控制函数法与2SLS系数的联系

我们知道，在2SLS中

$$\hat{\beta}_{\mathrm{2SLS}} = (\hat{\mathrm{X}}^{\mathrm{T}}\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}^{\mathrm{T}}\mathrm{y}$$

其中

$$\hat{\mathrm{X}} = (\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{X}$$

Z是$N \times L$的工具变量矩阵，X是$N \times K$的原始变量矩阵，一般有$L > K$。

而控制函数（Control Function, CF）法的思路与2SLS相反：2SLS将X对Z回归的预测值（投影）当做"干净的"变量重新与y做回归；而CF将X对Z回归的残差当做内生性的来源，与X一并与y做回归，其DGP为：

$$\mathrm{y} = [\mathrm{X} \quad \mathrm{M_Z X}]\,\beta_{\mathrm{CF}} + \epsilon_{\mathrm{CF}}$$

其中$\mathrm{M_Z} \equiv \mathrm{I} - \mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}$，是Z的残差矩阵。

我们也可以同时将β_{CF}分块，即：

$$\mathrm{y} = [\mathrm{X} \quad \mathrm{M_Z X}]\begin{bmatrix}\beta_{\mathrm{CF1}} \\ \beta_{\mathrm{CF2}}\end{bmatrix} + \epsilon_{\mathrm{CF}}$$

注意：虽然$\hat{\beta}_{\mathrm{2SLS}}$可以通过$\mathrm{y} \sim \hat{\mathrm{X}}$的最小二乘求出，但这只是数学上的相等（$\hat{\mathrm{X}}^{\mathrm{T}}\hat{\mathrm{X}} = \hat{\mathrm{X}}^{\mathrm{T}}\mathrm{X}$），其DGP依然是$\mathrm{y} = \mathrm{X}\beta_{\mathrm{2SLS}} + \epsilon_{\mathrm{2SLS}}$，而不是$\mathrm{y} = \hat{\mathrm{X}}\beta_{\mathrm{2SLS}} + \epsilon_{\mathrm{2SLS}}$。

有趣的是，$\hat{\beta}_{\mathrm{2SLS}} = \hat{\beta}_{\mathrm{CF1}}$，接下来我们就要证明这一等式。

由OLS知

$$\begin{aligned}\begin{bmatrix}\hat{\beta}_{\mathrm{CF1}} \\ \hat{\beta}_{\mathrm{CF2}}\end{bmatrix} &= [\mathrm{X} \quad \mathrm{M_Z X}]^{\mathrm{T}}[\mathrm{X} \quad \mathrm{M_Z X}][\mathrm{X} \quad \mathrm{M_Z X}]^{\mathrm{T}}\mathrm{y} \\ &= \begin{bmatrix}\mathrm{X}^{\mathrm{T}}\mathrm{X} & \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X} \\ \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X} & \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X}\end{bmatrix}^{-1}\begin{bmatrix}\mathrm{X}^{\mathrm{T}} \\ \mathrm{X}^{\mathrm{T}}\mathrm{M_Z}\end{bmatrix}\mathrm{y}\end{aligned}$$

这里用到了残差矩阵$\mathrm{M_Z}$的对称幂等性。

由分块逆矩阵公式知

$$\begin{bmatrix}\mathrm{X}^{\mathrm{T}}\mathrm{X} & \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X} \\ \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X} & \mathrm{X}^{\mathrm{T}}\mathrm{M_Z X}\end{bmatrix}^{-1} = \begin{bmatrix}(\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1} & -(\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1} \\ -(\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1} & (\mathrm{X}^{\mathrm{T}}\mathrm{M_Z X})^{-1} - (\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1}\end{bmatrix}$$

代入上式知

$$\begin{aligned}\hat{\beta}_{\mathrm{CF1}} &= ((\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}} - (\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}}\mathrm{M_Z})\mathrm{y} \\ &= (\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{y} \\ &= (\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})(\mathrm{I}-\mathrm{M_Z})\mathrm{X})^{-1}\mathrm{X}^{\mathrm{T}}(\mathrm{I}-\mathrm{M_Z})\mathrm{y} \\ &= (\underbrace{\mathrm{X}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}}\,\underbrace{\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}\mathrm{X}})^{-1}\,\underbrace{\mathrm{X}^{\mathrm{T}}\mathrm{Z}(\mathrm{Z}^{\mathrm{T}}\mathrm{Z})^{-1}\mathrm{Z}^{\mathrm{T}}}\,\mathrm{y} \\ &= (\hat{\mathrm{X}}^{\mathrm{T}}\hat{\mathrm{X}})^{-1}\hat{\mathrm{X}}^{\mathrm{T}}\mathrm{y} \\ &= \hat{\beta}_{\mathrm{2SLS}}\end{aligned}$$

原命题得证，这里同时用到了矩阵$\mathrm{I} - \mathrm{M_Z}$的对称幂等性。

另外可以关注$\mathrm{Var}(\hat{\beta}_{\mathrm{CF}})$与$\mathrm{Var}(\hat{\beta}_{\mathrm{2SLS}})$。我们知道

$$
\begin{aligned}
\mathrm{Var}(\hat{\beta}_{\mathrm{CF}}) &= \sigma^2_{\mathrm{CF}}\begin{bmatrix} \mathrm{X^TX} & \mathrm{X^TM_ZX} \\ \mathrm{X^TM_ZX} & \mathrm{X^TM_ZX} \end{bmatrix}^{-1} \\
&= \sigma^2_{\mathrm{CF}}\begin{bmatrix} (\mathrm{X^T(I-M_Z)X})^{-1} & -(\mathrm{X^T(I-M_Z)X})^{-1} \\ -(\mathrm{X^T(I-M_Z)X})^{-1} & (\mathrm{X^TM_ZX})^{-1}-(\mathrm{X^T(I-M_Z)X})^{-1} \end{bmatrix}
\end{aligned}
$$

即$\mathrm{Var}(\hat{\beta}_{\mathrm{CF1}}) = \sigma^2_{\mathrm{CF}}(\mathrm{X^T(I-M_Z)X})^{-1} = \sigma^2_{\mathrm{CF}}(\hat{\mathrm{X}}^{\mathrm{T}}\hat{\mathrm{X}})^{-1}$ 。而$\mathrm{Var}(\hat{\beta}_{\mathrm{2SLS}}) = \sigma^2_{\mathrm{2SLS}}(\hat{\mathrm{X}}^{\mathrm{T}}\hat{\mathrm{X}})^{-1}$，两者只相差一个常数。这说明，无论用CF还是2SLS，其变量对应的回归系数估计值都相同，同时任意量系数的协方差估计之比也相同，即：对$\forall 1 \le i, j, m, n \le K$，有

$$
\begin{gathered}
\hat{\beta}_{\mathrm{CF1},i} = \hat{\beta}_{\mathrm{2SLS},i} \\
\frac{cov(\hat{\beta}_{\mathrm{CF1},i}, \hat{\beta}_{\mathrm{CF1},j})}{cov(\hat{\beta}_{\mathrm{2SLS},i}, \hat{\beta}_{\mathrm{2SLS},j})} = \frac{cov(\hat{\beta}_{\mathrm{CF1},m}, \hat{\beta}_{\mathrm{CF1},n})}{cov(\hat{\beta}_{\mathrm{2SLS},m}, \hat{\beta}_{\mathrm{2SLS},n})}
\end{gathered}
$$