ECON471

Fall 2020

Problem Set 6

Due Wednesday December 9, by Midnight CST

Name:

Section:

1. Consider the model $y_i = \beta x_i + u_i$ with $E(u_i|x_i) = 0$ and $Var(u_i|x_i) = \sigma^2 x_i^2$. An estimator of β is obtained as follows: $\tilde{\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i}$.

(i) Derive the expected value of $\tilde{\beta}$ and show that it is unbiased.

$$E(\tilde{\beta}) = \frac{1}{n}\sum_{i=1}^{n} E\left(\frac{\beta x_i + u_i}{x_i}\right)$$

$$= \frac{1}{n}\sum_{i=1}^{n} E\left(\beta + \frac{u_i}{x_i}\right) = \frac{1}{n}\sum_{i=1}^{n}(\beta + 0)$$

$$= \beta.$$

(ii) Derive the weighted least squares estimator of β and show that it is identical to $\tilde{\beta}$. Is $\tilde{\beta}$ BLUE? Without any explicit derivations, compare the efficiency of $\tilde{\beta}$ to the OLS estimator of β.

$$Var(u_i|x_i) = \sigma^2 x_i^2 \Rightarrow \sqrt{h(x)} = x_i$$

$$\frac{y_i}{x_i} = \beta\frac{x_i}{x_i} + \frac{u_i}{x_i} \Rightarrow \text{homoscedasticity.}$$

$$\min \sum_{i=1}^{n}\left(\frac{y_i}{x_i} - \tilde{\beta}\right)^2 \Rightarrow \frac{\partial \sum_{i=1}^{n}\left(\frac{y_i}{x_i} - \tilde{\beta}\right)^2}{\partial \tilde{\beta}} = \sum_{i=1}^{n} -2\left(\frac{y_i}{x_i} - \tilde{\beta}\right) = 0$$

$$\tilde{\beta} = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i}{x_i}$$

$\tilde{\beta}$ is BLUE $\Rightarrow$ $\tilde{\beta}$ has a lower variance than any other linear unbiased estimator of β, including OLS estimator of β.

2. There are different ways to combine features of the Breusch-Pagan and White tests for heteroskedasticity. One possibility not covered in the lectures is to run the regression

$$\hat{u}_i^2 \text{ on } x_{i1}, x_{i2}, \ldots\ldots.., x_{ik}, \hat{y}_i^2, \quad i = 1, \ldots\ldots, n,$$

where the $\hat{u}_i$ are the OLS residuals and the $\hat{y}_i$ are the OLS fitted values. Then, we would test joint significance of $x_{i1}, x_{i2}, \ldots\ldots.., x_{ik}$ and $\hat{y}_i^2$. (Of course, we always include an intercept in this regression.)

(i) What are the degrees of freedom associated with the proposed F test for heteroskedasticity?

$k+1,\ n-k-2.$

(ii) Explain why the R-squared from the regression above will always be at least as large as the R-squared from the Breusch-Pagan regression and the special case of the White test.

the number of explanatory variables in this method is more than Breusch-Pagan regression.

special case: $\hat{u}_i^2$ on $\cancel{\hat{y}_i}, \hat{y}_i^2$

Put a restriction on how orginal explanatory variables appear in regression

(iii) Does part (ii) imply that the new test always delivers a smaller p-value than either the Breusch-Pagan or special case of the White statistic? Explain.

$$F = \frac{R^2_{\hat{u}^2}/(k+1)}{(1-R^2_{\hat{u}^2})/(n-k-2)} \qquad \frac{n-(k+1)-1}{k+1}$$

$$F_{BP} = \frac{R^2_{\hat{u}^2_{BP}}/k}{(1-R^2_{\hat{u}^2_{BP}})/(n-k-1)} \qquad F_W = \frac{R^2_{\hat{u}^2_W}/2}{(1-R^2_{\hat{u}^2_W})/n-3}$$

$\Longrightarrow$ No

(iv) Suppose someone suggests also adding $\hat{y}_i$ to the newly proposed test. What do you think of this idea?

No need $\hat{y}_i$ is linear combination of the original regressors

$\Rightarrow$ perfect collinearity.

3.

(i) Use the data in HPRICE1.TXT to obtain the heteroskedasticity-robust standard errors for the following model

$$\text{price} = \beta_0 + \beta_1 lotsize + \beta_2 sqrft + \beta_3 bdrms + u.$$

Discuss any important differences with the usual standard errors.

(ii) Repeat part (i) for the following model

$$\log(\text{price}) = \beta_0 + \beta_1 \log(lotsize) + \beta_2 \log(sqrft) + \beta_3 bdrms + u.$$

(iii) What does this example suggest about heteroskedasticity and the transformation used for the dependent variable?

4. Use VOTE1.TXT for this exercise.
   (i) Estimate a model with *voteA* as the dependent variable and *prtystrA*, *democA*, log*(expendA)*, and log*(expendB)* as independent variables. Obtain the OLS residuals, $\hat{u}_i$, and regress this on all of the independent variables. Explain why you obtain $R^2 = 0$.

$\hat{u}_i$ is uncorrelated with all independent vars.

   (ii) Now, compute the Breusch-Pagan test for heteroskedasticity. Use the *F*-statistic version and report the *p*-value.

(iii) Compute the special case of the White test for heteroskedasticity, again using the *F* statistic form. How strong is the evidence against heteroskedasticity now?

5. Use the data in MEAP00_01.TXT to answer this question.

(i) Estimate the model

$$math4 = \beta_0 + \beta_1 lunch + \beta_2 \log(enroll) + \beta_3 \log(exppp) + u$$

by OLS and obtain the usual standard errors and the fully robust standard errors. How do they generally compare?

(ii) Apply the special case of the White test for heteroskedasticity. What is the value of the *F* test. What do you conclude?