The linear regression model with multiple regressors, and the ceteris paribus interpretation of slope coefficients

Econometrics for minor Finance, Lecture 5

Tunga Kantarcı, Fall 2025

Linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

The linear regression model with no column of ones

$$1,$$

and hence no intercept term

$$\beta_0,$$

but with one explanatory variable is given by

$$y = \beta x + u$$

Since there is only one β, we do not subscript it with 1 as β_1.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

In this model the OLS estimator of

$$\beta$$

is given by

$$\hat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^* y_i^*}{\frac{1}{n}\sum_{i=1}^{n} (x_i^*)^2}$$

where

$$x_i^* = x_i - 0$$

and

$$y_i^* = y_i - 0$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

That is,

$$\hat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - 0)(y_i - 0)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2}$$

The numerator is the sample covariance between x and y, and the denominator is the sample variance of x.

Hence, the slope measures the amount of movement in y that is aligned with the movement in x, per unit of variation in x.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Here

$$\hat{\beta} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - 0)(y_i - 0)}{\frac{1}{n}\sum_{i=1}^{n}(x_i - 0)^2}$$

it is imposed that the variation in x is a spread from a mean of 0. Same for y. But why should x have a mean of 0? In practice, there is no reason that an explanatory variable should have a mean of 0. This is restrictive. If the mean is indeed not 0, then the OLS estimator is biased.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

The linear regression model with a column of ones

$$1$$

and one explanatory variable

$$x$$

is given by

$$y = \beta_0 \cdot 1 + \beta_1 x + u$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

In this model the OLS estimator of

$$\beta_1$$

is given by

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^* y_i^*}{\frac{1}{n}\sum_{i=1}^{n} (x_i^*)^2}$$

where

$$x_i^* = x_i - \bar{x}$$

and

$$y_i^* = y_i - \bar{y}$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

That is,

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{x})}{\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

The numerator is the sample covariance between x and y, and the denominator is the sample variance of x.

Hence, the slope measures the amount of movement in y that is aligned with the movement in x, per unit of variation in x.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

So now we allow

$$\bar{x}$$

and

$$\bar{y}$$

to be non zero.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

What does

$$x_i^* = x_i - \bar{x}$$

represent?

From each x_i in the data, we subtract the mean. What does this buy us?

If we do not subtract the mean, we end up with

$$x_i^* = x_i - \bar{x} + \bar{x}$$

meaning that the slope would not just reflect $x_i - \bar{x}$, the variation in x_i, but partly also $\bar{x}$, the average level of x.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

With the subtraction

$$x_i^* = x_i - \bar{x}$$

we net out the average level of x_i from x_i, so that the slope is not contaminated by it. In other words, the slope now measures the effect of changes in x_i, controlling for the average level of x_i.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

This is what including a column of ones in the model buys us. When we include it, we control for the average level of x_i, and the OLS estimator works as intended: it is unbiased. Let's elaborate a little further.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Recall from an earlier lecture that in the current regression model:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

This ensures that the fitted line passes through the point $(\bar{x}, \bar{y})$. Let's see why.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

The fitted line is

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

Evaluating x at $x = \bar{x}$ gives

$$\hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

By definition we know that

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Substituting this gives

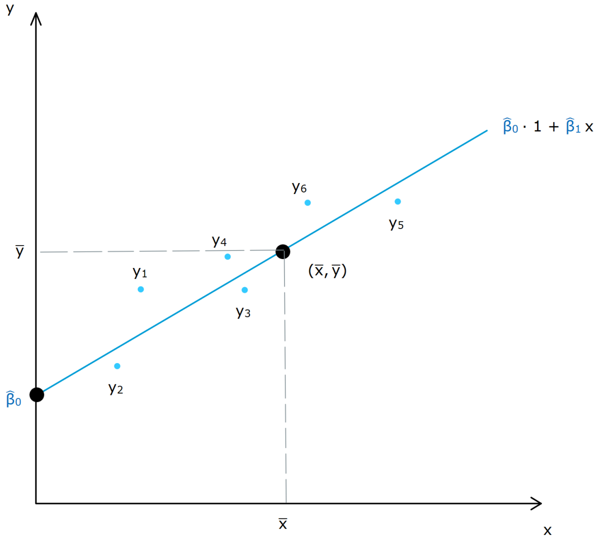
$$\bar{y} - \hat{\beta}_1 \bar{x} + \hat{\beta}_1 \bar{x}$$

where two same terms cancel, leaving

$$\bar{y}$$

Hence, when $x = \bar{x}$, $\hat{y} = \bar{x}$.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients



The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

The intercept is chosen so that the regression line goes through the sample means, thereby accounting for the average levels of x and y. This adjustment allows the line to shift vertically, so the slope

$$\hat{\beta}_1$$

reflects only how deviations in x are associated with deviations in y.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

Notice also that if we have assumed that

$$\bar{x} = 0, \bar{y} = 0$$

we would have been forcing

$$\hat{\beta}_0$$

to 0. But, again, in practice there is no reason that these means should be 0.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Without the intercept term, the regression line is forced to pass through the origin, even though the scatter plot of y against x need not suggest such an intersection. This means the line cannot be shifted vertically to match the data's average levels. In the figure below, the scatter plot indicates that the OLS fitted line is

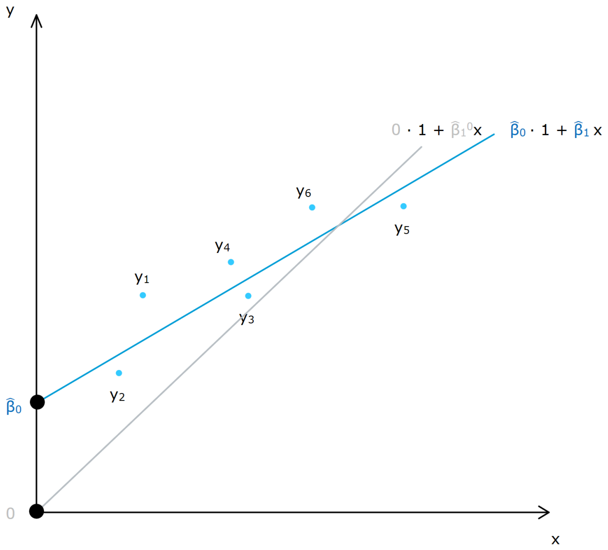
$$\hat{\beta}_1 x,$$

and it would make no sense to impose a fit such as

$$\hat{\beta}_1^0 x$$

that is constrained to pass through the origin.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients



The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

This explains why we should always include a column of ones in the regression model to avoid a biased estimator.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

The linear model with a column of ones

$$1$$

and two explanatory variables

$$x_1, x_2$$

is given by

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

In this model the OLS estimator of

$$\beta_1$$

is given by

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n} x_{1i}^* y_i^*}{\frac{1}{n}\sum_{i=1}^{n} (x_{1i}^*)^2}$$

where

$$x_{1i}^* = x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}$$

and

$$y_i^* = y_i - \hat{\gamma}_0 - \hat{\gamma}_1 x_{2i}$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

That is,

$$\hat{\beta}_1 = \frac{\frac{1}{n}\sum_{i=1}^{n}\left(x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}\right)\left(y_i - \hat{\gamma}_0 - \hat{\gamma}_1 x_{2i}\right)}{\frac{1}{n}\sum_{i=1}^{n}\left(x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}\right)^2}$$

The numerator is the sample covariance between x_1 and y, and the denominator is the sample variance of x_1, where both x_1 and y are netted out of the effect of x_2 on them.

Hence, the slope measures the amount of movement in y that is aligned with the movement in x_1, per unit of variation in x_1, where the slope nets out the effect of x_2 at the same time.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

What do

$$x_{1i}^{*} = x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}$$

represent? They are the residuals from the regression

$$x_{1i} = \alpha_0 + \alpha_1 x_{2i} + \nu_i$$

That is,

$$\hat{\nu}_i = x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}$$

and we simply define

$$\hat{\nu}_i := x_{1i}^{*}$$

This definition is used for consistency in the slides, where x^{*} denotes a transformed regressor.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

That is, with

$$x_{1i}^* = x_{1i} - \hat{\alpha}_0 - \hat{\alpha}_1 x_{2i}$$

we net out the effects of α_0 and x_{2i} on x_{1i}, ensuring that the slope for x_{1i} is not contaminated by their influence. In other words, the slope now measures the effect of changes in x_{1i}, controlling for the effects of α_0 and x_{2i}.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Consider the transformed regression

$$y - \gamma_0 - \gamma_1 x_2 = \beta_1 \left(x_1 - \alpha_0 - \alpha_1 x_2 \right) + u$$

Here, both the independent and dependent variables are expressed in their transformed form, x^* and y^* as discussed. Now compare with the original regression

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

The key point is that the OLS estimator of the slope coefficient

$$\beta_1$$

is identical in both regressions:

$$\hat{\beta}_1$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

This result shows that, for the estimate

$$\hat{\beta}_1$$

first netting out the effects of the constant and x_2 from both y and x_1 and then running the regression is equivalent to estimating

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$$

Hence,

$$\hat{\beta}_1$$

represents the partial effect of x_{1i} on y, controlling for the constant and x_2.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

This is the power of the multiple regression analysis. It allows to do in a non-experimental economic setting what natural scientists are able to do in a controlled laboratory setting: keeping other factors fixed. It provides this ceteris paribus interpretation although the data have not been collected in a ceteris paribus fashion.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

This result is fundamental in regression analysis and originates from publications in the first issue of Econometrica based on Frisch and Waugh (1933) and Lovell (1963) which led to the Frisch-Waugh-Lovell theorem.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Yet, this does not guarantee a causal effect. It guarantees conditional correlation.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients

Consider once again the regression

$$x_1 = \alpha_0 + \alpha_1 x_2 + \nu$$

and its residuals

$$\hat{\nu} = x_1 - \hat{\alpha}_0 - \hat{\alpha}_1 x_2$$

If

$$\hat{\alpha}_1$$

is statistically insignificant, this means that x_2 does not correlate with x_1. In that case, it does not matter whether we control for x_2 when estimating

$$\hat{\beta}_1$$

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Consider the regression of hourly wage on a constant term and schooling in years.

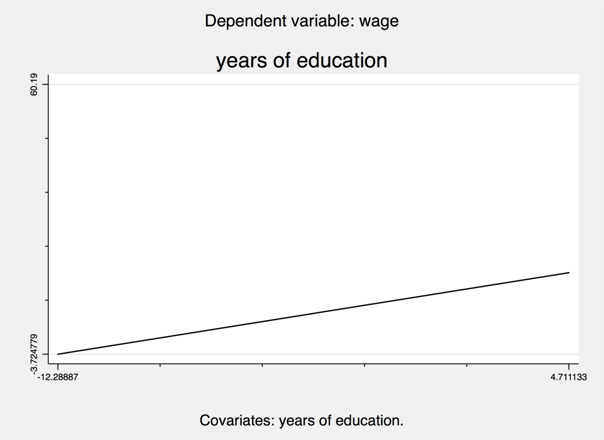
The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

```
. regress wage educ

      Source |       SS           df       MS      Number of obs   =       997
-------------+----------------------------------   F(1, 995)       =    251.46
       Model |  7842.35455         1  7842.35455   Prob > F        =    0.0000
    Residual |  31031.0745       995  31.1870095   R-squared       =    0.2017
-------------+----------------------------------   Adj R-squared   =    0.2009
       Total |   38873.429       996  39.0295472   Root MSE        =    5.5845

------------------------------------------------------------------------------
        wage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        educ |   1.135645   .0716154    15.86   0.000     .9951106     1.27618
       _cons |  -4.860424   .9679821    -5.02   0.000    -6.759944   -2.960903
------------------------------------------------------------------------------
```

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

The figure below shows the fitted line from the regression of wage on education.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example



The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Now consider the regression of wage on a constant, education, and experience.

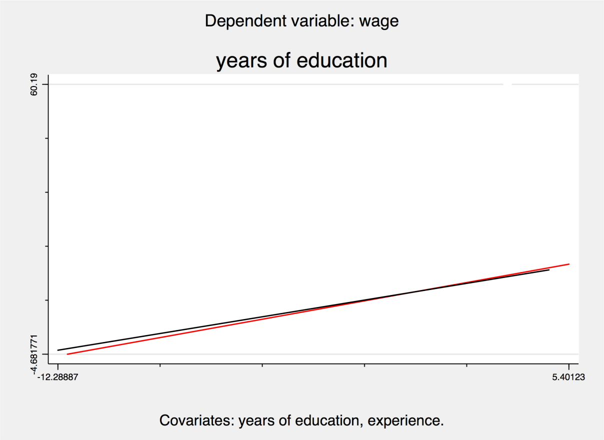
The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

```
. regress wage educ exper

      Source |       SS           df       MS      Number of obs   =       997
-------------+----------------------------------   F(2, 994)       =    172.32
       Model |  10008.3629         2  5004.18147   Prob > F        =    0.0000
    Residual |  28865.0661       994  29.0393019   R-squared       =    0.2575
-------------+----------------------------------   Adj R-squared   =    0.2560
       Total |   38873.429       996  39.0295472   Root MSE        =    5.3888

------------------------------------------------------------------------------
        wage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        educ |   1.246932   .0702966    17.74   0.000     1.108985    1.384879
       exper |   .1327808   .0153744     8.64   0.000     .1026108    .1629509
       _cons |  -8.833768   1.041212    -8.48   0.000    -10.87699   -6.790542
------------------------------------------------------------------------------
```

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

The figure below adds the fitted line from the regression of wage on education, after partialling out the effect of experience: red line. It demonstrates how the fitted line changes when we control for experience.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example



The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

The coefficient of education has changed, meaning that education and experience are correlated, and that we should control for experience in the model while analyzing the effect of education on wages.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

The interpretation of the coefficient estimate of education is as follows. An additional year of schooling increases hourly wage by $1.25, on average, holding experience constant. Holding experience constant means imagining comparing individuals who differ only in education, but have the same level of experience.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Consider the regression of education on experience.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

```
. regress educ exper

      Source |       SS           df       MS      Number of obs   =       997
-------------+----------------------------------   F(1, 995)       =     34.59
       Model |  204.317954         1  204.317954   Prob > F        =    0.0000
    Residual |  5876.48847       995  5.90601856   R-squared       =    0.0336
-------------+----------------------------------   Adj R-squared   =    0.0326
       Total |  6080.80642       996  6.10522733   Root MSE        =    2.4302

------------------------------------------------------------------------------
        educ | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       exper |  -.0400901    .006816    -5.88   0.000    -.0534655   -.0267147
       _cons |   14.04201   .1493993    93.99   0.000     13.74884    14.33519
------------------------------------------------------------------------------
```

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Education and experience are negatively correlated. Obtain the residuals of this model, and call them **Reduc**:

```
. predict Reduc, resid
```

These residuals represent education where the impact of experience on education is netted out.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Consider the regression of wage on residualized education.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

```
. regress wage Reduc

      Source |       SS           df       MS      Number of obs   =       997
-------------+----------------------------------   F(1, 995)       =    305.73
       Model |  9136.99562         1  9136.99562   Prob > F        =    0.0000
    Residual |  29736.4334       995  29.8858627   R-squared       =    0.2350
-------------+----------------------------------   Adj R-squared   =    0.2343
       Total |   38873.429       996  39.0295472   Root MSE        =    5.4668

------------------------------------------------------------------------------
        wage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
       Reduc |   1.246932   .0713139    17.49   0.000     1.106989    1.386875
       _cons |   10.23101   .1731352    59.09   0.000     9.891261    10.57077
------------------------------------------------------------------------------
```

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

Consider the regression of wage on education and experience.

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

```
. regress wage educ exper

      Source |       SS           df       MS      Number of obs   =       997
-------------+----------------------------------   F(2, 994)       =    172.32
       Model |  10008.3629         2  5004.18147   Prob > F        =    0.0000
    Residual |  28865.0661       994  29.0393019   R-squared       =    0.2575
-------------+----------------------------------   Adj R-squared   =    0.2560
       Total |   38873.429       996  39.0295472   Root MSE        =    5.3888

------------------------------------------------------------------------------
        wage | Coefficient  Std. err.      t    P>|t|     [95% conf. interval]
-------------+----------------------------------------------------------------
        educ |   1.246932   .0702966    17.74   0.000     1.108985    1.384879
       exper |   .1327808   .0153744     8.64   0.000     .1026108    .1629509
       _cons |  -8.833768   1.041212    -8.48   0.000    -10.87699   -6.790542
------------------------------------------------------------------------------
```

The linear regression model with multiple regressors: The ceteris paribus interpretation of slope coefficients: Example

The coefficient of residualized education in the first regression and the coefficient of education in the second regression are the same, as the Frisch-Waugh-Lowell theorem requires.

Estimating the standard deviation of the OLS estimator: SD estimator

The standard error estimator of OLS estimator $\hat{\beta}_j$ in the linear regression model with multiple predictors is given by

$$\text{SEE}\left[\hat{\beta}_j \,\middle|\, x\right] = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 \left(1 - R_j^2\right)}}$$

if errors are homoskedastic. Here

$$R_j^2$$

is the R^2 from a regression of x_j on all other regressors.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants

$$\text{SEE}\left[\hat{\beta}_j \,\middle|\, x\right] = \sqrt{\frac{\hat{\sigma}^2}{\sum_{i=1}^{n} (x_{ij} - \bar{x}_j)^2 (1 - R_j^2)}}$$

The expression shows that the SEE of the OLS estimator is

i. higher if the estimated variance of the regression error $\hat{\sigma}^2$ is higher,
ii. lower if the sample size n is larger,
iiii. lower if the sample variation in the predictor $x_i - \bar{x}$ is larger,
iv. larger if correlation between x_j and other regressors is larger meaning if R_j^2 is larger.

Estimating the standard deviation of the OLS estimator: SD estimator: Determinants: Intuition

The intuition behind the last determinant is straightforward. The more a predictor is correlated with other predictors, the more information it shares with them. With less unique variation left to identify its effect, the variance of its OLS estimator increases.