QUEEN'S UNIVERSITY AT KINGSTON
Department of Economics

ECONOMICS 351* - Section A

Introductory Econometrics

Fall Term 2001 **FINAL EXAMINATION** M.G. Abbott

DATE: **Friday December 14, 2001**

TIME: **Three (3) hours (180 minutes); 2:00 p.m. - 5:00 p.m.**

INSTRUCTIONS: The examination is divided into two parts.

PART A contains two questions; students are required to answer **ONE** of the two questions 1 and 2 in Part A.

PART B contains three questions; students are required to answer **ALL THREE** of the questions 3, 4 and 5 in Part B.

- Answer all questions in the exam booklets provided. Be sure your ***name*** and ***student number*** are printed clearly on the front of all exam booklets used.
- **Do not write answers to questions on the front page of the first exam booklet**.
- **Please label clearly** each of your answers in the exam booklets with the appropriate number and letter.
- **Please write legibly.** GOOD LUCK! Happy Holidays!

If doubt exists as to the interpretation of any question, then for any question requiring a written answer the candidate is urged to submit with the answer paper a clear statement of any assumptions made.

MARKING: The marks for each question are indicated in parentheses immediately above each question. **Total marks** for the exam **equal 200**.

PART A: Questions 1 and 2 (30 marks for each question) .... **30 marks**
Answer ***either one of*** **Questions 1 and 2**.

PART B: Questions 3 (70 marks), 4 (50 marks) and 5 (50 marks) .... **170 marks**
Answer ***all parts of*** **Questions 3, 4 *and* 5**.

TOTAL MARKS .... **200 marks**

PART A (30 marks)

Instructions: Answer **EITHER ONE (1) of questions 1 and 2** in this part. Total marks for each question equal 30; marks for each part are given in parentheses.

(30 marks)

1. Consider the simple (two-variable) linear regression model for which the population regression equation can be written in conventional notation as:

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad \textbf{(1)}$$

where Y_i and X_i are observable variables, β_1 and β_2 and unknown (constant) regression coefficients, and u_i is an unobservable random error term. The Ordinary Least Squares (OLS) sample regression equation corresponding to regression equation (1) is

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_i + \hat{u}_i \qquad (i = 1, ..., N) \qquad \textbf{(2)}$$

where $\hat{\beta}_1$ is the OLS estimator of the intercept coefficient β_1, $\hat{\beta}_2$ is the OLS estimator of the slope coefficient β_2, $\hat{u}_i$ is the OLS residual for the i-th sample observation, and N is sample size (the number of observations in the sample).

(20 marks)

(a) Stating explicitly all required assumptions, derive the expression (or formula) for $\text{Var}(\hat{\beta}_2)$, the variance of the OLS slope coefficient estimator $\hat{\beta}_2$. How do you compute an unbiased estimator of $\text{Var}(\hat{\beta}_2)$?

(10 marks)

(b) Define the p-value of the t-statistic for $\hat{\beta}_2$ when the null hypothesis H_0: $\beta_2 = 0$ is tested against each of the following three alternative hypotheses:

(1) H_1: $\beta_2 \neq 0$;

(2) H_1: $\beta_2 > 0$;

(3) H_1: $\beta_2 < 0$.

(30 marks)

2. Consider the multiple linear regression model for which the population regression equation can be written in conventional notation as:

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad \textbf{(1)}$$

where Y_i, X_{2i} and X_{3i} are observable variables; β_1, β_2 and β_3 are unknown (constant) regression coefficients; and u_i is an unobservable random error term. The Ordinary Least Squares (OLS) sample regression equation corresponding to regression equation (1) is

$$Y_i = \hat{\beta}_1 + \hat{\beta}_2 X_{2i} + \hat{\beta}_3 X_{3i} + \hat{u}_i \qquad (i = 1, ..., N) \qquad \textbf{(2)}$$

where $\hat{\beta}_1$ is the OLS estimator of the intercept coefficient β_1, $\hat{\beta}_2$ is the OLS estimator of the slope coefficient β_2, $\hat{\beta}_3$ is the OLS estimator of the slope coefficient β_3, $\hat{u}_i$ is the OLS residual for the i-th sample observation, and N is sample size (the number of observations in the estimation sample).

(20 marks)

(a) State the Ordinary Least Squares (OLS) estimation criterion. Derive the OLS normal equations for regression equation (1) from the OLS estimation criterion.

(10 marks)

(b) Explain what is meant by each of the following statements about the estimator $\hat{\beta}_2$ of the slope coefficient β_2.

(1) $\hat{\beta}_2$ is an unbiased estimator of β_2.

(2) $\hat{\beta}_2$ is an efficient estimator of β_2.

For $\hat{\beta}_2$ to be an efficient estimator of β_2, must it be an unbiased estimator of β_2, yes or no?

PART B (170 marks)

Instructions: Answer ***all*** **parts of questions 3, 4 and 5** in this part. Question 3 is worth a total of 70 marks. Questions 4 and 5 are each worth a total of 50 marks. Marks for each part are given in parentheses. **Show explicitly all formulas and calculations**.

(70 marks)

3. You are conducting an empirical investigation into the median prices of houses in 506 communities of a large metropolitan area. The sample data consist of 506 observations on the following observable variables:

P_i = the median house price in community i, in dollars;

NOX_i = the level of nitrous oxide in the air of community i, in parts per 100 million;

$DIST_i$ = the weighted distance of community i from 5 employment centres, in miles;

$ROOMS_i$ = the average number of rooms per house in community i;

$STRAT_i$ = the average student-teacher ratio of schools in community i.

Your research assistant estimates the following model of median house prices on the sample of N = 506 observations. The OLS estimation results for the model are given below (with standard errors given in parentheses below coefficient estimates).

Regression Model

$$\ln P_i = \beta_1 + \beta_2 \ln NOX_i + \beta_3 \ln DIST_i + \beta_4 ROOMS_i + \beta_5 STRAT_i + u_i \qquad \textbf{(1)}$$

where the β_j (j = 1, 2, 3, 4, 5) are regression coefficients, $\ln X_i$ denotes the natural logarithm of the variable X_i, and u_i is a random error term.

OLS Estimates of Equation (1) , Question 3: (standard errors in parentheses below coefficient estimates)

$\hat\beta_1 = 11.08$	$\hat\beta_2 = -0.9535$	$\hat\beta_3 = -0.1343$	$\hat\beta_4 = 0.2545$	$\hat\beta_5 = -0.05245$
(0.3181)	(0.1167)	(0.04310)	(0.01853)	(0.005897)

$$RSS = \sum_{i=1}^{N} \hat u_i^2 = 35.1835; \qquad TSS = \sum_{i=1}^{N} \left(\ln P_i - \overline{\ln P}\right)^2 = 84.5822; \qquad N = 506$$

where RSS is the Residual Sum-of-Squares and TSS is the Total Sum-of-Squares from OLS estimation of regression equation (1).

3. (continued)

(10 marks)

(a) Interpret each of the slope coefficient estimates $\hat{\beta}_2$ and $\hat{\beta}_4$ in regression equation (1); that is, explain in words what the numerical values of the slope coefficient estimates $\hat{\beta}_2$ and $\hat{\beta}_4$ mean.

(10 marks)

(b) Use the estimation results for regression equation (1) to test the *individual* significance of each of the slope coefficient estimates $\hat{\beta}_2$ for $\ln NOX_i$ and $\hat{\beta}_4$ for $ROOMS_i$. For each test, state the null and alternative hypotheses, show how you calculate the required test statistic, and state its null distribution. Which of these two slope coefficient estimates are individually significant at the 5 percent significance level? Which of these two slope coefficient estimates are individually significant at the 1 percent significance level?

(10 marks)

(c) Use the estimation results for regression equation (1) to test the *joint* significance of all the slope coefficient estimates at the 1 percent significance level (i.e., for significance level $\alpha = 0.01$). State the null and alternative hypotheses, show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test.

(10 marks)

(d) Use the estimation results for regression equation (1) to construct a two-sided 95 percent confidence interval for the slope coefficient β_3 of $\ln DIST_i$. Explain how you would use the two-sided 95 percent confidence interval you have computed for β_3 to perform a two-tail test of the hypothesis that $\beta_3 = 0$ at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$).

(10 marks)

(e) Use the estimation results for regression equation (1) to test the proposition that $\beta_2 = -1$, i.e., to test the proposition that the marginal effect of $\ln NOX_i$ on $\ln P_i$ is equal to -1. Explain in words what this proposition means. Perform the test at the 5 percent significance level. State the null and alternative hypotheses. Show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Would your inference be the same at the 10 percent significance level?

3. (continued)

(10 marks)

(f) Use the estimation results for regression equation (1) to test the proposition that $\beta_5 < 0$, i.e., to test the proposition that the marginal effect of STRAT_i on $\ln P_i$ is *less* than zero. Explain in words what this proposition means. Perform the test at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$). State the null and alternative hypotheses. Show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test.

(10 marks)

(g) Use the estimation results for regression equation (1) to test the proposition that $\beta_2 = \beta_3$, i.e., that the marginal effect of $\ln \text{NOX}_i$ on $\ln P_i$ equals the marginal effect of $\ln \text{DIST}_i$ on $\ln P_i$. Explain in words what this proposition means. Perform the test at the 1 percent significance level. State the null hypothesis H_0 and the alternative hypothesis H_1. Write the *restricted* regression equation implied by the null hypothesis H_0. OLS estimation of this *restricted* regression equation yields a Residual Sum-of-Squares value of **RSS = 41.9532**. Use this information, together with the results from OLS estimation of equation (1), to calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test.

(50 marks)

4. You are conducting an econometric investigation into the prices of recently sold houses in a single urban area. The sample data consist of 88 observations on the following variables:

P_i = the selling price of house i, in *thousands* of dollars;

HS_i = the house size of house i, in *hundreds* of square feet;

YS_i = the yard size of house i, in *hundreds* of square feet;

DC_i = an indicator variable defined such that $DC_i = 1$ if house i is a colonial-style house, and $DC_i = 0$ if house i is not a colonial-style house.

The regression model you propose to use is the log-log (double-log) regression equation

$$\ln P_i = \beta_1 + \beta_2 \ln HS_i + \beta_3 \ln YS_i + \beta_4 (\ln HS_i)^2 + \beta_5 (\ln YS_i)^2 + \beta_6 (\ln HS_i)(\ln YS_i) + \beta_7 DC_i + u_i \quad \textbf{(1)}$$

where the β_j (j = 1, 2, …, 7) are regression coefficients, $\ln X_i$ denotes the natural logarithm of the variable X_i, and u_i is a random error term.

Using the sample data described above, your research assistant computes OLS estimates of regression equation (1) and of three restricted versions of equation (1). For each of the sample regression equations estimated on the N = 88 observations, the following table contains the OLS coefficient estimates (with estimated standard errors in parentheses below the coefficient estimates) and the summary statistics RSS (residual sum-of-squares), TSS (total sum-of-squares), and N (number of sample observations).

4. (continued)

<u>Question 4</u>: OLS Sample Regression Equations for $\ln P_i$

Regressors		(1)	(2)	(3)	(4)
Intercept	$\hat{\beta}_1$	9.321 (2.500)	2.767 (0.5926)	8.456 (2.133)	2.639 (0.2443)
$\ln HS_i$	$\hat{\beta}_2$	−3.029 (1.398)	0.7489 (0.08122)	−3.045 (1.385)	0.7500 (0.08063)
$\ln YS_i$	$\hat{\beta}_3$	−0.2492 (0.5826)	0.1123 (0.2383)	0.1468 (0.03756)	0.1681 (0.03815)
$(\ln HS_i)^2$	$\hat{\beta}_4$	0.5509 (0.2536)	---	0.6217 (0.2266)	---
$(\ln YS_i)^2$	$\hat{\beta}_5$	0.01278 (0.02715)	0.006081 (0.02563)	---	---
$(\ln HS_i)(\ln YS_i)$	$\hat{\beta}_6$	0.09207 (0.1431)	---	---	---
DC_i	$\hat{\beta}_7$	0.09456 (0.04353)	0.06752 (0.04344)	0.09559 (0.04258)	0.06607 (0.04277)
	RSS =	2.59184	2.84127	2.60674	2.84320
	TSS =	8.01760	8.01760	8.01760	8.01760
	N =	88	88	88	88

<u>Note</u>: The symbol "---" means that the corresponding regressor was omitted from the estimated sample regression equation. The figures in parentheses below the coefficient estimates are the estimated standard errors. RSS is the Residual Sum-of-Squares, TSS is the Total Sum-of-Squares, and N is the number of sample observations.

(10 marks)

(a) Compare the goodness-of-fit to the sample data of the four sample regression equations (1), (2), (3) and (4) in the table. Calculate the value of an appropriate goodness-of-fit measure for each of the sample regression equations (1), (2), (3) and (4) in the table. Which of the four sample regression equations provides the best fit to the sample data? Which of the four sample regression equations provides the worst fit to the sample data?

(10 marks)

(b) Use the estimation results for regression equation (1) in the above table to perform a test of the proposition that colonial-style houses of any given house size and yard size have *higher* prices than houses of the same house size and yard size that are not colonial style. Perform the test at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$). State the null hypothesis H_0 and the alternative hypothesis H_1. Show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test.

4. (continued)

(10 marks)

(c) State the coefficient restrictions that regression equation (2) in the table imposes on regression equation (1). Explain in words what the restrictions mean. Use the estimation results given in the table to perform a test of these coefficient restrictions at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$). State the null and alternative hypotheses, show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Would your inference be the same at the 10 percent significance level (i.e., for significance level $\alpha = 0.10$)? Based on the outcome of the test, which of the two regression equations would you choose, equation (1) or equation (2)?

(10 marks)

(d) State the coefficient restrictions that regression equation (3) in the table imposes on regression equation (1). Explain in words what the restrictions mean. Use the estimation results given in the table to perform a test of these coefficient restrictions at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$). State the null and alternative hypotheses, show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Would your inference be the same at the 10 percent significance level (i.e., for significance level $\alpha = 0.10$)? Based on the outcome of the test, which of the two regression equations would you choose, equation (1) or equation (3)?

(10 marks)

(e) State the coefficient restrictions that regression equation (4) in the table imposes on regression equation (1). Explain in words what the restrictions mean. Use the estimation results given in the table to perform a test of these coefficient restrictions at the 5 percent significance level (i.e., for significance level $\alpha = 0.05$). State the null and alternative hypotheses, show how you calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Would your inference be the same at the 10 percent significance level (i.e., for significance level $\alpha = 0.10$)? Based on the outcome of the test, which of the two regression equations would you choose, equation (1) or equation (4)?

(50 marks)

5. You are conducting an econometric investigation into the hourly wage rates of male and female employees. Your particular interest is in comparing the determinants of wage rates for female and male workers. The sample data consist of a random sample of observations on 526 paid employees, 252 of whom are females and 274 of whom are males. The sample data provide observations on the following variables:

W_i = the hourly wage rate of employee i, measured in dollars per hour;
S_i = the number of years of formal education completed by employee i, in years;
A_i = the age of employee i, in years;
T_i = the firm tenure of employee i, in years;
F_i = a female indicator variable, defined such that $F_i = 1$ if employee i is female and $F_i = 0$ if employee i is male.

Your astute client proposes that the following pooled regression equation be estimated on the full sample of 526 observations for male and female employees:

$$\begin{aligned} \ln W_i = {} & \beta_1 + \beta_2 S_i + \beta_3 A_i + \beta_4 A_i^2 + \beta_5 T_i + \beta_6 S_i T_i + \beta_7 F_i \\ & + \beta_8 F_i S_i + \beta_9 F_i A_i + \beta_{10} F_i A_i^2 + \beta_{11} F_i T_i + \beta_{12} F_i S_i T_i + u_i \end{aligned} \tag{1}$$

where the β_j (j = 1, 2, …, 12) are regression coefficients, $\ln W_i$ denotes the natural logarithm of the variable W_i, and u_i is a random error term.

Using the sample data described above, your research assistant computes OLS estimates of regression equation (1) and OLS estimates of several restricted versions of equation (1) on the full sample of N = 526 observations for male and female employees.

OLS Estimates of Equation (1), Question 5: (standard errors in parentheses)

$\hat{\beta}_1 = -0.5667$ (0.2385) $\hat{\beta}_7 = 0.3593$ (0.3373)

$\hat{\beta}_2 = 0.05937$ (0.01104) $\hat{\beta}_8 = 0.01062$ (0.01684)

$\hat{\beta}_3 = 0.07980$ (0.01216) $\hat{\beta}_9 = -0.03847$ (0.01715)

$\hat{\beta}_4 = -0.0009314$ (0.0001548) $\hat{\beta}_{10} = 0.0004224$ (0.0002192)

$\hat{\beta}_5 = -0.01057$ (0.01128) $\hat{\beta}_{11} = 0.01850$ (0.02652)

$\hat{\beta}_6 = 0.002279$ (0.0008765) $\hat{\beta}_{12} = -0.002107$ (0.002187)

$\text{RSS} = \sum_{i=1}^{N} \hat{u}_i^2 = 80.5693$; $\text{TSS} = \sum_{i=1}^{N} \left(\ln W_i - \overline{\ln W}\right)^2 = 148.330$; $N = 526$

5. (continued)

The RSS (Residual Sum-of-Squares) and TSS (Total Sum-of-Squares) from OLS estimation of regression equation (1) on the N = 526 observations are:

$$\text{RSS} = \sum_{i=1}^{N} \hat{u}_i^2 = 80.5693; \qquad \text{TSS} = \sum_{i=1}^{N} \left(\ln W_i - \overline{\ln W}\right)^2 = 148.330; \qquad N = 526$$

(10 marks)

(a) Use the estimation results for regression equation (1) to compute OLS estimates of the slope coefficients of the regressors S_i, A_i and T_i for *male* employees. Use the estimation results for regression equation (1) to compute OLS estimates of the slope coefficients of the regressors S_i, A_i and T_i for *female* employees.

(10 marks)

(b) Write the expression (or formula) for the marginal effect of A_i on $\ln W_i$ for *male* employees implied by regression equation (1). Write the expression (or formula) for the marginal effect of A_i on $\ln W_i$ for *female* employees implied by regression equation (1). Compute a test of the null hypothesis that the marginal effect of A_i on $\ln W_i$ for *male* employees is equal to the marginal effect of A_i on $\ln W_i$ for *female* employees for any values of S_i, A_i and T_i. State the null hypothesis H_0 and the alternative hypothesis H_1. Write the *restricted* regression equation implied by the null hypothesis H_0. OLS estimation of this *restricted* regression equation yields a Residual Sum-of-Squares value of **RSS = 81.7242**. Use this information, together with the results from OLS estimation of equation (1), to calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Choose an appropriate significance level for the test.

(15 marks)

(c) Compute a test of the null hypothesis that *both* the marginal effect of S_i on $\ln W_i$ *and* the marginal effect of T_i on $\ln W_i$ are equal for male and female employees in regression equation (1). State the null hypothesis H_0 and the alternative hypothesis H_1. Write the *restricted* regression equation implied by the null hypothesis H_0. OLS estimation of this *restricted* regression equation yields a Residual Sum-of-Squares value of **RSS = 80.8747**. Use this information, together with the results from OLS estimation of equation (1), to calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Choose an appropriate significance level for the test.

5. (continued)

(15 marks)

(d) Compute a test of the proposition that the mean log-wage of *female* employees with any given values of S_i, A_i and T_i equals the mean log-wage of *male* employees with the same values of S_i, A_i and T_i. State the coefficient restrictions on regression equation (1) implied by this proposition; that is, state the null hypothesis H_0 and the alternative hypothesis H_1. Write the *restricted* regression equation implied by the null hypothesis H_0. OLS estimation of this *restricted* regression equation yields a Residual Sum-of-Squares value of **RSS = 93.1805**. Use this information, together with the results from OLS estimation of equation (1), to calculate the required test statistic, and state its null distribution. State the decision rule you use, and the inference you would draw from the test. Choose an appropriate significance level for the test.

Selected Formulas

For the Simple (Two-Variable) Linear Regression Model

$$Y_i = \beta_1 + \beta_2 X_i + u_i \qquad (i = 1, \ldots, N)$$

- Deviations from sample means are defined as:

$$y_i \equiv Y_i - \overline{Y}; \qquad x_i \equiv X_i - \overline{X};$$

where

$\overline{Y} = \Sigma_i Y_i / N = \dfrac{\Sigma_i Y_i}{N}$ is the sample mean of the Y_i values;

$\overline{X} = \Sigma_i X_i / N = \dfrac{\Sigma_i X_i}{N}$ is the sample mean of the X_i values.

- Formulas for the variance of the OLS intercept coefficient estimator $\hat{\beta}_1$ and the covariance of the OLS coefficient estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ in the two-variable linear regression model:

$$\mathrm{Var}\left(\hat{\beta}_1\right) = \frac{\sigma^2 \Sigma_i X_i^2}{N\,\Sigma_i \left(X_i - \overline{X}\right)^2} = \frac{\sigma^2 \Sigma_i X_i^2}{N\,\Sigma_i x_i^2};$$

$$\mathrm{Cov}\left(\hat{\beta}_1, \hat{\beta}_2\right) = -\overline{X}\left(\frac{\sigma^2}{\Sigma_i x_i^2}\right).$$

- Formulas for the variance of the conditional predictor $\hat{Y}_0 = \hat{\beta}_1 + \hat{\beta}_2 X_0$:

  - When $\hat{Y}_0$ is used as a mean predictor of $E\left(Y_0 \middle| X_0\right) = \beta_1 + \beta_2 X_0$,

$$\mathrm{Var}\left(\hat{Y}_0^m\right) = \sigma^2\left[\frac{1}{N} + \frac{\left(X_0 - \overline{X}\right)^2}{\Sigma_i x_i^2}\right].$$

  - When $\hat{Y}_0$ is used as an individual predictor of $Y_0 \big| X_0 = \beta_1 + \beta_2 X_0 + u_0$,

$$\mathrm{Var}\left(\hat{Y}_0\right) = \sigma^2 + \sigma^2\left[\frac{1}{N} + \frac{\left(X_0 - \overline{X}\right)^2}{\Sigma_i x_i^2}\right].$$

Selected Formulas (continued)

For the Multiple (Three-Variable) Linear Regression Model

$$Y_i = \beta_1 + \beta_2 X_{2i} + \beta_3 X_{3i} + u_i \qquad (i = 1, \ldots, N)$$

- Deviations from sample means are defined as:

$$y_i \equiv Y_i - \overline{Y}; \qquad x_{2i} \equiv X_{2i} - \overline{X}_2; \qquad x_{3i} \equiv X_{3i} - \overline{X}_3;$$

where

$\overline{Y} = \Sigma_i Y_i / N = \dfrac{\Sigma_i Y_i}{N}$ is the sample mean of the Y_i values;

$\overline{X}_{2i} = \Sigma_i X_{2i} / N = \dfrac{\Sigma_i X_{2i}}{N}$ is the sample mean of the X_{2i} values;

$\overline{X}_{3i} = \Sigma_i X_{3i} / N = \dfrac{\Sigma_i X_{3i}}{N}$ is the sample mean of the X_{3i} values.

- The OLS slope coefficient estimators $\hat{\beta}_2$ and $\hat{\beta}_3$ in deviation-from-means form are:

$$\hat{\beta}_2 = \frac{\left(\Sigma_i x_{3i}^2\right)\left(\Sigma_i x_{2i} y_i\right) - \left(\Sigma_i x_{2i} x_{3i}\right)\left(\Sigma_i x_{3i} y_i\right)}{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i}^2\right) - \left(\Sigma_i x_{2i} x_{3i}\right)^2};$$

$$\hat{\beta}_3 = \frac{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i} y_i\right) - \left(\Sigma_i x_{2i} x_{3i}\right)\left(\Sigma_i x_{2i} y_i\right)}{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i}^2\right) - \left(\Sigma_i x_{2i} x_{3i}\right)^2}.$$

- Formulas for the variances and covariances of the slope coefficient estimators $\hat{\beta}_2$ and $\hat{\beta}_3$:

$$\mathrm{Var}\left(\hat{\beta}_2\right) = \frac{\sigma^2 \Sigma_i x_{3i}^2}{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i}^2\right) - \left(\Sigma_i x_{2i} x_{3i}\right)^2};$$

$$\mathrm{Var}\left(\hat{\beta}_3\right) = \frac{\sigma^2 \Sigma_i x_{2i}^2}{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i}^2\right) - \left(\Sigma_i x_{2i} x_{3i}\right)^2};$$

$$\mathrm{Cov}\left(\hat{\beta}_2, \hat{\beta}_3\right) = \frac{\sigma^2 \Sigma_i x_{2i} x_{3i}}{\left(\Sigma_i x_{2i}^2\right)\left(\Sigma_i x_{3i}^2\right) - \left(\Sigma_i x_{2i} x_{3i}\right)^2}.$$

Percentage Points of the t-Distribution

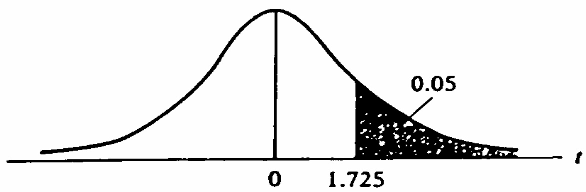
TABLE D.2
Percentage points of the *t* distribution

Example

Pr(t > 2.086) = 0.025

Pr(t > 1.725) = 0.05 for df = 20

Pr($|t|$ > 1.725) = 0.10



df \ Pr	0.25 0.50	0.10 0.20	0.05 0.10	0.025 0.05	0.01 0.02	0.005 0.010	0.001 0.002
1	1.000	3.078	6.314	12.706	31.821	63.657	318.31
2	0.816	1.886	2.920	4.303	6.965	9.925	22.327
3	0.765	1.638	2.353	3.182	4.541	5.841	10.214
4	0.741	1.533	2.132	2.776	3.747	4.604	7.173
5	0.727	1.476	2.015	2.571	3.365	4.032	5.893
6	0.718	1.440	1.943	2.447	3.143	3.707	5.208
7	0.711	1.415	1.895	2.365	2.998	3.499	4.785
8	0.706	1.397	1.860	2.306	2.896	3.355	4.501
9	0.703	1.383	1.833	2.262	2.821	3.250	4.297
10	0.700	1.372	1.812	2.228	2.764	3.169	4.144
11	0.697	1.363	1.796	2.201	2.718	3.106	4.025
12	0.695	1.356	1.782	2.179	2.681	3.055	3.930
13	0.694	1.350	1.771	2.160	2.650	3.012	3.852
14	0.692	1.345	1.761	2.145	2.624	2.977	3.787
15	0.691	1.341	1.753	2.131	2.602	2.947	3.733
16	0.690	1.337	1.746	2.120	2.583	2.921	3.686
17	0.689	1.333	1.740	2.110	2.567	2.898	3.646
18	0.688	1.330	1.734	2.101	2.552	2.878	3.610
19	0.688	1.328	1.729	2.093	2.539	2.861	3.579
20	0.687	1.325	1.725	2.086	2.528	2.845	3.552
21	0.686	1.323	1.721	2.080	2.518	2.831	3.527
22	0.686	1.321	1.717	2.074	2.508	2.819	3.505
23	0.685	1.319	1.714	2.069	2.500	2.807	3.485
24	0.685	1.318	1.711	2.064	2.492	2.797	3.467
25	0.684	1.316	1.708	2.060	2.485	2.787	3.450
26	0.684	1.315	1.706	2.056	2.479	2.779	3.435
27	0.684	1.314	1.703	2.052	2.473	2.771	3.421
28	0.683	1.313	1.701	2.048	2.467	2.763	3.408
29	0.683	1.311	1.699	2.045	2.462	2.756	3.396
30	0.683	1.310	1.697	2.042	2.457	2.750	3.385
40	0.681	1.303	1.684	2.021	2.423	2.704	3.307
60	0.679	1.296	1.671	2.000	2.390	2.660	3.232
120	0.677	1.289	1.658	1.980	2.358	2.617	3.160
∞	0.674	1.282	1.645	1.960	2.326	2.576	3.090

Note: The smaller probability shown at the head of each column is the area in one tail; the larger probability is the area in both tails.

Source: From E. S. Pearson and H. O. Hartley, eds., *Biometrika Tables for Statisticians*, vol. 1, 3d ed., table 12, Cambridge University Press, New York, 1966. Reproduced by permission of the editors and trustees of *Biometrika*.

Source: Damodar N. Gujarati, *Basic Econometrics*, Third Edition. New York: McGraw-Hill, 1995, p. 809.

Selected Upper Percentage Points of the F-Distribution

TABLE D.3
Upper percentage points of the *F* distribution (*continued*)

df for denominator N_2	Pr	df for numerator N_1: 1	2	3	4	5	6	7	8	9	10	11	12
	.25	1.40	1.48	1.47	1.45	1.44	1.42	1.41	1.40	1.39	1.39	1.38	1.37
22	.10	2.95	2.56	2.35	2.22	2.13	2.06	2.01	1.97	1.93	1.90	1.88	1.86
	.05	4.30	3.44	3.05	2.82	2.66	2.55	2.46	2.40	2.34	2.30	2.26	2.23
	.01	7.95	5.72	4.82	4.31	3.99	3.76	3.59	3.45	3.35	3.26	3.18	3.12
	.25	1.39	1.47	1.46	1.44	1.43	1.41	1.40	1.39	1.38	1.38	1.37	1.36
24	.10	2.93	2.54	2.33	2.19	2.10	2.04	1.98	1.94	1.91	1.88	1.85	1.83
	.05	4.26	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25	2.21	2.18
	.01	7.82	5.61	4.72	4.22	3.90	3.67	3.50	3.36	3.26	3.17	3.09	3.03
	.25	1.38	1.46	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.37	1.36	1.35
26	.10	2.91	2.52	2.31	2.17	2.08	2.01	1.96	1.92	1.88	1.86	1.84	1.81
	.05	4.23	3.37	2.98	2.74	2.59	2.47	2.39	2.32	2.27	2.22	2.18	2.15
	.01	7.72	5.53	4.64	4.14	3.82	3.59	3.42	3.29	3.18	3.09	3.02	2.96
	.25	1.38	1.46	1.45	1.43	1.41	1.40	1.39	1.38	1.37	1.36	1.35	1.34
28	.10	2.89	2.50	2.29	2.16	2.06	2.00	1.94	1.90	1.87	1.84	1.81	1.79
	.05	4.20	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19	2.15	2.12
	.01	7.64	5.45	4.57	4.07	3.75	3.53	3.36	3.23	3.12	3.03	2.96	2.90
	.25	1.38	1.45	1.44	1.42	1.41	1.39	1.38	1.37	1.36	1.35	1.35	1.34
30	.10	2.88	2.49	2.28	2.14	2.05	1.98	1.93	1.88	1.85	1.82	1.79	1.77
	.05	4.17	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16	2.13	2.09
	.01	7.56	5.39	4.51	4.02	3.70	3.47	3.30	3.17	3.07	2.98	2.91	2.84
	.25	1.36	1.44	1.42	1.40	1.39	1.37	1.36	1.35	1.34	1.33	1.32	1.31
40	.10	2.84	2.44	2.23	2.09	2.00	1.93	1.87	1.83	1.79	1.76	1.73	1.71
	.05	4.08	3.23	2.84	2.61	2.45	2.34	2.25	2.18	2.12	2.08	2.04	2.00
	.01	7.31	5.18	4.31	3.83	3.51	3.29	3.12	2.99	2.89	2.80	2.73	2.66
	.25	1.35	1.42	1.41	1.38	1.37	1.35	1.33	1.32	1.31	1.30	1.29	1.29
60	.10	2.79	2.39	2.18	2.04	1.95	1.87	1.82	1.77	1.74	1.71	1.68	1.66
	.05	4.00	3.15	2.76	2.53	2.37	2.25	2.17	2.10	2.04	1.99	1.95	1.92
	.01	7.08	4.98	4.13	3.65	3.34	3.12	2.95	2.82	2.72	2.63	2.56	2.50
	.25	1.34	1.40	1.39	1.37	1.35	1.33	1.31	1.30	1.29	1.28	1.27	1.26
120	.10	2.75	2.35	2.13	1.99	1.90	1.82	1.77	1.72	1.68	1.65	1.62	1.60
	.05	3.92	3.07	2.68	2.45	2.29	2.17	2.09	2.02	1.96	1.91	1.87	1.83
	.01	6.85	4.79	3.95	3.48	3.17	2.96	2.79	2.66	2.56	2.47	2.40	2.34
	.25	1.33	1.39	1.38	1.36	1.34	1.32	1.31	1.29	1.28	1.27	1.26	1.25
200	.10	2.73	2.33	2.11	1.97	1.88	1.80	1.75	1.70	1.66	1.63	1.60	1.57
	.05	3.89	3.04	2.65	2.42	2.26	2.14	2.06	1.98	1.93	1.88	1.84	1.80
	.01	6.76	4.71	3.88	3.41	3.11	2.89	2.73	2.60	2.50	2.41	2.34	2.27
	.25	1.32	1.39	1.37	1.35	1.33	1.31	1.29	1.28	1.27	1.25	1.24	1.24
∞	.10	2.71	2.30	2.08	1.94	1.85	1.77	1.72	1.67	1.63	1.60	1.57	1.55
	.05	3.84	3.00	2.60	2.37	2.21	2.10	2.01	1.94	1.88	1.83	1.79	1.75
	.01	6.63	4.61	3.78	3.32	3.02	2.80	2.64	2.51	2.41	2.32	2.25	2.18

Source: Damodar N. Gujarati, *Basic Econometrics*, Third Edition. New York: McGraw-Hill, 1995, p. 814.