ECON 452* -- Addendum to NOTE 8

A Full Interaction Regression Model with Higher Order Terms in X_1 and X_2

Model 5.6: Higher order terms in the continuous explanatory variables X_1 and X_2

Expand Model 5.5 to include **cubic (3rd order)** and **quartic (4th order)** terms in the two continuous explanatory variables X_1 and X_2.

The **population regression *equation* for Model 5.6** is:

$$\begin{aligned} Y_i = {} & \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\ & + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\ & + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\ & + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i + u_i \end{aligned} \quad \textbf{(5.6)}$$

The **population regression *function* for Model 5.6** is:

$$\begin{aligned} & E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i) \\ & = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_3 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\ & \quad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\ & \quad + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\ & \quad + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i \end{aligned} \quad \textbf{(5.6')}$$

$$
\begin{aligned}
&E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i) \\
&= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_3 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\
&\qquad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\
&\quad + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\
&\qquad + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i
\end{aligned}
\qquad \textbf{(5.6')}
$$

- The ***female*** **population regression function for Model 5.6** is obtained by setting the female indicator $F_i = 1$ in (5.6'):

$$
\begin{aligned}
&E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i) \\
&= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\
&\quad + \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \\
&= (\beta_0 + \delta_0) + (\beta_1 + \delta_1) X_{i1} + (\beta_2 + \delta_2) X_{i1}^2 + (\beta_3 + \delta_3) X_{i1}^3 + (\beta_4 + \delta_4) X_{i1}^4 + (\beta_5 + \delta_5) X_{i2} + (\beta_6 + \delta_6) X_{i2}^2 \\
&\quad + (\beta_7 + \delta_7) X_{i2}^3 + (\beta_8 + \delta_8) X_{i2}^4 + (\beta_9 + \delta_9) X_{i1} X_{i2} + (\beta_{10} + \delta_{10}) IN2_i + (\beta_{11} + \delta_{11}) IN3_i + (\beta_{12} + \delta_{12}) IN4_i
\end{aligned}
\qquad \textbf{(5.6f)}
$$

$$
\begin{aligned}
&E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i) \\
&\quad = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\
&\qquad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\
&\qquad + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\
&\qquad + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i \qquad \textbf{(5.6')}
\end{aligned}
$$

- The ***male*** **population regression function for Model 5.6** is obtained by setting the female indicator $F_i = 0$ in (5.6'):

$$
\begin{aligned}
&E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i) \\
&\quad = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\
&\qquad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}
\end{aligned}
$$

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$
\begin{aligned}
&E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T) \\
&\quad = \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} \\
&\qquad + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}
\end{aligned}
$$

The **population regression *equation* for Model 5.6** is:

$$\begin{aligned} Y_i = {} & \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\ & + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\ & + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\ & + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i + u_i \end{aligned} \qquad \textbf{(5.6)}$$

***Stata* command for OLS estimation of Model 5.6:**

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

The Marginal Effect of X_1 in Model 5.6

$$
\begin{aligned}
&E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i) \\
&\quad = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\
&\qquad\qquad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\
&\qquad + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\
&\qquad\qquad + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i + u_i
\end{aligned}
\qquad \textbf{(5.6')}
$$

- The **marginal effect of X_1 in Model 5.6** is:

$$
\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \,\middle|\, F_i, x_i^T\right)}{\partial X_{i1}}
$$

$$
= \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 F_i + 2\delta_2 F_i X_{i1} + 3\delta_3 F_i X_{i1}^2 + 4\delta_4 F_i X_{i1}^3 + \delta_9 F_i X_{i2}
$$

- The **marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial Y_i}{\partial X_{i1}} = \frac{\partial E\left(Y_i \mid F_i, x_i^T\right)}{\partial X_{i1}}$$

$$= \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 F_i + 2\delta_2 F_i X_{i1} + 3\delta_3 F_i X_{i1}^2 + 4\delta_4 F_i X_{i1}^3 + \delta_9 F_i X_{i2}$$

- The **marginal effect of X_1 for *females* in Model 5.6** is obtained by setting $F_i = 1$:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}}$$

$$= \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

$$= (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2) X_{i1} + 3(\beta_3 + \delta_3) X_{i1}^2 + 4(\beta_4 + \delta_4) X_{i1}^3 + (\beta_9 + \delta_9) X_{i2}$$

- The **marginal effect of X_1 for *males* in Model 5.6** is obtained by setting $F_i = 0$:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

- The ***female-male difference*** **in the marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}}$$

$$\begin{aligned} &= \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2} \\ &\quad - \beta_1 - 2\beta_2 X_{i1} - 3\beta_3 X_{i1}^2 - 4\beta_4 X_{i1}^3 - \beta_9 X_{i2} \\ &= \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2} \end{aligned}$$

Computing Estimates of the Marginal Effect of the *Continuous* Explanatory Variable X_1 in Model 5.6

- First, **select specific values of the two *continuous* explanatory variables X_1 and X_2** at which to compute estimates of the marginal effect of X_1 for males and females, and the corresponding female-male difference. To illustrate, select the **sample *median* values** – or **50-th percentile values** – of the variables **X_1 and X_2**.

  ***Stata* commands** for defining as scalars the **sample *median* values** of **X_1 and X_2:**

```
summarize x1, detail
return list
scalar x1med = r(p50)

summarize x2, detail
return list
scalar x2med = r(p50)

scalar list x1med x2med
```

- **Recall that the *Stata* command for OLS estimation of Model 5.6:**

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

- The **marginal effect of X_1 for *males* in Model 5.6** is given by the following function:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

Stata* command** for computing an ***estimate* of the marginal effect of X_1 for *males at the **sample *median* values of X_1 and X_2**:

```
lincom _b[x1] + 2*_b[x1sq]*x1med + 3*_b[x13rd]*x1med*x1med +
4*_b[x14th]*x1med*x1med*x1med + _b[x1x2]*x2med
```

- The **marginal effect of X_1 for *females* in Model 5.6** is given by the following function:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} = (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2)X_{i1} + 3(\beta_3 + \delta_3)X_{i1}^2 + 4(\beta_4 + \delta_4)X_{i1}^3 + (\beta_9 + \delta_9)X_{i2}$$

Stata* command** for computing an ***estimate* of the marginal effect of X_1 for *females at the **sample *median* values of X_1 and X_2**:

```
lincom _b[x1] + _b[fx1] + 2*(_b[x1sq] + _b[fx1sq])*x1med
+ 3*(_b[x13rd] + _b[fx13rd])*x1med*x1med
+ 4*(_b[x14th] + _b[fx14th])*x1med*x1med*x1med
+ (_b[x1x2] + _b[fx1x2])*x2med
```

- The ***female-male difference*** **in the marginal effect of X_1 in Model 5.6** is given by the following function:

$$\frac{\partial E\left(Y_i \middle| F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \middle| F_i = 0, x_i^T\right)}{\partial X_{i1}} = \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

Stata **command** for computing an ***estimate*** **of the *female-male difference* in the marginal effect of X_1** at the **sample *median* values of X_1 and X_2**:

```
lincom _b[fx1] + 2*_b[fx1sq]*x1med + 3*_b[fx13rd]*x1med*x1med
+ 4*_b[fx14th]*x1med*x1med*x1med + _b[fx1x2]*x2med
```

Hypothesis Tests Respecting the Marginal Effect of X_1 for *Males* in Model 5.6

- The **marginal effect of X_1 for *males* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

♦ ***<u>Test 1m</u>:*** Test the hypothesis that the ***marginal* effect of X_1 on Y for *males* is *zero*** for all values of X_1 and X_2.

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = 0$ for all i for males are $\beta_1 = 0$ *and* $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_9 = 0$.

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

  H_0: $\beta_1 = 0$ *and* $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_9 = 0$

  H_1: $\beta_1 \neq 0$ *and/or* $\beta_2 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_9 \neq 0$

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

```
test x1 x1sq x13rd x14th x1x2
```

- The **marginal effect of X_1 for *males* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

♦ ***<u>Test 2m</u>:*** Test the hypothesis that the ***marginal* effect of X_1 on Y for *males* is *constant*** – i.e., is unrelated to the values of X_1 and X_2.

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = \beta_1$ (a constant) for all males are $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_9 = 0$.

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

  H_0: $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$ *and* $\beta_9 = 0$

  H_1: $\beta_2 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$ *and/or* $\beta_9 \neq 0$

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

```
test x1sq x13rd x14th x1x2
```

- ***Test 3m:*** Test the hypothesis that the **marginal effect of X_1 on Y for *males* is unrelated to**, or **does not depend upon**, **X_1**.

- The **marginal effect of X_1 for *males* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

- Sufficient conditions for $\partial Y_i / \partial X_{i1}$ to be unrelated to X_{i1} for all males are $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$.

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

  H_0: $\beta_2 = 0$ *and* $\beta_3 = 0$ *and* $\beta_4 = 0$

  H_1: $\beta_2 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_4 \neq 0$

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

```
test x1sq x13rd x14th
```

- ***<u>Test 4m</u>:*** Test the hypothesis that the **marginal effect of X_1 on Y for *males* is unrelated to**, or **does not depend upon**, $\mathbf{X_2}$.

- The **marginal effect of X_1 for *males* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=0} = \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2}$$

- A sufficient condition for $\partial Y_i / \partial X_{i1}$ to be unrelated to X_{i2} for all males is $\beta_9 = 0$.

- The ***null* and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_9 = 0$

  H_1: $\beta_9 \neq 0$

- 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]
```

Hypothesis Tests Respecting the Marginal Effect of X_1 for *Females* in Model 5.6

- The **marginal effect of X_1 for *females* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \,\middle|\, F_i = 1, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

$$= (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2)X_{i1} + 3(\beta_3 + \delta_3)X_{i1}^2 + 4(\beta_4 + \delta_4)X_{i1}^3 + (\beta_9 + \delta_9)X_{i2}$$

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

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = 0$ for all females are $\beta_1 + \delta_1 = 0$ *and* $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_9 + \delta_9 = 0$.

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

  H_0: $\beta_1 + \delta_1 = 0$ *and* $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_9 + \delta_9 = 0$

  H_1: $\beta_1 + \delta_1 \neq 0$ *and/or* $\beta_2 + \delta_2 \neq 0$ *and/or* $\beta_3 + \delta_3 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$ *and/or* $\beta_9 + \delta_9 \neq 0$

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

```
test x1 + fx1 = 0, notest
test x1sq + fx1sq = 0, accumulate notest
test x13rd + fx13rd = 0, accumulate notest
test x14th + fx14th = 0, accumulate notest
test x1x2 + fx1x2 = 0, accumulate
```

- ***<u>Test 2f</u>:*** Test the hypothesis that the ***marginal* effect of X_1 on Y for *females* is *constant*** – i.e., is unrelated to the values of X_1 and X_2.

- The **marginal effect of X_1 for *females* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

$$= (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2) X_{i1} + 3(\beta_3 + \delta_3) X_{i1}^2 + 4(\beta_4 + \delta_4) X_{i1}^3 + (\beta_9 + \delta_9) X_{i2}$$

- Sufficient conditions for $\partial Y_i / \partial X_{i1} = \beta_1 + \delta_1$ (a constant) for all females are $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_9 + \delta_9 = 0$.

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

  H_0: $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_9 + \delta_9 = 0$

  H_1: $\beta_2 + \delta_2 \neq 0$ *and/or* $\beta_3 + \delta_3 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$ *and/or* $\beta_9 + \delta_9 \neq 0$

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

```
test x1sq + fx1sq = 0, notest
test x13rd + fx13rd = 0, accumulate notest
test x14th + fx14th = 0, accumulate notest
test x1x2 + fx1x2 = 0, accumulate
```

- ***<u>Test 3f</u>:*** Test the hypothesis that the **marginal effect of X_1 on Y for *females* is unrelated to**, or **does not depend upon**, **X_1**.

- The **marginal effect of X_1 for *females* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

$$= (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2)X_{i1} + 3(\beta_3 + \delta_3)X_{i1}^2 + 4(\beta_4 + \delta_4)X_{i1}^3 + (\beta_9 + \delta_9)X_{i2}$$

- Sufficient conditions for $\partial Y_i / \partial X_{i1}$ to be unrelated to X_{i1} for all females are $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$.

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

  H_0: $\beta_2 + \delta_2 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_4 + \delta_4 = 0$

  H_1: $\beta_2 + \delta_2 \neq 0$ *and/or* $\beta_3 + \delta_3 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$

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

```
test x1sq + fx1sq = 0, notest
test x13rd + fx13rd = 0, accumulate notest
test x14th + fx14th = 0, accumulate
```

- ***Test 4f:*** Test the hypothesis that the **marginal effect of X_1 on Y for *females* is unrelated to**, or **does not depend upon**, **X_2**.

- The **marginal effect of X_1 for *females* in Model 5.6** is:

$$\left.\frac{\partial Y_i}{\partial X_{i1}}\right|_{F_i=1} = \frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} = \beta_1 + 2\beta_2 X_{i1} + 3\beta_3 X_{i1}^2 + 4\beta_4 X_{i1}^3 + \beta_9 X_{i2} + \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

$$= (\beta_1 + \delta_1) + 2(\beta_2 + \delta_2)X_{i1} + 3(\beta_3 + \delta_3)X_{i1}^2 + 4(\beta_4 + \delta_4)X_{i1}^3 + (\beta_9 + \delta_9)X_{i2}$$

- A sufficient condition for $\partial Y_i / \partial X_{i1}$ to be unrelated to X_{i2} for all females is $\beta_9 + \delta_9 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_9 + \delta_9 = 0$

  H_1: $\beta_9 + \delta_9 \neq 0$

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

```
test x1x2 + fx1x2 = 0
```

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

```
lincom _b[x1x2] + _b[fx1x2]
```

Hypothesis Tests for *Female-Male Differences* in the Marginal Effect of X_1 in Model 5.6

- The ***female-male difference* in the marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial E\left(Y_i \middle| F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \middle| F_i = 0, x_i^T\right)}{\partial X_{i1}} = \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

- ***<u>Test 5</u>:*** Test the hypothesis that the ***marginal* effect of X_1 on Y for *females* equals** the ***marginal* effect of X_1 on Y for *males*** for any values of X_1 and X_2 **–** i.e., the ***female-male difference*** in the ***marginal* effect of X_1 on Y is *zero*** for any values of X_1 and X_2.

- Sufficient conditions for the **female-male *difference*** in the marginal effect of X_1 on Y **to equal *zero*** for all values of X_1 and X_2 are $\delta_1 = 0$ *and* $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_9 = 0$.

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

  H_0: $\delta_1 = 0$ *and* $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_9 = 0$

  H_1: $\delta_1 \neq 0$ *and/or* $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\delta_9 \neq 0$

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

```
test fx1 fx1sq fx13rd fx14th fx1x2
```

- ***Test 6*:** Test the hypothesis that the ***female-male difference*** in the ***marginal*** **effect of X_1 on Y is *a constant*** for any values of X_1 and X_2.

- The ***female-male difference*** **in the marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial E\left(Y_i \mid F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \mid F_i = 0, x_i^T\right)}{\partial X_{i1}} = \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

- Sufficient conditions for the **female-male *difference*** in the marginal effect of X_1 on Y **to equal the constant δ_1** for all values of X_1 and X_2 are $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_9 = 0$.

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

  H_0: $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_9 = 0$

  H_1: $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\delta_9 \neq 0$

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

```
test fx1sq fx13rd fx14th fx1x2
```

- ***Test 7:*** Test the hypothesis that the ***female-male difference*** in the ***marginal*** **effect of X_1 on Y is unrelated to**, or **does not depend upon**, **X_1**.

- The ***female-male difference* in the marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial E\left(Y_i \middle| F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \middle| F_i = 0, x_i^T\right)}{\partial X_{i1}} = \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

- Sufficient conditions for the **female-male *difference*** in the marginal effect of X_1 on Y **to be unrelated to X_1** are $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$.

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

  H_0: $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$

  H_1: $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$

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

```
test fx1sq fx13rd fx14th
```

- ***Test 8:*** Test the hypothesis that the ***female-male difference*** in the ***marginal*** **effect of X_1 on Y is unrelated to**, or **does not depend upon**, **X_2**.

- The ***female-male difference*** **in the marginal effect of X_1 in Model 5.6** is:

$$\frac{\partial E\left(Y_i \,\middle|\, F_i = 1, x_i^T\right)}{\partial X_{i1}} - \frac{\partial E\left(Y_i \,\middle|\, F_i = 0, x_i^T\right)}{\partial X_{i1}} = \delta_1 + 2\delta_2 X_{i1} + 3\delta_3 X_{i1}^2 + 4\delta_4 X_{i1}^3 + \delta_9 X_{i2}$$

- A sufficient condition for the **female-male *difference*** in the marginal effect of X_1 on Y **to be unrelated to X_2** is $\delta_9 = 0$.

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

  H_0: $\delta_9 = 0$

  H_1: $\delta_9 \neq 0$

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

```
test fx1x2
```
or
```
test fx1x2 = 0
```

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

```
lincom _b[fx1x2]
```

Hypothesis Tests Respecting the Effects of Industry in Model 5.6

- ***Test 1-Industry:*** Test the hypothesis of **no industry effects for *males***. This is equivalent to the hypothesis that conditional mean Y for males is unrelated to industry, i.e., that there **are no inter-industry differences in conditional mean Y for *males***.

- The ***male* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \quad \textbf{(5.6m)}$$

- Sufficient conditions for the conditional mean value of Y for males to be unrelated to industry are $\beta_{10} = 0$ *and* $\beta_{11} = 0$ *and* $\beta_{12} = 0$.

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

  H_0: $\beta_{10} = 0$ *and* $\beta_{11} = 0$ *and* $\beta_{12} = 0$

  H_1: $\beta_{10} \neq 0$ *and/or* $\beta_{11} \neq 0$ *and/or* $\beta_{12} \neq 0$

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

```
test in2 in3 in4
```

- ***Test 2-Industry:*** Test the hypothesis of **no industry effects for *females***. This is equivalent to the hypothesis that conditional mean Y for females is unrelated to industry, i.e., that there **are no inter-industry differences in conditional mean Y for *females***.

- The ***female* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$

$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \quad \text{… (5.6f)}$$

- Sufficient conditions for the conditional mean value of Y for females to be unrelated to industry are $\beta_{10} + \delta_{10} = 0$ *and* $\beta_{11} + \delta_{11} = 0$ *and* $\beta_{12} + \delta_{12} = 0$.

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

  H_0: $\beta_{10} + \delta_{10} = 0$ *and* $\beta_{11} + \delta_{11} = 0$ *and* $\beta_{12} + \delta_{12} = 0$

  H_1: $\beta_{10} + \delta_{10} \neq 0$ *and/or* $\beta_{11} + \delta_{11} \neq 0$ *and/or* $\beta_{12} + \delta_{12} \neq 0$

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

```
test in2 + fin2 = 0, notest
test in3 + fin3 = 0, accumulate notest
test in4 + fin4 = 0, accumulate
```

♦ ***<u>Test 3-Industry</u>:*** Test the hypothesis of **no *female-male differences* in industry effects** – i.e., that the ***female-male difference*** in conditional mean Y **is unrelated to industry**.

This is equivalent to the hypothesis that **industry effects are *equal*** for ***females* and *males***, i.e., that **inter-industry differences in conditional mean Y for *females* equal inter-industry differences in conditional mean Y for *males***.

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

… (5.6d)

- Sufficient conditions for the female-male difference in conditional mean Y to be unrelated to industry (for equal industry effects for males and females) are $\delta_{11} = 0$ *and* $\delta_{12} = 0$ *and* $\delta_{13} = 0$.

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

  H_0: $\delta_{10} = 0$ *and* $\delta_{11} = 0$ *and* $\delta_{12} = 0$

  H_1: $\delta_{10} \neq 0$ *and/or* $\delta_{11} \neq 0$ *and/or* $\delta_{12} \neq 0$

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

```
test fin2 fin3 fin4
```

Hypothesis Tests Respecting *Female-Male Differences* in Conditional Mean Y in Model 5.6

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \quad \text{… (5.6d)}$$

- ***Test 1:*** The ***female-male difference* in conditional mean Y *equals zero* for all observations**, i.e., for any given values of the explanatory variables X_1, X_2, and industry.

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

H_0: $\delta_0 = 0$ *and* $\delta_1 = 0$ *and* $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_5 = 0$ *and* $\delta_6 = 0$ *and* $\delta_7 = 0$ *and* $\delta_8 = 0$ *and* $\delta_9 = 0$ *and* $\delta_{10} = 0$ *and* $\delta_{11} = 0$ *and* $\delta_{12} = 0$

or $\delta_j = 0$ for all j = 0, 1, …, 12

H_1: $\delta_0 \neq 0$ *and/or* $\delta_1 \neq 0$ *and/or* $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\delta_5 \neq 0$ *and/or* $\delta_6 \neq 0$ *and/or* $\delta_7 \neq 0$ *and/or* $\delta_8 \neq 0$ *and/or* $\delta_9 \neq 0$ *and/or* $\delta_{10} \neq 0$ *and/or* $\delta_{11} \neq 0$ *and/or* $\delta_{12} \neq 0$

or $\delta_j \neq 0$ j = 0, 1, …, 12

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

```
test f fx1 fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3
fin4
```

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

… (5.6d)

♦ ***<u>Test 2</u>*:** The ***female-male difference* in conditional mean Y *equals a constant***, i.e., it **does not depend on** the values of the explanatory variables **X_1, X_2, and industry**.

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

H_0: $\delta_1 = 0$ *and* $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_5 = 0$ *and* $\delta_6 = 0$ *and* $\delta_7 = 0$ *and* $\delta_8 = 0$ *and* $\delta_9 = 0$ *and* $\delta_{10} = 0$ *and* $\delta_{11} = 0$ *and* $\delta_{12} = 0$

or $\delta_j = 0$ for all j = 1, 2, …, 12

H_1: $\delta_1 \neq 0$ *and/or* $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\delta_5 \neq 0$ *and/or* $\delta_6 \neq 0$ *and/or* $\delta_7 \neq 0$ *and/or* $\delta_8 \neq 0$ *and/or* $\delta_9 \neq 0$ *and/or* $\delta_{10} \neq 0$ *and/or* $\delta_{11} \neq 0$ *and/or* $\delta_{12} \neq 0$

or $\delta_j \neq 0$ j = 1, 2, …, 12

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

```
test fx1 fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3
fin4
```

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

… (5.6d)

- ***<u>Test 3</u>:*** The ***female-male difference*** in conditional mean Y **does not depend on X_1** – i.e., the **marginal effect of X_1 is *equal*** for ***males*** **and *females***.

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

H_0: $\delta_1 = 0$ *and* $\delta_2 = 0$ *and* $\delta_3 = 0$ *and* $\delta_4 = 0$ *and* $\delta_9 = 0$

or $\delta_j = 0$ for all j = 1, 2, 3, 4, 9

H_1: $\delta_1 \neq 0$ *and/or* $\delta_2 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\delta_9 \neq 0$

or $\delta_j \neq 0$ j = 1, 2, 3, 4, 9

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

```
test fx1 fx1sq fx13rd fx14th fx1x2
```

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

… (5.6d)

♦ ***<u>Test 4</u>:*** The ***female-male difference*** in conditional mean Y **does not depend on X_2** – i.e., the **marginal effect of X_2 is *equal*** for ***males*** **and *females***.

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

  H_0: $\delta_5 = 0$ *and* $\delta_6 = 0$ *and* $\delta_7 = 0$ *and* $\delta_8 = 0$ *and* $\delta_9 = 0$

  H_1: $\delta_5 \neq 0$ *and/or* $\delta_6 \neq 0$ *and/or* $\delta_7 \neq 0$ *and/or* $\delta_8 \neq 0$ *and/or* $\delta_9 \neq 0$

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

```
test fx2 fx2sq fx23rd fx24th fx1x2
```

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

… (5.6d)

♦ ***<u>Test 5</u>:*** The ***female-male difference*** in conditional mean Y **does not depend on industry** – i.e., **industry effects** are ***equal*** for ***males*** **and** ***females***.

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

  H_0: $\delta_{10} = 0$ *and* $\delta_{11} = 0$ *and* $\delta_{12} = 0$

  H_1: $\delta_{10} \neq 0$ *and/or* $\delta_{11} \neq 0$ *and/or* $\delta_{12} \neq 0$

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

```
test fin2 fin3 fin4
```

Hypothesis Tests for Selecting the Order of Polynomial for X_1 in Model 5.6

- The ***male*** **population regression function for Model 5.6**:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \quad \textbf{(5.6m)}$$

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1) X_{i1} + (\beta_2 + \delta_2) X_{i1}^2 + (\beta_3 + \delta_3) X_{i1}^3 + (\beta_4 + \delta_4) X_{i1}^4 + (\beta_5 + \delta_5) X_{i2} + (\beta_6 + \delta_6) X_{i2}^2$$

$$+ (\beta_7 + \delta_7) X_{i2}^3 + (\beta_8 + \delta_8) X_{i2}^4 + (\beta_9 + \delta_9) X_{i1} X_{i2} + (\beta_{10} + \delta_{10}) IN2_i + (\beta_{11} + \delta_{11}) IN3_i + (\beta_{12} + \delta_{12}) IN4_i \quad \textbf{(5.6f)}$$

Again, the *Stata* command for OLS estimation of Model 5.6:

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

Tests for Selecting the Order of Polynomial for X_1 for *Males* in Model 5.6

- ***Test 5m:*** Test the hypothesis that a ***third-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *males***.

- The ***male*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$

$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

- A sufficient condition for the male population regression function to be a third-order polynomial in X_{i1} is $\beta_4 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition 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 x14th` *or* **`test x14th = 0`**

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

```
lincom _b[x14th]
```

- ***Test 6m:*** Test the hypothesis that a ***second-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *males***.

- The ***male*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$

$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

- Sufficient conditions for the male population regression function to be a second-order polynomial in X_{i1} are $\beta_4 = 0$ *and* $\beta_3 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 = 0$ *and* $\beta_3 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_3 \neq 0$

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

```
test x14th x13rd
```

- ***<u>Test 7m</u>:*** Test the hypothesis that a ***first-order* polynomial** is adequate for representing the **conditional effect of X_1 on Y for *males***.

- The ***male* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$

$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

- Sufficient conditions for the male population regression function to be a second-order polynomial in X_{i1} are $\beta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_2 = 0$.

- The ***null* and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_2 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_2 \neq 0$

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

```
test x14th x13rd x1sq
```

- ***Test 8m:*** Test the hypothesis that a ***zero-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *males***.

- The ***male*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$
$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$
$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

- Sufficient conditions or the male population regression function to be a second-order polynomial in X_{i1} are $\beta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_2 = 0$ *and* $\beta_1 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_2 = 0$ *and* $\beta_1 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_2 \neq 0$ *and/or* $\beta_1 \neq 0$

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

```
test x14th x13rd x1sq x1
```

Tests for Selecting the Order of Polynomial for X_1 for *Females* in Model 5.6

- ***Test 5f:*** Test the hypothesis that a ***third-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *females***.

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$

$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \quad \textbf{(5.6f)}$$

- A sufficient condition for the female population regression function to be a third-order polynomial in X_{i1} is $\beta_4 + \delta_4 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 + \delta_4 = 0$

  H_1: $\beta_4 + \delta_4 \neq 0$

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

```
test x14th + fx14th = 0
```

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

```
lincom _b[x14th] + _b[fx14th]
```

- ***Test 6f:*** Test the hypothesis that a ***second-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *female***.

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$

$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \qquad \textbf{(5.6f)}$$

- Sufficient conditions for the female population regression function to be a second-order polynomial in X_{i1} are $\beta_4 + \delta_4 = 0$ *and* $\beta_3 + \delta_3 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 + \delta_4 = 0$ *and* $\beta_3 + \delta_3 = 0$

  H_1: $\beta_4 + \delta_4 = 0$ *and/or* $\beta_3 + \delta_3 = 0$

- Compute an **F-test** of H_0 against H_1 using the following series of linked *Stata* **test** commands:

```
test x14th + fx14th = 0, notest
test x13rd + fx13rd = 0, accumulate
```

- ***Test 7f:*** Test the hypothesis that a ***first-order* polynomial** is adequate for representing the **conditional effect of X_1 on Y for *females***.

- The ***female* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$

$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \qquad \textbf{(5.6f)}$$

- Sufficient conditions for the female population regression function to be a second-order polynomial in X_{i1} are $\beta_4 + \delta_4 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_2 + \delta_2 = 0$.

- The ***null* and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_2 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_2 \neq 0$

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

```
test x14th + fx14th = 0, notest
test x13rd + fx13rd = 0, accumulate notest
test x1sq + fx1sq = 0, accumulate
```

- ***Test 8f:*** Test the hypothesis that a ***zero-order*** **polynomial** is adequate for representing the **conditional effect of X_1 on Y for *females***.

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$

$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \qquad \textbf{(5.6f)}$$

- Sufficient conditions for the female population regression function to be a second-order polynomial in X_{i1} are $\beta_4 + \delta_4 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_2 + \delta_2 = 0$ *and* $\beta_1 + \delta_1 = 0$.

- The ***null*** **and *alternative* hypotheses** for this proposition are:

  H_0: $\beta_4 + \delta_4 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_2 + \delta_2 = 0$ *and* $\beta_1 + \delta_1 = 0$

  H_1: $\beta_4 + \delta_4 = 0$ *and/or* $\beta_3 + \delta_3 = 0$ *and/or* $\beta_2 + \delta_2 = 0$ *and/or* $\beta_1 + \delta_1 = 0$

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

```
test x14th + fx14th = 0, notest
test x13rd + fx13rd = 0, accumulate notest
test x1sq + fx1sq = 0, accumulate notest
test x1 + fx1 = 0, accumulate
```

A Symmetric Simplification Strategy for Selecting the Order of Polynomial in X_1 for *Both* Males and Females

General Nature: This model simplification procedure **tests *jointly* for *both* males and females** the proposition that some specific order of polynomial in X_1 is adequate.

Advantages of Keeping the Selected Order of Polynomial for X_1 the Same for Both Males and Females

- It results in **the *same* order of polynomial in X_1** being adopted **for both males and females**.
- It is **relatively straightforward** to implement.

Propositions in the Hypothesis Testing Sequence on Model 5.6

- ***<u>Test 1</u>:*** Test that a **3rd-order polynomial** (or cubic) **in X_1** is adequate for ***both* males and females**, i.e., that both the ***male*** and ***female* slope coefficients** on the regressor X_{i1}^4 are equal to zero.
- ***<u>Test 2</u>:*** Test that a **2nd-order polynomial** (or quadratic) **in X_1** is adequate for ***both* males and females**, i.e., that both the ***male*** and ***female* slope coefficients** on the regressors X_{i1}^3 and X_{i1}^4 are equal to zero.
- ***<u>Test 3</u>:*** Test that a **1st-order polynomial** (or linear function) **in X_1** is adequate for ***both* males and females**, i.e., that both the ***male*** and ***female* slope coefficients** on the regressors X_{i1}^2, X_{i1}^3 and X_{i1}^4 are equal to zero.

- ***Test 1:*** A **3rd-order polynomial (or cubic) in X_1** is adequate for representing the partial, or conditional, relationship of X_1 to Y **for *both* males and females**, i.e., that ***both* the *male* and *female* slope coefficients** on the regressor X_{i1}^4 are equal to zero.

- The ***null* and *alternative* hypotheses** to test on Model 5.6 are:

  H_0: $\beta_4 = 0$ *and* $\beta_4 + \delta_4 = 0$ OR H_0: $\beta_4 = 0$ *and* $\delta_4 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$ OR H_1: $\beta_4 \neq 0$ *and/or* $\delta_4 \neq 0$

  Note: Imposing the 2 coefficient restrictions in H_0 on Model 5.6 implies a 3rd-order polynomial in X_1 for both males and females.

- Compute an **F-test** of H_0 against H_1 on Model 5.6 using the following *Stata* **test** commands:

```
test x14th = 0, notest
test x14th + fx14th = 0, accumulate
```

  OR

```
test x14th fx14th
```

- ***How to Proceed:***

  If **H_0 is *retained***, proceed to Test 2, the next test in the testing sequence.

  If **H_0 is *rejected***, choose a **4th-order polynomial in X_1 for <u>both</u> *males* and *females***.

- ***Test 2:*** A **2nd-order polynomial (or quadratic) in X_1** is adequate for representing the partial, or conditional, relationship of X_1 to Y **for *both* males and females**, i.e., that ***both* the *male* and *female* slope coefficients** on the regressors X_{i1}^3 and X_{i1}^4 are equal to zero.

- The ***null* and *alternative* hypotheses** to test on Model 5.6 are:

  H_0: $\beta_4 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_3 + \delta_3 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_3 + \delta_3 \neq 0$

  OR

  H_0: $\beta_4 = 0$ *and* $\delta_4 = 0$ *and* $\beta_3 = 0$ *and* $\delta_3 = 0$

  H_1: $\beta_4 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\delta_3 \neq 0$

  Note: Imposing the 4 coefficient restrictions in H_0 on Model 5.6 implies a 2nd-order polynomial in X_1 for both males and females.

- Compute an **F-test** of H_0 against H_1 on Model 5.6 using the following *Stata* **test** commands:

```
test x14th = 0, notest
test x14th + fx14th = 0, notest accumulate
test x13rd = 0, notest accumulate
test x13rd + fx13rd = 0, accumulate
```

  OR

```
test x14th fx14th x13rd fx13rd
```

- ***How to Proceed:***

  If $\mathbf{H_0}$ **is *retained***, proceed to Test 3, the next test in the testing sequence.

  If $\mathbf{H_0}$ **is *rejected***, choose a **3rd-order polynomial in X_1 for <u>both</u> *males* and *females***.

- ***Test 3:*** A **1st-order polynomial (or linear function) in X_1** is adequate for representing the partial, or conditional, relationship of X_1 to Y **for *both* males and females**, i.e., that ***both* the *male* and *female* slope coefficients** on the regressors X_{i1}^2, X_{i1}^3 and X_{i1}^4 are equal to zero.

- The ***null* and *alternative* hypotheses** to test on Model 5.6 are:

H_0: $\beta_4 = 0$ *and* $\beta_4 + \delta_4 = 0$ *and* $\beta_3 = 0$ *and* $\beta_3 + \delta_3 = 0$ *and* $\beta_2 = 0$ *and* $\beta_2 + \delta_2 = 0$ *and*

H_1: $\beta_4 \neq 0$ *and/or* $\beta_4 + \delta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\beta_3 + \delta_3 \neq 0$ *and/or* $\beta_2 \neq 0$ *and/or* $\beta_2 + \delta_2 \neq 0$

OR

H_0: $\beta_4 = 0$ *and* $\delta_4 = 0$ *and* $\beta_3 = 0$ *and* $\delta_3 = 0$ *and* $\beta_2 = 0$ *and* $\delta_2 = 0$

H_1: $\beta_4 \neq 0$ *and/or* $\delta_4 \neq 0$ *and/or* $\beta_3 \neq 0$ *and/or* $\delta_3 \neq 0$ *and/or* $\beta_2 \neq 0$ *and/or* $\delta_2 \neq 0$

Note: Imposing the 6 coefficient restrictions in H_0 on Model 5.6 implies a 1st-order polynomial in X_1 for both males and females.

- Compute an **F-test** of H_0 against H_1 on Model 5.6 using the following *Stata* **test** commands:

```
test x14th = 0, notest
test x14th + fx14th = 0, notest accumulate
test x13rd = 0, notest accumulate
test x13rd + fx13rd = 0, notest accumulate
test x1sq = 0, notest accumulate
test x1sq + fx1sq = 0, accumulate
```

OR

```
test x14th fx14th x13rd fx13rd x1sq fx1sq
```

- ***How to Proceed:***

  If **H_0 is *retained***, adopt a **1st-order polynomial in X_1 for <u>both</u> *males* and *females***.

  If **H_0 is *rejected***, choose a **2nd-order polynomial in X_1 for <u>both</u> *males* and *females***.

Evaluating the *Marginal* Effects of the *Categorical* Explanatory Variable in Model 5.6

General Nature: The **marginal effects of a *categorical* explanatory variable such as industry** consist of **the differences in conditional mean values of Y between *pairs* of industry categories** – e.g., the **conditional mean Y difference between *males* in industries 4 and 2**, and the **conditional mean Y difference between *females* in industries 4 and 2.**

Recall that the **population regression *function* for Model 5.6** is:

$$E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_3 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$
$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i$$
$$+ \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2}$$
$$+ \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i \qquad \textbf{(5.6')}$$

$$E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_3 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$
$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i$$
$$+ \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2}$$
$$+ \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i \qquad \textbf{(5.6')}$$

- The ***female*** **population regression function for Model 5.6** is obtained by setting the female indicator $F_i = 1$ in (5.6'):

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i$$
$$+ \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1) X_{i1} + (\beta_2 + \delta_2) X_{i1}^2 + (\beta_3 + \delta_3) X_{i1}^3 + (\beta_4 + \delta_4) X_{i1}^4 + (\beta_5 + \delta_5) X_{i2} + (\beta_6 + \delta_6) X_{i2}^2$$
$$+ (\beta_7 + \delta_7) X_{i2}^3 + (\beta_8 + \delta_8) X_{i2}^4 + (\beta_9 + \delta_9) X_{i1} X_{i2} + (\beta_{10} + \delta_{10}) IN2_i + (\beta_{11} + \delta_{11}) IN3_i + (\beta_{12} + \delta_{12}) IN4_i \qquad \textbf{(5.6f)}$$

$$E(Y_i \mid X_{i1}, X_{i2}, F_i, IN2_i, IN3_i, IN4_i)$$

$$\begin{aligned} &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\ &\quad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \\ &\quad + \delta_0 F_i + \delta_1 F_i X_{i1} + \delta_2 F_i X_{i1}^2 + \delta_3 F_i X_{i1}^3 + \delta_4 F_i X_{i1}^4 + \delta_5 F_i X_{i2} + \delta_6 F_i X_{i2}^2 + \delta_7 F_i X_{i2}^3 + \delta_8 F_i X_{i2}^4 + \delta_9 F_i X_{i1} X_{i2} \\ &\quad + \delta_{10} F_i IN2_i + \delta_{11} F_i IN3_i + \delta_{12} F_i IN4_i \end{aligned} \qquad \textbf{(5.6')}$$

- The ***male*** **population regression function for Model 5.6** is obtained by setting the female indicator $F_i = 0$ in (5.6'):

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$\begin{aligned} &= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2} \\ &\quad + \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \end{aligned} \qquad \textbf{(5.6m)}$$

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$\begin{aligned} &= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2} \\ &\quad + \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \end{aligned} \qquad \textbf{(5.6d)}$$

Marginal Effects of Industry for Males in Model 5.6

- The ***male*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$
$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

The ***Stata*** **command** for computing OLS estimates of Model 5.6 is:

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

1. The **industry 2-industry 1 difference** in **conditional mean Y for *males*** equals β_{10} in Model 5.6. To display an estimate of β_{10} in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in2]
```

2. The **industry 3-industry 1 difference** in **conditional mean Y for *males*** equals β_{11} in Model 5.6. To display an estimate of β_{11} in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in3]
```

- The ***male* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 0, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$
$$= \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i1}^2 + \beta_3 X_{i1}^3 + \beta_4 X_{i1}^4 + \beta_5 X_{i2} + \beta_6 X_{i2}^2 + \beta_7 X_{i2}^3 + \beta_8 X_{i2}^4 + \beta_9 X_{i1} X_{i2}$$
$$+ \beta_{10} IN2_i + \beta_{11} IN3_i + \beta_{12} IN4_i \qquad \textbf{(5.6m)}$$

3. The **industry 4-industry 1 difference** in **conditional mean Y for *males*** equals β_{12} in Model 5.6. To display an estimate of β_{12} in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4]
```

4. The **industry 3-industry 2 difference** in **conditional mean Y for *males*** equals $\beta_{11} - \beta_{10}$ in Model 5.6. To compute an estimate of $\beta_{11} - \beta_{10}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in3] - _b[in2]
```

5. The **industry 4-industry 2 difference** in **conditional mean Y for *males*** equals $\beta_{12} - \beta_{10}$ in Model 5.6. To compute an estimate of $\beta_{12} - \beta_{10}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4] - _b[in2]
```

6. The **industry 4-industry 3 difference** in **conditional mean Y for *males*** equals $\beta_{12} - \beta_{11}$ in Model 5.6. To compute an estimate of $\beta_{12} - \beta_{11}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4] - _b[in3]
```

Marginal Effects of Industry for Females in Model 5.6

- The ***female* population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$
$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \quad \textbf{(5.6f)}$$

Again, the ***Stata* command** for computing OLS estimates of regression equation (5.5) is:

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

1. The **industry 2-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{10} + \delta_{10}}$ in Model 5.6. To compute an estimate of $\beta_{10} + \delta_{10}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in2] + _b[fin2]
```

2. The **industry 3-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{11} + \delta_{11}}$ in Model 5.6. To compute an estimate of $\beta_{11} + \delta_{11}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in3] + _b[fin3]
```

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$
$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \quad \textbf{(5.6f)}$$

3. The **industry 4-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{12} + \delta_{12}}$ in Model 5.6. To compute an estimate of $\beta_{12} + \delta_{12}$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4] + _b[fin4]
```

4. The **industry 3-industry 2 difference** in **conditional mean Y for *females*** equals $\boldsymbol{(\beta_{11} + \delta_{11}) - (\beta_{10} + \delta_{10})}$ in Model 5.6. To compute an estimate of $(\beta_{11} + \delta_{11}) - (\beta_{10} + \delta_{10})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in3] + _b[fin3] – (_b[in2] + _b[fin2])
```

5. The **industry 4-industry 2 difference** in **conditional mean Y for *females*** equals $\boldsymbol{(\beta_{12} + \delta_{12}) - (\beta_{10} + \delta_{10})}$ in Model 5.6. To compute an estimate of $(\beta_{12} + \delta_{12}) - (\beta_{10} + \delta_{10})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4] + _b[fin4] – (_b[in2] + _b[fin2])
```

- The ***female*** **population regression function for Model 5.6** is:

$$E(Y_i \mid F_i = 1, X_{i1}, X_{i2}, IN2_i, IN3_i, IN4_i)$$

$$= (\beta_0 + \delta_0) + (\beta_1 + \delta_1)X_{i1} + (\beta_2 + \delta_2)X_{i1}^2 + (\beta_3 + \delta_3)X_{i1}^3 + (\beta_4 + \delta_4)X_{i1}^4 + (\beta_5 + \delta_5)X_{i2} + (\beta_6 + \delta_6)X_{i2}^2$$
$$+ (\beta_7 + \delta_7)X_{i2}^3 + (\beta_8 + \delta_8)X_{i2}^4 + (\beta_9 + \delta_9)X_{i1}X_{i2} + (\beta_{10} + \delta_{10})IN2_i + (\beta_{11} + \delta_{11})IN3_i + (\beta_{12} + \delta_{12})IN4_i \qquad \textbf{(5.6f)}$$

6. The **industry 4-industry 3 difference** in **conditional mean Y for *females*** equals $(\boldsymbol{\beta_{12} + \delta_{12}}) - (\boldsymbol{\beta_{11} + \delta_{11}})$ in Model 5.6. To compute an estimate of $(\beta_{12} + \delta_{12}) - (\beta_{11} + \delta_{11})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[in4] + _b[fin4] – (_b[in3] + _b[fin3])
```

Female-Male Differences in the *Marginal* Effects of Industry in Model 5.6

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2}$$
$$+ \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}$$

Again, the ***Stata* command** for computing OLS estimates of regression equation (5.6) is:

```
regress y x1 x1sq x13rd x14th x2 x2sq x23rd x24th x1x2 in2 in3 in4 f fx1
fx1sq fx13rd fx14th fx2 fx2sq fx23rd fx24th fx1x2 fin2 fin3 fin4
```

1. The **industry 2-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{10} + \delta_{10}}$ in Model 5.6. The **industry 2-industry 1 difference** in **conditional mean Y for *males*** equals $\boldsymbol{\beta_{10}}$.

   The **female-male difference** in the **industry 2-industry 1 difference** in **conditional mean Y** is therefore:

   $= \boldsymbol{\beta_{10} + \delta_{10} - \beta_{10} = \delta_{10}}$

   To display an estimate of δ_{10} in Model 5.6, use the following *Stata* **limcom** command:

   ```
   lincom _b[fin2]
   ```

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2}$$
$$+ \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}$$

2. The **industry 3-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{11} + \delta_{11}}$ in Model 5.6. The **industry 3-industry 1 difference** in **conditional mean Y for *males*** equals $\boldsymbol{\beta_{11}}$.

   The **female-male difference** in the **industry 3-industry 1 difference** in **conditional mean Y** is therefore:

   $= \boldsymbol{\beta_{11} + \delta_{11} - \beta_{11} = \delta_{11}}$

   To display an estimate of δ_{11} in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[fin3]
```

3. The **industry 4-industry 1 difference** in **conditional mean Y for *females*** equals $\boldsymbol{\beta_{12} + \delta_{12}}$ in Model 5.6. The **industry 4-industry 1 difference** in **conditional mean Y for *males*** equals $\boldsymbol{\beta_{12}}$.

   The **female-male difference** in the **industry 4-industry 1 difference** in **conditional mean Y** is therefore:

   $= \boldsymbol{\beta_{12} + \delta_{12} - \beta_{12} = \delta_{12}}$

   To display an estimate of δ_{12} in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[fin4]
```

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2}$$
$$+ \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}$$

4. The **industry 3-industry 2 difference** in **conditional mean Y for *females*** equals $(\beta_{11} + \delta_{11}) - (\beta_{10} + \delta_{10})$ in Model 5.6. The **industry 3-industry 2 difference** in **conditional mean Y for *males*** equals $\beta_{11} - \beta_{10}$.

   The **female-male difference** in the **industry 3-industry 2 difference** in **conditional mean Y** is therefore:

   $= (\beta_{11} + \delta_{11}) - (\beta_{10} + \delta_{10}) - (\beta_{11} - \beta_{10})$

   $= \beta_{11} + \delta_{11} - \beta_{10} - \delta_{10} - (\beta_{11} - \beta_{10})$

   $= (\beta_{11} - \beta_{10}) + (\delta_{11} - \delta_{10}) - (\beta_{11} - \beta_{10})$

   $= (\delta_{11} - \delta_{10})$

   To compute an estimate of $(\delta_{11} - \delta_{10})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[fin3] – _b[fin2]
```

- The ***female-male difference* in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2}$$
$$+ \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}$$

5. The **industry 4-industry 2 difference** in **conditional mean Y for *females*** equals $(\boldsymbol{\beta_{12}} + \boldsymbol{\delta_{12}}) - (\boldsymbol{\beta_{10}} + \boldsymbol{\delta_{10}})$ in Model 5.6. The **industry 4-industry 2 difference** in **conditional mean Y for *males*** equals $\boldsymbol{\beta_{12}} - \boldsymbol{\beta_{10}}$.

   The **female-male difference** in the **industry 4-industry 2 difference** in **conditional mean Y** is therefore:

   $$= (\beta_{12} + \delta_{12}) - (\beta_{10} + \delta_{10}) - (\beta_{12} - \beta_{10})$$
   $$= \beta_{12} + \delta_{12} - \beta_{10} - \delta_{10} - (\beta_{12} - \beta_{10})$$
   $$= (\beta_{12} - \beta_{10}) + (\delta_{12} - \delta_{10}) - (\beta_{12} - \beta_{10})$$
   $$= (\delta_{12} - \delta_{10})$$

   To compute an estimate of $(\delta_{12} - \delta_{10})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[fin4] – _b[fin2]
```

- The ***female-male difference*** **in conditional mean Y for Model 5.6** is:

$$E(Y_i \mid F_i = 1, x_i^T) - E(Y_i \mid F_i = 0, x_i^T)$$

$$= \delta_0 + \delta_1 X_{i1} + \delta_2 X_{i1}^2 + \delta_3 X_{i1}^3 + \delta_4 X_{i1}^4 + \delta_5 X_{i2} + \delta_6 X_{i2}^2 + \delta_7 X_{i2}^3 + \delta_8 X_{i2}^4 + \delta_9 X_{i1} X_{i2}$$
$$+ \delta_{10} IN2_i + \delta_{11} IN3_i + \delta_{12} IN4_i \qquad \textbf{(5.6d)}$$

6. The **industry 4-industry 3 difference** in **conditional mean Y for *females*** equals $\boldsymbol{(\beta_{12} + \delta_{12}) - (\beta_{11} + \delta_{11})}$ in Model 5.6. The **industry 4-industry 3 difference** in **conditional mean Y for *males*** equals $\boldsymbol{\beta_{12} - \beta_{11}}$.

   The **female-male difference** in the **industry 4-industry 2 difference** in **conditional mean Y** is therefore:

   $$\boldsymbol{= (\beta_{12} + \delta_{12}) - (\beta_{11} + \delta_{11}) - (\beta_{12} - \beta_{11})}$$
   $$\boldsymbol{= \beta_{12} + \delta_{12} - \beta_{11} - \delta_{11} - (\beta_{12} - \beta_{11})}$$
   $$\boldsymbol{= (\beta_{12} - \beta_{11}) + (\delta_{12} - \delta_{11}) - (\beta_{12} - \beta_{11})}$$
   $$\boldsymbol{= (\delta_{12} - \delta_{11})}$$

   To compute an estimate of $(\delta_{12} - \delta_{11})$ in Model 5.6, use the following *Stata* **limcom** command:

```
lincom _b[fin4] – _b[fin3]
```