Deriving the OLS estimators

The Method of Moments and Minimizing the Sum of Squared Residuals both produce these two equations:

① $\sum_{i=1}^{n} (y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$

② $\sum_{i=1}^{n} [x_i(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)] = 0$

Solve for $\hat{\beta}_1$:

① $\Rightarrow \bar{y} - \hat{\beta}_0 - \hat{\beta}_1 \bar{x} = 0$, or $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$

Sub ① into ②: $\sum_{i=1}^{n} [x_i(y_i - (\bar{y} - \hat{\beta}_1 \bar{x}) - \hat{\beta}_1 x_i)] = 0$

Transform: $\sum_{i=1}^{n} [x_i(y_i - \bar{y}) - \hat{\beta}_1 x_i(x_i - \bar{x})] = 0$

$$\Rightarrow \sum_{i=1}^{n} [x_i(y_i - \bar{y})] = \hat{\beta}_1 \sum_{i=1}^{n} [x_i(x_i - \bar{x})]$$

Rewrite using properties of summation operator: (see p. 705 in Wooldridge 6th ed.)

$$\sum_{i=1}^{n} [(x_i - \bar{x})(y_i - \bar{y})] = \hat{\beta}_1 \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$\Rightarrow \hat{\beta}_1 = \frac{\sum_{i=1}^{n} [(x_i - \bar{x})(y_i - \bar{y})]}{\sum_{i=1}^{n} (x_i - \bar{x})^2} \quad \blacksquare$$

To get $\hat{\beta}_0$, plug $\hat{\beta}_1$ formula into $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$.

Proof of unbiasedness of $\hat{\beta}_1$:

Want to show that in expectation, $\hat{\beta}_1 = \beta_1$, the true population parameter. Requires assumptions 1-4.

Rewrite $\hat{\beta}_1$ as $\dfrac{\sum (x_i - \bar{x}) y_i}{\sum (x_i - \bar{x})^2} = \dfrac{\sum (x_i - \bar{x}) y_i}{\sum SST_x}$

where SST_x is the sum of squared deviation in x.

$$\Rightarrow \hat{\beta}_1 = \frac{1}{SST_x} \sum (x_i - \bar{x})(\beta_0 + \beta_1 x_i + u_i) \quad \text{(sub in for } y_i\text{)}$$

$$= \frac{1}{SST_x}\left[\beta_0 \sum (x_i - \bar{x}) + \beta_1 \sum (x_i - \bar{x}) x_i + \sum (x_i - \bar{x}) u_i\right]$$

by properties of summation.

Now, $\sum (x_i - \bar{x}) = 0$, and $\sum (x_i - \bar{x}) x_i = \sum (x_i - \bar{x})^2$ also by properties of summation.

$$\Longrightarrow \hat{\beta}_1 = \frac{1}{SST_x}\left[\beta_1 \sum (x_i - \bar{x})^2 + \sum (x_i - \bar{x}) u_i\right]$$

$$= \frac{1}{SST_x}\left[\beta_1 SST_x + \sum (x_i - \bar{x}) u_i\right]$$

$$= \beta_1 + \frac{1}{SST_x} \sum (x_i - \bar{x}) u_i$$

Let $d_i = (x_i - \bar{x})$, so $\hat{\beta}_1 = \beta_1 + \frac{1}{SST_x} \sum d_i u_i$

(This substitution just helps us see that $x_i - \bar{x}$ is a constant that we can take outside the expectations operator.)

Take expectation: $E(\hat{\beta}_1) = \beta_1 + \frac{1}{SST_x} \sum d_i E(u_i)$

$\Longrightarrow E(\hat{\beta}_1) = \beta_1$, because $E(u_i) = 0$ ∎

Proof of unbiasedness of $\hat{\beta_0}$

$$\hat{\beta_0} = \bar{y} - \hat{\beta_1}\bar{x}, \text{ so } E(\hat{\beta_0}) = \bar{y} - E(\hat{\beta_1})\bar{x}$$
$$= \beta_0 + \beta_1 \bar{x} - E(\hat{\beta_1})\bar{x} = \beta_0 \quad \blacksquare$$

where we have used the fact that the sample regression line fits exactly at the mean $(\bar{x}, \bar{y})$, as does the population line.

Deriving the variance of $\hat{\beta_1}$

Will use variance properties: $Var(aX+b) = a^2(Var\,X)$.

$$Var(\hat{\beta_1}) = Var\left[\beta_1 + \frac{1}{SST_x}\sum d_i u_i\right]$$

This is one of the expressions of $\hat{\beta_1}$ we saw when proving unbiasedness.

$$= \left(\frac{1}{SST_x}\right)^2 Var\left(\sum (d_i u_i)\right)$$

$$= \left(\frac{1}{SST_x}\right)^2 \sum \left[d_i^2 \, Var(u_i)\right]$$

$$= \left(\frac{1}{SST_x}\right)^2 \sum d_i^2 \sigma^2 \quad \text{(by homoskedasticity assumption)}$$

$$= \sigma^2 \left(\frac{1}{SST_x}\right)^2 SST_x \quad \text{(because } d_i = (x_i - \bar{x})\text{)}$$

$$= \frac{\sigma^2}{SST_x} \quad \blacksquare$$

Sample equivalent of this is: $\hat{\sigma}^2 = \frac{1}{n-2}\sum(\hat{u}_i)^2 = \frac{SSR}{n-2}$

Standard error of $\hat{\beta_1} = \frac{\hat{\sigma}}{\sqrt{SST_x}}$.

Why changes in logs are approximately equal to percent changes

Let $a = b + \Delta$, where Δ is a small change.

$\log(a) - \log(b) = \log\left(\frac{a}{b}\right)$ by properties of log.

$$\log\left(\frac{a}{b}\right) = \log\left(\frac{b+\Delta}{b}\right) = \log\left(1 + \frac{\Delta}{b}\right) \approx \frac{\Delta}{b}$$

where the last step is also by properties of log: $\log(1+x) \approx x$

$\frac{\Delta}{b} = \frac{a-b}{b}$, which is the percent change when going from b to a.

So, a change in logs $(\log(a) - \log(b))$ is $\approx$ % change. ■