Regression Variable Transformations

Ani Katchova

Outline

- Marginal or partial effects
- Regression with quadratic terms
- Regression with interaction terms
- Rescaling variables
- Rescaling logged variables

Marginal or partial effect

- A marginal or partial effect shows the change in the dependent variable y for a one unit change in the independent variable x_j.
- Regression model, $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$
- $\frac{\Delta y}{\Delta x_1} = \frac{\partial y}{\partial x_1} = \beta_1$
- The marginal effect is the coefficient, so the coefficient can be interpreted as the increase in y by β_1 units associated with an increase in x_1 by 1 unit.

Marginal or partial effect with nonlinearities

A model with quadratic terms and interaction terms (still a linear in parameters model)

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_2^2 + \beta_4 x_3 + \beta_5 x_3 x_4 + u$$

The marginal effect is obtained by taking a derivative with respect to a variable.

- $\frac{\Delta y}{\Delta x_1} = \frac{\partial y}{\partial x_1} = \beta_1$
- $\frac{\Delta y}{\Delta x_2} = \frac{\partial y}{\partial x_2} = \beta_2 + 2\beta_3 x_2$
- $\frac{\Delta y}{\Delta x_3} = \frac{\partial y}{\partial x_3} = \beta_4 + \beta_5 x_4$
- $\frac{\Delta y}{\Delta x_4} = \frac{\partial y}{\partial x_4} = \beta_5 x_3$
- In models that include quadratic terms, interaction terms, or other nonlinear terms, the partial effect depends on the values of the independent variables and is not constant.

Partial effect at the mean and average partial effect

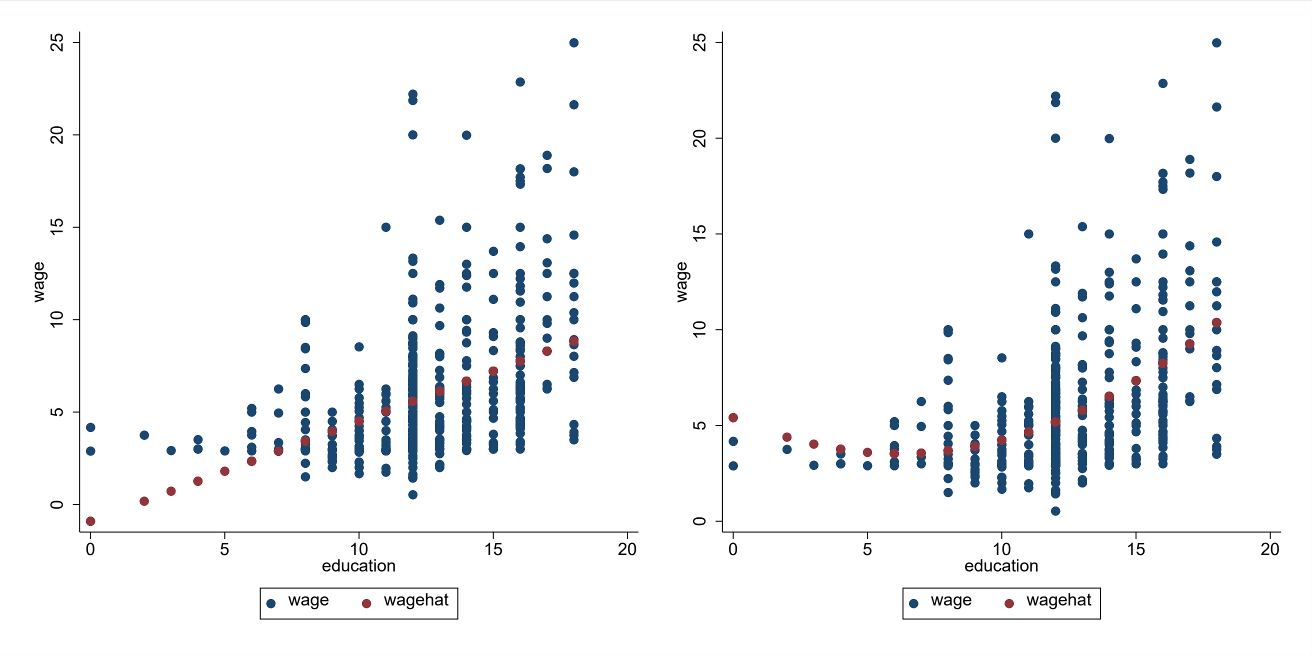
- A partial effect can be estimated for each observation i by substituting x_{ji}. $\frac{\Delta y}{\Delta x_2} = \frac{\partial y}{\partial x_2} = \beta_2 + 2\beta_3 x_{2i}$
- A partial effect can also be estimated for any value of the variable x_2.
- A partial effect at the mean substitutes x_{2i} with its mean $\bar{x}_2 = \frac{1}{n}\sum_{i=1}^{n} x_{2i}$, so PEM =$\beta_2 + 2\beta_3\bar{x}_2$. It explains the effect of the independent variable on the dependent variable at the mean of the independent variable.
- The average partial effect averages the partial effects across the i observations in the sample. APE = $\frac{1}{n}\sum_{i=1}^{n}(\beta_2+2\beta_3 x_{2i})$. It explains the average effect of the independent variable on the dependent variable.
- In this case, both methods will result in the same partial effect.

Regression with a quadratic term

- Regression model with a quadratic term allows for a nonlinear effect of an independent variable on the dependent variable.
- The model is still linear in parameters, but not in variables.
- Regression model with a quadratic term: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_1^2 + u$
- The marginal effect is $\frac{\Delta y}{\Delta x_1} = \beta_1 + 2\beta_2 x_1$
- The marginal effect is not a constant but depends on the variable x_1.
- To find the max/min point when the marginal effect changes from increasing to decreasing and vice versa, we can put this expression to zero: $x_1^* = -\beta_1 / 2\beta_2$.

Regression with a quadratic term

	Regression	Regression with quadratic term
VARIABLES	wage	wage
educ	0.541***	-0.607**
	(0.0532)	(0.241)
educsq		0.0491***
		(0.0101)
Constant	-0.905	5.408***
	(0.685)	(1.459)



The minimum marginal effect of education on experience is when it changes from decreasing to increasing at education $x_1^* = -\beta_1/2\beta_2$=-(-0.607)/(2*0.0491)=6.18

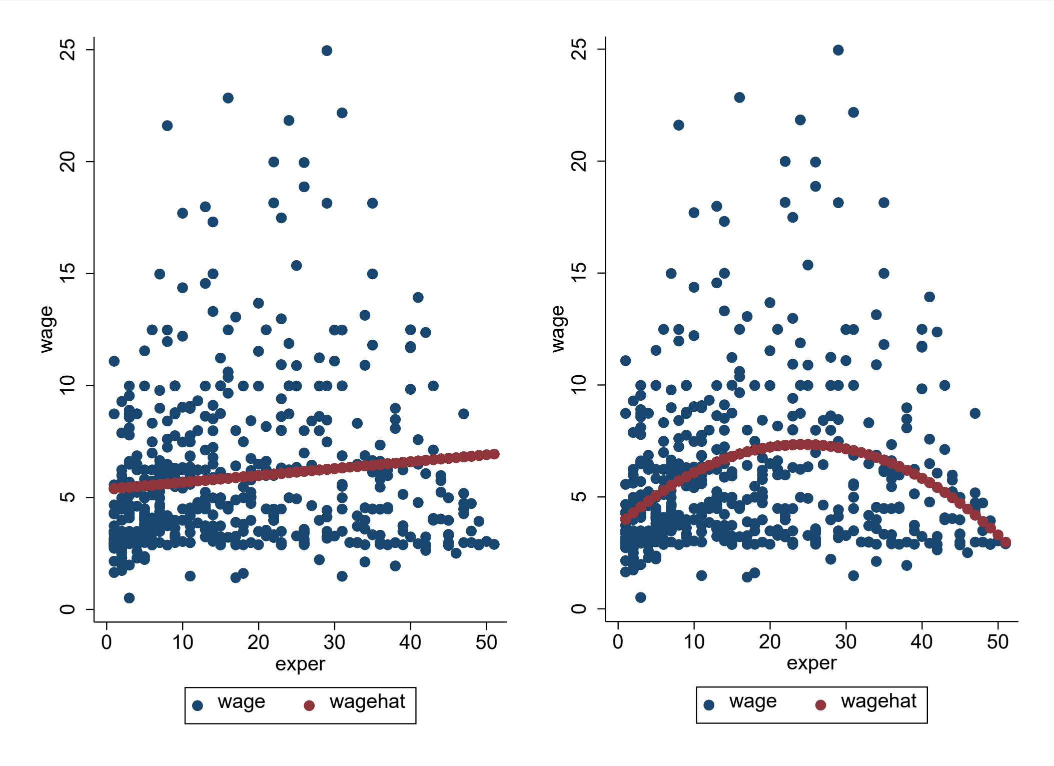
Marginal effects

Education	Marginal effect of education on wage calculation
5	-0.607+2*(0.0491)*5=-0.11
6.18 (turning point)	-0.607+2*(0.0491)*6.18=0
10	-0.607+2*(0.0491)*10=0.37
12.6 (mean)	-0.607+2*(0.0491)*12.6=0.63
15	-0.607+2*(0.0491)*15=0.86

- In the model of wage regressed on education, the marginal effect of education on wage (the coefficient) is constant at 0.54 (one additional year of education is associated with $0.54 increase in wage).
- In the model with quadratic term of education, the marginal effect of education on wage depends on education, marginal effect = $\frac{\Delta wage}{\Delta educ} = \beta_1 + 2\beta_2 educ = -0.607 + 2(0.0491)educ$.
- The average education is 12.6 years. The marginal effect at the mean is 0.63 (same as the average partial effect).
- At low levels of education (5 years), the marginal effect of education on wage is -0.11, but at high levels of education (15 years), the marginal effect of education on wage is 0.86.
- Higher education is associated with decreasing wages until education=6.18 years. After that, the effect of education on wage is increasing. In other words, we have positive and increasing returns to education.

Regression with a quadratic term

	Regression	Regression with quadratic term
VARIABLES	wage	wage
exper	0.0307***	0.298***
	(0.0118)	(0.0410)
expersq		-0.00613***
		(0.000903)
Constant	5.373***	3.725***
	(0.257)	(0.346)



The maximum effect of experience on wage is when marginal effect of experience on wage changes from increasing to decreasing at experience $x_1^* = -\beta_1/2\beta_2$= - 0.298/(2*(-0.00613)) = 24.3

Marginal effects

Experience	Marginal effect of experience on wage calculation
10	0.298+2*(-0.0061)*10 = 0.18
17 (mean)	0.298+2*(-0.0061)*17 = 0.09
20	0.298+2*(-0.0061)*20 = 0.05
24.3 (turning point)	0.298+2*(-0.0061)*24.3 = 0
30	0.298+2*(-0.0061)*30 = -0.07

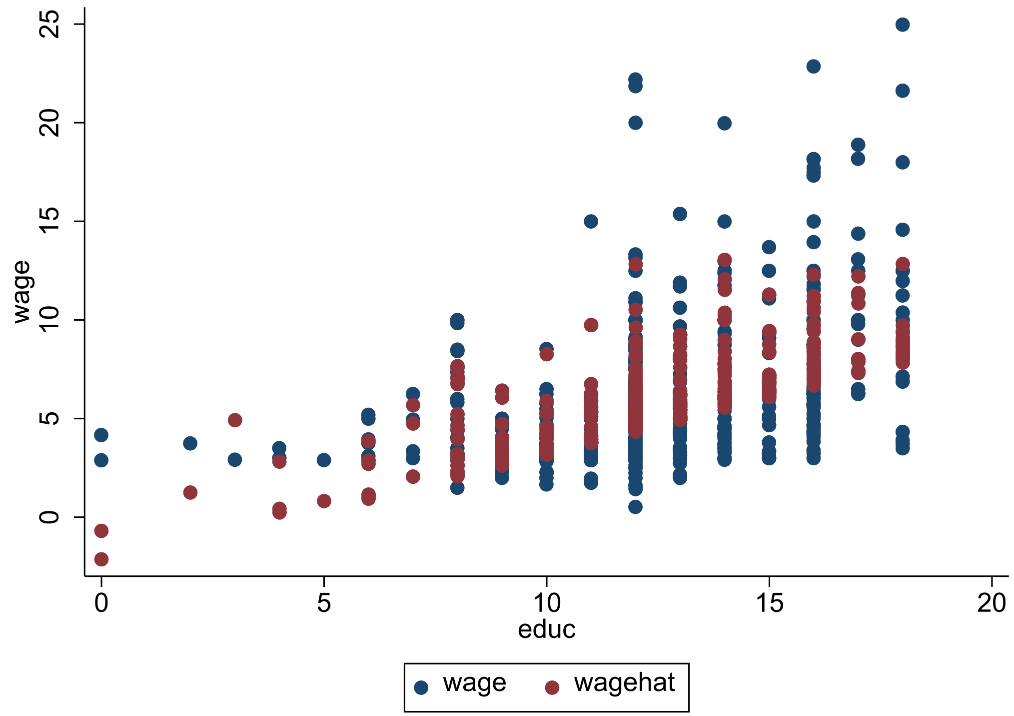
- In the model of wage regressed on experience, the marginal effect of experience on wage (the coefficient) is constant at 0.03.
- In the model with quadratic term of experience, the marginal effect of experience on wage depends on experience, marginal effect = $\frac{\Delta wage}{\Delta exper} = \beta_1 + 2\beta_2 exper = 0.298 + 2(-0.0061)\ exper$
- The average experience is 17 years. The marginal effect at the mean is 0.09 (same as the average partial effect).
- At low values of experience (10 years), the marginal effect of experience on wage is 0.18, but at high value of experience (30 years), the marginal effect of experience on wage is -0.07.
- Higher experience is associated with higher but decreasing wages until experience = 24.3 years. After that, the effect of experience on wages is negative.

Interaction terms

- Regression model with an interaction term allows for a nonlinear effect of an independent variable on the dependent variable.
- $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$
- The marginal effect is $\frac{\Delta y}{\Delta x_1} = \frac{\partial y}{\partial x_1} = \beta_1 + \beta_3 x_2$
- The marginal effect of x_1 on y also depends on x_2.
- If two positive individual effects also have a positive interaction effect, then they boost each other. If two positive individual effects have a negative interaction effect, then they work against each other.

Regression with interaction term - example

- Regression model with interaction term: $wage = \beta_0 + \beta_1 educ + \beta_2 exper + \beta_3 tenure + \beta_4 educ * exper + u$
- The marginal effect is $\frac{\Delta wage}{\Delta educ} = \beta_1 + \beta_4 exper$
- The marginal effect of education on wage depends on experience.



Regression with interaction terms

	Regression model	Regression model with interaction term	Regression model with interaction term
VARIABLES	wage	wage	wage
educ	0.599***	0.581***	0.564***
	(0.0513)	(0.0852)	(0.0509)
exper	0.0223*	0.0123	0.0452***
	(0.0121)	(0.0406)	(0.0128)
tenure	0.169***	0.169***	0.398***
	(0.0216)	(0.0217)	(0.0544)
educ x exper		0.000855	
		(0.00331)	
exper x tenure			-0.00734***
			(0.00161)
Constant	-2.873***	-2.653**	-2.995***
	(0.729)	(1.119)	(0.716)

Marginal effect of education on wage also depends on the value of experience.
The coefficient on the interaction term of education and experience is not significant,
but the coefficient on the interaction term of experience and tenure is negative and significant.

Marginal effects

exper	Marginal effect of education on wage calculation
10	0.581+0.0009*10=0.59
17 (mean)	0.581+0.0009*17=0.60
30	0.581+0.0009*30=0.61

- In the model with interaction term between education and experience, the marginal effect of education on wage depends on experience, marginal effect = $\frac{\Delta wage}{\Delta educ} = \beta_1 + \beta_4 exper = 0.581 + 0.0009\ exper$.
- The marginal effect for education depends on another variable experience.
- The average experience is 17 years. The marginal effect at the mean experience is 0.60 (same as the average partial effect).
- At low levels of experience (10 years), the marginal effect of education on wage is 0.59, but at high levels of experience (30 years), the marginal effect of education on wage is 0.61.

Rescaling variables

- A variable should be rescaled if its magnitude is too large or too small and the variable or the coefficient does not display well in a table.
- For example, a variable can be rescaled from dollars to thousands of dollars (by dividing by 1,000) or to millions of dollars (by dividing by 1,000,000).
- Regression equation: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$
- Rescaling a dependent variable
  - If y is multiplied by c then all $\hat{\beta}$ coefficients are multiplied by c

$$(c * y) = (c * \beta_0) + (c * \beta_1)x_1 + (c * \beta_2)x_2 + c * u$$

  - If the dependent variable is rescaled from thousands of dollars into dollars by multiplying it by 1,000, then all coefficients will increase by 1,000.
- Rescaling an independent variable
  - If x_j is multiplied by c then its coefficient $\hat{\beta}_j$ is divided by c
  - $y = \beta_0 + (\beta_1/c) * (c * x_1) + \beta_2 x_2 + u$
  - If an independent variable is rescaled from thousands of dollars into dollars by multiplying it by 1,000 then its coefficient will decrease by 1,000.

Rescaling variables – summary statistics

Variable	Units	Mean	Std. Dev.	Min	Max
salary	in thousand $	865 or 865.864	588	100	5299
salary_d	in $	865,864	587,589	100,000	5,299,000
sales	in million $	3,529 or 3,529.463	6,089	29	51,300
sales_k	in thousand $	3,529,463	6,088,654	29,000	5.13E+07
profits	in million $	208	404	-463	2,700

Salary is measured in thousand dollars and salary_d is measured in dollars. Salary_d=salary*1000
Sales is measured in million dollars and salary_k is measured in thousand dollars. Sales_k=sales*1000
Rescaling variables is a matter of preference. There is loss of information only if there is rounding.

Regressions with rescaled variables

VARIABLES	salary (thousand $)	salary_d ($)	salary (thousand $)	salary_d ($)
sales (mil $)	0.0175	17.49		
	(0.0111)	(11.09)		
sales_k (thous $)			1.75e-05	0.0175
			(1.11e-05)	(0.0111)
profits	0.362**	362.1**	0.362**	362.1**
	(0.167)	(166.9)	(0.167)	(166.9)
Constant	728.9***	728,872***	728.9***	728,872***
	(47.02)	(47,017)	(47.02)	(47,017)

If a dependent variable is multiplied by 1,000, then all coefficients are multiplied by 1,000.
If an independent variable is multiplied by 1,000, then only its coefficient is divided by 1,000.
Rescaling variables avoids the problem of the very small coefficient of 1.75e-05 in the third regression.

Log transformation (logged variables)

- Sometimes variables (y or x) are expressed as logs, log(y) or log(x)
- With logs, coefficient interpretation is in percentage/elasticity
- Variables measured in dollars can be logged
- Variables such as age and education that are measured in units such as years should not be logged
- Variables measured in percentage points (e.g. interest rates) should not be logged
- Logs cannot be used if variables have zero or negative values
- Taking logs often reduces problems with large values or outliers
- Taking logs helps with homoskedasticity and normality

Regressions with logged variables

VARIABLES	Linear form salary	Log-linear form lsalary	Linear-log form salary	Log-log form lsalary
sales	0.0175	2.56e-05**		
	(0.0111)	(1.13e-05)		
lsales			125.2***	0.194***
			(34.68)	(0.0340)
profits	0.362**	0.000287*	0.304**	0.000179
	(0.167)	(0.000171)	(0.123)	(0.000120)
Constant	728.9***	6.433***	-102.3	5.145***
	(47.02)	(0.0481)	(239.4)	(0.235)
Observations	177	177	177	177
R-squared	0.167	0.181	0.214	0.290

The log-log regression provides best fit (highest R-squared).

Rescaling logged variables

- Rescaling a logged dependent variable
  - If y is multiplied by c then all slope coefficients are the same but the intercept is higher by log(c)
  - $\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + u$
  - $\log(c * y) = \log(c) + \log(y) = [\log(c) + \ \beta_0] + \beta_1 x_1 + \beta_2 x_2 + u$
- Rescaling a logged independent variable
  - If x_j is multiplied by c then all slope coefficients are the same but the intercept is lower by $\beta_j \log(c)$
  - $y = \beta_0 + \beta_1 \log(x_1) + \beta_2 x_2 + u$
  - $y = [\beta_0 - \beta_1 \log(c)] \ + \beta_1 \log(c * x_1) + \beta_2 x_2 + u$
- Rescaling logged variables will not change the slope coefficients.

Rescaled logged variables – summary statistics

Variable	Units	Mean	Std. Dev.	Min	Max
lsalary	in thousand $	6.58	0.61	4.61	8.58
lsalary_d	in $	13.49	0.61	11.51	15.48
lsales	in million $	7.23	1.43	3.37	10.85
lsales_k	in thousand $	14.14	1.43	10.28	17.75
profits	in million $	207.83	404.45	-463.00	2700.00

lsalary is log of salary measured in thousand dollars and lsalary_d is log of salary measured in dollars. lsalary_d=log(salary_d)
lsales is log of sales measured in million dollars and lsalary_k is log of sales measured in thousand dollars. lsales_k=log(salary_k)

Regression with rescaled logged variables

VARIABLES	lsalary (thous $)	lsalary_d ($)	lsalary (thous $)	lsalary_d ($)
lsales (mil $)	0.194***	0.194***		
	(0.0340)	(0.0340)		
lsales_k (thous $)			0.194***	0.194***
			(0.0340)	(0.0340)
profits	0.000179	0.000179	0.000179	0.000179
	(0.000120)	(0.000120)	(0.000120)	(0.000120)
Constant	5.145***	12.05***	3.807***	10.71***
	(0.235)	(0.235)	(0.468)	(0.468)

Rescaling a dependent or independent variable that is logged does not change the slope coefficients, only the intercept.

Review questions

- Explain marginal or partial effects. Describe how the partial effect at the mean versus the average partial effect is calculated.
- How does adding a quadratic term change the marginal effect?
- How does adding an interaction term change the marginal effect?
- If a dependent or independent variables get rescaled, how do the coefficients change? What if the variables are in logs?