Econometrics: Deriving OLS Estimators

Ryan Safner

Proof. Begin with the formula we derived for $\hat{\beta}_1$:

$$\hat{\beta}_1 = \frac{\displaystyle\sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X})}{\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{1}$$

Recall from **Rule 6** of summations[1], we can rewrite the numerator as

$$\begin{aligned} &= \sum_{i=1}^{n}(Y_i - \bar{Y})(X_i - \bar{X}) \\ &= \sum_{i=1}^{n} Y_i(X_i - \bar{X}) \end{aligned}$$

$$\hat{\beta}_1 = \frac{\displaystyle\sum_{i=1}^{n} Y_i(X_i - \bar{X})}{\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{2}$$

We know the true population relationship is expressed as

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i$$

Substituting this in for Y_i in equation 2:

$$\hat{\beta}_1 = \frac{\displaystyle\sum_{i=1}^{n}(\beta_0 + \beta_1 X_i + \epsilon_i)(X_i - \bar{X})}{\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{3}$$

Breaking apart the sums in the numerator:

$$\hat{\beta}_1 = \frac{\displaystyle\sum_{i=1}^{n} \beta_0(X_i - \bar{X}) + \sum_{i=1}^{n} \beta_1 X_i(X_i - \bar{X}) + \sum_{i=1}^{n} \epsilon_i(X_i - \bar{X})}{\displaystyle\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{4}$$

We can simplify equation 4 using **Rules 4 and 5** of summations

[1] See the **handout** on the summation operator

1. The first term in the numerator $\left[\sum_{i=1}^{n} \beta_0(X_i - \bar{X})\right]$ has the constant β_0, which can be pulled out of the summation. This gives us the summation of deviations, which add up to 0 as per **Rule 4**:

$$\begin{aligned}\sum_{i=1}^{n} \beta_0(X_i - \bar{X}) &= \beta_0 \sum_{i=1}^{n}(X_i - \bar{X}) \\ &= \beta_0(0) \\ &= 0\end{aligned}$$

2. The second term in the numerator $\left[\sum_{i=1}^{n} \beta_1 X_i(X_i - \bar{X})\right]$ has the constant β_1, which can be pulled out of the summation. Additionally, **Rule 5** tells us $\sum_{i=1}^{n} X_i(X_i - \bar{X}) = \sum_{i=1}^{n}(X_i - \bar{X})^2$:

$$\begin{aligned}\sum_{i=1}^{n} \beta_1 X_1(X_i - \bar{X}) &= \beta_1 \sum_{i=1}^{n} X_i(X_i - \bar{X}) \\ &= \beta_1 \sum_{i=1}^{n}(X_i - \bar{X})^2\end{aligned}$$

   When placed back in the context of being the numerator of a fraction, we can see this term simplifies to just β_1:

$$\begin{aligned}\frac{\beta_1 \sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} &= \frac{\beta_1}{1} \times \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \\ &= \beta_1\end{aligned}$$

Thus, we are left with:

$$\hat{\beta_1} = \beta_1 + \frac{\sum_{i=1}^{n} \epsilon_i(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{5}$$

Now, take the expectation of both sides:

$$E[\hat{\beta_1}] = E\left[\beta_1 + \frac{\sum_{i=1}^{n} \epsilon_i(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}\right]$$

We can break this up, using properties of expectations. First, recall $E[a+b] = E[a] + E[b]$, so we can break apart the two terms.

$$E[\hat{\beta_1}] = E[\beta_1] + E\left[\frac{\sum_{i=1}^{n} \epsilon_i (X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}\right]$$

Second, the true population value of β_1 is a constant, so $E[\beta_1] = \beta_1$.

Third, since we assume X is also "fixed" and not random, the variance of X, $\sum_{i=1}^{n}(X_i - \bar{X})$, in the denominator, is just a constant, and can be brought outside the expectation.

$$E[\hat{\beta_1}] = \beta_1 + \frac{E\left[\sum_{i=1}^{n} \epsilon_i (X_i - \bar{X})\right]}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \tag{6}$$

Thus, the properties of equation 6 are primarily driven by the expectation $E\left[\sum_{i=1}^{n} \epsilon_i (X_i - \bar{X})\right]$. We now turn to this term.

Take equation 5, and use the property of summation operators to expand the numerator term:

$$\hat{\beta_1} = \beta_1 + \frac{\sum_{i=1}^{n} \epsilon_i (X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

$$\hat{\beta_1} = \beta_1 + \frac{\sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})(X_i - \bar{X})}{\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

Now divide the numerator and denominator of the second term by $\frac{1}{n}$. Realize this gives us the covariance between X and ϵ in the numerator and variance of X in the denominator, based on their respective definitions.

$$\hat{\beta_1} = \beta_1 + \frac{\frac{1}{n}\sum_{i=1}^{n} (\epsilon_i - \bar{\epsilon})(X_i - \bar{X})}{\frac{1}{n}\sum_{i=1}^{n} (X_i - \bar{X})^2}$$

$$\hat{\beta_1} = \beta_1 + \frac{cov(X, \epsilon)}{var(X)}$$

$$\hat{\beta_1} = \beta_1 + \frac{s_{X,\epsilon}}{s_X^2}$$

By the Zero Conditional Mean assumption of OLS, $cov(X, \epsilon) = 0$.

Alternatively, we can express the bias in terms of correlation instead of covariance:

$$E[\hat{\beta}_1] = \beta_1 + \frac{cov(X, \epsilon)}{var(X)}$$

From the definition of correlation:

$$corr(X, \epsilon) = \frac{cov(X, \epsilon)}{s_X s_\epsilon}$$
$$corr(X, \epsilon) s_X s_\epsilon = cov(X, \epsilon)$$

Plugging this in:

$$E[\hat{\beta}_1] = \beta_1 + \frac{cov(X, \epsilon)}{var(X)}$$
$$E[\hat{\beta}_1] = \beta_1 + \frac{\left[corr(X, \epsilon) s_x s_\epsilon\right]}{s_X^2}$$
$$E[\hat{\beta}_1] = \beta_1 + \frac{corr(X, \epsilon) s_\epsilon}{s_X}$$
$$E[\hat{\beta}_1] = \beta_1 + corr(X, \epsilon) \frac{s_\epsilon}{s_X}$$

□

0.1 Proof of Unbiasedness

Begin with equation:

$$\hat{\beta}_1 = \frac{\sum Y_i X_i}{\sum X_i^2} \tag{7}$$

Substitute for Y_i:

$$\hat{\beta}_1 = \frac{\sum (\beta_1 X_i + \epsilon_i) X_i}{\sum X_i^2} \tag{8}$$

Distribute X_i in the numerator:

$$\hat{\beta}_1 = \frac{\sum \beta_1 X_i^2 + \epsilon_i X_i}{\sum X_i^2} \tag{9}$$

Separate the sum into additive pieces:

$$\hat{\beta}_1 = \frac{\sum \beta_1 X_i^2}{\sum X_i^2} + \frac{\epsilon_i X_i}{\sum X_i^2} \tag{10}$$

β_1 is constant, so we can pull it out of the first sum:

$$\hat{\beta}_1 = \beta_1 \frac{\sum X_i^2}{\sum X_i^2} + \frac{\epsilon_i X_i}{\sum X_i^2} \tag{11}$$

Simplifying the first term, we are left with:

$$\hat{\beta}_1 = \beta_1 + \frac{\epsilon_i X_i}{\sum X_i^2} \tag{12}$$

Now if we take expectations of both sides:

$$E[\hat{\beta}_1] = E[\beta_1] + E\left[\frac{\epsilon_i X_i}{\sum X_i^2}\right] \tag{13}$$

β_1 is a constant, so the expectation of β_1 is itself.

$$E[\hat{\beta}_1] = \beta_1 + E\left[\frac{\epsilon_i X_i}{\sum X_i^2}\right] \tag{14}$$

Using the properties of expectations, we can pull out $\frac{1}{\sum X_i^2}$ as a constant:

$$E[\hat{\beta}_1] = \beta_1 + \frac{1}{\sum X_i^2} E\left[\sum \epsilon_i X_i\right] \tag{15}$$

Again using the properties of expectations, we can put the expectation inside the summation operator (the expectation of a sum is the sum of expectations):

$$E[\hat{\beta}_1] = \beta_1 + \frac{1}{\sum X_i^2} \sum E[\epsilon_i X_i] \tag{16}$$

Under the exogeneity condition, the correlation between X_i and ϵ_i is 0.

0.2 Variance of $\hat{\beta}_1$