

# Panel Data Analysis

## Part III – Modern Moment Estimation

James J. Heckman  
University of Chicago

Econ 312, Spring 2023

## Review Moments and Identification:

- $Y = X\beta + U$
- $E^*(U|X) = 0 \Rightarrow \text{Cov}(Y - X\beta, X) = 0$
- $\Rightarrow \hat{\beta} = (X'X)^{-1}X'Y$
- Key idea: orthogonality condition  $\Rightarrow$  moment condition

## Review Moments and Identification: IV Conditions

- $$Y_{(T \times 1)} = X_{(T \times K)} \beta_{(K \times 1)} + U_{T \times 1}$$
- $E^*(U|X) \neq 0$
- $E^*(U|Z) = 0$
- $Z$  is  $M \times K$   
( $M \geq K$ )
- $E^*(X|Z)$  non-degenerate
- $\therefore \text{Cov}(Z'X)$  rank =  $K$
- $Z'(Y - X\beta) = 0$ : These are the moments in GMM.
- $Z'Y = (Z'X)\beta$  if  $M = K$
- $\hat{\beta} = (Z'X)^{-1}Z'Y$  otherwise GMM

## Panel Data Model

- Suppose  $y_{it} = \beta X_{it} + \eta_i + v_{it} \quad i = 1, \dots, I$
- $U_{it} = \eta_i + v_{it} \quad t = 1, \dots, T$
- $X_{it}$  is strictly exogenous if

$$E^*(U_{it} | X_i^T) = 0 \quad \forall t$$
$$X_i^T = (X_{i1}, \dots, X_{iT})$$

- $\therefore$  OLS identifies  $\beta$  under rank conditions and  $E^*(\eta_i | X_i^T) = 0$
- $E^*$  is linear projection.

## Fixed Effects

- $X_{it}$  is **strictly exogenous** given  $\eta_i$  if

$$E^*(v_{it} \mid X_i^T, \eta_i) = 0 \quad t = 1, \dots, T$$

for all  $X_i^T$ .

- First difference eliminates fixed effects:

$$E^*(v_{it} - v_{i,t-1} \mid X_i^T) = 0.$$

- Multivariate regression with cross equation restrictions.
- Assume  $X_i^T$  that this is essentially all the information.

## Partial Adjustment Model With Strictly Exogenous Variable

$$y_{it} = \alpha y_{i,(t-1)} + \beta_0 X_{i,t} + \beta_1 X_{i,t-1} + \eta_i + v_{it}.$$

- Assume  $E^*(v_{it} | X_i^T) = 0$ ,  $t = 2, \dots, T$ .
- Does not restrict serial correlation in  $v_{it}$ , just its dependence on  $X_i^T$ .
- Model identified if  $T \geq 3$  and rank condition is established.
- $E^*(\Delta v_{it} | X_i^T) = 0$ .

- Model identified for  $T \geq 3$ .
- $T = 3$  case; acquire orthogonality restrictions

$$E(X_{is}(\Delta y_{i3} - \alpha \Delta y_{i2} - \beta_0 \Delta X_{i3} - \beta_1 \Delta X_{i2})) = 0$$

$$\Leftrightarrow E(X_{is}(\Delta v_{i3})) = 0, s = 1, 2, 3$$

- Use these in GMM to identify model.
- 3 equations in 3 unknowns and we acquire exact identification.
- **Note:** Strict exogeneity enables us to identify dynamic effect of  $X$  on  $y$  with arbitrary serial correlation in the errors;
- **Price:** Assumes  $X$  not influenced by past values of  $y$  and  $v$ .



- **Definition:**  $X$  is **predetermined** if
- $E^*(v_{it} \mid X_i^t, y_i^{t-1}) = 0, \quad t = 2, \dots, T$  (A)
- $X_{it} = (X_{it}, \dots, X_{it}), y_{i,t-1} = (y_{i1}, \dots, y_{i,t-1})$ .
- Current shocks are uncorrelated with past values of  $y$  and current and past values of  $X$ .
- Feedback from lagged dependent variables to future  $X$  not ruled out.
- *E.g.*, Euler equations. (Information set of agents uncorrelated with current and future idiosyncratic shocks but not past shocks).

## Example: Euler Equation:

$$E \left[ \frac{U_c^t(c_t)}{U_c^{t-1}(c_{t-1})} \beta R_t \mid \mathcal{I}_{t-1} \right] = 1$$

- $\beta$  = subjective discount rate
- $R_t = 1 + r_t$ ; interest rate discount factor

## Special Case: Power Utility:

$$U_t = \frac{(c_t)^{1-\gamma} - 1}{1-\gamma}; U_{c,t} = (c_t)^{-\gamma}$$
$$E \left[ \left( \frac{c_t}{c_{t-1}} \right)^{-\gamma} \beta R_t \mid \mathcal{I}_{t-1} \right] = 1$$

- $Z_{t-1}$  is in the information set.
- Crucial that instruments don't include variables that cause the innovation.
- $E_{t-1} \left[ Z_{t-1} \left( \beta \left( \frac{C_t}{C_{t-1}} \right)^{-\gamma} R_t - 1 \right) \right] = 0$
- $\beta \frac{C_t^{-\gamma}}{C_{t-1}^{-\gamma}} R_t - 1 = \varepsilon_t$
- $\varepsilon_t = \left[ \beta \frac{C_t^{-\gamma}}{(C_{t-1})^{-\gamma}} R_{t-1} \right] - \left[ E_{t-1} \frac{C_t^{-\gamma}}{C_{t-1}^{-\gamma}} R_{t+1} - 1 \right]$
- $\varepsilon_t$  is forecast error.
- $E(\varepsilon_t Z_{t-1}) = 0$ .
- $Z_t$  has to be relevant in forecasting future returns or consumption growth.
- Need at least 2 instruments for  $(\beta, \gamma)$  parameters.

## Implication of Predeterminedness:

- $E^*(v_{i,t} - v_{i,t-1} \mid X_i^{t-1}, y_i^{t-2}) = 0, t = 3, \dots, T$
- For  $T = 3$ , we acquire

$$0 = E \left[ \begin{pmatrix} y_{i1} \\ X_{i1} \\ X_{i2} \end{pmatrix} (\Delta y_{i3} - \alpha \Delta y_{i2} - \beta_0 \Delta X_{i3} - \beta_1 \Delta X_{i2}) \right]$$

- This condition is not the same as that in strictly exogenous models:
- We acquire 3 moments only 2 in common with last ones given.
- Standard errors are consistent with arbitrary serial correlation.

- Observe that in the predetermined case we can have special cases of serial correlation.
- e.g. for  $T = 4$

$$E(\Delta v_{i,t} \Delta v_{i,t-j}) = 0 \quad j > 2$$

- Valid for first order MA.
- Valid orthogonality conditions derived from:
- $\Delta y_{i,3} - \alpha \Delta y_{i,2} - \beta_0 \Delta X_{i,3} - \beta_1 \Delta X_{i,2} = v_{i,3} - v_{i,2}$
- $\Delta y_{i,4} - \alpha \Delta y_{i,3} - \beta_0 \Delta X_{i,4} - \beta_1 \Delta X_{i,3} = v_{i,4} - v_{i,3}$

- Orthogonality conditions:

$$E(y_{i1} \Delta v_{i,4}) = 0$$

$$E(x_{i1} \Delta v_{i,4}) = 0$$

$$E(x_{i2} \Delta v_{i,4}) = 0.$$

- Other orthogonality conditions from:

$$y_{i,4} = \alpha y_{i,3} + \beta_0 X_{i,4} + \beta_1 X_{i,3} + \eta_i + v_{i,4}$$

$$y_{i,3} = \alpha y_{i,2} + \beta_0 X_{i,2} + \beta_1 X_{i,1} + \eta_i + v_{i,3}$$

$$\begin{aligned} \Delta y_{i,4} &= \alpha(y_{i,3} - y_{i,2}) + \beta_0(X_{i,4} - X_{i,3}) \\ &\quad + \beta_1(X_{i,3} - X_{i,2}) + (v_{i,4} - v_{i,3}) \end{aligned}$$



- Many, many orthogonality conditions.

### Suppose Uncorrelated Fixed Effects

- Some  $X_{it}$  uncorrelated with  $\eta_i$

$$E[X_i^T (y_{i2} - \alpha y_{i1} - \beta_0 X_{i2} - \beta_1 X_{i1})] = 0$$

- $T$  orthogonality conditions for each regressor.
- Predetermined variables could be uncorrelated with fixed effects

$$X_{it} = \rho X_{i,(t-1)} + \gamma v_{i,(t-1)} + \varphi \eta_i + \varepsilon_{i,t}$$

if  $\phi = 0$ ,  $X$  would be uncorrelated with  $\eta$ .

- Adds more orthogonality restrictions:

$$E(X_{i1}(y_{i2} - \alpha y_{i1} - \beta_0 X_{i2} - \beta_1 X_{i1})) = 0$$

$$E(X_{it}(y_{it} - \alpha y_{i,t-1} - \beta_0 X_{it} - \beta_1 X_{i,t-1})) = 0, t = 2, \dots, T.$$

- Only identified when  $T \geq 3$ .

## AR-1 Models Balestra - Nerlove Problem

- $y_{it} = \alpha y_{i,t-1} + \eta_i + v_{i,t}$
- $i = 1, \dots, I; t = 2, \dots, T$
- (A-1)  $E^*(v_{it} | y_i^{t-1}) = 0 \quad t = 2, \dots, T$   
 $E(\eta_i) = \gamma, E(v_{it}^2) = \sigma_t^2$   
 $Var(\eta_i) = \sigma_\eta^2$
- $\eta_i$  and  $v_{it}$  freely correlated
- $E(v_{it}^2 | y_i^{t-1})$  need not coincide with  $\sigma_t^2$ .

- We get  $(T - 1)(T - 2)/2$  moment restrictions:

$$E(y_i^{t-2}(\Delta y_{it} - \alpha \Delta y_{i,t-1})) = 0$$

- Use minimum distance methods: GMM

- Define  $\omega_{ts} = \alpha\omega_{(t-1),s} + c_s$ ,  $t = 2, \dots, T$ ,  $s = 1, \dots, t-1$   $c_s = E(Y_{is}\eta_{it})$
- For  $s < t$ , we obtain

$$y_{i,t}y_{i,s} = \alpha y_{i,t-1}y_{i,s} + \eta_i y_{i,s} + y_{i,s}v_{i,t}$$

$$E(y_{it}y_{is}) = \alpha E(y_{i,t-1}y_{i,s}) + E(\eta_i y_{i,s}) + E(y_{i,s}v_{i,t})$$

$= 0$

$$E(y_{it}y_{is}) = \omega_{ts}$$

$$E(y_{i,t-1}y_{is}) = \omega_{t-1,s}$$

$$E(y_{is}\eta_{i,t}) = c_s$$

- We take  $T \times \left(\frac{T+1}{2}\right)$  distinct elements of

$$\Omega = E(y_i y_i').$$

- For  $T = 3$ , we obtain  $\omega_{31} = \alpha\omega_{21} + c_1$

- $$\omega_{21} = \alpha\omega_{11} + c_1$$

$$\alpha = \frac{\omega_{31} - \omega_{21}}{\omega_{21} - \omega_{11}} = \frac{\alpha(\omega_{21} - \omega_{11})}{(\omega_{21} - \omega_{11})}$$

$$c_1 = \omega_{31} - \alpha\omega_{21}$$

$$c_2 = \omega_{32} - \alpha\omega_{22}$$

- $\therefore$  model just identified.
- Fit discrepancies between the population moments and fitted moments.

$$y_{it} = \alpha y_{i,t-1} + \beta X_{i,t} + \eta_i + U_{it}$$

$$U_{it} = \rho U_{it-1} + \varepsilon_{it}$$

$$\varepsilon_{i,t} \perp\!\!\!\perp \varepsilon_{i,t'} \quad \forall t, t'$$

$$\varepsilon_{it} \perp\!\!\!\perp X_{i,t'} \quad \forall t, t'$$

$$\eta_i \perp\!\!\!\perp X_{i,t'} \varepsilon_{it} \quad ? \text{ maybe}$$

$$y_t = \alpha y_{t-1} + \beta X_t + \eta_i + \rho U_{i,t-1} + \varepsilon_{i,t}$$

$$\begin{aligned} y_t &= \alpha y_{t-1} + \beta X_t + \eta_i + \rho(y_{t-1} - \alpha y_{t-2} - \beta X_{t-1} - \eta_i) + \varepsilon_{it} \\ &= (\alpha + \rho)y_{t-1} + \beta X_t - \rho\beta X_{t-1} - \rho\alpha y_{t-2} + (1 - \rho)\eta_i + \varepsilon_{i,t} \end{aligned}$$

- What parameters are identified?
- If we work with  $\Delta y_{it}$ : eliminates fixed effect.

- I. Suppose  $\rho = 1$ : (errors are random walks)

$$y_t = (\alpha + 1)y_{t-1} - \alpha y_{t-2} + \beta(X_t - X_{t-1}) + \varepsilon_{it}$$

- II. Suppose  $\alpha = 1$

$$y_t = (1 + \rho)y_{t-1} - \rho y_{t-2} + \beta X_t - \rho\beta X_{t-1} + (1 - \rho)\eta_i + \varepsilon_{it}$$

- In I.,  $\eta_i$  vanishes
- In II., it does not



# GMM Estimation

- Consider

$$y_{it} = X_{it}\beta_0 + U_{it}$$

$$U_{it} = \eta_i + v_{it} \quad \text{iid across } i$$

$$E^*(v_{it} \mid Z_i^t) = 0$$

$Z^S$  are the instruments:  $(P \times 1)$

$$y_i = X_i\beta_0 + U_i$$

$$y_i = (y_{i1}, \dots, y_{iT})' \quad X_i = (X'_{i1}, \dots, X'_{iT})$$

$$U_i = (U_{i1}, \dots, U_{iT})'$$

- We get I.V. for models in differences.
- Let  $K$  be an upper triangular  $(T - 1) \times T$  transformation matrix of rank  $T - 1$ .
- Such that  $K\iota = 0$   $\iota$  is  $T \times 1$  vector of 1's.
- $K$  is first difference operator or forward difference operator.

- Orthogonality restrictions:

$$E [Z_i' K U_i] = 0.$$

- $Z_i$  is block diagonal

$$\begin{bmatrix} Z^1 & & \dots & & 0 \\ & \ddots & & & \\ \vdots & & Z^t & & \vdots \\ & & & \ddots & \\ 0 & & \dots & & Z^T \end{bmatrix}.$$

- Optimal GMM:

$$\hat{\beta}_{GMM} = (M'_{ZX} A_N M_{ZX})^{-1} M'_{ZX} A_N M_{Zy}$$

$$M_{ZX} = \sum_{i=1}^N Z'_i K X_i$$

$$M_{Zy} = \sum_{i=1}^N Z'_i K y_i.$$

- $A_N$  is estimate of inverse of

$$E(Z_i'K U_i U_i' K' Z_i)$$

- Just an application of GMM.
- Robust version use  $\tilde{U}$  in place of  $U_i$  (as in Eicker-White)

$$\tilde{U}_i = y_i - X_i \tilde{\beta}.$$

- Optimal variance:

$$\text{Var}(\hat{\beta})_R = \left[ E(X_i' K' Z_i) [E(Z_i' K U_i U_i' K' Z_i)]^{-1} E(Z_i' K X_i) \right]$$

- Can be shown to be invariant to  $K$ .

## Orthogonal Deviations (Forward Differencing)

- $U_{1t}^* = c_t \left[ U_{it} - \frac{1}{T-t} (U_{i,t+1} + \dots + U_{i,T}) \right]$
- $c_t^2 = \frac{T-t}{T-t+1}$
- (a) Preserves the orthogonality of transformed errors, gets rid of fixed effect.

$$K_0 = \text{diag} \left[ \left( \frac{(T-1)}{T}, \dots, \frac{1}{2} \right) \right]^{1/2} .$$



- $$M^+ = \begin{bmatrix} 1 & -(T-1)^{-1} & -(T-1)^{-1} & -(T-1)^{-1} \\ 0 & 1 & -(T-2)^{-1} & -(T-2)^{-1} \\ & 1 & -1/2 & -1/2 \\ & & 1 & -1 \end{bmatrix}$$
- $K_0 K_0' = I_{T-1}$
- $K_0' K_0 = I_T - \frac{ii'}{T} = F$ ; within operator.

## Consider the Following Model:

- $y_{it} = \alpha y_{i,t-1} + \eta_i + v_{i,t}$
- $\iff y_{it} = \eta_i^* + w_{i,t}$

- $w_{i,t} = \alpha w_{i,t-1} + v_{i,t}$
- $y_{i,t} - \eta_i^* = w_{i,t}$
- $(y_{i,t} - \eta_i^*) = \alpha(y_{i,t-1} - \eta_i^*) + v_{i,t}$
- $y_{i,t} = \alpha y_{i,t-1} + (1 - \alpha)\eta_i^* + v_{i,t}$ .

- Consider  $\alpha = 1$
- I.  $y_{i,t} = \eta_i^* + w_{i,t}$
- $w_{i,t} = w_{i,t-1} + v_{it}$

## Random Initial Condition Model?

- Or II.  $y_{i,t} = y_{i,t-1} + \eta_i + v_{i,t}$

## Random Walk with Heterogenous Drift.

- Model II has autocorrelation  $> 1$

$$\rho = \frac{\text{Cov}(y_{i,t}, y_{i,t-1})}{\text{Var}(y_{i,t})} = \alpha + \frac{\text{Cov}(\eta_i, y_{i,t-1})}{\text{Var}(y_{i,t})} > 1$$

when  $\alpha > 1$ .

$$\Delta y_{i,t} = \eta_i + v_{i,t} \quad \text{Correl}(\Delta y_{i,t}, \Delta y_{i,t-1}) > 0.$$

- $\text{Correl}(\Delta y_{i,t}, \Delta y_{i,t-1}) = \text{Correl}(\Delta \omega_{i,t}, \Delta \omega_{i,t-1})$   
 $= (v_{i,t}, v_{i,t-1})$ .

## Link to Other Restrictions

## Other Restrictions

- Lack of correlation between effects and errors:

$$E^*(v_{it} \mid y_i^{t-1}, \eta_i) = 0, \quad t = 2, \dots, T$$
$$0 = E[(y_{it} - \alpha y_{i,t-1}) [\Delta y_{i,t-1} - \alpha \Delta y_{i,t-2}]]$$

quadratic (in  $\alpha$ ) restrictions: because

$$E(\eta_i \Delta v_{i,t-1}) = 0.$$



- When  $\eta_i \perp\!\!\!\perp v_{it}$  (Ahn-Schmidt).
- Multiply

$$y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it}$$

by  $\eta_i$

$$\eta_i y_{it} = \alpha \eta_i y_{i,t-1} + \eta_i^2 + \eta_i v_{it} \quad c_t = \alpha c_{t-1} + \sigma_\eta^2.$$

- For  $T = 3$ , imposes no further restrictions.

$$\sigma_\eta^2 = (\omega_{32} - \omega_{21}) - \alpha(\omega_{22} - \omega_{11}).$$



- Change in  $y_{it}$  uncorrelated with fixed effect

$$E^*(y_{i,t} - y_{i,t-1} \mid \eta_i) = 0 \quad t = 2, \dots, T$$

adds (to A-1) the following moment conditions:

- $E[(y_{i,t} - \alpha y_{i,t-1})(\Delta y_{i,t-1})] = 0, t = 3, \dots, T$
- $\alpha$  will satisfy

$$\alpha = (\omega_{22} - \omega_{21})^{-1}(\omega_{32} - \omega_{31})$$

- Full stationarity:

$$\omega_{11} = \frac{\sigma_{\eta}^2}{(1 - \alpha)^2} + \frac{\sigma^2}{1 - \alpha^2}.$$

## Unit Roots Case

- $y_{it} = \alpha y_{i,t-1} + \eta_i + v_{it}$
- For  $|\alpha| < 1$ , we can write

$$y_{it} = \eta_i^* + \omega_{it}$$

$$\omega_{it} = \alpha \omega_{i,t-1} + v_{it}$$

- $\eta_i^* = \frac{\eta_i}{1 - \alpha}$ .

- Substitute:

$$y_{it} - \eta_i^* = \alpha(y_{i,t-1} - \eta_i^*) + v_{i,t}$$
$$y_{i,t} = \alpha(y_{i,t-1} + \eta_i) + v_{i,t}.$$

- Now when  $\alpha = 1$ , we get a distinction:

## Model I:

$$y_{it} = \eta_i^{**} + \omega_{it}$$
$$\omega_{it} = \omega_{i,t-1} + v_{it}.$$

- A model with an initial random intercept.

## Model II:

$$y_{it} = y_{i,t-1} + \eta_i + v_{it}$$

(heterogenous linear growth).

## Empirical Features of II:

- $\rho = \frac{\text{Cov}(y_{it}, y_{i,t-1})}{\text{Var}(y_{it})} = \alpha + \frac{\text{Cov}(\eta_i, y_{i,t-1})}{\text{Var}(y_{i,t})} > 1$

- And

$$\text{Cov}(\Delta y_t, \Delta y_{t-1}) > 0$$

- Observe that we fail the rank condition for application of (A-1) in Model I using:

$$E(y_{i,t-j}[\Delta y_{it} - \alpha \Delta y_{i,t-1}]) = 0 \quad j > 1$$

- Why?

$$\begin{aligned} & Cov(y_{i,t-j}, \Delta y_{i,t-1}) \\ &= Cov(y_{i,t-j}, v_{i,t-1}) = 0 \end{aligned}$$

- $\therefore$  we divide by zero in forming this function.



**Model II:** Passes the rank condition:

$$E(y_{i,t-j}(\Delta y_{it} - \alpha \Delta y_{i,t-1})) = \text{Cov}(y_{i,t-j}, v_{i,t-1} - v_{i,t-2})$$

for  $j = 2$  satisfy rank condition.

- Observe that with mean stationarity rank condition is satisfied:

$$E[(\Delta y_{i,t-1})(y_{i,t} - \alpha y_{i,t-1})] = 0$$

$$E[(\Delta y_{i,t-1})(y_{i,t-1})] = E[(\omega_{i,t-1} - \omega_{i,t-2})[\eta_i + \omega_{i,t-1}]] \neq 0.$$
$$(v_{i,t-1} - v_{i,t-2}) = \omega_{i,t-1} - \omega_{i,t-2}$$

$$\eta_i + \omega_{i,t-1} = y_{it} - \alpha y_{i,t-1}.$$

- $\therefore$  if maintained, can test null hypothesis of random walk without drift against mean stationarity.

## Using Stationarity Restrictions:

- Consider a model

$$y_{i,t} = \delta' w_{i,t} + \eta_i + v_{i,t}$$

if  $E^*(v_{i,t} | w_i^t) = 0$

$$w_i^t = (w_{i,t}, w_{i,t-1}, \dots, w_{i,1}).$$

- Moments are

$$E[w_i^{t-1}(\Delta y_{i,t} - \delta' \Delta w_{i,t})] = 0.$$

- If  $E^*(v_{it} \mid w_i^t, \eta_i) = 0$

- Add to this Ahn-Schmidt

$$E(y_{i,t} - \delta' w_{i,t})(\Delta y_{i,t-1} - \delta \Delta w_{i,t-1}) = 0$$

- If

$$E^*(\Delta w_{i,t} | \eta_i) = 0$$

- We get

$$E((\Delta w_{i,t})(y_{i,t} - \delta' w_{i,t})) = 0.$$

[Return to main text](#)