Solutions to HW2

Guanxi Li

May 4, 2024

1 P164.11

1.1

$$\mathcal{L}(\beta_1,\, \beta_2) = \min_{\beta_1,\, \beta_2} \sum_{i=1}^{n} (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i})^2 \tag{1}$$

1.2

$$\frac{\partial \sum (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{\partial \beta_1} = -2 \sum X_{1i}(Y_i - \beta_1 X_{1i} - \beta_2 X_{2i}) = 0 \tag{2}$$

$$\frac{\partial \sum (Y_i - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{\partial \beta_2} = -2 \sum X_{2i}(Y_i - \beta_1 X_{1i} - \beta_2 X_{2i}) = 0 \tag{3}$$

1.3

Proof. From 2 and 3,

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_{1i}Y_i - \hat{\beta}_2 \sum X_{1i}X_{2i}}{X_{1i}^2} \\ &= \frac{\sum X_{1i}Y_i}{X_{1i}^2}\end{aligned} \tag{4}$$

□

1.4

Proof. From 2 and 3,

$$\hat{\beta}_2 = \frac{\sum X_{1i}Y_i - \hat{\beta}_1 \sum X_{1i}X_{2i}}{X_{2i}^2} \tag{5}$$

From 4 and 5,

$$\hat{\beta}_1 = \frac{\sum X_{1i}Y_i - \frac{\sum X_{1i}Y_i - \hat{\beta}_1 \sum X_{1i}X_{2i}}{X_{2i}^2} \sum X_{1i}X_{2i}}{X_{1i}^2}$$

Solving for $\hat{\beta}_1$,

$$\hat{\beta}_1 = \frac{\sum X_{2i}^2 \sum X_{1i}Y_i - \sum X_{1i}X_{2i} \sum X_{2i}Y_i}{\sum X_{1i}^2 \sum X_{2i}^2 - \left(\sum X_{1i}X_{2i}\right)^2} \tag{6}$$

□

1.5

Proof. The objective function is

$$\mathcal{L}(\beta_0,\, \beta_1,\, \beta_2) = \min_{\beta_0,\, \beta_1,\, \beta_2} \sum_{i=1}^{n} (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2$$

The partial deviation with respect to β_0 is

$$\frac{\partial \sum (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{\partial \beta_0} = -2 \sum (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}) = 0$$

Solving for β_0,

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}_1 - \hat{\beta}_2 \bar{X}_2 \tag{7}$$

□

1.6

Proof. The partial deviation with respect to β_1 is

$$\frac{\partial \mathcal{L}(\beta_0,\, \beta_1,\, \beta_2)}{\partial \beta_1} = \frac{\partial \sum (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{\partial \beta_1} = -2 \sum X_{1i} (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}) = 0 \tag{8}$$

The partial deviation with respect to β_2 is

$$\frac{\partial \mathcal{L}(\beta_0,\, \beta_1,\, \beta_2)}{\partial \beta_2} = \frac{\partial \sum (Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i})^2}{\partial \beta_2} = -2 \sum X_{2i}(Y_i - \beta_0 - \beta_1 X_{1i} - \beta_2 X_{2i}) = 0 \tag{9}$$

From 7, 8 and 9,

$$\frac{\partial \mathcal{L}(\beta_0,\, \beta_1,\, \beta_2)}{\partial \beta_1} = -2 \sum X_{1i}(Y_i - \bar{Y} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) = 0 \tag{10}$$

$$\frac{\partial \mathcal{L}(\beta_0,\, \beta_1,\, \beta_2)}{\partial \beta_2} = -2 \sum X_{2i}(Y_i - \bar{Y} + \hat{\beta}_1 \bar{X}_1 + \hat{\beta}_2 \bar{X}_2 - \hat{\beta}_1 X_{1i} - \hat{\beta}_2 X_{2i}) = 0 \tag{11}$$

From 10 and 11,

$$\hat{\beta}_1 = \frac{\sum X_{1i}(Y_i - \bar{Y}) - \hat{\beta}_2 \sum X_{1i}(X_{2i} - \bar{X}_2)}{\sum X_{1i}(X_{1i} - \bar{X}_1)} \tag{12}$$

$$\hat{\beta}_2 = \frac{\sum X_{2i}(Y_i - \bar{Y}) - \hat{\beta}_1 \sum X_{2i}(X_{1i} - \bar{X}_1)}{\sum X_{2i}(X_{2i} - \bar{X}_2)} \tag{13}$$

From 12 and 13,

$$\begin{aligned}
\hat{\beta}_1 &= \frac{\sum X_{1i}(Y_i - \bar{Y}) - \frac{\sum X_{2i}(Y_i - \bar{Y}) - \hat{\beta}_1 \sum X_{2i}(X_{1i} - \bar{X}_1)}{\sum X_{2i}(X_{2i} - \bar{X}_2)} \sum X_{1i}(X_{2i} - \bar{X}_2)}{\sum X_{1i}(X_{1i} - \bar{X}_1)} \\
&= \frac{\sum X_{1i}(Y_i - \bar{Y}) \sum X_{2i}(X_{2i} - \bar{X}_2) - \sum X_{2i}(Y_i - \bar{Y}) \sum X_{1i}(X_{2i} - \bar{X}_2)}{\sum X_{1i}(X_{1i} - \bar{X}_1) \sum X_{2i}(X_{2i} - \bar{X}_2) - \sum X_{2i}(X_{1i} - \bar{X}_1) \sum X_{1i}(X_{2i} - \bar{X}_2)}
\end{aligned} \tag{14}$$

From

$$\begin{aligned}\sum X_{1i}(X_{2i} - \bar{X}_2) &= \sum X_{1i}(X_{2i} - \bar{X}_2) + \sum(X_{2i} - \bar{X}_2) \\ &= \sum X_{1i}(X_{2i} - \bar{X}_2) + \sum \bar{X}_1(X_{2i} - \bar{X}_2) \\ &= \sum(X_{1i} - \bar{X}_1)(X_{2i} - \bar{X}_2) = 0\end{aligned}$$

Thus, 14 can be expressed as

$$\begin{aligned}\hat{\beta}_1 &= \frac{\sum X_{1i}(Y_i - \bar{Y}) \sum X_{2i}(X_{2i} - \bar{X}_2)}{\sum X_{1i}(X_{1i} - \bar{X}_1) \sum X_{2i}(X_{2i} - \bar{X}_2)} \\ &= \frac{\sum X_{1i}(Y_i - \bar{Y})}{\sum X_{1i}(X_{1i} - \bar{X}_1)} \\ &= \frac{\sum(X_{1i} - \bar{X}_1)(Y_i - \bar{Y})}{\sum(X_{1i} - \bar{X}_1)(X_{1i} - \bar{X}_1)} \end{aligned} \tag{15}$$

□

Replace X_{1i} with X_i in 15 to obtain the OLS regression coefficient with only one explanatory variable.

2 P190.3

2.1

Yes. The t-statistic for coefficient Age is $0.29/0.04 = 7.25 > 1.96$, which means it is significant at 1% level.

2.2

$\Delta \text{Income} = \Delta \text{Age} \times [0.29 \pm 1.96 \times 0.04] = [1.06, 1.84]$

3 P190.4

3.1

The F-statistic for regional coefficients is $6.1 > F_{3,\infty} = 3.78$, which means it is significant at 1% level.

3.2

3.2.1

The confidence interval is $[-0.27 - 1.96 \times 0.26,\ -0.27 + 1.96 \times 0.26] = [-0.78,\ 0.24]$.

3.2.2

The confidence interval is $[(-0.6 - 0.27) - 1.96 \times \sqrt{0.28^2 + 0.26^2},\ (-0.6 - 0.27) + 1.96 \times \sqrt{0.28^2 + 0.26^2}] = [-0.61,\ -0.13]$. The confidence interval can be computed directly from the coefficient and the standard deviation of $South$ when replacing $Midwest$ with $West$.

4 P191.8

4.1

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1}(1-R^2),$$

Thus

$$R^2 = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2)$$

.

$$R^2_{(1)} = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2) = 1 - \frac{n-k-1}{n-1}(1-0.049^2) = 0.051$$
$$R^2_{(2)} = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2) = 1 - \frac{n-k-1}{n-1}(1-0.424^2) = 0.427$$
$$R^2_{(3)} = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2) = 1 - \frac{n-k-1}{n-1}(1-0.773^2) = 0.775$$
$$R^2_{(4)} = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2) = 1 - \frac{n-k-1}{n-1}(1-0.626^2) = 0.629$$
$$R^2_{(5)} = 1 - \frac{n-k-1}{n-1}(1-\bar{R}^2) = 1 - \frac{n-k-1}{n-1}(1-0.773^2) = 0.775$$

4.2

$$R^2_{unrestricted} = 0.775$$

$$R^2_{restricted} = 0.427$$

$k_{unrestricted} = 4$, $q = 2$,

$$F = \frac{(R^2_{\text{unrestricted}} - R^2_{\text{restricted}})/q}{(1 - R^2_{\text{unrestricted}})/(n - k_{\text{unrestricted}} - 1)} = 320.93 > F^{1\%}_{2,\infty} = 4.61$$

4.3

The confidence interval is $[-1.01 \pm 2.58 \times 0.27] = [-1.71, -0.31]$.

Table 1: Main Results.

VARIABLES	(1) a1 growth	(2) a2 growth	(3) a3 growth	(4) a4 growth	(5) a5 growth
tradeshare	1.898**	1.749**	1.104	1.883	-5.702
	(0.866)	(0.794)	(0.748)	(1.344)	(8.183)
lnyearsschool		1.016***	2.161***	2.525***	2.133***
		(0.204)	(0.400)	(0.639)	(0.416)
rev_coups			-2.300**	-2.350**	-2.035**
			(0.921)	(0.921)	(0.938)
assasinations			0.228	0.224	0.102
			(0.335)	(0.337)	(0.364)
lnrgdp60			-1.621***	-1.641***	-1.584***
			(0.440)	(0.440)	(0.477)
yearsschool	0.243***				
	(0.0759)				
trdslnys				-0.690	
				(0.809)	
tradeshare2					8.488
					(15.83)
tradeshare3					-2.760
					(8.527)
Constant	-0.122	-0.186	11.75***	11.50***	12.93***
	(0.691)	(0.566)	(3.297)	(3.345)	(3.055)
Observations	64	64	64	64	64
R-squared	0.161	0.287	0.453	0.457	0.471

* Significant at the 10 percent level.
** Significant at the 5 percent level.
*** Significant at the 1 percent level.

5

5.1

Figure 1 and 2 depicts a nonlinear relationship between *growth* and *yearsschool*, which is why nonlinear regression (3) works better.

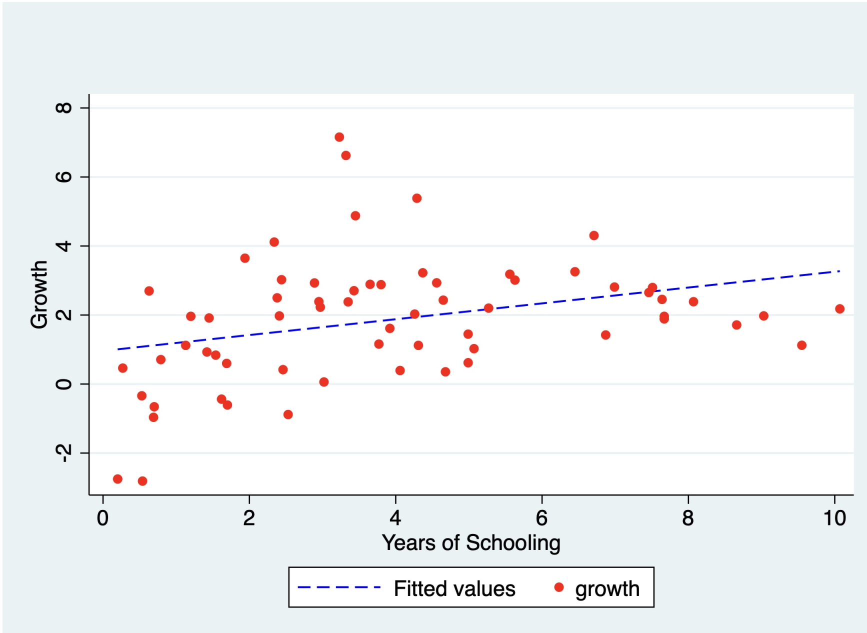

Figure 1: Scatter Plot Results 1

Notes: This figure depicts the scatter plot results of $growth$ and $yearsschool$.

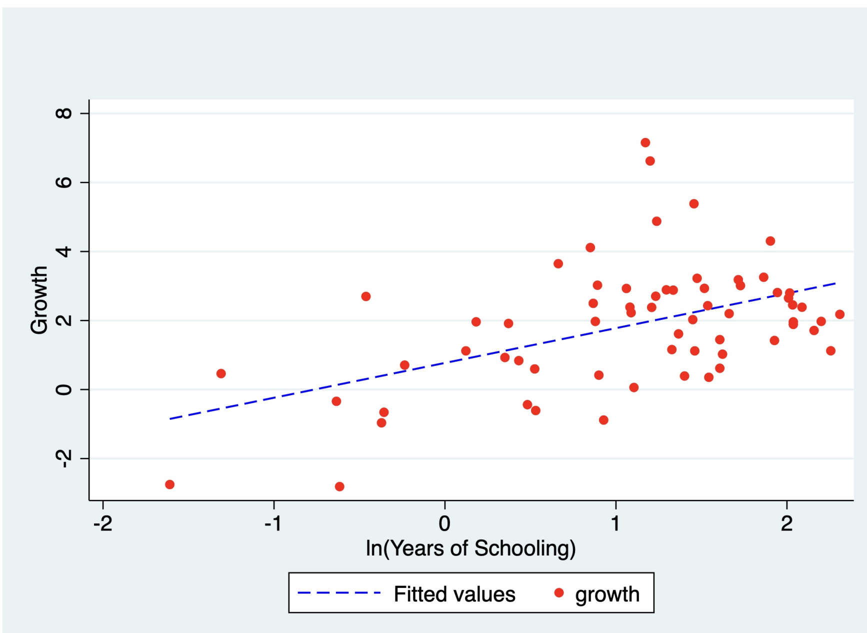

Figure 2: Scatter Plot Results 2

Notes: This figure depicts the scatter plot results of $growth$ and $ln(yearsschool)$.

5.2

The result of regression (1): growth $= 0.243 \times 2 = 0.49$, The result of regression (2): growth $= 1.016 \times ln(6/4) = 0.41$.

5.3

The coefficient of Rev_Coups is significant at 5% level, while the coefficient of $Assasinationa$ is not significant at 10% level.

5.4

The coefficient of $Tradeshare\_ln(YearsSchool)$ is not significant at 10% level.

5.5

The F-statistic for the joint test of the coefficients of quadratic and cubic $tradeshare$ is 1.96, which means the coefficients are not significant at 10% level.

5.6

The result of regression (3): growth $= 1.104 \times 0.5 = 0.55$, The result of regression (5): growth $= -5.702 \times 0.5 + 8.488 \times (1 - 0.5^2) - 2.760 \times (1 - 0.5^3) = 1.1$.