AME

Week 4: Dynamic Linear Panel Models

Sophie Bindslev, October 2022

UNIVERSITY OF COPENHAGEN

Today's Plan

- Today's Problem
- Dynamic Linear Panel Data Models
- Arellano Bond
- Sargan Test
- Robust Standard Errors (coding tips)
- Your time to shine!

Do firm profits exhibit state dependence?

- Consider a model of firm profits:

$$\pi_{it} = c_i + \rho\pi_{it-1} + \boldsymbol{x_{it}}\boldsymbol{\beta} + u_{it} \tag{1}$$

- For policy intervention a relevant question is *whether differences in firm performance (as measured by profits) are due to firm fixed effects or due to state dependence*
- Why do we care whether correlations in profits over time are due to c_i or $\rho \neq 0$?
- Today we will look to answer the question whether *conditional* on c_i last period profits help predict current profits i.e. $\rho \neq 0$

Dynamic Panel Data Model

- Including the lagged dependent variable, y_{it-1}, as a regressor

$$y_{it} = c_i + \rho y_{it-1} + \boldsymbol{x_{it}}\boldsymbol{\beta} + u_{it} \tag{2}$$

- Can we use FE/FD or POLS to estimate ρ ?
- FE/FD require strict exogeneity whilst POLS requires no confounding time-invariant heterogeneity
- Both fail as $y_{it-1} = c_i + \rho y_{it-2} + \boldsymbol{x_{it-1}}\boldsymbol{\beta} + u_{it-1}$ means that

$$E(y_{it-1}u_{it-1}) = \sigma_u^2 \neq 0 \tag{3}$$

$$\begin{aligned} E(y_{it-1}c_i) &= E((c_i + \rho y_{it-2} + \boldsymbol{x_{it-1}}\boldsymbol{\beta} + u_{it-1})c_i) \\ &= E(c_i^2 + \rho c_i^2 + ... + \rho^J c_i^2) \neq 0 \end{aligned} \tag{4}$$

- FE, FD and POLS on (2) will be **inconsistent**, what to do?

FD-IV AR(1) Model

- Let's lose $\mathbf{x}_{it}$ for a moment and consider the FD AR(1) Model

$$\Delta y_{it} = \rho \Delta y_{it-1} + \Delta u_{it} \tag{5}$$

- We've just seen that $E(y_{it-1}u_{it-1}) = \sigma_u^2 \neq 0$ which implies that $E(\Delta y_{it-1}\Delta u_{it-1}) \neq 0$. FD removes c_i but still isn't consistent
- **Solution**: Use instrumental variable (IV) to instrument Δy_{it-1}
- Requirements for z_{it} to be a valid IV for x_{it} :

$$\textit{Exogeneity}: \quad E(\mathbf{z}_{it}u_{it}) = 0 \tag{6}$$

$$\textit{Relevance}: \quad E(\mathbf{z}_{it}\mathbf{x}_{it}) \neq 0 \tag{7}$$

- In this context $\mathbf{x}_{it} = y_{it-1}$ is the endogenous variable we'd like to instrument. What are some contender IVs, $\mathbf{z}_{it}$? What exogeneity assumption can we make that might help find IVs in this setting?

Sequential Exogeneity

- If we assume that the idiosyncratic error is uncorrelated with all past realisations of the dependent variable:

$$\textit{Sequential Exogeneity}: \quad E(y_{is}u_{it}) = 0 \quad \forall \quad s < t \tag{8}$$

then we can use $\mathbf{z}_{it} = (y_{it-2}, y_{it-3}, \ldots, y_{i0})$ as IVs for y_{it-1}

- NB the instrument matrix $\mathbf{z}_{it}$ becomes larger the later the period we consider due to more lags becoming available. $\mathbf{z}_{it}$ is telescoping

Two Stage Least Squares (2SLS)

- Given the considerations above, a consistent estimator of ρ is provided by the 2SLS estimator
- The resulting estimator is the 2SLS estimator of ρ:

$$\rho_{2SLS} = (\hat{X}'\hat{X})^{-1}\hat{X}'y \tag{9}$$

where $\hat{X} = Z'(Z'Z)^{-1}Z'X$. In this case $X = y_{it-1}$,$y = y_{it}$ and $Z = (y_{it-2}, y_{it-3}, \ldots, y_{i0})$.

- Steps: 1) obtain IVs $\mathbf{z}_{it}$ for y_{it-1}, 2) first stage: get predicted values $\hat{y}_{it-1}$, 3) 2nd stage regress y_{it} on $\hat{y}_{it-1}$
- How do we handle re-introducing x_{it}? We can instrument them in a similar fashion to y_{it-1} (i.e. using lags, depending on what we assume about $E(x_{it}u_{is})$)

Arellano Bond

- Under sequential exogeneity 2SLS is consistent but inefficient
- To address this inefficiency, Arellano and Bond proposed the GMM estimator which uses an optimal weighting matrix, W:

$$\rho_{GMM} = (\hat{X}'ZWZ'\hat{X})^{-1}\hat{X}'ZWZ'y \tag{10}$$

- W puts greater weight on observations with more variation to increase precision of the estimation
- Optimal weighting matrix would be $W = S^{-1} = [E(Z'uu'Z)]^{-1}$ but in reality we start off with $W = (Z'Z)^{-1}$ (i.e. 2SLS), obtain residuals $\hat{u}$, and use these to update $\hat{W} = Z'\hat{u}\hat{u}'Z$
- AB's GMM estimator is 2SLS with a feasible weighting matrix, $\hat{W}$, which is updated at each step

Arellano Bond, a few notes of caution

- Doesn't work if $\rho = 1$ since then first differences are simply white noise and cannot be predicted using past levels
- **Weak instruments** and **instrument proliferation**: risk having too many poor IVs essentially no longer purging the regressors off their endogeneity
- Also relies on no serial correlation in the error terms (bonus: can you see why?)
- Even with those caveats this estimator is **hugely** popular - has 37,319 citations (v unusual for a method paper) - and is widely used for both growth and labour market applications

Sargan Test: Overidentifying Restrictions

- Due to the telescoping nature of the instrument matrix we have more IVs than strictly needed for identification
- This allows us to test these "over-identifying restrictions" under the null that all of our IVs are valid
- That is, we can test whether $E(\mathbf{Z}u) = 0$ holds
- Sargan Test Statistic

$$\mathbf{J} := \hat{\mathbf{u}}'\mathbf{Z}\mathbf{W}\mathbf{Z}'\hat{\mathbf{u}} \sim \chi^2_M$$

  where $M = r - K$ is the number of overidentifying restrictions, r is the number of IVs and K is the number of regressors
- Note! This is not a test of a subset of IVs, the null hypothesis should be rejected if any one of the IVs is correlated with the errors

Robust Standard Errors

- Robust standard errors for OLS:

$$\widehat{Avar}(\hat{\beta}_{OLS}) = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Omega\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} \tag{11}$$

  where $\Omega = diag(\hat{u}_i)^2$

- Panel robust standard errors for FE:

$$\widehat{Avar}(\hat{\beta}_{FE}) = (\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1}\left(\sum_{i}^{N}\ddot{\mathbf{X}}_i'\hat{\ddot{\mathbf{u}}}_i\hat{\ddot{\mathbf{u}}}_i'\ddot{\mathbf{X}}_i\right)(\ddot{\mathbf{X}}'\ddot{\mathbf{X}})^{-1} \tag{12}$$

- **Looping** over cross-sectional units, $i \in N$, is useful when computing these. You can loop using `for i in range(N):`
- Make sure you are **pulling out the correct portion** of the dataset for each i i.e. the first T elements for $i = 0$ if the data, y, is stacked into an $NT \times 1$ vector, `slice` function is useful for this

Your time to shine!

- Have a look at the toolbox LinearDynamic_ante.py and fill in missing pieces
- Solve the problem set and use functions from the toolboxes LinearDynamic_ante.py and gmm_ante.py where necessary