OLS bias for econometric models with errors-in-variables. The Lucas-critique Supplementary note to Lecture 17 in ECON 4150 spring term 2013

RNy

September 17, 2013

1 Properties of OLS in RE models

In Lecture 17 we discussed the following example of a rational expectations (RE) model:

$$Y_t = \beta_1 E(X_{t+1} \mid \mathcal{I}_{t-1}) + \varepsilon_t \tag{1}$$

$$X_t = \lambda X_{t-1} + \epsilon_{xt}, \quad -1 < \lambda < 1 \tag{2}$$

$$\varepsilon_t \sim IID(0, \sigma^2) \tag{3}$$

$$\epsilon_{xt} \sim IID(0, \sigma_x^2) \tag{4}$$

$$Cov(\varepsilon_t, \epsilon_{xs}) = 0 \text{ for all } t \text{ and } s \tag{5}$$

(1) is the structural equation and β_1 is the parameter of interest of this model:[1] β_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 $\mathcal{I}_{t-1}$ in (1).

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

The disturbances ε_t and ϵ_{xt} are independent from each other, as stated in assumption (5) and they have classical properties conditional on the agents'

[1] The model is specified without an intercept, without loss of generality, and because it simplifies the algebra.

information set $\mathcal{I}_{t-1}$. Without loss of generality, this is conveyed by the assumptions that each disturbance is IID with expectation zero and constant variance.

As shown in the lectures, given the model specification, the solution for X_{t+1} is

$$X_{t+1} = \lambda^2 X_{t-1} + \lambda\epsilon_{xt} + \epsilon_{xt+1} \tag{6}$$

with conditional mathematical expectation

$$E(X_{t+1} \mid \mathcal{I}_{t-1}) = \lambda^2 X_{t-1}. \tag{7}$$

(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

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

The expectation error is

$$X_{t+1} - E(X_{t+1} \mid \mathcal{I}_{t-1}) = \lambda\epsilon_{xt} + \epsilon_{xt+1} \tag{8}$$

We know that OLS gives (at least) consistent estimators of the conditional expectation of Y_t given X_{t+1}. But is β_1 a parameter in the conditional expectation? To answer that question we can investigate the probability limit of the OLS estimator

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

for

$$Y_t = \beta_1 X_{t+1} + u_t \tag{10}$$

Note that the disturbance u_t is

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

as a result of using (7) to replace

$$E(X_{t+1} \mid \mathcal{I}_{t-1})$$

in the structural equation by the right hand side of

$$E(X_{t+1} \mid \mathcal{I}_{t-1}) = X_{t+1} - \epsilon_{xt+1} - \lambda\epsilon_{xt}.$$

The detailed algebra for finding $plim(\hat{\beta}_1)$:

$$
\begin{aligned}
plim(\hat{\beta}_1) &= plim(\frac{\sum_t (X_{t+1})Y_t}{\sum_t X_{t+1}^2}) \\
&= \beta_1 + plim(\frac{\sum_t (X_{t+1})u_t}{\sum_t X_{t+1}^2}) \\
&= \beta_1 + \frac{1}{Var(X_t)} plim(\frac{1}{T}\sum_t (X_{t+1})(\varepsilon_t - \beta_1 \epsilon_{xt+1} - \beta_1 \lambda \epsilon_{xt})) \\
&= \beta_1 + \frac{1}{Var(X_t)} plim(\frac{1}{T}\sum_t (\lambda^2 X_{t-1} + \lambda\epsilon_{xt} + \epsilon_{xt+1})(\varepsilon_t - \beta_1 \epsilon_{xt+1} - \beta_1 \lambda \epsilon_{xt}) \\
&= \beta_1 + \frac{-\beta_1\lambda^2\sigma_x^2 - \beta_1\sigma_x^2}{Var(X_t)} = \beta_1 + \frac{-\beta_1\lambda^2\sigma_x^2 - \beta_1\sigma_x^2}{\frac{\sigma_x^2}{1-\lambda^2}} \\
&= \beta_1(1 + \frac{-\lambda^2\sigma_x^2 - \sigma_x^2}{\frac{\sigma_x^2}{1-\lambda^2}}) = \beta_1(1 - (\lambda^2+1)(1-\lambda^2) = \beta_1\left[1 - \lambda^2(1-\lambda^2) - (1-\lambda^2)\right] \\
&= \beta_1\left[-\lambda^2(1-\lambda^2) + \lambda^2\right] = \beta_1\lambda^4
\end{aligned}
$$

This means that the OLS estimator of β_1 in (10) is inconsistent when the true model is (1)-(5). The asymptotic bias is:

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

Note that in Lecture 17 we found that in the case where (1) is replaced by

$$Y_t = \beta_1 E(X_t \mid \mathcal{I}_{t-1}) + \varepsilon_t$$

and the model is otherwise kept unchanged, the bias is

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

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-varaibles* bias. As noted in Lecture 17, another famous example is the *measurement-error* bias.

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

$$Y_i = \beta_0 + \beta_1 X_i^* + \varepsilon_i^* \; i = 1, 2, \ldots, n \tag{12}$$

where $Cov(\varepsilon_i^*, X_i^*) = 0$, and ε_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 e_i and ε_i^* are independent, OLS on

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \;\; i = 1, 2, \ldots, n \tag{13}$$

will produce an inconsistent estimator of β_1, because $Cov(\varepsilon_i,\ X_i) \neq 0$ as a consequence of

$$\varepsilon_i = \varepsilon_i^* - \beta_1 e_i$$

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

$$plim(\hat{\beta}_1 - \beta_1) = plim \frac{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})\varepsilon_i}{\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2} \tag{14}$$

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,σ_e^2:

$$Var(X) = \sigma^2_{X^*} + \sigma_e^2 \tag{15}$$

The numerator is

$$\begin{aligned} plim\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})\varepsilon_i\right] &= plim\left[\frac{1}{n}\sum_{i=1}^{n}(X_i^* - \overline{X^*}) + (e_i - \bar{e}))(\varepsilon_i^* - \beta_1 e_i)\right] \\ &= -\beta_1 Var(e) = -\beta_1 \sigma_e^2 \end{aligned}$$

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

$$plim(\hat{\beta}_1 - \beta_1) = \frac{-\beta_1 \sigma_e^2}{Var(X)} = \frac{-\beta_1 \sigma_e^2}{\sigma^2_{X^*} + \sigma_e^2} \tag{16}$$

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

$$\begin{aligned} \sigma_e^2 &\equiv Var(u_t) = Var(\epsilon_{xt+1} + \lambda \epsilon_{xt}) = \sigma_x^2 + \lambda^2 \sigma_x^2 \\ Var(X_t) &\equiv \frac{\sigma_x^2}{1 - \lambda^2} \end{aligned}$$

and insert in (16):

$$\begin{aligned} plim(\hat{\beta}_1 - \beta_1) = \frac{-\beta_1 \sigma_e^2}{Var(X)} &= \frac{-\beta_1(\sigma_x^2 + \lambda^2 \sigma_x^2)}{\frac{\sigma_x^2}{1-\lambda^2}} \\ &= -\beta_1(1 + \lambda^2)(1 - \lambda^2) \\ &= -\beta_1(1 - \lambda^4) \end{aligned}$$

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 λ. We see that we if λ 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. The solution, as we have discussed in Lectures 18 and 19, is to use IV estimation.

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.

[2] As noted above, and as shown in Lecture 17, the same applies if the stuctural equation is specified with $E(X_t \mid \mathcal{I}_{t-1})$ as the explanatory varaible for Y_t..