Economics 471: Econometrics

Department of Economics, Finance and Legal Studies
University of Alabama

Problem Set #7 – Answers

1. With $V(u \mid inc, price, educ, female) = \sigma^2 inc^2$, $h(x) = inc^2$, where $h(x)$ is the heteroskedasticity function defined in class. Therefore, $\sqrt{h(x)} = inc$, and so the transformed equation is obtained by dividing the original equation by *inc*.

$$\frac{beer}{inc} = \alpha\frac{1}{inc} + \beta_1 + \beta_2\frac{price}{inc} + \beta_3\frac{educ}{inc} + \beta_4\frac{female}{inc} + \frac{u}{inc}$$

Notice that β_1, which is the slope on *inc* in the original model, is now a constant in the transformed equation. This is simply a consequence of the form of the heteroskedasticity and the functional forms of the explanatory variables in the original equation.

2. (a)

$$\begin{aligned}\widehat{trmgpa} &= -2.12 + 0.900crsgpa + 0.193cumgpa + 0.0014tothrs + 0.0018sat - 0.0039hsperc \\ &\quad +0.351female - 0.157season\end{aligned}$$

The homoskedastic standard errors are 0.55, 0.175, 0.064, 0.0012, 0.0002, 0.0018, 0.085, 0.098. The robust standard errors are 0.55, .166, 0.74, 0.0012, 0.0002, 0.0019, 0.079 and 0.080.

(b) These coefficients have the anticipated signs. If a student takes courses where grades are, on average, higher – as reflected by higher *crsgpa* – then his/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. The t statistic for *crsgpa* is very large, over five using the usual standard error (which is the largest of the two). Using the robust standard error for *cumgpa*, its t statistic is about 2.61, which is also significant at the 5% level. The t statistic for *tothrs* is only about 1.17 using either standard error, so it is not significant at the 5% level.

(c) This is easiest to see without other explanatory variables in the model. If *crsgpa* were the only explanatory variable, $H_0 : \beta_{crsgpa} = 1$ means that, without any information about the student, the best predictor of term GPA is the average GPA in the students' courses; this holds essentially by definition. (The intercept would be zero in this case.) With additional explanatory variables it is not necessarily true that $\beta_{crsgpa} = 1$ because *crsgpa* could be correlated with characteristics of the student. (For example, perhaps the courses students take are influenced by ability – as measured by test scores – and past college performance.) But it is still interesting to test this hypothesis. The t statistic using the usual standard error is $t = (0.900 - 1)/0.175 \approx -0.57$; using the heteroskedasticity-robust standard error gives $t \approx -0.60$. In either case we fail to reject $H_0 : \beta_{crsgpa} = 1$ at any reasonable significance level, certainly including 5%.

(d) The in-season effect is given by the coefficient on *season*, which implies that, other things equal, an athlete's GPA is about .16 points lower when his/her sport is competing. The t statistic using the usual standard error is about -1.60, while that using the robust standard error is about -1.96. Against a two-sided alternative, the t statistic using the robust standard error is just significant at the 5% level (the standard normal critical value is 1.96), while using the usual standard error, the t statistic is not quite significant at the 10% level (cv 1.65). So the standard error used makes a difference in this case. This example is somewhat unusual, as the robust standard error is more often the larger of the two.

3. (a) $\widehat{colGPA} = 1.36 + 0.412 hsGPA + 0.013 ACT + 0.071 skipped + 0.124 PC$

   (b) The F statistic obtained for the White test is about 3.58. With 2 and 138 df, this gives p-value $\approx$ .031. So, at the 5% level, we conclude there is evidence of heteroskedasticity in the errors of the $colGPA$ equation.

   (c) In fact, the smallest fitted value from the regression in part (b) is about .027, while the largest is about .165. Using these fitted values as the $\widehat{h}_i$ in a weighted least squares regression gives the following:

$$\widehat{colGPA} = 1.40 + 0.402 hsGPA + 0.013 ACT - 0.076 skipped + 0.126 PC$$

   There is very little difference in the estimated coefficient on PC, and the OLS t statistic and WLS t statistic are also very close. The R^2 in the weighted least squares estimation is larger than that from the OLS regression in part (a), but, remember, these are not comparable.