Panel Data Econometrics[1]
(Yonsei University)

Min Seong Kim
Ryerson University

[1]This handout is based on the lecture notes by Prof. Yixiao Sun (UCSD).

1 Introduction to Panel Data Model

Basic Model

$$y_{it} = X_{it}\beta + \varepsilon_{it},\ i = 1, ..., n,\ t = 1, ..., T$$

X_{it} includes

- variables that vary across individuals and time periods. Denote them as X_{it}.
- variables that are time invariant. ie. vary only across individuals. Denote them as X_i.
- variables that vary only over time, not across individuals. Denote them as X_t

Type of panel data

- Large n, small T (micro-panel data)
  - Rely on the large n asympototics in which $n \to \infty$ with fixed T.
- Large T, small n
  - Rely on the large T asymptotics in which $T \to \infty$ with fixed n.
- Large n, large T
  - Joint asymptotics: $n, T \to \infty$ at the same time. Sometimes control the relative rate of expansion (eg. $\sqrt{n}/T \to 0$)
  - Sequential asymptotics: $n \to \infty$ first, then $T \to \infty$. Or $T \to \infty$ first, then $n \to \infty$.
- Small n, small T
  - Not many things to do. Collect more data.

Benefits of using panel data

- Efficiency
  - For $y_{it} = X_{it}\beta + \varepsilon_{it}$,

$$s.e.(\hat{\beta}) = \frac{\sigma}{\sqrt{\sum_{i=1}^{n}\sum_{t=1}^{T}(X_{it} - \mu)^2}}$$

- Dynamics
  - 25% of women are unemployed.
    - Senario 1: This 25% women are determined.
    - Senario 2: This 25% women are randomly selected.

    With panel we can tell between these senario.

- Identification (Example: Agricultural production function)

$$y_{it} = \beta_1 K_{it} + \beta_2 L_{it} + \gamma Z_i + u_{it}$$

y_{it} : farm output, K_{it} : Capital, L_{it} : Labor, Z_i : Soil quality, u_{it} : random shock (rainfall, ..)

  - This is the production function from a farmer's point of view.
  - Econometricians cannot observe Z_i which is correlated with K_{it}, L_{it}. Let

$$X_{it} = K_{it} L_{it},\ \beta = \beta_1 \beta_2.$$

and

$$y_{\iota t} = X_{it}' \beta + \varepsilon_{it},\ \text{with } \varepsilon_{it} = \gamma Z_i + u_{it}.$$

Then,

$$\hat{\beta} = \beta + \underbrace{\left(\sum_{i=1}^{n} \sum_{t=1}^{T} X_{it} X_{it}' \right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} X_{it} Z_i \gamma}_{\nrightarrow 0} + \left(\sum_{i=1}^{n} \sum_{t=1}^{T} X_{it} X_{it}' \right)^{-1} \sum_{i=1}^{n} \sum_{t=1}^{T} X_{it} u_{it}$$

$\Longrightarrow$ Omitted variable bias.

  - But in panel data, this bias can be removed.

$$\Delta y_{it} = \beta_1 \Delta K_{it} + \beta_2 \Delta L_{it} + \Delta u_{it}.$$

2 Static Panel Data Model

Model

$$y_{it} = X_{it} \beta + \varepsilon_{it},\ i = 1, ..., n,\ t = 1, ..., T.$$

In static models, X_{it} does not include lagged dependent variables.

Fixed Effects vs. Random Effects

$$y_{it} = X_{it}' \beta + \varepsilon_{it}.$$

Let $\varepsilon_{it} = \alpha_i + u_{it}$. In our example, $\alpha_i = \gamma Z_i$.

- α_i : individual effect or individual heterogeneity

  u_{it} : idiosyncratic error.

When $E(\alpha_i | X_i) = 0$, we say the model is a random effect model. Otherwise, it is a fixed effects model.

Random Effects Approach

Assumptions

1. $E(\alpha_i|X_i) = 0$, $E(u_{it}|X_i, \alpha_i) = 0$
2. $\alpha_i \sim iid$, $u_{it} \sim iid$. u_{it} is independent of α_j.
3. $\text{Rank}\left(E\left[X_i'\Omega^{-1}X_i\right]\right) = k$ with $E(\varepsilon_i\varepsilon_i') = \Omega$.
4. $E\left(\alpha_i^2|X_i\right) = \sigma_\alpha^2$, $E(u_iu_i'|X_i) = \sigma_u^2 I_T$

Under these assumptions,

$$
\begin{aligned}
E\varepsilon_{it}\varepsilon_{js} &= E(\alpha_i + u_{it})(\alpha_j + u_{js}) \\
&= \begin{cases} \sigma_\alpha^2 + \sigma_u^2 & \text{if } i = j \text{ and } t = s, \\ \sigma_\alpha^2 & \text{if } i = j \text{ and } t \neq s, \\ 0 & \text{if } i \neq j \text{ and } t \neq s. \end{cases}
\end{aligned}
$$

Thus,

$$\Omega = var(\varepsilon_i) = \sigma_u^2 I_T + \sigma_\alpha^2 J_T$$

where J_T is the $T \times T$ matrix with unity in every element.

Let's use matrix notation

$$y = X\beta + \varepsilon$$

where

$$\varepsilon = \begin{bmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \text{ and } \varepsilon_i = \begin{bmatrix} \varepsilon_{i1} \\ \vdots \\ \varepsilon_{iT} \end{bmatrix}.$$

Then, we have

$$
\begin{aligned}
var(\varepsilon|X) &= I_n \otimes \Omega \\
&= V
\end{aligned}
$$

GLS estimator

$$
\begin{aligned}
y &= X\beta + \varepsilon, \quad \varepsilon|X \sim (0, V) \\
V^{-1/2}y &= V^{-1/2}X\beta + V^{-1/2}\varepsilon.
\end{aligned}
$$

Therefore, the GLS estimator for random effects model is

$$
\begin{aligned}
\hat{\beta}_{REGLS} &= \left(X'V^{-1}X\right)^{-1} X'V^{-1}y \\
&= \left(X'\left(I_n \otimes \Omega^{-1}\right)X\right)^{-1} X'\left(I_n \otimes \Omega^{-1}\right)y \\
&= \left(\sum_{i=1}^n X_i'\Omega^{-1}X_i\right)^{-1} \sum_{i=1}^n X_i'\Omega^{-1}y_i.
\end{aligned}
$$

Asymptotic Inference

- Under Assumptions 1,2 and 3, $\hat{\beta}_{REGLS}$ is consistent.
- Under Asymptotics 1,2,3 and 4, $\hat{\beta}_{REGLS}$ is the BLUE.

Feasible GLS

We need to estimate Ω, or σ_u^2 and σ_α^2.

$$\hat{\beta}_{OLS} = \left(\sum_{i=1}^{n}\sum_{t=1}^{T} X_{it}X_{it}'\right)^{-1}\sum_{i=1}^{n}\sum_{t=1}^{T} X_{it}y_{it}$$

is consistent under the assumtpions above. So, we can estimate σ_ε^2 ($\varepsilon_{it} = \alpha_i + u_{it}$) with

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{nT-k}\sum_{i=1}^{n}\sum_{t=1}^{T}\left(y_{it} - X_{it}'\hat{\beta}_{OLS}\right)^2$$

As $\sigma_\alpha^2 = E\left(\varepsilon_{it}\varepsilon_{is}\right)$ for $t \neq s$,

$$\hat{\sigma}_\alpha^2 = \frac{1}{nT\left(T-1\right)/2-k}\sum_{i=1}^{n}\sum_{t=1}^{T-1}\sum_{s=t+1}^{T}\left(y_{it} - X_{it}'\hat{\beta}_{OLS}\right)\left(y_{is} - X_{is}'\hat{\beta}_{OLS}\right)$$

Plugging $\hat{\sigma}_\varepsilon^2$ and $\hat{\sigma}_\alpha^2$ in Ω, we obtain $\hat{\Omega}$ and

$$\hat{\beta}_{REFGLS} = \left(\sum_{i=1}^{n} X_i'\hat{\Omega}^{-1}X_i\right)^{-1}\sum_{i=1}^{n} X_i'\hat{\Omega}^{-1}y_i.$$

Understanding the GLS estimator

Recall $V = var\left(\varepsilon\right) = I_n \otimes \Omega$. As

$$\Omega = \sigma_\alpha^2 J_T + \sigma_u^2 I_T,$$

$$V = \left(T\sigma_\alpha^2 + \sigma_u^2\right)P + \sigma_u^2 Q,$$

where $P = I_n \otimes J_T/T$ and $Q = I_{nT} - P$. Let $\sigma_1^2 = T\sigma_\alpha^2 + \sigma_u^2$. Then,

$$V^{-1} = \sigma_1^{-2}P + \sigma_u^{-2}Q$$

and

$$V^{-1/2} = \sigma_u^{-1}\left(I_{nT} - \left(1 - \frac{\sigma_u}{\sigma_1}\right)P\right).$$

If we premultiply the regression model

$$y = X\beta + \varepsilon$$

by

$$\sigma_u V^{-1/2} = \left(\sigma_u/\sigma_1\right)P + Q,$$

then we have

$$y_{it} - \theta \bar{y}_i = \left(X_{it} - \theta \bar{X}_i\right)' \beta + (\varepsilon_{it} - \theta \bar{\varepsilon}_i)$$

where

$$\theta = 1 - \frac{\sigma_u}{\sigma_1}.$$

The error term $(\varepsilon_{it} - \theta \bar{\varepsilon}_i)$ are uncorrelated and have the same variance across all i and t. Thus, OLS based on the above regression model (quasi-demeaned equation) is the BLUE.

Fixed Effects Approach

$$y_{it} = X_{it}'\beta + \alpha_i + u_{it}$$

Now, we assume that X_{it} and α_i are correlated. Then, the random effects estimator is biased.

$$\hat{\beta}_{REGLS} = \beta + \left(\sum_{i=1}^{n} X_i'\Omega^{-1} X_i\right)^{-1} \sum_{i=1}^{n} X_i'\Omega^{-1} (\alpha_i + u_{it})$$

eg. y_{it} = farm output, X_{it} = capital, labor, α_i = soil quality; y_{it} = birth weight, X_{it} = smoking habit, α_i = other health related habits

Assumptions

1. $E(u_{it}|X_i, \alpha_i) = 0$
2. Rank$[E(X'QX)] = k$.
3. $E[u_i u_i|X_i, \alpha_i] = \sigma_u^2 I_T$

Estimation strategy

The idea is to transform the equation to eliminate the unobserved effect α_i. Here we consider fixed effects transformation, which is also called within transformation.

1. Average equation $y_{it} = X_{it}'\beta + \alpha_i + u_{it}$ over t to get $\bar{y}_i = \bar{X}_i'\beta + \alpha_i + \bar{u}_i$ with $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$.
2. Take difference.

$$y_{it} - \bar{y}_i = \left(X_{it} - \bar{X}_i\right)' \beta + u_{it} - \bar{u}_i \Leftrightarrow Qy = QX\beta + Qu$$

Note that under Assumption 1,

$$E(u_{it} - \bar{u}_i)\left(X_{it} - \bar{X}_i\right) = 0$$

for all t and s. Thus, OLS estimator from the transformed model is consistent and unbiased.

$$\begin{aligned}
\hat{\beta}_{FE} &= \left(\sum_{i=1}^{n}\sum_{t=1}^{T}\left(X_{it} - \bar{X}_i\right)\left(X_{it} - \bar{X}_i\right)'\right)^{-1} \sum_{i=1}^{n}\sum_{t=1}^{T}\left(X_{it} - \bar{X}_i\right)(y_{it} - \bar{y}_i) \\
&= \left(X'QX\right)^{-1} X'Qy
\end{aligned}$$

Asymptotic inference

$$\sqrt{n}(\hat{\beta}_{FE}-\beta) \quad \to^{d} \quad N\left(0,\sigma_u^2\left(EX'QX\right)^{-1}\right)$$

As

$$\begin{aligned} E\hat{u}'\hat{u} &= E\left[u'Qu-u'QX\left(X'QX\right)^{-1}X'Q'u\right] \\ &= \left(N\left(T-1\right)-k\right)\sigma_u^2. \end{aligned}$$

This implies an unbiased estimator for σ_u^2 is

$$\hat{\sigma}_u^2=\frac{SSR}{N\left(T-1\right)-k}.$$

Robust variance matrix estimator

Assumption 3 may be very strong. The fixed effects estimator is consistent and asymptotically normal under assumptions 1 and 2. In the presence of possible heteroskedasticity and autocorrelation,

$$var\left(\hat{\beta}_{FE}\right)=E\left[\left(X'QX\right)^{-1}X'Quu'Q'X\left(X'QX\right)^{-1}\right].$$

But,

$$\begin{aligned} X'Quu'Q'X &= \sum_{i=1}^{n}\sum_{j=1}^{n}\left(X_i-\bar{X}_i\right)'u_iu_j'\left(X_j-\bar{X}_j\right) \\ X'QX &= \sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\left(X_j-\bar{X}_j\right) \end{aligned}$$

Under the cross-sectional independence assumption $\left(E\left(u_iu_j'|X\right)=0 \text{ for } i\neq j\right)$

$$\begin{aligned} var\left(\hat{\beta}_{FE}\right) &= E\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\left(X_j-\bar{X}_j\right)\right)^{-1}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'u_iu_i'\left(X_i-\bar{X}_i\right)\right) \\ &\quad\times\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\left(X_j-\bar{X}_j\right)\right)^{-1}. \end{aligned}$$

Based on this, we can estimate the $var\left(\hat{\beta}_{FE}\right)$ with

$$\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\left(X_j-\bar{X}_j\right)\right)^{-1}\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\hat{u}_i\hat{u}_i'\left(X_i-\bar{X}_i\right)\right)\left(\sum_{i=1}^{n}\left(X_i-\bar{X}_i\right)'\left(X_j-\bar{X}_j\right)\right)^{-1}.$$

Random Effects Estimator vs. Fixed Effects Estimator

Comparison between two estimators

Note that

$$\hat{\beta}_{REGLS} = \left(X'V^{-1}X\right)^{-1} XV^{-1}y$$

where

$$V^{-1} = \sigma_1^{-2}P + \sigma_u^{-2}Q \text{ and } \sigma_1^2 = T\sigma_\alpha^2 + \sigma_u^2.$$

- If $\sigma_\alpha^2 = 0$, then $\sigma_1^2 = \sigma_u^2$ so $\hat{\beta}_{REGLS} = \hat{\beta}_{POLS}$.
- If $T \to \infty$, then $\hat{\beta}_{REGLS} \to \hat{\beta}_{FE}$.
- If $\sigma_\alpha^2 \to \infty$, then $\hat{\beta}_{REGLS} \to \hat{\beta}_{FE}$. The larger σ_α^2, the closer $\hat{\beta}_{REGLS}$ is to $\hat{\beta}_{FE}$.
- $var\left(\hat{\beta}_{REGLS}\right) = \left(\sigma_1^{-2}X'PX + \sigma_u^{-2}X'QX\right)^{-1}$ and $var\left(\hat{\beta}_{FE}\right) = \left(\sigma_u^{-2}X'QX\right)^{-1}$. Hence, $var\left(\hat{\beta}_{REGLS}\right) \leq var\left(\hat{\beta}_{FE}\right)$.

Hausman-Wu test

$$y_{it} = X_{it}'\beta + \alpha_i + u_{it}$$

Empirical Question: Which estimator should we use?

- $cov(\alpha_i, X_{it}) = 0$ for all t vs. $cov(\alpha_i, X_{it}) \neq 0$ for some t

Assumptions

1. $E\left(\alpha_i\right) = 0$, $E\left(u_{it}|\alpha_i, X_i\right) = 0$.
2. $E\left(\alpha_i^2|X_i\right) = \sigma_\alpha^2$, $E\left(u_i u_i'|X_i\right) = \sigma_u^2 I_T$
3. $\alpha_i \sim iid$ over i, $u_{it} \sim iid$ over i and t.

$$H_0\text{: } cov\left(\alpha_i, X_{it}\right) = 0 \text{ for all } t. \text{ vs. } H_1\text{: } cov\left(a_i, X_{it}\right) \neq 0 \text{ for some } t$$

	$\hat{\beta}_{REGLS}$	$\hat{\beta}_{FE}$	$\hat{\beta}_{REGLS} - \hat{\beta}_{FE}$
H_0	efficient	inefficient	small $\to 0$
	consistent	consistent	
H_1	inconsistent	consistent	Not small $\nrightarrow 0$

We consider a test statistic

$$q = \left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right)' \hat{W}^{-1} \left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right),$$

where

$$W = var\left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right)$$

If H_0 is true,

$$\sqrt{n}\left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right) \to^d N\left(0, Q_0\right) \text{ with } Q_0 = \lim_{n\to\infty} var\left(\sqrt{n}\left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right)\right)$$

and thus

$$q \to^d \chi^2_k \text{ with } k = \dim(\beta).$$

We can show

$$cov\left(\hat{\beta}_{FE}, \hat{\beta}_{REGLS}\right) = var\left(\hat{\beta}_{REGLS}\right).$$

Therefore, under H_0

$$var\left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right) = var\left(\hat{\beta}_{FE}\right) - var\left(\hat{\beta}_{REGLS}\right).$$

Based on this, we construct the test statistic

$$q = \left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right)' \left(\hat{\sigma}_u^2 \left(X'QX\right)^{-1} - \left(X'\hat{V}X\right)^{-1}\right)^{-1} \left(\hat{\beta}_{FE} - \hat{\beta}_{REGLS}\right)$$

3 Dynamic Panel Data Model

Models with sequentially exogenous variables

Consider

$$y_{it} = X_{it}'\beta + \alpha_i + u_{it}$$

Assumptions

1. $E\alpha_i = Eu_{it} = 0$
2. $E\left(u_{it}|X_{it}, \ldots, X_{i1}, \alpha_i\right) = 0$: sequential exogeneity ($cov\left(X_{is}, u_{it}\right) \neq 0$ for $s > t$).
   $\not\Leftrightarrow E\left(u_{it}|X_i, \alpha_i\right) = 0$: strict exogeneity

Fixed effect estimator

$$y_{it} - \bar{y}_i = \left(X_{it} - \bar{X}_i\right)\beta + \left(u_{it} - \bar{u}_i\right)$$

$$\hat{\beta}_{FE} \to^p \beta + \left(\frac{1}{T}\sum_{t=1}^{T} E\left(X_{it} - \bar{X}_i\right)\left(X_{it} - \bar{X}_i\right)'\right)^{-1} \underbrace{\frac{1}{T}\sum_{t=1}^{T} E\left(X_{it} - \bar{X}_i\right)u_{it}}_{(*)}$$

$$(*) = \frac{1}{T}\sum_{t=1}^{T} E\left(X_{it}u_{it} - \bar{X}_i u_{it}\right) = -E\bar{X}_i\bar{u}_i$$

For simplicity, assume $E\bar{X}_i = 0$ and X_{it} is a scalar. Then,

$$E\bar{X}_i\bar{u}_i = O\left(1/T\right).$$

First difference estimator

$$\hat{\beta}_{FD} \quad \rightarrow^p \quad \beta + \left(\frac{1}{T-1} \sum_{t=2}^{T} E\Delta X_{it} \Delta X_{it}' \right)^{-1} \underbrace{\frac{1}{T-1} \sum_{t=1}^{T} E\Delta X_{it} \Delta u_{it}}_{(**)}$$

$$(**) \quad = \quad -EX_{it}u_{it-1} = O(1)$$

Anderson-Hsiao (AH) estimator

Consider the FD model.

$$\Delta y_{it} = \Delta X_{it}'\beta + \Delta u_{it}.$$

Under the sequential exogeneity assumption,

$$EX_{is}u_{it} = 0 \ \text{ for } s = 1, \ldots t.$$

and

$$E\Delta X_{is}\Delta u_{it} = 0 \ \text{ for } s = 1, \ldots t-1.$$

Thus, we can use ΔX_{it-1} as an instrument for ΔX_{it}. We can also use X_{it-1} and X_{it-2} as instruments.

Arellano-Bond estimator (Panel GMM estimator)

AH IV estimator is consistent, but it is not efficient because it does not take into account all the available moment restrictions.

$$\Delta y_{it} = \Delta X_{it}'\beta + \Delta u_{it}.$$

Let $X_{it}^o = \underbrace{\left(X_{i1}' \ X_{i2}' \ \ldots X_{iT}'\right)}_{1\times kt}$. Define

$$Z_i = \underbrace{\begin{bmatrix} X_{i1}^o & 0 & \cdots & 0 \\ 0 & X_{i2}^o & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & X_{iT-1}^o \end{bmatrix}}_{KT(T-1)/2}$$

Then,

$$EZ_i\Delta u_i = 0 \ \text{ where } \Delta u_i = [\Delta u_2 \ \Delta u_3 \ldots \Delta u_T]'$$

In addition to the moment condition, we need the following rank condition to identify β.

Assumption $\quad \text{Rank}(EZ_i')\Delta X_i) = k$.

$$\hat{\beta}_{GMM} = \arg\min \left(\sum_{i=1}^{n} Z_i'(\Delta y_i - \Delta X_i \beta) \right)' W \left(\sum_{i=1}^{n} Z_i'(\Delta y_i - \Delta X_i \beta) \right)$$

The GMM estimator of β is

$$\begin{aligned} \hat{\beta}_{GMM} &= \left[\left(\frac{1}{n} \sum_{i=1}^{n} \Delta X_i' Z_i \right) W \left(\frac{1}{n} \sum_{i=1}^{n} Z_i' \Delta X_i \right) \right]^{-1} \\ &\quad \times \left(\frac{1}{n} \sum_{i=1}^{n} \Delta X_i' Z_i \right) W \left(\frac{1}{n} \sum_{i=1}^{n} Z_i' \Delta y_i \right) \\ &= (\Delta X' Z W Z' \Delta X)^{-1} \Delta X' Z W Z' \Delta y. \end{aligned}$$

Asymptotics

Under the orthogonality and rank conditions, we can show that $\hat{\beta}_{GMM}$ is consistent. We can also show that $\hat{\beta}_{GMM}$ is asymptotically normal

$$\sqrt{n}(\hat{\beta}_{GMM} - \beta) \to^d N(0, V_\beta)$$

where

$$V_\beta = (C'WC)^{-1} C'W \Lambda W C (C'WC)^{-1},$$

using the fact that

$$\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i' \Delta u_i \to^d N(0, \Lambda).$$

Selecting the weighting matrix

The optimal W is

$$W_{opt} = \Lambda^{-1} = var \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i' \Delta u_i \right) = \frac{1}{n} \sum_{i=1}^{n} E \left(Z_i' \Delta u_i \Delta u_i' Z_i \right).$$

Without homoskedasticity,

$$var \left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} Z_i' \Delta u_i \right) = E[Z_i' \Delta u_i \Delta u_i' Z_i]$$

A feasible GMM procedure

1 Let $\hat{\beta}$ be an initial consistent estimator of β, for example, $\hat{\beta} = \hat{\beta}_{2SLS}$.

2 Define $\Delta \tilde{u}_i = \Delta y_i - \Delta X_i \hat{\beta}_{2SLS}$

3 Construct a consistent estimator of $\hat{\Lambda} = n^{-1} \sum_{i=1}^{n} Z_i' \tilde{u}_i \tilde{u}_i' Z_i$ and choose $\hat{W} = \hat{\Lambda}^{-1}$

4 Use $\hat{W}$ to construct the GMM estimator.