Supplemental Materials to Lab 1 of QM3

Shane Xinyang Xuan
ShaneXuan.com

Department of Political Science
University of California, San Diego

April 4, 2018

1 Preliminary: Matrix set-up

We can write the linear regression in matrix algebra:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ 1 & x_{21} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix}_{n\times(k+1)} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}_{(k+1)\times 1} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}_{n\times 1} \tag{1}$$

And we can write

$$y = \mathbf{X}\beta + e, \tag{2}$$

where y is $n \times 1$, $\mathbf{X}$ is $n \times (k+1)$, β is $(k+1) \times 1$, and e is $n \times 1$.

2 Omitted Variable Bias (OVB)

Consider the regression of y on $\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix} \in \mathbb{R}^{n\times(k+1)}$, where X_i is an $[1 \times (k+1)]$ row vector, controlling for $A = \begin{bmatrix} A_1 \\ \vdots \\ A_n \end{bmatrix}$, where A_i is a scalar, and is an unobservable:

$$y_i = \alpha + X_i\rho + A_i\gamma + e_i \tag{3}$$

What happens if we use OLS estimator when there is an omitted variable bias? Recall that

$$\hat{\rho} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T y \tag{4}$$

where $y = \mathbf{X}\rho + A\gamma + \xi$, and it follows that

$$\begin{aligned} \hat{\rho} &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T(\mathbf{X}\rho + A\gamma + \xi) && (5) \\ &= (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{X}\rho + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T A\gamma + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\xi && (6) \\ &= \rho + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T A\gamma + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\xi && (7) \end{aligned}$$

Taking expectations, we obtain

$$\mathbb{E}[\hat{\rho}|\mathbf{X}] = \rho + (\mathbf{X}^T\mathbf{X})^{-1}\mathbb{E}[\mathbf{X}^T A|\mathbf{X}]\gamma \tag{8}$$

That is, the bias comes from $\mathbf{X}^T A$ (which is non-zero if the omitted variable A is correlated with X_i) and γ (which is non-zero if A has an effect on y).

Another way to understand the situation is to think of a short regression obtained after we leave A_i out from Eq (3). The short regression coefficient is given by

$$\underbrace{\frac{\text{cov}(y_i, X_i)}{\text{var}(X_i)}}_{\text{short coef.}} = \rho + \delta_{AX}\gamma \tag{9}$$

where

$$\gamma = \text{effect of the omitted on the outcome}$$
$$\delta_{AX} = \text{vector of coefficients from regressions of the elements of } A_i \text{ on } X_i$$

That is, *short equals long plus the effect of omitted times the regression of omitted on included.*

Remark The direction of bias can be summarized:

	$\text{cov}(X_i, A_i) > 0$	$\text{cov}(X_i, A_i) < 0$
$\gamma > 0$	positive	negative
$\gamma < 0$	negative	positive

Table 1: Direction of bias due to omitted variables.

2.1 Omitted Variable Bias: A Simulation

```
clear
set obs 300
gen iq = rnormal(100,20)
gen educ = iq/10 + uniform()*2
gen income = iq*500 + educ*1000 + rnormal(0,3000)

* short
reg income educ
scalar short=_b[educ]

* delta is the the regression of omitted on included
reg iq educ
scalar delta=_b[educ]

* long
reg income educ iq
* gamma is the effect of omitted
scalar gamma=_b[iq]
* rho is long
scalar rho=_b[educ]

* Short equals long plus the effect of omitted times the regression of omitted on included
disp rho + delta * gamma
disp short
```

3 Simultaneity Bias

Consider the following model:

$$\begin{bmatrix} z \\ x \end{bmatrix} = \begin{bmatrix} a_z \\ a_x \end{bmatrix} + \begin{bmatrix} 0 & b_{zx} \\ b_{xz} & 0 \end{bmatrix} \begin{bmatrix} z \\ x \end{bmatrix} + \begin{bmatrix} \varepsilon_z \\ \varepsilon_x \end{bmatrix} \tag{10}$$

which can be written in a more compact way:

$$y = a + \mathbf{B}y + \varepsilon \tag{11}$$

We try to solve for y:

$$y - \mathbf{B}y = a + \varepsilon \tag{12}$$

$$(\mathbf{I} - \mathbf{B})y = a + \varepsilon \tag{13}$$

$$y = (\mathbf{I} - \mathbf{B})^{-1}(a + \varepsilon) \tag{14}$$

We now write the matrix-inverse multiplier in its general form:

$$(\mathbf{I} - \mathbf{B})^{-1} = \frac{1}{1 - b_{zx}b_{xz}} \begin{bmatrix} 1 & b_{zx} \\ b_{xz} & 1 \end{bmatrix} \tag{15}$$

And we can re-write y as

$$\begin{bmatrix} z \\ x \end{bmatrix} = \frac{1}{1 - b_{zx}b_{xz}} \begin{bmatrix} 1 & b_{zx} \\ b_{xz} & 1 \end{bmatrix} \left(\begin{bmatrix} a_z \\ a_x \end{bmatrix} + \begin{bmatrix} \varepsilon_z \\ \varepsilon_x \end{bmatrix} \right) \tag{16}$$

and z as

$$z = \frac{1}{1 - b_{zx}b_{xz}} \left[a_z + e_z + b_{zx}(a_x + e_x) \right] \tag{17}$$

In the case of an exogenous shock in x, there will be feedback effect as illustrated in Eq. (17). For example, a one-unit change in a_x will result in $\Delta z^{t=1} = \frac{1}{1 - b_{zx}b_{xz}} b_{zx}$ at time $t = 1$. Note that the change in z will further affect x, which recursively affects $\Delta z^{t=j}$, where $j > 1$. As a result of the *simultaneity bias*, the full response of z to the exogenous shock in x is larger than the "true" value.