

Quantitative Method

Assignment 10

Due January 10, 2006

1. Consider the multiple regression of y on K variables X . Prove that $E(\hat{\beta}'\hat{\beta}) = \beta'\beta + \sigma^2 \sum_{k=1}^K (1/\lambda_k)$ where $\hat{\beta}$ is the vector of the OLS estimator and λ_k is an eigenvalue of $X'X$.

2. Use the data GPA2.xls for this exercise.

- (i) Consider the equation

$$\begin{aligned} colgpa = & \beta_0 + \beta_1 hsize + \beta_2 hsize^2 + \beta_3 hsperc + \beta_4 sat \\ & + \beta_5 female + \beta_6 athlete + u, \end{aligned}$$

where *colgpa* is cumulative college grade point average, *hsize* is size of high school graduating class, in hundreds, *hsperc* is academic percentile in graduating class, *sat* is combined SAT score, *female* is binary gender variable, which is one for female, and *athlete* is a binary variable, which is one for student-athletes. What are your expectations for the coefficients in this equation? Which ones are you unsure about?

- (ii) Estimate the equation in part (i) and report the results in the usual form. What is the estimated GPA differential between athletes and nonathletes? Is it statistically significant?
 - (iii) Drop *sat* from the model and re-estimate the equation. Now, what is the estimated effect of being an athlete? Discuss why the estimate is different than that obtained in part (ii).
 - (iv) In the model from part (i), allow the effect of being an athlete to differ by gender and test the null hypothesis that there is no ceteris paribus difference between women athletes and women nonathletes.
 - (v) Does the effect of *sat* on *colgpa* differ by gender? Justify your answer.
3. Use the data in SLEEP75.xls to estimate the following sleep equation:

$$\text{Sleep} = \beta_0 + \beta_1 totwrk + \beta_2 educ + \beta_3 age + \beta_4 age^2 + \beta_5 yngkid + \beta_6 male + u.$$

- (i) Write down a model that allows the variance of u to differ between men and women. The variance should not depend on other factors.
- (ii) Estimate the parameters of the model for heteroskedasticity. (You have to estimate the *sleep* equation by OLS, first, to obtain the OLS residuals.) Is the estimated variance of u higher for men or for women?
- (iii) Is the variance of u statistically different for men and for women?
- (iv) Assume the variance of u to differ between men and women. Re-estimate the model to incorporate this information.