Chapter 15
Instrumental Variables Estimation

A basic assumption in analyzing the performance of estimators in multiple regression is that the explanatory variables and disturbance terms are independently distributed. The violation of such assumption disturbs the optimal properties of the estimators. The instrumental variable estimation method helps in estimating the regression coefficients in the multiple linear regression model when such violation occurs.

Consider the multiple linear regression model

$$y = X\beta + \varepsilon$$

where y is $(n \times 1)$ vector of observation on study variable, X is $(n \times k)$ matrix of observations on $X_1, X_2, ..., X_k$, β is a $(k \times 1)$ vector of regression coefficient and ε is a $(n \times 1)$ vector of disturbances.

Suppose one or more explanatory variables is correlated with the disturbances in the limit, then we can write

$$\text{plim}\left(\frac{1}{n}X'\varepsilon\right) \neq 0.$$

The consequences of such an assumption on ordinary least squares estimator are as follows:

$$\begin{aligned} b &= \left(X'X\right)^{-1}X'y \\ &= \left(X'X\right)^{-1}X'\left(X\beta+\varepsilon\right) \\ b-\beta &= \left(X'X\right)^{-1}X'\varepsilon \\ &= \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'\varepsilon}{n}\right) \\ \text{plim}\left(b-\beta\right) &= \text{plim}\left(\frac{X'X}{n}\right)^{-1}\text{plim}\left(\frac{X'\varepsilon}{n}\right) \\ &\neq 0 \end{aligned}$$

assuming $\text{plim}\left(\frac{X'X}{n}\right) = \Sigma_{XX}$ exists and is nonsingular. Consequently $\text{plim}\, b \neq \beta$ and thus the OLSE becomes an inconsistent estimator of β.

To overcome this problem and to obtain a consistent estimator of β, the instrumental variable estimation can be used.

Consider the model

$$y = X\beta + \varepsilon \text{ with } \operatorname{plim}\left(\frac{1}{n}X'\varepsilon\right) \neq 0.$$

Suppose that it is possible to find a data matrix Z of order $(n \times k)$ with the following properties.:

(i) $\operatorname{plim}\left(\frac{Z'X}{n}\right) = \Sigma_{ZX}$ is a finite and nonsingular matrix of full rank. This interprets that the variables in Z are correlated with those in X, in the limit.

(ii) $\operatorname{plim}\left(\frac{Z'\varepsilon}{n}\right) = 0,$

i.e., the variables in Z are uncorrelated with ε, in the limit.

(iii) $\operatorname{plim}\left(\frac{Z'Z}{n}\right) = \Sigma_{ZZ}$ exists.

Thus $Z-$variables are postulated to be

- uncorrelated with ε, in the limit and
- to have a nonzero cross product with X.

Such variables are called **instrumental variables**.

If some of X variables are likely to be uncorrelated with ε, then these can be used to form some of the columns of Z and extraneous variables are found only for the remaining columns.

First, we understand the role of the term $X'\varepsilon$ in the OLS estimation. The OLSE b of β is derived by solving the equation

$$\frac{\partial (y - X\beta)'(y - X\beta)}{\partial \beta} = 0$$

or $\quad X'y = X'Xb$

or $\quad X'(y - Xb) = 0.$

Now we look at this normal equation as if it is obtained by pre-multiplying $y = X\beta + \varepsilon$ by X' as

$$X'y = X'X\beta + X'\varepsilon$$

where the term $X'\varepsilon$ is dropped and β is replaced by b. The disappearance of $X'\varepsilon$ can be explained when X and ε are uncorrelated as follows:

$$\begin{aligned} X'y &= X'X\beta + X'\varepsilon \\ \frac{X'y}{n} &= \frac{X'X}{n}\beta + \frac{X'\varepsilon}{n} \\ \text{plim}\left(\frac{X'y}{n}\right) &= \text{plim}\left(\frac{X'X}{n}\right)\beta + \text{plim}\left(\frac{X'\varepsilon}{n}\right) \\ \Rightarrow \beta &= \text{plim}\left(\frac{X'X}{n}\right)^{-1}\left[\text{plim}\left(\frac{X'y}{n}\right) - \text{plim}\left(\frac{X'\varepsilon}{n}\right)\right]. \end{aligned}$$

Let

$$\begin{aligned} \text{plim}\left(\frac{X'X}{n}\right) &= \Sigma_{XX} \\ \text{plim}\left(\frac{X'y}{n}\right) &= \Sigma_{Xy} \end{aligned}$$

where population cross moments Σ_{XX} and Σ_{Xy} are finite, Σ_{XX} is finite and nonsingular.

If X and ε are uncorrelated so that

$$\text{plim}\left(\frac{X'\varepsilon}{n}\right) = 0,$$

then

$$\beta = \Sigma_{XX}^{-1}\Sigma_{Xy}.$$

If Σ_{XX} is estimated by sample cross moment $\frac{X'X}{n}$ and Σ_{Xy} is estimated by sample cross moment $\frac{X'y}{n}$, then the OLS estimator of β is obtained as

$$\begin{aligned} b &= \left(\frac{X'X}{n}\right)^{-1}\left(\frac{X'y}{n}\right) \\ &= (X'X)^{-1}X'y. \end{aligned}$$

Such an analysis suggests to use Z to pre-multiply the multiple regression model as follows:

$$Z'y = Z'X\beta + Z'\varepsilon$$

$$\frac{Z'y}{n} = \left(\frac{Z'X}{n}\right)\beta + \left(\frac{Z'\varepsilon}{n}\right)$$

$$\text{plim}\left(\frac{Z'y}{n}\right) = \text{plim}\left(\frac{Z'X}{n}\right)\beta + \text{plim}\left(\frac{Z'\varepsilon}{n}\right)$$

$$\Rightarrow \beta = \text{plim}\left(\frac{Z'X}{n}\right)^{-1}\left[\text{plim}\left(\frac{Z'y}{n}\right) - \text{plim}\left(\frac{Z'\varepsilon}{n}\right)\right]$$

$$= \Sigma_{ZX}^{-1}\Sigma_{Zy}.$$

Substituting the sample cross moment

- $\frac{Z'X}{n}$ of Σ_{ZX} and
- $\frac{Z'y}{n}$ of Σ_{Zy},

Thus the following instrumental variable estimator of β is obtained:

$$\hat{\beta}_{IV} = (Z'X)^{-1} Z'y$$

which is termed as an **instrumental variable estimator** of β and this method is called an **instrumental variable method.**

Since

$$\hat{\beta}_{IV} - \beta = (Z'X)^{-1} Z'(X\beta + \varepsilon) - \beta$$

$$= (Z'X)^{-1} Z'\varepsilon$$

$$\text{plim}\left(\hat{\beta}_{IV} - \beta\right) = \text{plim}\left[(Z'X)^{-1} Z'\varepsilon\right]$$

$$= \text{plim}\left[\left(\frac{Z'X}{n}\right)^{-1}\frac{Z'\varepsilon}{n}\right]$$

$$= \text{plim}\left(\frac{Z'X}{n}\right)^{-1}\text{plim}\left(\frac{Z'\varepsilon}{n}\right)$$

$$= \Sigma_{ZX}^{-1}.0$$

$$\Rightarrow \text{plim}\hat{\beta}_{IV} = \beta.$$

Thus the instrumental variable estimator is consistent. Note that the variables $Z_1, Z_2, ..., Z_k$ in Z are chosen such that they are uncorrelated with ε and correlated with X, at least asymptotically, so that the second order moment matrix Σ_{ZX} exists and is nonsingular.

Asymptotic distribution:

The asymptotic distribution of

$$\sqrt{n}\left(\hat{\beta}_{IV}-\beta\right)=\left(\frac{1}{n}Z'X\right)^{-1}\frac{1}{\sqrt{n}}Z'\varepsilon$$

is normal with mean vector 0 and the asymptotic covariance matrix is given by

$$\begin{aligned}
AsyVar\left(\hat{\beta}_{IV}\right) &= \text{plim}\left[\left(\frac{Z'X}{n}\right)^{-1}\frac{1}{n}Z'E\left(\varepsilon\varepsilon'\right)Z\left(\frac{X'Z}{n}\right)^{-1}\right] \\
&= \sigma^2\,\text{plim}\left[\left(\frac{Z'X}{n}\right)^{-1}\left(\frac{Z'Z}{n}\right)\left(\frac{X'Z}{n}\right)^{-1}\right] \\
&= \sigma^2\,\text{plim}\left(\frac{Z'X}{n}\right)^{-1}\text{plim}\left(\frac{Z'Z}{n}\right)\text{plim}\left(\frac{X'Z}{n}\right)^{-1} \\
&= \sigma^2\Sigma_{XZ}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1}.
\end{aligned}$$

For a large sample,

$$V\left(\hat{\beta}_{IV}\right)=\frac{\sigma^2}{n}\Sigma_{XZ}^{-1}\Sigma_{ZZ}\Sigma_{ZX}^{-1}$$

which can be estimated by

$$\hat{V}\left(\hat{\beta}_{IV}\right)=\frac{\hat{\sigma}^2}{n}\hat{\Sigma}_{XZ}^{-1}\hat{\Sigma}_{ZZ}\hat{\Sigma}_{ZX}^{-1}.$$

$$=\frac{s^2}{n}\left(Z'X\right)^{-1}Z'Z\left(X'Z\right)^{-1}$$

where

$$s^2=\frac{1}{n-k}\left(y-Xb\right)'(y-Xb),$$
$$b=\left(X'X\right)^{-1}X'y.$$

The variance of $\hat{\beta}_{IV}$ is not necessarily a minimum asymptotic variance because there can be more than one sets of instrumental variables that fulfil the requirement of being uncorrelated with ε and correlated with stochastic regressors.