$lm(y \sim x1 + x2 + x3 + \ldots + xk)$

$lm(y \sim$ all other reg except x_k) → resid.

$lm(x_k \sim$) → resid.

General case

Y_t regressand $\qquad Y_t = \log(\text{total comp})$

$X_t = (1, x_{1t}, \ldots, x_{kt})$

$x_{1t} = \log(\text{net sales})$

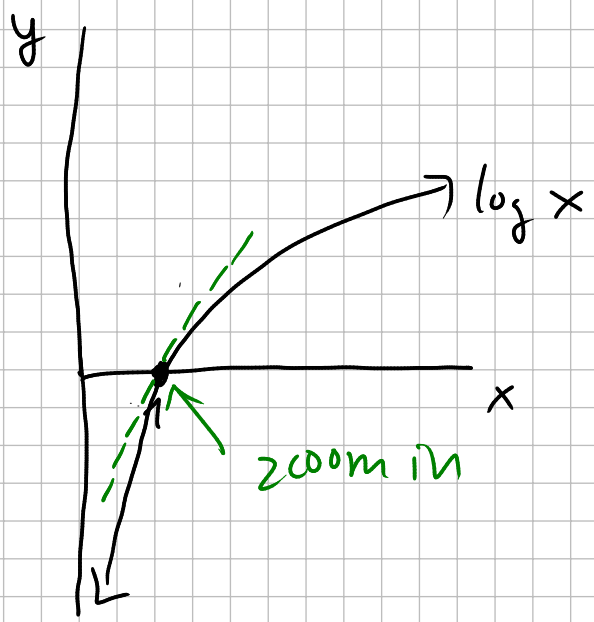
$\log(\text{net sales})$ differ by 1

$\log(\text{new net sales}) - \log(\text{old net sales}) = 1$

$\log x_2 - \log x_1 = \log\left(\frac{x_2}{x_1}\right) = \log\left(1 + \left(\frac{x_2}{x_1} - 1\right)\right)$

↑ algebra ↑ algebra

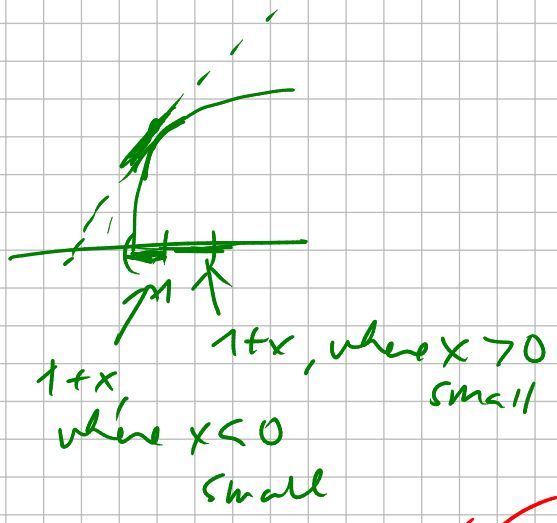
$\log(1 + \text{something})$ ≈ small



$$\approx \frac{\text{total comp}\ \text{new net sales} - \text{old}\ \text{total comp}\ \text{net sales}}{\text{old net sales}} = \underbrace{0.4}_{1}$$

log net sales differ by

new net sales = 2 * old net sales



$$\text{course eval} = \underbrace{4.069}_{\text{matches average course eval males}} - \underbrace{0.168}_{\text{diff of means}} \text{female}$$

dummy variable

average course evals of females — average of course evals of males.

as.factor(Industry)

↑ treat Industry as a categorical/grouping variable

female → 1 dummy

2 groups

base group = male

9 dummies.

10 groups

base group = Discretionary Industry

$\widehat{\text{conseval}} = \underline{3.99} + \underline{0.168}\,\text{male}$

"dummy where = 1 if male
= 0 if female

female as base group

---

average $\log_{10}(\text{totalcomp})$ in discret. industry = 6.43482

totalcomp average dis. ind = $10^{6.43482}$

≈ 2721573

$\Leftarrow$ 2.7 million

average (log (comp))

does not match with summaries of CEO comp averages

$\widehat{\text{TotalCompMil}} = 4.684 + 0.399\,\text{Energy} + 1.45\,\text{Financial} + \ldots$

$\ln(\text{TotalCompMil} \sim 1)$

average log compensation differ by 0.02629

$$\log(\text{comp energy}) - \log(\text{comp discrete}) = 0.02629$$

$$\approx \frac{\text{comp energy} - \text{comp discre}}{\text{comp discre}} = 0.02629$$

$$\text{comp energy} = 1.02629 * \text{comp discre.}$$

2.629% more

list of R commands (lines R could actually execute)

→ R Script

script_data_clean.R
script_minimal_example.R

different QMD file

$\widehat{\text{Total Comp}}_{Mil} = 4.6841 + 0.399\,\text{Energy}_i + \ldots$

"1 if CEO is energy sector, 0 otherwise"

$\widehat{\text{Total Comp}}_{Mil}{}_t = 4.6841$ t indexes CEOs in the discretionary industry

$\widehat{\text{Total Comp}}_{Mil}{}_t = 4.6841 + 0.399 \Rightarrow$ t indexes CEOs in Energy industry

$\widehat{\log_{10}(\text{Total comp})} = 2.48 + 0.43\log_{10}(\text{net sales}) + 0.082\,\text{Energy}_i + \ldots$

$\widehat{\log_{10}(\text{Total comp})}_t = 2.48 + 0.43\log_{10}(\text{net sales})_t$ for CEOs in discretionary ind.

$\widehat{\log_{10}(\text{Total comp})}_t = 2.48 + 0.43\log_{10}(\text{net sales})_t + 0.082$ for CEOs in energy industry

intercept

with interaction terms

$\log_{10}(\text{net sales}) * \text{Industry}$

$$\log_{10}(\text{TotalCompMil})_t = 2.05 + 0.47 * \log_{10}(\text{net sales}) + 0.487 * \text{Energy} + \dots - 1.21 * \text{Utility} - 0.043 * \log_{10}(\text{net sales}) * \text{Energy} + \dots$$

Discretionary industry: $\log_{10}(\text{TotalCompMil})_t = 2.05 + 0.47 \log_{10}(\text{Netsales})$

Energy industry: $\log_{10}(\text{TotalCompMil})_t = 2.05 + 0.47 \log_{10}(\text{netsales}) + 0.487 - 0.043 * \log_{10}(\text{net sales})$

$= 2.537 + 0.427 * \log_{10}(\text{net sales})$