Economics 113 Professor Spearot
Introduction to Econometrics
Fall 2007 – Final Exam
Name ________________________

Final Exam Version A – 80 Points

You must answer all the questions. The exam is closed book and closed notes. You may use calculators, but they must not be graphing calculators. Do not use your own scratch paper.

<u>You must show your work to receive full credit</u>

You have plenty of time to finish. Take your time and relax. And, have a safe and Happy Holiday!

1. You have four kids that that weigh 50, 60, 70, and 80 pounds. Their respective heights are 2, 3, 4, and 3 ft.

a. What is the covariance between height and weight? **(5 points)**

b. Suppose I estimate $Height = \beta_0 + \beta_1 Weight + u$ <u>using a different sample</u> (not your answer from a). The sample covariance of *Height* and *Weight* is 10. The sample variance of *Weight* is 2. What is the estimate of β_1? **(5 points)**

2. You wish to predict the effects of education, experience, and tenure on wage outcomes. Specifically, you estimate the following specification:

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 tenure + \beta_3 \text{exper} + u$$

The results from running this regression are below:

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  3,   931) =   56.97
       Model |  25.6953278     3  8.56510927           Prob > F      =  0.0000
    Residual |  139.960966   931  .150334013           R-squared     =  0.1551
-------------+------------------------------           Adj R-squared =  0.1524
       Total |  165.656294   934  .177362199           Root MSE      =  .38773
------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0748638   .0065124     xxxx   xxxxx     xxxxxxxx    xxxxxxxx
      tenure |   .0133748   .0025872     xxxx   xxxxx     xxxxxxxx    xxxxxxxx
       exper |   .0153285   .0033696     xxxx   xxxxx     xxxxxxxx    xxxxxxxx
       _cons |   5.496696   .1105282     xxxx   xxxxx     xxxxxxxx    xxxxxxxx
```

a. Do a two sided t-test at the 5% level to determine if experience **(exper)** is a statistically significant determinant of the log wage. Please state the null and alternative hypotheses, and interpret the result. **(5 points)**

b. Perhaps you've heard the phrase, "I was trained in the school of hard knocks…it's just as good as school". Write down the hypothesis that states that education (**educ**) and experience (**exper**) have equal effects on the log wage. Also provide a two-sided alternative. **(2 Points)**

c. Please manipulate the regression equation in (a) so that your null hypothesis in (b) can be tested using a t-test. <u>Show your work!!</u> **(8 Points)**

Suppose that I adjust the specification in (a) as follows,

$$\log(wage) = \beta_0 + \beta_1 educ + \beta_2 tenure + \beta_3 \text{exper} + \beta_4 \text{exper}^2 + \beta_5 age + \beta_6 age^2 + u$$

where age is the age of the respondent (in years).

```
      Source |       SS       df       MS              Number of obs =     935
-------------+------------------------------           F(  6,   928) =   29.17
       Model |  26.2843766     6  4.38072944           Prob > F      =  0.0000
    Residual |  139.371918   928  .150185256           R-squared     =  0.1587
-------------+------------------------------           Adj R-squared =  0.1532
       Total |  165.656294   934  .177362199           Root MSE      =  .38754
------------------------------------------------------------------------------
       lwage |      Coef.   Std. Err.      t    P>|t|     [99% Conf. Interval]
-------------+----------------------------------------------------------------
        educ |   .0703793   .0069319    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
      tenure |   .0127232   .0026224    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       exper |   .0230826    .013667    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
      exper2 |  -.0005216      .0006    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
         age |   .0069779   .1034685    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
        age2 |   .0000507   .0015552    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
       _cons |    5.26504   1.697521    xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx
```

d. Please construct a 99% confidence interval for the coefficient on tenure, β_2. **(5 points)**

e. What is the p-value for the linear experience term (**exper**)? Please draw the distribution under the null and compute the two-sided p-value (please give the range and an approximate value). **(5 points)**

f. Please test the hypothesis that adding *age*, age^2, and $exper^2$ makes no difference in predicting the log wage. That is, please test whether these variables are jointly insignificant. Do this at the 95% level, stating the null and alternative hypotheses. **(10 points)**

g. Professor Spearot will be 29 in January. He feels old, though is shamelessly hoping that his wage makes up for it. Assuming that he gains no additional experience and no additional tenure, what is the predicted effect on the wage in going from age 28 to age 29? Interpret briefly. **(5 points)**

h. At what experience level is the wage maximized? Show your work! **(5 points)**

3. Suppose that I run the following regression predicting the effects of candidate expenditures and other factors on election outcomes:

$$voteA = \beta_0 + \beta_1 \text{lexpendA} + \beta_2 \text{lexpendB} + \beta_3 prtystr + u$$

In the regression equation, lexpendA is the log expenditures of candidate A, lexpendB is the log expenditures of candidate B, and *prtystr* is the relative strength of party A The results from running this regression are below:

```
. regress  votea  lexpenda  lexpendb  prtystra

      Source |       SS       df       MS              Number of obs =     173
-------------+------------------------------           F(  3,   169) =  215.23
       Model |  38405.1089     3   12801.703           Prob > F      =  0.0000
    Residual |  10052.1396   169  59.4801161           R-squared     =  0.7926
-------------+------------------------------           Adj R-squared =  0.7889
       Total |  48457.2486   172  281.728189           Root MSE      =  7.7123

------------------------------------------------------------------------------
       votea |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    lexpenda |   6.083316     .38215    15.92   0.000     5.328914    6.837719
    lexpendb |  -6.615417   .3788203   -17.46   0.000    -7.363247   -5.867588
    prtystra |   .1519574   .0620181     2.45   0.015     .0295274    .2743873
       _cons |   45.07893   3.926305    11.48   0.000     37.32801    52.82985
```

a. Are the variables in the model (lexpendA, lexpendB, and *prtystr*) a significant determinant of the votes candidate A receives? Please test this hypothesis at the 95% level. Write the null hypothesis, the alternative, and briefly interpret the result. **(5 Points)**

b. Suppose that Candidate A is Mike Gravel and Candidate B is Mitt Romney. Mitt spends much more than Mike. I want to predict the outcome of this hypothetical race, and produce confidence intervals for this prediction. Please write an equation for the prediction if lexpendA=1 (Mike), lexpendB=10 (Mitt), and prtystra=50 (citizens hate both parties equally!!). <u>Derive</u> a new estimating equation to generate the prediction and its standard error. Please also write the necessary commands to generate any new variables in STATA. **(10 Points)**

c. Suppose that as campaign expenditures rise, unobserved factors affecting voting outcomes tend to become more variable. What kind of problem is this? What should be done about it? **(5 Points)**

d. Suppose that candidate A spends more money because he/she has better ideas, and better ideas get you more votes. What assumption is violated in our current model? In what direction is β_1 biased? **(5 Points)**

Extra Credit:

Look back at the full model in problem #2. Age discrimination is perceived to be commonplace in American society. Age discrimination occurs if there exists an age above which wages go down, independent of other attributes (education, experience, tenure). Using the results in Problem #2, briefly discuss whether there is evidence of age discrimination. (**2 points)**

Helpful formulas

$$\hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2 \qquad \hat{\sigma}_{xy} = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y) \qquad \hat{\rho}_{xy} = \frac{\hat{\sigma}_{xy}}{\hat{\sigma}_x \hat{\sigma}_y}$$

$$\hat{\beta}_0 = \hat{\mu}_y - \hat{\beta}_1 \hat{\mu}_x \qquad \hat{\beta}_1 = \frac{\sum_{i=1}^{n}(x_i - \hat{\mu}_x)(y_i - \hat{\mu}_y)}{\sum_{i=1}^{n}(x_i - \hat{\mu}_x)^2}$$

$$R^2 = 1 - \frac{SSR}{SST} \qquad SSR = \sum_{i=1}^{n}(\hat{u}_i)^2 \qquad SST = \sum_{i=1}^{n}(y_i - \hat{\mu}_y)^2$$

$$R^2 = 1 - \frac{\frac{SSR}{n-k-1}}{\frac{SST}{n-1}} \qquad \text{Fstat} = \frac{\frac{\text{SSR}_{\text{R}} - SSR_{UR}}{q}}{\frac{SSR_{UR}}{n-k-1}}$$