ECO375 Tutorial 5
Wooldridge: Chapter 8 and Misc

Matt Tudball

University of Toronto St. George

October 13, 2017

ECO375 Tutorial 5

Welcome back!

Today's coverage:

- Chapter 8, #1 (in slides)
- Chapter 8, #4 (in slides)
- Chapter 4, #6 (in slides)
- Chapter 6, #5 (in slides)

Chapter 8, #1

Which of the following are consequences of heteroskedasticity?
i) The OLS estimators $\hat{\beta}_j$ are inconsistent.

i) We haven't covered consistency yet in this class so it may be difficult to answer this question. I will outline the answer briefly anyway.
$\hat{\beta} = (X'X)^{-1}X'y = (X'X)^{-1}X'(X\beta + u) = \beta + (X'X)^{-1}X'u$
$= \beta + (X'X/n)^{-1}X'u/n$
With consistency we are asking what happens to $\hat{\beta}$ as n goes to infinity.
Recall that we assume that $E(u_i|X_i) = 0$, regardless of heteroskedasticity.
Also note that $E(u_iX_i) = E(X_iE(u_i|X_i)) = 0$. By the Law of Large Numbers, $X'u/n \xrightarrow{p} E(u_iX_i) = 0$ and so $\hat{\beta} \xrightarrow{p} \beta$.
Heteroscedasticity therefore **does not** affect consistency.

Chapter 8, #1

Which of the following are consequences of heteroskedasticity?
ii) The usual F statistic no longer has an F distribution.

ii) Remember that the F distribution is the ratio of two independent χ^2 distributions each divided by their degrees of freedom. $F = \frac{X_1/d_1}{X_2/d_2}$ where $X_1 \sim \chi^2_{d_1}$ and $X_2 \sim \chi^2_{d_2}$. Remember also that the χ^2 distribution on d degrees of freedom is the sum of squares of d standard normal distributions. Remember that the F statistic can be written $F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$. If u is not homoskedastic than the SSR_r and SSR_{ur} will not be the sum of squares of residuals with a common variance. This means that they will not be χ^2 distributed and therefore the F statistic will not follow an F distribution.
Heteroskedasticity therefore **does** affect the distribution of the F statistic.

Chapter 8, #1

Which of the following are consequences of heteroskedasticity?
iii) The OLS estimators are no longer BLUE.

iii) Homoskedasticity is one of the Gauss-Markov assumptions needed for OLS to be BLUE.
Heteroskedasticity therefore **does** affect whether OLS is BLUE.

Chapter 8, #4

Using the data in GPA3.dta, the following equation was estimated for the fall and second semester students:

$$\begin{aligned}\widehat{trmgpa} = &-2.12 + .900\ crsgpa + .193\ cumgpa + .0014\ tothrs \\ &\quad (.55) \quad (.175) \quad (.064) \quad (.0012) \\ &\quad [.55] \quad [.166] \quad [.074] \quad [.0012] \\ &+ .0018\ sat - .0039\ hsperc - .351\ female - .157\ season \\ &\quad (.0002) \quad (.0018) \quad (.085) \quad (.098) \\ &\quad [.0002] \quad [.0019] \quad [.079] \quad [.080] \\ n = &\ 269,\ R^2 = 0.465\end{aligned}$$

Here *trmgpa* is term GPA, *crsgpa* is a weighted average of overall GPA in courses taken, *cumgpa* is GPA prior to the current semester, *tothrs* is total credit hours prior to the semester, *sat* is SAT score, *hsperc* is graduating percentile in high school class, *female* is a gender dummy and *season* is a dummy variable equal to one if the student's sport is in season during the fall. The usual and heteroskedasticity-robust standard errors are reported in parentheses and square brackets, respectively.

Chapter 8, #4

i) Do the variables *crsgpa*, *cumgpa* and *tothrs* have the expected estimated effects? Which of these variables are statistically significant at the 5% level? Does it matter which standard errors are used?

These coefficients have the anticipated signs. If a student takes courses where grades are, on average, higher, as reflected by higher *crsgpa*, then his or her grades will be higher. The better the student has been in the past, as measured by *cumgpa*, the better the student does (on average) in the current semester. Finally, *tothrs* is a measure of experience, and its coefficient indicates an increasing return to experience.

crsgpa and *cumgpa* are significant at the 5% level using either standard errors and *tothrs* is insignificant at the 5% levels using either standard errors.

Chapter 8, #4

ii) Why does the hypothesis $H_0 : \beta_{crsgpa} = 1$ make sense? Test this hypothesis against the two-sided alternative at the 5% level, using both standard errors. Describe your conclusions.

Suppose there were no other explanatory variables in the model. $\beta_{crsgpa} = 1$ indicates that the best predictor of an individual student's term GPA is the average GPA in the student's courses. We can calculate the t statistic under homoskedasticity $t_{hom} = \frac{.9-1}{.175} = -.571$ and under heteroskedasticity $t_{het} = \frac{.9-1}{.166} = -.602$.
We fail to reject the null hypothesis in both cases.

Chapter 8, #4

iii) Test whether there is an in-season effect on term GPA using both standard errors. Does the significance level at which the null can be rejected depend on the standard error used?

We want to test $H_0 : \beta_{season} = 0$ against the two-sided alternative. $t_{hom} = \frac{-.157}{.098} = -1.602$ and $t_{het} = \frac{-.157}{.0.080} = -1.962$. We therefore cannot reject the null hypothesis under homoscedasticity but we can reject it under heteroscedasticity.

This is unusual since usually heteroskedasticity-robust standard errors are larger. Conventional standard errors will be too big whenever covariate values far from the mean of the covariate distribution are associated with lower variance residuals (so small residuals for small and big values of X, and large residuals in the middle of the X range). The most common case is where the variance of the residuals grows with X.

Chapter 4, #6

In Section 4.5, we used as an example testing the rationality of assessments of housing prices. There, we used as a log-log model in *price* and *assess* [see equation (4.47)]. Here, we use a level-level formulation.

In the simple regression model

$$price = \beta_0 + \beta_1 assess + u$$

the assessment is rational if $\beta_1 = 1$ and $\beta_0 = 0$. The estimated equation is

$$\widehat{price} = -14.47 + .976\ assess + u$$
$$(16.27) \quad (.049)$$
$$n = 88,\ SSR = 165,644.51,\ R^2 = 0.820$$

Chapter 4, #6

i) First, test the hypothesis that $H_0 : \beta_0 = 0$ against the two-sided alternative. Then, test $H_0 : \beta_1 = 1$ against the two-sided alternative. What do you conclude?

For $H_0 : \beta_0 = 0$, $t = \frac{-14.47}{16.27} = -0.889$ and so we fail to reject the null hypothesis. For $H_0 : \beta_1 = 1$, $t = \frac{0.976-1}{0.049} = -0.490$ and so we fail to reject the null hypothesis. In both cases there is insufficient evidence against rationality.

Chapter 4, #6

ii) To test the joint hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$, we need the SSR in the restricted model. This amounts to computing $\sum_{i=1}^{n}(price_i - assess_i)^2$, where $n = 88$, since the residuals in the restricted model are just $price_i - assess_i$. (No estimation is needed for the restricted model because both parameters are specified under H_0). This turns out to yield $SSR_r = 209,448.99$. Carry out the F test for the joint hypothesis.

Recall $F = \frac{(SSR_r - SSR_{ur})/q}{SSR_{ur}/(n-k-1)}$. Note that $q = 2$ since we have 2 restrictions, $n = 88$ and $k = 1$. $SSR_r = 209,448.99$ as above and $SSR_{ur} = 165,644.51$ as given in the initial question. We have all of the information we need to calculate our F statistic,

$F = \frac{(209448.99 - 165644.51)/2}{165644.51/(88-1-1)} = 11.382$

We therefore have very strong evidence to reject the null hypothesis of rationality.

Chapter 4, #6

iii) Now, test $H_0 : \beta_2 = 0, \beta_3 = 0, \beta_4 = 0$ in the model

$$price = \beta_0 + \beta_1 assess + \beta_2 lotsize + \beta_3 sqrft + \beta_4 bdrms + u$$

The R^2 from estimating this model using the same 88 houses is 0.829.

Recall $F = \frac{(R^2_{ur} - R^2_r)/q}{(1-R^2_{ur})/(n-k-1)}$. Note that $q = 3$ since we have 3 restrictions, $n = 88$ and $k = 4$. $R^2_{ur} = 0.829$ and $R^2_r = 0.820$. We have all of the information we need to calculate our F statistic,
$F = \frac{(0.829-0.820)/3}{(1-0.829)/(88-4-1)} = 1.456$.
The 10% critical value (again using 90 denominator df in Table G.3a) is 2.15, so we fail to reject H0 at even the 10% level.

Chapter 4, #6

iv) If the variance of *price* changes with *assess*, *lotsize*, *sqrft* or *bdrms*, what can you say about the F test from part (iii)?

Chapter 4, #6

iv) If the variance of *price* changes with *assess*, *lotsize*, *sqrft* or *bdrms*, what can you say about the F test from part (iii)?

As we showed above, if heteroskedasticity were present, the F statistic would not have an F distribution under the null hypothesis. Therefore, comparing the F statistic against the usual critical values from the F distribution would not be especially meaningful.

Chapter 6, #5

In example 4.2, where the percentage of students receiving a passing score on a tenth grade math exam (*math*10) is the dependent variable, does it make sense to include *sci*11 - the percentage of eleventh graders passing a science exam - as an additional explanatory variable?

Example 4.2 uses a school-level dataset and is interested in the effect of school size on math scores. The estimated equation in example 4.2 is,

$$math10 = \beta_0 + \beta_1\ totcomp + \beta_2\ staff + \beta_3\ enroll + u$$

totcomp is average annual teacher compensation, *staff* is the number of staff per one thousand students and *enroll* is student enrolment.

Chapter 6, #5

Probably not. Think about the causal pathways in this question. We want to know the effect of school size on student performance as measured by scores on a tenth grade math exam. School size is also going to affect performance on eleventh grade science exams. By controlling for this, our estimate on *enroll* (our school size measure) is going to be interpreted as the effect of school size on tenth grade math scores other than the effects which also determine eleventh grade science scores. The interpretation becomes really odd. It is also true that *sci*11 is an output of the education system. It makes little sense to control for another output when we are interested in the effects of inputs into education.