Econ 203B: Single Equation Models

Model Assumptions

Michael Powell
Department of Economics, UCLA
February 10th, 2006

So far in this class, we have dealt with several different econometric models, all of which have been similar and all of which have been characterized by a series of assumptions. The purpose of these notes is to help make clear which model is accompanied by which set of assumptions and to highlight the implications of these assumptions.

1 The Classical Linear Regression (CLR)

For all that follows, define $Y = \begin{bmatrix} Y_1 \\ \vdots \\ Y_n \end{bmatrix}$, $X = \begin{bmatrix} X_1 \\ \vdots \\ X_n \end{bmatrix}$ where $X_j = \begin{bmatrix} X_{j1} & \cdots & X_{jk} \end{bmatrix}$, $\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix}$. The classical linear regression model consists of the following assumptions: For all i,

Assumption 1 (Linearity) $Y_i = X_i\beta + \varepsilon_i$

Assumption 2 (Strict Exogeneity) $E\left[\varepsilon_i | X\right] = 0$

Assumption 3 (Spherical Errors) $E\left[\varepsilon_i^2 \middle| X\right] = \sigma^2 > 0$, $E\left[\varepsilon_i\varepsilon_j | X\right] = 0$

Assumption 4 (Full Rank) The $n \times K$ matrix X has rank K with probability 1.

1.1 Discussion

Assumption 1 (Linearity) is carried thoughout most of these models. This allows us to use a linear estimator for the $\beta_j' s$ with good conscience. (That is, without this assumption, any linear estimator of the $\beta_j' s$ would probably not be consistent.)

Assumption 2 (Strict Exogeneity) is a very strong assumption which also implies many other weaker characteristics of the data:

Proposition 1 *Strict exogeneity implies zero unconditional means of the error terms. That is,* $E\left[\varepsilon_i | X\right] = 0$ $\forall i \Rightarrow E\left[\varepsilon_i\right] = 0$.

Proof.

$$E\left[\varepsilon_i\right] = E\left[E\left[\varepsilon_i | X\right]\right] = E\left[0\right] = 0$$

By the law of iterated expectations. In particular, $E\left[\varepsilon_i\right] = 0$. ■

Proposition 2 *Strict exogeneity implies orthogonality between each regressor and each error term. That is,* $E\left[\varepsilon_i | X\right] = 0$ $\forall i \Rightarrow E\left[X_{ik}\varepsilon_j\right] = 0$ $\forall i, j$.

Proof.

$$E\left[X_{ik}\varepsilon_j\right] = E\left[E\left[X_{ik}\varepsilon_j | X\right]\right] = E\left[X_{ik}E\left[\varepsilon_j | X\right]\right] = E\left[X_{ik} \cdot 0\right] = 0$$

Where I made use of the law of iterated expectations. Indeed, $E\left[X_{ik}\varepsilon_j\right] = 0$. Note that this holds both when $i = j$ and when $i \neq j$. ■

Proposition 3 *Strict exogeneity implies that each regressor is unconditionally uncorrelated with each error term. That is,* $E\left[\varepsilon_i | X\right] = 0$ $\forall i \Rightarrow Cov\left(X_{ik}, \varepsilon_j\right) = 0$ $\forall i, j$.

Proof.

$$\begin{aligned} Cov\left(X_{ik},\varepsilon_j\right) &= E\left[X_{ik}\varepsilon_j\right]-E\left[X_{ik}\right]E\left[\varepsilon_j\right] \\ &= 0-0=0 \end{aligned}$$

Where the second equality holds since $E\left[X_{ik}\varepsilon_j\right]=0\ \forall i,j$ by proposition 2 and $E\left[\varepsilon_j\right]=0$ by proposition 3. ■

Assumption 3 (Spherical Errors) combines two assumptions: 1) conditional homoskedasticity and 2) no conditional serial correlation. It can be shown that conditional homoskedasticity also implies unconditional homoskedasticity when assumption 2 (Strict Exogeneity) holds and that no conditional serial correlation implies no unconditional serial correlation.

Proposition 4 *Conditional homoskedasticity implies unconditional homoskedasticity. That is,* $E\left[\left.\varepsilon_i^2\right|X\right]=0\ \forall i\Rightarrow E\left[\varepsilon_i^2\right]=0.$

Proof. First note that under strict exogeneity,

$$\begin{aligned} Var\left(\left.\varepsilon_i\right|X\right) &= E\left[\left.\varepsilon_i^2\right|X\right]+E\left[\left.\varepsilon_i\right|X\right]\underbrace{E\left[\left.\varepsilon_i\right|X\right]}_{=0} \\ &= E\left[\left.\varepsilon_i^2\right|X\right] \end{aligned}$$

And, using the result from the notes on unconditional variance,

$$Var\left(\varepsilon_i\right)=Var\left(E\left[\left.\varepsilon_i\right|X\right]\right)+E\left[Var\left(\left.\varepsilon_i\right|X\right)\right]$$

We have:

$$\begin{aligned} Var\left(\varepsilon_i\right) &= Var\left(0\right)+E\left[Var\left(\left.\varepsilon_i\right|X\right)\right]=0+E\left[E\left[\left.\varepsilon_i^2\right|X\right]\right] \\ &= E\left[\left.\varepsilon_i^2\right|X\right]=0 \end{aligned}$$

But note that

$$\begin{aligned} Var\left(\varepsilon_i\right) &= E\left[\varepsilon_i^2\right]-E\left[\varepsilon_i\right]\underbrace{E\left[\varepsilon_i\right]}_{=0} \\ &= E\left[\varepsilon_i^2\right] \end{aligned}$$

Therefore, $E\left[\varepsilon_i^2\right]=0$. ■

Proposition 5 *No conditional serial correlation implies no unconditional serial correlation. That is,* $E\left[\left.\varepsilon_i\varepsilon_j\right|X\right]=0\Rightarrow E\left[\varepsilon_i\varepsilon_j\right]=0.$

Proof.

$$E\left[\varepsilon_i\varepsilon_j\right]=E\left[E\left[\left.\varepsilon_i\varepsilon_j\right|X\right]\right]=E\left[0\right]=0$$

By the law of iterated expectations. ■

Finally, assumption 4 (Full Rank) rules out perfect collinearity, which allows us to derive a unique estimator for this model. It does so by ensuring that the matrix $X'X$ is invertible. (That is, $rank\left(X'X\right)=K$)

1.2 Alternative Characterizations

As has been shown in class, there is more than one way of writing down all these assumptions, which can be useful depending on the situation. For example, sometimes it is convenient not to have to write the ε_i's explicity and other times, we would prefer to have a matrix characterization of the model.

1.2.1 Without mention of ε_i

For all i and for all $i \neq j$,

Assumption 1 $E\left[Y_i | X\right] = X_i \beta$

Assumption 2 $V\left(Y_i | X\right) = \sigma^2$

Assumption 3 $Cov\left(Y_i, Y_j | X\right) = 0$

Assumption 4 (Full Rank) The $n \times K$ matrix X has rank K with probability 1.

1.2.2 Matrix notation with mention of ε

Assumption 1 $Y = X\beta + \varepsilon$

Assumption 2 $E\left[\varepsilon | X\right] = 0$

Assumption 3 $V\left(\varepsilon | X\right) = \sigma^2 I_n$

Assumption 4 (Full Rank) The $n \times K$ matrix X has rank K with probability 1.

1.2.3 Matrix notation without mention of ε

Assumption 1 $E\left[Y | X\right] = X\beta$

Assumption 2 $V\left(Y | X\right) = \sigma^2 I_n$

Assumption 3 (Full Rank) The $n \times K$ matrix X has rank K with probability 1.

Here, the linearity assumption was combined with the strict exogeneity assumption to produce assumption 1, which is described as "linearity of conditional expectations."