Part 1

Basics

Two-way Error Components Model #flashcard

Two-way Error Components Model is the most general panel data model formulation:

$$y_{it} = \underbrace{x_{it}\beta + w_i\gamma + s_t\delta}_{\text{no restriction, can drop terms}} + (\eta_i + f_t + v_{it})$$

where the error term contains:

- η_i is individual-specific, time-invaryving
- f_t is common across individual, time-invariant
- v_{it} is individual-specific, time-varying

Deal with f_t: One-way Error Components Model / Time Dummies #flashcard

We can use time dummies to estimate the common (same across individual) but time-varying component. The corresponding model is known as a **Standard/One-way Error Components Model**:

$$y_{it} = x_{it}\beta + w_i\gamma + \sum_{s=1}^{T} \phi_s \underbrace{\mathbb{1}\{t = s\}}_{D_t^s} + (\eta_i + v_{it})$$

where:

- $\phi_s = s_t\delta + f_t$ captures all common (same across individual) but time-varying component, including the time error term f_t
  There is no loss in generality in this approach, but be aware that we will not be able to separately identify parameters on specific common time-varying variables.

One-way Error Components Model #flashcard

In the one-way error components model, we suppress f_t since we can deal with easily with time dummies:

$$\underbrace{\begin{bmatrix} y_{i1} \\ \cdots \\ y_{iT} \end{bmatrix}}_{y_i(T\times 1)} = \underbrace{\begin{bmatrix} x_{1,i,1} & \cdots & x_{K,i,1} \\ \cdots & \cdots & \cdots \\ x_{1,i,T} & \cdots & x_{K,i,T} \end{bmatrix}}_{X_i(T\times K)} \underbrace{\begin{bmatrix} \beta_1 \\ \cdots \\ \beta_K \end{bmatrix}}_{\beta(K\times 1)} + \underbrace{\begin{bmatrix} w_{1,i} & \cdots & w_{G,i} \\ \cdots & \cdots & \cdots \\ w_{1,i} & \cdots & w_{G,i} \end{bmatrix}}_{W_i(T\times G)} \underbrace{\begin{bmatrix} \gamma_1 \\ \cdots \\ \gamma_G \end{bmatrix}}_{\gamma} + \left(\underbrace{\begin{bmatrix} \eta_i \\ \cdots \\ \eta_i \end{bmatrix}}_{\eta_i j_T(T\times 1)} + \underbrace{\begin{bmatrix} v_{i1} \\ \cdots \\ v_{iT} \end{bmatrix}}_{v_i(T\times 1)} \right)$$

where:

- $X_i(T \times K)$ contains time-varying and individual-specific variables
  - there are T rows due to the panel structure
- $W_i(T \times G)$ contains time-invariant but individual-specific variables
  - all T rows are the same
- $\eta_i j_T$ is the individual specific constant. It's not time-varying $\implies$ a vector of the same constant
- v_i contains the individual-specific and time-varying error terms
  Stack them together vertically:

$$\underbrace{y}_{NT\times 1} = \underbrace{X}_{NT\times K} \underbrace{\beta}_{K\times 1} + \underbrace{W}_{NT\times G} \underbrace{\gamma}_{G\times 1} + (\underbrace{\eta}_{NT\times 1} + \underbrace{v}_{NT\times 1})$$

  Typically, we are interested in β, so we can set the number of time-invariant but individual-specific variables to 0 ($G = 0$)

$\iff$ merge them into η, and obtain the following model:

$$\underbrace{y}_{NT\times 1} = \underbrace{X}_{NT\times K}\underbrace{\beta}_{K\times 1} + (\underbrace{\eta}_{NT\times 1} + \underbrace{v}_{NT\times 1})$$

Model with Predetermined Regressors / Random Effects

Pooled OLS

Pooled OLS (Predetermined Reg + RE Consistent) #flashcard
Estimator: the standard OLS estimator:

$$\hat{\beta}_{\text{pooled}} = \hat{\beta}_{OLS} = (X^T X)^{-1} X^T y$$

Assumptions:

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- Random Effect 1 - **Predetermined Regressors**:

$$\mathbb{E}\left[x_{it} v_{it}\right] = 0 \iff x_{it} \perp v_{it} \text{ for given } t$$

- Random Effect 2 - **Random Effects/Uncorrelated Individual Effects**:

$$\mathbb{E}\left[x_{it} \eta_i\right] = 0 \iff x_{it} \perp \eta_i \text{ for given } t$$

  - This is *very restrictive* because it requires the explanatory variables x_{it} to be uncorrelated with all individual-specific effects
  Under those 5 assumptions, we have the consistency result:

$$\text{plim}_{NT\to\infty} \hat{\beta}_{\text{pooled}} = \beta$$

- Note that this allows $u_{it} = \eta_i + v_{it}$ to be serially correlated (indeed will be correlated if $\eta_i \neq 0$)
  Avar for valid inference: Cluster-Roboust Standard Errors with individual as clusters

Part 2

Models with Strictly Exogenous Regressors / Fixed Effects

WGOLS

Within Group (WG) Estimator (Strictly Exogenous Regressors + FE Consistent) #flashcard
Estimator: OLS estimator on **within transformed variables**:

$$\underbrace{y_{it} - y_i}_{\tilde{y}_{it}} = \underbrace{(x_{it} - x_i)}_{\tilde{x}_{it}} + \underbrace{v_{it} - v_i}_{\tilde{v}_{it}}$$

note that η_i drops out after within transformation

$$\hat{\beta}_{WG} = \left(\tilde{X}^T \tilde{X}\right)^{-1} \tilde{X}^T \tilde{y}$$

Assumptions:

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \dots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- **Strict Exogeneity**:

$$\mathbb{E}\left[x_{it} v_{is}\right] = 0 \, \forall \, t, s \in \{1, 2, \dots, T\}$$

  - This is *also very restrictive* because it requires the current values of x_{it} to be uncorrelated with all past, present, and future values of the time-varying component of the error term v_{is} for the same individual
  - It rules out any **feedback** from the past shocks v_{is} onto later x_{it}
  - It rules out the presence of any lagged dependent variable $y_{i,t-1}$ in x_{it} since it's correlated with $v_{i,t-1}$
  - v_{is}, y_{is} cannot appear in the expression for x_{it}!
- Relaxed assumption: we allow correlation between regressors and the individual-specific but time-varying components (**Correlated Individual Effects / Fixed Effects**):

$$\mathbb{E}\left[x_{it} \eta_i\right] \neq 0$$

  Under those 4 assumptions, we have the consistency result:

$$\text{plim}_{N \to \infty} \hat{\beta}_{\text{W}} = \beta$$

  which is valid for small fixed T.
  Avar for valid inference: standard errors based on within transformed variables:

$$\hat{\beta}_{BG} \sim^a N\left(\beta, \hat{\sigma}_v^2 (\tilde{X}^T \tilde{X})^{-1}\right)$$

  where the consistent estimator $\hat{\sigma}_v^2$ can be obtained using:

$$\hat{\sigma}_v^2 = \frac{\hat{\tilde{v}}^T \hat{\tilde{v}}}{NT - N - K}$$

- We typically report the cluster-robust SE to account for serial correlations in v_{it}

LSDVOLS

Least Squares Dummy Variables (LSDV) Estimator (Strictly Exogenous Regressors + FE Consistent) #flashcard
This is an equivalent to the Within Group Estimator.
Estimator: OLS estimator of the original model plus individual dummies:

$$y_{it} = x_{it} \beta + \sum_{j=1}^{N} \eta_j \mathbb{1}\left\{i = j\right\} + v_{it}$$

and

$$\hat{\beta}_{LSDV} = \hat{\beta}_{WG}$$

Assumptions (same as WG):

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \dots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- **Strict Exogeneity**:

$$\mathbb{E}\left[x_{it}v_{is}\right] = 0 \; \forall \; t, s \in \{1, 2, \ldots, T\}$$

- Relaxed assumption: we allow correlation between regressors and the individual-specific but time-varying components (**Correlated Individual Effects / Fixed Effects**):

$$\mathbb{E}\left[x_{it}\eta_i\right] \neq 0$$

  Under those 4 assumptions, we have the consistency result:

$$\text{plim}_{N\to\infty}\hat{\beta}_{\text{W}} = \beta$$

  which is valid for small fixed T.
  Avar for valid inference: standard errors based on within transformed variables:

$$\hat{\beta}_{LSDV} \sim^a N\left(\beta, \hat{\sigma}_v^2(\tilde{X}^T\tilde{X})^{-1}\right)$$

  where the consistent estimator $\hat{\sigma}_v^2$ can be obtained using:

$$\hat{\sigma}_v^2 = \frac{\hat{\tilde{v}}^T\hat{\tilde{v}}}{NT - N - K}$$

FDOLS

First-Difference OLS (FDOLS) Estimator (Strict Exogeneity Consistent (weaker version / Orthogonality) + FE Consistent #flashcard
Estimator: OLS estimator after the **first-differenced transformation**:

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta v_{it}$$

Assumptions (same as WG):

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- **Strict Exogeneity (weaker version)**:

$$\mathbb{E}\left[\Delta x_{it}\Delta v_{it}\right] = 0$$

  - This is less restrictive than Strict Exogeneity Assumption ($\mathbb{E}\left[x_{it}v_{is}\right] = 0 \; \forall \; t, s \in \{1, 2, \ldots, T\}$), but still stronger than the Predetermined Regressors Assumption ($\mathbb{E}\left[x_{it}v_{it}\right] = 0$)
  - *It only rules feedback from 1 period lag ($v_{i,t-1}$ on x_{it}), but allows for longer-period feedback ($v_{i,t-k}$ on x_{it} for $k \geq 2$)*
- We allow correlation between regressors and the individual-specific but time-varying components (**Correlated Individual Effects / Fixed Effects**):

$$\mathbb{E}\left[x_{it}\eta_i\right] \neq 0$$

  Under those 4 assumptions, we have the consistency result:

$$\text{plim}_{N\to\infty}\hat{\beta}_{\text{FDOLS}} = \beta$$

  which is valid for small fixed T.
  Avar for valid inference:
- If v_{it} follows a Random Walk (not likely), standard Avar can be used
- If v_{it} is serially uncorrelated, then Δv_{it} will follow a MA(1) process, and we need to use Cluster-Robust SE

Misc

Efficiency Comparison between WGOLS/LSDV and FDOLS #flashcard

- If v_{it} is serially uncorrelated, then WGOLS will be more efficient than FDOLS.
- If all K regressors in x_{it} are correlated with η_i and v_{it} is serially uncorrelated as implied by the **classical assumption**:

$$v_{it}|x_{i1}, x_{i2}, \ldots, x_{iT} \sim^{iid} (0, \sigma_v^2) \,\forall\, i, t \implies \text{Strict Exogeneity} + \text{Serially Uncorrelated } v_{it}$$

  Then, *WGOLS will be efficient*.
  - If we have the additional **Normality** assumption:

$$v_{it}|x_{i1}, x_{i2}, \ldots, x_{iT} \sim^{iid} N(0, \sigma_v^2)$$

    Then $\hat{\beta}_{WG} = \hat{\beta}_{\text{Max Likelihood}}$
- If all K regressors in x_{it} are correlated with η_i and v_{it} follows a **Random Walk** process:

$$v_{it} = v_{i,t-1} + \epsilon_{it}, \epsilon_{it}|x_{i1}, x_{i2}, \ldots, x_{iT} \sim^{iid} (0, \sigma_\epsilon^2)$$

  Then, *FDOLS will be efficient* (FD model satisfies all Gauss-Markov Assumptions).
- **Special Cases**:
  - $T = 2 \implies$ WGOLS/FDOLS are exactly the same.
  - One variable in x_{it} is a step-function (switches from 0 to 1 once in the sample period) $\implies$ FDOLS essentially only uses the switching period to estimate the coefficient on the step variable while WGOLS uses the whole sample. $\implies$ WGOLS is more efficient, and FDOLS will be biased if the response of y_{it} to the change in the step variable is gradual ($>$ 1 period).
  - Note that the switching time has to be different if we also wanna to include time dummies.

WGOLS and FDOLS in Big T Panels #flashcard

As $T \to \infty$:

- *WGOLS is consistent even if we have Predetermined Coefficients but NOT Strict Exogeneity*
  - Intuition: as $T \to \infty$, the contribution of each observation to the within transformed variable $\to 0$. Therefore, in the limit $\mathbb{E}[x_{it}v_{it}] = 0 \implies \mathbb{E}[\tilde{x}_{it}\tilde{v}_{it}] = 0$
- FDOLS is still inconsistent without Strict Exogeneity
  - Intuition: the requirement $\mathbb{E}[\Delta x_{it}\Delta v_{it}] = 0$ does not change as $T \to \infty$

Models with Predetermined and Strict Exogenous Regressors

BGOLS

Between Group OLS (BGOLS) Estimator (Strictly Exogenous Regressors + Uncorrelated Individual Effects (RE) => Consistent) #flashcard

Uncorrelated Individual Effects $\iff$ *Predetermined Regressors*

Estimator: OLS estimator for the cross-section equation:

$$y_i = x_i\beta + (\eta_i + v_i)$$

where $y_i = \frac{1}{T}\sum_{t=1}^{T} y_{it}$

Assumptions:

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$

- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right]=\mathbb{E}\left[\eta_i\right]=\mathbb{E}\left[v_{it}\right]=0$$

- **Strict Exogeneity**:

$$\mathbb{E}\left[x_{it} v_{is}\right]=0 \ \forall \ t, s \in\{1,2, \ldots, T\}$$

- **Uncorrelated Individual Effects / Random Effects**:

$$\mathbb{E}\left[x_{it} \eta_i\right]=0 \Longleftrightarrow x_{it} \perp \eta_i \text{ for given } t$$

  Under those 5 assumptions, we have the consistency result:

$$\operatorname{plim}_{N \rightarrow \infty} \hat{\beta}_{\mathrm{B}}=\beta$$

  BGOLS is NOT efficient! It's main purpose is to get a consistent estimator for σ_η^2 (in order to carry out REGLS):

- Construct the residuals:

$$\bar{u}=(\widehat{\eta_i+v_i})=y_i-x_i \hat{\beta}_{BG}$$

- Estimator for $\hat{\sigma}^2_{\bar{u}}$:

$$\hat{\sigma}^2_{\bar{u}}=\frac{\bar{u}^T \bar{u}}{N-K}$$

- Backout $\widehat{\sigma}_\eta^2$ using the following relation:

$$\widehat{\sigma}^2_{\bar{u}}=\widehat{\sigma}_\eta^2+\frac{\widehat{\sigma}_v^2}{T} \Longrightarrow \widehat{\sigma}_\eta^2=\widehat{\sigma}^2_{\bar{u}}-\frac{\widehat{\sigma}_v^2}{T}$$

  where $\widehat{\sigma}^2_{\bar{u}}$ is obtained from BGOLS just now, and $\widehat{\sigma}_v^2$ is obtained from WGOLS.

REGLS

Random Effect GLS (REGLS) Estimator (2 Classical Assumptions => Consistent) #flashcard
Estimator: **Theta Transformation**:

- Obtain $\widehat{\sigma}^2_{\bar{u}}$ is obtained from BGOLS, $\widehat{\sigma}_v^2$ from WGOLS, and calculate $\widehat{\sigma}_\eta^2$.
- Define:

$$\theta^2:=\frac{\widehat{\sigma}_v^2}{\widehat{\sigma}_v^2+T \widehat{\sigma}_\eta^2}$$

- Perform **theta differencing**:

$$y_{it}^*=y_{it}-(1-\theta) y_i, x_{it}^*=x_{it}-(1-\theta) x_i, u_{it}^*=u_{it}-(1-\theta) u_i$$

- REGLS estimator is the OLS estimator of the transformed model:

$$y_{it}^*=x_{it}^* \beta+u_{it}^*$$

  Assumptions (the Strongest Set of Assumptions):

- Panel Common Assumption 1 - **Cross-sectional Independence**: observations $\left(y_i, X_i\right)$ are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right]=\mathbb{E}\left[\eta_i\right]=\mathbb{E}\left[v_{it}\right]=0$$

- **2 "Classical" Assumption**:

$$\begin{cases} \eta_i | x_{i1}, x_{i2}, \ldots, x_{iT} & \sim^{iid} (0, \sigma_\eta^2) \\ v_{it} | x_{i1}, x_{i2}, \ldots, x_{iT} & \sim^{iid} (0, \sigma_v^2) \end{cases}$$

  - Those are stronger than strict exogeneity and predetermined regressors.
- Under those 5 assumptions, REGLS is consistent and efficient.

Testing for Correlated Individual Effects

Tests for Correlated Individual Effects #flashcard

- Idea: compare estimates consistent with correlated individual effects (WG/FD) and those inconsistent with correlated individual effects (POLS/GLS)
- **Hausman Test**:
  - Assumptions: the same as REGLS
  - Null Hypothesis: Uncorrelated Individual Effects:

$$H_0 : \mathbb{E}\left[\eta_i | X_i\right] = 0$$

  - Test statistics and distribution under null
  - Option 1: **comparing WG and REGLS**:

$$\left(\hat{\beta}_{WG} - \hat{\beta}_{REGLS}\right)^T \left[\widehat{Var}(\hat{\beta}_{WG}) - \widehat{Var}(\hat{\beta}_{REGLS})\right]^{-1} \left(\hat{\beta}_{WG} - \hat{\beta}_{REGLS}\right) \sim^a \chi_K^2$$

  - If we only focus on one parameter β^k, this simplifies to

$$\frac{\sqrt{\hat{\beta}_{WG}^k - \hat{\beta}_{REGLS}^k}}{\sqrt{\widehat{Var}(\hat{\beta}_{WG}^k) - \widehat{Var}(\hat{\beta}_{REGLS}^k)}} \sim^a N(0,1)$$

  - Option 2: **comparing WG and BG**:

$$\left(\hat{\beta}_{WG} - \hat{\beta}_{BG}\right)^T \left[\widehat{Var}(\hat{\beta}_{WG}) + \widehat{Var}(\hat{\beta}_{BG})\right]^{-1} \left(\hat{\beta}_{WG} - \hat{\beta}_{BG}\right) \sim^a \chi_K^2$$

Time-Invariant Variables Estimation

Estimating Time-Invariant Variables #flashcard

Setup: we would like to estimate γ in:

$$y_{it} = x_{it}\beta + w_i\gamma + (\eta_i + v_{it})$$

- **Case 1: Uncorrelated Time-Invariant Regressors** $w_i \perp \eta_i$
  - *Assumptions*:
  - Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
  - Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$
  - Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

  - **Uncorrelated Time-Invariant Regressors** $w_i \perp \eta_i$
  - + **Assumptions** on the correlations between x_{it} and η_i, v_{it}
  - *Procedures*:
  - First, estimate β using a consistent estimator (depends on the assumptions on the correlations between x_{it} and η_i, v_{it})
  - Then, regress the unexplained component on w_i, adjusted for measurement error

- *Example*:
- Suppose we have strictly exogenous x_{it} and it's correlated with η_i (FE)
1. Estimate β using WGOLS
2. Construct a new dependent variable:

$$Y_{it} = y_{it} - x_{it}\hat{\beta}_{WG}$$

3. Choose one of the options:
- POLS: consistent, but with SE adjustment issues:

$$Y_{it} = w_i\gamma + (\eta_i + v_{it} + e_{it})$$

- BGOLS: consistent, but with SE adjustment issues:

$$Y_i = w_i\gamma + (\eta_i + v_i + e_i)$$

- GMM with augmented system of equations: consistent and will produce the correct SE: form a system of equations by attaching WG equations with either of the above (use the second/BG as an example):$$\underbrace{ \begin{bmatrix}
\tilde{y}{i1} \
\tilde{y}{i2} \
\vdots \
\tilde{y}{iT} \
\textcolor{ red }{ \bar{y}{i} }
\end{bmatrix} }{ y{i}^{+} } = \underbrace{ \begin{bmatrix}
\tilde{x}{i1} \
\tilde{x}{i2} \
\vdots \
\tilde{x}{iT} \
\textcolor{ red }{ \bar{x}{i} }
\end{bmatrix} }{ x{i}^{+} }\beta

- \underbrace{ \begin{bmatrix}
0 \
0 \
\vdots \
0 \
\textcolor{ red }{ w{i} }
\end{bmatrix} }{ w_{i}^{+} }\textcolor{ red }{ \gamma }
- \underbrace{ \begin{bmatrix}
\tilde{v}{i1} \
\tilde{v}{i2} \
\vdots \
\tilde{v}{iT} \
\textcolor{ red }{ \eta{i} + \bar{v}{i} }
\end{bmatrix} }{ u{i}^{+} } \text{ and } \underbrace{ \textcolor{ red }{ \begin{bmatrix}
\tilde{x}{i1} & 0 \
\tilde{x}{i2} & 0 \
\vdots & \vdots \
\tilde{x}{iT} & 0 \
0 & w{i}
\end{bmatrix} } }{ z{i}^{+} } \tag{Augmented System of Equations}

$$and use GMM estimator with the moment condition :$$

\mathbb{E}\left[(z{i}^{+})^{T}u_{i}^{+} \right] = 0$$

- **Case 2: Correlated Time-Invariant Regressors** $w_i \perp \eta_i$
  - *Assumptions*:
  - Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
  - Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \ldots N$
  - Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

  - **Correlated Time-Invariant Regressors** $w_i \perp \eta_i$
  - **Valid Instruments**: We will need time-invariant instruments z_i for w_i with the standard IV assumptions:
  - Exogeneity:

$$\mathbb{E}\left[z_i \eta_i\right] = \mathbb{E}\left[z_i v_{it}\right] = 0$$

  - Informative: $\theta \neq 0$ in the first-stage regression (take a scalar example):

$$w_i = x_{it}\delta + z_i\theta + r_i$$

  - *Procedures*:
  - Choose one from these:
  - Standard 2SLS for the following equations, but with SE adjustment issues:
  -

$$Y_{it} = (y_{it} - x_{it}\hat{\beta}_{WG}) = w_i\gamma + (\eta_i + v_{it} + e_{it})$$

  - or:

$$Y_i = (y_i - x\hat{\beta}_{WG}) = w_i\gamma + (\eta_i + v_i + e_i)$$

  - If x_{it} is strictly exogenous wrt both v_{it} and η_i (classical assumptions), we can also use z_i to instrument w_i
  :$$\underbrace{ \begin{bmatrix}
  \tilde{y}*{i1} *
  \tilde{y}{i2} \
  \vdots \
  \tilde{y}*{iT} *
  \textcolor{ red }{ \bar{y}{i} }
  \end{bmatrix} }*{ y*{i}^{+} } = \underbrace{ \begin{bmatrix}
  \tilde{x}*{i1} *
  \tilde{x}{i2} \
  \vdots \
  \tilde{x}*{iT} *
  \textcolor{ red }{ \bar{x}{i} }
  \end{bmatrix} }*{ x*{i}^{+} }\beta

- \underbrace{ \begin{bmatrix}
  0 \
  0 \
  \vdots \
  0 \
  \textcolor{ red }{ w*{i} }*
  \end{bmatrix} }{ w_{i}^{+} }\textcolor{ red }{ \gamma }
- \underbrace{ \begin{bmatrix}
  \tilde{v}*{i1} *
  \tilde{v}{i2} \
  \vdots \
  \tilde{v}*{iT} *
  \textcolor{ red }{ \eta{i} + \bar{v}*{i} }*

\end{bmatrix} }{ u{i}^{+} } \text{ and } \underbrace{ \textcolor{ red }{ \begin{bmatrix}
\tilde{x}{i1} & 0 \
\tilde{x}{i2} & 0 \
\vdots & \vdots \
\tilde{x}{iT} & 0 \
0 & z{i}
\end{bmatrix} } }{ z{i}^{+} } \tag{Augmented System of Equations}

$$and use GMM estimator with the moment condition:$$

\mathbb{E}\left[(z{i}^{+})^{T}u_{i}^{+} \right] = 0$$

Part 3

Models WITHOUT Strict Exogeneity NOR Random Effect Assumptions

Predetermined Regressors + Correlated Individual Effects (FDX-2SLS/LagX-2SLS/Arellano-Bond GMM)

FDX-2SLS/LagX-2SLS/Arellano-Bond GMM Estimator (Predetermined Regressors + Correlated Individual Effects Consistent) #flashcard

Procedure:

- **FD+LagX-2SLS**
  - Use $x_{i,t-1}$ or $\Delta x_{i,t-1}$ as IV for Δx_{it} in FD regression:

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta v_{it}$$

    - Both satisfy exogeneity condition ($\mathbb{E}\left[x_{i,t-1}v_{it}\right] = \mathbb{E}\left[x_{i,t-1}v_{i,t-1}\right] = 0, \mathbb{E}\left[x_{i,t-1}\Delta v_{it}\right] = 0$) under our assumptions, but the relevance condition rules out $x_{it} \sim RW$.
      - We can also use WG transformation when having $T \to \infty$
    - 1st-stage Projection:

$$\Delta x_{it} = x_{i,t-1}\theta + \epsilon_{it}$$

    - 2nd-stage Regression:

$$\Delta y_{it} = \widehat{\Delta x}_{it}\beta + \Delta v_{it}$$

- **Arellano-Bond GMM**:
  - For $T \geq 3$, we can use both $x_{i,t-1}$ and $\Delta x_{i,t-1}$
  - Base model is still the FD regression:

$$\Delta y_{it} = \Delta x_{it}\beta + \Delta v_{it}$$

  - Implement GMM using Sequential Moment Conditions:

$$\begin{aligned}
&\mathbb{E}\left[x_{i1}\Delta v_{i2}\right] = 0 \\
&\mathbb{E}\left[(x_{i2}, x_{i1})\Delta v_{i3}\right] = 0 \\
&\mathbb{E}\left[(x_{i3}, x_{i2}, x_{i1})\Delta v_{i4}\right] = 0 \\
&\dots
\end{aligned}$$

  or more concisely:

$$\mathbb{E}\left[Z_i^T \Delta v_i\right] = 0$$

where Z_i is defined in the first stage projection:

$$\underbrace{\begin{bmatrix} \Delta x_{i2} \\ \Delta x_{i3} \\ \Delta x_{i4} \\ \\ \end{bmatrix}}_{\Delta x_i} = \underbrace{\begin{bmatrix} x_{i1} & 0 & 0 & 0 & \dots \\ 0 & x_{i2} & x_{i1} & 0 & \dots \\ 0 & 0 & 0 & x_{i3} & \dots \\ \\ \end{bmatrix}}_{Z_i} \underbrace{\begin{bmatrix} \pi_1 \\ \pi_2 \\ \pi_3 \\ \\ \end{bmatrix}}_{\pi} + r_i$$

- Beware that: if T is not very small or N is not very large, we need to be careful about "too-many instruments" / overfitting bias.
- *Lagged dependent variable is a special case of this method.* e.g. we may use $y_{i,t-2}, x_{i,t-1}$ as IVs for $\Delta y_{i,t-1}, \Delta x_{it}$
Assumptions:
- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \dots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- **Predetermined Regressors**:

$$\mathbb{E}\left[x_{it} v_{is}\right] \begin{cases} = 0 & \text{for } s \geq t & \text{(uncorrelated with currentfuture time-varying error)} \\ \neq 0 & \text{for } s < t & \text{(correlated with past time-varying error)} \end{cases}$$

- **(Possibly) Correlated Individual Effects**: we allow correlation between regressors and the individual-specific but time-varying components

$$\mathbb{E}\left[x_{it} \eta_i\right] \neq 0$$

- Extra: **Serial Uncorrelated Time-Varying Error**: v_{it} is serially uncorrelated
- They satisfies non of assumptions for our previous estimators.
- Under those assumptions:

$$\lim_{N \to \infty} \hat{\beta}_{2SLS} = \beta$$

  - Note that $\hat{\beta}_{WG}$ is also consistent if we allow $T \to \infty$.
- *Avar*: we can use clustered SE, but since we know the serially correlation is precisely MA(1), we also have specialised tools.

Endogenous Regressors + Correlated Individual Effects

Models with Endogenous Regressors and Correlated Individual Effects #flashcard
Procedure:

- Exactly the same as FDX-2SLS/LagX-2SLS/Arellano-Bond GMM Estimator, but we need 2 period levels/differences as IVs (or other IVs satisfying exogeneity and the rank condition).
Assumptions:
- Panel Common Assumption 1 - **Cross-sectional Independence**: observations (y_i, X_i) are cross-sectionally independent
- Panel Common Assumption 2 - **Slope Parameter Homogeneity**: parameters β are common to all $1, \dots N$
- Panel Common Assumption 3 - **Zero Error Mean**: both error components have expected value of zero:

$$\mathbb{E}\left[u_{it}\right] = \mathbb{E}\left[\eta_i\right] = \mathbb{E}\left[v_{it}\right] = 0$$

- **Endogenous Regressors**:

$$\mathbb{E}\left[x_{it}v_{is}\right]\begin{cases}=0 & \text{for } s>t \quad \text{(uncorrelated with future time-varying error)}\\ \neq 0 & \text{for } s\leq t \quad \text{(correlated with currentpast time-varying error)}\end{cases}$$

- **(Possibly) Correlated Individual Effects**: we allow correlation between regressors and the individual-specific but time-varying components

$$\mathbb{E}\left[x_{it}\eta_i\right]\neq 0$$

- Extra: **Serial Uncorrelated Time-Varying Error**: v_{it} is serially uncorrelated
- Note that $\hat{\beta}_{WG}$ is now inconsistent even when $T\to\infty$ since the correlation comes from current period v_{it}, which cannot be averaged out in the limit.

Ensemble of Endogenous and Strictly Exogenous Regressors

Ensemble of Lagged Y, Endogenous, and Strictly Exogenous Regressors #flashcard

Setup:

$$y_{it}=\alpha y_{i,t-1}+\beta_1 x_{1it}+\beta_2 x_{2it}+(\eta_i+v_{it})$$

where:

- **Panel Common Assumption 1+2+3**
- **Serial Uncorrelated Time-Varying Error**: v_{it} is serially uncorrelated
- **Mixture of Endogenous / Strictly Exogenous Regressors**
  - x_{1it} is endogenous wrt v_{it}
  - x_{2it} is strictly exogenous wrt v_{it}
- **Correlated Individual Effects**: both x_{1it}, x_{2it} are correlated with η_i

  Consistent Estimator:
- Focus on FD equation:

$$\Delta y_{it}=\alpha\Delta y_{i,t-1}+\beta_1\Delta x_{1it}+\beta_2\Delta x_{2it}+\Delta v_{it}$$

- Use $(y_{i,t-2}, x_{1,t-2}, x_{2it})$ or $(\Delta y_{i,t-2}, \Delta x_{1,t-2}, \Delta x_{2it})$ as IVs for $(\Delta y_{i,t-1}, \Delta x_{1it}, \Delta x_{2it})$ and perform 2SLS
- Alternatively, use sequential moment conditions and BAGMM.

Ensemble of Endogenous, Strictly Exogenous Regressors, and Extra Valid IVs #flashcard

Setup:

$$y_{it}=\beta_1 x_{1it}+\beta_2 x_{2it}+(\eta_i+v_{it})$$

where:

- **Panel Common Assumption 1+2+3**
- **Mixture of Endogenous / Strictly Exogenous Regressors**
  - x_{1it} is endogenous wrt v_{it}
  - x_{2it} is strictly exogenous wrt v_{it}
- **Valid IV**: z_{it} satisfying strict exogeneity and rank condition
  - For large T panels, we can relax the strict exogeneity and only require x_{2it}, z_{it} to be predetermined.
- **Correlated Individual Effects**: both x_{1it}, x_{2it} are correlated with η_i
- We no longer require v_{it} to be serially uncorrelated!

  Consistent Estimator:
- Focus on with-in transformed equation:

$$\tilde{y}_{it}=\beta_1\tilde{x}_{1it}+\beta_2\tilde{x}_{2it}+\tilde{v}_{it}$$

- Use $(\tilde{x}_{2it}, \tilde{z}_{it})$ as IVs for $(\tilde{x}_{1it}, \tilde{x}_{2it})$ and perform 2SLS

Summary

- Two-way Error Component Model
  - Deal with Common Time Effects → Time Dummies
  - One-way Error Component Model
    - Estimate Effects of x_{it}
      - Predetermined X + Random Effects (Uncorrelated Individual Effects) → Pooled OLS
      - Strictly Exogenous X + Correlated Individual Effects → WGOLS, LSDVOLS
      - (Weaker) Strictly Exogenous X + Correlated Individual Effects → FDOLS
      - Strictly Exogenous X + Random Effects (Uncorrelated Individual Effects) → BGOLS, REGLS
      - Predetermined X + Correlated Individual Effects → FDX-2SLS/LagX-2SLS/ABGMM
    - Estimate Effects of w_i

All Estimators for Parameters on x_{it} #flashcard

Estimator	Assumption on x_{it}, v_{is}	Assumption on x_{it}, η_i	Additional Assumptions	Method	Remarks
Pooled OLS	$\mathbb{E}[x_{it}v_{it}] = 0$ (Predetermined Regressor)	$\mathbb{E}[x_{it}\eta_i] = 0$ (Random Effects/Uncorrelated Individual Effects)		Standard Pooled OLS	
WG/LSDVOLS	$\mathbb{E}[x_{it}v_{is}] = 0 \, \forall \, s, t$ (Strict Exogeneity)	$\mathbb{E}[x_{it}\eta_i] \neq 0$ (Fixed Effects/Correlated Individual Effects)		With-in transformation / individual dummies	Efficient if v_{it} is serially uncorrelated; consistently estimate $\hat{\sigma}_v^2$
FDOLS	$\mathbb{E}[\Delta x_{it} \Delta v_{it}] = 0$ (Weaker Strict Exogeneity)	$\mathbb{E}[x_{it}\eta_i] \neq 0$ (Fixed Effects/Correlated Individual Effects)		First difference	Efficient if $v_{it} \sim RW$
BGOLS	$v_{it} x_{i1}, \ldots, x_{iT} \sim^{iid} (0, \sigma_v^2)$ (Classical Assumption)	$\mathbb{E}[x_{it}\eta_i] = 0$ (Random Effects/Uncorrelated Individual Effects)		Individual average	Not Efficient; just to get $\hat{\sigma}_\eta^2$
REGLS	$v_{it} x_{i1}, \ldots, x_{iT} \sim^{iid} (0, \sigma_v^2)$ (Classical Assumption)	$\eta_i x_{i1}, \ldots, x_{iT} \sim^{iid} (0, \sigma_\eta^2)$ (Classical Assumption)		WGOLS + BGOLS → Theta-differencing	Most Efficient Under Assumptions
FDX-2SLS/LagX-2SLS/Arellano-Bond GMM	$\mathbb{E}[x_{it}v_{is}] \begin{cases} = 0 & \text{for } s \geq t \\ \neq 0 & \text{for } s < t \end{cases}$ (Predetermined Regressors)	$\mathbb{E}[x_{it}\eta_i] \neq 0$ (Fixed Effects/Correlated Individual Effect)	v_{it} is serially uncorrelated (Serial Uncorrelated Time-Varying Error)	Use previous FD/level as IV for Δx_{it} / Sequential Moment Conditions	WGOLS will aslo be consistent if $T \to \infty$
FDX-2SLS/LagX-2SLS/Arellano-Bond GMM (Endogenous Version)	$\mathbb{E}[x_{it}v_{is}] \begin{cases} = 0 & \text{for } s > t \\ \neq 0 & \text{for } s \leq t \end{cases}$ (Endegenous Regressors)	$\mathbb{E}[x_{it}\eta_i] \neq 0$ (Fixed Effects/Correlated Individual Effect)	v_{it} is serially uncorrelated (Serial Uncorrelated Time-Varying Error)	Use previous FD/level as IV for Δx_{it} / Sequential Moment Conditions	FDX-2SLS/LagX-2SLS are also consistent as $T \to \infty$, but ABGMM will suffer from too many IVs. WGOLS will be inconsistent even if $T \to \infty$
WG2SLS	$\mathbb{E}[x_{it}v_{is}] \begin{cases} = 0 & \text{for } s > t \\ \neq 0 & \text{for } s \leq t \end{cases}$ (Endegenous Regressors)	$\mathbb{E}[x_{it}\eta_i] \neq 0$ (Fixed Effects/Correlated Individual Effect)	v_{it} is serially uncorrelated (Serial Uncorrelated	Use previous level as IV for $\tilde{x}_{it}$ in the WG transformed equation	WG2SLS is consistent ONLY when $T \to \infty$

Estimator	Assumption on x_{it}, v_{is}	Assumption on x_{it}, η_i	Additional Assumptions	Method	Remarks
			Time-Varying Error)		