QUEEN'S UNIVERSITY AT KINGSTON
Department of Economics

ECONOMICS 351* - Winter Term 2008

Introductory Econometrics

Winter Term 2008 **FINAL EXAMINATION** M.G. Abbott

DATE: **Tuesday April 15, 2008**

TIME: **Three (3) hours (180 minutes); 9:00 a.m. – 12 noon**

INSTRUCTIONS: The examination is divided into two parts.

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

PART B contains **3** questions; students are required to **answer ALL THREE** questions in Part B.

- Answer all questions in the exam booklets provided. Be sure **your *student number*** is printed clearly and legibly on the front page 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.
- This exam is **CONFIDENTIAL**. This question paper must be submitted in its entirety with your answer booklet(s); otherwise your exam will not be marked.
- **Please write legibly.** GOOD LUCK!

QUESTION 5 – Winter Term 2008 (30 marks)

(30 marks)

5. You are investigating the relationship between the final exam grades of university students in an introductory economics course and those students' class attendance, as measured by the percentage of classes each student attended during the term. You also have sample data on two additional explanatory variables: each student's cumulative GPA (Grade Point Average) prior to the term in which the introductory economics course was taken; and each student's score on a standardized college entrance exam, the ACT exam. You have sample data for 680 students on the following variables:

finalpct_i = final exam grade of the i-th student, measured in *percentage points*;

attrate_i = percentage of classes attended by the i-th student during the term, measured in *percentage points*;

GPA_i = cumulative Grade Point Average (GPA) of the i-th student prior to the term in which the introductory economics course was taken, measured *out of 4.0*;

ACT_i = ACT score of the i-th student on the ACT college entrance exam, measured in *points*.

Using the given sample data on 680 students, your trusty research assistant has estimated regression equation (1) and obtained the following estimation results (with estimated *standard errors* given in parentheses below the coefficient estimates):

$$\text{finalpct}_i = \beta_0 + \beta_1\text{attrate}_i + \beta_2\text{GPA}_i + \beta_3\text{GPA}_i^2 + \beta_4\text{GPA}_i\text{attrate}_i + \beta_5\text{ACT}_i + u_i \qquad \ldots \textbf{(1)}$$

$\hat{\beta}_0 = 62.261$ (10.026) $\quad \hat{\beta}_1 = -0.061858$ (0.12175) $\quad \hat{\beta}_2 = -20.117$ (5.7271)

$\hat{\beta}_3 = 3.8356$ (1.1959) $\quad \hat{\beta}_4 = 0.057852$ (0.051337) $\quad \hat{\beta}_5 = 0.90182$ (0.13359)

$\hat{Cov}(\hat{\beta}_1, \hat{\beta}_2) = 0.13006 \quad \hat{Cov}(\hat{\beta}_1, \hat{\beta}_3) = 0.069642 \quad \hat{Cov}(\hat{\beta}_1, \hat{\beta}_4) = -0.0060842$

$\hat{Cov}(\hat{\beta}_2, \hat{\beta}_3) = -5.0408 \quad \hat{Cov}(\hat{\beta}_2, \hat{\beta}_4) = -0.066730 \quad \hat{Cov}(\hat{\beta}_2, \hat{\beta}_5) = 0.073538$

$RSS = 73080.413 \quad TSS = 94137.169 \quad N = 680$

RSS is the Residual Sum-of-Squares and TSS is the Total Sum-of-Squares for sample regression equation (1). Sample size N = 680. $\hat{Cov}(\hat{\beta}_j, \hat{\beta}_h)$ is the estimated covariance between coefficient estimates $\hat{\beta}_j$ and $\hat{\beta}_h$. Estimated *standard errors* are given in parentheses below the coefficient estimates $\hat{\beta}_j$ (j = 0, 1, …, 5).

Question 5(a): (10 marks)

(a) Compute a test of the proposition that class attendance as measured by attrate_i, the percentage of classes attended by the i-th student during the term, has no effect on students' final exam grade. Perform the test at the **5 percent significance level** (i.e., for significance level $\alpha = 0.05$). 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 = 73934.237**. 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 at the **5 percent significance level**.

Answer to Question 5(a): (10 marks)

The **marginal effect of attrate$_i$** is:

$$\text{finalpct}_i = \beta_0 + \beta_1\text{attrate}_i + \beta_2\text{GPA}_i + \beta_3\text{GPA}_i^2 + \beta_4\text{GPA}_i\text{attrate}_i + \beta_5\text{ACT}_i + u_i \quad \ldots \textbf{(1)}$$

$$\frac{\partial\,\text{finalpct}_i}{\partial\,\text{attrate}_i} = \frac{\partial\,E(\text{finalpct}_i \mid \text{attrate}_i, \text{GPA}_i, \text{ACT}_i)}{\partial\,\text{attrate}_i} = \beta_1 + \beta_4\text{GPA}_i$$ **(1 mark)**

Test that the **marginal effect of attrate$_i$** is ***equal* to *zero***.

<u>Test</u>: H_0: $\beta_1 = 0$ and $\beta_4 = 0$ ***versus*** **H_1: $\beta_1 \neq 0$ and/or $\beta_4 \neq 0$** **(1 mark)**

Answer to Question 5(a) -- continued: (10 marks)

Restricted regression implied by H_0: **(1 mark)**

$$\text{finalpct}_i = \beta_0 + \beta_2 \text{GPA}_i + \beta_3 \text{GPA}_i^2 + \beta_5 \text{ACT}_i + u_i$$

Compute *sample value* of general F-statistic:

$$F = \frac{(\text{RSS}_0 - \text{RSS}_1)/(\text{df}_0 - \text{df}_1)}{\text{RSS}_1/\text{df}_1} \sim F[\text{df}_0 - \text{df}_1, \text{df}_1] = F[K - K_0, N - K] = \mathbf{F[2, 674]}$$ **(1 mark)**

where: $\text{RSS}_0 = 73934.237$ with $\text{df}_0 = N - K_0 = 680 - 4 = 676$
$\text{RSS}_1 = 73080.413$ with $\mathbf{df_1 = N - K} = 680 - 6 = \mathbf{674}$

$\text{RSS}_1/\text{df}_1 = 73080.413/674 = \mathbf{108.4279}$

$\text{RSS}_0 - \text{RSS}_1 = 73934.237 - 73080.413 = \mathbf{853.824}$ and $\mathbf{df_0 - df_1 = 2}$

$$F_0 = \frac{853.824/2}{108.4279} = \frac{426.912}{108.4279} = \mathbf{3.9373} = \underline{\mathbf{3.94}}$$ **(3 marks)**

Answer to Question 5(a) -- continued: (10 marks)

Decision Rule: at significance level α **(1 mark)**

- If $F_0 > F_\alpha[2, 674]$, ***reject*** **H$_0$** at the 100α percent significance level.
- If $F_0 \le F_\alpha[2, 674]$, ***retain*** **(do not reject) H$_0$** at the 100α percent significance level.

Critical value of F[2, 674] at **5% significance level** (α = 0.05) is **$F_{0.05}[2, \infty]$ = 3.00** ← **use 5%** **(1 mark)**

Critical value of F[2, 674] at 1% significance level (α = 0.01) is **$F_{0.01}[2, \infty]$ = 4.61**

Inference: at the **5%** significance level. **(1 mark)**

- Since **F_0 = 3.94 > 3.00 = $F_{0.05}[2, \infty]$, *reject* H$_0$ at 5% significance level.** ← **use 5%**
- Since **F_0 = 3.94 < 4.61 = $F_{0.01}[2, \infty]$, *retain* H$_0$ at 1% significance level.**

Question Q(b): (10 marks)

(b) Compute a test of the proposition that students' cumulative Grade Point Average as measured by GPA_i has no effect on students' final exam grade. Perform the test at the **5 percent significance level** (i.e., for significance level $\alpha = 0.05$). 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 an R-squared value of $\mathbf{R^2 = 0.1701}$. 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 at the **5 percent significance level**.

<u>Answer to Question 5(b)</u>: (10 marks)

$$\text{finalpct}_i = \beta_0 + \beta_1\text{attrate}_i + \beta_2\text{GPA}_i + \beta_3\text{GPA}_i^2 + \beta_4\text{GPA}_i\text{attrate}_i + \beta_5\text{ACT}_i + u_i \quad \ldots \textbf{(1)}$$

The **marginal effect of $\mathbf{GPA_i}$** is:

$$\frac{\partial\,\text{finalpct}_i}{\partial\,\text{GPA}_i} = \frac{\partial\,E(\text{finalpct}_i\,|\text{attrate}_i, \text{GPA}_i, \text{ACT}_i)}{\partial\,\text{GPA}_i} = \beta_2 + 2\beta_3\text{GPA}_i + \beta_4\text{attrate}_i$$

Test that the **marginal effect of $\mathbf{GPA_i}$** is ***equal* to *zero*** at the **5% significance level**.

<u>Test</u>: H_0: $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$ **(1 mark)**
H_1: $\beta_2 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ **(1 mark)**

Answer to Question 5(b) -- continued: (10 marks)

$$\text{finalpct}_i = \beta_0 + \beta_1 \text{attrate}_i + \beta_2 \text{GPA}_i + \beta_3 \text{GPA}_i^2 + \beta_4 \text{GPA}_i \text{attrate}_i + \beta_5 \text{ACT}_i + u_i \quad \ldots \textbf{(1)}$$

Restricted regression implied by H_0: **(1 mark)**

$$\text{finalpct}_i = \beta_0 + \beta_1 \text{attrate}_i + \beta_5 \text{ACT}_i + u_i$$

$\mathbf{K_0 = 3}$,
$\mathbf{df_1 = N - K} = 680 - 6 = \mathbf{674}$
$\mathbf{df_0 - df_1 = 677 - 674 = 3}$

Compute *sample value* of general F-statistic:

$$F = \frac{(R_U^2 - R_R^2)/(df_0 - df_1)}{(1 - R_U^2)/df_1} \sim F[df_0 - df_1, df_1] = F[K - K_0, N - K] = \mathbf{F[3, 674]}$$ **(1 mark)**

where: $R_R^2 = \mathbf{0.1701}$ with $df_0 = N - K_0 = 680 - 3 = 677$

$$R_U^2 = 1 - \frac{RSS_1}{TSS} = 1 - \frac{73080.413}{94137.169} = 1 - 0.77632 = \underline{\mathbf{0.22368}}$$

$$F_0 = \frac{(0.22368 - 0.1701)/3}{0.77632/674} = \frac{0.05358/3}{0.00115181} = \frac{0.01786}{0.00115181} = \mathbf{15.5061} = \underline{\mathbf{15.51}}$$ **(3 marks)**

Answer to Question 5(b) -- continued: (10 marks)

Decision Rule: at significance level α **(1 mark)**

- If $F_0 > F_\alpha[3, 674]$, ***reject* H₀** at the 100α percent significance level.
- If $F_0 \le F_\alpha[3, 674]$, ***retain* (do not reject) H₀** at the 100α percent significance level.

Critical value of F[3, 674] at **5%** sig level (α = 0.05) is $\mathbf{F_{0.05}[3, \infty] = \underline{2.60}}$ ← **use 5%** **(1 mark)**

Critical value of F[3, 674] at 1% sig level (α = 0.01) is $\mathbf{F_{0.01}[3, \infty] = 3.78}$

Inference: is **at the 5%** significance level. **(1 mark)**

- Since $\mathbf{F_0 = \underline{15.51} > \underline{2.60} = F_{0.05}[3, \infty]}$, ***reject* H_0 at 5% significance level.** ← **use 5%**
- Since $\mathbf{F_0 = 15.51 > 3.78 = F_{0.01}[3, \infty]}$, ***reject* H_0 at 1% significance level.**

Question 5(c): (10 marks)

(c) The average student in the course attends 80 percent of the classes and has a cumulative Grade Point Average of 2.6; that is, the sample mean value of attrate_i equals 80, and the sample mean value of GPA_i equals 2.6. Write the expression (or formula) for the marginal effect on finalpct_i of attrate_i implied by regression equation (1). Use the estimation results for regression equation (1) to test the proposition that the marginal effect of class attendance (attrate_i) on students' final exam grade (finalpct_i) equals zero for the average student whose class attendance rate is 80 percent and GPA is 2.6. Perform the test at the **1 percent significance level** (i.e., for significance level $\alpha = 0.01$). State the null hypothesis H_0 and the alternative hypothesis H_1 implied by this proposition. 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 at the **1 percent significance level**.

Answer to Question 5(c): (10 marks)

Proposition to test: The **marginal effect of attrate_i** with respect to **finalpct_i** is ***equal* to *zero*** for **$\text{attrate}_i = 80$** and **$\text{GPA}_i = 2.6$**.

$$\text{finalpct}_i = \beta_0 + \beta_1\text{attrate}_i + \beta_2\text{GPA}_i + \beta_3\text{GPA}_i^2 + \beta_4\text{GPA}_i\text{attrate}_i + \beta_5\text{ACT}_i + u_i \quad \ldots \textbf{(1)}$$

$$\frac{\partial\,\text{finalpct}_i}{\partial\,\text{attrate}_i} = \frac{\partial\,E(\text{finalpct}_i\,|\text{attrate}_i, \text{GPA}_i, \text{ACT}_i)}{\partial\,\text{attrate}_i} = \beta_1 + \beta_4\text{GPA}_i$$ **(1 mark)**

Test: H_0: $\beta_1 + \beta_4 2.6 = \beta_1 + 2.6\beta_4 = 0$ versus H_1: $\beta_1 + \beta_4 2.6 = \beta_1 + 2.6\beta_4 \neq 0$ **(1 mark)**

<u>Answer to Question 5(c) -- continued</u>: (10 marks)

Compute *sample value* of t-statistic or F-statistic:

$\hat{\beta}_1 + 2.6\hat{\beta}_4 = -0.061858 + 2.6(0.057852) = -0.061858 + 0.150415 = 0.088557$ **= <u>0.088557</u>** **(1 mark)**

$$t_0(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \frac{\hat{\beta}_1 + 2.6\hat{\beta}_4}{\hat{se}(\hat{\beta}_1 + 2.6\hat{\beta}_4)} = \frac{\hat{\beta}_1 + 2.6\hat{\beta}_4}{\sqrt{\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4)}} \quad \textbf{or} \quad F_0(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \frac{(\hat{\beta}_1 + 2.6\hat{\beta}_4)^2}{\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4)}$$

Compute $\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4)$ **or** $\hat{se}(\hat{\beta}_1 + 2.6\hat{\beta}_4)$: **(2 marks)**

$$\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \hat{Var}(\hat{\beta}_1) + (2.6)^2\,\hat{Var}(\hat{\beta}_4) + 5.2\hat{Cov}(\hat{\beta}_1, \hat{\beta}_4)$$

$$\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4) = (0.12175)^2 + (6.76)(0.051337)^2 + 5.2(-0.0060842)$$

$$= 0.0149231 + 0.0178159 - 0.0316378 = \mathbf{0.001001118} = \underline{\mathbf{0.00100112}}$$

$$\hat{se}(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \sqrt{\hat{Var}(\hat{\beta}_1 + 2.6\hat{\beta}_4)} = \sqrt{0.001001118} = \underline{\mathbf{0.03164045}}$$

Compute sample value of test statistic: **(2 marks)**

$$t_0(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \frac{0.088557}{0.03164045} = \underline{\mathbf{2.7989}} \sim \mathbf{t[674]}$$

$$F_0(\hat{\beta}_1 + 2.6\hat{\beta}_4) = \frac{(0.088557)^2}{0.001001118} = \frac{0.00784234}{0.001001118} = \underline{\mathbf{7.8336}} \sim \mathbf{F[1, 674]}$$

<u>Answer to Question 5(c) -- continued</u>: (10 marks)

<u>Decision rule</u> for a *two-tail* t-test and F-test: **(1 mark)**

1. If $|t_0| \le t_{\alpha/2}[N-K] = t_{\alpha/2}[\infty]$, ***retain* (do not reject) H_0** at the 100α % significance level.
If $F_0 \le F_\alpha[1, N-K] = F_\alpha[1, \infty]$, ***retain* (do not reject) H_0** at the 100α % significance level.

2. If $|t_0| > t_{\alpha/2}[N-K] = t_{\alpha/2}[\infty]$, ***reject* H_0** at the 100α % significance level.
If $F_0 > F_\alpha[1, N-K] = F_\alpha[1, \infty]$, ***reject* H_0** at the 100α % significance level.

<u>Critical values</u> at significance level α = 0.01, α/2 = 0.005: **(1 mark)**

$t_{0.005}[674] = \mathbf{t_{0.005}[\infty] = \underline{2.576}}$;

$F_{0.01}[1,674] = \mathbf{F_{0.01}[1,\infty] = \underline{6.63}}$

<u>Inference</u>: **(1 mark)**

$\mathbf{|t_0| = \underline{2.7989} > 2.576 = t_{0.005}[\infty]}$ ⇒ ***reject* H_0** at the **1% significance level**.

$\mathbf{F_0 = \underline{7.8336} > 6.63 = F_{0.01}[1,\infty]}$ ⇒ ***reject* H_0** at the **1% significance level**.