Professor Nagler **Spring, 2001**

G53.2127 : Quantitative Political Analysis II

Practice Final: Spring, 2001

You will be given all the tables that will be required to answer these tests, and you may use a calculator. If you get real confused by a problem, take your time. Good luck!

You must show all your work for all problems.

1) Say you believe that the true model determining the attrition rate (dropout rate) in school districts is given by:

$$attrit_i = \beta_0 + \beta_1 * wealth_i + \beta_2 * ppupexp_i + \beta_3 * phisp_i + \beta_4 * clssize_i + u_i \quad (1)$$

where

- $attrit_i$ denotes the dropout rate in the i^{th} district as a percentage of students enrolled
- $wealth_i$ denotes the total wealth the i^{th} district in thousands of dollars
- $ppupexp_i$ denotes the per pupil expenditure in the i^{th} district in thousand-dollars per pupil
- $phisp_i$ denotes the percentage of hispanic students in the i^{th} district
- $clssize_i$ denotes the average class-size in the i^{th} district

Now because you cannot get data on wealth, you are forced to estimate the model with this variable omitted:

$$attrit_i \;=\; \beta_{20} \;+\; \beta_{22} * ppupexp_i \;+\; \beta_{23} * phisp_i \;+\; \beta_{24} * clssize_i \;+\; v_i \qquad (2)$$

You run OLS and find that:

$$\hat{\beta_{20}} \;=\; 2.89 \quad \sigma_{\hat{\beta_{20}}} \;=\; 0.48$$
$$\hat{\beta_{22}} \;=\; -1.50 \quad \sigma_{\hat{\beta_{22}}} \;=\; 0.24$$
$$\hat{\beta_{23}} \;=\; 0.50 \quad \sigma_{\hat{\beta_{23}}} \;=\; 0.35$$
$$\hat{\beta_{24}} \;=\; 0.20 \quad \sigma_{\hat{\beta_{24}}} \;=\; 0.42$$

$$R^2 \;=\; 0.21$$
$$N \;=\; 1000$$

a) [**15 points**] You have a strong prior that β_1 is negative, and that *ppupexp* and *wealth* are positively correlated. Can you give a good guess as to whether the true population parameter β_2 is less than 0? Explain your answer carefuly. Write down an equation for the auxiliary regression you would want to run, and explain how the results of it would affect your answer. Is $\hat{\beta}_{22}$ an unbiased estimator of β_2? If not, write down an equation that would indicate the bias.

Now, for parts (b) through (e) assume that equation (2) – the equation you estimated generating β_{20}, β_{22}, β_{23}, and β_{24} – was correctly specified (i.e., ignore equation 1 for the rest of the problem), and that all of the Gauss-Markov assumptions are met. [**This means to ignore the omitted variable problem for parts (b) through (e)!!**]

b) [**5 points**] Using a one-tailed test, are you 90% certain that increases in the percentage of hispanic students in a district would lead to an increase in the attrition rate? Please include equations for H0 and H1. [**One-Tailed Test.**]

c) [**15 points**] Assume that you find out that there is a high correlation between the percentage of hispanc students in a district and class-size, and say that you would be content to show that the effect of either is significant at the 99% level. Is **either** *phisp* or $classize$ significant at the 99% level (this is a question about **joint significance**? Write down the null and research hypotheses involved here, as well as equations for any test-statistic you use.

Big Hint: You get an R^2 of .13 from the following regression

$$attrit_i \;=\; \beta_{30} \;+\; \beta_{32} * ppupexp_i \;+\; \vartheta_i \tag{3}$$

d) [**10 Points**] If per-pupil expenditure were measured as hundred-dollars spent per student, rather than thousand-dollars spent per student, what would its estimated coefficient β_{22} be in equation (2)? What would the t-statistic be?

e) [**10 Points**] If multicollinearity were present, and the correlation between the *phisp* and *ppupexp* were -.7, how would that affect your answer to (b)? Would your estimate of β_{23} be unbiased? Would your estimated standard-error ($\sigma_{\hat{\beta_{23}}}$) be correct?

3) Say you estimate the following model:

$$Wage_i = \alpha_0 + \beta_1 * education_i + \beta_2 * experience_i + \beta_3 * women_i + u_i \quad (4)$$

where

- *wage* is an individual's wage in dollars-per-hour
- *education* is the number of years of education an individual has
- *experience* is the number of years of experience an individual has at his/her present job
- *women* is a dummy variable coded 1 for female individuals and 0 for male individuals

a) [**5 points**] Assume that all of the Gauss-Markov assumptions are met. If $\beta_3 = -1.50$, what would be the difference in wages between a man and women with equal education and experience? [Be sure your answer includes a direction to it.]

Now say your estimates generate the following results:

$\hat{\alpha_0}$	=	2.00	$\sigma_{\hat{\alpha_0}}$	=	0.50
$\hat{\beta_1}$	=	0.50	$\sigma_{\hat{\beta_1}}$	=	0.10
$\hat{\beta_2}$	=	0.10	$\sigma_{\hat{\beta_2}}$	=	0.02
$\hat{\beta_3}$	=	−1.60	$\sigma_{\hat{\beta_3}}$	=	0.20

R^2	=	0.33
N	=	1000

Standard Error of the Regression = 1.50

b) [**5 points**] What is your best estimate of how much one's wage would increase by completing four years of high school? What is your best estimate of how much one's wage would increase by completing four years of college?

c) [**5 points**] If you were earning \$10 per hour with a job that you had 5 years of experience at, and were fired and had to start a job where you had zero years of experience, what would you expect your new wage to be?

4) [**15 points**] What are the properties of the OLS estimators?

5) [**10 points**] When is it meaningful to compare R^2 values across different regressions?

6) [**5 points**] What is the goal of good coding style?