ECON 452* -- NOTE 3

Marginal Effects of Continuous Explanatory Variables: Constant or Variable?

❑ Constant Marginal Effects of Explanatory Variables: A Starting Point

Nature: A continuous explanatory variable has a ***constant*** **marginal effect** on the dependent variable if it enters the regressor set only linearly and additively.

Model 1:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i \qquad \textbf{(1)}$$

- Model 1 contains only **two** ***explanatory variables*** -- X_1 and X_2 -- and **two** ***regressors***.

- The population regression function, or conditional mean function, $f(X_{i1}, X_{i2})$ in Model 1 takes the form

$$E(Y_i \mid X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}\,.$$

***Model 1*:**

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i \qquad \textbf{(1)}$$

$$E(Y_i \mid X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}.$$

- The ***marginal* effects on Y of the two explanatory variables X_1 and X_2** in equation (1) are obtained analytically by partially differentiating Y, or the conditional mean of Y given X_1 and X_2, with respect to each of the explanatory variables X_1 and X_2.

1. The **marginal effect of X_1** in Model 1 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 = \text{a } \textit{constant}$$

2. The **marginal effect of X_2** in Model 1 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 = \text{a } \textit{constant}$$

Example of Model 1:

$$price_i = \beta_0 + \beta_1 wgt_i + \beta_2 mpg_i + u_i$$

where

$price_i$ = the price of the i-th car (in US dollars);
wgt_i = the weight of the i-th car (in pounds);
mpg_i = the miles per gallon (fuel efficiency) for the i-th car (in miles per gallon).

❑ Variable Marginal Effects and Interaction Terms: Squares and Cross Products of Continuous Explanatory Variables

Nature: ***Interactions* between two *continuous* variables** refer to products of pairs of explanatory variables.

- If X_{ij} and X_{ih} are two continuous explanatory variables, the interaction term between them is the product $X_{ij}X_{ih}$.
- The interaction of the variable X_{ij} with itself is simply the product $X_{ij}X_{ij} = X_{ij}^2$.
- Inclusion of these regressors in a linear regression model allows for ***variable* -- or *nonconstant* -- marginal effects** of the explanatory variables on the conditional mean of the dependent variable Y.

Usage: ***Interaction terms*** between ***continuous* variables** allow the marginal effect of one explanatory variable to be a linear function of both itself and other explanatory variables.

Model 2:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + u_i \qquad \textbf{(2)}$$

- Model 2 contains only **two** ***explanatory variables*** -- X_1 and X_2 -- but **five** ***regressors***.

- Formally, the population regression function $E(Y_i \,|\, X_{i1}, X_{i2}) = f(X_{i1}, X_{i2})$ in PRE (2) can be derived as a second-order Taylor series approximation to the function $f(X_{i1}, X_{i2})$. A second-order Taylor series approximation to the population regression function $f(X_{i1}, X_{i2})$ takes the form

$$E(Y_i \,|\, X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2}\,.$$

- The ***marginal*** **effects on Y of the two explanatory variables X_1 and X_2** in equation (2) are obtained analytically by partially differentiating Y, or the conditional mean of Y given X_1 and X_2, with respect to each of the explanatory variables X_1 and X_2.

1. The **marginal effect of X_1** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

= a linear function of *both* X_{i1} and X_{i2}

2. The **marginal effect of X_2** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

= a linear function of *both* X_{i1} and X_{i2}

<u>Squares of Continuous Explanatory Variables</u>

Purpose: Allow for ***increasing* or *decreasing* marginal effects** of an explanatory variable on the dependent variable -- sometimes called ***increasing* or *decreasing* marginal returns**.

Determining whether the *marginal* effect of X_1 is *increasing* or *decreasing*

- Whether the ***marginal* effect of X_1 is *increasing* or *decreasing*** -- i.e., whether X_1 exhibits ***increasing* or *decreasing* marginal returns --** is determined by the sign of the regression coefficient β_3 on the regressor X_{i1}^2 in Model 2.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + u_i \qquad \textbf{(2)}$$

- We previously saw that the ***marginal* effect of X_1 in Model 2** is given by the ***first-order* partial derivative** of Y_i, or $E(Y_i \,|\, X_{i1}, X_{i2})$, with respect to X_{i1}:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

- To determine whether the ***marginal* effect of X_1 in Model 2** is ***increasing* or *decreasing* in X_1**, we need to examine the ***second-order* partial derivative** of Y_i, or $E(Y_i \,|\, X_{i1}, X_{i2})$, with respect to X_{i1}:

$$\frac{\partial^2 Y_i}{\partial X_{i1}^2} = \frac{\partial}{\partial X_{i1}} \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \frac{\partial^2 E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial^2 X_{i1}^2} = 2\beta_3$$

$$\frac{\partial^2 Y_i}{\partial X_{i1}^2} = \frac{\partial}{\partial X_{i1}} \frac{\partial E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i1}} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i1}^2} = 2\beta_3$$

1. The ***marginal*** **effect of X_1** is ***increasing*** **in X_1** -- meaning X_1 exhibits ***increasing*** **marginal returns** -- when

$$\frac{\partial^2 Y_i}{\partial X_{i1}^2} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i1}^2} = 2\beta_3 > 0 \qquad \text{i.e., when } \beta_3 > 0$$

2. The ***marginal*** **effect of X_1** is ***decreasing*** **in X_1** -- meaning X_1 exhibits ***decreasing*** **marginal returns** -- when

$$\frac{\partial^2 Y_i}{\partial X_{i1}^2} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i1}^2} = 2\beta_3 < 0 \qquad \text{i.e., when } \beta_3 < 0$$

Determining whether the *marginal* effect of X_2 is *increasing* or *decreasing*

- Whether the ***marginal* effect of X_2 is *increasing* or *decreasing* --** i.e., whether X_2 exhibits ***increasing* or *decreasing* marginal returns --** is determined by the sign of the regression coefficient β_4 on the regressor X_{i2}^2 in Model 2.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + u_i \qquad \textbf{(2)}$$

- We previously saw that the ***marginal* effect of X_2 in Model 2** is given by the ***first-order* partial derivative** of Y_i, or $E(Y_i \,|\, X_{i1}, X_{i2})$, with respect to X_{i2}:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

- To determine whether the ***marginal* effect of X_2 in Model 2** is ***increasing* or *decreasing* in X_2**, we need to examine the ***second-order* partial derivative** of Y_i, or $E(Y_i \,|\, X_{i1}, X_{i2})$, with respect to X_{i2}:

$$\frac{\partial^2 Y_i}{\partial X_{i2}^2} = \frac{\partial}{\partial X_{i2}} \frac{\partial E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \frac{\partial^2 E\left(Y_i \,\middle|\, X_{i1}, X_{i2}\right)}{\partial^2 X_{i2}^2} = 2\beta_4$$

$$\frac{\partial^2 Y_i}{\partial X_{i2}^2} = \frac{\partial}{\partial X_{i2}} \frac{\partial E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i2}} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i2}^2} = 2\beta_4$$

1. The ***marginal*** **effect of X_2** is ***increasing*** **in X_2** -- meaning X_2 exhibits ***increasing*** **marginal returns** -- when

$$\frac{\partial^2 Y_i}{\partial X_{i2}^2} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i2}^2} = 2\beta_4 > 0 \quad \text{i.e., when } \beta_4 > 0$$

2. The ***marginal*** **effect of X_2** is ***decreasing*** **in X_2** -- meaning X_2 exhibits ***decreasing*** **marginal returns --** when

$$\frac{\partial^2 Y_i}{\partial X_{i2}^2} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial^2 X_{i2}^2} = 2\beta_4 < 0 \quad \text{i.e., when } \beta_4 < 0$$

Products of Two Continuous Explanatory Variables

Purpose: Allow for relationships of ***complementarity*** **or** ***substitutability*** **between X_1 and X_2** in determining Y.

- Whether **X_1 and X_2 are** ***complementary*** **or** ***substitutable*** is determined by the sign of the regression coefficient β_5 on the interaction term $X_{i1}X_{i2}$ in Model 2.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + u_i \qquad \textbf{(2)}$$

- We previously saw that the ***marginal*** **effect of X_1 in Model 2** is given by the ***first-order*** **partial derivative** of Y_i, or $E(Y_i \mid X_{i1}, X_{i2})$, with respect to X_{i1}:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

- To determine whether the ***marginal* effect of X_1 (X_2) in Model 2** is ***increasing* or *decreasing* in X_2 (X_1)**, we need to examine the ***second-order* cross partial derivative** of Y_i, or $E(Y_i \mid X_{i1}, X_{i2})$, with respect to X_{i1} and X_{i2}:

$$\frac{\partial^2 Y_i}{\partial X_{i2} \partial X_{i1}} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i2} \partial X_{i1}} = \beta_5$$

1. The ***marginal* effect of X_1** is ***increasing* in X_2** (or the ***marginal* effect of X_2** is ***increasing* in X_1**) **--** meaning **X_1 and X_2 are *complementary*** -- when

$$\frac{\partial^2 Y_i}{\partial X_{i2} \partial X_{i1}} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i2} \partial X_{i1}} = \beta_5 > 0$$

2. The ***marginal* effect of X_1** is ***decreasing* in X_2** (or the ***marginal* effect of X_2** is ***decreasing* in X_1**) **--** meaning **X_1 and X_2 are *substitutable*** -- when

$$\frac{\partial^2 Y_i}{\partial X_{i2} \partial X_{i1}} = \frac{\partial^2 E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i2} \partial X_{i1}} = \beta_5 < 0$$

Example of Model 2:

$$\text{price}_i = \beta_0 + \beta_1 \text{wgt}_i + \beta_2 \text{mpg}_i + \beta_3 \text{wgt}_i^2 + \beta_4 \text{mpg}_i^2 + \beta_5 \text{wgt}_i \text{mpg}_i + u_i .$$

where

price_i = the price of the i-th car (in US dollars);
wgt_i = the weight of the i-th car (in pounds);
wgt_i^2 = the square of wgt_i;
mpg_i = the miles per gallon (fuel efficiency) for the i-th car (in miles per gallon);
mpg_i^2 = the square of mpg_i;
wgt_impg_i = the product of wgt_i and mpg_i for the i-th car.

❑ Review: Models 1 and 2 in Two *Continuous* Explanatory Variables X_1 and X_2

***Model 1*: *Constant* marginal effects** of the explanatory variables X_1 and X_2

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + u_i \qquad \textbf{(1)}$$

$$E(Y_i \mid X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2}$$

- Model 1 contains only **two** ***continuous explanatory variables*** -- X_1 and X_2 -- and **two** ***regressors***.

- The ***marginal*** **effects on Y of the explanatory variables X_1 and X_2** are:

  1. The **marginal effect of X_1** in Model 1 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i1}} = \beta_1 = \text{a } \textit{constant}$$

  2. The **marginal effect of X_2** in Model 1 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i2}} = \beta_2 = \text{a } \textit{constant}$$

Model 2:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + u_i \qquad \textbf{(2)}$$

$$E(Y_i \,|\, X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2}$$

- Model 2 contains only **two** ***continuous explanatory variables*** -- X_1 and X_2 -- but **five** ***regressors***.
- The ***marginal*** **effects on Y of the two explanatory variables X_1 and X_2** are:

1. The **marginal effect of X_1** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \,|\, X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$ = a linear function of *both* X_{i1} and X_{i2}

2. The **marginal effect of X_2** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \,|\, X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$ = a linear function of *both* X_{i1} and X_{i2}

<u>Example of Model 1</u>:

$$price_i = \beta_0 + \beta_1 wgt_i + \beta_2 mpg_i + u_i$$

where

$price_i$ = the price of the i-th car (in US dollars)
wgt_i = the weight of the i-th car (in pounds)
mpg_i = the fuel efficiency of the i-th car (in miles per gallon)

<u>Example of Model 2</u>:

$$price_i = \beta_0 + \beta_1 wgt_i + \beta_2 mpg_i + \beta_3 wgt_i^2 + \beta_4 mpg_i^2 + \beta_5 wgt_i mpg_i + u_i$$

where

$price_i$ = the price of the i-th car (in US dollars)
wgt_i = the weight of the i-th car (in pounds)
wgt_i^2 = the square of wgt_i for the i-th car
mpg_i = the fuel efficiency of the i-th car (in miles per gallon)
mpg_i^2 = the square of mpg_i for the i-th car
$wgt_i mpg_i$ = the product (interaction) of wgt_i and mpg_i for the i-th car

❑ Hypothesis Tests on Model 2

♦ ***Test 1*:** Test the hypothesis that the ***marginal* effect of X_1 on Y in Model 2 is *zero*** for all values of X_1 and X_2.

- The **marginal effect of X_1** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = 0$ for all i are $\beta_1 = 0$ *and* $\beta_3 = 0$ *and* $\beta_5 = 0$. We therefore want to test these three coefficient exclusion restrictions.

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

  H_0: $\beta_1 = 0$ *and* $\beta_3 = 0$ *and* $\beta_5 = 0$

  H_1: $\beta_1 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_5 \neq 0$

- First, estimate Model 2 by OLS using the following *Stata* **regress** command:

```
regress y x1 x2 x1sq x2sq x1x2
```

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x1 x1sq x1x2
```

♦ ***Test 2:*** Test the hypothesis that the ***marginal*** **effect of X_1 on Y in Model 2 is *constant***.

- The **marginal effect of X_1** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E(Y_i | X_{i1}, X_{i2})}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

$$= \beta_1 \text{ (a constant)} \qquad \text{if } \beta_3 = 0 \text{ and } \beta_5 = 0$$

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = \beta_1$ (a constant) for all i are $\beta_3 = 0$ *and* $\beta_5 = 0$. We therefore want to test these two coefficient exclusion restrictions.

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

H_0: $\beta_3 = 0$ *and* $\beta_5 = 0$

H_1: $\beta_3 \neq 0$ *and/or* $\beta_5 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x1sq x1x2
```

- ***Test 3:*** Test the hypothesis that the ***marginal* effect of X_2 on Y in Model 2 is *zero*** for all values of X_1 and X_2.

- The **marginal effect of X_2** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

- Sufficient conditions for $\partial Y_i / \partial X_{i2} = 0$ for all i are $\beta_2 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$. We therefore want to test these three coefficient restrictions.

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

H_0: $\beta_2 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$

H_1: $\beta_2 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x2 x2sq x1x2
```

- ***Test 4:*** Test the hypothesis that the ***marginal*** **effect of X_2 on Y in Model 2 is *constant***.

- The **marginal effect of X_2** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

$$= \beta_2 \text{ (a constant)} \qquad \text{if } \beta_4 = 0 \text{ and } \beta_5 = 0$$

- Sufficient conditions for $\partial Y_i / \partial X_{i2} = \beta_2$ (a constant) for all i are $\beta_4 = 0$ *and* $\beta_5 = 0$. We therefore want to test these two coefficient exclusion restrictions.

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

  H_0: $\beta_4 = 0$ *and* $\beta_5 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x2sq x1x2
```

- ***Test 5:*** Test the hypothesis that ***both*** the **marginal effect of X_1** on Y ***and*** the ***marginal effect of X_2*** on Y **are *constants***.

  This is equivalent to testing whether the three additional regressors that Model 2 introduces into Model 1 – namely X_{i1}^2, X_{i2}^2 and $X_{i1}X_{i2}$ – are necessary in order to adequately represent the relationship of Y_i to the two continuous explanatory variables X_{i1} and X_{i2}.

- The **marginal effects of X_1 and X_2** in Model 2 are:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \middle| X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

$$= \beta_1 \qquad \text{if} \quad \beta_3 = 0 \text{ and } \beta_5 = 0$$

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \middle| X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

$$= \beta_2 \qquad \text{if} \quad \beta_4 = 0 \text{ and } \beta_5 = 0$$

- Sufficient conditions for both marginal effects to be constants for all i are $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$. We therefore want to test jointly these three coefficient restrictions.

- ***Test 5 (continued):*** Test the hypothesis that ***both*** the **marginal effect of X_1** on Y ***and*** the ***marginal effect of X_2*** on Y **are *constants***.

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

  H_0: $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_5 = 0$ ⇒ Model 1

  H_1: $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_5 \neq 0$ ⇒ Model 2

  Note that the null hypothesis H_0 implies that Model 1 is empirically adequate, whereas the alternative hypothesis H_1 implies rejection of Model 1 in favor of Model 2.

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x1sq x2sq x1x2
```

- ***Test 6:*** Test the hypothesis that the **marginal effect of X_1** on Y ***is unrelated to***, or ***does not depend upon***, **X_1**.

- The **marginal effect of X_1** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E(Y_i \mid X_{i1}, X_{i2})}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

- A sufficient condition for $\partial Y_i / \partial X_{i1}$ to be unrelated to X_{i1} for all i is $\beta_3 = 0$. We therefore want to test this coefficient exclusion restriction on Model 2.

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

H_0: $\beta_3 = 0$

H_1: $\beta_3 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

test x1sq *or* **test x1sq = 0**

- Equivalently, compute a **two-tail t-test** of H_0 against H_1 using the following *Stata* **lincom** command:

```
lincom _b[x1sq]
```

- ***Test 7:*** Test the hypothesis that the **marginal effect of X_2** on Y ***is unrelated to***, or ***does not depend upon***, **X_2**.

- The **marginal effect of X_2** in Model 2 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \middle| X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

- A sufficient condition for $\partial Y_i / \partial X_{i2}$ to be unrelated to X_{i2} for all i is $\beta_4 = 0$. We therefore want to test this coefficient exclusion restriction on Model 2.

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

H_0: $\beta_4 = 0$

H_1: $\beta_4 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

```
test x2sq
```

or

```
test x2sq = 0
```

- Equivalently, compute a **two-tail t-test** of H_0 against H_1 using the following *Stata* **lincom** command:

```
lincom _b[x2sq]
```

- ***Test 8:*** Consider the following two propositions:

1. The ***marginal*** **effect of X_1 on Y is unrelated to X_2** in Model 2.
2. The ***marginal*** **effect of X_2 on Y is unrelated to X_1** in Model 2.

- Recall that the marginal effects on Y of X_1 and X_2 in Model 2 are:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2}$$

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \mid X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1}$$

Propositions 1 and 2 imply the same coefficient exclusion restriction on Model 2, namely the restriction $\beta_5 = 0$.

1. The ***marginal*** **effect of X_1 on Y is unrelated to X_2** if $\beta_5 = 0$ in Model 2.
2. The ***marginal*** **effect of X_2 on Y is unrelated to X_1** if $\beta_5 = 0$ in Model 2.

- The ***null*** **and** ***alternative*** **hypotheses** for both these propositions are:

H_0: $\beta_5 = 0$

H_1: $\beta_5 \neq 0$

♦ ***<u>Test 8 (continued)</u>:*** Consider the following two propositions:

H_0: $\beta_5 = 0$

H_1: $\beta_5 \neq 0$

- Compute an **F-test** of H_0 against H_1 using the following *Stata* **test** command:

`test x1x2` *or* **`test x1x2 = 0`**

- Equivalently, compute a **two-tail t-test** of H_0 against H_1 using the following *Stata* **lincom** command:

```
lincom _b[x1x2]
```

❑ Determining the Order of Polynomial for a Continuous Explanatory Variable: Model 3

In practice, it is often unclear what order of polynomial function in a continuous explanatory variable is required to adequately represent the conditional effect of that explanatory variable on the regressand Y.

Model 3:

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

$$E(Y_i \mid X_{i1}, X_{i2}) = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4$$

- Model 3 contains only **two** ***continuous explanatory variables*** -- X_1 and X_2 -- but **nine** ***regressors***.

- The ***marginal*** **effects on Y of the two explanatory variables X_1 and X_2** in Model 3 are:

1. The **marginal effect of X_1** in Model 3 is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \middle| X_{i1}, X_{i2}\right)}{\partial X_{i1}} = \beta_1 + 2\beta_3 X_{i1} + \beta_5 X_{i2} + 3\beta_6 X_{i1}^2 + 4\beta_8 X_{i1}^3$$

= a cubic function of X_{i1} and linear function of X_{i2}

2. The **marginal effect of X_2** in Model 3 is:

$$\frac{\partial Y_i}{\partial X_{i2}} = \frac{\partial E\left(Y_i \middle| X_{i1}, X_{i2}\right)}{\partial X_{i2}} = \beta_2 + 2\beta_4 X_{i2} + \beta_5 X_{i1} + 3\beta_7 X_{i2}^2 + 4\beta_9 X_{i2}^3$$

= a cubic function of X_{i2} and linear function of X_{i1}

- The ***conditional*** **effects on Y of the two explanatory variables X_1 and X_2** in Model 3:

1. The **conditional effect of X_1** in Model 3 is the partial, or *ceteris paribus*, relationship between X_1 and the conditional mean value of Y for any given value of the other explanatory variable X_2.

Let $\bar{X}_2$ = the sample mean value of the explanatory variable X_2. Then the **conditional effect of X_1** in Model 3 is:

$$E(Y_i \mid X_{i1}, \bar{X}_2) = \beta_0 + \beta_1 X_{i1} + \beta_2 \bar{X}_2 + \beta_3 X_{i1}^2 + \beta_4 \bar{X}_2^2 + \beta_5 X_{i1}\bar{X}_2 + \beta_6 X_{i1}^3 + \beta_7 \bar{X}_2^3 + \beta_8 X_{i1}^4 + \beta_9 \bar{X}_2^4$$

$$= (\beta_0 + \beta_2 \bar{X}_2 + \beta_4 \bar{X}_2^2 + \beta_7 \bar{X}_2^3 + \beta_9 \bar{X}_2^4) + (\beta_1 + \beta_5 \bar{X}_2) X_{i1} + \beta_3 X_{i1}^2 + \beta_6 X_{i1}^3 + \beta_8 X_{i1}^4$$

2. The **conditional effect of X_2** in Model 3 is the partial, or *ceteris paribus*, relationship between X_2 and the conditional mean value of Y for any given value of the other explanatory variable X_1.

Let $\bar{X}_1$ = the sample mean value of the explanatory variable X_1. Then the **conditional effect of X_2** in Model 3 is:

$$E(Y_i \mid \bar{X}_1, X_{i2}) = \beta_0 + \beta_1 \bar{X}_1 + \beta_2 X_{i2} + \beta_3 \bar{X}_1^2 + \beta_4 X_{i2}^2 + \beta_5 \bar{X}_1 X_{i2} + \beta_6 \bar{X}_1^3 + \beta_7 X_{i2}^3 + \beta_8 \bar{X}_1^4 + \beta_9 X_{i2}^4$$

$$= (\beta_0 + \beta_1 \bar{X}_1 + \beta_3 \bar{X}_1^2 + \beta_6 \bar{X}_1^3 + \beta_8 \bar{X}_1^4) + (\beta_2 + \beta_5 \bar{X}_1) X_{i2} + \beta_4 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_9 X_{i2}^4$$

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \qquad \textbf{(3)}$$

```
regress y x1 x2 x1sq x2sq x1x2 x13rd x23rd x14th x24th
```

Determining the Order of Polynomial for the Continuous Explanatory Variable X_1

Perform the following sequence of hypothesis tests on the OLS estimates of regression equation (3), in particular, on the slope coefficients of the polynomial terms in the explanatory variable X_1.

- **Test 1.1:** Determine whether a ***third-order*** **polynomial** is adequate for representing the conditional effect on Y of X_1.

  Perform the following two-tail t-test or F-test:

  H_0: $\beta_8 = 0$

  H_1: $\beta_8 \neq 0$

  Stata commands:

  ```
  lincom _b[x14th]
  test x14th = 0    or    test x14th
  ```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **fourth-order polynomial** for the conditional effect of X_1 on Y.
  - **If H_0 is *retained***, proceed to Test 1.2 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \qquad \textbf{(3)}$$

- **Test 1.2:** Determine whether a ***second-order*** **polynomial** is adequate for representing the conditional effect on Y of X_1.

  Perform the following joint F-test:

  H_0: $\beta_8 = 0$ *and* $\beta_6 = 0$

  H_1: $\beta_8 \neq 0$ *and/or* $\beta_6 \neq 0$

  Stata commands:

```
test x14th = 0, notest
test x13rd = 0, accumulate
```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **third-order polynomial** for the conditional effect of X_1 on Y. The one coefficient exclusion restriction $\beta_8 = 0$ is a candidate for imposition in obtaining a simplified regression model.
  - **If H_0 is *retained***, proceed to Test 1.3 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

- **<u>Test 1.3</u>:** Determine whether a ***first-order* polynomial** is adequate for representing the conditional effect on Y of X_1.

  Perform the following joint F-test:

  H_0: $\beta_8 = 0$ *and* $\beta_6 = 0$ *and* $\beta_3 = 0$

  H_1: $\beta_8 \neq 0$ *and/or* $\beta_6 \neq 0$ *and/or* $\beta_3 \neq 0$

  Stata commands:

```
test x14th = 0, notest
test x13rd = 0, notest accumulate
test x1sq = 0, accumulate
```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **second-order polynomial** for the conditional effect of X_1 on Y. The two coefficient exclusion restrictions $\beta_8 = 0$ and $\beta_6 = 0$ are candidates for imposition in obtaining a simplified regression model.

  - **If H_0 is *retained***, proceed to Test 1.4 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

- **Test 1.4:** Determine whether a ***zero-order*** **polynomial** is adequate for representing the conditional effect on Y of X_1.

  Perform the following joint F-test:

  H_0: $\beta_8 = 0$ *and* $\beta_6 = 0$ *and* $\beta_3 = 0$ *and* $\beta_1 = 0$

  H_1: $\beta_8 \neq 0$ *and/or* $\beta_6 \neq 0$ *and/or* $\beta_3 \neq 0$ *and* $\beta_1 \neq 0$

  Stata commands:

```
test x14th = 0, notest
test x13rd = 0, notest accumulate
test x1sq = 0, notest accumulate
test x1 = 0, accumulate
```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **first-order polynomial** for the conditional effect of X_1 on Y. The three coefficient exclusion restrictions $\beta_8 = 0$, $\beta_6 = 0$ and $\beta_3 = 0$ are candidates for imposition in obtaining a simplified regression model.

  - **If H_0 is *retained***, stop testing. The four coefficient exclusion restrictions $\beta_8 = 0$, $\beta_6 = 0$, $\beta_3 = 0$ and $\beta_1 = 0$ are candidates for imposition in obtaining a simplified regression model.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \qquad \textbf{(3)}$$

Determining the Order of Polynomial for the Continuous Explanatory Variable X_2

Perform the following sequence of hypothesis tests on the OLS estimates of regression equation (3), in particular, on the slope coefficients of the polynomial terms in the explanatory variable X_2.

- **Test 2.1:** Determine whether a ***third-order*** **polynomial** is adequate for representing the conditional effect on Y of X_2.

  Perform the following two-tail t-test or F-test:

  H_0: $\beta_9 = 0$

  H_1: $\beta_9 \neq 0$

  Stata command:

  ```
  lincom _b[x24th]
  test x24th
  ```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **fourth-order polynomial** for the conditional effect of X_2 on Y.

  - **If H_0 is *retained***, proceed to Test 2.2 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

- **Test 2.2:** Determine whether a ***second-order*** **polynomial** is adequate for representing the conditional effect on Y of X_2.

  Perform the following joint F-test:

  H_0: $\beta_9 = 0$ *and* $\beta_7 = 0$

  H_1: $\beta_9 \neq 0$ *and/or* $\beta_7 \neq 0$

  Stata command:

  ```
  test x24th x23rd
  ```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **third-order polynomial** for the conditional effect of X_2 on Y. The one coefficient exclusion restriction $\beta_9 = 0$ is a candidate for imposition in obtaining a simplified regression model.

  - **If H_0 is *retained***, proceed to Test 2.3 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

- **<u>Test 2.3</u>:** Determine whether a ***first-order* polynomial** is adequate for representing the conditional effect on Y of X_2.

  Perform the following joint F-test:

  H_0: $\beta_9 = 0$ *and* $\beta_7 = 0$ *and* $\beta_4 = 0$

  H_1: $\beta_9 \neq 0$ *and/or* $\beta_7 \neq 0$ *and/or* $\beta_4 \neq 0$

  Stata command:

  ```
  test x24th x23rd x2sq
  ```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **second-order polynomial** for the conditional effect of X_2 on Y. The two coefficient exclusion restrictions $\beta_9 = 0$ and $\beta_7 = 0$ are candidates for imposition in obtaining a simplified regression model.

  - **If H_0 is *retained***, proceed to Test 2.4 below.

$$Y_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \beta_3 X_{i1}^2 + \beta_4 X_{i2}^2 + \beta_5 X_{i1} X_{i2} + \beta_6 X_{i1}^3 + \beta_7 X_{i2}^3 + \beta_8 X_{i1}^4 + \beta_9 X_{i2}^4 + u_i \quad \textbf{(3)}$$

- **Test 2.4:** Determine whether a ***zero-order*** **polynomial** is adequate for representing the conditional effect on Y of X_2.

  Perform the following joint F-test:

  H_0: $\beta_9 = 0$ *and* $\beta_7 = 0$ *and* $\beta_4 = 0$ *and* $\beta_2 = 0$

  H_1: $\beta_9 \neq 0$ *and/or* $\beta_7 \neq 0$ *and/or* $\beta_4 \neq 0$ *and* $\beta_2 \neq 0$

  Stata command:

```
test x24th x23rd x2sq x2
```

  Outcomes and Implied Decisions:

  - **If H_0 is *rejected***, stop testing. Choose a **first-order polynomial** for the conditional effect of X_2 on Y. The three coefficient exclusion restrictions $\beta_9 = 0$, $\beta_7 = 0$ and $\beta_4 = 0$ are candidates for imposition in obtaining a simplified regression model.

  - **If H_0 is *retained***, stop testing. The four coefficient exclusion restrictions $\beta_9 = 0$, $\beta_7 = 0$, $\beta_4 = 0$ and $\beta_2 = 0$ are candidates for imposition in obtaining a simplified regression model.

❑ Determining the Order of Polynomial for a Continuous Explanatory Variable: An Example of Model 3

Consider the following linear regression model for the average hourly wage rates of a cross-sectional sample of 252 female employees in the United States in 1976.

***Model 3*:**

$$\ln w_i = \beta_0 + \beta_1 ed_i + \beta_2 \exp_i + \beta_3 ed_i^2 + \beta_4 \exp_i^2 + \beta_5 ed_i \exp_i + \beta_6 ed_i^3 + \beta_7 \exp_i^3 + \beta_8 ed_i^4 + \beta_9 \exp_i^4 + u_i \qquad \textbf{(3)}$$

$$E(\ln w_i \mid ed_i, \exp_i) = \beta_0 + \beta_1 ed_i + \beta_2 \exp_i + \beta_3 ed_i^2 + \beta_4 \exp_i^2 + \beta_5 ed_i \exp_i + \beta_6 ed_i^3 + \beta_7 \exp_i^3 + \beta_8 ed_i^4 + \beta_9 \exp_i^4$$

The observable variables in Model 3 are defined as follows:

$\ln w_i$ = the natural logarithm of w_i, where is the average hourly wage rate of worker i in 1976;

ed_i = the years of completed formal education for worker i;

$\exp_i$ = the years of work experience accumulated by worker i.

- **OLS estimation of Model 3 in *Stata*:**

```
. *
. * Model 3 for Females
. *
. regress lnw ed exp edsq expsq edexp ed3rd exp3rd ed4th exp4th if female == 1

      Source |       SS       df       MS              Number of obs =     252
-------------+------------------------------           F(  9,   242) =   12.38
       Model |  15.6135128     9  1.73483476           Prob > F      =  0.0000
    Residual |  33.9200943   242  .140165679           R-squared     =  0.3152
-------------+------------------------------           Adj R-squared =  0.2897
       Total |  49.5336071   251  .197345048           Root MSE      =  .37439

------------------------------------------------------------------------------
         lnw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ed |  -.0415939   .1858768    -0.22   0.823    -.4077367     .324549
         exp |   .1028106   .0321942     3.19   0.002      .039394    .1662271
        edsq |   .0052875   .0408062     0.13   0.897    -.0750932    .0856682
       expsq |  -.0076544   .0027291    -2.80   0.005    -.0130303   -.0022786
       edexp |   .0002434   .0009553     0.25   0.799    -.0016383     .002125
       ed3rd |  -.0003684   .0031869    -0.12   0.908    -.0066461    .0059093
      exp3rd |   .0002109    .000087     2.42   0.016     .0000395    .0003824
       ed4th |   .0000204    .000082     0.25   0.804    -.0001411    .0001819
      exp4th |  -1.96e-06   9.01e-07    -2.18   0.031    -3.73e-06   -1.85e-07
       _cons |   .9050268   .4626025     1.96   0.052    -.0062147    1.816268
------------------------------------------------------------------------------
```

Tests to determine the order of polynomial for *ed* in Model 3 for female employees

- **Test 1.1:** Determine whether a ***third-order*** **polynomial** is adequate for representing the conditional effect on $\ln w_i$ of ed_i.

  Perform on the OLS SRE for Model 3 the following two-tail t-test or F-test:

  H_0: $\beta_8 = 0$

  H_1: $\beta_8 \neq 0$

  Stata commands and results for Test 1.1:

```
. test ed4th

 ( 1)  ed4th = 0

       F(  1,   242) =    0.06
            Prob > F =    0.8039

. lincom _b[ed4th]

 ( 1)  ed4th = 0

------------------------------------------------------------------------------
         lnw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   .0000204    .000082     0.25   0.804    -.0001411    .0001819
------------------------------------------------------------------------------
```

```
. return list

scalars:
                r(df) =  242
                r(se) =  .0000820003024539
          r(estimate) =  .0000203847630542

. display r(estimate)/r(se)
.24859375

. display 2*ttail(r(df), abs(r(estimate)/r(se)))
.80388587
```

Outcome and Implied Decision of Test 1.1:

- **p-value of F_0** and **two-tail p-value of t_0** = **0.8039**

Since the p-value of the calculated F and t statistics for test 1.1 equals 0.8039, we can infer that a ***third-order*** **polynomial** may be adequate for representing the conditional effect on $\ln w_i$ of ed_i, and proceed to Test 1.2 below.

- **<u>Test 1.2</u>:** Determine whether a ***second-order*** **polynomial** is adequate for representing the conditional effect on $\ln w_i$ of ed_i.

  Perform on the OLS SRE for Model 3 the following joint F-test:

  H_0: $\beta_8 = 0$ *and* $\beta_6 = 0$

  H_1: $\beta_8 \neq 0$ *and/or* $\beta_6 \neq 0$

  Stata commands and results for Test 1.2:

```
. test ed4th ed3rd

 ( 1)  ed4th = 0
 ( 2)  ed3rd = 0

       F(  2,   242) =    0.88
            Prob > F =    0.4152

. return list

scalars:
              r(drop) =  0
              r(df_r) =  242
                 r(F) =  .8821689268440828
                r(df) =  2
                 r(p) =  .4152109068470182
```

Outcome and Implied Decision of Test 1.2:

- **p-value of F_0 = 0.4152**

Since the p-value of the calculated F statistic for test 1.2 equals 0.4152, we can infer that a ***second-order*** **polynomial** may be adequate for representing the conditional effect on $\ln w_i$ of ed_i, and proceed to Test 1.3 below.

- **<u>Test 1.3</u>:** Determine whether a ***first-order*** **polynomial** is adequate for representing the conditional effect on $\ln w_i$ of ed_i.

  Perform on the OLS SRE for Model 3 the following joint F-test:

  H_0: $\beta_8 = 0$ *and* $\beta_6 = 0$ *and* $\beta_3 = 0$

  H_1: $\beta_8 \neq 0$ *and/or* $\beta_6 \neq 0$ *and/or* $\beta_3 \neq 0$

  Stata commands and results for Test 1.3:

```
. test ed4th ed3rd edsq

 ( 1)  ed4th = 0
 ( 2)  ed3rd = 0
 ( 3)  edsq = 0

       F(  3,   242) =    5.48
            Prob > F =    0.0012

. return list

scalars:
              r(drop) =  0
              r(df_r) =  242
                 r(F) =  5.482515565912297
                r(df) =  3
                 r(p) =  .001163612607363
```

Outcome and Implied Decision of Test 1.3:

- **p-value of F_0 = 0.0012**

Since the p-value of the calculated F statistic for test 1.3 equals 0.0012, we **reject** the null hypothesis that a ***first-order* polynomial** is adequate for representing the conditional effect on $\ln w_i$ of ed_i.

<u>Results of Tests 1.1, 1.2 and 1.3</u>

A second-order polynomial may be adequate for representing the conditional effect on $\ln w_i$ of ed_i.

In other words, the coefficient exclusion restrictions $\beta_8 = 0$ *and* $\beta_6 = 0$ may be imposed on Model 3 for female employees.

Tests to determine the order of polynomial for *exp* in Model 3 for female employees

- **Test 2.1:** Determine whether a ***third-order* polynomial** is adequate for representing the conditional effect on $\ln w_i$ of exp_i.

  Perform on the OLS SRE for Model 3 the following two-tail t-test or F-test:

  H_0: $\beta_9 = 0$

  H_1: $\beta_9 \neq 0$

  Stata commands and results for Test 2.1:

```
. test exp4th

 ( 1)  exp4th = 0

       F(  1,   242) =    4.73
            Prob > F =    0.0306

. return list

scalars:
              r(drop) =  0
              r(df_r) =  242
                 r(F) =  4.733306717938799
                r(df) =  1
                 r(p) =  .0305521435661833
```

```
. lincom _b[exp4th]

 ( 1)  exp4th = 0

------------------------------------------------------------------------------
         lnw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |  -1.96e-06   9.01e-07    -2.18   0.031    -3.73e-06   -1.85e-07
------------------------------------------------------------------------------

. return list

scalars:
                r(df) =  242
                r(se) =  9.00524794364e-07
          r(estimate) =  -1.95919651081e-06

. display r(estimate)/r(se)
-2.1756164

. display 2*ttail(r(df), abs(r(estimate)/r(se)))
.03055214
```

Outcome and Implied Decision of Test 2.1:

- **p-value of F_0** and **two-tail p-value of t_0 = 0.03055**

Since the p-value of the calculated F and t statistics for test 2.1 equals 0.03055, we **reject** the null hypothesis that a ***third-order*** **polynomial** is adequate for representing the conditional effect on $\ln w_i$ of exp_i.

Results of Tests 2.1

A ***fourth-order* polynomial** is required to adequately represent the conditional effect on $\ln w_i$ of exp_i.

Results of Tests 1.1, 1.2 and 1.3

A ***second-order* polynomial** may be adequate for representing the conditional effect on $\ln w_i$ of ed_i.

In other words, the coefficient exclusion restrictions $\beta_8 = 0$ *and* $\beta_6 = 0$ may be imposed on Model 3 for female employees.

Final Step: Jointly Testing the Coefficient Restrictions

The results of these two sequences of hypothesis tests on the OLS estimates of Model 3 for female employees identify only two coefficient exclusion restrictions that might be imposed on the fourth-order polynomials in ed_i and exp_i for female employees.

These candidate exclusion restrictions should always be **jointly tested** before they are imposed on Model 3, the unrestricted model.

In this case, Test 1.2 of the coefficient exclusion restrictions $\beta_8 = 0$ *and* $\beta_6 = 0$ is the required joint test. But recall that the p-value of the calculated F statistic for test 1.2 equals **0.4152**. We can therefore infer from these two sequences of hypothesis tests that a ***second-order* polynomial** is adequate for representing the conditional effect on $\ln w_i$ of ed_i, but that a ***fourth-order* polynomial** is required to adequately represent the conditional effect on $\ln w_i$ of exp_i.

The Implied Restricted Model for Female Employees

Model 3 for female employees is given by the population regression equation

$$\ln w_i = \beta_0 + \beta_1 ed_i + \beta_2 \exp_i + \beta_3 ed_i^2 + \beta_4 \exp_i^2 + \beta_5 ed_i \exp_i + \beta_6 ed_i^3 + \beta_7 \exp_i^3 + \beta_8 ed_i^4 + \beta_9 \exp_i^4 + u_i \qquad \textbf{(3)}$$

Imposition of the coefficient exclusion restrictions $\beta_8 = 0$ *and* $\beta_6 = 0$ on Model 3 gives the following ***restricted*** **regression model for *female* employees**:

$$\ln w_i = \beta_0 + \beta_1 ed_i + \beta_2 \exp_i + \beta_3 ed_i^2 + \beta_4 \exp_i^2 + \beta_5 ed_i \exp_i + \beta_7 \exp_i^3 + \beta_9 \exp_i^4 + u_i \qquad \textbf{(3f)}$$

- **OLS estimation of Restricted Model 3f in *Stata*:**

```
. regress lnw ed exp edsq expsq edexp exp3rd exp4th if female == 1

      Source |       SS       df       MS              Number of obs =     252
-------------+------------------------------           F(  7,   244) =   15.68
       Model |  15.3662132     7  2.19517332           Prob > F      =  0.0000
    Residual |  34.1673939   244  .140030303           R-squared     =  0.3102
-------------+------------------------------           Adj R-squared =  0.2904
       Total |  49.5336071   251  .197345048           Root MSE      =  .37421

------------------------------------------------------------------------------
         lnw |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          ed |  -.1052975   .0588168    -1.79   0.075     -.221151    .0105561
         exp |   .0993418    .032072     3.10   0.002     .0361686     .162515
        edsq |   .0078393   .0020448     3.83   0.000     .0038115    .0118671
       expsq |  -.0074103    .002714    -2.73   0.007    -.0127561   -.0020645
       edexp |   .0003417   .0009348     0.37   0.715    -.0014997    .0021831
      exp3rd |   .0002016   .0000865     2.33   0.021     .0000313     .000372
      exp4th |  -1.85e-06   8.94e-07    -2.07   0.040    -3.61e-06   -8.59e-08
       _cons |    1.10083   .4382619     2.51   0.013     .2375704    1.964089
------------------------------------------------------------------------------
```

$$E(\ln w_i \mid ed_i, \exp_i) = \beta_0 + \beta_1 ed_i + \beta_2 \exp_i + \beta_3 ed_i^2 + \beta_4 \exp_i^2 + \beta_5 ed_i \exp_i + \beta_7 \exp_i^3 + \beta_9 \exp_i^4 \qquad \textbf{(3f)}$$

- **The *conditional* effects on $\ln w_i$ of ed$_i$ and exp$_i$ in Model 3f for females**:

1. The **conditional effect of *ed*** in Model 3f is the partial, or *ceteris paribus*, relationship between **ed** and the **conditional mean value of $\ln w$** for any given value of the other explanatory variable **exp**.

Set $\exp_i$ = 13 = the sample median value of the explanatory variable ***exp*** for female employees. Then the corresponding **conditional effect of *ed*** in Model 3f is the following **quadratic function of *ed***:

$$E(\ln w_i \mid ed_i, \exp_i = 13) = \beta_0 + \beta_1 ed_i + \beta_2 13 + \beta_3 ed_i^2 + \beta_4 13^2 + \beta_5 ed_i 13 + \beta_7 13^3 + \beta_9 13^4$$

$$= (\beta_0 + \beta_2 13 + \beta_4 13^2 + \beta_7 13^3 + \beta_9 13^4) + (\beta_1 + \beta_5 13) ed_i + \beta_3 ed_i^2$$

2. The **conditional effect of *exp*** in Model 3f is the partial, or *ceteris paribus*, relationship between **exp** and the **conditional mean value of $\ln w$** for any given value of the other explanatory variable **ed**.

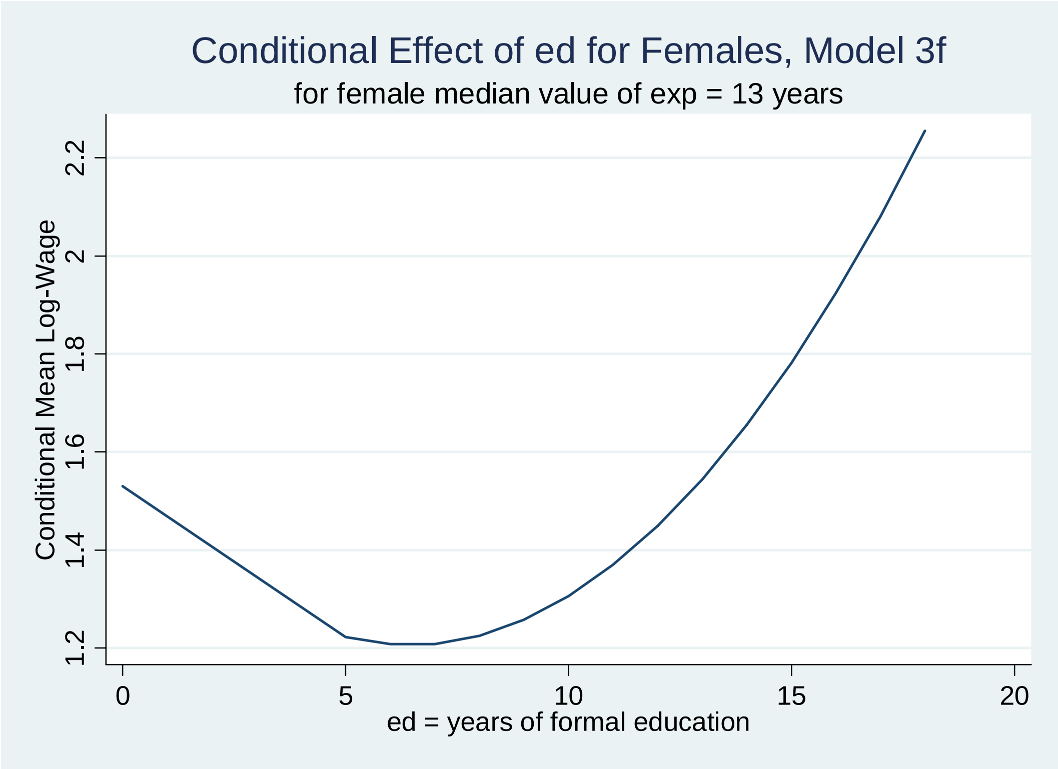
Set ed_i = 12 = the sample median value of the explanatory variable ***ed*** for female employees. Then the corresponding **conditional effect of *exp*** in Model 3f is the following **quartic function of *exp***:

$$E(\ln w_i \mid ed_i = 12, \exp_i) = \beta_0 + \beta_1 12 + \beta_2 \exp_i + \beta_3 12^2 + \beta_4 \exp_i^2 + \beta_5 12 \exp_i + \beta_7 \exp_i^3 + \beta_9 \exp_i^4$$

$$= (\beta_0 + \beta_1 12 + \beta_3 12^2) + (\beta_2 + \beta_5 12) \exp_i + \beta_4 \exp_i^2 + \beta_7 \exp_i^3 + \beta_9 \exp_i^4$$

- **Conditional effect on $\ln w_i$ of ed_i for exp_i = 13 in Model 3f for females**:

$$E(\ln w_i \mid ed_i, exp_i = 13) = (\beta_0 + \beta_2 13 + \beta_4 13^2 + \beta_7 13^3 + \beta_9 13^4) + (\beta_1 + \beta_5 13)ed_i + \beta_3 ed_i^2$$



- **Conditional effect on $\ln w_i$ of $\exp_i$ for ed_i = 12 in Model 3f for females**:

$$E(\ln w_i \mid ed_i = 12, \exp_i) = (\beta_0 + \beta_1 12 + \beta_3 12^2) + (\beta_2 + \beta_5 12)\exp_i + \beta_4 \exp_i^2 + \beta_7 \exp_i^3 + \beta_9 \exp_i^4$$