Exercises Week 3

Econometrics

1. In the lecture we argued that in case of heteroskedastiity/autocorrelation, GLS is more efficient than OLS by looking at their covariance matrices. You are going to show that this is indeed the case. That is, show that

$$Var^{-1}(\tilde{\beta}_{GLS}) - Var^{-1}(\hat{\beta}_{OLS}) = X^T\Omega^{-1}X - X^TX(X^T\Omega X)^{-1}X^TX,$$

   is a positive semidefinite matrix.

   Hint: Write $\Omega = PP^T$ and show that the difference can be written as $Z^T M_Q Z$, where M_Q is the matrix that projects off the space generated by Q, for appropriate Z and Q matrices.

2. Generate a sample of size 30 from the model

$$y_t = \beta_0 + \beta_1 x_t + u_t,$$

   with $\beta_0 = 1$ and $\beta_1 = 1$. For simplicity, assume that x_t are $NID(2,2)$. Moreover, make $u_t = N(0, \exp(0.6x_t))$; that is, the errors are heteroskedastic.

   For this exercise you can only use the *lm()* function and the HCCME function from your favorite package (*vcovHC()* from *sandwich* or *hccm()* from *car*).

   (a) Estimate the model by OLS; that is, without correcting for heteroskedasticity. Compute the t-statistic associated to the test $H_0 : \beta_1 = 0$ and determine if you reject the null using a critical value of 2.0423. Why use this value?

   (b) Compute the t-statistic associated to the test $H_0 : \beta_1 = 0$ using a HCC matrix and determine if you reject the null using the same critical value as before.

   (c) Estimate the model by GLS using the true weights and compute the t-statistic associated to the test $H_0 : \beta_1 = 0$. Determine if you reject the null using the same critical value as before.

   (d) Estimate the model by FGLS using weights obtained from the auxiliar regression

$$\log(\hat{u}_t^2) = x_t\gamma + \nu_t.$$

   Compute the t-statistic associated to the test $H_0 : \beta_1 = 0$. Determine if you reject the null using the same critical value as before.

   (e) Repeat the exercise at least 100 times and compute the average number of rejections of the null for all methods consider. What do you observe? Explain

3. **Ex. 7.14 in ETM**: The dataset *Money* from the *Ecdat* package contains seasonally adjusted quarterly data for the logarithm of the real money supply, m_t, real GDP, y_t, and the 3-month Treasury Bill rate, r_t, for Canada for the period 1967:1 to 1998:4. A conventional demand for money function is

$$m_t = \beta_1 + \beta_2 r_t + \beta_3 y_t + \beta_4 m_{t-1} + u_t.$$

   Estimate this model over the period 1968:1 to 1998:4, and then test it for AR(1) errors using a Durbin-Watson test.

4. **Exercise 4.1 in AGME**: The dataset *Airq* from the *Ecdat* package contains observations for 30 standard metropolitan statistical areas (SMSAs) in California for 1972 on the following variables:

   airq: indicator for air quality (the lower the better);

   vala: value added of companies (in 1000 US$);

   rain: amount of rain (in inches);

   coas: dummy variable, 1 for SMSAs at the coast; 0 for others;

   dens: population density (per square mile);

   medi: average income per head (in US$).

   (a) Estimate a linear regression model that explains *airq* from the other variables using ordinary least squares. Interpret the coefficient estimates.

   (b) Test the null hypothesis that average income does not effect the air quality. Test the joint hypothesis that none of the variables has an effect upon air quality.

   (c) Perform a Breusch–Pagan test for heteroskedasticity related to all five explanatory variables.

   (d) Perform a White test for heteroskedasticity.

   (e) Assuming that we have heteroskedasticity related to *coas* and *medi*, estimate the coefficients by running a regression of $\log u_t^2$ upon these two variables. Test the null hypothesis of homoskedasticity on the basis of this auxiliary regression.

   (f) Using the results from (e), compute a FGLS estimator for the linear model. Compare your results with those obtained under (a). Redo the tests from (b).