Econometrics I
Midterm Exam
Fall 1996
Total Points: 100
Time: 1 hr. and 15 min.

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

1.(50 points) A sample of 57 quarterly observations on household expenditures for natural gas and electricity is to be used to estimate a linear demand function for gas. The dependent variable (Y) is quantity of gas consumed, while the independent variables are average electricity cost (X_2) and average gas cost (X_3) per quarter. The relevant sample moments are as follows:

$$\frac{1}{n}x'x = \frac{1}{n}\begin{pmatrix} x_2'x_2 & x_2'x_3 \\ x_3'x_2 & x_3'x_3 \end{pmatrix} = \begin{pmatrix} 350 & 170 \\ 170 & 320 \end{pmatrix}, \quad \frac{1}{n}x'y = \frac{1}{n}\begin{pmatrix} x_2'y \\ x_3'y \end{pmatrix} = \begin{pmatrix} 0 \\ -100 \end{pmatrix}, \quad \frac{1}{n}y'y = 60.$$

(a) Since electricity cost is uncorrelated with gas consumption ($x_2'y = 0$), can it be ignored from the multiple regression? Demonstrate your answer by comparing b_{y2}, the slope coefficient of X_2 in the simple regression Y_2 on X_2, to b_2, the slope coefficient of X_2 in the multiple regression Y_2 on X_2 and X_3.

(b) Find the standardized coefficients (beta coefficients) b_2^* and b_3^*. Compare these with b_2 and b_3. Calculate the standard errors of b_2^*, b_3^*, b_2, and b_3.

(c) Show that $R_*^2 = R^2$ where R_*^2 is the R^2 in the standardized regression.

(d) Test the hypothesis of "no regression" at the 5% level of significance.

(e) Test the hypothesis that $\beta_2 + \beta_3 = 1$.

(f) Suppose that you know a priori that $\beta_2 + \beta_3 = 1$. Obtain the OLS estimates of β_2 and β_3 using this restriction.

2. (32 points) Suppose you have observations on $N = 100$ Canadians for the following regression:

$$Y = \beta_1 Z_1 + \beta_2 Z_2 + \cdots + \beta_6 Z_6 + X\gamma + u, \qquad (1)$$

where Y = log wage,

Z_1 = intercept,

$Z_2 = 1$ if female, $= 0$ otherwise,

$Z_3 = 1$, if lives in Ontario, $= 0$, otherwise

$Z_4 = 1$, if lives in Quebec, $= 0$, otherwise

$Z_5 = Z_2 Z_3$,

$Z_6 = Z_2 Z_4$,.

X is a $N \times (K - 6)$ matrix of relevant socioeconomic-demographic variables. *State clearly and explicitly the null and two-sided alternative hypotheses for the following tests.* State whether a t-test or a F-test is appropriate and the corresponding degrees of freedom.

(a) Test whether (everything else being equal) a male living outside of Ontario and Quebec has the same wage rate as a female living outside of Ontario and Quebec.

(b) Test whether (everything else being equal) females have the same wage rates as males.

(c) Test whether (everything else being equal) wage rates are the same in all regions of Canada (not necessarily the same for males and females).

(d) Test whether (everything else being equal) females in Quebec have the same wage rates as males in Ontario.

3. (18 points) You are given data on the gross domestic product, X, labor input, L, and capital input, K, over the period 1929 to 1978. You decide to estimate a Cobb-Douglas production function:

Econometrics I
Midterm Exam
Fall 1997
Total Points: 100
Time: 1 hr. and 15 min.

Answer all parts. Note that each part has different weight. Good Luck!!

Data on a three-variable problem yield the following results:

$$X'X = \begin{pmatrix} 33 & 0 & 0 \\ 0 & 40 & 20 \\ 0 & 20 & 60 \end{pmatrix}, \quad X'Y = \begin{pmatrix} 132 \\ 24 \\ 92 \end{pmatrix}, \quad y'y = 150.$$

Note the partioned matrix $X_2'X_2 = x_2'x_2$ since $\bar{X}_2 = \bar{X}_3 = 0$.

(a) Find the OLS estimate of β_2 and β_3 (call them b_2 and b_3) for the model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u. \tag{1}$$

(b) Calculate the standard errors of b_2 and b_3. Find R^2.

(c) Test $\beta_2 = 0$ at the 5% level of significance. Test $\beta_2 = \beta_3 = 0$ at the 1% level of significance.

(d) Test $\beta_2 + \beta_3 = 1$ at the 5% level of significance.

(e) Consider the model in (1) with the restriction $\beta_2 + \beta_3 = 1$. Obtain the restricted least squares estimators of β_2 and β_3.

(f) Calculate the standard error of b_2 and R^2 from the restricted model in part (e).

(g) Compare standard error of b_2 and R^2 in the restricted and the unrestricted model. Comment on the differences.

(h) Assume that there is no intercept in the model. Now estimate β_2 and β_3 (call them $\tilde{b}_2$ and $\tilde{b}_3$). Compare $\tilde{b}_2$ and $\tilde{b}_3$ with b_2 and b_3. Comment on the results.

(i) Calculate the standard errors of $\tilde{b}_2$ and $\tilde{b}_3$. Compare them with those of b_2 and b_3 obtained in part (b). Any comment??

(j) Rewrite the model in (1) as

$$Y = \beta_1 + \beta_2^* X_2^* + \beta_3^* X_3^* + u, \tag{2}$$

where $X_2^* = x_2/s_2$ and $X_3^* = x_3/s_3$, s_2 and s_3 being the sample standard deviations of X_2 and X_3. Show that the OLS estimators of the slope coefficients in (2) (call them b_2^* and b_3^*) are related to b_2 and b_3 as $b_2^* = b_2 s_2$ and $b_3^* = b_3 s_3$. (Hint: you are not supposed to calculate b_2^* and b_3^* directly. The above relation can be established algebraically).

(k) Refer to the model in (2). Show that R^2 in this model is the same as the one obtained in part (b).

Econometrics I
Midterm Exam 1
Fall 1998
Total Points: 100
Time: 1 hr. and 20 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Data on a three-variable problem yield the following results:

$$X'X = \begin{pmatrix} 10 & 30 & 40 \\ 30 & 92 & 119 \\ 40 & 119 & 163 \end{pmatrix}, \quad X'Y = \begin{pmatrix} 20 \\ 59 \\ 88 \end{pmatrix}, \quad Y'Y = 80.$$

(a) (**10 points**) Find the OLS estimate of β_1, β_2 and β_3 (call them b_1, b_2 and b_3) for the model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u. \tag{1}$$

(b) (**10 points**) Calculate the standard errors of b_2 and b_3. Find R^2.

(c) (**10 points**) Test $\beta_2 = 2$ at the 5% level of significance. Test $\beta_2 = \beta_3 = 0$ at the 1% level of significance.

(d) (**10 points**) Consider the model in (1) with the restriction $\beta_2 + \beta_3 = 5$. Obtain the restricted least squares estimators of β_2 and β_3.

2. The model

$$Y = \beta_1 + \beta_2 E_2 + u \tag{2}$$

is estimated by OLS where E_2 is a dummy variable for college graduates (the omitted group is the high school graduates).

(a) (**10 points**) Show that the OLS estimates of β_1 and β_2 are

$$b_1 = \bar{Y}_1, b_2 = \bar{Y}_2 - \bar{Y}_1,$$

where $\bar{Y}_1$ and $\bar{Y}_2$ are the mean earning (Y) for the high school and college graduates, respectively. Assume that there are n_1 high school and $n_2 = n - n_1$ college graduates in your sample.

(b) (**10 points**) Extend the above model in such a way that you can test the hypothesis that there is racial difference in earnings between high school and college graduates. (Assume at least two racial groups). State explicitly how you would test such an hypothesis in terms of restrictions in the parameters of the model.

(c) (**10 points**) Consider another extension to separate earnings of males from females. Show how you can test hypotheses relating to gender and racial differences in earnings.

3. Consider two K variable regression models

$$Y = X\beta + u \tag{3}$$

and

$$Y = Z\gamma + v \tag{4}$$

where $Z = XP$, P being a $K \times K$ non-singular matrix. Both the models satisfy the standard assumptions of the classical regression model.

(a) (**5 points**) Find the OLS estimates of β and γ. Are they identical?

(b) (**5 points**) Show that the OLS residual vector from the model in (3) is exactly the same as the one in model (4).

(c) (**5 points**) Show that R^2 from model (3) is the same as R^2 from model (4).

(d) (**3 points**) What lesson did you learn from this exercise? Explain in a few lines.

4. (**12 points**) You estimated the K variable regression model

$$Y = X\beta + u \tag{5}$$

and your friend estimated it by adding one extra variable (Z) in it so that he has $K + 1$ regressors in his model. Assuming that the variance of the error term in both models is the same, show that $V(b) - V(b^*)$ is a positive semi-definite matrix when $V(b)$ and $V(b^*)$ are the variance-covariance matrices corresponding to the OLS estimators of the coefficient of X variables in your model (equation (5)) and your friend's model (which differs from your model in terms of the extra variable Z).

Econometrics I
Midterm Exam 1
Spring 1998
Total Points: 100
Time: 1 hr. and 15 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Consider the following data for the two-variable regression model

$$Y = \alpha + \beta X + u \tag{1}$$

which satisfies all the standard assumptions of the classical regression model.

$$n = 10, \sum X = 30, \sum Y = 20, \sum X^2 = 92, \sum Y^2 = 50, \sum XY = 64.$$

(a) Find the OLS estimators of α and β. Calculate R^2 and test the hypothesis that $\beta = 0$ at the 5% level of significance.

(b) Show that a change in the unit of X (X being multiplied by 10, for example) will reduce the OLS estimator of the slope coefficient by a factor of 10.

(c) Show that R^2 is not affected due the change in unit of X.

(d) How is the slope coefficient (γ) of the following standardized regression

$$Y^* = \gamma X^* + \epsilon \tag{2}$$

related to β in (1)? What is the relationship between R^2 from (1) and (2)?

2. In the following earning (EARN) regression with dummy variables

$$EARN = \alpha + \beta_F Female + \beta_B Black + \beta_H Hispanic + u \tag{3}$$

(a) Interpret the coefficients β_F, β_B, and β_H.

(b) Can you test whether there is any racial difference in earnings from this model? Explain.

(c) Now add interactions of Female dummy with Black and Hispanic dummies so that the model becomes

$$\begin{aligned} EARN =& \alpha + \beta_F Female + \beta_B Black + \beta_H Hispanic + \beta_{FB} Female * Black \\ &+ \beta_{FH} Female * Hispanic + u. \end{aligned} \tag{4}$$

(d) Interpret the interaction terms β_{FB} and β_{FH}. How can you test the hypothesis that there is no racial difference in earnings using model (4)? Explain. How about testing no gender difference in earnings?

3. Suppose that you have estimated the parameters of the multiple regression model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \tag{5}$$

by OLS method. Denote the OLS residuals by $e_i, i = 1, \ldots, n$ and the predicted values by $\hat{Y}_i$.

(a) What is the R^2 of the regression of e_i on a constant, X_2 and X_3?

(b) If you regress Y on a constant and $\hat{Y}$, what are the estimated intercept and slope coefficient? What is the relationship between R^2 of this regression and R^2 of the original regression in (5)?

(c) If you regress Y on a constant and e, what are the estimated intercept and slope coefficient? What is the relationship between R^2 of this regression and R^2 of the original regression in (5)?

Econometrics I
Midterm Exam 1
Fall 1999
Total Points: 100
Time: 1 hr. and 20 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Data from a sample of 20 observations yield the following sample moments:

$$\frac{1}{n}x'x = \begin{pmatrix} 4 & 2 \\ 2 & 6 \end{pmatrix}, \quad \frac{1}{n}x'y = \begin{pmatrix} 2 \\ 5 \end{pmatrix}, \quad \frac{1}{n}y'y = 10.$$

(a) (**8 points**) Find the OLS estimate of β_2 and β_3 (call them b_2 and b_3) for the model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u. \tag{1}$$

(b) (**8 points**) Calculate the standard errors of b_2 and b_3.

(c) (**6 points**) Test the hypothesis $3\beta_2 + \beta_3 = 1$ at the 5% level of significance.

(d) (**6 points**) Form a 99% confidence interval for σ^2.

• Now consider the model in (1) with the restriction $\beta_1 = 0$, i.e.,

$$Y = \beta_2 X_2 + \beta_3 X_3 + u. \tag{2}$$

(e) (**10 points**) Find the OLS estimate of β_2 and β_3 (call them $\tilde{b}_2$ and $\tilde{b}_3$). (**Hint:** You need to use the following information in addition to those listed in the top to solve this problem. Assume that $\bar{X}_2 = 2$, $\bar{X}_3 = 3$, and $\bar{Y} = 4$).

(f) (**8 points**) Calculate the standard errors of $\tilde{b}_2$ and $\tilde{b}_3$.

(g) (**12 points**) Find R^2 from model (1) and the following **two** R^2 for the model in (2):

$$(i) R^2 = 1 - RSS/TSS = 1 - e'e/y'y$$

and

$$(ii) R^2 = ESS/TSS = [\tilde{b}_2^2 \sum x_2^2 + \tilde{b}_3^2 \sum x_3^2]/y'y.$$

Comment on these **three** measures of R^2.

(h) (**6 points**) Use the model in (2) to forecast Y given $X_2 = 2.5$ and $X_3 = 3$. Find a 99% confidence interval for σ^2.

2. You may use the following partition inverse formula to solve some parts of the problems listed below.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B_{22} \\ -B_{22}A_{21}A_{11}^{-1} & B_{22} \end{pmatrix}$$

where $B_{22} = \{A_{22} - A_{21}A_{11}^{-1}A_{12}\}^{-1}$.

- Write the K variable regression model

$$Y = X\beta + u \text{ as } Y = X_1\beta_1 + X_2\beta_2 + u \tag{3}$$

where $u \sim (0, \sigma^2 I)$.

(a.1) (**6 points**) Show that the OLS estimator of β_2 is $\hat{\beta}_2 = (X_2' M_1 X_2)^{-1} X_2' M_1 Y$ where $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

(a.2) (**4 points**) Assume *(for this part only)* that $X_1'Y = 0$ i.e., X_1 variables are orthogonal to Y. Can you drop these variables (X_1) from the regression without affecting the estimate of β_2? Explain.

- Now consider the regression

$$Y = \delta_0 + \delta_1 \hat{Y} + v \qquad (4),$$

where $\hat{Y}$ is the predicted value of Y from (3), and v is an error term with zero mean.

(b) (**5 points**) Prove that the OLS estimator of $\delta_0 = 0$ and $\delta_1 = 1$.

- For parts c.1-c.3 assume that X_2 contains a single variable (i.e., X_2 is a $n \times 1$ vector). Estimate γ_1 using OLS from

$$X_2 = X_1\gamma_1 + \eta, \tag{4}$$

where $\eta \sim (0, \sigma_\eta^2 I)$. Define $\hat{\eta} = X_2 - X_1\hat{\gamma}_1$.

(c.1) (**4points**) Show that $\hat{\eta} = M_1 X_2$.

(c.2) (**3 points**) Show that $X_1'\hat{\eta} = 0$, i.e., X_1 and $\hat{\eta}$ are orthogonal.

- Now regress Y on X_1 and $\hat{\eta}$, i.e., use OLS to the following regression

$$Y = X_1\mu_1 + \hat{\eta}\mu_2 + \psi, \tag{5}$$

assuming $E(\psi) = 0$.

(c.3) (**8 points**) Show that $\hat{\mu}_2 = \hat{\beta}_2$.

(c.4) (**6 points**) How would you generalize the result when X_2 contains $K_2 > 1$ regressors? Explain.

Econometrics I
Midterm Exam 1
Spring 1999
Total Points: 75
Time: 1 hr. and 20 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Suppose that application of OLS to 63 observations on a three-variable regression

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \tag{1}$$

yields the following results:

$$(X'X)^{-1} = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{pmatrix}, \quad X'Y = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad s^2 = 4.$$

(a) (**5 points**) Find b.

(b) (**8 points**) Find 95% confidence interval estimates for b_2 and b_3.

(c) (**8 points**) Write the hypotheses $\beta_2 = \beta_3 = 0$ in the form $R\beta = r$ and use the formula

$$\frac{\{R(b-\beta)\}'\{R(X'X)^{-1}R'\}^{-1}\{R(b-\beta)\}/q}{e'e/(n-K)} \sim F_{q,(n-K)}$$

to test the above hypotheses.

(d) (**10 points**) Assume that $\beta_1 = 0$ in (1). Use the above information to calculate the restricted OLS estimators of β using the formula

$$b^* = b + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - Rb)$$

where the restriction $\beta_1 = 0$ is written as $R\beta = r$.

(e) (**10 points**) Suppose that the X observations for a future year are $X_0' = (1\ 2\ 3)$. Construct a 95% confidence interval for the future value of Y denoted by Y_0.

2. (**7 points**) You estimated the K variable regression model

$$Y = X\beta + u \tag{2}$$

and your friend estimated it by adding one extra observation so that he has $n+1$ observations in his model. Assuming that the variance of the error term in both models is the same, show that $V(b) - V(b^*)$ is a positive semi-definite matrix when $V(b)$ and $V(b^*)$ are the variance-covariance matrices corresponding to the OLS estimators of the coefficient of X variables in your model (estimated with n data points) and your friend's model (estimated with $n+1$ data points).

3. Write the K variable regression model

$$Y = X\beta + u \text{ as } Y = X_1\beta_1 + X_2\beta_2 + u \tag{3}$$

where $u \sim (0, \sigma^2 I)$.

Use the partition inverse rule

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B_{22} \\ -B_{22}A_{21}A_{11}^{-1} & B_{22} \end{pmatrix}$$

to show that the OLS estimator of β_2 is

(a) (**10 points**) $b_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$ where $B_{22} = \{A_{22} - A_{21}A_{11}^{-1}A_{12}\}^{-1}$ and $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$. Note that M_1 is an idempotent matrix.

(b) (**5 points**) Find $V(b_2)$.

(c) Assume that $X_1 = \iota$ which is a column vector of ones. Show that

(i)(**5 points**) $e'e = Y'AY - b_2'X_2'AX_2b_2 \equiv y'y - b_2'x_2'x_2b_2 = TSS - ESS$

(iii) (**6 points**) $\frac{b_2'\hat{V}(b_2)^{-1}b_2}{K-1} = \frac{R^2}{(1-R^2)}\frac{(n-K)}{(K-1)}$ where $\hat{V}(b_2)$ is the *estimated* variance-covariance matrix of b_2.

Econometrics I
Midterm Exam
Fall 2000
Total Points: 100
Time: 1 hr. and 15 min.

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

1.(50 points) A sample of 50 quarterly observations on household expenditures for natural gas and electricity is to be used to estimate a linear demand function for gas. The dependent variable (Y) is quantity of gas consumed, while the independent variables are average electricity cost (X_2) and average gas cost (X_3) per quarter. The relevant information is as follows:

$$x'x = \begin{pmatrix} x_2'x_2 & x_2'x_3 \\ x_3'x_2 & x_3'x_3 \end{pmatrix} = \begin{pmatrix} 35 & 17 \\ 17 & 32 \end{pmatrix}, \quad x'y = \begin{pmatrix} x_2'y \\ x_3'y \end{pmatrix} = \begin{pmatrix} 0 \\ -10 \end{pmatrix}, \quad y'y = 6.$$

(a) Since electricity cost is uncorrelated with gas consumption ($x_2'y = 0$), can it be ignored from the multiple regression? Demonstrate your answer by comparing b_{y2}, the slope coefficient of X_2 in the simple regression Y_2 on X_2, to b_2, the slope coefficient of X_2 in the multiple regression Y_2 on X_2 and X_3. Assume that there are intercepts in both models.

(b) Find the standard errors b_2 and b_3. Calculate the standard errors of b_2 and b_3 in the multiple regression Y_2 on X_2 and X_3.

(c) Find R^2 of the regression in part(b). Test the hypothesis of "no regression" at the 5% level of significance.

(d) Test the hypothesis that $3\beta_2 - \beta_3 = 1$.

(e) Suppose that you know a priori that $3\beta_2 - \beta_3 = 1$. Obtain the OLS estimates of β_2 and β_3 using this restriction. Comment on R^2 of this regression.

2. (?? Points) **Notations used in this problem are different from the ones used in class and lecture notes.** Let the regression model be $Y = X\beta + u$, where X is a $n \times k$ matrix whose rank is k. Let

$$Q = X'X, A = Q^{-1}X', N = XA, M = I - N.$$

Recall that $b = AY, \hat{Y} = NY, e = MY$ are the vector of OLS estimators, predicted values of Y, and the OLS residual vector. Using the above definitions and results, show that

$$(a) AN = A, AM = 0, MN = 0, NM = 0.$$

$$(b) NX = X, MX = 0, MM' = M, rank(M) = n - k.$$

$$(c) N\hat{Y} = \hat{Y}, Ne = 0, M\hat{Y} = 0, Me = e, X'\hat{Y} = X'Y.$$

$$(d) Y'\hat{Y} = Y'Xb = b'X'Y = b'Qb = \hat{Y}'\hat{Y}.$$

$$(e) e'e = Y'MY = Y'Y - \hat{Y}'\hat{Y}.$$

Econometrics I
Midterm Exam 1
Spring 2000
Total Points: 100
Time: 1 hr. and 20 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Data from a sample of 25 observations yield the following:

$$x'x = \begin{pmatrix} 15 & 10 \\ 10 & 12 \end{pmatrix}, \quad x'y = \begin{pmatrix} 16 \\ 12 \end{pmatrix}, \quad y'y = 20.$$

(a) (**8 points**) Find the OLS estimate of β_2 and β_3 (call them b_2 and b_3) for the model

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u. \tag{1}$$

(b) (**8 points**) Calculate the standard errors of b_2 and b_3.

(c) (**6 points**) Test the hypothesis $\beta_2 = \beta_3 = 0$ at the 5% level of significance.

(d) (**8 points**) Construct a 99% confidence interval for β_2.

(e) (**14 points**) In addition to the above data, now assume that $\bar{Y} = 4$, $\bar{X}_2 = 2$ and $\bar{X}_3 = 3$. Furthermore, let $X_0 = \{1, 3, 5\}'$. Construct a 95% confidence interval for $Y_0 = \beta' X_0 + u_0$, where $\beta = \{\beta_1, \beta_2, \beta_3\}'$. Would the 95% confidence interval for $E(Y_0)$ be smaller? Explain.

2. The parameters of the regression below

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \tag{2}$$

satisfies the following restrictions in terms of another parameter α:

$$\begin{aligned} \beta_1 + \beta_2 &= \alpha \\ \beta_1 + \beta_3 &= -\alpha \end{aligned} \tag{3}$$

(a) (**4 points**) Use the above restrictions to write equation (2) in terms of parameters β_1 and α.

(b) (**15 points**) Derive an explicit formula for the OLS estimator of α and its variance (assuming $u \sim i.i.d.(0, \sigma^2)$.

(c) (**4 points**) Can you use the data in problem # 1 to get an estimate of α? Explain. (You are NOT asked to use the data to perform the computation).

3. In the regression $Y = \beta_1 \iota + X_2 \beta_2 + u$, where ι is a column vector of ones, the regressors $(X_2, \ldots, X_K)$ are independent (i.e., the sample covariances of each pair of X's are zero).

(a) (**8 points**) Show that $b_2 = (x_2' x_2)^{-1} x_2' y = (\tilde{b}_2, \ldots, \tilde{b}_K)'$, where $\tilde{b}_k = \sum x_k y / \sum x_k^2$, $k = 2, \ldots, K$. Note that the lowercase Y and X variables are all in deviation forms.

(b) (**8 points**) Show that $R^2 = \sum_k r_{ky}^2$, $k = 2, \ldots, K$, where r_{ky} is the simple correlation coefficient between X_k and Y.
What can you say about this result when X's are correlated?

(c) (**4 points**) If another regressor X_{K+1} which is uncorrelated with $(X_2, \ldots, X_K)$ is added to the regression, show that its marginal contribution to R^2 equals r^2 between X_{K+1} and Y.

(d) ((**8+5) points**) Let $y^* = (Y - \bar{Y})/s_Y$ and $x_k^* = (X_k - \bar{X}_k)/s_k$, $k = 2, \ldots, K$, where s_Y^2 and s_k^2 are sample variances of Y and $X_k, k = 2, \ldots, K$. Assume that $X_2, \ldots, X_k$ are uncorrelated. Show that the regression coefficients $(b_k^*, k = 2, \ldots, K)$ from the regression y^* on $x_2^*, \ldots, x_K^*$ are nothing but the simple correlation coefficients between Y and $X_k, k = 2, \ldots, K)$, i.e., $b_k^* = r_{ky}, k = 2, \ldots, K$.
How do you interpret these β_k^* coefficients? Is it better to run a regression y on $x_k, k = 2, \ldots, K$, or y^* on $x_k^*, k = 2, \ldots, K$? Explain.

Econ 616
Midterm Exam 1
Spring 2001
Total Points: 75
Time: 1 hr. and 20 min.

Answer all parts. Note that each part has different weight. Good Luck!!

1. Suppose that application of OLS to 63 observations on a three-variable regression

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u \tag{1}$$

yields the following results:

$$(X'X)^{-1} = \begin{pmatrix} 3 & -2 & 1 \\ -2 & 4 & 0 \\ 1 & 0 & 3 \end{pmatrix}, \quad X'Y = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}, \quad s^2 = 4.$$

(a) (**5 points**) Find b.

(b) (**8 points**) Find 95% confidence interval estimates for b_2 and b_3.

(c) (**8 points**) Write the hypotheses $\beta_2 = \beta_3 = 0$ in the form $R\beta = r$ and use the formula

$$\frac{\{R(b-\beta)\}'\{R(X'X)^{-1}R'\}^{-1}\{R(b-\beta)\}/q}{e'e/(n-K)} \sim F_{q,(n-K)}$$

to test the above hypotheses.

(d) (**10 points**) Assume that $\beta_1 = 0$ in (1). Use the above information to calculate the restricted OLS estimators of β using the formula

$$b^* = b + (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r - Rb)$$

where the restriction $\beta_1 = 0$ is written as $R\beta = r$.

(e) (**10 points**) Suppose that the X observations for a future year are $X_0' = (1\ 2\ 3)$. Construct a 95% confidence interval for the future value of Y denoted by Y_0.

2. You may use the following partition inverse formula to solve some parts of the problems listed below.

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}B_{22}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}B_{22} \\ -B_{22}A_{21}A_{11}^{-1} & B_{22} \end{pmatrix}$$

where $B_{22} = \{A_{22} - A_{21}A_{11}^{-1}A_{12}\}^{-1}$.

- Write the K variable regression model

$$Y = X\beta + u \text{ as } Y = X_1\beta_1 + X_2\beta_2 + u \tag{3}$$

where $u \sim (0, \sigma^2 I)$.

(a.1) (**6 points**) Show that the OLS estimator of β_2 is $\hat{\beta}_2 = (X_2'M_1X_2)^{-1}X_2'M_1Y$ where $M_1 = I - X_1(X_1'X_1)^{-1}X_1'$.

(a.2) (**4 points**) Assume *(for this part only)* that $X_1'Y = 0$ i.e., X_1 variables are orthogonal to Y. Can you drop these variables (X_1) from the regression without affecting the estimate of β_2? Explain.

- Now consider the regression

$$Y = \delta_0 + \delta_1\hat{Y} + v \tag{4},$$

where $\hat{Y}$ is the predicted value of Y from (3), and v is an error term with zero mean.

(b) (**5 points**) Prove that the OLS estimator of $\delta_0 = 0$ and $\delta_1 = 1$.

- For parts c.1-c.3 assume that X_2 contains a single variable (i.e., X_2 is a $n \times 1$ vector).

Estimate γ_1 using OLS from

$$X_2 = X_1\gamma_1 + \eta, \tag{4}$$

where $\eta \sim (0, \sigma_\eta^2 I)$. Define $\hat{\eta} = X_2 - X_1\hat{\gamma}_1$.

(c.1) (**4points**) Show that $\hat{\eta} = M_1X_2$.

(c.2) (**3 points**) Show that $X_1'\hat{\eta} = 0$, i.e., X_1 and $\hat{\eta}$ are orthogonal.

- Now regress Y on X_1 and $\hat{\eta}$, i.e., use OLS to the following regression

$$Y = X_1\mu_1 + \hat{\eta}\mu_2 + \psi, \tag{5}$$

assuming $E(\psi) = 0$.

(c.3) (**8 points**) Show that $\hat{\mu}_2 = \hat{\beta}_2$.

(c.4) (**6 points**) How would you generalize the result when X_2 contains $K_2 > 1$ regressors? Explain.