

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

$$\beta,$$

but with one explanatory variable is given by

$$y = \beta x + u$$

# 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,

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

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

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 to assume that an explanatory variable has a mean of zero. This is restrictive. If this mean is 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,

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

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

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  $x$ . But we do not want the latter.

# 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  $x_i$ , 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. We can be a bit more specific.

# 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 the fitted line passes through the point  $(\bar{x}, \bar{y})$ . By **aligning** the line with  $(\bar{x}, \bar{y})$ , the intercept **absorbs** the average levels of  $x$  and  $y$ . This alignment allows the regression 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

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. In the figure below, the scatter plot indicates that the OLS fitted line is

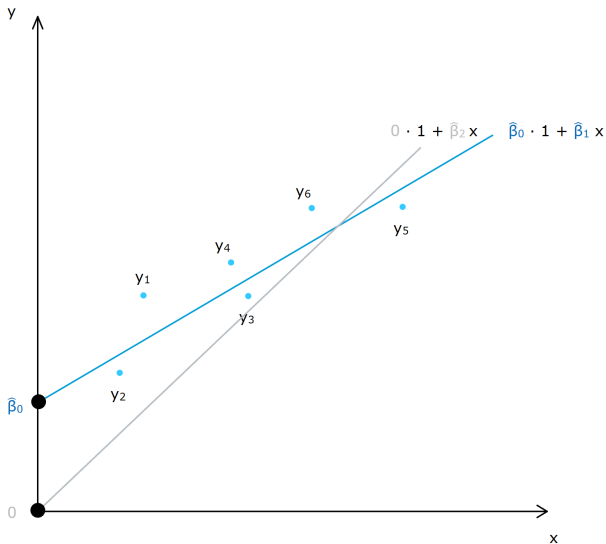
$$\hat{\beta}_1 x,$$

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

$$\hat{\beta}_2 x$$

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,

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

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

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  $\alpha_0$  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  $\alpha_0$  and  $x_{2i}$ .

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

Consider the transformed regression

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

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
Model	<b>7842.35455</b>	<b>1</b>	<b>7842.35455</b>	F(1, 995)	=	<b>251.46</b>
Residual	<b>31031.0745</b>	<b>995</b>	<b>31.1870095</b>	Prob > F	=	<b>0.0000</b>
				R-squared	=	<b>0.2017</b>
				Adj R-squared	=	<b>0.2009</b>
Total	<b>38873.429</b>	<b>996</b>	<b>39.0295472</b>	Root MSE	=	<b>5.5845</b>

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

# 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
Model	<b>10008.3629</b>	<b>2</b>	<b>5004.18147</b>	F(2, 994)	=	<b>172.32</b>
Residual	<b>28865.0661</b>	<b>994</b>	<b>29.0393019</b>	Prob > F	=	<b>0.0000</b>
				R-squared	=	<b>0.2575</b>
				Adj R-squared	=	<b>0.2560</b>
Total	<b>38873.429</b>	<b>996</b>	<b>39.0295472</b>	Root MSE	=	<b>5.3888</b>

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

# 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
Model	<b>204.317954</b>	<b>1</b>	<b>204.317954</b>	F(1, 995)	=	<b>34.59</b>
Residual	<b>5876.48847</b>	<b>995</b>	<b>5.90601856</b>	Prob > F	=	<b>0.0000</b>
				R-squared	=	<b>0.0336</b>
				Adj R-squared	=	<b>0.0326</b>
Total	<b>6080.80642</b>	<b>996</b>	<b>6.10522733</b>	Root MSE	=	<b>2.4302</b>

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

# 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
Model	9136.99562	1	9136.99562	F(1, 995)	=	305.73
Residual	29736.4334	995	29.8858627	Prob > F	=	0.0000
				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
Model	<b>10008.3629</b>	<b>2</b>	<b>5004.18147</b>	F(2, 994)	=	<b>172.32</b>
Residual	<b>28865.0661</b>	<b>994</b>	<b>29.0393019</b>	Prob > F	=	<b>0.0000</b>
				R-squared	=	<b>0.2575</b>
				Adj R-squared	=	<b>0.2560</b>
Total	<b>38873.429</b>	<b>996</b>	<b>39.0295472</b>	Root MSE	=	<b>5.3888</b>

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

# 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-Lowe theorem requires.