Economics 466
Exam II, Spring 2004

Total Points: 100, Time 1.5 hrs

Answer all questions. Note that each question has different weight. **Good Luck!**

1. Consider the following model of education determinants

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

where *GPA* is cumulative college *GPA*, *hsize* is the size of the student's high school graduating class, in 100s, *SAT* is the student's SAT score, *Fem* is a dummy variable equal to one if the student is a female, and *Ath* is a binary variable set to one for student athletes. The following table list several variants of the model. Please note there are 4137 observations. .

	Model I	Model II	Model III	Model IV	Model V
Int	0.458529	0.460781	0.411691	0.413463	0.372752
hsize	-0.02137	-0.02126	-0.02142	0.023222	0.0235
SAT	0.002085	0.002083	0.002129	0.002075	0.002114
Fem	0.232824	0.23086	0.346687	0.232732	0.334425
Ath	0.031671	0.018146	0.035301	0.036592	0.027365
*Ath * Fem*	---	0.05534	---	---	0.051353
*SAT * Fem*	---	---	-0.00011	---	-0.0001
hsize²	---	---	---	-0.00645	-0.00648
RSS	1435.577	1435.475	1435.348	1432.925	1432.610

Use the above table to answer the following questions.

a. Use Model I to estimate the mean GPA differential between (i) athletes and non-athletes, and (ii) female and male students. **(5+5 points)**
b. Interpret the *SAT*Fem* coefficient in Model III. Test the statistical significance of the coefficient? What is the advantage of using Model III over Model I? **(5+4+4 points)**
c. What is the advantage of using Model IV over Model I? Find the value of *hsize* that maximizes GPA in Models IV and V? **(4+4 points)**
d. Perform a test of functional form for Model I vs. Model V. **(4 points)**
e. Describe the steps of the RESET test. **(5 points)**

2. The following are the estimation results for a model that determines the effect of a company's sales on Research and Development intensity. The numbers in parentheses are the standard errors.

$$\widehat{RD} = \underset{(0.429)}{2.613} + \underset{(0.00014)}{0.00030 sales} - \underset{(0.0000000037)}{0.000000007 sales^2}$$
$$n = 32, \quad R^2 = 0.1484$$

   a. At what point does the marginal effect of *sales* on *RD* become negative? **(6 points)**
   b. Would you keep the quadratic term in the model? Answer both from a statistical perspective as well as from an intuitive approach. **(6 points)**
   c. Let $salesbil = sales/1000$. Rewrite the estimated equation with *salesbil* and *salesbil²* in the estimated equation. Be sure to report the new standard errors as well as the new R^2. **(10 points)**
   d. Which equation is easier on the eyes? **(3 points)**

3. Consider the model $S_t = \beta_0 + \beta_1 Y_t + \beta_2 A_t + u_t$ where S is the sales of a firm in state t, Y is the GDP of state t, and A is the amount of advertising expenditures for the firm in state t.

   a. You suspect the error term is heteroskedastic with respects to the population of the states. Describe step by step how you would go about testing this proposition. Be sure to include your model, null and alternative hypotheses, as well as any test statistics you will create including the degrees of freedom, and your acceptance criterion. **(10 points)**
   b. Assume that the heteroskedasticity is of the form $\sigma_u^2 = \sigma^2 P_t^2$. Describe step by step how you would obtain estimates corrected for heteroskedasticity. **(10 points)**
   c. Suppose your test results confirm that there is heteroskedasticity in your model but you decide to ignore it. Are your OLS estimates unbiased, consistent, and/or efficient? **(5 points)**
   d. Now consider the model $S_t = \beta_0 + \beta_1 A_t + u_t$ and assume that u_t is homoskedastic but A_t is measured with error. Explain whether the OLS estimator of β_1 will be consistent or not. **(10 points)**