OLS bias for econometric models with errors-in-variables. The Lucas-critique

RNy

September 19, 2014

1 Properties of OLS in Rational Expectations models

Consider the following example of a rational expectations (RE) model:

\[ Y_t = \beta_1 E(X_{t+1} \mid I_{t-1}) + \varepsilon_t \]  
\[ X_t = \lambda X_{t-1} + \epsilon_{xt}, \quad -1 < \lambda < 1 \]  
\[ \varepsilon_t \sim IID(0, \sigma^2) \]  
\[ \epsilon_{xt} \sim IID(0, \sigma_x^2) \]  
\[ Cov(\varepsilon_t, \epsilon_{xs}) = 0 \text{ for all } t \text{ and } s \]

(1) is the structural equation and \( \beta_1 \) is the parameter of interest of this model. \(^1\) \( \beta_1 \) shows by how much \( Y \) is adjusted in period \( t \) when the expectation about \( X \) in period \( t+1 \) is changed based on information that the agents have available in period \( t-1 \). Hence the agents are forward-looking in this model, but like other forecasters they must build on history in order to formulate their expectations about the future \( X \). The information set used by the agents to form expectations is denoted by \( I_{t-1} \) in (1).

In this note, we focus on the salient implications of this type of expectations for econometric modelling. Algebraic details aside, those consequences are the same if we have \( E(X_{t+1} \mid I_t) \) in (1), allowing agents to form expectations on the basis of period \( t \) information, and also if \( E(X_t \mid I_{t-1}) \) which would only remove the forward-looking aspect of the model, not the econometric consequences of rational expectations formation.

The disturbances \( \varepsilon_t \) and \( \epsilon_{xt} \) are independent from each other, as stated in assumption (5) and they have classical properties conditional on the agents’ information set \( I_{t-1} \). Without loss of generality, this is conveyed by the assumptions that each disturbance is \( IID \) with expectation zero and constant variance.

\(^1\) The model is specified without an intercept, without loss of generality, and because it simplifies the algebra.
The solution for $X_{t+1}$ is

$$X_{t+1} = \lambda^2 X_{t-1} + \lambda \epsilon_{xt} + \epsilon_{xt+1}$$  \hspace{1cm} (6)$$

with conditional mathematical expectation

$$E(X_{t+1} \mid I_{t-1}) = \lambda^2 X_{t-1}.$$  \hspace{1cm} (7)$$

(The unconditional expectation of $X_{t+1}$ is zero, it does not play a role in the derivations below.) The unconditional variance of $X_{t+1}$ does not depend on $t$ and it is therefore the same as $Var(X_t)$. From the lecture on dynamic regression we found it to be

$$Var(X_t) = \frac{\sigma^2_x}{1 - \lambda^2}.$$$$

The expectation error is

$$X_{t+1} - E(X_{t+1} \mid I_{t-1}) = \lambda \epsilon_{xt} + \epsilon_{xt+1}$$  \hspace{1cm} (8)$$

We know that OLS gives consistent estimators of the conditional expectation of $Y_t$ given $X_{t+1}$. But is $\beta_1$ a parameter in the conditional expectation? To answer that question we can investigate the probability limit of the OLS estimator

$$\hat{\beta}_1 = \frac{\sum_t (X_{t+1}) Y_t}{\sum_t X_{t+1}^2}$$  \hspace{1cm} (9)$$

for

$$Y_t = \beta_1 X_{t+1} + u_t$$  \hspace{1cm} (10)$$

Note that the disturbance $u_t$ is

$$u_t = \varepsilon_t - \beta_1 \varepsilon_{xt+1} - \beta_1 \lambda \epsilon_{xt}$$

as a result of using (7) to replace

$$E(X_{t+1} \mid I_{t-1})$$

in the structural equation by the right hand side of

$$E(X_{t+1} \mid I_{t-1}) = X_{t+1} - \epsilon_{xt+1} - \lambda \epsilon_{xt}.$$
The detailed algebra for finding \( \text{plim}(\hat{\beta}_1) \):

\[
\text{plim}(\hat{\beta}_1) = \text{plim}(\frac{\sum_t (X_{t+1}) Y_t}{\sum_t X^2_{t+1}})
\]

\[
= \beta_1 + \text{plim}(\frac{\sum_t (X_{t+1}) u_t}{\sum_t X^2_{t+1}})
\]

\[
= \beta_1 + \frac{1}{\text{Var}(X_t)} \text{plim}(\frac{1}{T} \sum_t (X_{t+1})(\epsilon_t - \beta_1 \epsilon_{xt+1} - \beta_1 \lambda \epsilon_{xt}))
\]

\[
= \beta_1 + \frac{1}{\text{Var}(X_t)} \text{plim}(\frac{1}{T} \sum_t (\lambda^2 X_{t-1} + \lambda \epsilon_{xt} + \epsilon_{xt+1})(\epsilon_t - \beta_1 \epsilon_{xt+1} - \beta_1 \lambda \epsilon_{xt}))
\]

\[
= \beta_1 + \frac{-\beta_1 \lambda^2 \sigma^2_x - \beta_1 \sigma^2_x}{\text{Var}(X_t)} = \beta_1 + \frac{-\beta_1 \lambda^2 \sigma^2_x - \beta_1 \sigma^2_x}{\frac{\sigma^2_x}{1-\lambda^2}}
\]

\[
= \beta_1 + \frac{-\beta_1 \lambda^2 \sigma^2_x - \sigma^2_x}{1-\lambda^2} = \beta_1 (1 - \frac{-\lambda^2 \sigma^2_x - \sigma^2_x}{1-\lambda^2}) = \beta_1 \lambda^4
\]

This means that the OLS estimator of \( \beta_1 \) in (10) is inconsistent when the true model is (1)-(5). The asymptotic bias is:

\[
\text{plim}(\hat{\beta}_1) - \beta_1 = \beta_1 (\lambda^4 - 1) < 0 \tag{11}
\]

Exercise:

Assume that (1) is replaced by:

\[
Y_t = \beta_1 E(X_t | I_{t-1}) + \epsilon_t
\]

but that the model is otherwise kept unchanged. Show that the bias becomes:

\[
\text{plim}(\hat{\beta}_1) - \beta_1 = \beta_1 (\lambda^2 - 1) < 0.
\]

2 Errors-in-variables bias

The source of the OLS bias is that we “contaminate” the disturbance term by the forecast error for the \( X_{t+1} \) (or \( X_t \), depending on the specification of (1)). Therefore, the bias is a special case of the errors-in-variables bias. Another famous example is the measurement-error bias.

The measurement-error bias is often presented for the cross-section data model:

\[
Y_i = \beta_0 + \beta_1 X_i^* + \epsilon_i^*, \quad i = 1, 2, \ldots, n \tag{12}
\]

where \( \text{Cov}(\epsilon_i^*, X_i^*) = 0 \), and \( \epsilon_i^* \) has the other classical properties as well. We assume that \( X_i^* \) is an unobservable random variable which is replaced by the observable \( X_i \) in the estimation of (12). The difference between \( X_i \) and \( X_i^* \) is random:

\[
e_i = X_i - X_i^*
\]
Even if all \( \epsilon_i \) and \( \epsilon_i^* \) are independent, OLS on

\[
Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad i = 1, 2, \ldots, n
\]  

(13)

will produce an inconsistent estimator of \( \beta_1 \), because \( \text{Cov}(\epsilon_i, X_i) \neq 0 \) as a consequence of

\[
\epsilon_i = \epsilon_i^* - \beta_1 e_i
\]

As usual, the probability limit of \( \hat{\beta}_1 - \beta_1 \) is

\[
\text{plim}(\hat{\beta}_1 - \beta_1) = \frac{\text{plim}}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \epsilon_i
\]  

(14)

The denominator of (14) is equal to the theoretical variance of \( X \), which is the sum of the variance of \( X^* \), \( \sigma^2_{X^*} \), and the measurement-error variance, \( \sigma^2_e \):

\[
\text{Var}(X) = \sigma^2_{X^*} + \sigma^2_e
\]  

(15)

The numerator is

\[
\text{plim} \left[ \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}) \epsilon_i \right] = \text{plim} \left[ \frac{1}{n} \sum_{i=1}^{n} (X_i^* - \bar{X}^*) + (e_i - \bar{e})(\epsilon^*_i - \beta_1 e_i) \right]
\]

\[
= -\beta_1 \text{Var}(e) = -\beta_1 \sigma^2_e
\]

since the probability limit of all the other terms are zero by the assumptions of the model. Collecting results, we have the compact expressions

\[
\text{plim}(\hat{\beta}_1 - \beta_1) = \frac{-\beta_1 \sigma^2_e}{\text{Var}(X)} = -\beta_1 \frac{\sigma^2_e}{\sigma^2_{X^*} + \sigma^2_e}
\]  

(16)

for the asymptotic bias of the OLS estimator in the measurement-error model.

To see the errors-in-variables interpretation of the RE bias of the OLS estimator (11) in the model (1)-(5), set

\[
\sigma^2_e \equiv \text{Var}(u_t) = \text{Var}(\epsilon_{xt+1} + \lambda \epsilon_{xt}) = \sigma^2_x + \lambda^2 \sigma^2_x
\]

\[
\text{Var}(X_t) \equiv \frac{\sigma^2_x}{1 - \lambda^2}
\]

and insert in (16):

\[
\text{plim}(\hat{\beta}_1 - \beta_1) = \frac{-\beta_1 \sigma^2_e}{\text{Var}(X)} = -\beta_1 (\sigma^2_x + \lambda^2 \sigma^2_x)
\]

\[
= -\beta_1 (1 + \lambda^2)(1 - \lambda^2)
\]

\[
= -\beta_1 (1 - \lambda^4)
\]

which is the same expression as in (11) above.
3 The Lucas-critique

The Lucas-critique attacks the idea that if there is a change in the expectations about $X_{t+1}$, the effect of this change on $Y_t$ can be predicted by using the OLS estimate $\hat{\beta}_1$ from a regression model where $Y_t$ is regressed on $X_{t+1}$.$^2$ By “change in the expectations” the critique means a change in a parameter of the equations that are used to form the mathematical expectation of $X_{t+1}$. By looking at (11) we see that the relevant parameter must be $\lambda$. We see that we if $\lambda$ change—we call this a structural break in the $X_t$ process—the probability limit of $\hat{\beta}_1$ must also change. This means that the estimated effect of a change in $X$ on $Y$ will be wrongly estimated by the use of OLS estimation of a conditional model.

More generally, the critique implies that policy analysis cannot be based on OLS estimated conditional expectations (regression models). Empirically, the relevance of the critique can be confirmed by finding proof of a structural break in the conditional model that occurs at the same time as a structural break in model for $X$. However, the relevance can also be refuted if the structural break in the conditional expectation for $X$ does not lead to a structural break in the conditional model.

4 Causality

If indeed the conditional relationship between $Y_t$ and $X_t$ is invariant to breaks in the marginal model of $X_t$, then $X_t$ is deemed to cause the resultant changes in $Y_t$. In effect the billions of real-world economic decisions that underlie our measured time series is a running laboratory that churns out a relentless flow of structural breaks. The power of econometric tests to detect changes in the parameters of the conditional model when a break in the marginal model occurs depends on the magnitude of the change in the latter, and is a practical matter, which does not raise any issue of principle: when breaks in the $X_t$ process are large and intermittent, while the parameters of the conditional model remain constant, it seems reasonable to conclude that $X_t$ causes $Y_t$.

$^2$As noted above, and as shown in Lecture 17, the same applies if the structural equation is specified with $E(X_t \mid I_{t-1})$ as the explanatory variable for $Y_t$. 

5