Answer three of the following questions. You must answer question 5. The questions are weighted equally. You have 75 minutes. You may use a calculator. Brevity is recommended. The exam is closed-book and closed-notes. You may utilize both sides of an 8.5 by 11 "cheat" sheet as an aid.

1. Answer true or false and state why.
   a. The power of the test increases as the variability of the disturbance term increases.
   b. Suppose the RHS regressors are stochastic, then the usual $t$-ratios will have a $t$-distribution in finite samples.
   c. If the omitted variables are orthogonal to the included variables in a regression, then the OLS estimates on the included variables will be biased and inconsistent.
   d. If multicollinearity is known to be present, then we have no hope of obtaining significant estimates for the coefficients on the collinear terms.
   e. If the DW statistic has a realization between $d_L$ and $d_U$ then we should reject the null hypothesis of no serial correlation.
   f. If omitted variables move smoothly then the error terms on the specified equations will be serially correlated.
   g. When the model has evidence of a nonscalar covariance matrix, then we can correct any difficulties through the use of instrumental variables.

2. Consider the following multivariate regression model (for $t = 1, 2, ..., n$):
   $$ y_t = x_t^* \beta + u_t $$
   where $x_t$ is a $k \times 1$ vector of observations on the $k$ independent variables for period $t$, $y_t$ is the value of the dependent variable and $\beta$ is the $k \times 1$ vector of unknown coefficients. The disturbance term satisfies the classical stochastic assumptions. Suppose we seek to predict the value of the dependent variable given by
   $$ y_* = x_0^* \beta + u_* $$
   for some observation * outside the sample. We treat $x_*$ as given for the purposes of these predictions.
   a. Show that $x_0^* \beta$ is the minimum mean-squared error predictor of $y_*$. 
   b. Determine the mean and variance of the least-squares based predictor $\hat{y}_* = x_0^* \hat{\beta}$ and the prediction error $y_* - \hat{y}_*$. 
   c. Suppose $u_t$ i.i.d. $N(0, \sigma^2)$, then what is the distribution of
      $$ (y_* - \hat{y}_*)/\left\{\sigma^2[1 + x_*^*(X'X)^{-1}x_*]\right\}^{1/2} $$
   d. What is the distribution of this statistic if we replace $\sigma^2$ by $s^2$? How might this statistic be used to obtain a prediction interval for $y_*$?
   a. Discuss the values of $x_*$ for which the predictions will be more accurate.

3. Consider the following model of aggregate consumption
   $$ C_t = \alpha + \beta Y_t + u_t $$
   where $C_t$ is aggregate consumption and $Y_t$ is aggregate income. The unobservable disturbances are assumed to be i.i.d. $N(0, \sigma^2)$. 

---

Econ 409
B. Brown  
Spring 2011

Second Midterm Exam

Answer three of the following questions. You must answer question 5. The questions are weighted equally. You have 75 minutes. You may use a calculator. Brevity is recommended. The exam is closed-book and closed-notes. You may utilize both sides of an 8.5 by 11 “cheat” sheet as an aid.

1. Answer true or false and state why.
   a. The power of the test increases as the variability of the disturbance term increases.
   b. Suppose the RHS regressors are stochastic, then the usual $t$-ratios will have a $t$-distribution in finite samples.
   c. If the omitted variables are orthogonal to the included variables in a regression, then the OLS estimates on the included variables will be biased and inconsistent.
   d. If multicollinearity is known to be present, then we have no hope of obtaining significant estimates for the coefficients on the collinear terms.
   e. If the $DW$ statistic has a realization between $d_L$ and $d_U$ then we should reject the null hypothesis of no serial correlation.
   f. If omitted variables move smoothly then the error terms on the specified equations will be serially correlated.
   g. When the model has evidence of a nonscalar covariance matrix, then we can correct any difficulties through the use of instrumental variables.

2. Consider the following multivariate regression model (for $t = 1, 2, ..., n$):
   $$ y_t = x_t^* \beta + u_t $$
   where $x_t$ is a $k \times 1$ vector of observations on the $k$ independent variables for period $t$, $y_t$ is the value of the dependent variable and $\beta$ is the $k \times 1$ vector of unknown coefficients. The disturbance term satisfies the classical stochastic assumptions. Suppose we seek to predict the value of the dependent variable given by
   $$ y_* = x_0^* \beta + u_* $$
   for some observation * outside the sample. We treat $x_*$ as given for the purposes of these predictions.
   a. Show that $x_0^* \beta$ is the minimum mean-squared error predictor of $y_*$. 
   b. Determine the mean and variance of the least-squares based predictor $\hat{y}_* = x_0^* \hat{\beta}$ and the prediction error $y_* - \hat{y}_*$. 
   c. Suppose $u_t$ i.i.d. $N(0, \sigma^2)$, then what is the distribution of
      $$ (y_* - \hat{y}_*)/\left\{\sigma^2[1 + x_*^*(X'X)^{-1}x_*]\right\}^{1/2} $$
   d. What is the distribution of this statistic if we replace $\sigma^2$ by $s^2$? How might this statistic be used to obtain a prediction interval for $y_*$?
   a. Discuss the values of $x_*$ for which the predictions will be more accurate.

3. Consider the following model of aggregate consumption
   $$ C_t = \alpha + \beta Y_t + u_t $$
   where $C_t$ is aggregate consumption and $Y_t$ is aggregate income. The unobservable disturbances are assumed to be i.i.d. $N(0, \sigma^2)$.
a. Assume the $Y_t$ are nonstochastic and nonconstant, then what will be the distribution of the OLS estimates of $\alpha$ and $\beta$.
b. Alternatively, assume the $Y_t$ are stochastic, what can be said about the distribution of the OLS estimates in small samples? In large samples?
Suppose consumption and income must satisfy the identity
$$Y_t = C_t + G_t$$
where $G_t$ is autonomous expenditures by the government.
c. Show that $Y_t$ will now be correlated with $u_t$. What problems will this cause for the OLS estimates of $\alpha$ $\beta$?
d. Outline a procedure for estimating $\alpha$ and $\beta$ which will avoid the difficulties introduced by (ii).

4. Professor I. M. Economist asked his research assistant to estimate the parameters of the regression model
$$Y_i = \alpha + \beta X_i + \gamma Z_i + \delta W_i + u_i.$$  
He suspected that multicollinearity might be a problem and suggested an analysis of the possibility. Suppose you are the research assistant.

a. Explain exactly what is meant by multicollinearity in this model.
b. What would the consequences of collinearity be for estimation and inference?
c. How might you go about detecting the presence and severity of collinearity?
d. Given that collinearity is present and poses serious complications, what are some possible approaches to solving or circumventing the problem.

5. Prof. I.M. Wright is interested in determining the relationship between housing expenditures and income. Accordingly, he proposed the model
$$H_i = \alpha + \beta Y_i + u_i$$
where $H_i$ is housing expenditures, $Y_i$ is gross annual income, and $u_i$ is a well-behaved disturbance. Using a cross-section of 20 families, his research assistant obtained the following estimates of the model (estimated standard errors in parentheses).
$$H = 0.89 + 0.237Y$$  
(i)
with $R^2 = .934$, $s = 0.373$, and $d.f. = 18$. It was suggested by his know-it-all colleague, Dr. U.R. Wong, that housing expeditur behavior may differ depending upon whether the income was small or large, so the model was reestimated for two subsamples. For the families with incomes below $10,000 he obtained
$$H = 0.60 + 0.276Y$$  
(ii)
with $R^2 = .940$, $s = 0.194$, and $d.f. = 8$. While for families with incomes above $10,000$ the regression was
$$H = 1.54 + 0.20Y$$  
(iii)
with $R^2 = .552$, $s = 0.503$, and $d.f. = 8$.

a. Making explicit your assumptions, briefly describe the properties of coefficient estimates for (i), given that the model does not change between samples.
b. Calculate an approximate prob-value for the income coefficients in (i). What informations does this prob-value provide us with regard to hypothesis testing?
c. Test the hypothesis that the variance of the disturbance term is the same in both samples. What are the implications for the estimate (i) if the hypothesis is rejected?
d. Test the hypothesis that the coefficients are the same in both samples (given that the variance is constant.) What are the consequences for the estimates (i) if the hypothesis is rejected?

e. Suppose the true model is

\[ H_i = a + \beta Y_i + \gamma r_i + u_i \]

where \( r_i \) is the interest rate. What are the consequences for the estimates of all three regressions?