PROBLEM SET 6

PAUL SCHRIMPF

DUE: TUESDAY MARCH 17TH, 2015 AT THE START OF LECTURE

UNIVERSITY OF BRITISH COLUMBIA

ECONOMICS 326

Problem 1 (Combination of Wooldridge (2013) 8.1 and Stock and Watson (2009) review question 17.1)**:**

(i) Suppose there is homoskedasticity, $\text{Var}(\epsilon_i|x) = \sigma^2$, but you construct a 95% confidence interval for β_1 using the heteroskedasticity robust standard error. Would this confidence interval be valid (have the correct coverage probability) in large samples?

(ii) Suppose that there is not homoskedasticity, but you construct a 95% confidence interval for β_1 using the homoskedasticity-only standard error, $\sqrt{\frac{\sum_{i=1}^n \hat{\epsilon}_i^2}{\sum_{i=1}^n \tilde{x}_{1,i}^2}}$. Would this confidence interval be valid in large samples?

(iii) Which of the following are consequences of heteroskedasticity?

(a) The OLS estimates, $\hat{\beta}_j$, are inconsistent.

(b) The usual F statistics no longer has an F distribution.

(c) The OLS estimators are no longer BLUE.

Problem 2 (Wooldridge (2013) 8.4)**:** Using the data in GPA3.RAW, student athlete's term GPA was regressed on, average GPA in courses taken, past GPA (GPA prior to the current semester), total credit hours, SAT score, percentile in high school, female, and a dummy variable equal to one if the student's sport is in season during the fall for fall and second semester students. The results were

Variable	**Coefficient**	**Homoskedastic Std. Err.**	**Heteroskedastic Std. Err.**
Intercept	-2.12	(.55)	[.55]
Avg course GPA	0.900	(.175)	[.166]
Past GPA	0.193	(.064)	[.074]
Credit hours	.0014	(.0012)	[.0012]
SAT	.0018	(.0002)	[.0002]
H.S. %tile	.0039	(.0018)	[.0019]
Female	.351	(.085)	[.079]
In season	-.157	(.098)	[.080]

(i) Do average course GPA, past GPA, and credit hours have the expected estimated effects? Which of these variables are statistically significant at the 5% level? Does it matter which standard errors are used?

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

(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?

Problem 3 (Wooldridge (2013) 8.C2)**:** Code to get started with this problem.

(i) Use the data in hprice1.Rdata to obtain the heteroskedasticity-robust standard errors for the regression of *price* on *lotsize*, $sqrft$m and *bdrms*. Discuss any important differences with the usual standard errors.
(ii) Repeat part i for the regression of $\log(price)$ on $\log(lotsize)$, $\log(sqrft)$ and *bdrms*.
(iii) What does this example suggest about heteroskedasticity and the transformation (i.e. taking log) used for the dependent variable?

Problem 4: Suppose we have a bivariate regression model

$$y_i = \beta_0 + \beta_1 x_i + \epsilon_i$$

but $\mathrm{E}[\epsilon_i x_i] \neq 0$.

(i) Show that $\hat{\beta}_1$ is inconsistent so $\operatorname{plim} \hat{\beta}_1 \neq \beta_1$.
(ii) Suppose we observe a variable z_i such that $\mathrm{E}[\epsilon_i | z_i] = 0$ and $\mathrm{Cov}(x, z) \neq 0$. Show that

$$\hat{\beta}_1^z = \frac{\sum_{i=1}^n (z_i - \bar{z}) y_i}{\sum_{i=1}^n (z_i - \bar{z}) x_i}$$

is a consistent estimator for β_1.

References

Stock, J.H. and M.W. Watson. 2009. "Introduction to Econometrics, 2/E." *Instructor* :10.
Wooldridge, J.M. 2013. *Introductory econometrics: A modern approach*. South-Western.