Lab 10 – OLS Regressions II

This lab will cover how to perform a simple OLS regression using different functional forms.

LAB 10 QUICK VIEW

- Non-linear relationships between variables include:
  - Log-Lin: $\ln(Y_t) = \beta_1 + \beta_2 X_t + \varepsilon_t$
  - Lin-Log: $Y_t = \beta_1 + \beta_2 \ln(X_t) + \varepsilon_t$
  - Log-Log: $\ln(Y_t) = \beta_1 + \beta_2 \ln(X_t) + \varepsilon_t$
  - Reciprocal: $Y_t = \beta_1 + \frac{\beta_2}{X_t} + \varepsilon_t$
  - Simple Quadratic: $Y_t = \beta_1 + \beta_2 X_t^2 + \varepsilon_t$
  - Simple Power: $Y_t = \beta_1 + \beta_2 X_t^A + \varepsilon_t$, where A is any power
  - Square Root: $Y_t = \beta_1 + \beta_2 \sqrt{X_t} + \varepsilon_t$
  - Other Root: $Y_t = \beta_1 + \beta_2 \sqrt[A]{X_t} + \varepsilon_t$, where A represents the A-th root.

- Useful excel formulas to modify data to reflect the above relationships include:
  - =ln(CELL) (to take the natural log of a variable)
  - =CELL^2 (to square a variable)
  - =CELL^A (to raise a variable to the power A)
  - =sqrt(CELL) (to take a square root of a variable)
  - =CELL^1/A (to take the A-th root of a variable)

- Proper functional form can be determined by plotting data (see scatter graphs and trend lines in chapter 4) and referring to existing economic research and theory.

Written by Lorne Priemaza

A) Different Functional Forms

As introduced in class, the relationship between two variables can be expressed in many non-linear ways. Often the relationship between two variables can be estimated by plotting those two variables data (see scatter graphs and trend lines in chapter 4) or referring to existing research and theory.

- Log-Lin: $\ln(Y_t) = \beta_1 + \beta_2 X_t + \varepsilon_t$
- Lin-Log: $Y_t = \beta_1 + \beta_2 \ln(X_t) + \varepsilon_t$
- Log-Log: $\ln(Y_t) = \beta_1 + \beta_2 \ln(X_t) + \varepsilon_t$
- Reciprocal: $Y_t = \beta_1 + \frac{\beta_2}{X_t} + \varepsilon_t$
- Simple Quadratic: $Y_t = \beta_1 + \beta_2 X_t^{\,2} + \varepsilon_t$
- Simple Power: $Y_t = \beta_1 + \beta_2 X_t^{\,A} + \varepsilon_t$, where A is any power
- Square Root: $Y_t = \beta_1 + \beta_2 \sqrt{X_t} + \varepsilon_t$
- Other Root: $Y_t = \beta_1 + \beta_2 \sqrt[A]{X_t} + \varepsilon_t$, where A represents the A-th root.
- Plus others

B) Adjusting variables

Both the X and Y variables can be modified to fit the above functional forms using the following equations (Where CELL refers to the excel cell location of the data):

=ln(CELL)	(to take the natural log of a variable)
=CELL^2	(to square a variable)
=CELL^A	(to raise a variable to the power A)
=sqrt(CELL)	(to take a square root of a variable)
=CELL^1/A	(to take the A-th root of a variable)

C) OLS

An OLS estimation of the non-linear relationship between two variables can be run by choosing “Data Analysis” as seen in lab 9, then selecting the new variables created using the formulas under B above.

D) Trend Lines

One way to see if a non-linear functional form is required is to plot your data using a TREND LINE to do this, select your data to plot using a “Scatter Graph/Chart” and under “Chart Tools” and “Design” select “Add Chart Element -> Trendline - > Linear”. If the linear trend line does not fit the data, another functional form may be more appropriate (See Practice Lab for an Example).

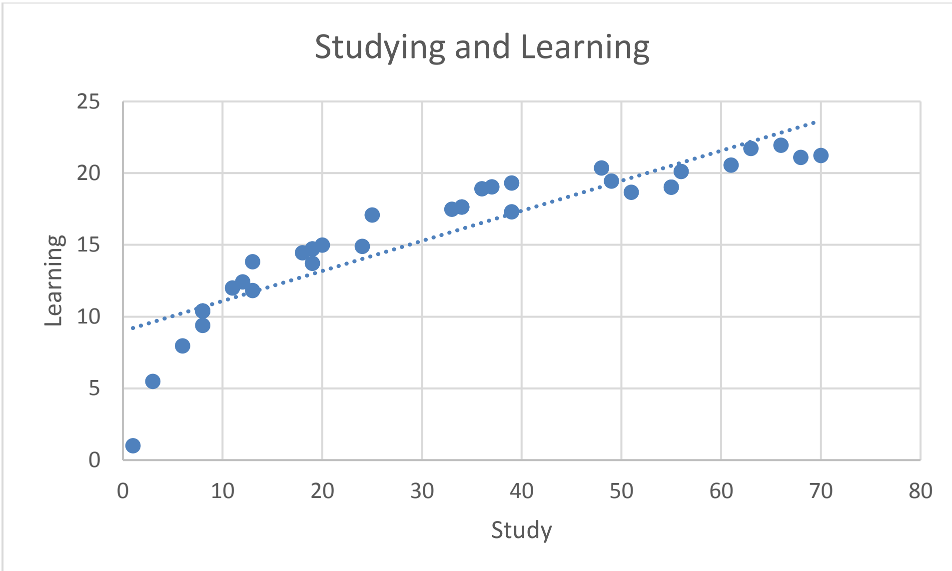
Econ 299 Practice Lab 10

A) Download (highly recommended) or copy the following Data Sets into excel:

A	B
Study	Learning
55	19.03667
13	11.82475
37	19.05459
56	20.12676
55	19.03667
33	17.48254
6	7.958797
20	14.97866
11	11.98948
36	18.91759
63	21.71567
34	17.6318
51	18.65913
19	14.72219
70	21.24248
8	10.39721
18	14.45186
19	13.72219
49	19.4591
48	20.35601
61	20.55437
13	13.82475
24	14.89027
66	21.94827
39	17.31781
12	12.42453
8	10.39721
34	17.6318
19	14.72219
39	19.31781
8	10.39721
68	21.09754
8	9.397208
25	17.09438
1	0
3	5.493061

Written by Lorne Priemaza

B) Insert a Scatter Plot of these variables, with "Study" as your explanatory (x) variable. (See lab 4 for a refresher if needed.) Inserting a linear trend line, we see that a linear estimation isn't the best estimation:



C) In column C, calculate the square root of the Study variable, along with an appropriate title.

D) In column D, calculate the natural log of the Learning variable, along with an appropriate title.

E) In column E, calculate the inverse of the Study variable, along with an appropriate title.

F) In column F, calculate the natural log of the Study variable, along with an appropriate title.

G) Use the "Regression" option under "Data Analysis" to estimate the real lin-log function $L_i = \beta_1 + \beta_2 \ln(S_i) + \varepsilon_i$. Insert an appropriate title and report the OLS estimation. (For a real lab assignment, this can be written by hand afterwards.)

H) Use the "Regression" option under "Data Analysis" to estimate the real log-lin function $\ln(L)_i = \beta_1 + \beta_2 S_i + \varepsilon_i$. Insert an appropriate title and report the OLS estimation. (For a real lab assignment, this can be written by hand afterwards.)

I) Use the "Regression" option under "Data Analysis" to estimate the real log-log function $\ln(L)_i = \beta_1 + \beta_2 \ln(S_i) + \varepsilon_i$. Insert an appropriate title and report the OLS estimation. (For a real lab assignment, this can be written by hand afterwards.)

J) Use the "Regression" option under "Data Analysis" to estimate the real reciprocal function $L_i = \beta_1 + {\beta_2}/{S_i} + \varepsilon_i$. Insert an appropriate title and report the OLS estimation. (For a real lab assignment, this can be written by hand afterwards.)

K) Use the "Regression" option under "Data Analysis" to estimate the real square root function $L_i = \beta_1 + \beta_2 \sqrt{S_i} + \varepsilon_i$. Insert an appropriate title and report the OLS estimation. (For a real lab assignment, this can be written by hand afterwards.)

Written by Lorne Priemaza

Econ 299 Practice Lab 10 Answers

A	B	C	D	E	F
Study	Learning	Root(Study)	Ln(Learn)	1/Study	Ln(Study)
55	19.03667	7.41619849	2.946367	0.018182	4.007333
13	11.82475	3.60555128	2.470195	0.076923	2.564949
37	19.05459	6.08276253	2.947308	0.027027	3.610918
56	20.12676	7.48331477	3.00205	0.017857	4.025352
55	19.03667	7.41619849	2.946367	0.018182	4.007333
33	17.48254	5.74456265	2.861203	0.030303	3.496508
6	7.958797	2.44948974	2.074278	0.166667	1.791759
20	14.97866	4.47213595	2.706627	0.05	2.995732
11	11.98948	3.31662479	2.484029	0.090909	2.397895
36	18.91759	6	2.940092	0.027778	3.583519
63	21.71567	7.93725393	3.078034	0.015873	4.143135
34	17.6318	5.83095189	2.869704	0.029412	3.526361
51	18.65913	7.14142843	2.926335	0.019608	3.931826
19	14.72219	4.35889894	2.689356	0.052632	2.944439
70	21.24248	8.36660027	3.056003	0.014286	4.248495
8	10.39721	2.82842712	2.341537	0.125	2.079442
18	14.45186	4.24264069	2.670823	0.055556	2.890372
19	13.72219	4.35889894	2.619015	0.052632	2.944439
49	19.4591	7	2.968315	0.020408	3.89182
48	20.35601	6.92820323	3.013376	0.020833	3.871201
61	20.55437	7.81024968	3.023074	0.016393	4.110874
13	13.82475	3.60555128	2.62646	0.076923	2.564949
24	14.89027	4.89897949	2.700708	0.041667	3.178054
66	21.94827	8.1240384	3.088688	0.015152	4.189655
39	17.31781	6.244998	2.851735	0.025641	3.663562
12	12.42453	3.46410162	2.519673	0.083333	2.484907
8	10.39721	2.82842712	2.341537	0.125	2.079442
34	17.6318	5.83095189	2.869704	0.029412	3.526361
19	14.72219	4.35889894	2.689356	0.052632	2.944439
39	19.31781	6.244998	2.961027	0.025641	3.663562
8	10.39721	2.82842712	2.341537	0.125	2.079442
68	21.09754	8.24621125	3.049156	0.014706	4.219508
8	9.397208	2.82842712	2.240413	0.125	2.079442
25	17.09438	5	2.83875	0.04	3.218876
1	1	1	0	1	0
3	5.493061	1.73205081	1.703486	0.333333	1.098612

Written by Lorne Priemaza

(Note: Important parts of the regression output have been highlighted, and the regression reporting can be written in by hand afterwards.)

Learn=f(ln(Study))

$$\hat{Learn}_i = -\underset{(0.412)}{0.0638} + \underset{(0.127)}{5.02}\ln(Study)_i$$

$$N = 36, R^2 = 0.9788$$

Regression Statistics	
Multiple R	0.989365
R Square	0.978844
Adjusted R Sq	0.978222
Standard Erro	0.727326
Observations	36

ANOVA

	df	SS	MS	F	gnificance F
Regression	1	832.1825082	832.18	1573.1	5E-30
Residual	34	17.98611312	0.529		
Total	35	850.1686213			

	Coefficients	Standard Error	t Stat	P-value	ower 95%	Jpper 95%	ower 95.0%	Jpper 95.0%
Intercept	-0.06384	0.412224809	-0.155	0.8778	-0.902	0.7739	-0.902	0.7739
Ln(Study)	5.020509	0.126580744	39.663	5E-30	4.7633	5.2778	4.7633	5.2778

ln(Learn)=f(Study)

$$\ln(\hat{Learn})_i = \underset{(0.124)}{2.10} + \underset{(0.00330)}{0.0177}\ Study_i$$

$$N = 36, R^2 = 0.4595$$

Regression Statistics	
Multiple R	0.677884
R Square	0.459527
Adjusted R Sq	0.44363
Standard Erro	0.412916
Observations	36

ANOVA

	df	SS	MS	F	gnificance F
Regression	1	4.928767952	4.9288	28.908	6E-06
Residual	34	5.796979989	0.1705		
Total	35	10.72574794			

	Coefficients	Standard Error	t Stat	P-value	ower 95%	Jpper 95%	ower 95.0%	Jpper 95.0%
Intercept	2.095483	0.124230013	16.868	4E-18	1.843	2.3479	1.843	2.3479
Study	0.017732	0.003297916	5.3766	6E-06	0.011	0.0244	0.011	0.0244

Written by Lorne Priemaza

ln(Learn)=f(ln(Study)

Regression Statistics	
Multiple R	0.915671
R Square	0.838454
Adjusted R S	0.833703
Standard Erro	0.225747
Observations	36

$$\ln(\hat{Learn})_i = \underset{(0.128)}{1.03} + \underset{(0.0392)}{0.521}\ln(Study)_i$$

$$N = 36, R^2 = 0.8385$$

ANOVA

	df	SS	MS	F	gnificance F
Regression	1	8.993047049	8.993	176.47	5E-15
Residual	34	1.732700892	0.051		
Total	35	10.72574794			

	Coefficients	Standard Error	t Stat	P-value	ower 95%	Jpper 95%	ower 95.0%	pper 95.0%
Intercept	1.027068	0.127946163	8.0273	2E-09	0.7671	1.2871	0.7671	1.2871
Ln(Study)	0.521906	0.039288078	13.284	5E-15	0.4421	0.6017	0.4421	0.6017

Learn=F(1/Study)

Regression Statistics	
Multiple R	0.748056
R Square	0.559587
Adjusted R S	0.546634
Standard Erro	3.318509
Observations	36

$$\hat{Learn}_i = \underset{(0.621)}{17.4} - \underset{(3.33)}{21.8}/(Study)_i$$

$$N = 36, R^2 = 0.5596$$

ANOVA

	df	SS	MS	F	gnificance F
Regression	1	475.7436546	475.74	43.2	2E-07
Residual	34	374.4249668	11.012		
Total	35	850.1686213			

	Coefficients	Standard Error	t Stat	P-value	ower 95%	Jpper 95%	ower 95.0%	pper 95.0%
Intercept	17.42223	0.621215072	28.045	4E-25	16.16	18.685	16.16	18.685
1/Study	-21.8725	3.327784157	-6.573	2E-07	-28.64	-15.11	-28.64	-15.11

Written by Lorne Priemaza

Learn=f(Study^1/2)

Regression Statistics	
Multiple R	0.960029
R Square	0.921656
Adjusted R S	0.919351
Standard Erro	1.399643
Observations	36

$$\hat{Learn}_i = \underset{(0.647)}{3.50} - \underset{(0.115)}{2.31}\sqrt{Study_i}$$

$$N = 36,\ R^2 = 0.9217$$

ANOVA

	df	*SS*	*MS*	*F*	*gnificance F*
Regression	1	783.5625903	783.56	399.98	2E-20
Residual	34	66.60603106	1.959		
Total	35	850.1686213			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*.ower 95%*	*Jpper 95%*	*.ower 95.0%*	*Jpper 95.0%*
Intercept	3.50237	0.646597882	5.4166	5E-06	2.1883	4.8164	2.1883	4.8164
Root(Study)	2.309182	0.115461862	20	2E-20	2.0745	2.5438	2.0745	2.5438