Exercise 7

Problem 1. To examine the quantity theory of money, Brumm (2005) [''Money Growth, Output Growth, and Inflation: A Reexamination of the Modern Quantity Theory's Linchpin Prediction,'' *Southern Economic Journal*, 71(3), 661–667] specifies the equation:

$$Inflation = \beta_0 + \beta_1 * Money + \beta_2 * Output + u$$

where *INFLAT* is the growth rate of the general price level, *MONEY* is the growth rate of the money supply, and *OUTPUT* is the growth rate of national output. According to theory we should observe that $\beta_0 = 0, \beta_1 = 1$, and $\beta_2 = -1$. The data used in this paper is contained in the file *brumm.gdt*. It consists of 1995 year data on 76 countries.

a) Estimate the model by OLS and interpret all the parameters.

ols Inflation const Money Output

```
Model 1: OLS, using observations 1-76
Dependent variable: Inflation

             coefficient   std. error    t-ratio     p-value
  -----------------------------------------------------------
  const       -0.234214    0.979925      -0.2390    0.8118
  Money        1.03313     0.00904221   114.3       4.65e-084 ***
  Output      -1.66201     0.250566      -6.633     4.95e-09  ***

Mean dependent var   25.35395   S.D. dependent var   58.94767
Sum squared resid    1356.034   S.E. of regression   4.309966
R-squared            0.994797   Adjusted R-squared   0.994654
F(2, 73)             6978.325   P-value(F)           4.41e-84
Log-likelihood      -217.3396   Akaike criterion     440.6792
Schwarz criterion    447.6714   Hannan-Quinn         443.4736
```

b) Test the joint hypothesis that $\beta_0 = 0, \beta_1 = 1$ and $\beta_2 = -1$. What do you conclude?

restrict
b[1] = 0
b[2] = 1
b[3] = -1
end restrict

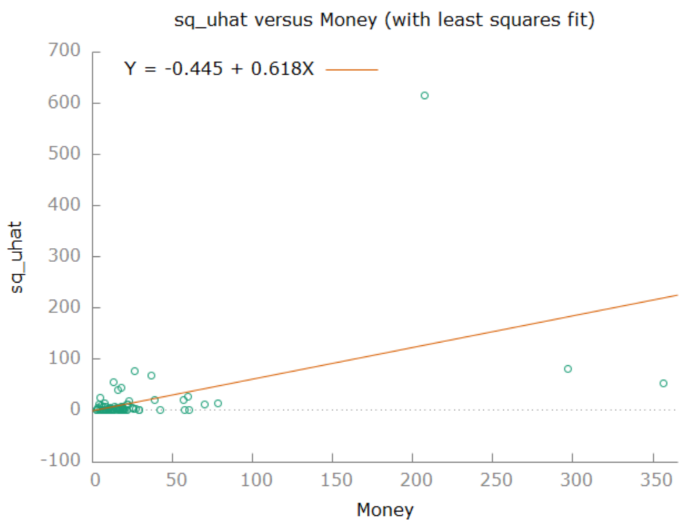
```
Restriction set
 1: b[const] = 0
 2: b[Money] = 1
 3: b[Output] = -1

Test statistic: F(3, 73) = 10.5158, with p-value = 7.88962e-006

Restricted estimates:

             coefficient   std. error   t-ratio   p-value
  -------------------------------------------------------
  const       0.000000     0.000000       NA        NA
  Money       1.00000      0.000000       NA        NA
  Output     -1.00000      0.000000       NA        NA

  Standard error of the regression = 5.05503
```

We reject H0, therefore, restrictions do not hold

c) Examine the least squares residuals for the presence of heteroskedasticity related to the variable *Money*.

series resid=$uhat

gnuplot sq_resid Money



modtest –white   (this tests all explanatory variables for heteroskedasticity

Manually:

genr mout=Money*Output

ols sq_resid const Money Output sq_Money sq_Output mout

```
Model 2: OLS, using observations 1-76
Dependent variable: sq_resid

              coefficient     std. error     t-ratio    p-value
  -------------------------------------------------------------
  const        -16,4206       20,9162        -0,7851    0,4351
  Money          2,36862       0,540649       4,381     4,06e-05 ***
  Output        -5,08322      10,6932        -0,4754    0,6360
  sq_Money      -0,00569956    0,00156983    -3,631     0,0005   ***
  sq_Output      0,254486      1,13671        0,2239    0,8235
  mout          -0,0176535     0,136937      -0,1289    0,8978

Mean dependent var    17,84255   S.D. dependent var    71,65330
Sum squared resid     238967,9   S.E. of regression    58,42797
R-squared             0,379408   Adjusted R-squared    0,335080
F(5, 70)              8,559119   P-value(F)            2,30e-06
Log-likelihood       -413,8667   Akaike criterion      839,7333
Schwarz criterion     853,7177   Hannan-Quinn          845,3222

Excluding the constant, p-value was highest for variable 8 (mout)
```

LM= R^2*n=0.3794*76=28.83

Critical value at 1% significance $\chi^2(5) = 15.086$

Therefore, we reject the hypothesis that there is no heteroskedasticity with respect to the variable Money

d) Obtain robust standard errors for the model and compare them to the OLS standard errors. Does your conclusion change in part (b) after using robust standard errors?

ols Inflation const Money Output –robust

```
Model 3: OLS, using observations 1-76
Dependent variable: Inflation
Heteroskedasticity-robust standard errors, variant HC1

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      -0.234214     0.619615     -0.3780    0.7065
  Money       1.03313      0.0236942    43.60      5.08e-054 ***
  Output     -1.66201      0.175914     -9.448     2.71e-014 ***

Mean dependent var   25.35395   S.D. dependent var   58.94767
Sum squared resid    1356.034   S.E. of regression   4.309966
R-squared            0.994797   Adjusted R-squared   0.994654
F(2, 73)             956.8215   P-value(F)           4.26e-53
Log-likelihood      -217.3396   Akaike criterion     440.6792
Schwarz criterion    447.6714   Hannan-Quinn         443.4736
```

Conclusion does not change – they are jointly not equal to the theoretical parameters

e) It is argued that *Output* may be endogenous. Four instrumental variables are proposed, *INITIAL* = initial level of real GDP, *SCHOOL* = a measure of the population's educational attainment, *INVEST* = average investment as a share of GDP, and *POPRATE* = average population growth rate. Using these instruments, obtain instrumental variables (2SLS) estimates of the inflation equation (do the two stage procedure).

First stage:

ols Output const initial poprate school invest Money

series Output_hat=$yhat

Second stage:

ols Inflation const Money Output_hat

```
Model 6: OLS, using observations 1-76
Dependent variable: Inflation

             coefficient   std. error   t-ratio    p-value
  ---------------------------------------------------------
  const      -1,09398      2,26858      -0,4822    0,6311
  Money       1,03506      0,0119309    86,75      2,16e-075 ***
  Out_hat    -1,39420      0,673298     -2,071     0,0419    **

Mean dependent var   25,35395   S.D. dependent var   58,94767
Sum squared resid    2052,737   S.E. of regression   5,302800
R-squared            0,992123   Adjusted R-squared   0,991908
F(2, 73)             4597,479   P-value(F)           1,65e-77
Log-likelihood      -233,0948   Akaike criterion     472,1896
Schwarz criterion    479,1818   Hannan-Quinn         474,9840
```

Alternatively, we can use Gretl command

tsls Inflation 0 Output Money ; 0 initial invest poprate school Money

OR

tsls Inflation const Output Money ; const initial invest poprate school Money

f) Are the instruments strong? **Only invest predicts the Output significantly, other variables are weak instruments. The theoretical parameters are again jointly rejected. The impact of output on the inflation is now lower than before.**

According to F test, instrument is weak because it falls below 10.27, where bias could have been roughly around 10% only but this is beyond the scope of our course

```
Weak instrument test -
  First-stage F-statistic (4, 70) = 4,64206
  Critical values for TSLS bias relative to OLS:

     bias       5%      10%      20%      30%
    value    16,85    10,27     6,71     5,34
```

Problem 2. Consider a simple model to estimate the effect of personal computer (PC) ownership on college grade point average for graduating seniors at a large public university:

$$GPA = \beta_0 + \beta_1 PC + u$$

where PC is a binary variable indicating PC ownership.

(i) Why might PC ownership be correlated with u?

It has been fairly well established that socioeconomic status affects student performance. The error term *u* contains, among other things, family income, which has a positive effect on *GPA* and is also very likely to be correlated with PC ownership

(ii) Explain why PC is likely to be related to parents' annual income. Does this mean parental income is a good IV for PC? Why or why not?

Families with higher incomes can afford to buy computers for their children. Therefore, family income certainly satisfies the second requirement for an instrumental variable: it is correlated with the endogenous explanatory variable. But as we suggested in part (i), *faminc* has a positive affect on *GPA*, so the first requirement for a good IV fails for *faminc*. If we had *faminc* we would include it as an explanatory variable in the equation; if it is the only important omitted variable correlated with *PC*, we could then estimate the expanded equation by OLS.

(iii) Suppose that, four years ago, the university gave grants to buy computers to roughly one-half of the incoming students, and the students who received grants were randomly chosen. Carefully explain how you would use this information to construct an instrumental variable for PC.

This is a natural experiment that affects whether or not some students own computers. Some students who buy computers when given the grant would not have without the grant.

(Students who did not receive the grants might still own computers.) Define a dummy variable, *grant*, equal to one if the student received a grant, and zero otherwise. Then, if *grant* was randomly assigned, it is uncorrelated with *u*. In particular, it is uncorrelated with family income and other socioeconomic factors in *u*. Further, *grant* should be correlated with *PC*: the probability of owning a PC should be significantly higher for student receiving grants. Incidentally, if the university gave grant priority to low-income students, *grant* would be negatively correlated with *u*, and IV would be inconsistent.