ECON 546: Themes in Econometrics

Lab Exercises #6
(24 February, 2010)

In this Lab. class we will use Monte Carlo simulation to investigate the finite-sample performance of the Breusch-Godfrey LM for serial independence, against the alternative hypothesis of AR(1) errors in our regression model. We are going to do the following:

(a) Investigate the distortion in the significance level (or "size") of the test when it is applied in small samples.

(b) Investigate the ability of the test to reject false hypotheses (*i.e.*, its "power", when there is no adjustment for any size distortion) when it is applied in small samples.

(c) Examine the effect on each of the above results, in both small and large sample sizes, when the test is wrongly applied in a situation where the model's errors follow a simple form of heteroskedasticity.

In a paper in the latest issue of *Applied Economics Letters*, Hyun *et al.* (2010) have undertaken a similar study. However, note that they looked only at an "F-test" version of the LM test, rather than the nR^2 version that we usually use. In addition, they looked only at the effect of the heteroskedasticity on the test's significance level – they did not consider power. A copy of their paper is attached, and you should look at it before the Lab. class.

Recall how the Breusch-Godfrey LM test statistic is constructed in the case where the null hypothesis is that the errors are serially independent, and the alternative hypothesis is that there is first-order autocorrelation:

- Estimate the model by OLS: $y = X\beta + \varepsilon$
- Regress the residuals on *X* and lagged residuals: $e = X\gamma + \delta e_{-1} + u$
- Use the R^2 from this second regression to calculate LM: $LM = nR^2$
- Then $LM \to \chi^2_{(1)}$ as $n \to \infty$, if the null hypothesis is true.

(**Note:** Strictly speaking, the R^2 that is used here should be the "uncentered" R^2, but this will just be the usual R^2 as long as the model includes an intercept.)

The EViews workfile and program file that you need are on the server, and are called **lab6.wfl** and **lab6.prg** respectively.

Here is what I would like you to do:

1. Look at the program code and answer the following questions:
   (i) What is the sample size?
   (ii) What (nominal) significance level is being used for the LM test?
   (iii) Are the data being generated with the null hypothesis true or false?
   (iv) Are the errors in the data generating process homoskedastic or not?
2. Now run the program:
   (i) Is the value of **reject_rate** (close to) what you expected?
   (ii) Look at the sampling distribution for the LM statistic. Are its mean and variance (close to) what you expected?

(iii) Construct a Quantile-Quantile plot of LMS against the appropriate theoretical asymptotic distribution. Interpret the graph.

3. Leaving everything else the same, now consider sample sizes of n = 10, 25, 50 and 100. What do you observe?

4. Now, set n = 1000 and run the program with different non-zero values for Rho. What are you examining here? What do you observe?

5. Repeat question 4, with n = 25. Compare your results with those when n = 1000.

6. Set Rho = 0, and n = 1000. Leave **!tau = 0.5**, but set **!delta=0.25**. What situation are you now considering for the data-generating process and for the test? Run the program and discuss what you find. What happens if you set Rho = 0.1?

7. Finally, set Rho = 0 and n = 50. (Leave **!tau = 0.5**, and **!delta=0.25**.) Run the program, and compare your result with that in Table 1 of Hyun *et al*.'s paper.

As a follow-on exercise, you could now consider different values of Rho and different values of **!tau** and **!delta**, with n = 50, to see how the power of the test is affected by the degree and location of the heteroskedasticity when the sample size is quite small.

Reference:

Hyun, J-Y.,H. H. Mun, T-H. Kim and J. Jeong (2010), “The effect of a variance shift on the Breusch-Godfrey’s LM test” *Applied Economics Letters*, 17, 399-404.

PROGRAM CODE

```
' A PROGRAM TO EXAMINE THE SIZE AND POWER OF THE (nR-SQUARED) VERSION OF THE LM
TEST FOR SERIAL INDEPENDENCE, AGAINST AR(1) / MA(1) ERRORS, WHEN THE SAMPLE SIZE IS
SMALL.

' NOTE THAT IN THE CASE OF THE POWER CALCULATIONS, THE PROGRAM IS SET UP ONLY TO
LOOK AT AR(1) ERRORS, AND NOT MA(1) ERRORS, UNDER THE ALTERNATIVE HYPOTHESIS. (OF
COURSE THIS COULD BE EXTENDED QUITE EASILY)

' THE PROGRAM ALSO ALLOWS US TO EXAMINE THE EFFECTS OF HETEROSKEDASTICITY ON THE
PROPERTIES OF THE TEST. THESE RESULTS CAN BE COMPARED WITH THOSE OF HYUN et al.
("APPLIED ECONOMICS LETTERS", 2010, vol.17, pp. 399-404) - THEY LOOK AT AN F-TEST VARIANT
OF THE LM TEST.
'=================================================================

' INITIALIZE VARIOUS VALUES
' ---------------------------------------------
rndseed 123456
!nrep=1000
' CHOOSE THE (NOMINAL)  SIGNIFICANCE LEVEL FOR THE LM TEST, AND USE IT TO CALCULATE
THE (NOMINAL) CRITICAL VALUE:
scalar sig_level=0.05
scalar s=1-sig_level
scalar crit=@qchisq(s,1)
' CHANGE THE NEXT LINE TO GO FROM THE NULL HYPOTHESIS TO THE ALTERNATIVE
HYPOTHESIS
!rho= 0.0
' SET THE SAMPLE SIZE FOR THE CASE WHERE THE NULL IS TRUE,  AND THEN ADD ONE TO
ALLOW FOR THE FACT THAT WE WILL "LOSE" AN OBSERVATION WHEN WE GENERATE
AUTOCORRELATED ERRORS
!n=1000
!n1=!n+1
```

```
smpl 1 !n1
scalar b0=1
scalar b1=1
scalar b2=1
series y
series eps
series res
series sig=1
vector(!nrep) lm
' CHANGE THE NEXT LINE TO ALTER THE DEGREE OF HETEROSKEDASTICITY
!delta=1
' CHANGE THE NEXT LINE TO ALTER THE LOCATION OF THE BREAK-POINT IN THE VARIANCE,
FROM SIGMA1^2 TO SIGMA2^2
!tau=0.5

' SET UP THE HETEROSKEDASTICITY IN THE ERRORS
' MAKE SURE THAT THE BREAK OCCURS AT AN INTEGER VALUE IN THE SAMPLE
' NOTE THAT DELTA ALTERS THE VARIANCE SO WE NEED TO TAKE TEH SQUARE ROOT HERE
WHEN ALTERING THE STANDARD DEVIATION

!break1=@floor(!n1*!tau)+1
smpl !break1 !n1
sig=@sqrt(!delta)*sig
' GENERATE THE DATA FOR THE REGRESSORS - THESE WILL THEN BE FIXED IN REPEATED
SAMPLES
series x1=@rnorm
series x2=@rnorm

' START THE MONTE CARLO LOOP
for !i = 1 to !nrep
smpl 1 !n1
eps=0
smpl 2 !n1
' HERE IS THE DATA-GENERATING PROCESS
eps=!rho*eps(-1)+@rnorm*sig
y=b0+b1*x1+b2*x2+eps
equation eq1.ls y c x1 x2
' RE-NAME "RESID" BECAUSE IT IS A RESERVED NAME AND CAN'T BE USED AS A VARIABLE IN A
REGRESSION
res=resid
' FIT THE ARTIFICAL REGRESSION TO GET THE LM TEST STATISTIC. STORE THE LM STATISTIC
VALUES AND ALSO
equation eq2.ls res c x1 x2 res(-1)
lm(!i)=@regobs*@r2
next

' END OF THE MONTE CARLO LOOP

' COMPUTE THE   "REFECTION RATE" OF THE TEST. THIS WILL BE THE TRUE "SIZE" OF THE TEST
IF THE NUL HYPOTHESIS IS TRUE, OR THE (NOMINAL) POWER IN THE CASE WHERE THE
ALTERNATIVE HYPOTHESIS OF AUTOCORRELATION HAS BEEN IMPOSED ON THE DATA-
GENERATING PROCESS.
' ALSO, PLOT THE SAMPLING DISTRIBUTION OF THE LM TEST STATISTIC

smpl 1 !nrep
' THE FOLLOWING WAY OF CREATING THE REJECTION RATE IS FASTER THAN TESTING WITHIN
EACH ITERATION OF THE MONTE CARLO LOOP TO SEE IF lm EXCEEDS THE CRITICAL VALUE
mtos(lm,lms)
series dum=(lms>crit)
scalar reject_rate=@sum(dum)/!nrep
lms.hist
```