EPE - Lecture 6
Estimating Precision In Group Designs With Autocorrelated Errors

Sylvain Chabé-Ferret

Toulouse School of Economics, Inra

January 2016

In a nutshell

In this lecture, we are going to study how to estimate the precision of an estimator when the treatment has been allocated to groups of observations whose outcomes are correlated.

Basic Motivation: Failure of i.i.d.

Up to now, we have maintained the i.i.d. assumption. We are going to relax it now. Failure of i.i.d. matters for the sampling noise of our estimator if and only if:

- Treatment is allocated at the level of groups of observations
- Outcomes are correlated within groups

Failure of i.i.d.: Examples

- Cross section: treatment randomized at village level, and individual income and consumption correlated at village level (price shocks)
- Panel: treatment varies at the individual level but we observe several observations for each individual before and after the treatment date and observations are correlated over time (fixed effects of permanent shocks)
- Both panel and cross-section: treatment varies at the state/group of years level and we have individual observations with several periods before and after the treatment and outcomes are correlated across individuals within state and across time within individuals

Basic Intuition: Failure of i.i.d. and Sampling Noise

- With i.i.d. observations, averaging decreases sampling noise because shocks on one side of the mean are compensated by equivalent shocks on the other side
- The larger the number of shocks, the more the shocks compensate each other
- With e.g. group shocks, only shocks to other groups can start compensating the shock to one group
- The number of groups does not grow as fast as the number of observations: we thus have more noise than if the observations were i.i.d.
- Obviously, the more variance is accounted for by the group shocks, the more severe the problem

Outline

Example: RCT with Clustered Design

- Randomization at cluster level
- Cluster level shocks make observations autocorrelated
- Key parameter: Intra-Cluster Correlation (ICC)

$$ICC = \frac{\mathbb{V}[Y_i]_{\text{inter}}}{\mathbb{V}[Y_i]}$$

Intra-cluster correlation

- Let's assume that unobserved shocks are the sum of an idiosyncratic shock and of a common village shock

$$U_{ic} = \nu_c + \epsilon_{ic}$$

- Let's also assume that all these shocks are i.i.d with finite variances σ_ν^2 and σ_ϵ^2 and are independent of each other
- As a consequence $\sigma_U^2 = \sigma_\nu^2 + \sigma_\epsilon^2$
- We define intra cluster correlation as follows:

$$ICC = \frac{\sigma_\nu^2}{\sigma_\nu^2 + \sigma_\epsilon^2}$$

  - It is the fraction of the total variance that is due to common village-level shocks
  - It measures how correlated two observations from the same village are

Failure of i.i.d. and Sampling Noise: Illustration

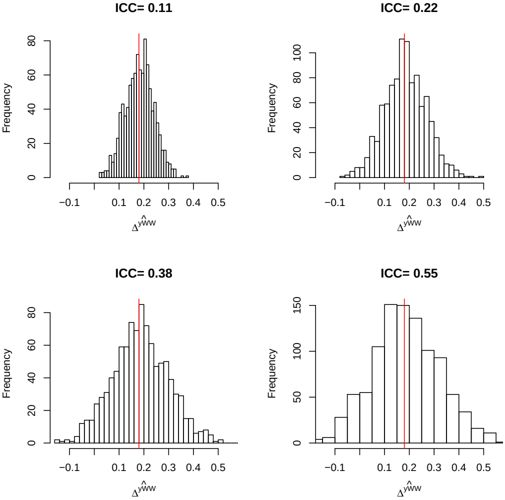

Figure: Distribution of the *WW* estimator in a Brute Force Design for various levels of ICC

Failure of i.i.d. and Sampling Noise: Illustration

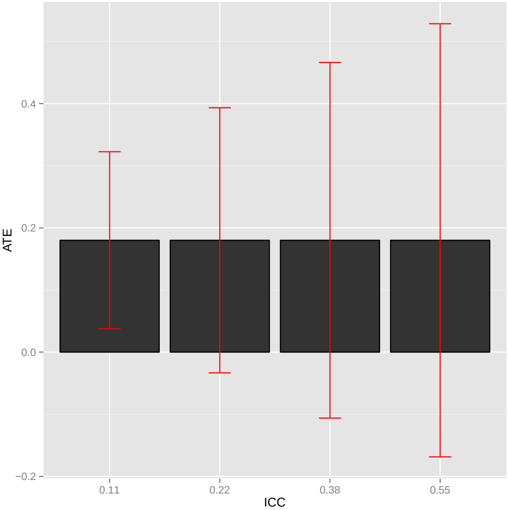

Figure: Effect of ICC on the 99% sampling noise of the WW estimator in a Brute Force design

Block diagonal covariance matrix

- Example of three clusters with three individuals

$$\mathbb{E}[UU'] = \sigma_U^2 \begin{pmatrix} 1 & \rho & \rho & & & & & & \\ \rho & 1 & \rho & & 0 & & & 0 & \\ \rho & \rho & 1 & & & & & & \\ & & & 1 & \rho & \rho & & & \\ & 0 & & \rho & 1 & \rho & & 0 & \\ & & & \rho & \rho & 1 & & & \\ & & & & & & 1 & \rho & \rho \\ & 0 & & & 0 & & \rho & 1 & \rho \\ & & & & & & \rho & \rho & 1 \end{pmatrix}$$

Variance of the OLS estimator with cluster effects

- It is very useful to start from the base case of fixed regressors
- I am also assuming away heteroskedasticity for the moment (results does not change much)
- I also assume that each village/cluster has the same number of observations and that there are J clusters
- As a consequence, sample size is $N = n_c J$

$$\begin{aligned}
\mathbb{V}[\hat{\Theta}_{OLS} - \Theta|X] &= \mathbb{V}[(X'X)^{-1}X'U|X] \\
&= (X'X)^{-1}X'\mathbb{E}[UU']X(X'X)^{-1} \\
&= (1 + (n_c - 1)\rho)\sigma_U^2(X'X)^{-1}
\end{aligned}$$

Clustered standard errors and the design effect

- The standard error of the OLS estimate of the average treatment effect when the treatment is randomized at the cluster level is approximately equal to:

$$\text{s.e.}_c(\hat{\beta}_{OLS}) \approx \sqrt{(1+(n_c-1)\rho)}\frac{1}{\sqrt{N}}\sqrt{\frac{\sigma^2_{U^1}}{p}+\frac{\sigma^2_{U^0}}{1-p}}$$

- We thus have

$$\text{s.e.}_c(\hat{\beta}_{OLS}) = \sqrt{(1+(n_c-1)\rho)}\text{s.e.}_r(\hat{\beta}_{OLS})$$

- $\sqrt{(1+(n_c-1)\rho)}$ is called the *design effect*
  - It is always larger than one
  - We always decrease the precision if we adopt a clustered design
  - If we do not take it into account when estimating standard errors, we underestimate standard errors

Clustered standard errors in practice

- Use clustered standard errors a la White/Arellano

$$\hat{V}_c(\hat{\beta}_{OLS}) = (X'X)^{-1}(\sum_c X_c'\hat{\Phi}_c X_c)(X'X)^{-1}$$

with $\hat{\Phi}_c = a\hat{U}_c\hat{U}_c'$, and a a correction for degrees of freedom

- Use FGLS
- Use block bootstrap
- Use randomization inference
  - Draw placebo treatments at the cluster level
  - Compute OLS estimate of treatment effect
  - It gives sampling noise

Estimation of Standard Errors: Illustration

- True 99% sampling noise is 0.4214
- Standard OLS 99% sampling noise is 0.2886
- Cluster robust OLS 99% sampling noise is 0.2904

Outline

RCT with Clustered Design

Panel Data with Temporal Persistence

Panel with Temporal and Cross-Sectional Persistence

Outline

RCT with Clustered Design

Panel Data with Temporal Persistence

Panel with Temporal and Cross-Sectional Persistence