Econ 466 Exam II - Answer Key

1. $GPA = \beta_0 + \beta_1 hsize + \beta_2 SAT + \beta_3 Fem + \beta_4 Ath + u$

a) On average females have a .23 higher GPA than men, holding other factors fixed 5
On average athletes have a .03 higher GPA than non-athletes, holding other factors fixed. 5

b) SAT · Fem is the additional boost to GPA due to SAT score for women. In this model (III) women's SAT scores affect college GPA, on average, less than a man's SAT score. 5

$H_0: \beta_{SAT \cdot Fem} = 0$ Need to use F-test since no standard errors are given.

$H_1: \beta_{SAT \cdot Fem} \neq 0$

$\alpha = .05$

4 $$F = \frac{(RSS_R - RSS_{UR})/1}{RSS_{UR}/(4137-6)} = \frac{(1435.577 - 1435.348)(4131)}{1435.348}$$

$$F = \frac{(.229)(4131)}{1435.348} = \frac{945.999}{1435.348} = .659073$$

$F = t^2 \quad t = .811833 \quad C_\alpha = 1.96$

$F_\alpha = 3.84$ Fail to reject - statistically insignifica[nt]

4 The advantage of model III over model I is that model III allows the presence of a more general relationship b/w SAT scores and college GPA. Model III doesn't make the tenuous assumption that men & women's SAT scores have equivalent effects on college GPA.

c) The advantage of model IV over model I is that we can now provide a more intuitive relationship b/w number of students in a school and a students GPA. With relatively few students each one gets a lot of attention and is well schooled - resulting in a high GPA. As the size of a school increases each student gets less and less attention until he/she receives so little that it makes learning more difficult, causing GPA to decrease.

4

Optimal hsize

Model IV: $hsize^{opt} = \left|\frac{\beta_{hsize}}{2\beta_{hsize^2}}\right| = \left|\frac{.023222}{2(-.00645)}\right| = 1.80$

4 2 2

which is 180 students

Model V: $hsize^{opt} = \left|\frac{\beta_{hsize}}{2\beta_{hsize^2}}\right| = \left|\frac{.0235}{2(-.00648)}\right| = 1.81$

about 181 students

d) $H_0: \beta_{Ath \cdot fem} = \beta_{SAT \cdot fem} = \beta_{hsize^2} = 0$ $\alpha = .05$

H_1: Not H_0:

$$F = \frac{(RSS_R - RSS_{UR})/2}{RSS_{UR}/(n-k-1)} = \frac{(1435.577 - 1432.610)/3}{1432.61/(4137-8)}$$

4

$$F = \frac{(2.967)(4129)}{(1432.61)(3)} = \frac{12250.743}{4297.83} = 2.8505$$

$F_{.05}(3, \infty) = 2.60$

reject null, parameters are jointly significant

e) Describe steps of the RESET test

For general linear model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + U$

5

1. Run OLS to model
2. Find $\hat{Y}$ and create variables $\hat{Y}^2$, $\hat{Y}^3$
3. Create new model $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k + \gamma_1 \hat{Y}^2 + \gamma_2 \hat{Y}^3 + U$
4. Run OLS to model in 3
5. Perform F-test on $H_0: \gamma_1 = \gamma_2 = 0$ against H_1: not H_0
   F is distributed $F(2, n-k-3)$

6.3 in wooldridge

2.

$$\widehat{RD} = \underset{(0.424)}{2.613} + \underset{(0.00014)}{0.00030}\ sales - \underset{(0.000000037)}{0.00000007}\ sales^2$$

$n = 32 \quad R^2 = 0.1484$

a) Find optimal sales to answer question.

(6)

$$sales^{opt} = \left| \frac{.0003}{2(-.00000007)} \right| = 21,428.57$$

from b) we know sales are in millions so this says that after sales reaches 21.42857 billion dollars there is a negative marginal effect on R&D intensity

b) Intuitive: I would keep the quadratic term for two reasons: (1) the form is more general, allowing for the level of R&D to depend on sales, (2) This form makes sense because when firms are small they cannot put as much into R&D and after firms become "large" R&D intensity is hard to maintain or even increase due to the fact that RD intensity $= \frac{\text{R\&D expenditure}}{\text{total expenditure}}$

(3)

Statistical: A one-sided test $(H_0: \beta_{sales^2} = 0,\ H_1: \beta_{sales^2} < 0)$ yields a t-statistic of -1.89184 with $df = 29$ the critical value for $\alpha = .05$ is 1.699 so we would reject the null and claim that $sales^2$ is statistically significant.

(3)

c) let sales bil = sales / 1000

New equation is

$\widehat{RD} = 2.613 + .3 \text{ salesbil} - .0070 \text{ salesbil}^2$
$(.429) \quad (.14) \quad\quad (.0037)$

10

$n = 32 \quad R^2 = .1484$

The intercept doesn't change because there is no origin change and R^2 is invariant to scale changes.

d) The second equation is nicer because there are fewer zeros for the coefficients.

3

3. $S_i = \beta_0 + \beta_1 Y + \beta_2 A + U$

S - sales in state
Y - State GDP
A - State Advertising expenditure

a) we suspect Heteroskedasticity

10

1. Estimate original model and obtain residuals, $\hat{u}$
2. square residuals $\hat{u}^2$
3. New model $\hat{u}^2 = \gamma_0 + \gamma_1 Y + \gamma_2 A + \gamma_3 A^2 + \gamma_4 Y^2 + \gamma_5 YA + \varepsilon$
4. OLS to model in 3
5. $H_0: \gamma_1 = \gamma_2 = \gamma_3 = \gamma_4 = \gamma_5$, H_1: Not H_0
   F-test - distributed as $F(5, n-k-6)$ $\alpha = .05$
   if we reject then we have evidence for heteroskedasticity

b) Assume $V(u) = \sigma^2 P_t^2$ | describe step by step how to obtain estimates that are corrected for heteroskedasticity

we can do one step here since the form of heteroskedasticity is know

(10) ① OLS to model $S^* = \beta_0 \left(\frac{1}{P}\right) + \beta_1 Y^* + \beta_2 A^* + u^*$

$S^* = S/P$, $Y^* = Y/P$, $A^* = A/P$, $u^* = u/P$

$E(u^*) = \frac{1}{P} E(u) = 0$

$V(u^*) = \frac{1}{P^2} v(u) = \frac{\sigma^2 P^2}{P^2} = \sigma^2$ (homoskedastic)

c) If we ignore heteroscedasticity then OLS is unbiased and consistent, but no longe efficient. The reason for this is that if we know there is heteroskedasticity
(5) there is information that we are ignoring when we do OLS. Thus we can always benefit from additional information

d) $S = \beta_0 + \beta_1 A + u$ where $v(u) = \sigma^2$ but A is measured with error. Explain whether $\hat{\beta}_1$ (OLS estimator) is
(10) consistent or not.

The answer to this depends on whether we make the assumption that if $A = A^* + e$, $\operatorname{Cov}(A, e) = 0$ then OLS provides a consistent estimate of $\hat{\beta}_1$. If we assume $\operatorname{Cov}(A^*, e) = 0$ then OLS is not consistent.