Impact of Autocorrelation on OLS Estimates
ECON 30331/Evans

Consider a simple bivariate time-series model of the form: $y_t = \beta_0 + x_t\beta_1 + \varepsilon_t$

The four key assumptions about ε in this model are

i) $E(\varepsilon_t) = E[\varepsilon_t | x_t]=0$

ii) $Var(\varepsilon_t) = Var(\varepsilon_t | x_t) = \sigma_\varepsilon^2$

iii) $Cov(\varepsilon_t, \varepsilon_j) = 0$ for $t\neq j$

There are n observations in the sample, so t=1,2,3….n

Assume that the third assumption is violated and there is first order auto-correlation, which is characterized by the following equation

$$\varepsilon_t = \rho\varepsilon_{t-1} + v_t$$

Where $|\rho|<1$ and we make the following assumptions about v

i) $E(v_t) = 0$

ii) $Var(v_t) = \sigma_v^2$

iii) $Cov(v_t, v_j) = 0$ for $t\neq j$

iv) $Cov(\varepsilon_{t-1}, v_t) = 0$

With autocorrelation, consider the value of $Cov(\varepsilon_t, \varepsilon_{t-1})$. By definition

$$(1) \quad Cov(\varepsilon_t, \varepsilon_{t-1}) = E[(\varepsilon_t - E(\varepsilon_t))(\varepsilon_{t-1} - E(\varepsilon_{t-1}))]$$

but because $E(\varepsilon_t) = E(\varepsilon_{t-1}) = 0$, equation (1) reduces to

$$(2) \quad Cov(\varepsilon_t, \varepsilon_{t-1}) = E[\varepsilon_t\varepsilon_{t-1}]$$

Using the fact that we have first-order autocorrelation, $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$, we can substitute this equation into (2) for ε_t

$$(3) \quad Cov(\varepsilon_t, \varepsilon_{t-1}) = E[\varepsilon_t\varepsilon_{t-1}] = E[(\rho\varepsilon_{t-1} + v_t)\varepsilon_{t-1}] = E[\rho\varepsilon_{t-1}^2] + E[v_t\varepsilon_{t-1}]$$

Recall that by definition $Var(\varepsilon_{t-1}) = E[(\varepsilon_{t-1} - E(\varepsilon_{t-1}))^2] = \sigma_\varepsilon^2$, but because $E(\varepsilon_t) = 0$, then $Var(\varepsilon_{t-1}) = E[\varepsilon_{t-1}^2] = \sigma_\varepsilon^2$. Notice also that we assume v_t and ε_{t-1} are uncorrelated so $E[v_t\varepsilon_{t-1}]=0$, and therefore

$$(4) \quad Cov(\varepsilon_t, \varepsilon_{t-1}) = E[\rho\varepsilon_{t-1}^2] + E[v_t\varepsilon_{t-1}] = \rho\sigma_\varepsilon^2$$

This method can easily be generalized to consider the degree of covariance between any two time periods in the data. For example, suppose we are interested in $Cov(\varepsilon_t, \varepsilon_{t-2})$ which by definition equals

(5) $Cov(\varepsilon_t, \varepsilon_{t-2}) = E[\varepsilon_t \varepsilon_{t-2}]$.

Note that we can write ε_t as a function of ε_{t-1} and ε_{t-1} as a function of ε_{t-2}.

(6) $\varepsilon_t = \rho\varepsilon_{t-1} + v_t$

(7) $\varepsilon_{t-1} = \rho\varepsilon_{t-2} + v_{t-1}$

And therefore, substituting (7) into (6) we get

(8) $\varepsilon_t = \rho(\rho\varepsilon_{t-2} + v_{t-1}) + v_t = \rho^2\varepsilon_{t-2} + \rho v_{t-1} + v_t$

Substituting (8) into (5) for ε_t , we get

(9) $Cov(\varepsilon_t, \varepsilon_{t-2}) = E[(\rho^2\varepsilon_{t-2} + \rho v_{t-1} + v_t)\varepsilon_{t-2}] = E[\rho^2\varepsilon_{t-2}^2] + E[\rho v_{t-1}\varepsilon_{t-2}] + E[v_t\varepsilon_{t-2}]$

And using the facts that $E[\rho v_{t-1}\varepsilon_{t-2}] = E[v_t\varepsilon_{t-2}] = 0$ and $E[\rho^2\varepsilon_{t-2}^2] = \rho^2\sigma_\varepsilon^2$ we find that

(10) $Cov(\varepsilon_t, \varepsilon_{t-2}) = E[\rho^2\varepsilon_{t-2}^2] = \rho^2\sigma_\varepsilon^2$

We can easily generalize this result to note that if there is a j unit difference between the errors, then

(11) $Cov(\varepsilon_t, \varepsilon_{t-t}) = E[\rho^j\varepsilon_{t-j}^2] = \rho^j\sigma_\varepsilon^2$

In what follows, we will examine the bias and variance of the OLS estimate for β_1

a) assuming autocorrelation is present, but
b) ignoring this fact in the model

Is $\hat{\beta}_1$ Unbiased?

Recall the definition for $\hat{\beta}_1$

(1) $$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(y_t - \bar{y})(x_t - \bar{x})}{\sum_{i=1}^{n}(x_t - \bar{x})^2}$$

Note also that we can drop the $\bar{y}$ in the numerator and we can substitute the true definition of y $(\beta_0 + x_t\beta_1 + \varepsilon_t)$ into the model

(2) $$\hat{\beta}_1 = \frac{\sum_{t=1}^{n}(y_t - \bar{y})(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{\sum_{t=1}^{n}y_t(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2} = \frac{\sum_{t=1}^{n}(\beta_0 + \beta_1 x_t + \varepsilon_t)(x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

Break apart the terms in the numerator

$$(3) \qquad \hat{\beta}_1 = \frac{\beta_0 \sum_{t=1}^{n}(x_t - \bar{x}) + \beta_1 \sum_{t=1}^{n} x_t (x_t - \bar{x}) + \sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

We can simplify the terms in (3) using the properties of summations:

$$\beta_0 \sum_{t=1}^{n}(x_t - \bar{x}) = \beta_0 (0) = 0$$

$$\sum_{i=1}^{n}(x_i - \bar{x})^2 = \sum_{i=1}^{n} x_i (x_i - \bar{x}) \text{ so}$$

And therefore, equation (3) reduces to our old friend

$$(4) \qquad \hat{\beta}_1 = \beta_1 + \frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

Taking the expectation of (4), notice that

$$(5) \qquad E[\hat{\beta}_1] = E\left[\beta_1 + \frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right] = E[\beta_1] + E\left[\frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{\sum_{t=1}^{n}(x_t - \bar{x})^2}\right] = \beta_1 + \frac{E\left[\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})\right]}{\sum_{t=1}^{n}(x_t - \bar{x})^2}$$

Note that the given the presence of autocorrelation, the key assumption that x is fixed has not been altered. Therefore, we can still maintain assumption iv) $E(\varepsilon_t | x_t) = 0$ and as a result,

(6) $E[\hat{\beta}_1] = \beta_1$

Even in the presence of autocorrelation, OLS estimates are unbiased.

The variance of $\hat{\beta}_1$ in the presence of autocorrelation:

By definition,

$$(7)\quad Var(\hat{\beta}_1) = E[(\hat{\beta}_1 - E[\hat{\beta}_1])^2]$$

Previously, we demonstrated that $\hat{\beta}_1$ is an unbiased estimate even in the presence of autocorrelation and therefore, $E[\hat{\beta}_1]=\beta_1$

$$(8)\quad Var(\hat{\beta}_1) = E[(\hat{\beta}_1 - \beta_1)^2]$$

Looking at equation (4) on the previous page, note that the difference $\hat{\beta}_1 - \beta_1$ is simply

$$(9)\quad \hat{\beta}_1 - \beta_1 = \frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{\sum_{t=1}^{n} (x_t - \bar{x})^2} = \frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{SST_x} \qquad where\ SST_x = \sum_{t=1}^{n} (x_t - \bar{x})^2$$

Using the definition of the variance and equation (9)

$$(10)\quad Var(\hat{\beta}_1) = E[(\hat{\beta}_1 - \beta_1)^2] = E\left[\left(\frac{\sum_{t=1}^{n} \varepsilon_t (x_t - \bar{x})}{SST_x}\right)^2\right] = E\left[\left(\frac{\sum_{t=1}^{n} \varepsilon_t \tilde{x}_t}{SST_x}\right)^2\right]$$

Where $\tilde{x}_t = x_t - \bar{x}$. Because SST_x is a constant (x is considered fixed) we can bring it outside the expectation. Therefore

$$(11)\quad Var(\hat{\beta}_1) = \frac{1}{SST_x^2} E\left[\left(\sum_{t=1}^{n} \varepsilon_t \tilde{x}_t\right)^2\right]$$

Let's work with the numerator in the far right hand term in equation (8). Complete the square on this term generates.

$$\left[\sum_{t=1}^{n} \varepsilon_t \tilde{x}_t\right]^2 = \left[\varepsilon_1 \tilde{x}_1 + \varepsilon_2 \tilde{x}_2 + \dots \varepsilon_n \tilde{x}_n\right]^2 = [\varepsilon_1^2 \tilde{x}_1^2 + \varepsilon_2^2 \tilde{x}_2^2 + \dots \varepsilon_n^2 \tilde{x}_n^2 + 2\varepsilon_1 \tilde{x}_1 \varepsilon_2 \tilde{x}_2 + 2\varepsilon_1 \tilde{x}_1 \varepsilon_3 \tilde{x}_3 + \dots 2\varepsilon_{n-1} \tilde{x}_{n-1} \varepsilon_n \tilde{x}_n]$$

Which can be simplified to read.

$$(12)\quad \sum_{t=1}^{n} \varepsilon_t^2 \tilde{x}_t^2 + 2\sum_{t=1}^{n-1} \sum_{j=t+1}^{n} \varepsilon_t \varepsilon_j \tilde{x}_t \tilde{x}_j$$

Notice in the second term, the summations go from t=1 to n-1 and j=t+1 to n. So when t=1, we consider the cross terms $\varepsilon_1 \varepsilon_2 \tilde{x}_1 \tilde{x}_2$ to $\varepsilon_1 \varepsilon_n \tilde{x}_1 \tilde{x}_n$ while when t=n-1, we only consider the term, $\varepsilon_{n-1} \varepsilon_n \tilde{x}_{n-1} \tilde{x}_n$. Substituting (12) into (11), we obtain that

(13)
$$Var(\hat{\beta}_1) = \frac{1}{SST_x^2} E\left[\sum_{t=1}^{n} \varepsilon_t^2 \tilde{x}_t^2 + 2\sum_{t=1}^{n-1} \sum_{j=t+1}^{n} \varepsilon_t \varepsilon_j \tilde{x}_t \tilde{x}_j \right]$$
$$= \frac{1}{SST_x^2} E\left[\sum_{t=1}^{n} \varepsilon_t^2 \tilde{x}_t^2 \right] + \frac{2}{SST_x^2} E\left[\sum_{t=1}^{n-1} \sum_{j=t+1}^{n} \varepsilon_t \varepsilon_j \tilde{x}_t \tilde{x}_j \right]$$

By definition, $\tilde{x}_t^2$ is fixed and therefore, the first term reduces to something we already know:

(14)
$$\frac{1}{SST_x^2} E\left[\sum_{t=1}^{n} \varepsilon_t^2 \tilde{x}_t^2 \right] = \frac{1}{SST_x^2} \left[\sum_{t=1}^{n} \tilde{x}_t^2 E[\varepsilon_t^2] \right] = \frac{1}{SST_x^2} \sum_{t=1}^{n} \tilde{x}_t^2 \sigma_\varepsilon^2 = \frac{\sigma_\varepsilon^2}{SST_x^2} \sum_{t=1}^{n} \tilde{x}_t^2$$
$$= \frac{\sigma_\varepsilon^2}{SST_x^2} \sum_{t=1}^{n} (x_t - \bar{x})^2 = \frac{\sigma_\varepsilon^2}{SST_x^2} SST_x = \frac{\sigma_\varepsilon^2}{SST_x}$$

Looking back at our notes for the basic OLS model, recall that the second term in equation (13) $\frac{2}{SST_x^2} E\left[\sum_{t=1}^{n-1} \sum_{j=t+1}^{n} \varepsilon_t \varepsilon_j \tilde{x}_t \tilde{x}_j \right]$ drops out completely because we assume that Cov(ε_t, ε_j) =0. In this case, however, we are allowing the errors to be correlated. Using results from page (2) of this handout, the expectation of a term inside the summation is the following

(15) $$E[\varepsilon_t \varepsilon_j \tilde{x}_t \tilde{x}_j] = \tilde{x}_t \tilde{x}_j E[\varepsilon_t \varepsilon_j] = \tilde{x}_t \tilde{x}_j \sigma_\varepsilon^2 \rho^{j-t}$$

and therefore, adding (15) and (14) back into (13)

(16) $$Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{SST_x} + \frac{2\sigma_\varepsilon^2}{SST_x^2} \left[\sum_{t=1}^{n-1} \sum_{j=t+1}^{n} (x_t - \bar{x})(x_j - \bar{x}) \rho^{j-t} \right]$$

And for lack of a better term, call this variance $V_{ac(1)}$. Note in the second term in equation (11) that σ_ε^2 >0, SSTx>0, and assuming autocorrelation is positive, then ρ^{j-t} >0 as well. Finally, if the x's are positively correlated over time, then $\sum_{t=1}^{n-1} \sum_{j=t1}^{n-j} (x_t - \bar{x})(x_{t+j} - \bar{x}) \rho^j$ >0.

If there is no autocorrelation in the model, the true variance of the estimate would be $Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{SST_x}$ and for lack of a better term, call this $Var(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{SST_x}$ = V_{ols}. Therefore, if the data exhibits first-order autocorrelation and we ignore it, we will use V_{ols} as the variance. However, V_{ols} is contained in $V_{ac(1)}$ and the second term in equation (16) is positive, then by construction $V_{ac(1)}$ > V_{ols} and as result, if we have positive autocorrelation in the data and we ignire this fact, the estimated variance will be too small, leading to more Type I errors.