

---

## *Stata 10 Tutorial 8*

### **TOPIC: Using Dummy Variable Regressors to Test for Coefficient Differences in Multiple Linear Regression Models**

**DATA:** `auto1.dta` (a *Stata*-format data file you created in *Stata Tutorial 1*)

**TASKS:** *Stata Tutorial 8* is intended to introduce you to the use of *dummy variable regressors* to test for regression coefficient differences in multiple linear regression models.

- The *Stata commands* that constitute the primary subject of this tutorial are:
  - regress** Used to perform OLS estimation of multiple linear regression models.
  - test** Used to compute F-tests of linear coefficient equality restrictions after OLS estimation.
  - lincom** Used after estimation to compute linear combinations of coefficient estimates and associated statistics.
  
- The *Stata statistical functions* introduced in this tutorial are:
  - ttail(df, t<sub>0</sub>)** Computes *p-values* of *calculated t-statistics*.
  - invttail(df, p)** Computes *right-tail critical values* of *t-distributions*.
  - Ftail(df<sub>1</sub>, df<sub>2</sub>, F<sub>0</sub>)** Computes *p-values* of *calculated F-statistics*.
  - invFtail(df<sub>1</sub>, df<sub>2</sub>, α)** Computes *right-tail critical values* of *F-distributions*.

**IMPORTANT:** If you are to derive maximum benefit from this tutorial, it is not sufficient that you merely enter the *Stata* commands. You should carefully go over the *log* file for *Stata Tutorial 8* to learn about the use of dummy variable regressors in linear regression models.

**NOTE:** *Stata* commands are *case sensitive*. All *Stata* command names must be typed in the Command window in **lower case letters**.

---

## □ Preparing for Your *Stata* Session

---

Before beginning your *Stata* session, use Windows Explorer to copy the *Stata*-format dataset **auto1.dta** to the *Stata working directory* on the C:-drive or D:-drive of the computer at which you are working.

- **On the computers in Dunning 350**, the default *Stata* working directory is usually **C:\data**.
- **On the computers in MC B111**, the default *Stata* working directory is usually **D:\courses**.

---

## □ Start Your *Stata* Session

---

**To start your *Stata* session**, double-click on the ***Stata 10* icon** in the Windows desktop.

After you double-click the ***Stata 10* icon**, you will see the now familiar screen of four *Stata* windows.

---

## □ Record Your *Stata* Session -- log using

---

**To record your *Stata* session**, including all the *Stata* commands you enter and the results (output) produced by these commands, make a **.log** file named **351tutorial8.log**. To open (begin) the log file **351tutorial8.log**, enter in the Command window:

```
log using 351tutorial8.log
```

This command opens a text-format file called **351tutorial8.log** in the current *Stata* working directory. Remember that once you have opened the **351tutorial8.log** file, a copy of all the commands you enter during your *Stata* session and of all the results they produce is recorded in that **351tutorial8.log** file.

An alternative way to open the **.log** file **351tutorial8.log** is to click on the **Log** button; click on **Save as type:** and select **Log (\*.log)**; click on the **File name:** box and type the file name **351tutorial8**; and click on the **Save** button.

**To record only the *Stata* commands you type during your *Stata* session**, use the *Stata* `cmdlog using` command. To start (open) the command log file `351tutorial8.txt`, enter in the Command window:

```
cmdlog using 351tutorial8
```

This command opens a plain text-format (ASCII) file called `351tutorial8.txt` in the current *Stata* working directory. All commands you enter during your *Stata* session are recorded in this file.

### **❑ Loading a *Stata*-Format Data Set into *Stata* – use**

---

**Load, or read, into memory the data set you are using.** To load the *Stata*-format data file `auto1.dta` into memory, enter in the Command window:

```
use auto1
```

This command loads into memory the *Stata*-format data set `auto1.dta`.

### **❑ Familiarize Yourself with the Current Data Set**

---

**To familiarize (or re-familiarize) yourself with the contents of the current data set**, type in the Command window the following commands:

```
describe  
summarize
```

### **❑ The Nature of Dummy or Indicator Variables**

---

**Nature:** *Dummy or indicator variables* are *binary variables* that take only two values, 0 and 1. An indicator or dummy variable is used to distinguish between two categories of observations, namely those observations that exhibit some observed characteristic or property, and those that do not.

***Examples:***

- ◆ The data set **auto1.dta** (or **auto1.raw**) contains the *indicator (or dummy) variable* **foreign<sub>i</sub>** defined as follows:

$$\begin{aligned} \text{foreign}_i &= 1 && \text{for all foreign cars} \\ &= 0 && \text{for all domestic cars.} \end{aligned}$$

The variable **foreign<sub>i</sub>** is an *indicator variable* for foreign-made cars. It distinguishes between two categories of cars sold in North America: (1) **foreign cars**, for which the indicator variable **foreign<sub>i</sub> = 1** by definition, and (2) **domestic cars**, for which the indicator variable **foreign<sub>i</sub> = 0** by definition.

- Enter the following command to obtain details on the properties of the indicator variable **foreign<sub>i</sub>**:

```
codebook foreign
```

- ◆ An alternative but equivalent way to distinguish between foreign and domestic cars is to define an indicator variable for domestic cars. Thus, define the *indicator (or dummy) variable* **domestic<sub>i</sub>** as follows:

$$\begin{aligned} \text{domestic}_i &= 1 && \text{for all domestic cars} \\ &= 0 && \text{for all foreign cars.} \end{aligned}$$

- ◆ Note that the two indicator variables **foreign<sub>i</sub>** and **domestic<sub>i</sub>** satisfy the following *adding-up property* for all observations:

$$\text{domestic}_i + \text{foreign}_i = 1 \quad \text{for all } i = 1, \dots, N.$$

This adding-up property implies that

$$\begin{aligned} \text{domestic}_i &= 1 - \text{foreign}_i && (i = 1, \dots, N); \\ \text{foreign}_i &= 1 - \text{domestic}_i && (i = 1, \dots, N). \end{aligned}$$

- Create the new indicator variable **domestic<sub>i</sub>**. Enter the command:

```
generate domestic = 1 - foreign
```

- Label the new indicator variable **domestic<sub>i</sub>**. Enter the commands:

```
label variable domestic "Domestic car indicator"
codebook domestic
```

- The following *Stata* commands can be used to see how the two indicator variables  $foreign_i$  and  $domestic_i$  are related to one another. Enter the commands:

```
generate sumdum = domestic + foreign
list domestic foreign sumdum, nolabel
summarize domestic foreign sumdum
codebook domestic foreign
tab1 domestic foreign, nolabel
tab2 domestic foreign, nolabel
```

Note that the **nolabel option** on the above commands suppresses the display of the value labels for the indicator variable  $foreign_i$ , and results in the display of the numerical values of  $foreign_i$ . Carefully examine the results of these commands to see how the two dummy variables  $domestic_i$  and  $foreign_i$  related.

### □ Using Dummy Variable Regressors to Test for Differences in Means

Indicator or dummy variables can be used as regressors in linear regression models to perform **t-tests for differences in means** of the dependent variable between the categories of observations identified by the indicator variables.

To illustrate, suppose we want to perform a **two-tail test** of the hypothesis that the mean price of domestic cars equals the mean price of foreign cars.

➤ Let

$\mu_D \equiv$  the population mean price of **domestic cars**, and

$\mu_F \equiv$  the population mean price of **foreign cars**.

➤ The **null and alternative hypotheses** are:

$$H_0: \mu_D = \mu_F \quad \text{against} \quad H_1: \mu_D \neq \mu_F. \quad (1)$$

➤ Define the *sample means* of the variable  $price_i$  as follows:

$\overline{price}_D \equiv$  the sample mean price of *domestic cars*;

$\overline{price}_F \equiv$  the sample mean price of *foreign cars*.

- Enter the following **summarize** command to display the sample means of the variable  $price_i$  for domestic and foreign cars:

```
sort foreign
by foreign: summarize price
```

### Method 1: the ttest command

The *Stata ttest* command can be used to perform a *two-tail t-test* of hypothesis (1) on the assumption that the population standard deviations of  $price_i$  for domestic and foreign cars are unknown but equal ( $\sigma_D = \sigma_F$ ).

As you probably learned in your introductory statistics course, the test statistic appropriate for this test is the following t-statistic:

$$t = \frac{\overline{price}_D - \overline{price}_F}{\sqrt{\frac{(N_D - 1)s_D^2 + (N_F - 1)s_F^2}{N_D + N_F - 2}} \sqrt{\frac{1}{N_D} + \frac{1}{N_F}}} \sim t[N_D + N_F - 2] \text{ under } H_0 \quad (2)$$

where

$s_D^2 =$  the sample variance of  $price_i$  for domestic cars;

$s_F^2 =$  the sample variance of  $price_i$  for foreign cars;

$N_D =$  the number of sample observations for domestic cars;

$N_F =$  the number of sample observations for foreign cars.

- This two-tail t-test can be computed by entering either of the following **ttest** commands:

```
ttest price, by(foreign)
ttest price, by(domestic)
```

Carefully examine the output produced by these commands.

### Method 2: a regression test

An OLS regression can be used to perform a two-tail t-test of hypothesis (1) identical to that computed by above **ttest** commands. In fact, any of three alternative but equivalent simple regression equations can be used to perform this test.

#### ➤ Simple regression 1:

$$\text{price}_i = \beta_0 + \beta_1 \text{domestic}_i + u_i \quad (3)$$

- Estimate regression equation (3) by OLS. Enter the command:

```
regress price domestic
```

Verify that the OLS coefficient estimates for equation (3) have the following interpretation:

$$\hat{\beta}_0 = \overline{\text{price}}_F \quad \text{and} \quad \hat{\beta}_1 = \overline{\text{price}}_D - \overline{\text{price}}_F.$$

- ◆ **Result:** The null hypothesis  $H_0: \mu_D = \mu_F$  can be tested against the alternative hypothesis  $H_1: \mu_D \neq \mu_F$  by using the OLS estimates of equation (3) to perform a *two-tail t-test* or an *F-test* of

$$H_0: \beta_1 = \mu_D - \mu_F = 0 \quad \text{against} \quad H_1: \beta_1 = \mu_D - \mu_F \neq 0.$$

The *t-statistic* for this hypothesis test is:

$$t(\hat{\beta}_1) = \frac{\hat{\beta}_1}{\text{s}\hat{\text{e}}(\hat{\beta}_1)} = \frac{\overline{\text{price}}_D - \overline{\text{price}}_F}{\text{s}\hat{\text{e}}(\overline{\text{price}}_D - \overline{\text{price}}_F)} \sim t[N - 2] = t[72].$$

The equivalent *F-statistic* for this hypothesis test is:

$$F(\hat{\beta}_1) = \frac{\hat{\beta}_1^2}{\text{V}\hat{\text{a}}\text{r}(\hat{\beta}_1)} = \frac{(\overline{\text{price}}_D - \overline{\text{price}}_F)^2}{\text{V}\hat{\text{a}}\text{r}(\overline{\text{price}}_D - \overline{\text{price}}_F)} \sim F[1, N - 2] = F[1, 72].$$

- To display the t-statistic  $t(\hat{\beta}_1)$  and its two-tailed p-value, enter the command:

```
lincom _b[domestic]
```

- To compute the F-statistic  $F(\hat{\beta}_1)$  and its p-value, enter the command:

```
test domestic = 0
```

➤ **Simple regression 2:**

$$\text{price}_i = \alpha_0 + \alpha_1 \text{foreign}_i + u_i \quad (4)$$

- Estimate regression equation (4) by OLS. Enter the command:

```
regress price foreign
```

Verify that the OLS coefficient estimates for equation (4) have the following interpretation:

$$\hat{\alpha}_0 = \overline{\text{price}}_D \quad \text{and} \quad \hat{\alpha}_1 = \overline{\text{price}}_F - \overline{\text{price}}_D.$$

- ◆ **Result:** The null hypothesis  $H_0: \mu_D = \mu_F$  can be tested against the alternative hypothesis  $H_1: \mu_D \neq \mu_F$  by using the OLS estimates of equation (4) to perform a *two-tail t-test* or an *F-test* of

$$H_0: \alpha_1 = \mu_F - \mu_D = 0 \quad \text{against} \quad H_1: \alpha_1 = \mu_F - \mu_D \neq 0.$$

The *t-statistic* for this hypothesis test is:

$$t(\hat{\alpha}_1) = \frac{\hat{\alpha}_1}{\text{s}\hat{\text{e}}(\hat{\alpha}_1)} = \frac{\overline{\text{price}}_F - \overline{\text{price}}_D}{\text{s}\hat{\text{e}}(\overline{\text{price}}_F - \overline{\text{price}}_D)} \sim t[N - 2] = t[72].$$



The equivalent *F-statistic* for this hypothesis test is:

$$F(\hat{\alpha}_1) = \frac{\hat{\alpha}_1^2}{\text{Var}(\hat{\alpha}_1)} = \frac{(\overline{\text{price}_F} - \overline{\text{price}_D})^2}{\text{Var}(\overline{\text{price}_F} - \overline{\text{price}_D})} \sim F[1, N - 2] = F[1, 72].$$

- To display the t-statistic  $t(\hat{\alpha}_1)$  and its two-tailed p-value, enter the command:

```
lincom _b[foreign]
```

- To compute the F-statistic  $F(\hat{\alpha}_1)$  and its p-value, enter the command:

```
test foreign = 0
```

➤ **Simple regression 3:**

$$\text{price}_i = \mu_1 \text{domestic}_i + \mu_2 \text{foreign}_i + u_i \quad (5)$$

*Note:* Regression equation (5) does not contain an intercept coefficient.

- Estimate regression equation (5) by OLS, using the **noconstant option** on the **regress** command. Enter the command:

```
regress price domestic foreign, noconstant
```

Verify that the OLS coefficient estimates for equation (5) have the following interpretation:

$$\hat{\mu}_1 = \overline{\text{price}_D} \quad \text{and} \quad \hat{\mu}_2 = \overline{\text{price}_F}.$$

- ◆ **Result:** The null hypothesis  $H_0: \mu_D = \mu_F$  can be tested against the alternative hypothesis  $H_1: \mu_D \neq \mu_F$  by using the OLS estimates of equation (5) to perform a *two-tail t-test* or an *F-test* of

$$H_0: \mu_1 - \mu_2 = \mu_D - \mu_F = 0 \quad \text{against} \quad H_1: \mu_1 - \mu_2 = \mu_D - \mu_F \neq 0.$$

The *t-statistic* for this hypothesis test is:

$$t(\hat{\mu}_1 - \hat{\mu}_2) = \frac{\hat{\mu}_1 - \hat{\mu}_2}{\widehat{\text{se}}(\hat{\mu}_1 - \hat{\mu}_2)} = \frac{\overline{\text{price}}_D - \overline{\text{price}}_F}{\widehat{\text{se}}(\overline{\text{price}}_D - \overline{\text{price}}_F)} \sim t[N - 2] = t[72].$$

The equivalent **F-statistic** for this hypothesis test is:

$$F(\hat{\mu}_1 - \hat{\mu}_2) = \frac{(\hat{\mu}_1 - \hat{\mu}_2)^2}{\widehat{\text{var}}(\hat{\mu}_1 - \hat{\mu}_2)} = \frac{(\overline{\text{price}}_D - \overline{\text{price}}_F)^2}{\widehat{\text{var}}(\overline{\text{price}}_D - \overline{\text{price}}_F)} \sim F[1, N - 2] = F[1, 72].$$

- To compute the t-statistic  $t(\hat{\mu}_1 - \hat{\mu}_2)$  and its two-tailed p-value, enter the command:

```
lincom _b[domestic] - _b[foreign]
```

- To compute the F-statistic  $F(\hat{\mu}_1 - \hat{\mu}_2)$  and its p-value, enter the command:

```
test domestic - foreign = 0
```

### □ Using Dummy Variable Regressors to Test for Intercept Coefficient Differences

This section demonstrates how **dummy (or indicator) variables** can be used as **additive regressors** in multiple linear regression models to allow for, and test for, **intercept coefficient differences** between domestic and foreign cars in the multiple linear regression model given by PRE (6):

$$\text{price}_i = \beta_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + u_i \quad (6)$$

where

$\text{price}_i$  = the price of the i-th car (in US dollars);

$\text{wgt}_i$  = the weight of the i-th car (in pounds);

$\text{wgt}_i^2$  = the square of  $\text{wgt}_i$ ;

$\text{mpg}_i$  = the miles per gallon for the i-th car (in miles per gallon).

- Before estimating regression equation (6), it is necessary to do two things: first, to make subsequent models introduced in this tutorial easier to write, create the

new variable  $wgt_i = weight_i$ ; and second, create the new variable  $wgt_i^2$ . Enter the following commands:

```
generate wgt = weight
generate wgtsq = wgt^2
summarize wgt weight wgtsq
```

- To estimate regression model (6) by OLS for the full sample of  $N = 74$  observations in data set **auto1.dta**, enter in the Command window:

```
regress price wgt wgtsq mpg
```

- ◆ **Objective:** We want to modify regression equation (6) to allow for mean price differences between foreign and domestic cars of the same size and fuel efficiency sold in North America. The simplest way to do this is to include as an additional regressor in regression equation (6) either the *foreign-car indicator variable*  $foreign_i$  or the *domestic-car indicator variable*  $domestic_i$ .

**NOTE:** The important thing you must do is correctly interpret the regression coefficients in whichever of these two equivalent regression models you choose to use.

- ◆ **Formulation 1:** Include the *foreign-car indicator variable*  $foreign_i$  as an additional regressor in PRE (6).

$$price_i = \beta_0 + \beta_1 wgt_i + \beta_2 wgt_i^2 + \beta_3 mpg_i + \gamma_0 foreign_i + u_i \quad (7)$$

The **population regression function**, or **conditional mean function**, implied by regression equation (7) is:

$$E(\text{price}_i | wgt_i, mpg_i, foreign_i) = \beta_0 + \beta_1 wgt_i + \beta_2 wgt_i^2 + \beta_3 mpg_i + \gamma_0 foreign_i \quad (7.1)$$

The **conditional mean price of domestic cars** is obtained by setting the indicator variable  $foreign_i = 0$  in regression function (7.1):

$$E(\text{price}_i | wgt_i, mpg_i, foreign_i = 0) = \beta_0 + \beta_1 wgt_i + \beta_2 wgt_i^2 + \beta_3 mpg_i \quad (7.2)$$

The **conditional mean price of foreign cars** is obtained by setting the indicator variable  $\text{foreign}_i = 1$  in regression function (7.1):

$$\begin{aligned} E(\text{price}_i | \text{weight}_i, \text{mpg}_i, \text{foreign}_i = 1) \\ &= \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + \gamma_0 \\ &= \beta_0 + \gamma_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i \end{aligned} \quad (7.3)$$

**Interpreting the Regression Coefficients in PRE (7):**

1. The **slope coefficients**  $\beta_j$  ( $j = 1, 2, 3$ ) in PRE (7) are the common slope coefficients for both *domestic* cars and *foreign* cars; these are restricted to be equal for domestic and foreign cars. In other words, the marginal effect of the explanatory variable  $\text{wgt}_i$  for domestic cars is equal to the marginal effect of the explanatory variable  $\text{wgt}_i$  for foreign cars. Similarly, the marginal effect of the explanatory variable  $\text{mpg}_i$  for domestic cars is equal to the marginal effect of the explanatory variable  $\text{mpg}_i$  for foreign cars.
2. The **equation intercept coefficient**  $\beta_0$  in PRE (7) is the **intercept coefficient for domestic cars**, for which the foreign-car dummy variable  $\text{foreign}_i = 0$  by definition.
3. The **intercept coefficient for foreign cars**, for which the foreign-car dummy variable  $\text{foreign}_i = 1$  by definition, is  $\beta_0 + \gamma_0$ .
4. The **slope coefficient**  $\gamma_0$  of the **foreign car indicator variable** in PRE (7) is therefore the foreign-car/domestic-car difference in intercepts:

$$\gamma_0 = \text{the intercept coefficient for foreign cars } (\beta_0 + \gamma_0) \text{ minus the intercept coefficient for domestic cars } (\beta_0)$$

5. A more insightful interpretation of the slope coefficient  $\gamma_0$  in PRE (7) can be obtained by taking the difference between the **conditional mean price of foreign cars** given by (7.3) and the **conditional mean price of domestic cars** given by (7.2):

$$\begin{aligned}
& E(\text{price}_i | \text{wgt}_i, \text{mpg}_i, \text{foreign}_i = 1) - E(\text{price}_i | \text{wgt}_i, \text{mpg}_i, \text{foreign}_i = 0) \\
&= \beta_0 + \gamma_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i - \beta_0 - \beta_1 \text{wgt}_i - \beta_2 \text{wgt}_i^2 - \beta_3 \text{mpg}_i \\
&= \gamma_0
\end{aligned} \tag{7.4}$$

Equation (7.4) demonstrates that the **slope coefficient  $\gamma_0$  of the foreign car indicator variable** in PRE (7) is equal to: (1) the **conditional mean price of foreign cars** of some given size and fuel efficiency *minus* (2) the **conditional mean price of domestic cars** of the same size and fuel efficiency. In other words, the **slope coefficient  $\gamma_0$**  is mean price difference between foreign and domestic cars of the same size (i.e., the same value of  $\text{wgt}_i$ ) and fuel efficiency (i.e., the same value of  $\text{mpg}_i$ ).

- ◆ **Formulation 2:** Include the **domestic-car indicator variable**  $\text{domestic}_i$  as an additional regressor in PRE (6).

$$\text{price}_i = \alpha_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + \delta_0 \text{domestic}_i + u_i \tag{8}$$

The **population regression function, or conditional mean function**, implied by regression equation (8) is:

$$\begin{aligned}
& E(\text{price}_i | \text{wgt}_i, \text{mpg}_i, \text{domestic}_i) \\
&= \alpha_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + \delta_0 \text{domestic}_i
\end{aligned} \tag{8.1}$$

The **conditional mean price of foreign cars** is obtained by setting the indicator variable  $\text{domestic}_i = 0$  in regression function (8.1):

$$E(\text{price}_i | \text{wgt}_i, \text{mpg}_i, \text{domestic}_i = 0) = \alpha_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i \tag{8.2}$$

The **conditional mean price of domestic cars** is obtained by setting the indicator variable  $\text{domestic}_i = 1$  in regression function (8.1):

$$\begin{aligned}
& E(\text{price}_i | \text{weight}_i, \text{mpg}_i, \text{domestic}_i = 1) \\
&= \alpha_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + \delta_0 \\
&= \alpha_0 + \delta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i
\end{aligned} \tag{8.3}$$

**Interpreting the Regression Coefficients in PRE (8):**

1. The **slope coefficients**  $\beta_j$  ( $j = 1, 2, 3$ ) in PRE (8) are again the common slope coefficients for both *domestic* cars and *foreign* cars; these are restricted to be equal for domestic and foreign cars. In other words, the marginal effect of the explanatory variable  $wgt_i$  for domestic cars is equal to the marginal effect of the explanatory variable  $wgt_i$  for foreign cars. Similarly, the marginal effect of the explanatory variable  $mpg_i$  for domestic cars is equal to the marginal effect of the explanatory variable  $mpg_i$  for foreign cars.
2. The **equation intercept coefficient**  $\alpha_0$  in PRE (8) is the **intercept coefficient for foreign cars**, for which the domestic-car dummy variable  $domestic_i = 0$  by definition.
3. The **intercept coefficient for domestic cars**, for which the domestic-car dummy variable  $domestic_i = 1$  by definition, is  $\alpha_0 + \delta_0$ .
4. The **slope coefficient**  $\delta_0$  of the **domestic car indicator variable** in PRE (8) is therefore the domestic-car/foreign-car difference in intercepts:

$\delta_0 =$  the intercept coefficient for domestic cars ( $\alpha_0 + \delta_0$ ) *minus* the intercept coefficient for foreign cars ( $\alpha_0$ )

5. A more insightful interpretation of the **slope coefficient**  $\delta_0$  in PRE (8) can be obtained by taking the difference between the **conditional mean price of domestic cars** given by (8.3) and the **conditional mean price of foreign cars** given by (8.2):

$$\begin{aligned}
 & E(\text{price}_i | wgt_i, mpg_i, domestic_i = 1) - E(\text{price}_i | wgt_i, mpg_i, domestic_i = 0) \\
 &= \alpha_0 + \delta_0 + \beta_1 wgt_i + \beta_2 wgt_i^2 + \beta_3 mpg_i - \alpha_0 - \beta_1 wgt_i - \beta_2 wgt_i^2 - \beta_3 mpg_i \\
 &= \delta_0
 \end{aligned}
 \tag{8.4}$$

Equation (8.4) demonstrates that the *slope coefficient*  $\delta_0$  of the *domestic car indicator variable* in PRE (8) is equal to: (1) the **conditional mean price of domestic cars of some given size and fuel efficiency** *minus* (2) the **conditional mean price of foreign cars of the same size and fuel efficiency**. In other words, the *slope coefficient*  $\delta_0$  is mean price difference between domestic and foreign cars of the same size (i.e., the same value of  $wgt_i$ ) and the same fuel efficiency (i.e., the same value of  $mpg_i$ ).

- Now estimate by OLS regression model (7) containing the foreign-car indicator variable  $foreign_i$  on the full pooled sample of  $N = 74$  observations on both domestic and foreign cars in data set **auto1.dta**. Enter in the Command window:

```
regress price wgt wgtsq mpg foreign
```

- ◆ The results from OLS estimation of regression model (7) can be used to perform a two-tail test of the proposition that the mean price of foreign cars of any given size and fuel efficiency is equal to the mean price of domestic cars of the same size and fuel efficiency. The **null and alternative hypotheses** for regression model (7) are:

$$H_0: \gamma_0 = 0$$

$$H_1: \gamma_0 \neq 0$$

- To perform a **two-tail t-test** of  $H_0: \gamma_0 = 0$  versus  $H_1: \gamma_0 \neq 0$  in regression model (7), enter the following commands:

```
lincom _b[foreign]
return list
display r(estimate)/r(se)
display r(df)
display 2*ttail(r(df), abs(r(estimate)/r(se)))
```

Explain in words what the numeric value of the OLS estimate of  $\gamma_0$  means. What inference would you draw from the results of this two-tailed t-test?

- To perform an **F-test** of  $H_0: \gamma_0 = 0$  versus  $H_1: \gamma_0 \neq 0$  in regression model (7), enter the following commands:

```
test foreign = 0
return list
display r(F)
display sqrt(r(F))
```

What inference would you draw from the results of this F-test? Compare the results of this F-test with those of the preceding t-test. You should be able to explain why these two tests are equivalent.

- Now estimate by OLS regression model (8) containing the domestic-car indicator variable  $\text{domestic}_i$  on the full pooled sample of  $N = 74$  observations on both domestic and foreign cars in data set **auto1.dta**. Enter in the Command window:

```
regress price wgt wgtsq mpg domestic
```

- ◆ The results from OLS estimation of regression model (8) can be used to perform a two-tail test of the proposition that the mean price of domestic cars of any given size and fuel efficiency is equal to the mean price of foreign cars of the same size and fuel efficiency. The **null and alternative hypotheses** for regression model (8) are:

$$H_0: \delta_0 = 0$$

$$H_1: \delta_0 \neq 0$$

- To perform a **two-tail t-test** of  $H_0: \delta_0 = 0$  versus  $H_1: \delta_0 \neq 0$  in regression model (8), enter the following commands:

```
lincom _b[domestic]
return list
display r(estimate)/r(se)
display r(df)
display 2*ttail(r(df), abs(r(estimate)/r(se)))
```

Explain in words what the numeric value of the OLS estimate of  $\delta_0$  means. What inference would you draw from the results of this two-tailed t-test?



- To perform an **F-test** of  $H_0: \delta_0 = 0$  versus  $H_1: \delta_0 \neq 0$  in regression model (8), enter the following commands:

```
test domestic = 0
return list
display r(F)
display sqrt(r(F))
```

What inference would you draw from the results of this F-test? Compare the results of this F-test with those of the preceding t-test. You should be able to explain why these two tests are equivalent.

Finally, compare the results of the t- and F-tests of  $H_0: \gamma_0 = 0$  versus  $H_1: \gamma_0 \neq 0$  in regression model (7) with the results of the t- and F-tests of  $H_0: \delta_0 = 0$  versus  $H_1: \delta_0 \neq 0$  in regression model (8). They are identical. Can you explain why?

- Before finishing this section, re-estimate the regression equations (7) and (8). Enter in the Command window:

```
regress price wgt wgtsq mpg foreign
regress price wgt wgtsq mpg domestic
```

Carefully compare the OLS estimation results for models (7) and (8). These two models are simply two alternative formulations of essentially the same population regression equation. Can you explain why?

### □ Using Dummy Variable Regressors to Test for Both Intercept and Slope Coefficient Differences

---

This section demonstrates how *dummy (or indicator) variables* can be used as regressors in multiple linear regression models to test for **both intercept and slope coefficient differences** between domestic and foreign cars in the multiple linear regression model given by PRE (6):

$$\text{price}_i = \beta_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + u_i \quad (6)$$

- To re-estimate model (6) by OLS for the full sample of  $N = 74$  observations in data set **auto1.dta**, enter in the Command window:

```
regress price weight weightsq mpg
```

□ **Approach 1: Separate Regression Equations**

To allow all four of the regression coefficients in PRE (6) to take different values for domestic and foreign cars, we formulate separate regression equations for these two classes of cars.

Write the *population regression equation (PRE)* for domestic cars as

$$\text{price}_i = \beta_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + u_i \quad (9a)$$

Write the *PRE* for foreign cars as

$$\text{price}_i = \alpha_0 + \alpha_1 \text{wgt}_i + \alpha_2 \text{wgt}_i^2 + \alpha_3 \text{mpg}_i + u_i \quad (9b)$$

Note that the PREs (9a) and (9b) allow *all four* of the regression coefficients to take *different* values for *domestic and foreign cars*, i.e.,  $\alpha_j \neq \beta_j$  ( $j = 0, \dots, 3$ ).

- Compute separate OLS estimates of the PREs (9a) and (9b)*. OLS estimates of regression equations (9a) and (9b) can be computed using the following *Stata* commands:

```
sort foreign
by foreign: regress price wgt wgtsq mpg
```

- The same results could be obtained by entering the following two *Stata* commands:

```
regress price wgt wgtsq mpg if foreign==0
regress price wgt wgtsq mpg if foreign==1
```

The results of separate estimation of PREs (9a) and (9b) can be used together with the results from OLS estimation of PRE (6) to perform a *joint F-test* of the

hypothesis of *full coefficient equality* between the regression functions for domestic and foreign cars. The null and alternative hypotheses for this test are:

$$\begin{aligned} H_0: \alpha_j &= \beta_j & \forall j = 0, \dots, 3 \\ H_1: \alpha_j &\neq \beta_j & j = 0, \dots, 3 \end{aligned}$$

- ◆ **Exercise:** Use a calculator and the results from OLS estimation of regression equations (6), (9a), and (9b) to compute a joint F-test of the above hypothesis.

### □ **Approach 2: Pooled Regression Equations**

It is possible to compute OLS coefficient estimates for the two regression equations (9a) and (9b) using only one regression equation. Such a regression equation is called a *pooled full-interaction regression equation* because it includes as regressors *interaction terms* between either of the indicator variables  $\text{foreign}_i$  and  $\text{domestic}_i$  and all the other regressors in equation (6).

The pooled regression equation can be formulated in *three* equivalent ways.

#### ◆ **Formulation 1:**

$$\begin{aligned} \text{price}_i &= \beta_0 \text{domestic}_i + \beta_1 (\text{domestic}_i)(\text{wgt}_i) + \beta_2 (\text{domestic}_i)(\text{wgt}_i^2) \\ &+ \beta_3 (\text{domestic}_i)(\text{mpg}_i) + \alpha_0 \text{foreign}_i + \alpha_1 (\text{foreign}_i)(\text{wgt}_i) \\ &+ \alpha_2 (\text{foreign}_i)(\text{wgt}_i^2) + \alpha_3 (\text{foreign}_i)(\text{mpg}_i) + u_i \end{aligned} \quad \dots \text{ (10.1)}$$

*Note:* PRE (10.1) does not contain an intercept coefficient.

#### **Interpreting the Regression Coefficients in PRE (10.1):**

1. The regression coefficients  $\beta_j$  ( $j = 0, \dots, 3$ ) in PRE (10.1) are the regression coefficients for *domestic* cars. They are the coefficients of the regressors that include the indicator variable  $\text{domestic}_i$  either by itself or interacted with other regressors.

2. The regression coefficients  $\alpha_j$  ( $j = 0, \dots, 3$ ) in PRE (10.1) are the regression coefficients for *foreign* cars. They are the coefficients of the regressors that include the indicator variable  $\text{foreign}_i$  either by itself or interacted with other regressors.

◆ **Formulation 2:**

$$\begin{aligned} \text{price}_i = & \beta_0 + \beta_1 \text{wgt}_i + \beta_2 \text{wgt}_i^2 + \beta_3 \text{mpg}_i + \gamma_0 \text{foreign}_i \\ & + \gamma_1 (\text{foreign}_i)(\text{wgt}_i) + \gamma_2 (\text{foreign}_i)(\text{wgt}_i^2) + \gamma_3 (\text{foreign}_i)(\text{mpg}_i) + u_i \\ & \dots \quad \textbf{(10.2)} \end{aligned}$$

**Interpreting the Regression Coefficients in PRE (10.2):**

1. The regression coefficients  $\beta_j$  ( $j = 0, \dots, 3$ ) in PRE (10.2) are the regression coefficients for *domestic* cars.
2. The regression coefficients  $\gamma_j$  ( $j = 0, \dots, 3$ ) in PRE (10.2) are interpreted as follows:

$$\gamma_j = \alpha_j - \beta_j \quad j = 0, 1, 2, 3$$

The  $\gamma_j$  coefficients are thus coefficient differences between foreign cars and domestic cars, i.e., they equal a foreign-car regression coefficient  $\alpha_j$  minus the corresponding domestic-car regression coefficient  $\beta_j$ .

3. Regression equation (10.2) implies that the **foreign-car–domestic-car difference in mean price** for cars with any given values of the explanatory variables  $\text{wgt}_i$  and  $\text{mpg}_i$  is given by the following function:

$$\begin{aligned} & E(\text{price}_i | \text{foreign}_i = 1, \text{wgt}_i, \text{mpg}_i) - E(\text{price}_i | \text{foreign}_i = 0, \text{wgt}_i, \text{mpg}_i) \\ & = \gamma_0 + \gamma_1 \text{wgt}_i + \gamma_2 \text{wgt}_i^2 + \gamma_3 \text{mpg}_i \end{aligned}$$

◆ **Formulation 3:**

$$\begin{aligned} \text{price}_i = & \alpha_0 + \alpha_1 \text{wgt}_i + \alpha_2 \text{wgt}_i^2 + \alpha_3 \text{mpg}_i + \delta_0 \text{domestic}_i \\ & + \delta_1 (\text{domestic}_i)(\text{wgt}_i) + \delta_2 (\text{domestic}_i)(\text{wgt}_i^2) + \delta_3 (\text{domestic}_i)(\text{mpg}_i) + u_i \\ & \dots \quad \textbf{(10.3)} \end{aligned}$$

**Interpreting the Regression Coefficients in PRE (10.3):**

1. The regression coefficients  $\alpha_j$  ( $j = 0, \dots, 3$ ) in PRE (10.3) are the regression coefficients for *foreign* cars.
2. The regression coefficients  $\delta_j$  ( $j = 0, \dots, 3$ ) in PRE (10.3) are interpreted as follows:

$$\delta_j = \beta_j - \alpha_j \quad j = 0, 1, 2, 3$$

The  $\delta_j$  coefficients are thus coefficient differences between domestic cars and foreign cars, i.e., they equal a domestic-car regression coefficient  $\beta_j$  minus the corresponding foreign-car regression coefficient  $\alpha_j$ .

3. Regression equation (10.3) implies that the **domestic-car–foreign-car difference in mean price** for cars with any given values of the explanatory variables  $\text{wgt}_i$  and  $\text{mpg}_i$  is given by the following function:

$$\begin{aligned} & E(\text{price}_i | \text{domestic}_i = 1, \text{wgt}_i, \text{mpg}_i) - E(\text{price}_i | \text{domestic}_i = 0, \text{wgt}_i, \text{mpg}_i) \\ & = \delta_0 + \delta_1 \text{wgt}_i + \delta_2 \text{wgt}_i^2 + \delta_3 \text{mpg}_i \end{aligned}$$

The pooled regression equations (10.1), (10.2), and (10.3) contain *interaction terms* as regressors.

- ◆ The **domestic<sub>i</sub> interaction terms** are:

$$(\text{domestic}_i)(\text{wgt}_i), (\text{domestic}_i)(\text{wgt}_i^2), \text{ and } (\text{domestic}_i)(\text{mpg}_i).$$

- ◆ The **foreign<sub>i</sub> interaction terms** are:

$$(\text{foreign}_i)(\text{wgt}_i), (\text{foreign}_i)(\text{wgt}_i^2), \text{ and } (\text{foreign}_i)(\text{mpg}_i).$$

- Before you can estimate the pooled regression equations (10.1), (10.2), and (10.3), you must create the interaction terms. To create the required *dummy variable interaction terms*, enter the following commands:

```
generate dwgt = domestic*wgt
generate dwgtsq = domestic*wgtsq
generate dmpg = domestic*mpg
generate fwgt = foreign*wgt
generate fwgtsq = foreign*wgtsq
generate fmpg = foreign*mpg
```

- Examine these interaction terms by entering the following **list** commands:

```
list wgt domestic dwgt foreign fwgt, nolabel
list wgtsq domestic dwgtsq foreign fwgtsq, nolabel
list mpg domestic dmpg foreign fmpg, nolabel
```

Note the use of the **nolabel option** on the above **list** commands. The **nolabel** option suppresses the display of the value labels for the indicator variable *foreign<sub>i</sub>*, and results in the display of the numerical values of *foreign<sub>i</sub>*.

### *Compute OLS estimates of the pooled regression equation (10.1).*

- 1) Use the following **regress** command with the **noconstant option** to compute OLS estimates of the pooled regression equation (10.1):

```
regress price domestic dwgt dwgtsq dmpg foreign fwgt fwgtsq
fmpg, noconstant
```

Examine the results of this command. In particular, compare them with the separate OLS estimates of population regression equations (9a) and (9b).

- 2) Use the following sequence of **test** commands with the **accumulate** option to test the hypothesis of full coefficient equality between the domestic and foreign regression functions:

$$\begin{aligned}
 H_0: \alpha_j &= \beta_j & \forall j = 0, \dots, 3 & & (11.1) \\
 H_1: \alpha_j &\neq \beta_j & j = 0, \dots, 3 & &
 \end{aligned}$$

Note that the proposition being tested is that the mean price of foreign cars of given weight and fuel efficiency is equal to the mean price of domestic cars of the same weight and fuel efficiency.

- The null hypothesis (11.1) specifies **four linear coefficient restrictions** on PRE (10.1). A **joint F-test** is therefore appropriate.

```
test domestic = foreign
test dwgt = fwgt, accumulate
test dwgtsq = fwgtsq, accum
test dmpg = fmpg, accum
```

- The following commands display the 10 percent, 5 percent, and 1 percent **critical values of the F[4, 66]-distribution** that represents the null distribution of the F-statistic computed by the above **test** commands:

```
display invFtail(4,66, 0.10)
display invFtail(4,66, 0.05)
display invFtail(4,66, 0.01)
```

- 3) Test the hypothesis that the **marginal effect of  $wgt_i$  is the same** for **domestic and foreign cars**:

$$H_0: \alpha_j = \beta_j \quad \forall j = 1, 2 \quad \Leftrightarrow \quad \alpha_1 = \beta_1 \text{ and } \alpha_2 = \beta_2 \quad (12.1)$$

$$H_1: \alpha_j \neq \beta_j \quad j = 1, 2 \quad \Leftrightarrow \quad \alpha_1 \neq \beta_1 \text{ and/or } \alpha_2 \neq \beta_2$$

Enter the following **test** commands:

```
test dwgt = fwgt
test dwgtsq = fwgtsq, acc
```

- Again, use the following commands display the 10 percent, 5 percent, and 1 percent **critical values of the F[2, 66]-distribution**, which is the null distribution of the F-statistic computed by the two **test** commands:

```
display invFtail(2,66, 0.10)
display invFtail(2,66, 0.05)
display invFtail(2,66, 0.01)
```

- 4) Use the following **lincom** commands to compute the differences between the foreign-car coefficient estimates  $\hat{\alpha}_j$  and the domestic-car coefficient estimates  $\hat{\beta}_j$  implied by the OLS sample regression equation corresponding to equation (10.1):

```
lincom _b[foreign] - _b[domestic]
lincom _b[fwgt] - _b[dwgt]
lincom _b[fwgtsq] - _b[dwgtsq]
lincom _b[fmpg] - _b[dmpg]
```

Note that the estimated coefficient differences computed by these **lincom** commands take the form  $\hat{\alpha}_j - \hat{\beta}_j$  ( $j = 0, \dots, 3$ ); i.e., they are foreign-car coefficient estimates *minus* the corresponding domestic-car coefficient estimates.

- 5) The following **lincom** commands compute the differences between the domestic-car coefficient estimates  $\hat{\beta}_j$  and the foreign-car coefficient estimates  $\hat{\alpha}_j$  implied by the OLS sample regression equation corresponding to equation (10.1):

```
lincom _b[domestic] - _b[foreign]
lincom _b[dwgt] - _b[fwgt]
lincom _b[dwgtsq] - _b[fwgtsq]
lincom _b[dmpg] - _b[fmpg]
```

Note that the estimated coefficient differences computed by these **lincom** commands take the form  $\hat{\beta}_j - \hat{\alpha}_j$  ( $j = 0, \dots, 3$ ); i.e., they are domestic-car coefficient estimates *minus* the corresponding foreign-car coefficient estimates. Compare the results of these **lincom** commands with those for the **lincom** commands in subsection (4) above.

***Compute OLS estimates of the pooled regression equation (10.2).***

- 1) Use the following command to compute OLS estimates of the pooled regression equation (10.2):



```
regress price wgt wgtsq mpg foreign fwgt fwgtsq fmpg
```

- 2) Use the following **test** command to test the proposition that the mean price of foreign cars of given weight and fuel efficiency is equal to the mean price of domestic cars of the same weight and fuel efficiency. This proposition corresponds to the hypothesis of full coefficient equality between the domestic and foreign regression functions:

$$\begin{aligned} H_0: \gamma_j &= 0 & \forall j = 0, \dots, 3 \\ H_1: \gamma_j &\neq 0 & j = 0, \dots, 3 \end{aligned} \quad (11.2)$$

Remember the *interpretation of the  $\gamma_j$  coefficients* in PRE (10.2):

$$\gamma_j = \alpha_j - \beta_j \quad j = 0, 1, 2, 3$$

The estimates of the  $\gamma_j$  coefficients in PRE (10.2) therefore equal the coefficient estimates for foreign cars *minus* the corresponding coefficient estimates for domestic cars:

$$\hat{\gamma}_j = \hat{\alpha}_j - \hat{\beta}_j \quad j = 0, 1, 2, 3$$

- The null hypothesis (11.2) specifies **four linear coefficient restrictions** on PRE (10.2). A **joint F-test** is therefore appropriate. This joint F-test can be computed using only one **test** command:

```
test foreign fwgt fwgtsq fmpg
return list
```

- 3) Test the hypothesis that the **marginal effect of  $weight_i$**  is **the same** for **domestic and foreign cars**:

$$\begin{aligned} H_0: \gamma_j &= 0 \quad \forall j = 1, 2 & \Leftrightarrow & \gamma_1 = 0 \text{ and } \gamma_2 = 0 \\ H_1: \gamma_j &\neq 0 \quad j = 1, 2 & \Leftrightarrow & \gamma_1 \neq 0 \text{ and/or } \gamma_2 \neq 0 \end{aligned} \quad (12.2)$$

- Enter the following **test** command:

```
test fwgt fwgtsq
```

- 4) Use the following **lincom** commands to compute the *foreign-car coefficient estimates*  $\hat{\alpha}_j$  implied by the OLS sample regression equation corresponding to equation (10.2):

```
lincom _b[_cons] + _b[foreign]           (computes  $\hat{\alpha}_0$ )
lincom _b[wgt] + _b[fwgt]                (computes  $\hat{\alpha}_1$ )
lincom _b[wgtsq] + _b[fwgtsq]           (computes  $\hat{\alpha}_2$ )
lincom _b[mpg] + _b[fmpg]                (computes  $\hat{\alpha}_3$ )
```

Note that the *foreign-car coefficient estimates* computed by these **lincom** commands take the form

$$\hat{\alpha}_j = \hat{\beta}_j + \hat{\gamma}_j \quad j = 0, 1, 2, 3$$

- 5) OLS estimates of regression equation (10.2) can be used to compute the **estimated mean price difference** between a *foreign car* that weighs 2,300 pounds and has fuel efficiency of 25 miles per gallon and an **identical domestic car** that also weighs 2,300 pounds and has fuel efficiency of 25 miles per gallon, i.e., to compute an estimate of the mean price difference

$$\begin{aligned} & E(\text{price}_i | \text{foreign}_i = 1, \text{wgt}_i = 2300, \text{mpg}_i = 25) \\ & - E(\text{price}_i | \text{foreign}_i = 0, \text{wgt}_i = 2300, \text{mpg}_i = 25) \\ & = \gamma_0 + \gamma_1 2300 + \gamma_2 (2300)(2300) + \gamma_3 25 \end{aligned}$$

- Enter the following **lincom** command to compute the estimated **foreign-car–domestic-car mean price difference** for cars that weigh 2,300 pounds and get 25 miles to the gallon:

```
lincom _b[foreign] + _b[fwgt]*2300 + _b[fwgtsq]*2300*2300 +
_b[fmpg]*25
```

You should be able to explain and interpret all the output produced by this **lincom** command.

**Compute OLS estimates of the pooled regression equation (10.3).**

- 1) Use the following command to compute OLS estimates of the pooled regression equation (10.3):

```
regress price wgt wgtsq mpg domestic dwgt dwgtsq dmpg
```

- 2) Use the following **test** command to test the proposition that the mean price of foreign cars of given weight and fuel efficiency is equal to the mean price of domestic cars of the same weight and fuel efficiency. This proposition corresponds to the hypothesis of full coefficient equality between the domestic and foreign regression functions:

$$\begin{aligned} H_0: \delta_j &= 0 & \forall j = 0, \dots, 3 & & (11.3) \\ H_1: \delta_j &\neq 0 & j = 0, \dots, 3 & & \end{aligned}$$

Remember the *interpretation of the  $\delta_j$  coefficients* in PRE (10.3):

$$\delta_j = \beta_j - \alpha_j \quad j = 0, 1, 2, 3$$

The estimates of the  $\delta_j$  coefficients in PRE (10.3) therefore equal the coefficient estimates for domestic cars *minus* the corresponding coefficient estimates for foreign cars:

$$\hat{\delta}_j = \hat{\beta}_j - \hat{\alpha}_j \quad j = 0, 1, 2, 3$$

- The null hypothesis (11.3) specifies **four linear coefficient restrictions** on PRE (10.3). A **joint F-test** is therefore appropriate. This joint F-test can be computed using only one **test** command:

```
test domestic dwgt dwgtsq dmpg
```

- 3) Test the hypothesis that the **marginal effect of  $weight_i$  is the same for domestic and foreign cars**:

$$\begin{aligned} H_0: \delta_j &= 0 & \forall j = 1, 2 & \Leftrightarrow & \delta_1 = 0 \text{ and } \delta_2 = 0 & & (12.3) \\ H_1: \delta_j &\neq 0 & j = 1, 2 & \Leftrightarrow & \delta_1 \neq 0 \text{ and/or } \delta_2 \neq 0 & & \end{aligned}$$

Enter the following **test** command:

```
test dwgt dwgtsq
```

- 4) Use the following **lincom** commands to compute the *domestic-car coefficient estimates*  $\hat{\beta}_j$  implied by the OLS sample regression equation corresponding to equation (10.3):

```
lincom _b[_cons] + _b[domestic]           (computes  $\hat{\beta}_0$ )
lincom _b[wgt] + _b[dwgt]                 (computes  $\hat{\beta}_1$ )
lincom _b[wgtsq] + _b[dwgtsq]            (computes  $\hat{\beta}_2$ )
lincom _b[mpg] + _b[dmpg]                 (computes  $\hat{\beta}_3$ )
```

Note that the domestic-car coefficient estimates computed by these **lincom** commands take the form

$$\hat{\beta}_j = \hat{\alpha}_j + \hat{\delta}_j \quad j = 0, 1, 2, 3$$

- 5) OLS estimates of regression equation (10.3) can be used to compute the **estimated mean price difference** between a *domestic car* that weighs 2,300 pounds and has fuel efficiency of 25 miles per gallon and an identical *foreign car* that also weighs 2,300 pounds and has fuel efficiency of 25 miles per gallon, i.e., to compute an estimate of the mean price difference

$$\begin{aligned} & E(\text{price}_i | \text{domestic}_i = 1, \text{wgt}_i = 2300, \text{mpg}_i = 25) \\ & - E(\text{price}_i | \text{domestic}_i = 0, \text{wgt}_i = 2300, \text{mpg}_i = 25) \\ & = \delta_0 + \delta_1 2300 + \delta_2 (2300)(2300) + \delta_3 25 \end{aligned}$$

- Enter the following **lincom** command to compute the estimated **domestic-car–foreign-car mean price difference** for cars that weigh 2,300 pounds and get 25 miles to the gallon:

```
lincom _b[domestic] + _b[dwgt]*2300 + _b[dwgtsq]*2300*2300 +
_b[dmpg]*25
```

You should be able to explain and interpret all the output produced by this **lincom** command.

***Exercise:*** Use the log file for this tutorial to compare the OLS coefficient estimates of pooled regression equations (10.1), (10.2) and (10.3). You should be able to see from inspection of the log file exactly how the three sets of coefficient estimates for these three pooled regressions are related to each other.

---

## □ Preparing to End Your *Stata* Session

---

**Before you end your *Stata* session**, you should do three things.

- First, you may want to **save the current data set** (although you will not need it for future tutorials). Enter the following **save** command to save the current data set as *Stata*-format data set **auto8.dta**:

```
save auto8
```

- Second, **close the command log file** you have been recording. Enter the command:

```
cmdlog close
```

- Third, **close the log file** you have been recording. Enter the command:

```
log close
```

---

## □ End Your *Stata* Session -- exit

---

- **To end your *Stata* session**, enter either one of the following commands:

```
exit  
exit, clear
```

---

## □ Cleaning Up and Clearing Out

---

**After returning to Windows**, you should copy all the files you have used and created during your *Stata* session to your own diskette. These files will be found in the *Stata working directory*, which is usually **C:\data** on the computers in Dunning 350, and **D:\courses** on the computers in MC B111. There are two files you will want to be sure you have: the *Stata* log file **351tutorial8.log**, and the command log file **351tutorial8.txt**. If you saved the *Stata*-format data set **auto8.dta**, you will probably want to take it with you as well. Use the Windows **copy** command to copy any files you want to keep to your own portable electronic storage device (e.g., flash memory stick) in the E:-drive (or a diskette in A:-drive).

Finally, **as a courtesy to other users** of the computing classroom, please delete all the files you have used or created from the *Stata* working directory.

---