

Abstract

JEL Classification: H75, M14

Keywords:

1 Heterogeneity Analysis

Preliminary and Half-baked

Why are we getting near-zero std errors? Let's consider an extreme case where we cluster standard error at the pair level. I'll show that in this case, std errors can't be estimated at all. Clustering at classroom level makes it better, that's why we get non-zero but near-zero std errors.

When we cluster at the pair level,

$$\hat{var}(\hat{\beta}) = (X'X)^{-1}\hat{\Omega}(X'X)^{-1} \quad (1)$$

$$\hat{\Omega} = \sum_{g=1}^G s_g s_g' \quad (2)$$

$$s_g = X_g' e_g, \quad (3)$$

where X_g' is 4 by $|X|$ matrix of regressors for that pair, e_g is 4 by 1 vector of error terms for that pair.

Consider a case where individual 1 is mover and high CCEI in both pre and post. Then $X_g^{(m_i H_{it})} = (1, 0, 1, 0) = X_g^{(m_i)}$. When individual 1 is mover and high CCEI in only pre case, then $X_g^{(m_i H_{it} post_{it})} = (0, 0, 0, 0)$, which is 1- constant vector. Other cases can be walked through in a similar way. To sum up, in whichever case, we always have a result that one column of X is a perfect linear combination of the other. This is because mover dummy is summed to 1, and Higher CCEI dummy is summed to 1 within each pair. In addition, note that $e_1 + e_2 = 0$ since $y_1 + y_2 = 1$.

Getting rid of post dummy helps Post dummy does not make sense at all because $y_1 + y_2 = 1$ both in pre and post. If individual 1's post value of y goes up, it means that it's exactly compensated by the other in the pair.

Getting rid of post dummy or alternatively constant helps, but the interpretation of coefficients gets odd. Simulation shows that then coefficient captures combinations of DGP parameters, and interpretation gets really hard.

My suggestion would be, actually, keep only one of the randomly chosen individual from each pair and run regressions only with those.