Queen's University
Department of Economics

ECON 351* -- Introductory Econometrics

ASSIGNMENT 3 – ANSWERS

Winter Term 2009 M.G. Abbott

TOPIC: **OLS Estimation and Inference in the Multiple Linear Regression Model**

INSTRUCTIONS:

- Answer all questions on standard-sized 8.5 x 11-inch paper.
- Answers need not be typewritten (document processed), but if hand-written must be legible. Illegible assignments will be returned unmarked.
- Please label clearly each answer with the appropriate question number and letter. Securely staple all answer sheets together, and make certain that your ***name(s)*** and ***student number(s)*** are printed clearly at the top of each answer sheet.
- Students submitting joint assignments must ensure that each student's name and student number are printed clearly at the top of each answer sheet. ***Submit only one copy of the assignment***.

MARKING: Marks for each question are indicated in parentheses. **Total marks for the assignment equal *132***. Marks are given for both content and presentation.

SOFT DUE DATE: **Tuesday March 31, 2009 by 4:00 p.m.**

HARD DUE DATE: **Friday April 3, 2009 by 4:00 p.m.**

- Assignments submitted **on or before** the soft due date will receive a bonus of 5 points to a maximum total mark of 132.
- Assignments submitted **after** the hard due date will be penalized 20 points per day.
- Please submit your assignments either to me in class, or by depositing them in the ECON 351 slot of the **Assignment Collection Box** located immediately **inside the double doors** on the **second floor of Dunning Hall** (opposite the elevator).

DATA FILE: **351assn3w09.raw** (a text-format, or ASCII-format, data file)

- ***Data Description:*** A random sample of 321 houses that were sold in a single metropolitan area in the years 1978 and 1981.
- ***Objective of the Analysis:*** The primary objective of the research project for which this dataset was originally assembled was to estimate empirically the effect on house prices of proximity to an incinerator, which presumably generates negative externalities for homeowners and others located close to it. Keep this objective in mind as you work your way through the questions in this assignment.

- ***Variable Definitions:***

  $PRICE_i$ ≡ selling price of house i, in dollars.
  $HSIZE_i$ ≡ living area of house i, in square metres.
  $LSIZE_i$ ≡ area of the lot of house i, in square metres.
  $ROOMS_i$ ≡ number of rooms in house i.
  $BATHS_i$ ≡ number of bathrooms in house i.
  AGE_i ≡ age of house i, in years.
  $Y81_i$ ≡ an indicator variable defined to equal 1 if house i was sold in 1981, and 0 if house i was sold in 1978.
  CBD_i ≡ distance of house i from central business district, in metres.
  $DIST_i$ ≡ distance of house i from incinerator, in metres.

- ***Stata Infile Statement:*** Use the following *Stata* **infile** statement to read the text-format data file **351assn3w09.raw**:

  infile price hsize lsize rooms baths age y81 cbd dist using 351assn3w09.raw

QUESTIONS AND ANSWERS

(48 marks)

1. Compute and present OLS estimates of the following population regression equation for the full sample of 321 houses:

$$\text{PRICE}_i = \beta_0 + \beta_1\text{HSIZE}_i + \beta_2\text{LSIZE}_i + \beta_3\text{HSIZE}_i\text{LSIZE}_i + \beta_4\text{ROOMS}_i + \beta_5\text{BATHS}_i + \beta_6\text{AGE}_i + \beta_7\text{AGE}_i^2 + \beta_8\text{DIST}_i + \beta_9\text{DIST}_i^2 + \beta_{10}\text{CBD}_i + \beta_{11}\text{Y81}_i + u_i \quad \textbf{(1)}$$

(12 marks)

(a) Compute and report OLS estimates of regression equation (3) for the full sample of 321 houses. Present the estimation results in an appropriate table.

<u>ANSWER Question 1(a)</u>

```
. regress price hsize lsize hsizelsize rooms baths age agesq dist distsq cbd
y81

      Source |       SS       df       MS              Number of obs =     321
-------------+------------------------------           F( 11,   309) =   86.42
       Model |  4.5119e+11    11  4.1018e+10           Prob > F      =  0.0000
    Residual |  1.4666e+11   309   474626870           R-squared     =  0.7547
-------------+------------------------------           Adj R-squared =  0.7460
       Total |  5.9785e+11   320  1.8683e+09           Root MSE      =   21786

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       hsize |   93.80356    34.6208     2.71   0.007     25.68122    161.9259
       lsize |  -3.410538   .8284851    -4.12   0.000    -5.040724   -1.780352
  hsizelsize |   .0270578    .004151     6.52   0.000     .0188899    .0352257
       rooms |   3343.379   1839.101     1.82   0.070    -275.3674    6962.125
       baths |   11453.16   2860.263     4.00   0.000     5825.102    17081.22
         age |   -584.915   149.3339    -3.92   0.000     -878.755   -291.0749
       agesq |   2.120924   .9216408     2.30   0.022     .3074383     3.93441
        dist |     8.0328   2.712159     2.96   0.003     2.696163    13.36944
      distsq |  -.0004487   .0002048    -2.19   0.029    -.0008517   -.0000457
         cbd |  -2.833553   1.896985    -1.49   0.136    -6.566196    .8990891
         y81 |   37509.43   2574.096    14.57   0.000     32444.46    42574.41
       _cons |   -4241.35   13134.38    -0.32   0.747    -30085.49    21602.79
------------------------------------------------------------------------------
```

ANSWER Question 1(a) – continued **(total marks = 12)**

Note: The OLS estimates of equation (1) ***must*** be presented in a tabular format similar to the one below in Table 1. **Deduct 4 marks if only some version of the Stata output from the *regress* command is given.**

Table 1: OLS Estimates of Equation (1)

Regressor	$\hat{\beta}_j$	$\hat{se}(\hat{\beta}_j)$	$t(\hat{\beta}_j)$	Lower 95% limit	Upper 95% limit
Constant	**-4241.35**	**13134.38**	**-0.32**	-30085.49	21602.79
$HSIZE_i$	**93.80356**	**34.62080**	**2.71**	25.68122	161.9259
$LSIZE_i$	**-3.410538**	**0.8284851**	**-4.12**	-5.040724	-1.780352
$HSIZE_iLSIZE$	**0.0270578**	**0.004151**	**6.52**	0.0188899	0.0352257
$ROOMS_i$	**3343.379**	**1839.507**	**1.82**	-275.3674	6962.125
$BATHS_i$	**11453.16**	**2860.263**	**4.00**	5825.102	17081.22
AGE_i	**-584.915**	**149.3339**	**-3.92**	-878.755	-291.0749
AGE_i-sq	**2.120924**	**.9216408**	**2.30**	.3074383	3.93441
$DIST_i$	**8.032800**	**2.712159**	**2.96**	2.696163	13.36944
$DIST_i$-sq	**-0.0004487**	**0.0002048**	**-2.19**	-0.0008517	-0.0000457
CBD_i	**-2.833553**	**1.896985**	**-1.49**	-6.566196	0.8990891
$Y81_i$	**37509.43**	**2574.096**	**14.57**	32444.46	42574.41

N = **321**; $\hat{\sigma}^2$ = **474,626,870.0**; R^2 = **0.7547**; F(11, 309) = **86.42 (0.0000)**

(12 marks)

(b) Write the expression implied by equation (1) for the partial marginal effect of $DIST_i$ on house prices. Use the OLS estimation results for equation (1) to test the proposition that distance from the incinerator has no effect on mean house prices for all values of $DIST_i$. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the proposition?

ANSWER Question 1(b) **(total marks = 12)**

- The partial **marginal effect** of $DIST_i$ on house prices is:

$$\frac{\partial PRICE_i}{\partial DIST_i} = \frac{\partial E(PRICE_i | \bullet)}{\partial DIST_i} = \beta_8 + 2\beta_9 DIST_i$$ **(2 marks)**

Null hypothesis: H_0: $\beta_8 = 0$ *and* $\beta_9 = 0$ **(1 mark)**

Alternative hypothesis: H_1: $\beta_8 \neq 0$ *and/or* $\beta_9 \neq 0$ **(1 mark)**

- **Formula for general F-test statistic** is: **(1 mark)**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N - K)}$$

or

$$F = \frac{(R^2_U - R^2_R)/(df_0 - df_1)}{(1 - R^2_U)/df_1} = \frac{(R^2_U - R^2_R)/(K - K_0)}{(1 - R^2_U)/(N - K)}$$

where:

$df_0 = N - K_0$ = degrees-of-freedom for RSS_0 = 321 – 10 = 311;

$df_1 = N - K$ = degrees-of-freedom for RSS_1 = 321 – 12 = **309**;

number of restrictions specified by H_0 = $q = df_0 - df_1 = K - K_0$ = 12 – 10 = **2**

- **Sample value F_0** of the general F-test statistic **and p-value** under H_0 is:

F_0 = **4.3952 = <u>4.40</u>** **(2 marks)**

p-value of F_0 = 0.01312 = <u>0.0131</u> **(1 mark)**

ANSWER Question 1(b): continued **(total marks = 12)**

- **Decision rule and inference: *either formulation 1 or 2*** **(3 marks)**

(1) Decision Rule -- Formulation 1: This a ***two-tail* test**. Compare the sample value F_0 with the **α-level critical value of the F[2, N−12] = F[2, 309] distribution**.

1. If $F_0 \le F_\alpha[2, N-12]$, ***retain* (do not reject) H_0** at the 100α percent significance level.
2. If $F_0 > F_\alpha[2, N-12]$, ***reject* H_0** at the 100α percent significance level.

Inference:
Since F_0 = 4.40 > 3.025 = $F_{0.05}[2, 309]$, ***reject*** H_0 at the ***5 %*** significance level.
Since F_0 = 4.40 < 4.674 = $F_{0.01}[2, 309]$, ***retain*** H_0 at the ***1 %*** significance level.

(2) Decision Rule -- Formulation 2: This a ***two-tail* test**. Compare the **p-value for F_0** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain* (do not reject) H_0** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject* H_0** at significance level α.

Inference:

Since **p-value for F_0 = 0.0131 < 0.05**, ***reject*** H_0 at the ***5 %*** significance level.
Since **p-value for F_0 = 0.0131 > 0.01**, ***retain*** H_0 at the ***1 %*** significance level.

- **Result:** **(1 mark)**
The sample evidence **favours the *alternative* hypothesis H_1** that distance from the incinerator has a non-zero effect on mean house prices **at the 5% significance level**, but **favours the *null* hypothesis H_0 that** distance from the incinerator has no effect on mean house prices **at the 1%** significance level. It thus **provides moderately strong evidence *against* the *null* hypothesis H_0** that distance from the incinerator has no marginal effect on mean house prices.

1. (continued)

(12 marks)

(c) Use the OLS estimation results for equation (1) to test a real estate broker's conjecture that the partial marginal effect on house prices of $DIST_i$, distance from the incinerator, is smaller for homes located further from the incinerator than for homes located close to the incinerator. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the appropriate critical values of the null distribution of the test statistic for both the 5 percent and 1 percent significance levels. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the real estate broker's conjecture?

ANSWER Question 1(c) **(total marks = 12)**

Null hypothesis: H_0: $\beta_9 = 0$ (or $\beta_9 \geq 0$) **(1 mark)**
Alternative hypothesis: H_1: $\beta_9 < 0$ a ***left-tail*** **t-test** **(2 marks)**

- **Calculation of t-test statistic:** The ***sample value*** $\mathbf{t_0}$ under the null hypothesis H_0 is

$$t_0(\hat{\beta}_9) = \frac{\hat{\beta}_9 - b_9}{\hat{se}(\hat{\beta}_9)} = \frac{\hat{\beta}_9 - 0}{\hat{se}(\hat{\beta}_9)} = \frac{\hat{\beta}_9}{\hat{se}(\hat{\beta}_9)} = \frac{-0.000448674}{0.000204801} = \mathbf{-2.1908} = \mathbf{-2.19}$$ **(2 marks)**

- ***Left-tail*** **p-value of** $\mathbf{t_0}$ = Pr(t < –t_0) = Pr(t < –2.191) **= 0.0146072 = 0.0146** **(1 mark)**

- ***Left-tail*** **critical values** of the **t[N–12] = t[309]** distribution are:

$-t_\alpha[N-12] = -t_{0.05}[309]$ = **–1.650** at the **5%** significance level (α = 0.05) **(1 mark)**
$-t_\alpha[N-12] = -t_{0.01}[309]$ = **–2.338** at the **1%** significance level (α = 0.01) **(1 mark)**

- **Decision rule and inference**: ***either formulation 1 or 2*** **(3 marks)**

(1) Decision Rule -- Formulation 1: This a ***left-tail*** **test**. Compare the sample value t_0 with the ***lower*** **α-level critical value of the t[N–12] = t[309] distribution**.

1. If $t_0 < -t_\alpha[N-12]$, ***reject*** $\mathbf{H_0}$ at the 100α percent significance level.
2. If $t_0 \geq -t_\alpha[N-12]$, ***retain*** **(do not reject)** $\mathbf{H_0}$ at the 100α percent significance level.

Inference:

Since $t_0(\hat{\beta}_9)$ = **–2.19 < –1.650 = –$t_{0.05}$[309]**, ***reject*** $\mathbf{H_0}$ at the ***5 %*** significance level.
Since $t_0(\hat{\beta}_9)$ = **–2.19 > –2.338 = –$t_{0.01}$[309]**, ***retain*** $\mathbf{H_0}$ at the ***1 %*** significance level.

(2) Decision Rule -- Formulation 2: This a ***right-tail*** **test**. Compare the ***right-tail*** **p-value for** $\mathbf{t_0}$ with the chosen **significance level** α.

1. If ***left-tail*** **p-value** for $t_0 \geq \alpha$, ***retain*** **(do not reject)** $\mathbf{H_0}$ at significance level α.
2. If ***left-tail*** **p-value** for $t_0 < \alpha$, ***reject*** $\mathbf{H_0}$ at significance level α.

Inference:

Since ***left-tail*** **p-value for** $\mathbf{t_0}$ **= 0.0146 < 0.05**, ***reject*** $\mathbf{H_0}$ at the ***5 %*** significance level.
Since ***left-tail*** **p-value for** $\mathbf{t_0}$ **= 0.0146 > 0.01**, ***retain*** $\mathbf{H_0}$ at the ***1 %*** significance level.

Result: **(1 mark)**
The sample evidence **favours the *alternative* hypothesis** that $\mathbf{\beta_9 < 0}$ **at the 5% significance level, but *not* at the 1% significance level**. It thus provides only moderately strong evidence that the partial marginal effect on house prices of $DIST_i$ is decreasing in $DIST_i$.

(12 marks)

(d) Use the OLS estimation results for equation (1) to test the conjecture that distance from the incinerator has no marginal effect on mean prices for houses located 4,000 metres (4 kilometres) from the incinerator. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the conjecture?

ANSWER Question 1(d): **(total marks = 12)**

- **Null hypothesis:** H_0: $\beta_8 + 2\beta_9 4000 = 0$
  Alternative hypothesis: H_1: $\beta_8 + 2\beta_9 4000 \neq 0$ a ***two-tail*** **test** **(3 marks)**

- **Calculation of t-test statistic:** The ***sample value*** $\mathbf{t_0}$ under the null hypothesis H_0 is

$$t_0(\hat{\beta}_8 + 8000\hat{\beta}_9) = \frac{\hat{\beta}_8 + 8000\hat{\beta}_9}{\hat{se}\left(\hat{\beta}_8 + 8000\hat{\beta}_9\right)} = \frac{4.443406}{1.790786} = \underline{\mathbf{2.4813}} = \underline{\mathbf{2.48}}$$ **(4 marks)**

- ***Two-tail*** **p-value for $\mathbf{t_0}$ = <u>0.013624</u> = <u>0.0136</u>** **(1 mark)**

- **<u>Decision rule and inference</u>:** ***either formulation 1 or 2*** **(3 marks)**

(1) **<u>Decision Rule -- Formulation 1</u>:** This a ***two-tail*** **test**. Compare the sample value t_0 with the **α/2-level critical value of the t[N–12] = t[309] distribution**.

1. If $|t_0| \le t_{\alpha/2}[N-12]$, ***retain*** **(do not reject) $\mathbf{H_0}$** at the 100α percent significance level.
2. If $|t_0| > t_{\alpha/2}[N-12]$, ***reject*** **$\mathbf{H_0}$** at the 100α percent significance level.

<u>Inference</u>:

Since $\left|t_0(\hat{\beta}_8 + 8000\hat{\beta}_9)\right|$ = **2.48 > <u>1.968</u> = $\mathbf{t_{0.025}}$[309]**, ***reject*** **$\mathbf{H_0}$** at ***5 %*** significance level.

Since $\left|t_0(\hat{\beta}_8 + 8000\hat{\beta}_9)\right|$ = **2.48 < <u>2.592</u> = $\mathbf{t_{0.005}}$[309]**, ***retain*** **$\mathbf{H_0}$** at ***1 %*** significance level.

(2) **<u>Decision Rule -- Formulation 2</u>:** This a ***two-tail*** **test**. Compare the ***two-tail*** **p-value for $\mathbf{t_0}$** with the chosen **significance level** α.

1. If ***two-tail*** **p-value** for $t_0 \ge \alpha$, ***retain*** **(do not reject) $\mathbf{H_0}$** at significance level α.
2. If ***two-tail*** **p-value** for $t_0 < \alpha$, ***reject*** **$\mathbf{H_0}$** at significance level α.

<u>Inference</u>:

Since ***two-tail*** **p-value for $\mathbf{t_0}$ = <u>0.0136</u> < 0.05**, ***reject*** **$\mathbf{H_0}$** at the ***5 %*** significance level.
Since ***two-tail*** **p-value for $\mathbf{t_0}$ = <u>0.0136</u> > 0.01**, ***retain*** **$\mathbf{H_0}$** at the ***1 %*** significance level.

- **<u>Result</u>:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis $\mathbf{H_1}$** that $\beta_8 + 2\beta_{11} 4000 \neq 0$ **at the 5% significance level** but ***not*** **at the 1% significance level**; it provides moderately strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses located 4,000 metres (4 kilometres) from the incinerator.

<u>Alternative ANSWER 1 to Question 1(d):</u> two-sided confidence intervals (total marks = 12)

- **Null hypothesis:** H_0: $\beta_8 + 2\beta_9 4000 = 0$
  Alternative hypothesis: H_1: $\beta_8 + 2\beta_9 4000 \neq 0$ a ***two-tail*** **test** **(3 marks)**

- **Two-sided 100(1–α) percent confidence interval for β_8 + 2$\beta_9$4000 = β_8 + 8000β_9** is:

  Lower 100(1–α) percent confidence limit for β_8 + 2$\beta_9$4000
  $= \hat{\beta}_8 + 8000\hat{\beta}_9 - t_{\alpha/2}[N-12]\,\hat{se}\left(\hat{\beta}_8 + 8000\hat{\beta}_9\right)$

  Upper 100(1–α) percent confidence limit for β_8 + 2$\beta_9$4000
  $= \hat{\beta}_8 + 8000\hat{\beta}_9 + t_{\alpha/2}[N-12]\,\hat{se}\left(\hat{\beta}_8 + 8000\hat{\beta}_9\right)$

- Compute **two-sided <u>95 percent</u> confidence interval for β_8 + 2$\beta_9$4000**

  Lower **95 percent confidence limit for β_8 + 2$\beta_9$4000** = **<u>0.9197292</u>** = **<u>0.9197</u>** **(1 mark)**

  Upper **95 percent confidence limit for β_8 + 2$\beta_9$4000** = **<u>7.967083</u>** = **<u>7.967</u>** **(1 mark)**

- Compute **two-sided <u>99 percent</u> confidence interval for β_8 + 2$\beta_9$4000**

  Lower **99 percent confidence limit for β_8 + 2$\beta_9$4000** = **<u>–0.1980122</u>** = **<u>–0.1980</u>** **(1 mark)**

  Upper **99 percent confidence limit for β_8 + 2$\beta_9$4000** = **<u>9.084825</u>** = **<u>9.085</u>** **(1 mark)**

- **<u>Decision rule and inference</u>:** **(2 marks)**

  1. If **hypothesized value of β_8 + 2$\beta_9$4000** lies ***inside*** the **two-sided 100(1–α) percent confidence interval for β_8 + 2$\beta_9$4000**, ***retain*** **(do not reject) H_0** at significance level α.
  2. If **hypothesized value of β_8 + 2$\beta_9$4000** lies ***outside*** the **two-sided 100(1–α) percent confidence interval for β_8 + 2$\beta_9$4000**, ***reject*** **H_0** at significance level α.

  <u>Inference</u>: **(2 marks)**

  Since 0 lies **<u>outside</u>** the **two-sided 95 percent confidence interval for β_8 + 2$\beta_9$4000**, which is [**0.9197, 7.967**], ***reject*** **H_0** at the **5% significance level**.

  Since 0 lies **<u>inside</u>** the **two-sided 99 percent confidence interval for β_8 + 2$\beta_9$4000**, which is [**–0.1980, 9.085**], ***retain*** **H_0** at the **1% significance level**.

- **<u>Result</u>:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis H_1** that $\beta_8 + 2\beta_9 4000 \neq 0$ **at the 5% significance level** but ***not*** **at the 1% significance level**; it provides moderately strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses located 4,000 metres (4 kilometres) from the incinerator.

Alternative ANSWER 2 to Question 1(d): An F-test **(total marks = 12)**

- **Null hypothesis:** H_0: $\beta_8 + 2\beta_9 4000 = 0$
  Alternative hypothesis: H_1: $\beta_8 + 2\beta_9 4000 \neq 0$ a ***two-tail*** **F-test** **(3 marks)**

- **Calculation of F-test statistic:** The ***sample value*** $\mathbf{F_0}$ under the null hypothesis H_0 is

$$F_0(\hat{\beta}_8 + 2\hat{\beta}_9 4000) = \frac{(\hat{\beta}_8 + 8000\hat{\beta}_9)^2}{\hat{Var}\left(\hat{\beta}_8 + 8000\hat{\beta}_9\right)} = \frac{9179.4649}{764.7256} = \underline{\mathbf{6.1567}} = \underline{\mathbf{6.16}}$$ **(4 marks)**

- **(right-tail) p-value for F_0 = 0.013624 = 0.0136** **(1 mark)**

- **Decision rule and inference: *either formulation 1 or 2*** **(3 marks)**

(1) Decision Rule -- Formulation 1: This a ***two-tail*** **test**. Compare the sample value F_0 with the **α-level critical value of the F[1, N–12] = F[1, 309] distribution**.

1. If $F_0 \le F_\alpha[1, N-12]$, ***retain*** **(do not reject) H_0** at the 100α percent significance level.
2. If $F_0 > F_\alpha[1, N-12]$, ***reject*** **H_0** at the 100α percent significance level.

Inference:
Since **F_0 = 6.16 > 3.872 = $F_{0.05}$[1, 309]**, ***reject*** **H_0** at the ***5 %*** significance level.
Since **F_0 = 6.16 < 6.718 = $F_{0.01}$[1, 309]**, ***retain*** **H_0** at the ***1 %*** significance level.

(2) Decision Rule -- Formulation 2: This a ***two-tail*** **test**. Compare the **p-value for F_0** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain*** **(do not reject) H_0** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject*** **H_0** at significance level α.

Inference:

Since **p-value for F_0 = 0.0136 < 0.05**, ***reject*** **H_0** at the ***5 %*** significance level.
Since **p-value for F_0 = 0.0136 > 0.01**, ***retain*** **H_0** at the ***1 %*** significance level.

- **Result:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis H_1** that $\beta_8 + 2\beta_9 4000 \neq 0$ **at the 5% significance level** but ***not*** **at the 1% significance level**; it provides moderately strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses located 4,000 metres (4 kilometres) from the incinerator.

(84 marks)

2. An economics graduate suggests that the partial marginal effect on house prices of distance from the incinerator was different in 1981 that it was in 1978. To account for this possibility, she proposes that the interaction terms $Y81_i DIST_i$ and $Y81_i DIST_i^2$ be added to the set of regressors in equation (1) of question 3. The resulting population regression equation can be written as

$$\begin{aligned} PRICE_i = \beta_0 + \beta_1 HSIZE_i + \beta_2 LSIZE_i + \beta_3 HSIZE_i LSIZE_i + \beta_4 ROOMS_i + \beta_5 BATHS_i \\ + \beta_6 AGE_i + \beta_7 AGE_i^2 + \beta_8 DIST_i + \beta_9 DIST_i^2 + \beta_{10} CBD_i + \beta_{11} Y81_i \\ + \beta_{12} Y81_i DIST_i + \beta_{13} Y81_i DIST_i^2 + u_i \end{aligned} \quad \textbf{(2)}$$

(12 marks)

(a) Compute and report OLS estimates of regression equation (2) for the full sample of 321 houses. Present the estimation results in an appropriate table.

ANSWER Question 2(a)

```
. regress price hsize lsize hsizelsize rooms baths age agesq dist distsq cbd y81
y81dist y81distsq

      Source |       SS       df       MS              Number of obs =     321
-------------+------------------------------           F( 13,   307) =   76.29
       Model |  4.5653e+11    13  3.5118e+10           Prob > F      =  0.0000
    Residual |  1.4132e+11   307   460340273           R-squared     =  0.7636
-------------+------------------------------           Adj R-squared =  0.7536
       Total |  5.9785e+11   320  1.8683e+09           Root MSE      =   21456

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
       hsize |   93.15511    34.0963     2.73   0.007      26.0631    160.2471
       lsize |  -3.270685   .8187273    -3.99   0.000    -4.881712   -1.659658
  hsizelsize |   .0267907   .0040954     6.54   0.000     .0187321    .0348494
       rooms |   3547.095   1812.854     1.96   0.051    -20.09611    7114.287
       baths |    10942.8   2835.747     3.86   0.000     5362.838    16522.76
         age |  -628.9538   148.6912    -4.23   0.000    -921.5366    -336.371
       agesq |    2.41538   .9165659     2.64   0.009     .6118334    4.218926
        dist |   1.073421   3.408945     0.31   0.753    -5.634432    7.781275
      distsq |  -.0000138   .0002479    -0.06   0.956    -.0005017    .0004741
         cbd |  -2.286209   1.875692    -1.22   0.224    -5.977049    1.404631
         y81 |  -9376.649   14060.95    -0.67   0.505    -37044.68    18291.38
     y81dist |   15.20343   4.731664     3.21   0.001     5.892829    24.51402
   y81distsq |  -.0010536   .0003615    -2.91   0.004     -.001765   -.0003423
       _cons |   16675.51   14326.98     1.16   0.245    -11515.99    44867.01
------------------------------------------------------------------------------
```

ANSWER Question 2(a) -- continued

Table 2: OLS Estimates of Equation (2)

Regressor	$\hat{\beta}_j$	$\hat{se}(\hat{\beta}_j)$	$t(\hat{\beta}_j)$	Lower 95% limit	Upper 95% limit
Constant	**-4241.35**	**13134.38**	**-0.32**	-11515.99	44867.01
$HSIZE_i$	**93.1551**	**34.0963**	**2.73**	26.0631	160.2471
$LSIZE_i$	**-3.27069**	**0.818727**	**-3.99**	-4.881712	-1.659658
$HSIZE_i LSIZE$	**0.026907**	**0.0040954**	**6.54**	0.0187321	0.0348494
$ROOMS_i$	**3547.094**	**1812.854**	**1.96**	-20.09611	7114.287
$BATHS_i$	**10942.8**	**2835.757**	**3.86**	5362.838	16522.76
AGE_i	**-628.954**	**148.691**	**-4.23**	-921.5366	-336.371
AGE_i-sq	**2.41538**	**0.916566**	**2.64**	0.6118334	4.218926
$DIST_i$	**1.07342**	**3.40895**	**0.31**	-5.634432	7.781275
$DIST_i$-sq	**-0.0000138**	**0.0002479**	**-0.06**	-0.0005017	0.0004741
CBD_i	**-2.28621**	**1.87569**	**-1.22**	-5.977049	1.404631
$Y81_i$	**-9376.65**	**14060.95**	**-0.67**	-37044.68	18291.38
$Y81_i DIST_i$	**15.2034**	**4.73166**	**3.21**	5.892829	24.51402
$Y81_i DIST_i$-sq	**-0.0010536**	**0.0003615**	**-2.91**	-0.001765	-0.0003423

N = **321**; $\hat{\sigma}^2$ = **460,340,273.0**; R^2 = **0.7636**; F(13, 307) = **76.29 (0.0000)**

2. (continued)

(12 marks)

(b) Write the expression implied by regression equation (2) for the partial marginal effect of $DIST_i$ on the prices of houses sold in 1978. Use the OLS estimation results for equation (2) to test the proposition that distance from the incinerator had no effect on mean prices for all houses sold in 1978. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the proposition that the partial marginal effect of distance from the incinerator had no effect on mean prices for all houses sold in 1978?

ANSWER Question 2(b)

- The partial **marginal effect** of $DIST_i$ on house prices in 1978 is:

$$\frac{\partial PRICE_i}{\partial DIST_i} = \frac{\partial E(PRICE_i \mid \bullet)}{\partial DIST_i} = \beta_8 + 2\beta_9 DIST_i$$ **(2 marks)**

Null hypothesis: H_0: $\beta_8 = 0$ *and* $\beta_9 = 0$ **(1 mark)**

Alternative hypothesis: H_1: $\beta_8 \neq 0$ *and/or* $\beta_9 \neq 0$ **(1 mark)**

- **Formula for general F-test statistic** is: **(1 mark)**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N - K)}$$

or

$$F = \frac{(R_U^2 - R_R^2)/(df_0 - df_1)}{(1 - R_U^2)/df_1} = \frac{(R_U^2 - R_R^2)/(K - K_0)}{(1 - R_U^2)/(N - K)}$$

where:

$df_0 = N - K_0$ = degrees-of-freedom for RSS_0 = 321 – 12 = 309;

$df_1 = N - K$ = degrees-of-freedom for RSS_1 = 321 – 14 = **307**;

number of restrictions specified by H_0 = $q = df_0 - df_1 = K - K_0$ = 14 – 12 = **2**

- **Sample value F_0** of the general F-test statistic **and p-value** under H_0 is:

F_0 = **0.12554 = <u>0.13</u>** **(2 marks)**

p-value of F_0 = <u>0.8821</u> **(1 mark)**

ANSWER Question 1(b): continued **(total marks = 12)**

- **Decision rule and inference**: ***either formulation 1 or 2*** **(3 marks)**

(1) **Decision Rule -- Formulation 1:** This a ***two-tail*** **test**. Compare the sample value F_0 with the **α-level critical value of the F[2, N–14] = F[2, 307] distribution**.

1. If $F_0 \le F_\alpha[2, N-14]$, ***retain*** **(do not reject) H_0** at the 100α percent significance level.
2. If $F_0 > F_\alpha[2, N-14]$, ***reject*** **H_0** at the 100α percent significance level.

Inference:
Since **F_0 = 0.13 < 3.025 = $F_{0.05}$[2, 307]**, ***retain*** **H_0** at the ***5 %*** significance level.
Since **F_0 = 0.13 < 2.320 = $F_{0.10}$[2, 307]**, ***retain*** **H_0** at the ***10 %*** significance level.

(2) **Decision Rule -- Formulation 2:** This a ***two-tail*** **test**. Compare the **p-value for F_0** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain*** **(do not reject) H_0** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject*** **H_0** at significance level α.

Inference:

Since **p-value for F_0 = 0.8821 > 0.05**, ***retain*** **H_0** at the ***5 %*** significance level.
Since **p-value for F_0 = 0.8821 > 0.10**, ***retain*** **H_0** at the ***10 %*** significance level.

- **Result:** **(1 mark)**
The sample evidence **favours the *null* hypothesis H_0** that distance from the incinerator had no effect on mean prices for all houses sold in 1978 at both **the 5% and 10% significance levels**. It thus **provides no evidence** that the partial marginal effect on 1978 house prices of distance from the incinerator was non-zero.

2. (continued)

(12 marks)

(c) Write the expression implied by regression equation (2) for the partial marginal effect of $DIST_i$ on the prices of houses sold in 1981. Use the OLS estimation results for equation (2) to test the proposition that distance from the incinerator had no effect on mean prices for all houses sold in 1981. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the proposition that the partial marginal effect of distance from the incinerator had no effect on the mean prices of houses sold in 1981?

ANSWER Question 2(c)

- The partial **marginal effect** of $DIST_i$ on house prices in 1981 is:

$$\frac{\partial PRICE_i}{\partial DIST_i} = \frac{\partial E(PRICE_i | \bullet)}{\partial DIST_i} = \beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})DIST_i$$ **(2 marks)**

Null hypothesis: H_0: $\beta_8 + \beta_{12} = 0$ *and* $\beta_9 + \beta_{13} = 0$ **(1 mark)**

Alternative hypothesis: H_1: $\beta_8 + \beta_{12} \neq 0$ *and/or* $\beta_9 + \beta_{13} \neq 0$ **(1 mark)**

- **Formula for general F-test statistic** is: **(1 mark)**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N - K)}$$

or

$$F = \frac{(R_U^2 - R_R^2)/(df_0 - df_1)}{(1 - R_U^2)/df_1} = \frac{(R_U^2 - R_R^2)/(K - K_0)}{(1 - R_U^2)/(N - K)}$$

where:

$df_0 = N - K_0$ = degrees-of-freedom for RSS_0 = 321 – 12 = 309;

$df_1 = N - K$ = degrees-of-freedom for RSS_1 = 321 – 14 = **307**;

number of restrictions specified by H_0 = $q = df_0 - df_1 = K - K_0$ = 14 – 12 = **2**

- **Sample value F_0** of the general F-test statistic **and p-value** under H_0 is:

F_0 = **9.5600 = <u>9.56</u>** **(2 marks)**

p-value of F_0 = 0.0000942 = <u>0.0001</u> **(1 marks**

<u>ANSWER Question 2(c): continued</u> **(total marks = 12)**

- **<u>Decision rule and inference</u>:** ***either formulation 1 or 2*** **(3 marks)**

(1) **<u>Decision Rule -- Formulation 1</u>:** This a ***two-tail*** **test**. Compare the sample value F_0 with the **α-level critical value of the F[2, N–14] = F[2, 307] distribution**.

1. If $F_0 \le F_{\alpha}[2, N-14]$, ***retain*** **(do not reject) H_0** at the 100α percent significance level.
2. If $F_0 > F_{\alpha}[2, N-14]$, ***reject*** **$H_0$** at the 100α percent significance level.

<u>Inference</u>:
Since **F_0 = <u>9.56</u>** > **<u>4.675</u> = $F_{0.01}$[2, 307]**, ***reject*** **H_0** at the ***1 %*** significance level.
Since **F_0 = <u>9.56</u>** > **<u>3.025</u> = $F_{0.05}$[2, 307]**, ***reject*** **H_0** at the ***5 %*** significance level.

(2) **<u>Decision Rule -- Formulation 2</u>:** This a ***two-tail*** **test**. Compare the **p-value for F_0** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain*** **(do not reject) H_0** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject*** **H_0** at significance level α.

<u>Inference</u>:

Since **p-value for F_0 = 0.0001 < 0.01**, ***reject*** **H_0** at the ***1 %*** significance level.
Since **p-value for F_0 = 0.0001 < 0.05**, ***reject*** **H_0** at the ***5 %*** significance level.

- **<u>Result</u>:** **(1 mark)**
  The sample evidence **strongly favours the *alternative* hypothesis H_1** that distance from the incinerator had a non-zero effect on mean house prices in 1981 at both **the 5% and 10% significance levels**. It thus **provides strong evidence** that the partial marginal effect on house prices in 1981 of distance from the incinerator was non-zero.

2. (continued)

(12 marks)

(d) Use the OLS estimation results for equation (2) to test the economics graduate's conjecture that the partial marginal effect on house prices of distance from the incinerator was different in 1981 than it was in 1978. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the conjecture?

<u>ANSWER Question 2(d)</u>

- The 1981-1978 difference in the partial **marginal effect** of $DIST_i$ on house prices is:

$$\frac{\partial E\left(PRICE_i \mid Y81_i = 1\right)}{\partial DIST_i} - \frac{\partial E\left(PRICE_i \mid Y81_i = 0\right)}{\partial DIST_i} = \beta_{12} + 2\beta_{13}DIST_i$$ **(2 marks)**

Null hypothesis: H_0: $\beta_{12} = 0$ *and* $\beta_{13} = 0$ **(1 mark)**

Alternative hypothesis: H_1: $\beta_{12} \neq 0$ *and/or* $\beta_{13} \neq 0$ **(1 mark)**

- **Formula for general F-test statistic** is: **(1 mark)**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} = \frac{(RSS_0 - RSS_1)/(K - K_0)}{RSS_1/(N-K)}$$

<u>or</u>

$$F = \frac{(R_U^2 - R_R^2)/(df_0 - df_1)}{(1 - R_U^2)/df_1} = \frac{(R_U^2 - R_R^2)/(K - K_0)}{(1 - R_U^2)/(N-K)}$$

where:

$df_0 = N - K_0$ = degrees-of-freedom for RSS_0 = 321 – 12 = 309;

$df_1 = N - K$ = degrees-of-freedom for RSS_1 = 321 – 14 = **307**;

number of restrictions specified by H_0 = $q = df_0 - df_1 = K - K_0$ = 14 – 12 = **2**

- **Sample value F_0** of the general F-test statistic **and p-value** under H_0 is:

F_0 = **5.7949 = <u>5.79</u>** **(2 marks)**

p-value of F_0 = 0.003386 = <u>0.0034</u> **(1 mark)**

ANSWER Question 2(d): continued **(total marks = 12)**

- **Decision rule and inference**: ***either formulation 1 or 2*** **(3 marks)**

(1) **Decision Rule -- Formulation 1:** This a ***two-tail*** **test**. Compare the sample value F_0 with the **α-level critical value of the F[2, N−14] = F[2, 307] distribution**.

1. If $F_0 \le F_\alpha[2, N-14]$, ***retain*** **(do not reject) H_0** at the 100α percent significance level.
2. If $F_0 > F_\alpha[2, N-14]$, ***reject*** **H_0** at the 100α percent significance level.

Inference:
Since F_0 = **5.79** > **4.675** = $F_{0.01}[2, 307]$, ***reject*** H_0 at the ***1 %*** significance level.
Since F_0 = **5.79** > **3.025** = $F_{0.05}[2, 307]$, ***reject*** H_0 at the ***5 %*** significance level.

(2) **Decision Rule -- Formulation 2:** This a ***two-tail*** **test**. Compare the **p-value for F_0** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain*** **(do not reject) H_0** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject*** **H_0** at significance level α.

Inference:

Since **p-value for F_0 = 0.0034 < 0.01**, ***reject*** H_0 at the ***1 %*** significance level.
Since **p-value for F_0 = 0.0034 < 0.05**, ***reject*** H_0 at the ***5 %*** significance level.

- **Result:** **(1 mark)**
The sample evidence **strongly supports the conjecture (i.e., the *alternative* hypothesis H_1)** that distance from the incinerator had a different marginal effect on mean house prices in 1981 than in 1978.

2. (continued)

(10 marks)

(e) Use the OLS estimation results for equation (2) to test the proposition that the partial marginal effect of $DIST_i$ on house prices was positive for houses sold in 1981 that were located 6,000 metres from the incinerator, i.e., for houses sold in 1981 for which DIST = 6000. State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test. Does the sample evidence favour the conjecture?

<u>ANSWER Question 2(e)</u>

- The partial **marginal effect** of $DIST_i$ on house prices in 1981 is:

$$\frac{\partial PRICE_i}{\partial DIST_i} = \frac{\partial E(PRICE_i \mid \bullet)}{\partial DIST_i} = \beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})DIST_i$$

Null hypothesis: H_0: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})6000 = 0$ **(1 mark)**

Alternative hypothesis: H_1: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})6000 > 0$ **(1 mark)**

- **Calculation of t-test statistic:** The ***sample value*** $\mathbf{t_0}$ under the null hypothesis H_0 is

$$t_0\left(\hat\beta_8 + \hat\beta_{12} + 12000(\hat\beta_9 + \hat\beta_{13})\right) = \frac{\hat\beta_8 + \hat\beta_{12} + 12000(\hat\beta_9 + \hat\beta_{13})}{\hat{se}\left(\hat\beta_8 + \hat\beta_{12} + 12000(\hat\beta_9 + \hat\beta_{13})\right)} = \frac{3.467287}{1.815021} = \underline{\mathbf{1.91033}} = \underline{\mathbf{1.91}}$$

(3 marks)

- ***Right-tail*** **p-value of $\mathbf{t_0}$** = Pr(t > t_0) = Pr(t > 1.91033) **= <u>0.02851</u>** **(1 mark)**

- ***Right-tail*** **critical values** of the t[N−14] = t[307] distribution are:

$t_\alpha[N-14] = t_{0.05}[307]$ = **<u>1.6498</u>** = **<u>1.65</u>** at the **5%** significance level (α = 0.05) **(1 mark)**

$t_\alpha[N-14] = t_{0.01}[307]$ = **<u>2.3386</u>** = **<u>2.34</u>** at the **1%** significance level (α = 0.01)

ANSWER Question 2(e) -- continued **(total marks = 10)**

- **Decision rule and inference**: ***either formulation 1 or 2*** **(3 marks)**

(1) Decision Rule -- Formulation 1: This a ***right-tail*** **test**. Compare the sample value t_0 with the ***upper*** **α-level critical value of the t[N–14] = t[307] distribution**.

1. If $t_0 \le t_\alpha[N-14]$, ***retain*** **(do not reject) H_0** at the 100α percent significance level.
2. If $t_0 > t_\alpha[N-14]$, ***reject*** **H_0** at the 100α percent significance level.

Inference: Since t_0 = **1.91 > 1.65 = $t_{0.05}$[307]**, ***reject*** **H_0** at the ***5 %*** significance level.
Since t_0 = **1.91 < 2.34 = $t_{0.01}$[307]**, ***retain*** **H_0** at the ***1 %*** significance level.

OR

(2) Decision Rule -- Formulation 2: This a ***right-tail*** **test**. Compare the ***right-tail*** **p-value for t_0** with the chosen **significance level α**.

1. If ***right-tail*** **p-value** for $t_0 \ge \alpha$, ***retain*** **(do not reject) H_0** at significance level α.
2. If ***right-tail*** **p-value** for $t_0 < \alpha$, ***reject*** **H_0** at significance level α.

Inference:

Since ***right-tail*** **p-value for t_0 = 0.02851 < 0.05**, ***reject*** **H_0** at the ***5 %*** significance level.
Since ***right-tail*** **p-value for t_0 = 0.02851 > 0.01**, ***retain*** **H_0** at the ***1 %*** significance level.

The sample evidence **favours the *alternative* hypothesis H_1 at the 5% significance level**, **but *not* at the 1% significance level**. It provides moderately strong evidence supporting the conjecture that the partial marginal effect of $DIST_i$ on house prices was positive for houses sold in 1981 that were located 6,000 metres from the incinerator.

2. (continued)

(14 marks)

(f) Use the OLS estimation results for equation (2) to compute an estimate of the partial marginal effect of $DIST_i$ on house prices for houses sold in 1981 that were located 2,000 metres from the incinerator, i.e., for houses sold in 1981 for which DIST = 2000. Use the estimation results for equation (2) to test the proposition that **the partial marginal effect of** $DIST_i$ **on house prices was zero for houses sold in 1981 that were located 2,000 metres from the incinerator** (for which DIST = 2000). State the null and alternative hypotheses, and show how the sample value of the test statistic is calculated (give its formula). Report the sample value of the test statistic and its p-value. State the decision rule you use, and the inference you would draw from the test.

ANSWER Question 2(f) -- continued **(total marks = 14)**

- The partial **marginal effect** of $DIST_i$ on house prices in 1981 is:

$$\frac{\partial PRICE_i}{\partial DIST_i} = \frac{\partial E(PRICE_i \mid \bullet)}{\partial DIST_i} = \beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})DIST_i$$ **(2 marks)**

- Estimate of the partial **marginal effect** of $DIST_i$ on house prices in 1981 for $DIST_i$ = 2000:

$\hat{\beta}_8 + \hat{\beta}_{12} + 2(\hat{\beta}_9 + \hat{\beta}_{13})2000$ = 12.00699 = **$12.01 dollars per metre**
= **12.01 dollars per metre** **(2 marks)**

Null hypothesis: H_0: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 = 0$ **(1 mark)**
Alternative hypothesis: H_1: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **(1 mark)**

- **Calculation of t-test statistic:** The ***sample value*** $\mathbf{t_0}$ under the null hypothesis H_0 is

$$t_0\left(\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})\right) = \frac{\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})}{\widehat{se}\left(\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})\right)} = \frac{12.00699}{2.760975} = \mathbf{4.3488} = \mathbf{4.35}$$

(3 marks)

- ***Two-tail*** **p-value for** $\mathbf{t_0}$ **= 0.00001865 = 0.0000** **(1 mark)**

- **Decision rule and inference**: ***either formulation 1 or 2*** **(3 marks)**

(1) Decision Rule -- Formulation 1: This a ***two-tail*** **test**. Compare the sample value t_0 with the **α/2-level critical value of the t[N–14] = t[307] distribution**.

1. If $|t_0| \le t_{\alpha/2}[N-14]$, ***retain*** **(do not reject)** $\mathbf{H_0}$ at the 100α percent significance level.
2. If $|t_0| > t_{\alpha/2}[N-14]$, ***reject*** $\mathbf{H_0}$ at the 100α percent significance level.

Inference:
Since $|t_0|$ = **4.35 > 1.968 =** $\mathbf{t_{0.025}}$**[307]**, ***reject*** $\mathbf{H_0}$ at ***5 %*** significance level.
Since $|t_0|$ = **4.35 < 2.592 =** $\mathbf{t_{0.005}}$**[307]**, ***reject*** $\mathbf{H_0}$ at ***1 %*** significance level.

(2) Decision Rule -- Formulation 2: This a ***two-tail*** **test**. Compare the ***two-tail*** **p-value for** $\mathbf{t_0}$ with the chosen **significance level** α.

1. If ***two-tail*** **p-value** for $t_0 \ge \alpha$, ***retain*** **(do not reject)** $\mathbf{H_0}$ at significance level α.
2. If ***two-tail*** **p-value** for $t_0 < \alpha$, ***reject*** $\mathbf{H_0}$ at significance level α.

Inference:

Since ***two-tail*** **p-value for** $\mathbf{t_0}$ **= 0.0000 < 0.05**, ***reject*** $\mathbf{H_0}$ at the ***5 %*** significance level.
Since ***two-tail*** **p-value for** $\mathbf{t_0}$ **= 0.0000 < 0.01**, ***reject*** $\mathbf{H_0}$ at the ***1 %*** significance level.

ANSWER Question 2(f) -- continued **(total marks = 14)**

- **Result:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis H_1** that $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **at both the 5% and 1% significance levels**; it provides strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses sold in 1981 that were located 2,000 metres (2 kilometres) from the incinerator.

<u>Alternative ANSWER 1 to Question 2(f):</u> An F-test **(total marks = 14)**

- The partial **marginal effect** of DIST_i on house prices in 1981 is:

$$\frac{\partial \text{PRICE}_i}{\partial \text{DIST}_i} = \frac{\partial E(\text{PRICE}_i \mid \bullet)}{\partial \text{DIST}_i} = \beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})\text{DIST}_i$$

(2 marks)

- Estimate of the partial **marginal effect** of DIST_i on house prices in 1981 for $\text{DIST}_i = 2000$:

$\hat{\beta}_8 + \hat{\beta}_{12} + 2(\hat{\beta}_9 + \hat{\beta}_{13})2000 = 12.00699$ = **<u>$12.01 dollars per metre</u>**
= **<u>12.01 dollars per metre</u>** **(2 marks)**

- **Null hypothesis:** H_0: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 = 0$ **(1 mark)**
  Alternative hypothesis: H_1: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **(1 mark)**

- **Calculation of F-test statistic:** The ***sample value*** $\mathbf{F_0}$ under the null hypothesis H_0 is

$$F_0\left(\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})\right) = \frac{\left(\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})\right)^2}{\widehat{\text{Var}}\left(\hat{\beta}_8 + \hat{\beta}_{12} + 4000(\hat{\beta}_9 + \hat{\beta}_{13})\right)} = \frac{144.16781}{7.622983} = \underline{\mathbf{18.9123}} = \underline{\mathbf{18.91}}$$

(3 marks)

- **(right-tail) p-value for $\mathbf{F_0}$ = <u>0.00001865</u> = <u>0.0000</u>** **(1 mark)**

- **<u>Decision rule and inference</u>:** ***either formulation 1 or 2*** **(3 marks)**

(1) **<u>Decision Rule -- Formulation 1</u>:** This a ***two-tail* test**. Compare the sample value F_0 with the **α-level critical value of the F[1, N–14] = F[1, 307] distribution**.

1. If $F_0 \le F_\alpha[1, N-14]$, ***retain*** **(do not reject) $\mathbf{H_0}$** at the 100α percent significance level.
2. If $F_0 > F_\alpha[1, N-14]$, ***reject*** $\mathbf{H_0}$ at the 100α percent significance level.

<u>Inference</u>:
Since $\mathbf{F_0}$ **= 18.91 > <u>3.872</u> = $\mathbf{F_{0.05}}$[1, 307]**, ***reject*** $\mathbf{H_0}$ at the ***5 %*** significance level.
Since $\mathbf{F_0}$ **= 18.91 > <u>6.718</u> = $\mathbf{F_{0.01}}$[1, 307]**, ***reject*** $\mathbf{H_0}$ at the ***1 %*** significance level.

(2) **<u>Decision Rule -- Formulation 2</u>:** This a ***two-tail* test**. Compare the **p-value for $\mathbf{F_0}$** with the chosen **significance level α**.

1. If **p-value** for $F_0 \ge \alpha$, ***retain*** **(do not reject) $\mathbf{H_0}$** at significance level α.
2. If **p-value** for $F_0 < \alpha$, ***reject*** $\mathbf{H_0}$ at significance level α.

Alternative ANSWER 1 to Question 2(f): An F-test **(total marks = 14)**

Inference:

Since **p-value for F_0 = 0.0000 < 0.05**, ***reject*** H_0 at the ***5 %*** significance level.
Since **p-value for F_0 = 0.0000 < 0.01**, ***reject*** H_0 at the ***1 %*** significance level.

- **Result:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis H_1** that $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **at *both* the 5% and 1% significance levels**; it provides strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses sold in 1981 that were located 2,000 metres (2 kilometres) from the incinerator.

```
. * Question 2(f):
. *
. lincom _b[dist] + _b[y81dist] + 2*(_b[distsq] + _b[y81distsq])*2000

 ( 1)  dist + 4000 distsq + y81dist + 4000 y81distsq = 0

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   12.00699   2.760975     4.35   0.000     6.574164    17.43982
------------------------------------------------------------------------------

. lincom _b[dist] + _b[y81dist] + 2*(_b[distsq] + _b[y81distsq])*2000, level(99)

 ( 1)  dist + 4000 distsq + y81dist + 4000 y81distsq = 0

------------------------------------------------------------------------------
       price |      Coef.   Std. Err.      t    P>|t|     [99% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   12.00699   2.760975     4.35   0.000     4.850716    19.16327
------------------------------------------------------------------------------
```

<u>Alternative ANSWER 2 to Question 2(f):</u> two-sided confidence intervals (total marks = 14)

- Estimate of the partial **marginal effect** of DIST_i on house prices in 1981 for $\text{DIST}_i = 2000$:

  $\hat\beta_8 + \hat\beta_{12} + 2(\hat\beta_9 + \hat\beta_{13})2000 = 12.00699$ = **<u>$12.01 dollars per metre</u>**
  = **<u>12.01 dollars per metre</u>** **(3 marks)**

- **Null hypothesis:** H_0: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 = 0$ **(1 mark)**
  Alternative hypothesis: H_1: $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **(1 mark)**

- **Two-sided 100(1–α) percent confidence interval for β₈ + β₁₂ + 2(β₉ + β₁₃)2000 = β₈ + β₁₂ + 4000(β₉ + β₁₃)** is:

  Lower 100(1–α) percent confidence limit for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$
  $= \hat\beta_8 + \hat\beta_{12} + 4000(\hat\beta_9 + \hat\beta_{13}) - t_{\alpha/2}[N-14]\,\hat{se}\left(\hat\beta_8 + \hat\beta_{12} + 4000(\hat\beta_9 + \hat\beta_{13})\right)$

  Upper 100(1–α) percent confidence limit for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$
  $= \hat\beta_8 + \hat\beta_{12} + 4000(\hat\beta_9 + \hat\beta_{13}) + t_{\alpha/2}[N-14]\,\hat{se}\left(\hat\beta_8 + \hat\beta_{12} + 4000(\hat\beta_9 + \hat\beta_{13})\right)$

- Compute **two-sided <u>95 percent</u> confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**

  ***Lower* 95 percent confidence limit** = **6.574164** = **<u>6.5742</u>** **(1 mark)**

  ***Upper* 95 percent confidence limit** = **17.43982** = **<u>17.440</u>** **(1 mark)**

- Compute **two-sided <u>99 percent</u> confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**

  ***Lower* 99 percent confidence limit** = **4.850716** = **<u>4. 8507</u>** **(1 mark)**

  ***Upper* 99 percent confidence limit** = **19.16327** = **<u>19.163</u>** **(1 mark)**

- **<u>Decision rule and inference</u>:** **(2 marks)**

  1. If **hypothesized value of $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$** lies ***inside*** the **two-sided 100(1–α) percent confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**, ***retain*** **(do not reject) H_0 at significance level α**.
  2. If **hypothesized value of $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$** lies ***outside*** the **two-sided 100(1–α) percent confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**, ***reject*** **$H_0$ at significance level α**.

  <u>Inference</u>: **(2 marks)**

  Since 0 lies **<u>outside</u>** the **two-sided 95 percent confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**, which is [**6.5742, 17.440**], ***reject*** **H_0** at the **5% significance level**.

  Since 0 lies **<u>outside</u>** the **two-sided 99 percent confidence interval for $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000$**, which is [**4.8507, 19.163**], ***reject*** **H_0** at the **1% significance level**.

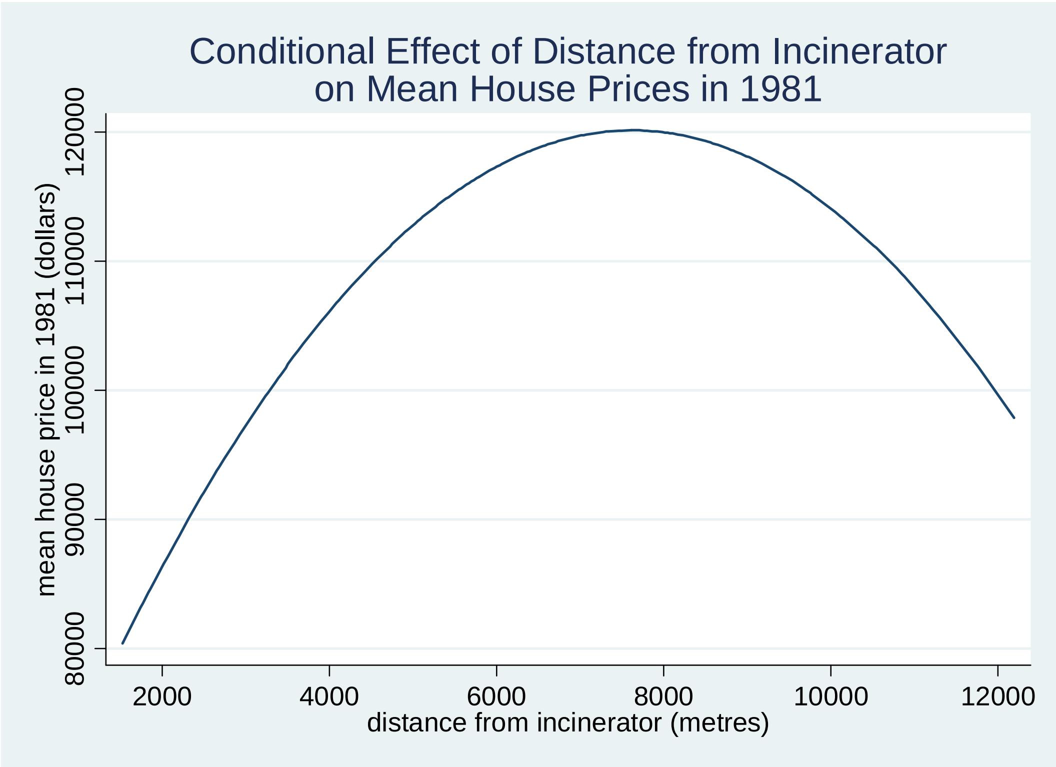
Alternative ANSWER 2 to Question 2(f) -- continued **(total marks = 14)**

- **Result:** **(1 mark)**
  The sample evidence **favours the *alternative* hypothesis H_1** that $\beta_8 + \beta_{12} + 2(\beta_9 + \beta_{13})2000 \neq 0$ **at *both* the 5% and 1% significance levels**; it provides strong evidence that distance from the incinerator has a non-zero marginal effect on mean prices for houses sold in 1981 that were located 2,000 metres (2 kilometres) from the incinerator.

2. (continued)

(12 marks)

(g) Use the *Stata* **graph** command to draw a line graph of the conditional relationship of estimated mean house price to $DIST_i$, distance from the incinerator, for houses sold in 1981 (for which Y81 = 1) that had the following observed characteristics: HSIZE = 196 square metres; LSIZE = 3700 square metres; ROOMS = 7; BATHS = 2; AGE = 18 years; and CBD = 4,800 metres. Note that these observed characteristics correspond to the approximate sample mean values of the explanatory variables $HSIZE_i$, $LSIZE_i$, $ROOMS_i$, $BATHS_i$, AGE_i and CBD_i in the estimation sample.



***Stata* commands used to create the above line graph:**

```
#delimit ;
generate pricehat = _b[_cons] + _b[hsize]*196 + _b[lsize]*3700
+ _b[hsizelsize]*196*3700 + _b[rooms]*7 + _b[baths]*2 + _b[age]*18 + _b[agesq]*18*18
+ (_b[dist] + _b[y81dist] )*dist + (_b[distsq] + _b[y81distsq] )*dist *dist
+ _b[cbd]*4800 + _b[y81] ;
#delimit cr
sort dist
summarize pricehat dist
list dist pricehat in 1/10
#delimit ;
graph twoway line pricehat dist,
ytitle("mean house price in 1981 (dollars)")
xtitle("distance from incinerator (metres)")
title("Conditional Effect of Distance from Incinerator" "on Mean House Prices in
1981") ;
#delimit cr
```