Econometrics I
Exam 3
Fall 1998
Total Points: 100
Time: 1 hr. 15 min.

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

1. Consider the standard linear multiple regression model given by

$$Y_t = \beta_1 + \beta_2 X_{2t} + \beta_3 X_{3t} + u_t, \quad t = 1, \ldots, 100. \tag{1}$$

Suppose that in deviation form:

$$x'x = \begin{pmatrix} 20 & 15 \\ 15 & 25 \end{pmatrix}, \quad x'y = \begin{pmatrix} 35 \\ 40 \end{pmatrix}, \quad y'y = 100.$$

(a) (**8 points**) Find the OLS estimator of β_2 and β_3.

(b) (**12 points**) It is often stated that multicollinearity is a serious problem if $R^2 < R_k^2$ where R^2 is the goodness-of-fit measure of the regression equation in (1) and R_k^2 is goodness-of-fit for the auxiliary regression X_k on an intercept and $X_2, \ldots, X_{k-1}$. Do you find serious multicollinearity problem, using this rule?

(c) (**10 points**) Assume that $X_2 = mX_3$ where m is a known constant. Show that R^2s from the regression Y on X_2 and Y on X_3 are the same. What is the lesson here? (**Hint**: Note that you don't have to use the above data to answer this question).

2. Consider the following regression model:

$$Y_t^* = \alpha + \beta X_t^* \tag{1}$$

where Y^* and X^* are not observed. Instead we observe Y and X which are related to Y^* and X^* in the following manner: $Y_t = Y_t^* + u_t$, and $X_t = X_t^* + v_t$ where u and v are random variables independent of each other as well as of X^* and Y^*. Assume that for all t: $E(u) = 0, E(v) = 0$, and $E(u^2) = \sigma_u^2$, $E(v^2) = \sigma_v^2$.

(a) (**10 points**) Rewrite (1) in the form $Y_t = \alpha + \beta X_t +$ an error term. Show that the OLS regression of this equation produces inconsistent estimators of α and β. (**Hint**: Note that X is a random variable because of v.)

(b) (**10 points**) Find plim $(\hat{\beta} - \beta)$, where $\hat{\beta}$ is the OLS estimator of β.

(c) (**16 points**) If the regression equation is of the form

$$Y_t = \beta_1 X_{1t} + \beta_2 X_{2t}^* + u_t, \tag{2}$$

where X_{1t} is a non-stochastic variable and $X_{2t} = X_{2t}^* + v_t$. Furthermore, $E(u_t) = E(v_t) = 0$. $E(u_t^2) = \sigma_u^2$, $E(v_t^2) = \sigma_v^2$. v_t is independent of X_{2t}^*, X_{1t} and u_t. Similarly u_t is independent of X_{2t}^* and X_{1t}.

Suggest an estimation procedure that will give you consistent estimators of β_1, and β_2. Briefly explain.

3. (**10 + 8 + 16 points**) Consider the two-variable regression model $Y = \alpha + \beta X + u$. The regressor X is non-stochastic, and $u \sim i.i.dN(0, \sigma^2)$ where σ^2 is *known*. The hypothesis you want to test is $H_0 : \beta = 1$. Show that the LR, W, and LM tests are identical, i.e., $LR = LM = W = \{\sum xy - \sum x^2\}^2/(\sigma^2 \sum x^2)$.

Econometrics I
Final Exam
Spring 1998
Total Points: 115
Time: 2 hrs.

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

1. Consider the linear regression model $Y = \beta X + u$ which satisfies all the standard assumptions. Furthermore, assume that σ^2 is known. Also note that X is a scalar. You are given the following information about the data:

$$\sum X = 11, \sum Y = 20, \sum XY = 24, \sum X^2 = 12, \sum Y^2 = 84, n = 20.$$

(a) Based on the above data calculate the following three estimators of β, viz.,

$$(i) \quad b = \frac{\sum XY}{\sum X^2}$$

$$(ii) \quad b^* = \frac{\sum xy}{\sum x^2}$$

$$(iii) \quad \tilde{b} = \bar{Y}/\bar{X}$$

where x and y are in deviation forms.

(b) Calculate variances of all three estimators and show that b the most efficient estimator of β.

(c) Show that $\tilde{b}$ is the most efficient estimator of β, when $Var(u_i) = \sigma^2 X_i$.

(d) Which one is the most efficient estimator of β if the model is $Y = \alpha + \beta X + u$ in which all the classical assumptions are satisfied? Explain.

2. Consider the linear regression model

$$Y_i = \beta X_i + \gamma D_i + u_i, i = 1, \ldots, n \tag{1}$$

which satisfies all the standard assumptions. X is a scalar and D is a one-time dummy variable ($D_i = 1$ if $i = 1$ and 0 otherwise) That is, $D = 1$ for the first observation and 0 for all other observations.

(a) Find the OLS estimators of β (say $\hat{\beta}_{OLS}$) and γ (say $\hat{\gamma}_{OLS}$) in (1), and show that the OLS residual corresponding to the *first observation* is zero.

(b) Now consider running the regression

$$Y_i = \delta X_i + u_i, i = 2, \ldots, n \tag{2}$$

(note that the first observation and the one-time dummy variable are deleted in this regression). Let $\hat{\delta}_{OLS}$ be the OLS estimator of δ in (2). Show that $\hat{\delta}_{OLS} = \hat{\beta}_{OLS}$. That is, use of a one-time dummy variable is equivalent to throwing that observation.

3. Consider the regression $Y = \alpha + \beta X + u$ in which all the classical assumptions are satisfied. Let $\hat{\beta}_{OLS}$ be the OLS estimator of β.

Now consider running the regression $Y = \delta\hat{v} + \epsilon$ where $\hat{v}$ is the OLS residual of the auxiliary regression $X = \gamma + v$.

(a) Show that the OLS estimator of δ in the regression $Y = \delta\hat{v} + \epsilon$ is the same as $\hat{\beta}_{OLS}$.

(b) How do you generalize the above result to the model $Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + u$? Explain.

4. Consider the simple model $Y_i = \alpha + u_i, i = 1, 2$, where $E(u_i) = 0, V(u_i) = \sigma^2$ and $Cov(u_1, u_2) = \rho\sigma^2$. Assume that ρ and σ^2 are known. Show that $\hat{\alpha}_{OLS} = \hat{\alpha}_{GLS}$.

5. Consider the linear regression model $Y = \beta X + u$ which satisfies all the standard assumptions. Construct the LR and W statistics to test the hypothesis $\beta = 1$ under the assumption that $u \sim i.i.d.N(0, \sigma^2)$.

6. Suppose you want to estimate the model

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t}^* + u_t, \tag{1}$$

where X_{2t}^* is not observed. (Both X_1 and X_2 are scalars). The measurement equation is

$$X_{2t} = X_{2t}^* + v_t, \tag{2}$$

where $E(u_t) = E(v_t) = 0$. $E(u_t^2) = \sigma_u^2$, $E(v_t^2) = \sigma_v^2$. v_t is independent of X_{2t}^*, X_{1t} and u_t. Similarly u_t is independent of X_{2t}^* and X_{1t}.

(a) Show that the OLS estimator of β_1 is inconsistent when X_2^* is dropped from the model in (1).

(b) Show that the OLS estimator of β_1 is also inconsistent when X_2^* is replaced by X_2 in (1).

(c) Which one of the above two is worse? Why?

(d) Suggest an estimation procedure to obtain consistent estimators of both β_1 and β_2. Briefly explain the procedure in the context of the model in (1).

Econometrics I
Final Exam
Fall 1999
Total Points: 100
Time: 3 hrs.

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

1. Consider the following regression model

$$Y_t = \alpha + u_t, \tag{1}$$

where $u_t \sim (0, \sigma_u^2 X_t^2)$. Assume σ_u^2 is known.

(a) (**2 points**) Derive the OLS estimator of α.

(b) (**5 points**) Derive the GLS estimator of α.

(c) (**5 points**) Prove that the GLS estimator is more efficient than the OLS estimator. [Hint: Use the Cauchy-Schwartz inequality $\sum z_1^2 \sum z_2^2 \geq (\sum z_1 z_2)^2$].

Now consider the model

$$Y_t = \alpha + u_t, \tag{2}$$

and

$$u_t = \rho u_{t-1} + \epsilon_t. \tag{3}$$

Furthermore, $\epsilon_t \sim (0, \sigma_\epsilon^2 X_t^2)$. Assume σ_ϵ^2 is known.

(d) (**15 points**) You are given the following 5 observations on $Y = (6, 3, 12, 15, 4)$ and $X = (4, 1, 5, 8, 2)$. Assume $\rho = 0.5$. Compute the estimator of α correcting for both heteroscedasticity and autocorrelation.

2. Data on a three variable regression

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u, \tag{4}$$

are given in the following form

$$x'x = \begin{pmatrix} 20 & 10 \\ 10 & 30 \end{pmatrix}, \quad x'y = \begin{pmatrix} 30 \\ 40 \end{pmatrix}, \quad y'y = 75.$$

Note that y and x are in deviation form.

(a) (**12 points**) Compute the OLS estimator of β_2 and β_3, and R^2.

(b) (**4 points**) It is often stated that there is serious multicollinearity problem if R_k^2 (R^2 from the regression X_k on all other X variables including a constant, for $k = 2, \ldots, K$) exceeds R^2 of the regression equation in (4). Using this rule, do you find muticollinearity a problem in the above data?

(c) (**6 points**) If X_3 is omitted from (4) show that $E(\tilde{b}_2) \neq \beta_2$, where $\tilde{b}_2$ is the OLS estimator of β_2 from $Y = \beta_1 + \beta_2 X_2 + v$.

(d) (**18 points**) Compute the LR, LM, and W statistics to test the hypothesis $H_0 : \beta_3 = 0$ in (4).

3. Consider the regression equation

$$Y = \beta_1 + \beta_2 D + u, \tag{5}$$

where D is a dummy variable for sex (male = 1) and $u \sim N(0, 10)$. There are 20 males and 30 females in the sample. The mean values of Y for the males and females are 2 and 3, respectively.

(a) (**6 points**) Write down the $X'X$ and $X'Y$ using the above information. matrics.

(b) (**8 points**) Compute the OLS estimates of β_1 and β_2.

(c) (**5 points**) Compute the value of the test statistic for testing the hypothesis H_0 : $3\beta_1 + \beta_2 = 37$.

(d) (**4 points**) Instead of using (5) if you run two separate regressions (one for males and another for females) on a constant, will these regressions be more general than the regression in (5)? Explain.

(e) (**10 points**) If the relationship in (4) is changed to

$$Y = \beta_1 + \beta_2 D + \beta_3 X + u, \tag{6}$$

and you regress Y on a constant and X separately for males and females, will these regressions be more general than the regression in (6)? Explain. What sort of hypothesis can you test in these regressions compared to the one in (6)? Explain.

Econometrics I
Exam III
Spring 1999
Total Points: 140
Time: 150 min.

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

1. Suppose that you are given data on three X variables (X_2, X_3, X_4), and the Y variable. The sample moment matrices for a sample of 24 observations in deviation form are:

$$x'x = \begin{pmatrix} 10 & 10 & 5 \\ 10 & 30 & 15 \\ 5 & 15 & 20 \end{pmatrix}, \quad x'y = \begin{pmatrix} 7 \\ -7 \\ -26 \end{pmatrix}, \quad y'y = 60.$$

(a) Let the true relationship be

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

in which X_3^* is unobserved. Use X_3 as a proxy for X_3^* (i.e., $X_3 = X_3^* + v$) and estimate β_2 and β_3. State the assumptions you need to make on the error terms u and v. **8 pts.**

(b) Find the OLS estimates of β_2 if X_3^* is dropped from the regression in (1). **6 pts.**

(c) Now consider the regression

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

and assume that X_3 is random and correlated with u. Find the IV estimates of β_2 and β_3 when X_4 is used as an instrument for X_3. **8 pts.**

(d) What are the properties of the estimators in (a), (b), and (c)? **6 pts.**

(e) Assume that the model is

$$Y^* = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + u, \tag{3}$$

where Y^* is measured with errors but the X variables are measured without errors. Furthermore, assume that $Y = Y^* + v$ where Y is the observed value, and v is a random variable with zero mean and constant variance, independent of Y^*, X_2 and X_3. If you use OLS to estimate the parameters of the following regression

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \eta, \tag{4}$$

where $\eta = u + v$, will they be consistent? Show your results explicitly. **7 pts.**

(f) Now assume that $v = 0$ in (4). It is often stated that multicollinearity is a serious problem if $R^2 < R_k^2$ where R^2 is the goodness-of-fit measure of the regression equation in (4) and R_k^2 is goodness-of-fit for the auxiliary regression X_k on an intercept and $X_2, \ldots, X_{k-1}$. Do you find serious multicollinearity problem, using this rule? **7 pts.**

2. Consider the model

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

where $u \sim i.i.d.N(0, \sigma^2)$. Test $H_0 : \beta_2 = \beta_3 = 0$ using the LR, W, and LM tests. **6+6+6 pts.**

Note: Use the relevant data from Problem #1 to perform the above tests. Assume that all the variables are observed and measured without errors.

Econometrics I
Final Exam
Spring 2000
Total Points: 150
Time: 3 hrs.

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

1. The usual two variable regression model

$$Y = \beta_1 + \beta_2 X + u, \tag{1}$$

which satisfied all classical assumptions, is postulated. A sample of 20 observations is drawn from an urban area, a sample of 10 observations is drawn from a rural area, and another sample of 15 observations is drawn from a metropolitan area. The sample information in raw form is summarized below:

Urban

$$X'X = \begin{pmatrix} 20 & 20 \\ 20 & 25 \end{pmatrix}, X'Y = \begin{pmatrix} 10 \\ 20 \end{pmatrix}, Y'Y = 30$$

Rural

$$X'X = \begin{pmatrix} 10 & 10 \\ 10 & 20 \end{pmatrix}, X'Y = \begin{pmatrix} 8 \\ 20 \end{pmatrix}, Y'Y = 24$$

Metropolitan

$$X'X = \begin{pmatrix} 15 & 15 \\ 15 & 24 \end{pmatrix}, X'Y = \begin{pmatrix} 10 \\ 16 \end{pmatrix}, Y'Y = 32$$

(a) (**15 points**) Test the hypothesis that the same relationship holds in all three areas, assuming that $V(u)$ is the same for all three groups.

(b) (**10 points**) Assume that the relationship in rural and metropolitan areas are the same (irrespective of your answer to part (a)) and you combined them and labeled it semi-urban. Test the hypothesis that the same relationship holds in the urban and semi-urban areas.

(c) (**8 points**) Assume that the three areas have the same relationship (irrespective of your answer to part (a)) but the error variances differ. Obtain GLS estimates of β_1 and β_2.

2(a). (**4+4+4 points**) Use the following data to estimate β_2 and β_3, their standard errors, and R^2 for the model

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

which satisfies all the assumptions of the classical linear regression model. The raw data is:

$$Y = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{pmatrix}, X = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ 1 & 4 & 12 \\ 1 & 6 & 0 \\ 1 & 8 & 16 \end{pmatrix}$$

(**4 points**) Do you suspect multicollinearity? Why?

Now consider a small change in the above data. Interchange the 3rd and 4th values of X_3. Now the X matrix becomes

$$X = \begin{pmatrix} 1 & 2 & 4 \\ 1 & 0 & 2 \\ 1 & 4 & 0 \\ 1 & 6 & 12 \\ 1 & 8 & 16 \end{pmatrix}$$

(b) (**4+4+4 points**) Recalculate estimators of β_2 and β_3, their standard errors, and R^2. Do you suspect multicollinearity? Why?

(c) (**4 points**) One effect of multicollinearity is that the estimates of the parameters and their standard errors are very sensitive to even a small change in the data. Do you find any evidence of this from your answers to parts (a) and (b). Explain.

3. Consider the usual two variable regression model

$$Y_t = \beta_1 + \beta_2 X_t + u_t, t = 1, \ldots, n, \tag{2}$$

which satisfied all classical assumptions, *except* that $E(u_t^2) = \sigma_t^2$. Assume further that σ_t^2 is known.

(a) (**8 points**) Find the GLS estimators of β_2 and its variance, denoted by b_2^G and $V(b_2^G)$.

(b) (**6 points**) Show that b_2^G and $V(b_2^G)$ can be expressed as

$$b_2^G = \frac{\sum_t w_t y_t^* x_t^*}{\sum_t w_t (x_t^*)^2},$$
$$V(b_2^G) = \frac{1}{\sum_t w_t (x_t^*)^2}$$

where $w_t = 1/\sigma_t^2$, $y_t^* = Y_t - \overline{Y^*}, x_t^* = X_t - \bar{X}^*, \bar{Y}^* = \sum_t w_t Y_t / \sum_t w_t, \bar{X}^* = \sum_t w_t X_t / \sum_t w_t$.

(c) (**6 points**) If $\sigma_t^2 = \sigma^2 \kappa_t$ (κ_t known), show that $V(b_2) = \frac{\sigma^2}{\sum_t x_t^2} \frac{\sum_t x_t^2 \kappa_t}{\sum_t x_t^2}$, where b_2 is the OLS estimator of β_2 from (2), and $x_t = X_t - \bar{X}$. What can you say about this variance compared to the one obtained when heteroskedasticity is ignored.

4. In the regression $Y_t = \alpha + \beta X_t + u_t, t = 1, \ldots, n$ (X a scalar), and $u_t = \rho u_{t-1} + \epsilon_t, |\rho| < 1$, $\epsilon \sim i.i.d\,(0, \sigma_\epsilon^2)$.

(a) (**4 points**) Show that $b - -$ the OLS estimator of β is consistent.

(b) (**6 points**) Denote the OLS residuals by e_t. Now consider the OLS estimator of ρ ($\hat{\rho}$) from the regression

$$e_t = \rho e_{t-1} + \epsilon_t. \tag{3}$$

Show that $plim \hat{\rho} = \rho$. (**Hint: Write** $e_t = u_t + (\beta - b) X_t$ **and use the result from part (a)**).

(c) (**6 points**) Write down the expression for the estimated variance of $\hat{\rho}$ from (3), and show that $var(\hat{\rho}) \approx (1 - \hat{\rho}^2)/n$.

(d) (**6 points**) Assume that the estimated model is

$$Y_t = -2.3 + 1.8 X_t + e_t, R^2 = .82, n = 25, DW - (d) = .79.$$

Use the result that $\hat{\rho}$ is approximately normally distributed, to test the hypothesis that $\rho = 0$. (**Hint: use the result that $\hat{\rho}$ is normal in conjunction with the results in parts (b) and (c) to derive the value of the t statistic.**)

5. In the regression $Y_t = \alpha + \beta X_t + u_t, t = 1, \ldots, n$ (X a scalr), and $u_t = \rho u_{t-1} + \epsilon_t, |\rho| < 1$, $\epsilon \sim (0, \sigma_\epsilon^2 X_t^2)$. Furthermore, $E(\epsilon_t \epsilon_\tau) = 0 \, \forall t \neq \tau$.

(a) (**8 points**) Derive the variance covariance matrix of u.

(a) (**4 points**) Show that the transformation

$$\frac{Y_t - \rho Y_{t-1}}{X_t} = \frac{\alpha(1-\rho)}{X_t} + \beta \frac{X_t - \rho X_{t-1}}{X_t} + \frac{\epsilon_t}{X_t} \tag{4}$$

corrects both heteroskedasticity and autocorrelation problems.

3. Consider the following regression:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + u_t, \quad t = 1, \ldots, n. \tag{1}$$

(a) Suppose $u_t \sim i.i.d.(0, \sigma^2)$. Show that the OLS estimators of β_1 and β_2 are consistent. **8 pts.**

(b) If $u_t = \rho u_{t-1} + \epsilon_t$ where $|\rho| < 1$ and $\epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$, show that the OLS estimators of β_1 and β_2 are inconsistent. Calculate $plim(b_1 - \beta_1)$ and $plim(b_2 - \beta_2)$, when b_1 and b_2 are OLS estimators of β_1 and β_2. **12 pts.**

(c) Suggest an alternative estimator, and show that it give consistent estimators of β_1 and β_2 in (b). **8 pts.**

(d) Use the following three data points on Y, Z_1, and Z_2 to derive IV estimators of β_1 and β_2 in (b). Use Z_1 and Z_2 as instruments for Y_{t-1}, and assume that $Y_0 = 2$.
$Y_t = 1, 4, 3$; $Z_{1t} = 2, 3, 1$; and $Z_{2t} = 1, 3, 5$. **10 pts.**

4. Consider the following four data points on Y and X,
$Y = 4, 7, 3, 9$; $X = 2, 3, 1, 5$; to estimate the model

$$Y_t = \beta_1 + \beta_2 X_{2t} + u_t, \quad t = 1, 2, 3, 4. \tag{1}$$

(a) Estimate β_1 and β_2 using OLS. **7 pts.**

(b) After estimating the model you discovered that u is heteroskedastic, viz., $\sigma_t^2 = \sigma^2 \; diag(.10, .05, .20, .25)$. Derive the heteroskedastic consistent variance covariance matrix of β_1 and β_2. **12 pts.**

(c) Assume that $u_t = \rho u_{t-1} + \epsilon_t$ where $|\rho| < 1$ and $\epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$.
Test $\rho = 0$ using (i) the DW test, and (ii) the LM (Breusch and Godfrey) test. **6+9 pts.**

(e) Estimate β_1 and β_2 when $u_t = \rho u_{t-1} + \epsilon_t$, $|\rho| < 1$ and $\epsilon \sim i.i.d.(0, \sigma_\epsilon^2)$. Use the Prais-Winsten procedure. **8 pts.**

Econ 616
Final Exam
Spring 2001
Total Points: 105
Time: 2 hrs.

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

1(a). Consider the following regression

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

in which X_3^* is unobserved. Consider using X_3 as a proxy for X_3^*.

(i) When do you obtain consistent estimators of β_2 and β_3 by regressing Y on X_2 and X_3? Explain. **6 points**.

(ii) Under what conditions/situations the regression Y on X_2 and X_3 will **not** produce consistent estimators of β_2 and β_3? Explain. **6 points**.

Suppose that you are given data on three X variables (X_2, X_3, X_4), and the Y variable. The sample moment matrices for a sample of 30 observations in deviation form are:

$$x'x = \begin{pmatrix} 10 & 10 & 5 \\ 10 & 15 & 20 \\ 5 & 20 & 30 \end{pmatrix}, \quad x'y = \begin{pmatrix} 5 \\ -5 \\ 25 \end{pmatrix}, \quad y'y = 60.$$

(iii) Use the above data to obtain numerical values of the estimators of β_2 and β_3 proposed in (i) above. **6 points**.

(b) Find the OLS estimates of β_2 if X_3^* is dropped from the regression in (1) and show that it is inconsistent. Show the degree of inconsistency algebraically (no need to compute numerical values). **8 points**.

(c) Now consider the regression

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

and assume that X_3 is random and correlated with u. Do you find this problem different from the one specified in either (i) or (ii) above? Find the IV estimates of β_2 and β_3 when X_4 is used as an instrument for X_3. **10 points**.

2. Let the salary of college teachers be given by the following regression

$$Y = \beta_1 + \beta_2 D_2 + \beta_3 D_3 + \beta_4 (D_2.D_3) + \beta_5 Experience + u, \tag{3}$$

where $D_2 = 1$ if male and 0 otherwise, $D_3 = 1$ if white and 0 is non-white.

(a) Interpret coefficients β_3 and β_4.

Is it possible to further generalize the model in (3)? If so, how? Explain. **6+4 points**.

(b) Assume that the variance of u is different for males and females but the same for whiles and non-whites. That is u is heteroskedastic across genders. How would you estimate the parameters in (3) correcting for heteroskedasticity? Explain. **6 points**.

(c) How can you test whether the variance of u is different for males and females (assuming that the variance is the same for whiles and non-whites)? Explain. **8 points**. (**Hint:** If you do a LR test, show the log-likelihood functions for the restricted and unrestricted models.)

3. It is often argued (Goldberger) that near perfect multicollinearity and micronumerosity (small sample size or small degrees of freedom) have the same effect on standard errors of OLS estimators. Illustrate this point by using a model of your own choice. **10 points**.

4. In the adaptive expectations model of capital stock growth, the level of capital is set at a multiple of anticipated output Y_t^*,

$$K_t = \delta Y_t^*, \tag{4}$$

where

$$Y_t^* - Y_{t-1}^* = \alpha(Y_t - Y_{t-1}^*) + u_t, \quad 0 < \alpha \leq 1 \tag{5}$$

(a) Express K_t as a function of observed variables. (Note that Y_t^* and Y_{t-1}^* are unobserved). **5 points**.

(b) Does OLS to the model in part (a) give consistent estimates of α and δ when u_t follows an AR(1) process? Explain. **8 points**.

(c) Now assume that

$$K_t = \delta Y_t^*$$

and

$$K_t - K_{t-1} = \gamma(K_t^* - K_{t-1}) + v_t, \quad 0 < \delta \leq 1,$$

$$Y_t^* - Y_{t-1}^* = \alpha(Y_t - Y_{t-1}^*) + u_t, \quad 0 < \alpha \leq 1.$$

Express K_t as a function of observed variables. Can the parameters α, γ and δ be estimated? Explain. **8 points**.

5. Consider the following regression:

$$Y_t = \beta_1 + \beta_2 Y_{t-1} + u_t, \quad t = 1, \ldots, n, \tag{4}$$

when $u_t = \epsilon_t - \phi\epsilon_{t-1}$ and $\epsilon_t \sim i.i.d.(0, \sigma_\epsilon^2)$. Assume that $|\phi| < 1$.
Calculate **either** $plim(b_1 - \beta_1)$ **or** $plim(b_2 - \beta_2)$, when b_1 and b_2 are OLS estimators of β_1 and β_2. **14 points**.