ECON 351* -- NOTE 22

Tests for Coefficient Differences: Examples 1

1. Introduction

Model and Data

The model we use to illustrate some common tests for coefficient differences is the model of North American car prices given by the PRE

$$\text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + u_i \qquad \textbf{(1)}$$

where

price_i = the price of the i-th car (in US dollars);
weight_i = the weight of the i-th car (in pounds);
mpg_i = the fuel efficiency of the i-th car (in miles per gallon);
N = 74 = the number of observations in the estimation sample.

The **estimation sample data** consist of observations on the prices and related variables of 74 cars sold in North America, of which $N_d = 52$ are domestically-made cars and $N_f = 22$ are foreign-made cars. To distinguish between domestic and foreign cars, the estimation sample includes an ***indicator (dummy) variable*** foreign_i, defined as follows:

$$\text{foreign}_i = 1 \text{ for all foreign cars}$$
$$= 0 \text{ for all domestic cars.}$$

We can alternatively define the indicator variable domestic_i such that

$$\text{domestic}_i = 1 \text{ for all domestic cars}$$
$$= 0 \text{ for all foreign cars.}$$

The two indicator variables domestic_i and foreign_i are related according to the following linear equation:

$$\text{domestic}_i = 1 - \text{foreign}_i \qquad \forall \;\; i = 1, \ldots, N.$$

<u>The Problem</u>

We wish to use the sample data to perform various tests for **differences in the regression coefficients of model (1) between *domestic* and *foreign* cars**. That is, we wish to investigate whether and how the population regression functions for domestic and foreign car prices differ from one another.

To allow all four of the regression coefficients in PRE (1) to take different values for domestic and foreign cars, write separate regression equations for domestic and foreign cars. Let β_j (j = 0, …, 3) denote the regression coefficients for domestic cars; and let α_j (j = 0, …, 3) denote the regression coefficients for foreign cars.

- The **population regression equation (PRE) for *domestic cars*** is

$$price_i = \beta_0 + \beta_1 weight_i + \beta_2 weight_i^2 + \beta_3 mpg_i + u_i, \quad i = 1, ..., N_d \quad \textbf{(1a)}$$
$$\forall \; i \in domestic_i = 1 \text{ or } foreign_i = 0.$$

  The $\boldsymbol{\beta_j}$ **(j = 0, …, 3)** are the **regression coefficients for *domestic cars***.

- The **population regression equation (PRE) for *foreign cars*** is

$$price_i = \alpha_0 + \alpha_1 weight_i + \alpha_2 weight_i^2 + \alpha_3 mpg_i + u_i, \quad i = 1, ..., N_f \quad \textbf{(1b)}$$
$$\forall \; i \in domestic_i = 0 \text{ or } foreign_i = 1.$$

  The $\boldsymbol{\alpha_j}$ **(j = 0, …, 3)** are the **regression coefficients for *foreign cars***.

- Note that the PREs (1a) and (1b) allow all four of the regression coefficients to take different values for domestic and foreign cars.

2. Tests for Full Coefficient Equality

- ***The null and alternative hypotheses for a joint test of full coefficient equality between PREs (1a) and (1b)*** are written in terms of the regression coefficients of PREs (1a) and (1b) as

$$H_0: \ \alpha_j = \beta_j \qquad \forall \ j = 0, \ldots, 3$$
$$H_1: \ \alpha_j \neq \beta_j . \qquad j = 0, \ldots, 3$$

- ***Two alternative but equivalent approaches*** may be used to test for full coefficient equality between the population regression functions for domestically-made and foreign-made cars:

  - The ***separate regressions approach***
  - The ***pooled regression approach***

  These two approaches are outlined in turn below.

1. *Separate* Regressions Approach – Test for Full Coefficient Equality

❑ **Step 1: Formulate the *restricted* model implied by H_0.**

- The ***unrestricted* model** consists of the domestic PRE (1a) and the foreign PRE (1b):

$$\text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + u_i \qquad i = 1, ..., N_d \qquad \textbf{(1a)}$$

$$\text{price}_i = \alpha_0 + \alpha_1 \text{weight}_i + \alpha_2 \text{weight}_i^2 + \alpha_3 \text{mpg}_i + u_i \qquad i = 1, ..., N_f \qquad \textbf{(1b)}$$

- To obtain the restricted model corresponding to the null hypothesis H_0, **substitute into PRE (1b)** for foreign cars **the four coefficient restrictions specified by H_0**: i.e., set $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$ and $\alpha_3 = \beta_3$ in equation (1b).

Under H_0, equations (1a) and (1b) can be written as:

$$\text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + u_i \qquad i = 1, ..., N_d \qquad \textbf{(1a)}$$

$$\text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + u_i \qquad i = 1, ..., N_f \qquad \textbf{(1b)}$$

Obviously, **regression equations (1a) and (1b) are identical under the null hypothesis H_0**: they have identical coefficients and identical regressors. In other words, equations (1a) and (1b) are identical for all observations -- both those for domestic cars and those for foreign cars.

- ***Result:*** The ***restricted*** **model** for all observations is simply PRE (1) below.

$$\text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i + u_i$$
$$i = 1, ..., N = N_d + N_f \qquad \textbf{(1)}$$

❑ **Step 2: Estimate the *restricted* model implied by H_0.**

- ***Compute OLS estimates of the restricted model (1) on all N observations***.

The restricted OLS SRE can be written in general as

$$\text{price}_i = \tilde{\beta}_0 + \tilde{\beta}_1 \text{weight}_i + \tilde{\beta}_2 \text{weight}_i^2 + \tilde{\beta}_3 \text{mpg}_i + \tilde{u}_i, \qquad i = 1, ..., N$$

and has RSS = RSS_0 with degrees of freedom $df_0 = N - K_0 = N - 4$.

Table 1: OLS Estimates of the Restricted Pooled Regression Equation (1)

$\tilde{\beta}_j$	$\tilde{\beta}_0 = 19804.8$	$\tilde{\beta}_1 = -9.03956$	$\tilde{\beta}_2 = .001679$	$\tilde{\beta}_3 = -124.77$
(t-ratios)	(3.443)	(-3.111)	(3.791)	(-1.531)
	$RSS_0 = 372311797.0$		$df_0 = N - K_0 = 74 - 4 = 70$	

❑ **<u>Step 3</u>: Estimate the *<u>unrestricted</u>* model implied by H_1.**

The ***unrestricted* model** consists of the domestic PRE (1a) and the foreign PRE (1b):

$$price_i = \beta_0 + \beta_1 weight_i + \beta_2 weight_i^2 + \beta_3 mpg_i + u_i \qquad i = 1, ..., N_d \qquad \textbf{(1a)}$$

$$price_i = \alpha_0 + \alpha_1 weight_i + \alpha_2 weight_i^2 + \alpha_3 mpg_i + u_i \qquad i = 1, ..., N_f \qquad \textbf{(1b)}$$

- ***Compute separate OLS estimates of equations (1a) and (1b)***.

The **OLS SRE for domestic cars** -- estimated on the subsample of observations for which $domestic_i = 1$ or $foreign_i = 0$ -- can be written in general as

$$price_i = \hat{\beta}_0 + \hat{\beta}_1 weight_i + \hat{\beta}_2 weight_i^2 + \hat{\beta}_3 mpg_i + \hat{u}_{di}, \qquad i = 1, ..., N_d$$

and has RSS = RSS_d with degrees of freedom $df_d = N_d - K_0 = N_d - 4$.

The **OLS SRE for foreign cars** -- estimated on the subsample of observations for which $domestic_i = 0$ or $foreign_i = 1$ -- can be written in general as

$$price_i = \hat{\alpha}_0 + \hat{\alpha}_1 weight_i + \hat{\alpha}_2 weight_i^2 + \hat{\alpha}_3 mpg_i + \hat{u}_{fi}, \qquad i = 1, ..., N_f$$

and has RSS = RSS_f with degrees of freedom $df_f = N_f - K_0 = N_f - 4$.

Table 2: OLS Estimates of Separate Regression Equations (1a) and (1b)

$\hat{\beta}_j$ (t-ratios)	$\hat{\beta}_0 = 11972.1$ (1.515)	$\hat{\beta}_1 = -8.8130$ (-2.654)	$\hat{\beta}_2 = .001898$ (4.112)	$\hat{\beta}_3 = 78.075$ (0.615)
		$RSS_d = 187040495.0$	$df_d = N_d - K_0 = 52 - 4 = 48$	
$\hat{\alpha}_j$ (t-ratios)	$\hat{\alpha}_0 = 9497.7$ (1.104)	$\hat{\alpha}_1 = -6.6685$ (-1.011)	$\hat{\alpha}_2 = .002324$ (1.808)	$\hat{\alpha}_3 = -22.077$ (-0.405)
		$RSS_f = 26208743.4$	$df_f = N_f - K_0 = 22 - 4 = 18$	

❑ **<u>Step 4</u>: Compute *sample value* of the F-test statistic.**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} \sim F[df_0 - df_1, df_1] \text{ under } H_0.$$

- $RSS_0 = 372311797.0$ with $df_0 = N - K_0 = 70$.
- $RSS_1 = RSS_d + RSS_f = 187040495.0 + 26208743.4 = 213249238.4$;
  with $df_1 = df_d + df_f = 48 + 18 = 66$.
- $df_0 - df_1 = 70 - 66 = 4$.

Thus, the sample value of the F-statistic under the null hypothesis H_0 is:

$$\begin{aligned} F_0 &= \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} \\ &= \frac{(372311797.0 - 213249238.4)/(70 - 66)}{213249238.4/66} \\ &= \frac{159062558.6/4}{213249238.4/66} \\ &= \frac{39765639.65}{3231049.067} \\ &= 12.307. \end{aligned}$$

<u>Result</u>: **F_0 = 12.31**.

❑ **<u>Step 5</u>: Apply the usual *decision rule* for an F-test.**

2. ***Pooled*** Regression Approach -- Test for Full Coefficient Equality

The ***pooled full-interaction regression equation*** can be formulated in ***three*** equivalent ways -- all of which use indicator variables as regressors.

- **Formulation 1: *no* base group**

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

... (2.0)

Note: PRE (2.0) *does not contain an intercept coefficient.*

The ***unrestricted* OLS SRE** obtained by OLS estimation of PRE (2.0) on the full sample of N observations can be written in general as

$$\begin{aligned}\text{price}_i = {} & \hat{\beta}_0 \text{domestic}_i + \hat{\beta}_1(\text{domestic}_i)(\text{weight}_i) + \hat{\beta}_2(\text{domestic}_i)(\text{weight}_i^2) \\ & + \hat{\beta}_3(\text{domestic}_i)(\text{mpg}_i) + \hat{\alpha}_0 \text{foreign}_i + \hat{\alpha}_1(\text{foreign}_i)(\text{weight}_i) \\ & + \hat{\alpha}_2(\text{foreign}_i)(\text{weight}_i^2) + \hat{\alpha}_3(\text{foreign}_i)(\text{mpg}_i) + \hat{u}_i\end{aligned}$$

and has RSS = RSS_1 with degrees of freedom $df_1 = N - K_1 = N - 8$.

Table 3: OLS Estimates of Pooled Regression Equation (2.0)

$\hat{\beta}_j$ (t-ratios)	$\hat{\beta}_0 = 11972.1$ (1.663)	$\hat{\beta}_1 =$ -8.8130 (-2.915)	$\hat{\beta}_2 = .001898$ (4.516)	$\hat{\beta}_3 = 78.075$ (0.676)
$\hat{\alpha}_j$ (t-ratios)	$\hat{\alpha}_0 = 9497.7$ (0.741)	$\hat{\alpha}_1 =$ -6.6685 (-0.679)	$\hat{\alpha}_2 = .002324$ (1.214)	$\hat{\alpha}_3 =$ -22.077 (-0.272)
$\hat{\delta}_j = \hat{\beta}_j - \hat{\alpha}_j$ (t-ratios)	$\hat{\delta}_0 = 2474.4$ (0.168)	$\hat{\delta}_1 =$ -2.1446 (-0.209)	$\hat{\delta}_2 =$ -000425 (-0.217)	$\hat{\delta}_3 = 100.15$ (0.709)
$\hat{\gamma}_j = \hat{\alpha}_j - \hat{\beta}_j$ (t-ratios)	$\hat{\gamma}_0 =$ -2474.4 (-0.168)	$\hat{\gamma}_1 = 2.1446$ (0.209)	$\hat{\gamma}_2 = .000425$ (0.217)	$\hat{\gamma}_3 =$ -100.15 (-0.709)
	$RSS_1 = 213249239.0$		$df_1 = N - 2K_0 = 74 - 8 = 66$	

♦ **<u>Formulation 2</u>: *domestic cars* as base group**

$$\begin{aligned} \text{price}_i = \beta_0 + \beta_1 \text{weight}_i + \beta_2 \text{weight}_i^2 + \beta_3 \text{mpg}_i \\ + \gamma_0 \text{foreign}_i + \gamma_1 (\text{foreign}_i)(\text{weight}_i) + \gamma_2 (\text{foreign}_i)(\text{weight}_i^2) \\ + \gamma_3 (\text{foreign}_i)(\text{mpg}_i) + u_i \end{aligned} \quad \textbf{... (2.1)}$$

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

The ***unrestricted*** **OLS SRE** obtained by OLS estimation of PRE (2.1) on the full sample of N observations can be written in general as

$$\begin{aligned} \text{price}_i = \hat{\beta}_0 + \hat{\beta}_1 \text{weight}_i + \hat{\beta}_2 \text{weight}_i^2 + \hat{\beta}_3 \text{mpg}_i \\ + \hat{\gamma}_0 \text{foreign}_i + \hat{\gamma}_1 (\text{foreign}_i)(\text{weight}_i) + \hat{\gamma}_2 (\text{foreign}_i)(\text{weight}_i^2) \\ + \hat{\gamma}_3 (\text{foreign}_i)(\text{mpg}_i) + \hat{u}_i \end{aligned}$$

and has RSS = RSS_1 with degrees of freedom $df_1 = N - K_1 = N - 8$.

Table 4: OLS Estimates of Pooled Regression Equation (2.1)

$\hat{\beta}_j$ (t-ratios)	$\hat{\beta}_0 = 11972.1$ (1.663)	$\hat{\beta}_1 = -8.8130$ (-2.915)	$\hat{\beta}_2 = .001898$ (4.516)	$\hat{\beta}_3 = 78.075$ (0.676)
$\hat{\gamma}_j$ (t-ratios)	$\hat{\gamma}_0 = -2474.4$ (-0.168)	$\hat{\gamma}_1 = 2.1446$ (0.209)	$\hat{\gamma}_2 = .000425$ (0.217)	$\hat{\gamma}_3 = -100.15$ (-0.709)
$\hat{\alpha}_j = \hat{\beta}_j + \hat{\gamma}_j$ (t-ratios)	$\hat{\alpha}_0 = 9497.7$ (0.741)	$\hat{\alpha}_1 = -6.6685$ (-0.679)	$\hat{\alpha}_2 = .002324$ (1.214)	$\hat{\alpha}_3 = -22.077$ (-0.272)
	$RSS_1 = 213249239.0$		$df_1 = N - 2K_0 = 74 - 8 = 66$	

1. The regression coefficients β_j (j = 0, ..., 3) in PRE (2.1) are the regression coefficients for ***domestic*** cars.

2. The regression coefficients γ_j (j = 0, ..., 3) in PRE (2.1) are interpreted as follows: $\gamma_j = \alpha_j - \beta_j$ (j = 0, 1, 2, 3).

♦ **<u>Formulation 3</u>: *foreign cars* as base group**

$$\text{price}_i = \alpha_0 + \alpha_1 \text{weight}_i + \alpha_2 \text{weight}_i^2 + \alpha_3 \text{mpg}_i + \delta_0 \text{domestic}_i + \delta_1 (\text{domestic}_i)(\text{weight}_i) + \delta_2 (\text{domestic}_i)(\text{weight}_i^2) + \delta_3 (\text{domestic}_i)(\text{mpg}_i) + u_i \quad \text{... (2.2)}$$

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

The ***unrestricted*** **OLS SRE** obtained by OLS estimation of PRE (2.2) on the full sample of N observations can be written in general as

$$\text{price}_i = \hat{\alpha}_0 + \hat{\alpha}_1 \text{weight}_i + \hat{\alpha}_2 \text{weight}_i^2 + \hat{\alpha}_3 \text{mpg}_i + \hat{\delta}_0 \text{domestic}_i + \hat{\delta}_1 (\text{domestic}_i)(\text{weight}_i) + \hat{\delta}_2 (\text{domestic}_i)(\text{weight}_i^2) + \hat{\delta}_3 (\text{domestic}_i)(\text{mpg}_i) + \hat{u}_i$$

and has RSS = RSS_1 with degrees of freedom $df_1 = N - K_1 = N - 8$.

Table 5: OLS Estimates of Pooled Regression Equation (2.2)

$\hat{\alpha}_j$ (t-ratios)	$\hat{\alpha}_0 = 9497.7$ (0.741)	$\hat{\alpha}_1 = -6.6685$ (-0.679)	$\hat{\alpha}_2 = .002324$ (1.214)	$\hat{\alpha}_3 = -22.077$ (-0.272)
$\hat{\delta}_j$ (t-ratios)	$\hat{\delta}_0 = 2474.4$ (0.168)	$\hat{\delta}_1 = -2.1446$ (-0.209)	$\hat{\delta}_2 = -.000425$ (-0.217)	$\hat{\delta}_3 = 100.15$ (0.709)
$\hat{\beta}_j = \hat{\alpha}_j + \hat{\delta}_j$ (t-ratios)	$\hat{\beta}_0 = 11972.1$ (1.663)	$\hat{\beta}_1 = -8.8130$ (-2.915)	$\hat{\beta}_2 = .001898$ (4.516)	$\hat{\beta}_3 = 78.075$ (0.676)
	$RSS_1 = 213249239.0$		$df_1 = N - 2K_0 = 74 - 8 = 66$	

1. The regression coefficients α_j (j = 0, ..., 3) in PRE (2.2) are the regression coefficients for ***foreign*** cars.

2. The regression coefficients δ_j (j = 0, ..., 3) in PRE (2.2) are interpreted as follows: $\delta_j = \beta_j - \alpha_j$ (j = 0, 1, 2, 3).

♦ **Joint Test for Full Coefficient Equality -- Pooled Regression Approach**

H0: $\alpha_j = \beta_j \quad \forall \; j = 0, \ldots, 3$ in pooled PRE (2.0)

$\gamma_j = \alpha_j - \beta_j = 0 \quad \forall \; j = 0, \ldots, 3$ in pooled PRE (2.1)

$\delta_j = \beta_j - \alpha_j = 0 \quad \forall \; j = 0, \ldots, 3$ in pooled PRE (2.2)

H1: $\alpha_j \neq \beta_j \quad j = 0, \ldots, 3$ in pooled PRE (2.0)

$\gamma_j = \alpha_j - \beta_j \neq 0 \quad j = 0, \ldots, 3$ in pooled PRE (2.1)

$\delta_j = \beta_j - \alpha_j \neq 0 \quad j = 0, \ldots, 3$ in pooled PRE (2.2)

❑ **Step 1: Formulate and estimate the *restricted* model implied by H_0.**

To obtain the restricted model corresponding to the null hypothesis H_0, **substitute into any one of the pooled full-interaction PREs (2.0), (2.1), or (2.2) the four coefficient restrictions specified by H_0**.

• ***In pooled regression equation (2.0)*, set $\alpha_0 = \beta_0$, $\alpha_1 = \beta_1$, $\alpha_2 = \beta_2$, and $\alpha_3 = \beta_3$.**

$$\begin{aligned} \text{price}_i = {} & \beta_0 \text{domestic}_i + \beta_1(\text{domestic}_i)(\text{weight}_i) + \beta_2(\text{domestic}_i)(\text{weight}_i^2) \\ & + \beta_3(\text{domestic}_i)(\text{mpg}_i) + \alpha_0 \text{foreign}_i + \alpha_1(\text{foreign}_i)(\text{weight}_i) \\ & + \alpha_2(\text{foreign}_i)(\text{weight}_i^2) + \alpha_3(\text{foreign}_i)(\text{mpg}_i) + u_i \end{aligned} \qquad \text{... (2.0)}$$

***Restricted* model implied by H_0** is:

$$\begin{aligned} \text{price}_i = {} & \beta_0 \text{domestic}_i + \beta_1(\text{domestic}_i)(\text{weight}_i) + \beta_2(\text{domestic}_i)(\text{weight}_i^2) \\ & + \beta_3(\text{domestic}_i)(\text{mpg}_i) + \beta_0 \text{foreign}_i + \beta_1(\text{foreign}_i)(\text{weight}_i) \\ & + \beta_2(\text{foreign}_i)(\text{weight}_i^2) + \beta_3(\text{foreign}_i)(\text{mpg}_i) + u_i \\ = {} & \beta_0(\text{domestic}_i + \text{foreign}_i) + \beta_1(\text{domestic}_i + \text{foreign}_i)(\text{weight}_i) \\ & + \beta_2(\text{domestic}_i + \text{foreign}_i)(\text{weight}_i^2) \\ & + \beta_3(\text{domestic}_i + \text{foreign}_i)(\text{mpg}_i) + u_i \\ = {} & \beta_0 + \beta_1(\text{weight}_i) + \beta_2(\text{weight}_i^2) + \beta_3(\text{mpg}_i) + u_i \end{aligned}$$

where we have used the adding-up property

$$\text{domestic}_i + \text{foreign}_i = 1 \quad \forall\ i = 1, \ldots, N.$$

- ***In pooled regression equation (2.1)*, set $\gamma_0 = 0$, $\gamma_1 = 0$, $\gamma_2 = 0$ and $\gamma_3 = 0$**.

$$\begin{aligned}\text{price}_i &= \beta_0 + \beta_1\text{weight}_i + \beta_2\text{weight}_i^2 + \beta_3\text{mpg}_i \\ &\quad + \gamma_0\text{foreign}_i + \gamma_1(\text{foreign}_i)(\text{weight}_i) + \gamma_2(\text{foreign}_i)(\text{weight}_i^2) \\ &\quad + \gamma_3(\text{foreign}_i)(\text{mpg}_i) + u_i\end{aligned} \qquad \textbf{... (2.1)}$$

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

<u>Restricted</u> model implied by H_0 is:

$$\text{price}_i = \beta_0 + \beta_1\text{weight}_i + \beta_2\text{weight}_i^2 + \beta_3\text{mpg}_i + u_i \qquad i = 1, ..., N \qquad \textbf{(1)}$$

- ***In pooled regression equation (2.2)*, set $\delta_0 = 0$, $\delta_1 = 0$, $\delta_2 = 0$ and $\delta_3 = 0$**.

$$\begin{aligned}\text{price}_i &= \alpha_0 + \alpha_1\text{weight}_i + \alpha_2\text{weight}_i^2 + \alpha_3\text{mpg}_i \\ &\quad + \delta_0\text{domestic}_i + \delta_1(\text{domestic}_i)(\text{weight}_i) + \delta_2(\text{domestic}_i)(\text{weight}_i^2) \\ &\quad + \delta_3(\text{domestic}_i)(\text{mpg}_i) + u_i\end{aligned} \qquad \textbf{... (2.2)}$$

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

<u>Restricted</u> model implied by H_0 is:

$$\text{price}_i = \alpha_0 + \alpha_1\text{weight}_i + \alpha_2\text{weight}_i^2 + \alpha_3\text{mpg}_i + u_i \qquad i = 1, ..., N.$$

➢ **<u>Result</u>:** The ***<u>restricted</u>*** **model** for all observations **is simply PRE (1) below**.

$$\text{price}_i = \beta_0 + \beta_1\text{weight}_i + \beta_2\text{weight}_i^2 + \beta_3\text{mpg}_i + u_i, \quad i = 1, ..., N = N_d + N_f. \qquad \textbf{(1)}$$

- **Compute OLS estimates of the *restricted* model (1) on all N observations**.

The ***restricted*** **OLS SRE** can be written in general as

$$\text{price}_i = \tilde{\beta}_0 + \tilde{\beta}_1\text{weight}_i + \tilde{\beta}_2\text{weight}_i^2 + \tilde{\beta}_3\text{mpg}_i + \tilde{u}_i, \qquad i = 1, ..., N$$

and has RSS = RSS_0 with degrees of freedom $df_0 = N - K_0 = N - 4$.

Values: $RSS_0 = 372311797.0$ with $df_0 = N - K_0 = 74 - 4 = 70$.

❑ **Step 2: Estimate the *unrestricted* model implied by H_1.**

The ***unrestricted* model** is given by any one of the three **pooled full-interaction regression equations (2.0), (2.1), or (2.2)**:

$$\begin{aligned}\text{price}_i = &\ \beta_0\text{domestic}_i + \beta_1(\text{domestic}_i)(\text{weight}_i) + \beta_2(\text{domestic}_i)(\text{weight}_i^2) \\ &+ \beta_3(\text{domestic}_i)(\text{mpg}_i) + \alpha_0\text{foreign}_i + \alpha_1(\text{foreign}_i)(\text{weight}_i) \\ &+ \alpha_2(\text{foreign}_i)(\text{weight}_i^2) + \alpha_3(\text{foreign}_i)(\text{mpg}_i) + u_i\end{aligned} \quad \text{... (2.0)}$$

$$\begin{aligned}\text{price}_i = &\ \beta_0 + \beta_1\text{weight}_i + \beta_2\text{weight}_i^2 + \beta_3\text{mpg}_i \\ &+ \gamma_0\text{foreign}_i + \gamma_1(\text{foreign}_i)(\text{weight}_i) + \gamma_2(\text{foreign}_i)(\text{weight}_i^2) \\ &+ \gamma_3(\text{foreign}_i)(\text{mpg}_i) + u_i\end{aligned} \quad \text{... (2.1)}$$

$$\begin{aligned}\text{price}_i = &\ \alpha_0 + \alpha_1\text{weight}_i + \alpha_2\text{weight}_i^2 + \alpha_3\text{mpg}_i \\ &+ \delta_0\text{domestic}_i + \delta_1(\text{domestic}_i)(\text{weight}_i) + \delta_2(\text{domestic}_i)(\text{weight}_i^2) \\ &+ \delta_3(\text{domestic}_i)(\text{mpg}_i) + u_i\end{aligned} \quad \text{... (2.2)}$$

The regression equations (2.0), (2.1), and (2.2) are ***observationally equivalent***: OLS estimation of them yields the same values for RSS, $\hat{\sigma}^2$, R^2, $\bar{R}^2$, and so on.

- **Compute OLS estimates of any *one* of the *unrestricted* models (2.0), (2.1) and (2.2) on all N observations**.

Values: $RSS_1 = 213249239.0$ with $df_1 = N - K_1 = N - 2K_0 = 74 - 8 = 66$.

❑ **<u>Step 3</u>: Compute *sample value* of the F-test statistic.**

$$F = \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} \sim F[df_0 - df_1, df_1] \text{ under } H_0.$$

- $RSS_0 = 372311797.0$ with $df_0 = N - K_0 = 70.$
- $RSS_1 = 213249239.0$ with $df_1 = N - K_1 = N - 2K_0 = 74 - 8 = 66.$
- $df_0 - df_1 = 70 - 66 = 4.$

Thus, the ***sample value of the F-statistic*** under the null hypothesis H_0 is:

$$\begin{aligned} F_0 &= \frac{(RSS_0 - RSS_1)/(df_0 - df_1)}{RSS_1/df_1} \\ &= \frac{(372311797.0 - 213249239.0)/(70 - 66)}{213249239.0/66} \\ &= \frac{159062558.6/4}{213249239.0/66} \\ &= \frac{39765639.65}{3231049.076} \\ &= 12.307. \end{aligned}$$

<u>Result</u>: $\mathbf{F_0 = 12.31}$.

❑ **<u>Step 4</u>: Apply the usual *decision rule* for an F-test.**

3. Tests for Individual Coefficient Equality

In this section, we consider how to test for equality of a single regression coefficient. It is important to recognize that ***in testing for equality of any subset of coefficients, only the pooled regression approach can be used.***

Proposition: The marginal effect of fuel efficiency on price is the same for domestic and foreign cars.

Formulate this proposition as a ***testable hypothesis***.

- Fuel efficiency is measured here by the explanatory variable mpg_i.

- The **marginal effects of mpg_i for *domestic* and *foreign* cars** are:

$$\frac{\partial price_i}{\partial mpg_i} = \beta_3 \qquad \text{for } \textit{domestic} \text{ cars;}$$

$$\frac{\partial price_i}{\partial mpg_i} = \alpha_3 \qquad \text{for } \textit{foreign} \text{ cars.}$$

- The ***null and alternative hypotheses*** can be formulated as a single coefficient restriction on the pooled regression equations (2.0), (2.1), and (2.2).

For PRE (2.0): H_0: $\alpha_3 = \beta_3$ vs. H_1: $\alpha_3 \neq \beta_3$.
For PRE (2.1): H_0: $\gamma_3 = \alpha_3 - \beta_3 = 0$ vs. H_1: $\gamma_3 = \alpha_3 - \beta_3 \neq 0$.
For PRE (2.2): H_0: $\delta_3 = \beta_3 - \alpha_3 = 0$ vs. H_1: $\delta_3 = \beta_3 - \alpha_3 \neq 0$.

Test Procedure: Since the null hypothesis specifies only ***one* coefficient restriction**, a **two-tailed t-test procedure** can be used to test it.

- **Why use a t-test?** Because **a t-test can be performed using *only* OLS estimates of the *unrestricted* model** corresponding to the alternative hypothesis H_1.

- In fact, we can perform the required t-test using OLS estimates of ***any one*** of the three formulations of the unrestricted model given by equations (2.0), (2.1), and (2.2).

<u>Test 1</u>: A *<u>t-test</u>* based on OLS estimates of the pooled regression equation (2.1)

$$\begin{aligned}\text{price}_i &= \beta_0 + \beta_1\text{weight}_i + \beta_2\text{weight}_i^2 + \beta_3\text{mpg}_i \\ &\quad + \gamma_0\text{foreign}_i + \gamma_1(\text{foreign}_i)(\text{weight}_i) + \gamma_2(\text{foreign}_i)(\text{weight}_i^2) \\ &\quad + \gamma_3(\text{foreign}_i)(\text{mpg}_i) + u_i\end{aligned} \qquad \textbf{... (2.1)}$$

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

- The **testable hypothesis** consists of the null and alternative hypotheses

  H_0: $\gamma_3 = \alpha_3 - \beta_3 = 0$ vs. H_1: $\gamma_3 = \alpha_3 - \beta_3 \neq 0$.

- The required **t-statistic** for the OLS estimate $\hat{\gamma}_3$ of the coefficient γ_3 is

$$t(\hat{\gamma}_3) = \frac{\hat{\gamma}_3 - \gamma_3}{\hat{se}(\hat{\gamma}_3)} \sim t[N-k] = t[N-k_1] = t[66]. \qquad (3)$$

  where $\hat{se}(\hat{\gamma}_3) = \sqrt{\hat{Var}(\hat{\gamma}_3)}$ = the estimated standard error of $\hat{\gamma}_3$.

- Calculate the ***sample value* of the t-statistic** (3) under the null hypothesis H_0.

  Set $\gamma_3 = 0$, $\hat{\gamma}_3 = -100.152$, and $\hat{se}(\hat{\gamma}_3) = 141.238$ in t-statistic (3):

$$t_0(\hat{\gamma}_3) = \frac{\hat{\gamma}_3 - \gamma_3}{\hat{se}(\hat{\gamma}_3)} = \frac{-100.152 - 0}{141.238} = \frac{-100.152}{141.238} = \mathbf{-0.709}.$$

- ***Null distribution*** of $t_0(\hat{\gamma}_3)$ is $t[N-K] = t[74-8] = t[66]$:

$$t_0(\hat{\gamma}_3) \sim t[N-K_1] = t[N-K] = t[66] \text{ under } H_0.$$

- **Decision Rule -- Formulation 1:** At significance level α,

  - ***reject*** **H_0** if $|t_0(\hat{\gamma}_3)| > t_{\alpha/2}[66]$,
    i.e., if either (1) $t_0(\hat{\gamma}_3) > t_{\alpha/2}[66]$ or (2) $t_0(\hat{\gamma}_3) < -t_{\alpha/2}[66]$;

  - ***retain*** **H_0** if $|t_0(\hat{\gamma}_3)| \le t_{\alpha/2}[66]$, i.e., if $-t_{\alpha/2}[66] \le t_0(\hat{\gamma}_3) \le t_{\alpha/2}[66]$.

- **Two-tailed *critical values* of t[66]-distribution:**

  - two-tailed 5 percent critical value = $t_{\alpha/2}[66]$ = $t_{0.025}[66]$ = 1.997;
  - two-tailed 10 percent critical value = $t_{\alpha/2}[66]$ = $t_{0.05}[66]$ = 1.668.

- **Inference:**

  - ***At 5 percent significance level***, i.e., for α = 0.05,

    $|t_0(\hat{\gamma}_3)| = 0.709 < 1.997 = t_{0.025}[72] \Rightarrow$ ***retain*** **H_0 at *5* percent level**.

  - ***At 10 percent significance level***, i.e., for α = 0.10,

    $|t_0(\hat{\gamma}_3)| = 0.709 < 1.668 = t_{0.05}[66] \Rightarrow$ ***retain*** **H_0 at *10* percent level**.

<u>Test 2</u>: A <u>*t-test*</u> based on OLS estimates of the pooled regression equation (2.0)

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

... (2.0)

- The **testable hypothesis** consists of the null and alternative hypotheses

H_0: $\alpha_3 = \beta_3 \quad \Rightarrow \quad \alpha_3 - \beta_3 = 0$
H_1: $\alpha_3 \neq \beta_3 \quad \Rightarrow \quad \alpha_3 - \beta_3 \neq 0$.

- The required **t-test statistic** for testing H_0 against H_1 takes the form

$$t\left(\hat{\alpha}_3 - \hat{\beta}_3\right) = \frac{(\hat{\alpha}_3 - \hat{\beta}_3) - (\alpha_3 - \beta_3)}{\hat{se}(\hat{\alpha}_3 - \hat{\beta}_3)} \sim t[N-K] = t[N-K_1]. \qquad (4)$$

where $\hat{se}(\hat{\alpha}_3 - \hat{\beta}_3) = \sqrt{\hat{Var}(\hat{\alpha}_3 - \hat{\beta}_3)}$.

<u>*Note*</u>*:* The t-statistic (4) requires computation of the estimated variance of the difference in coefficient estimates $\hat{\alpha}_3 - \hat{\beta}_3$.

- ♦ The formula for the **estimated variance** of the linear combination $\hat{\alpha}_3 - \hat{\beta}_3$ is

$$\hat{Var}(\hat{\alpha}_3 - \hat{\beta}_3) = \hat{Var}(\hat{\alpha}_3) + \hat{Var}(\hat{\beta}_3) - 2\,\hat{Cov}(\hat{\alpha}_3, \hat{\beta}_3)$$

where

$\hat{Var}(\hat{\alpha}_3) \equiv$ the estimated variance of the coefficient estimate $\hat{\alpha}_3$;
$\hat{Var}(\hat{\beta}_3) \equiv$ the estimated variance of the coefficient estimate $\hat{\beta}_3$;
$\hat{Cov}(\hat{\alpha}_3, \beta_3) \equiv$ the estimated covariance of $\hat{\alpha}_3$ and $\hat{\beta}_3$.

- ♦ From OLS estimation of the pooled regression equation (2.0), **retrieve the following values** for $\hat{\alpha}_3$, $\hat{\beta}_3$, $\hat{Var}(\hat{\alpha}_3)$, $\hat{Var}(\hat{\beta}_3)$, and $\hat{Cov}(\hat{\alpha}_3, \beta_3)$:

$$\hat{\alpha}_3 = -22.07724;$$
$$\hat{\beta}_3 = 78.07451;$$
$$\hat{Var}(\hat{\alpha}_3) = 13348.87;$$
$$\hat{Var}(\hat{\beta}_3) = 6599.275;$$
$$\hat{Cov}(\hat{\alpha}_3, \beta_3) = -0.001003425.$$

- ♦ Calculate the **difference in coefficient estimates** $\hat{\alpha}_3 - \hat{\beta}_3$:

$$\hat{\alpha}_3 - \hat{\beta}_3 = -22.07724 - 78.07451 = -100.15175.$$

- ♦ Calculate the **estimated variance of the coefficient difference** $\hat{\alpha}_3 - \hat{\beta}_3$ is

$$\begin{aligned}\hat{Var}(\hat{\alpha}_3 - \hat{\beta}_3) &= \hat{Var}(\hat{\alpha}_3) + \hat{Var}(\hat{\beta}_3) - 2\hat{Cov}(\hat{\alpha}_3, \hat{\beta}_3) \\ &= 13348.87 + 6599.275 - 2(-0.001003425) \\ &= 13348.87 + 6599.275 + 0.00200685 \\ &= 19948.147\end{aligned}$$

- ♦ The **estimated standard error of the coefficient difference** $\hat{\alpha}_3 - \hat{\beta}_3$ is therefore

$$\hat{se}(\hat{\alpha}_3 - \hat{\beta}_3) = \sqrt{\hat{Var}(\hat{\alpha}_3 - \hat{\beta}_3)} = \sqrt{19948.147} = 141.2379.$$

- Calculate the ***sample value* of the t-statistic** (4) under the null hypothesis H_0.

Set $\hat{\alpha}_3 - \hat{\beta}_3 = -100.15175$, $\alpha_3 - \beta_3 = 0$, and $\hat{se}(\hat{\alpha}_3 - \hat{\beta}_3) = 141.2379$ in (4):

$$t_0\left(\hat{\alpha}_3 - \hat{\beta}_3\right) = \frac{(\hat{\alpha}_3 - \hat{\beta}_3) - (\alpha_3 - \beta_3)}{\hat{se}(\hat{\alpha}_3 - \hat{\beta}_3)} = \frac{-100.15175 - 0}{141.2379} = \frac{-100.15175}{141.2379} = \mathbf{-0.709}.$$

- ***Null distribution*** of $t_0(\hat{\alpha}_3 - \hat{\beta}_3)$ is t[N–K] = t[74–8] = t[66]:

$$t_0(\hat{\alpha}_3 - \hat{\beta}_3) \sim t[N - K_1] = t[N - K] = t[66] \text{ under } H_0.$$

- **Decision Rule -- Formulation 1:** At significance level α,

  - ***reject*** $\mathbf{H_0}$ if $\left|t_0(\hat{\alpha}_3 - \hat{\beta}_3)\right| > t_{\alpha/2}[66]$,
    i.e., if either (1) $t_0(\hat{\alpha}_3 - \hat{\beta}_3) > t_{\alpha/2}[66]$ or (2) $t_0(\hat{\alpha}_3 - \hat{\beta}_3) < -t_{\alpha/2}[66]$;

  - ***retain*** $\mathbf{H_0}$ if $\left|t_0(\hat{\alpha}_3 - \hat{\beta}_3)\right| \le t_{\alpha/2}[66]$,
    i.e., if $-t_{\alpha/2}[66] \le t_0(\hat{\alpha}_3 - \hat{\beta}_3) \le t_{\alpha/2}[66]$.

- **Two-tailed *critical values* of t[66]-distribution:**

  - two-tailed 5 percent critical value = $t_{\alpha/2}[66]$ = $t_{0.025}[66]$ = 1.997;
  - two-tailed 10 percent critical value = $t_{\alpha/2}[66]$ = $t_{0.05}[66]$ = 1.668.

- **Inference:**

  - ***At 5 percent significance level***, i.e., for α = 0.05,

    $\left|t_0(\hat{\alpha}_3 - \hat{\beta}_3)\right|$ = 0.709 < 1.997 = $t_{0.025}[72]$ ⇒ ***retain*** $\mathbf{H_0}$ **at *5* percent level**.

  - ***At 10 percent significance level***, i.e., for α = 0.10,

    $\left|t_0(\hat{\alpha}_3 - \hat{\beta}_3)\right|$ = 0.709 < 1.668 = $t_{0.05}[66]$ ⇒ ***retain*** $\mathbf{H_0}$ **at *10* percent level**.

- <u>**Result**</u>**:** ***Test 2***, the t-test based on pooled regression equation (2.0), yields the same sample value of the t-statistic and hence the same inferences as ***Test 1***, the t-test based on pooled regression equation (2.1). This means that ***Test 1* and *Test 2* are *equivalent* tests** of the null hypothesis H_0 that the marginal effect of gasoline mileage (i.e., of mpg_i) on car prices is the same for domestic and foreign cars.