Econometric Methods I
Homework 1: Due 1 February 2012

1. We are interested in studying salary differences between two professions. Our sample consists of 60 individuals from profession A and 40 from profession B. The average salary in B is 1. The average salary in the entire sample is 0.85. If we compute a least squares regression of salaries on a constant and a variable that is simply 0 for all individuals in A and 1 for those in B, what will be the constant and the slope? What is R^2 in the regression?

2. Suppose a regression (with only one slope variable) is fit without a constant term and the computer program reports two measures of R^2:

$$R_1^2 = 1 - \frac{\sum e_t^2}{\sum (y_t - \bar{y})^2}$$

$$R_2^2 = \frac{b^2 \sum (x_t - \bar{x})^2}{\sum (y_t - \bar{y})^2}$$

where b is the estimated coefficient on the slope variable. Are these measures identical? Which of them do you think is the appropriate one?

3. Suppose **b** is the least squares coefficient vector in the regression of **Y** on **X** and **c** is any other **K** x 1 vector. Prove that the difference in the two sums of squared residuals is

(y-Xc)’ (Y – Xc) – (Y- Xb)’(Y – Xb) = (c – b)’X’X(c – b).

4 Consider the least squares regression of **Y** on **K** variables (and a constant),**X.** Consider an alternative set of regressors, **Z** = **XP**, where **P** is a nonsingular matrix. Thus, each column of **Z** is a mixture of some of the columns of **X.** Prove that the residual vectors in the regressions of **Y** on **X** and **Y** on **Z** are identical. What relevance does this have to the question of changing the fit of regression by changing units of measurement of the independent variables?

5. In the least squares regression of **y** on a constant and **X**, show that the regression coefficients on **X** can also be computed by (i) transforming **y** to deviations from the mean (ii) transform each column of **X** to deviations from the respective column mean and (iii) regress the transformed **y** on the transformed **X** without a constant. Do we get the same result if we only transform **y**? What if we only transform **X**?