ECOM20001: Econometrics 1

Tutorial 11: Suggested Solutions

Logarithmic Regressions

1. See the tute11.R code for the construction of the logarithmic variables.
2. The table below contains the regression results:

	Dependent variable:			
	AHE		Log(AHE)	
	(1)	(2)	(3)	(4)
Age	0.541***		0.026***	
	(0.026)		(0.001)	
Log(Age)		15.951***		0.780***
		(0.759)		(0.039)
Bachelor Degree	7.796***	7.794***	0.411***	0.410***
	(0.153)	(0.153)	(0.008)	(0.008)
Female	-3.530***	-3.530***	-0.178***	-0.178***
	(0.145)	(0.145)	(0.008)	(0.008)
1992 Dummy	0.294**	0.292**	0.038***	0.038***
	(0.147)	(0.147)	(0.008)	(0.008)
Constant	1.124	-36.834***	1.924***	0.066
	(0.777)	(2.567)	(0.041)	(0.134)
Observations	15,052	15,052	15,052	15,052
R2	0.188	0.188	0.189	0.190
Adjusted R2	0.187	0.188	0.189	0.189
Residual Std. Error (df = 15047)	8.985	8.984	0.462	0.462
F Statistic (df = 4; 15047)	869.114***	870.195***	879.134***	880.735***
Note:		*p<0.1; **p<0.05; ***p<0.01		

Interpreting the age-related coefficients in the columns:

- Column (1), Linear-Linear: a one year increase in age is associated with a 0.541= $0.54 increase in average hourly earnings, ahe, a relationship that is statistically significant at the 5% level.

- Column (2), Linear-Log: a one percent increase in age is associated with a 0.01 x 15.951 = $0.16 increase in average hourly earnings, ahe, a relationship that is statistically significant at the 5% level.
- Column (3), Log-Linear: a one year increase in age is associated with a 100 x 0.026 = 2.6 percent increase in average hourly earnings, ahe, a relationship that is statistically significant at the 5% level.
- Column (4), Log-Log: a one percent increase in age is associated with a 0.78 percent increase in average hourly earnings, ahe, a relationship that is statistically significant at the 5% level. That is, the elasticity of average hourly earnings ahe with respect to age is 0.78.

Interactions

3. See the tute11.R code for the construction of the interaction variables.
4. The table below contains the regression results:

	Dependent variable: AHE (1)	Dependent variable: AHE (2)
Age	0.648***	0.541***
	(0.036)	(0.026)
Bachelor Degree	7.773***	7.769***
	(0.153)	(0.222)
Female	3.977***	-3.560***
	(1.501)	(0.162)
Female x Age	-0.253***	
	(0.051)	
Female x Bachelor		0.062
		(0.296)
1992 Dummy	0.292**	0.295**
	(0.147)	(0.147)
Constant	-2.051*	1.129
	(1.080)	(0.778)
Observations	15,052	15,052
R2	0.189	0.188
Adjusted R2	0.189	0.187
Residual Std. Error (df = 15046)	8.979	8.986
F Statistic (df = 5; 15046)	700.963***	695.255***
Note:		*p<0.1; **p<0.05; ***p<0.01

Interpreting any age, female, bachelor-related coefficients in the columns, remembering that average hourly earnings:

- Regression 5 (or column (1) in the table) partial effects of interest:
  - Being 1 year older (change in age =1) yields a (0.648 - female x 1 x 0.253) increase in average household earnings
    - If female (female=1): 1 year older yields a (0.648 - 1 x 1 x 0.253) =$0.39 increase in ahe
    - If male (female=0): 1 year older yields a(0.648 - 0 x 1 x 0.253)=$0.65 increase in ahe
  - Being female (female=1) yields a (3.977 - 1 x age x 0.253) change in average household earnings
    - if 25 years old (age=25), being female yields a (3.977 - 1 x 25 x 0.253)= -$2.35 (decrease) in earnings in ahe
    - if 30 years old (age=30), being female yields a (3.977 - 1 x 30 x 0.253)= -$3.61 (decrease) in earnings in ahe
  - Having a bachelor's (bachelor=1) degree yields a $7.773=$7.77 increase in ahe
- Regression 6 (or column (2) in the table) partial effects of interest, remembering that average hourly earnings:
  - Being 1 year older (change in age=1) yields a $0.541 = $0.54 increase in ahe
  - Having a bachelor's degree (bachelor=1) yields a partial effect of $(7.769 + female x 1 x 0.062)
    - if female (female=1), having a bachelor degree yields a partial effect of $(7.769 + 1 x 1 x 0.062)=$7.83 increase in ahe relative to someone without a bachelor's degree.
    - if male (female=0), having a bachelor degree yields a partial effect of $(7.769 + 0 x 1 x 0.062)=$7.77 increase in ahe relative to someone without a bachelor's degree.
    - <u>Note</u>: the 0.062 coefficient on female_bachelor in column (2) of the table above is statistically insignificant, which implies males and females do not have statistically significantly different earnings gains from earning a bachelor's degree.

- Being female (female=1) yields a partial effect of \$(-3.560+0.062 x 1 x bachelor) on average hourly earnings
  - if someone has a bachelor's degree (bachelor=1), the partial effect of being female on earnings is \$(-3.560+0.062 x 1 x 1)=-\$3.50 (decrease) in ahe relative to being male.
  - if someone does not have a bachelor's degree (bachelor=0), the partial effect of being female on earnings is \$(-3.560+0.062 x 1 x 0)=-\$3.56 (decrease) in ahe relative to being male.

5. We follow the general approach to computing standard errors for partial effects with nonlinear models that is covered in slides 19-23 of the lecture note 8. In Regression 5, the estimated regression equation's predicted value can be written out as follows:

$$\widehat{Y} = \hat{\beta}_0 + \hat{\beta}_1 age + \hat{\beta}_2 bachelor + \hat{\beta}_3 female + \hat{\beta}_4 female \times age + \hat{\beta}_5 d1992$$

Using this equation, the partial effect on the outcome variable (Y=ahe) from changing female=0 (individual is male) to female=1 (individual is female) at age=28 is:

$$\begin{aligned}\Delta\widehat{Y} &= \left(\hat{\beta}_0 + \hat{\beta}_1 28 + \hat{\beta}_2 bachelor + \hat{\beta}_3 1 + \hat{\beta}_4 1 \times 28 + \hat{\beta}_5 d1992\right) - \left(\hat{\beta}_0 + \hat{\beta}_1 28 + \hat{\beta}_2 bachelor + \hat{\beta}_3 0 + \hat{\beta}_4 0 \times 28 + \hat{\beta}_5 d1992\right)\\ &= \hat{\beta}_3 + 28\hat{\beta}_4\end{aligned}$$

Given our estimates for the regression coefficients from column (1) in the table in question 4 in the second line of these equations, we obtain a partial effect of being female at age 28 on, ahe, of 3.997-28 x 0.253=-3.087 (note: in the tute11.R code the partial effect is computed as -3.106 because of rounding). This implies a partial effect of \$3.09 per hour in average hourly earnings.

Following the general approach to computing standard errors for nonlinear effects on slides 16-24 of lecture note 8, we obtain the F-statistic from the following joint test based on the partial effect for Y we just derived:

$$H_0 : \beta_3 + 28\beta_4 = 0; \quad H_1 : \beta_3 + 28\beta_4 \neq 0$$

Doing so yields an F-statistic of 386.258 (see lines 106 and 107 of tute11.R). The standard error for the partial effect is then computed as:

$$SE(\hat{Y}) = \frac{|\Delta \hat{Y}|}{\sqrt{F}} = \frac{|-3.087|}{\sqrt{386.258}} = 0.157$$

(note: in the tute11.R code the SE is computed as 0.158 because of rounding).

The 95% CI of the partial effect of being female at age 28 on earnings, ahe, is then computed as [-3.087-1.96*0.157,-3.087+1.96*0.157]=[-3.395,-2.779]. (note: in the tute11.R code the 95% CI is [-3.416, -2.796] because of rounding). This implies a 95% CI of [-$3.39,-$2.78] for the partial effect of being female at age 28 on average hourly earnings.

Combining Logarithmic Regression and Interactions

6. See the tute11.R code for the construction of the logarithmic variables with interactions.
7. Regression results are presented in the table below

	Dependent variable: Log(AHE) (1)	(2)
Age	0.031*** (0.002)	
Age x Female	-0.012*** (0.003)	
Log(Age)		0.924*** (0.053)
Log(Age) x Female		-0.341*** (0.079)
Female	0.175** (0.080)	0.975*** (0.268)
Bacelor Degree	0.409*** (0.008)	0.409*** (0.008)
1992 Dummy	0.038*** (0.008)	0.038*** (0.008)
Constant	1.775*** (0.054)	-0.422** (0.178)
Observations	15,052	15,052
R2	0.190	0.191
Adjusted R2	0.190	0.190
Residual Std. Error (df = 15046)	0.462	0.462
F Statistic (df = 5; 15046)	708.043***	709.088***
Note:		*p<0.1; **p<0.05; ***p<0.01

Interpreting any age, female, bachelor-related coefficients in the columns:

- Regression 7 (or column (1) in the table with a log-linear specification) partial effects of interest:
  - Being 1 year older (change in age=1) yields a 100 x (0.031 - female x 1 x 0.012) percent change in average hourly earnings, ahe
    - If female (female=1): 1 year older yields a 100 x (0.031 - 1 x 0.012)= 1.90% increase in average hourly earnings, ahe
    - If male (male=1): 1 year older yields a 100 x (0.031 - 0 x 0.012)= 3.10% increase in average hourly earnings, ahe
  - Being female (female=1) yields a partial effect of 100 x (0.175 - 1 x age x 0.012)
    - if someone is 25 years old (age=25), the partial effect of being female on earnings (female=1) is 100 x (0.175 - 1 x 25 x 0.012) =-12.5% reduction in earnings, ahe relative to 25 year old males
    - if someone is 35 years old (age=35), the partial effect of being female on earnings (female=1) is 100 x (0.175 - 1 x 35 x 0.012) =-24.5% reduction in earnings, ahe relative to 35 year old males
  - Having a bachelor's degree (bachelor=1) has a partial effect of 40.9% increase in earnings, ahe, holding fixed age, gender, and year.
- Regression 8 (or column (1) in the table with a log-log specification) partial effects (or in this case, elasticities) of interest:
  - Being 1 year older (change in age=1) yields a 0.924 - 0.341 x female percent change in average hourly earnings, ahe
    - If female (female=1): 1% increase in age yields a 0.924 - 0.341 x 1 = 0.583% increase in average hourly earnings, ahe. That is, if someone is female, the elasticity of earnings with respect to age is 0.583.
    - If male (female=0): 1% increase in age yields a 0.924 - 0.341 x 0 = 0.924% increase in average hourly earnings, ahe. That is, if someone is male, the elasticity of earnings with respect to age is 0.924.

8. The predicted value from running the regression in Regression 8 can be written as follows:

$$\widehat{Y} = \widehat{\log(ahe)} = \hat{\beta}_0 + \hat{\beta}_1 log(age) + \hat{\beta}_2 female \times \log(age) + \hat{\beta}_3 female + \hat{\beta}_4 bachelor + \hat{\beta}_5 d1992$$

Using this log-log equation, when the individual is female (female=1), the predicted value becomes:

$$\begin{aligned}\widehat{\log(ahe)} &= \hat{\beta}_0 + \hat{\beta}_1 log(age) + \hat{\beta}_2 \times 1 \times \log(age) + \hat{\beta}_3 \times 1 + \hat{\beta}_4 bachelor + \hat{\beta}_5 d1992 \\ &= \hat{\beta}_0 + (\hat{\beta}_1 + \hat{\beta}_2) log(age) + \hat{\beta}_3 + \hat{\beta}_4 bachelor + \hat{\beta}_5 d1992\end{aligned}$$

which implies that the elasticity of ahe the with respect to age for females is the sum of the coefficients on log(age). Computing this partial effect/elasticity based on the estimated coefficients from column (2) of the table in question 7 we obtain an elasticity of: 0.924-0.341=0.583.

Again following the general approach to computing standard errors for nonlinear effects on slides 16-24 of lecture note 8, we can obtain the F-statistic from the following joint test based on the elasticity for females that we just derived:

$$H_0 : \beta_1 + \beta_2 = 0; \;\; H_1 : \beta_1 + \beta_2 \neq 0$$

This test yields an F-statistic of 97.439 (see lines 158 and 159 of tute11.R). The standard error for the partial effect/elasticity for females is then computed as:

$$SE(\%\Delta\hat{Y}) = \frac{|\%\Delta\hat{Y}|}{\sqrt{F}} = \frac{|0.583|}{\sqrt{97.439}} = 0.059$$

which implies that the 95% CI for the elasticity of ahe the with respect to age for females is [0.583-1.96 x 0.059 , 0.583-1.96 x 0.059] = [0.467, 0.699].

<u>Note</u>: we work with percent changes in Y and not just changes in Y in computing the standard error because of the interpretation of the log-log specification; the partial effects are in terms of the percentage change in Y (ahe) associated with a percentage change in X (age).