Comments on Econometrics 1, Coursework

Ron Smith, EMS, Birkbeck

Christmas 2020-21

1 Exercise

- Do the exercise on your own sample of 100 years, which finishes in the year of your birth. For instance if you were born in 1990, your sample is 1891-1990.
  - **Do follow instructions! You lose marks if you chose another date for instance. Some had difficulty controlling sample when there were lags, so there were less than 100 observations. Although everyone had different samples there were some general features that came up quite often, not always.**
- When it says comment on your results, comment on any relevant statistical or economic aspects.
  - **You do have to interpret the results. You lose marks if you just give the output or do not discuss the economics.**

2 Steps

1. **Data generation and description**.
   (a) [5] Load the data and generate log GDP, LQ_t, log GDP deflator, LP_t, inflation, $INF_t = 100 * (LP_t - LP_{t-1})$, and growth, $G_t = 100 * (LQ_t - LQ_{t-1})$. Inflation and growth are in percent per annum.
   (b) [5] Suggest four economic relationships that you might expect between these variables? **E.G. Term structure, links short and long interest rates; Phillips Curve links inflation and unemployment, Fisher real interest relationship links interest rates and inflation, Okun's law links growth and unemployment; IS curve links output and interest rates.**
   (c) [5] For your sample, plot unemployment, growth, inflation and short and long interest rates. Calculate descriptive statistics. Comment on the main features. **Make sure that you comment on the**

variances. For instance. Inflation has much greater variance than interest rates or growth rates. They are all percent and in comparable units. Are short interest rates above or below long rates? Is the interest rate greater than the growth rate? $r - g$ is an important economic magnitude to some.

2. **Unrestricted model**

   (a) [5] Estimate over your sample

   $$U_t = \alpha_0 + \alpha_1 U_{t-1} + \alpha_2 U_{t-2} + \beta_0 LQ_t + \beta_1 LQ_{t-1} + \beta_2 LQ_{t-2} + \gamma t + \varepsilon_{1t}. \tag{1}$$

   (b) [5] Which coefficients are significant?

   (c) [5] Plot the residuals and comment on the results. **Some big spikes.**

   (d) [5] Carry out diagnostic tests on the residuals for (i) serial correlation, (ii) heteroskedasticity, (iii) normality, and (iv) non-linearity using a RESET test. **Note DW not strictly valid with a lagged dependent variable.**

   (e) [5] Briefly explain the form of the tests and what the results show. **For some programs you need to search to find the form of the tests and the explanation is not always clear, e.g. hettest in Stata.**

3. **Restricted model**

   (a) [5] Over your sample, estimate

   $$\Delta U_t = a_0 + a_1 \Delta U_{t-1} + b_0 \Delta LQ_t + \varepsilon_{2t}. \tag{2}$$

   Which coefficients are significant? Plot the residuals, repeat the four diagnostic tests and comment on the results. **You need to generate new variables ΔU_t and ΔLQ_t, to run this equation. This is an Okun's Law type equation.**

   (b) [5] What restrictions on (1) give (2) . Test the restrictions. **The unrestricted model (1) has 7 parameters the restricted model (2) 3, so we need 4 restrictions. The restricted model can be written**

   $$U_t = a_0 + (1 + a_1) U_{t-1} - a_1 U_{t-2} + b_0 LQ_t - b_0 LQ_t + \varepsilon_{2t}$$

   Matching coefficients**:** $\alpha_1 + \alpha_2 = 1$; $\beta_0 = -\beta_1$; $\beta_2 = 0$; $\gamma = 0$.

   (c) [5] Compare the models, (1) and (2) , and the results of the diagnostic tests. **Use LR tests and information criteria, the restricted is generally preferred on most samples. Strictly the LR test critical values are non-standard, but since they would be larger making it less likely to reject, if you do not reject it makes no difference.**

4. **Univariate models**

   (a) [5] Over your sample conduct an ADF unit root test on INF_t, U_t, RL_t and G_t. **Generally, INF_t and G_t rejected the unit root, looking I(0), U_t, RL_t did not looking I(1).**

   (b) [5] Estimate (i) a random walk with drift, and (ii) an ARMA(1,1) model for INF_t. **For RW with drift you can just regress the change in inflation on a constant or use ARIMA(0,1,0).**

   (c) [5] Use a LR test and information criteria to choose between the two models in 4(b). **The ARMA(1,1) model is**

   $$INF_t = \alpha + \rho INF_{t-1} + \varepsilon_t + \mu\varepsilon_{t-1}$$

   the restrictions to give the RW with drift are $\rho = 1, \mu = 0$. These restrictions are generally rejected since INF_t seems $I(0)$ over a century. Again the usual critical values for the LR test not strictly right since we are testing for a unit root and the distribution of the LR test would be non-standard with a larger critical value.

5. **VAR**

   (a) [5] Over your sample estimate a VAR between the four variables INF_t, RL_t, U_t and G_t.

   (b) [5] Test for (i) Granger causality, and (ii) cointegration. Comment on your results. (i) **Comment on the economics, e.g. because of expectations in some samples interest rates Granger cause, GC, inflation with a positive effect but inflation does not GC interest rates. With lag 1 can also use just the t stat to test for GC and look at the signs of the effects (ii) Generally find two or three cointegrating vectors, not surprising since two of the variables are I(0) and cointegrate with themselves.**

   (c) [10] Using the unrestricted VAR. Estimate impulse response functions for (i) a shock to G_t (ii) a shock to RL_t. Explain what assumptions you have made to estimate the impulse response functions. What economic interpretation would you give to the results? **If you use the Choleski orthogonalised IRFs, which are the standard ones, the main assumption is the causal ordering you used. See section 15.5 of the notes. If you used generalised you are looking at shocks to the reduced form not structural form and allowing for correlation between the errors.**

6. **Instrumental variables**

   (a) [5] Over your sample estimate by instrumental variables a regression of INF_t on a constant U_t and INF_{t-1} using as instruments a constant, INF_{t-1}, U_{t-1}, G_{t-1}, RL_{t-1}. Comment on your results. **The**

fact that U_t is not included in the list of instruments indicates that it is the endogenous variable.

(b) [5] Estimate the reduced form equation for unemployment. Does there seem to be a weak instrument problem? Save the residuals. Estimate by OLS a regression of INF_t on a constant U_t, INF_{t-1} and these residuals. Compare these estimates with those in 6a. What do your results tell you about the exogeneity of U_t? **The reduced form does an OLS regression of U_t on a constant, INF_{t-1}, U_{t-1}, G_{t-1}, RL_{t-1}. Save the residuals from this. Do OLS of INF_t** on a constant U_t, INF_{t-1} **and these residuals.The coefficient of U_t, INF_{t-1} in the equation with residuals are identical to the 2SLS/IV estimate you have alreadhy got. The t statistic on the residuals is a test for the null of exogeneity. Generally this t value is significant, U is endogenous.**