Homework 5

Due in class on Thursday, February 24, 2011.

Problem 1. (2 points) Suppose you use Gretl to estimate the following model using the finaid.txt dataset from the course homepage.

\[ \text{FINAID}_i = \beta_0 + \beta_1 \text{HSRANK}_i + \beta_2 \text{PARENT}_i + \beta_3 \text{MALE}_i + \epsilon_i. \]

Use the regression output below to answer the questions that follow.

```
Model 1: OLS estimates using the 50 observations 1-50
Dependent variable: FINAID

VARIABLE COEFFICIENT STDERROR T STAT P-VALUE
const 9813.02 1743.10 5.630 <0.00001 ***
HSRANK 83.2612 20.1479 4.132 0.00015 ***
PARENT -0.342754 0.0315054 -10.879 <0.00001 ***
MALE -1570.14 784.297 -2.002 0.05120 *

Mean dependent var 11676.26 S.D. dependent var 5365.233
Sum squared resid 3.32e+08 S.E. of regression 2686.575
R-squared 0.764613 Adjusted R-squared 0.749262
F(3, 46) 49.80764 P-value(F) 1.72e-14
Log-likelihood -463.6635 Akaike criterion 935.3270
Schwarz criterion 942.9751 Hannan-Quinn 938.2394
```

a. Use the reported \( t \) ratios to determine whether you would reject \( H_0 \) or fail to reject \( H_0 \) in the following tests. Show the calculation in the rejection rule that led you to reject or not.

(a) \( H_0 : \beta_1 = 0 \) vs \( H_A : \beta_1 \neq 0 \) with \( \alpha = 0.10 \).
(b) \( H_0 : \beta_2 = 0 \) vs \( H_A : \beta_2 \neq 0 \) with \( \alpha = 0.05 \).
(c) \( H_0 : \beta_3 = 0 \) vs \( H_A : \beta_3 \neq 0 \) with \( \alpha = 0.01 \).

b. Now, use the reported \( p \) values to determine whether you would reject \( H_0 \) or fail to reject \( H_0 \) in the following tests. Show the calculation in the rejection rule that led you to reject or not.

(a) \( H_0 : \beta_1 = 0 \) vs \( H_A : \beta_1 \neq 0 \) with \( \alpha = 0.01 \).
(b) \( H_0 : \beta_3 = 0 \) vs \( H_A : \beta_3 \neq 0 \) with \( \alpha = 0.10 \).
(c) \( H_0 : \beta_3 = 0 \) vs \( H_A : \beta_3 \neq 0 \) with \( \alpha = 0.05 \).
c. Calculate the 95% confidence interval for $\hat{\beta}_1$.

d. Calculate the 90% confidence interval for $\hat{\beta}_3$.

e. Gretl reports information about the $F$ test of overall significance. Write down $H_0$ and $H_A$ for this test for the regression you ran above.

f. For the $F$-test of overall significance, would you reject $H_0$ or fail to reject $H_0$ at the 10% level? At the 5% level? At 1%?

Problem 2. (2 points) (Studenmund, Exercise 6.4) Assume that you’ve been hired by the surgeon general of the United States to study the determinants of smoking behavior and that you estimate the following cross-sectional model based on data for all 50 states (standard errors in parentheses with $t$-ratios underneath):

$$
\hat{C}_i = 100 - 9.0 E_i + 1.0 I_i - 0.04 T_i - 3.0 V_i + 1.5 R_i
$$

where

- $C_i$ is the number of cigarettes consumed per day per person in the $i$-th state,
- $E_i$ is the average years of education for persons over 21 in the $i$-th state,
- $I_i$ is the average income in the $i$-th state (in thousands of dollars),
- $T_i$ is the tax per package of cigarettes in the $i$-th state (in cents),
- $V_i$ is the number of video ads against smoking aired on the three major networks in the $i$-th state,
- $R_i$ is the number of radio ads against smoking aired on the five largest radio networks in the $i$-th state.

a. Develop and test (at the 5-percent level) appropriate hypotheses for the coefficients of the variables in this equation.

b. Do you appear to have any irrelevant variables? Do you appear to have any omitted variables? Explain your answer.

c. Let’s assume that your answer above was yes to both. Which problem is more important to solve first—irrelevant variables or omitted variables? Why?

d. One of the purposes of running the equation was to determine the effectiveness of anti-smoking advertising on television and radio. What is your conclusion?
e. The surgeon general decides that tax rates are irrelevant to cigarette smoking and orders you to drop the variable from your equation. Given the following results, use our four specification criteria to decide whether you agree with her conclusion. Carefully explain your reasoning (standard errors in parentheses).

\[ \hat{C}_i = 101 - 9.1 E_i + 1.0 I_i - 3.5 V_i + 1.6 R_i \]

\[ \bar{R}^2 = 0.40 \quad N = 50 \]

**Problem 3.** (1 point) (Studenmund, Exercise 6.8) For each of the following situations, determine the sign of the expected bias introduced by omitting a variable:

a. In an equation for the demand for peanut butter, the impact on the coefficient of disposable income of omitting the price of peanut butter variable. *(Hint: Start by hypothesizing signs.)*

b. In an earnings equation for workers, the impact on the coefficient of experience of omitting the variable for age.

c. In a production function for airplanes, the impact on the coefficient of labor of omitting the capital variable.

d. In an equation for daily attendance at outdoor concerts, the impact on the coefficient of the weekend dummy variable (1 weekend) of omitting a variable that measures the probability of precipitation at concert time.