## Interpretation of coefficients in multiple regression

Consider the multiple regression model with $K = 2$ :

$$y_i = x_{1i}\beta_1 + x_{2i}\beta_2 + u_i \qquad i = 1, ..., n$$

$$\text{or} \quad y = X_1\beta_1 + X_2\beta_2 + u$$

$$= X\beta + u$$

where $y, X_1, X_2$ and $u$ are $n \times 1$ column vectors, $X$ is an $n \times 2$ matrix and $\beta$ is a $2 \times 1$ column vector

Notice that $X_1 = (x_{11}, ..., x_{1n})'$ and $X_2 = (x_{21}, ..., x_{2n})'$ denote the two columns of $X$

$$y = X_1\beta_1 + X_2\beta_2 + u = X\beta + u$$

Assume, for simplicity, that $\overline{y} = \overline{X}_1 = \overline{X}_2 = 0$ (not really restrictive, since we can always express variables as deviations from their sample means)

Recall

$$\widehat{\beta}_{\text{OLS}} = (X'X)^{-1}X'y = Ay$$

$$\widehat{y} = X\widehat{\beta}_{\text{OLS}} = X(X'X)^{-1}X'y = Py$$

$$\widehat{u} = y - \widehat{y} = (I - X(X'X)^{-1}X')y = (I - P)y = My$$

with $AX = I$, $MX = 0$ and $X'\widehat{u} = 0$

Note that

$$X'\widehat{u} = \begin{pmatrix} X_1'\widehat{u} \\ X_2'\widehat{u} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Also note that

$$\begin{aligned} y &= \widehat{y} + \widehat{u} \\ &= X\widehat{\beta}_{\text{OLS}} + \widehat{u} \\ &= X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + \widehat{u} \end{aligned}$$

How can we interpret the OLS estimates of the individual coefficients, $\widehat{\beta}_1$ and $\widehat{\beta}_2$, in the multiple regression model?

First consider the simple regression of $y$ on $X_1$ :

$$y = X_1\gamma + r_y$$

We have

$$\widehat{\gamma}_{\text{OLS}} = (X_1'X_1)^{-1}X_1'y = A_1y$$

$$\widetilde{y} = X_1\widehat{\gamma}_{\text{OLS}} = X_1(X_1'X_1)^{-1}X_1'y = P_1y$$

$$\widehat{r}_y = y - \widetilde{y} = (I - X_1(X_1'X_1)^{-1}X_1')y = (I - P_1)y = M_1y$$

with $A_1X_1 = I$, $M_1X_1 = 0$ and $X_1'\widehat{r}_y = 0$

Note that we can write

$$\widehat{\gamma}_{\text{OLS}} = A_1y = A_1[X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + \widehat{u}]$$

Consider also the simple regression of $X_2$ on $X_1$ :

$$X_2 = X_1\delta + r_2$$

We have

$$\widehat{\delta}_{\text{OLS}} = (X_1'X_1)^{-1}X_1'X_2 = A_1X_2$$

$$\widehat{X}_2 = X_1\widehat{\delta}_{\text{OLS}} = X_1(X_1'X_1)^{-1}X_1'X_2 = P_1X_2$$

$$\widehat{r}_2 = X_2 - \widehat{X}_2 = (I - X_1(X_1'X_1)^{-1}X_1')X_2 = (I - P_1)X_2 = M_1X_2$$

with $X_1'\widehat{r}_2 = 0$

Now write

$$\begin{aligned}
\widehat{\gamma}_{\text{OLS}} &= A_1 y = A_1[X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + \widehat{u}] \\
&= (X_1'X_1)^{-1}(X_1'X_1)\widehat{\beta}_1 + (X_1'X_1)^{-1}(X_1'X_2)\widehat{\beta}_2 + (X_1'X_1)^{-1}X_1'\widehat{u} \\
&= \widehat{\beta}_1 + \widehat{\delta}_{\text{OLS}}\widehat{\beta}_2 + 0 \\
&= \widehat{\beta}_1 + \widehat{\delta}_{\text{OLS}}\widehat{\beta}_2
\end{aligned}$$

So $\widehat{\beta}_1 \neq \widehat{\gamma}_{\text{OLS}}$ , unless either $\widehat{\beta}_2 = 0$ or $\widehat{\delta}_{\text{OLS}} = 0$ (or both)

In general, the estimated coefficient $(\widehat{\beta}_1)$ on $X_1$ in the multiple regression of $y$ on $X_1$ and $X_2$ **differs** from the estimated coefficient $(\widehat{\gamma}_{\text{OLS}})$ on $X_1$ in the simple regression of $y$ on $X_1$

Now consider the simple regression of $\widehat{r}_y$ on $\widehat{r}_2$:

$$\widehat{r}_y = \widehat{r}_2 b + v$$

where $\widehat{r}_y$ are the OLS residuals from the simple regression of $y$ on $X_1$, and $\widehat{r}_2$ are the OLS residuals from the simple regression of $X_2$ on $X_1$

$$\widehat{b}_{\text{OLS}} = (\widehat{r}_2'\widehat{r}_2)^{-1}\widehat{r}_2'\widehat{r}_y$$

Recall that $\widehat{r}_2 = M_1 X_2$ and $\widehat{r}_y = M_1 y$, with $M_1$ a symmetric and idempotent matrix

$$\widehat{r}_2'\widehat{r}_2 = (M_1 X_2)' M_1 X_2 = (X_2' M_1') M_1 X_2 = X_2' M_1 M_1 X_2 = X_2' M_1 X_2$$

$$\widehat{r}_2'\widehat{r}_y = (M_1 X_2)' M_1 y = (X_2' M_1') M_1 y = X_2' M_1 M_1 y = X_2' M_1 y$$

So $\widehat{b}_{\text{OLS}} = (\widehat{r}_2'\widehat{r}_2)^{-1}\widehat{r}_2'\widehat{r}_y = (X_2'M_1X_2)^{-1}X_2'M_1y$

Consider
$$\begin{aligned} M_1y &= M_1[X_1\widehat{\beta}_1 + X_2\widehat{\beta}_2 + \widehat{u}] \\ &= M_1X_1\widehat{\beta}_1 + M_1X_2\widehat{\beta}_2 + M_1\widehat{u} \\ &= 0 + M_1X_2\widehat{\beta}_2 + M_1\widehat{u} \qquad \text{since } M_1X_1 = 0 \end{aligned}$$

Moreover
$$\begin{aligned} M_1\widehat{u} &= (I - X_1(X_1'X_1)^{-1}X_1')\widehat{u} \\ &= \widehat{u} - X_1(X_1'X_1)^{-1}X_1'\widehat{u} \\ &= \widehat{u} \qquad \text{since } X_1'\widehat{u} = 0 \end{aligned}$$

Then we have $M_1y = M_1X_2\widehat{\beta}_2 + \widehat{u}$

And using $M_1 y = M_1 X_2 \widehat{\beta}_2 + \widehat{u}$, we can write

$$\begin{aligned}
\widehat{b}_{\text{OLS}} &= (X_2' M_1 X_2)^{-1} X_2' M_1 y \\
&= (X_2' M_1 X_2)^{-1} X_2' [M_1 X_2 \widehat{\beta}_2 + \widehat{u}] \\
&= (X_2' M_1 X_2)^{-1} (X_2' M_1 X_2) \widehat{\beta}_2 + (X_2' M_1 X_2)^{-1} X_2' \widehat{u} \\
&= \widehat{\beta}_2 \qquad\qquad \text{since } X_2' \widehat{u} = 0
\end{aligned}$$

The estimated coefficient $(\widehat{\beta}_2)$ on $X_2$ in the multiple regression of $y$ on $X_1$ and $X_2$ **equals** the estimated coefficient $(\widehat{b}_{\text{OLS}})$ on $\widehat{r}_2$ in the simple regression of $\widehat{r}_y$ on $\widehat{r}_2$, i.e. where we regress the OLS residuals $(\widehat{r}_y)$ from the regression of $y$ on $X_1$ on the OLS residuals $(\widehat{r}_2)$ from the regression of $X_2$ on $X_1$

$$
\begin{aligned}
y &= X_1\beta_1 + X_2\beta_2 + u \\
\widehat{r}_y &= \widehat{r}_2 b + v \\
\widehat{\beta}_2 &= \widehat{b}_{\text{OLS}}
\end{aligned}
$$

That is, we partial out or remove (in a linear regression sense) the effect of $X_1$ on both $y$ and $X_2$ before considering the relationship between $y$ and $X_2$

In effect, the estimated coefficient in the multiple regression considers the relationship between $y$ and $X_2$, controlling for the effect of $X_1$

Recall that we have $X_1'\widehat{r}_y = 0$ and $X_1'\widehat{r}_2 = 0$, so both $\widehat{r}_y$ and $\widehat{r}_2$ are orthogonal to $X_1$ by construction

This is closely related to the concept of **partial correlation** between $y$ and $X_2$, holding $X_1$ constant

$$\rho_{yX_2.X_1} = \text{corr}(\widehat{r}_y, \widehat{r}_2) = \frac{\text{cov}(\widehat{r}_y, \widehat{r}_2)}{\sqrt{\text{Var}(\widehat{r}_y)\text{Var}(\widehat{r}_2)}}$$

$$\widehat{\beta}_2 = \widehat{b}_{\text{OLS}} = (\widehat{r}_2'\widehat{r}_2)^{-1}\widehat{r}_2'\widehat{r}_y = \frac{\sum_{i=1}^{n} \widehat{r}_{yi}\widehat{r}_{2i}}{\sum_{i=1}^{n} \widehat{r}_{2i}^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} \widehat{r}_{yi}\widehat{r}_{2i}}{\frac{1}{n}\sum_{i=1}^{n} \widehat{r}_{2i}^2}$$

The numerator $\frac{1}{n}\sum_{i=1}^{n} \widehat{r}_{yi}\widehat{r}_{2i}$ estimates $\text{cov}(\widehat{r}_y, \widehat{r}_2)$

As a result:

- $\widehat{\beta}_2$ has the same sign as $\widehat{\rho}_{yX_2.X_1}$
- $\widehat{\beta}_2 = 0$ if and only if $\widehat{\rho}_{yX_2.X_1} = 0$
- the $R^2$ from the simple regression of $\widehat{r}_y$ on $\widehat{r}_2$ estimates $(\rho_{yX_2.X_1})^2$

The OLS estimates of coefficients in the multiple regression model are thus informative about patterns of partial correlations in the data

These ideas generalise to models with more than 2 explanatory variables

$$y_i = \beta_1 x_{1i} + \beta_2 x_{2i} + ... + \beta_{K-1} x_{(K-1)i} + \beta_K x_{Ki} + u_i$$

Consider

$$y_i = \gamma_1 x_{1i} + \gamma_2 x_{2i} + ... + \gamma_{K-1} x_{(K-1)i} + r_{yi}$$

$$x_{Ki} = \delta_1 x_{1i} + \delta_2 x_{2i} + ... + \delta_{K-1} x_{(K-1)i} + r_{Ki}$$

$$\widehat{r}_{yi} = b\widehat{r}_{Ki} + v_i$$

Again we have that

$$\widehat{\beta}_K = \widehat{b}_{\text{OLS}} = (\widehat{r}_K' \widehat{r}_K)^{-1} \widehat{r}_K' \widehat{r}_y = \frac{\sum_{i=1}^{n} \widehat{r}_{yi} \widehat{r}_{Ki}}{\sum_{i=1}^{n} \widehat{r}_{Ki}^2} = \frac{\frac{1}{n}\sum_{i=1}^{n} \widehat{r}_{yi} \widehat{r}_{Ki}}{\frac{1}{n}\sum_{i=1}^{n} \widehat{r}_{Ki}^2}$$

Notice that here we have partialled out the effects of *all* the remaining $K-1$ explanatory variables in the multiple regression model, in the construction of the OLS residuals $\widehat{r}_y$ and $\widehat{r}_K$

By construction, these OLS residuals are orthogonal to $(X_1, X_2, ..., X_{K-1})$

i.e. we have $X_j'\widehat{r}_y = X_j'\widehat{r}_K = 0$ for $j = 1, 2, ..., K-1$

The result that $\widehat{\beta}_K = \widehat{b}_{\text{OLS}}$ is sometimes called the Frisch-Waugh-Lovell theorem

Again the OLS estimate $(\widehat{\beta}_K)$ of the coefficient on $X_K$ in the multiple regression of $y$ on $(X_1, X_2, ..., X_{K-1}, X_K)$ is related to the partial correlation between $y$ and $X_K$, holding $(X_1, X_2, ..., X_{K-1})$ constant

$$\rho_{yX_K.X_1X_2...X_{K-1}} = \text{corr}(\widehat{r}_y, \widehat{r}_K) = \frac{\text{cov}(\widehat{r}_y, \widehat{r}_K)}{\sqrt{\text{Var}(\widehat{r}_y)Var(\widehat{r}_K)}}$$

The $R^2$ from the OLS regression of $\widehat{r}_y$ on $\widehat{r}_K$ estimates $(\rho_{yX_K.X_1X_2...X_{K-1}})^2$

This provides a measure of the strength of the relationship between $y$ and $X_K$, holding $(X_1, X_2, ..., X_{K-1})$ constant, and is sometimes referred to as the **partial** $R^2$ measure