POLS 6481. Research Design and Quantitative Methods II

Lecture 13. Logarithmic Functional Forms

Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Outline

Prelude: *Health Care Expenditures* (Lecture 15 health.R)

1. The Logarithmic Transformations
2. Interpreting Coefficients

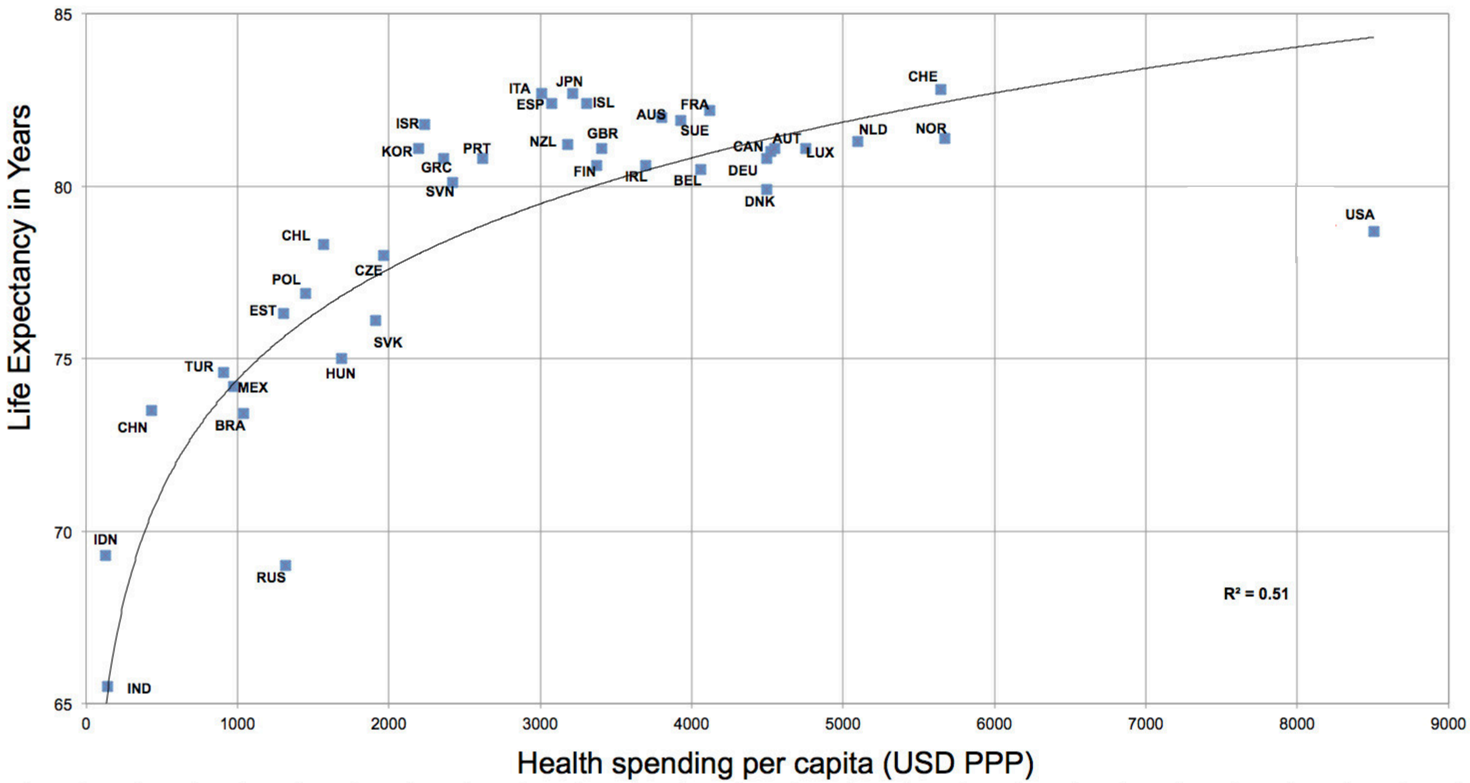
Interlude: *Wages and Salaries* (Lecture 15 wages salaries.R)

3. Transforming From *log*(y) Back to y

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Prelude: *Health Care Expenditures*

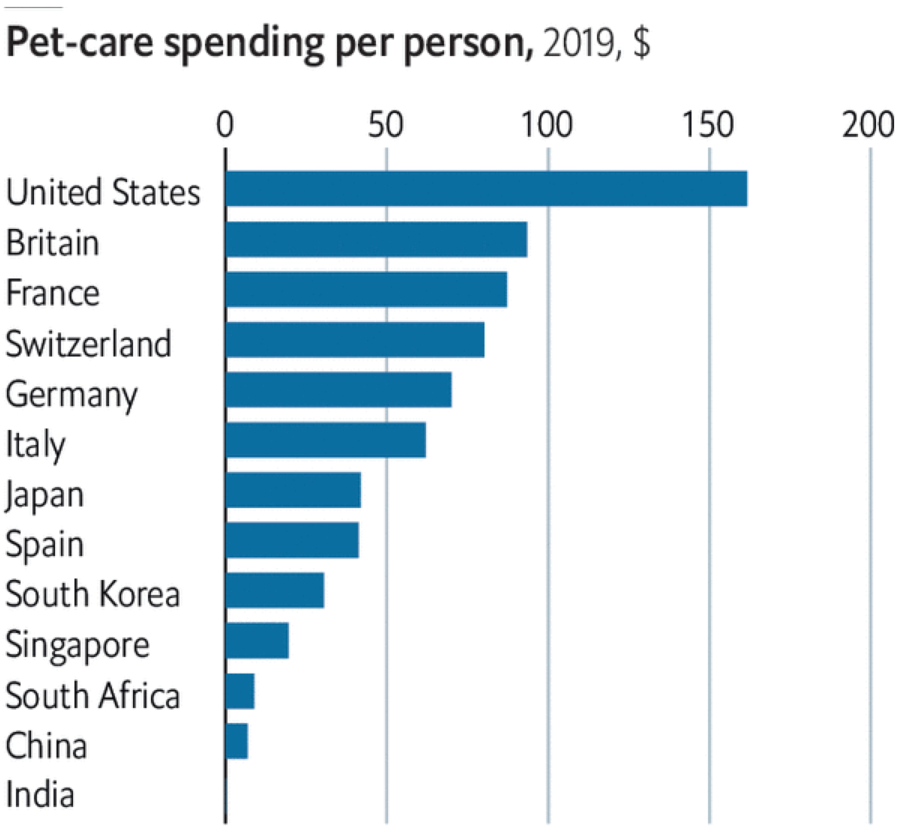
http://www.huffingtonpost.com/2013/11/22/american-health-care-terrible_n_4324967.html



POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Prelude: *Health Care Expenditures*

https://www.economist.com/graphic-detail/2020/02/11/which-country-spends-the-most-on-its-pets



POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

1. The Logarithmic Transformation

The 'semilog' model transforms only the left-hand side (y, the dependent variable)
The 'log-log' model transforms both the left- and right-hand sides (y and x's)

Some pro's of logs:

- often mitigates problems with outliers
- often secures normality and homoskedasticity
- "convenient interpretation in terms of percentages and elasticities"

Regarding the first two bullet points, a **Normal quantile plot** might help diagnose issues

Some con's of logs:

- can not be used if y takes on negative or zero values (see p. 193–4 in 5th ed.)
- should not be used if a variable is in years or percentage points (see p. 193 in 5th ed.)
- difficult to reverse the log operation when constructing predictions
- difficult to compare fit statistics ($RMSE$ or R^2) between models of y and models of log(y)

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

2. Interpreting Coefficients

Level-level bivariate model $y = \beta_0 + \beta_1 x + u$

$$\beta_1 = \frac{\partial\, y}{\partial\, x}$$

Level-log bivariate model $y = \beta_0 + \beta_1 \log(x) + u$

$$\beta_1 = \frac{\partial\, y}{\partial \log(x)}$$

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

2. Interpreting Coefficients

Level-level bivariate model

$$y = \beta_0 + \beta_1 x + u$$

$$\beta_1 = \frac{\partial\, y}{\partial\, x}$$

Level-log bivariate model

$$life\_expect = \beta_0 + \beta_1 \log(health\_expend) + u$$

$$\beta_1 = \frac{\partial\, life\_expect}{\partial \log(health\_expend)}$$

Log-level bivariate model

$$\log(wage) = \beta_0 + \beta_1 educ + u$$

$$\beta_1 = \frac{\partial \log(wage)}{\partial\, educ} = \frac{1}{wage} \cdot \frac{\partial\, wage}{\partial\, educ} = \frac{\partial\, wage / wage}{\partial\, educ}$$

Log-log bivariate model

$$\log(salary) = \beta_0 + \beta_1 \log(sales) + u$$

$$\beta_1 = \frac{\partial \log(salary)}{\partial \log(sales)} = \frac{\partial\, salary / salary}{\partial\, sales / sales}$$

Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

2. Interpreting Coefficients

Table 2.3 Summary of Functional Forms Involving Logarithms (p. 44)

Model	DV	IV	Interpretation of Coefficient
Level-level	y	x	$\Delta y = \beta_1 \, \Delta x$
Level-log	y	$\log(x)$	$\Delta y = (\beta_1 / 100) \cdot \% \Delta x$
Log-level	$\log(y)$	x	$\% \, \Delta y = (100 \cdot \beta_1) \, \Delta x$ ***
Log-log	$\log(y)$	$\log(x)$	$\% \, \Delta y = \beta_1 \cdot \% \Delta x$

*** a more precise equation is $\% \, \Delta y = 100 \cdot [\exp(\beta_1 \, \Delta x) - 1]$; see **7.2b**

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Prelude, revisited: *Health Care Expenditures*

R script: Lecture 15 health.R
Data: OECD_health.dta

Level-level model: $life\_expect = 72.33 + .002\ health\_expend$ ($R^2 = .354$)
interpretation: for every additional dollar spent, .0019773 additional years
or .72 additional days

Level-log model: $life\_expect = 42.34 + 4.64\ log(health\_expend)$ ($R^2 = .508$)
interpretation: for every 1% increase in spending, .046 additional years
or 17 additional days

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Interlude: *Wages and Salaries* (Lecture 15 wages salaries.R)

Examples 2.4, 2.10 (p. 33–34, 42 in 5th ed.)

Data: WAGE1.DTA

$$\widehat{wage} = -0.90 + 0.54 \; educ$$

$$\widehat{\log}(wage) = 0.584 + 0.083 \; educ$$

$$\beta_1 = \frac{\partial \log(wage)}{\partial educ} = \frac{1}{wage} \cdot \frac{\partial wage}{\partial educ} = \frac{\frac{\partial wage}{wage}}{\partial educ}$$

$$\frac{\frac{\partial wage}{wage}}{\partial educ} = \frac{\frac{+0.83\$}{10\$}}{+1 \text{ year}} = 0.083 = +8.3\%$$

POLS 6481. Research Design and Quantitative Methods II
Lecture 13. Logarithmic Functional Forms
Readings: Wooldridge, *Introductory Econometrics 5e*, 2.4 + 6.2b + 6.4 + 7.2b

Interlude: *Wages and Salaries* (Lecture 15 wages salaries.R)

Examples 2.3, 2.8, 2.11, 6.4, 6.7, 6.8 (p. 32–33, 39–40, 43, 204–205, 214–215 in 5th ed.)

Data: CEOSAL1.DTA

$$\widehat{salary} = 963.191 + 18.501\ roe \qquad n = 209, \qquad R^2 = 0.0132$$

$$\widehat{\log}(salary) = 4.822 + 0.257 \log(sales) \qquad n = 209, \qquad R^2 = 0.211$$

$$\beta_1 = \frac{\partial \log(salary)}{\partial \log(sales)} = \frac{\frac{\partial salary}{salary}}{\frac{\partial sales}{sales}}$$

$$\frac{\frac{\partial salary}{salary}}{\frac{\partial sales}{sales}} = \frac{\frac{+2{,}570\$}{1{,}000{,}000\$}}{\frac{+10{,}000{,}000\$}{1{,}000{,}000{,}000\$}} = \frac{+0.257\% \text{ salary}}{+1\% \text{ sales}} = 0.257$$

3. Transforming From *log*(*y*) Back to *y*

See p. 212-215 in Wooldridge, 5e.

http://healthcare-economist.com/2010/11/16/duans-smearing-estimator/

the below equations apply equally to log-log and semi-log (aka log-linear) models

$$\log(y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + u$$

$$\Rightarrow \quad y = \exp(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k)\exp(u)$$

$$\Rightarrow \quad E(y|\mathbf{x}) = \exp(\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k)E(\exp(u))$$

$$\Rightarrow \quad \hat{y} = \exp(\hat{\beta}_0+\hat{\beta}_1 x_1+\hat{\beta}_2 x_2+\cdots+\hat{\beta}_k x_k)\boxed{(\frac{1}{n}\sum_{i=1}^{n}\exp(\hat{u}_i))}$$

3. Transforming From *log*(*y*) Back to *y*

$$\widehat{salary} = \underset{(223.63)}{223.90} + \underset{(.0163)}{.0089}\ sales + \underset{(11.08)}{19.63}\ roe$$

$$n = 209, R^2 = .029, \bar{R}^2 = .020, TSS = 391,732,982$$

$$\widehat{lsalary} = \underset{(0.29)}{4.36} + \underset{(.033)}{.275}\ lsales + \underset{(.0040)}{.0179}\ roe$$

$$n = 209, R^2 = .282, \bar{R}^2 = .275, TSS = 66.72$$

3. Transforming From *log*(*y*) Back to *y*

$$\underset{}{\widehat{salary}} = \underset{(65.23)}{613.43} + \underset{(.0100)}{.0190}\ sales + \underset{(.0095)}{.0234}\ mktval + \underset{(5.61)}{12.70}\ ceoten$$

$$n = 177, R^2 = .201$$

$$\widehat{lsalary} = \underset{(.257)}{4.504} + \underset{(.039)}{.163}\ lsales + \underset{(.050)}{.109}\ mktval + \underset{(.0053)}{.0117}\ ceoten$$

$$n = 177, R^2 = .318$$

$$\tilde{R}^2 = .243$$