

ECON220C Discussion Section 4

Dynamic Panel Data & Staggered DID

Lapo Bini

UC San Diego

April 28, 2026

DYNAMIC PANEL WITH PERSISTENT RESIDUALS

► Framework

Consider a panel with $N = 500$ and $t = 0, \dots, 4$. The data generating process includes unobserved heterogeneity and **persistence** in the **dependent variable**:

$$y_{it} = \alpha y_{i,t-1} + \mu_i + \varepsilon_{it}, \quad t = 1, \dots, 4,$$

where ε_{it} is an idiosyncratic error.

► Questions

- (i) A colleague claims that the **within estimator** removes the fixed effects bias and yields a consistent estimate of α as $N \rightarrow \infty$. Is this claim correct?
- (ii) Propose an **IV strategy** based on moment conditions in differences. Describe how the transformation deals with μ_i , what moment conditions are required for identification, and how the approach depends on assumptions about the error structure.
- (iii) Suppose y_{it} follows a **highly persistent** process ($\alpha \approx 1$). Discuss the potential weaknesses of your proposed estimator under this condition.

ARELLANO BOND ESTIMATOR

► Framework

Consider a dynamic panel model with **MA(1) errors**:

$$y_{it} = y_{it-1}\beta + \mu_i + u_{it}, \quad u_{it} = e_{it} + \theta e_{it-1},$$

where, conditional on $\{y_{i0}\}_{i=1}^n$, $e_{it} \stackrel{iid}{\sim} (0, \sigma_e^2)$ across i and t ; μ_i is independent of $\{e_{it}\}$ with $\text{Var}(\mu_i) = \sigma_\mu^2 > 0$; $\theta \neq 0$, $\sigma_e^2 > 0$. Observations are $\{y_{i,0}, \dots, y_{iT}\}_{i=1}^n$. Consider $n \rightarrow \infty$.

► Questions

- (i) In the absence of weak instrument problems, is y_{it-2} a **valid instrument** for the first-differenced equation $\Delta y_{it} = \beta \Delta y_{it-1} + \Delta u_{it}$? Explain.
- (ii) Suppose $T = 3$. Is the panel ARMA(1,1) model **under, exactly, or over-identified**?
- (iii) Suppose $T = 4$. Design an **Arellano–Bond GMM estimator** of β that is as asymptotically efficient as possible. Be specific about the **form** and **number** of moment conditions, and write down the **GMM objective function**.

STAGGERED DID

Consider the **staggered DID** methodology for panel data in the potential outcomes framework. There are two periods $t = 1, 2$ and three potential treatment trajectories $\{d_1 = (1, 1); d_2 = (0, 1); d_3 = (0, 0)\}$ as listed in the table below:

d_i	t = 1	t=2
group 1	1	1
group 2	0	1
group 3	0	0

	t = 1	t=2
group 1	$\bar{Y}_{d_1,1} = 0.3$	$\bar{Y}_{d_1,2} = 1.2$
group 2	$\bar{Y}_{d_2,1} = 0$	$\bar{Y}_{d_2,2} = 0.6$
group 3	$\bar{Y}_{d_3,1} = 0.3$	$\bar{Y}_{d_3,2} = 0.3$

There are three potential outcomes $\{Y_{i1}(d_i), Y_{i2}(d_i)\}$ given the three different treatment status d_i . The sample is iid and each group has size $n_1 = n_2 = n_3 = 1/3$. The table on the right reports the average value of the outcome of interest for group d and time t.

EXERCISE - Q1

Let $ATT(d_2, 2) = E[Y_{i2}(d_2) - Y_{i2}(d_3) | Di = d_2]$ be the average treatment effect on treated at time 2 for the group of individuals who chose d_2 as the treatment trajectory. Explain how you would identify $ATT(d_2, 2)$.

EXERCISE - Q2

Explain how you would estimate $ATT(d_2, 2)$. Please first provide a formula for your proposed estimator and then give a numerical answer based on the information provided.

EXERCISE - Q3

Suppose we estimate the following model $Y_{it}(D_i) = \alpha_i + \lambda_t + D_{it}\beta + e_i$ by the TWFE estimator, which requires a two-way demeaned version of the dummy variable of the treatment. Find $\tilde{D}_{it}$ and the TWFE estimate $\hat{\beta}_{TWFE}$. Please provide a numerical answer.