation given the non-stationarity of *cons* and *inc*. A hint at spurious regression might also be the very high R^2. σ, the regression standard deviation, is low at 0.12%.

d) Solve for the static long-run solution.

This is given in PcGive by going to

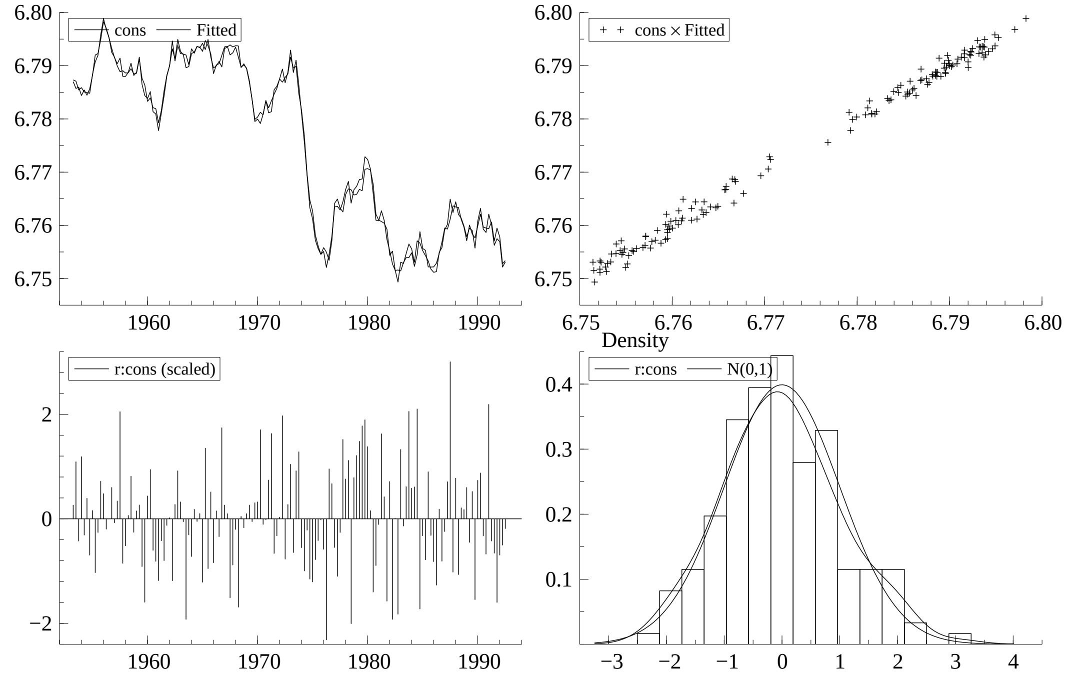

Figure 5: Graphical output from ADL of regression.

```
Test->Dynamic Analysis...->Static Long Run Solution
```

The long run solution is found by stripping out all the t subscripts, or more accurately assuming stationarity and taking expectations of (1), giving, where $E(X_t) = \mu_x$:

$$\begin{aligned}\mu_{cons} &= \beta_0 + \beta_1\mu_{cons} + (\beta_2 + \beta_3)\mu_{inc} + \beta_4\mu_{inflat}\\ \mu_{const} &= \frac{\beta_0}{1-\beta_1} + \frac{\beta_2+\beta_3}{1-\beta_1}\mu_{inc} + \frac{\beta_4}{1-\beta_1}\mu_{inflat}.\end{aligned}$$

Given our output above, we can calculate that:

$$\begin{aligned}\mu_{cons} &= -\frac{0.167215}{0.187453} + \frac{0.211856}{0.187453}\mu_{inc} - \frac{0.11327}{0.187453}\mu_{inflat}\\ &= -0.89 + 1.13\mu_{cons} - 0.60\mu_{inflat}.\end{aligned}$$

The output it produces here is:

```
Solved static long-run equation for cons
                  Coefficient  Std.Error  t-value  t-prob
Constant            -0.892039     0.3110    -2.87   0.005
inc                   1.13018    0.04574     24.7   0.000

inflat              -0.604260   0.06533 -9.25   0.000

Long-run sigma = 0.00661018
ECM = cons + 0.892039 - 1.13018*inc + 0.60426*inflat;
WALD test: Chi^2(2) = 848.673 [0.0000] **
```

e) For this part, we need to form the ECM model. This is done noting that $inflat_t$ is stationary, hence should not form part of the ECM relationship we found above. Instead, the rearranging of the model should go like the following, where y_t is con_t, z_t is inc_t and w_t is $inflat_t$:

$$\begin{aligned} y_t &= \beta_1 z_t + \beta_2 y_{t-1} + \beta_3 z_{t-1} + \beta_4 w_t \\ \Delta y_t &= (\beta_2 - 1)\, y_{t-1} + \beta_1 \Delta z_t + (\beta_1 + \beta_3)\, z_{t-1} + \beta_4 w_t + \epsilon_t \\ &= \beta_1 \Delta z_t + (\beta_2 - 1)\, [y_{t-1} - \kappa_1 z_{t-1}] + \beta_4 w_t + \epsilon_t, \end{aligned}$$

where $\kappa_1 = (\beta_2 - 1) / (\beta_1 + \beta_3)$, which we calculated earlier as 1.13. Hence we simply strip out the $inflat$ term in the ECM and estimate (1).

```
Modelling Dcons by OLS (using DATA.in7)
The estimation sample is: 1953 (2) to 1992 (3)

                  Coefficient  Std.Error  t-value  t-prob Part.R^2
Dinc                 0.515763    0.02751     18.8   0.000   0.6954
ECM_1               -0.187453    0.02448    -7.66   0.000   0.2757

inflat              -0.113270   0.009721    -11.7   0.000   0.4685
Constant            -2.03777e-006  0.0002212 -0.00921 0.993 0.0000

sigma              0.00123507  RSS            0.000234911463
R^2                  0.765594  F(3,154) =    167.7 [0.000]**
log-likelihood 835.901
DW                       1.95 no. of observations       158
no. of parameters           4
mean(Dcons)       -0.00024432 var(Dcons) 6.34275e-006

AR 1-5 test:      F(5,149) =   1.1264 [0.3489]
ARCH 1-4 test:    F(4,146) =  0.73602 [0.5688]
Normality test:   Chi^2(2) =   1.7384 [0.4193]
hetero test:      F(6,147) =   1.1921 [0.3137]
hetero-X test:    F(9,144) =   1.2531 [0.2677]
RESET test:       F(1,153) =   1.3282 [0.2509]
```

Note the exact correspondence between the coefficient on `Dinc` (β_1), and the coefficient on `inflat` (β_4), which is as expected given that this is simply a rearrangement of the earlier levels equation. Strictly speaking, the constant has been restricted to the cointegration space, i.e. inside the ECM term, but when the constant is omitted, the coefficients vary slightly, though not significantly.