Econometrics	**October 22, 2023**

Topic 5: Two-Way Cluster-Robust (TWCR) Standard Errors

by Sai Zhang

Key points: The validity of Two-Way Cluster-Robust (TWCR) standard errors

Disclaimer: *This note is compiled by Sai Zhang.*

5.1 One-Way Clustering

First, consider the case of one-way clustering. The linear model with one-way clustering

$$y_{ig} = \mathbf{x}_{ig}\boldsymbol{\beta} + u_{ig}$$

where i denotes the ith of the N individuals in the sample, j denotes the gth of the G clusters, assume that

- $\mathbb{E}\left[u_{ig} \mid \mathbf{x}_{ig}\right] = 0$
- error independence across clusters: for $i \neq j$

$$\mathbb{E}\left[u_{ig}u_{jg'} \mid \mathbf{x}_{ig}, \mathbf{x}_{jg'}\right] = 0 \tag{5.1}$$

  unless $g = g'$, that is, errors for individuals within the same cluster may be correlated.

Grouping observations by cluster, get

$$\mathbf{y}_g = \mathbf{X}_g\boldsymbol{\beta} + \mathbf{u}$$

where $\mathbf{X}_g$ has dimension $N_g \times K$ and $\mathbf{y}_g$ has dimension $N_g \times 1$, with N_g observations in cluster g. Stacking over cluster, get the matrix form of the model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

with $\mathbf{y}$, $\mathbf{u}$ being $N \times 1$ vectors, $\mathbf{X}$ being an $N \times K$ matrix. OLS estimator gives

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}'\mathbf{X}\right)^{-1}\mathbf{X}'\mathbf{y} = \left(\sum_{g=1}^{G}\mathbf{X}_g'\mathbf{X}_g\right)^{-1}\sum_{g=1}^{G}\mathbf{X}_g'\mathbf{y}_g \tag{5.2}$$

then, by CLT, we have that $\sqrt{G}\left(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}\right) \xrightarrow{d} \mathcal{N}(0, \boldsymbol{\Sigma})$ where the variance matrix of the limit normal distribution $\boldsymbol{\Sigma}$ is

$$\left(\lim_{G\to\infty}\frac{1}{G}\sum_{g=1}^{G}\mathbf{E}\left[\mathbf{X}_g'\mathbf{X}_g\right]\right)^{-1}\left(\lim_{G\to\infty}\frac{1}{G}\sum_{g=1}^{G}\mathbf{E}\left[\mathbf{X}_g'\mathbf{u}_g'\mathbf{u}_g\mathbf{X}_g\right]\right)\times\left(\lim_{G\to\infty}\frac{1}{G}\sum_{g=1}^{G}\mathbf{E}\left[\mathbf{X}_g'\mathbf{X}_g\right]\right)^{-1} \tag{5.3}$$

If the primary source of clustering is due to group-level common shocks, a useful approximation is that for the jth regressor, the default OLS variance estimate based on $s^2\left(\mathbf{X}'\mathbf{X}\right)^{-1}$ should be inflated by $\tau_j \simeq 1 + \rho_{x_j}\rho_u\left(\overline{N}_g - 1\right)$, where

- s is the estimated standard deviation of the error

- ρ_{x_j} is a measure of within-cluster correlation of x_j
- ρ_u is the within-cluster error correlation
- $\overline{N}_g$ is the average cluster size

It's easy to see the τ_j can be large even with small ρ_u (Kloek, 1981; Scott and Holt, 1982; Moulton, 1990).

References

A Colin Cameron, Jonah B Gelbach, and Douglas L Miller. Robust inference with multiway clustering. *Journal of Business & Economic Statistics*, 29(2):238–249, 2011.

Harold D Chiang and Yuya Sasaki. On using the two-way cluster-robust standard errors. *arXiv preprint arXiv:2301.13775*, 2023.

Teunis Kloek. Ols estimation in a model where a microvariable is explained by aggregates and contemporaneous disturbances are equicorrelated. *Econometrica: Journal of the Econometric Society*, pages 205–207, 1981.

Konrad Menzel. Bootstrap with cluster-dependence in two or more dimensions. *Econometrica*, 89(5):2143–2188, 2021.

Brent R Moulton. An illustration of a pitfall in estimating the effects of aggregate variables on micro units. *The review of Economics and Statistics*, pages 334–338, 1990.

Andrew J Scott and D Holt. The effect of two-stage sampling on ordinary least squares methods. *Journal of the American statistical Association*, 77(380):848–854, 1982.