Panel Econometrics – Notes, Chuxin Liu. Spring 2019

Chapter 2: One-way Error Component

(2.1) $y_{it} = \alpha + X_{it}'\beta + u_{it}$, $i=1,\cdots,N$; $t=1,\cdots,T$

Most of the panel data application utilize a one-way error component model for the disturbances, with.

(2.2) $u_{it} = \mu_i + \nu_{it}$

In vector form (2.1) can be written as

(2.3) $y = \alpha \iota_{NT} + X\beta + u = Z\delta + u$ where $Z = [\iota_{NT} \ X]$, $\delta' = [\alpha' \ \beta']$

(2.4) $u = Z_\mu \mu + \nu$

where $u' = (u_{11}, u_{12}, \cdots, u_{1T}, u_{21}, \cdots u_{2T}, \cdots, u_{N1} \cdots u_{NT})$

$$Z_\mu = I_N \otimes \iota_T$$

$$\begin{aligned} Z_\mu Z_\mu' = (I_N \otimes \iota_T)(I_N \otimes \iota_T)' &= (I_N \otimes \iota_T)(I_N' \otimes \iota_T') \\ &= I_N \otimes (\iota_T \iota_T') \\ &= I_N \otimes J_T \end{aligned}$$

✱ Define $P = Z_\mu (Z_\mu' Z_\mu)^{-1} Z_\mu' = I_N \otimes \bar{J}_T$ where $\bar{J}_T = J_T / T$

✱ Define $Q = I_{NT} - P$

$\Rightarrow$ P is the individual average across time, Q is deviation from individual average!

① $Qy = y_{it} - \bar{y}_{i\cdot}$

② P and Q are symmetric idempotent matrices, $P' = P$, $P^2 = P$
$\text{rank}(P) = tr(P) = N$ and $\text{rank}(Q) = tr(Q) = N(T-1)$

③ P and Q are orthogonal, $PQ = 0$

④ $P + Q = I_{NT}$

2.2 The fixed effects model

In this case. the μ_i are assumed to be fixed parameters to be estimated

$\nu_{it} \sim IID(0, \sigma_\nu^2)$

X_{it} are independent of the ν_{it} for all i and t

(2.5) $y = \alpha l_{NT} + X\beta + Z_\mu \mu + \nu = Z\delta + Z_\mu \mu + \nu$

LSDV (least squares dummy variables) estimator by:

(2.6) $Qy = QX\beta + Q\nu$

(2.7) $\tilde{\beta} = (X'QX)^{-1}X'Qy$ with $var(\tilde{\beta}) = \sigma_\nu^2 (X'QX)^{-1} = \sigma_\nu^2 (\tilde{X}'\tilde{X})^{-1}$

(2.8) $y_{it} = \alpha + \beta x_{it} + \mu_i + \nu_{it}$

β and $(\alpha + \mu_i)$ are estimable, not α and μ_i separately, unless a restriction like $\sum_{i=1}^N \mu_i = 0$ is imposed!

How to separate μ_i and α?

(2.9) $\bar{y}_{i\cdot} = \alpha + \beta \bar{X}_{i\cdot} + \mu_i + \bar{\nu}_{i\cdot}$

(2.10) $y_{it} - \bar{y}_{i\cdot} = \beta(X_{it} - \bar{X}_{i\cdot}) + (\nu_{it} - \bar{\nu}_{i\cdot}) \rightarrow$ obtain $\tilde{\beta}$, "within regression"

(2.11) $\bar{y}_{\cdot\cdot} = \alpha + \beta \bar{X}_{\cdot\cdot} + \bar{\nu}_{\cdot\cdot} \rightarrow$ obtain $\tilde{\alpha} \rightarrow$ obtain $\tilde{\mu}_i = \bar{y}_{i\cdot} - \tilde{\alpha} - \tilde{\beta}\bar{X}_{i\cdot}$

* Disadvantage of FE

① When N is very large, FE or LSDV, suffers large loss of degree of freedom

② FE estimator cannot estimate the effect of any time-invariant variables like sex, race, religion, schooling because they are wiped out by Q

③ If (2.5) is true, LSDV is BLUE as long as $\nu_{it} \sim IID(0, \sigma_\nu^2)$
$T \rightarrow \infty$. FE estimator is consistent.
T fixed, $N \rightarrow \infty$: FE of β is consistent, $(\alpha + \mu_i)$ are NOT consistent
"incidental parameter problem"

#1 Testing for FE

$$F_0 = \frac{(RRSS - URSS)/N-1}{URSS/(NT-N-k)} \overset{H_0}{\sim} F_{N-1,\ N(T-1)-k}$$

$H_0 : \mu_1 = \mu_2 = \cdots = \mu_{N-1} = 0$

#2. Computational Warning

#3. Robust estimates of standard errors

For the within estimator, Arellano (1987) suggests a simple method for obtaining robust estimates of the standard errors that allow for a general variance-covariance matrix on the v_{it} as in White (1980)

Stack the panel as an equation for each individual:

$$(2.13) \quad \underset{T\times 1}{y_i} = Z_i\delta + \mu_i \iota_T + v_i$$

$E(v_i v_i') = \Omega_i$ for $i = 1, \cdots, N$.

Assume $E(v_i v_j') = 0$ for $i \neq j$

$$(2.14) \quad \tilde{y}_i = \tilde{X}_i \beta + \tilde{v}_i \quad \text{wher } \tilde{y} = Qy,\ \tilde{y} = (\tilde{y}_1', \cdots, \tilde{y}_N')'$$

$$\tilde{y}_i = (I_T - \bar{J}_T) y_i$$

$$(2.15) \quad N^{1/2}(\hat{\beta} - \beta) \sim N(0, M^{-1} V M^{-1})$$

where $M = \operatorname{plim}(\tilde{X}'\tilde{X})/N$ and $V = \operatorname{plim} \sum_{i=1}^N (\tilde{X}_i' \Omega_i \tilde{X}_i)/N$

Note that $\tilde{X}_i = (I_T - \bar{J}_T) X_i$, $\tilde{X}' Q \operatorname{diag}[\Omega_i] Q \tilde{X} = \tilde{X}' \operatorname{diag}[\Omega_i] \tilde{X}$

V is estimated by $\hat{V} = \sum_{i=1}^N \tilde{X}_i' \tilde{u}_i \tilde{u}_i' \tilde{X}_i / N$ where $\tilde{u}_i = \tilde{y}_i - \tilde{X}_i \tilde{\beta}$

Robust $\operatorname{var}(\tilde{\beta}) = (\tilde{X}'\tilde{X})^{-1} \left[\sum_{i=1}^N \tilde{X}_i' \tilde{u}_i \tilde{u}_i' \tilde{X}_i \right] (\tilde{X}'\tilde{X})^{-1}$

2.3 The Random Effects Model

$\mu_i \sim IID(0, \sigma_\mu^2)$

$\nu_{it} \sim IID(0, \sigma_\nu^2)$

μ_i independent of the ν_{it}

X_{it} independent of the μ_i and ν_{it} for all i and t

(2.17) $\Omega = E(uu') = Z_\mu E(\mu\mu') Z_\mu' + E(\nu\nu')$

$$= \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\nu^2 (I_N \otimes I_T)$$

This implies a homoskedastic variance $var(u_{it}) = \sigma_\mu^2 + \sigma_\nu^2 \ \forall\ i$ and t

This implies an equicorrelated block-diagonal covariance matrix which exhibits serial correlation over time only between the disturbance of the same individual.

$$\Rightarrow Cov(u_{it}, u_{js}) = \sigma_\mu^2 + \sigma_\nu^2 \quad i=j,\ s=t$$
$$= \sigma_\mu^2 \quad i=j,\ s\neq t$$

$$\Rightarrow \rho = correl(u_{it}, u_{js}) = 1 \quad i=j,\ t=s$$
$$= \frac{\sigma_\mu^2}{\sigma_\mu^2+\sigma_\nu^2} \quad i=j,\ t\neq s$$

How to get Ω^{-1} for GLS?

Define $E_T = I_T - \bar{J}_T$

$$\Omega = \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\nu^2 (I_N \otimes I_T)$$
$$= \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\nu^2 (I_N \otimes (E_T + \bar{J}_T))$$
$$= \sigma_\mu^2 (I_N \otimes J_T) + \sigma_\nu^2 (I_N \otimes E_T) + \sigma_\nu^2 (I_N \otimes \bar{J}_T)$$
$$= (T\sigma_\mu^2 + \sigma_\nu^2)(I_N \otimes \bar{J}_T) + \sigma_\nu^2 (I_N \otimes E_T) = \sigma_1^2 P + \sigma_\nu^2 Q$$

where $\sigma_1^2 = T\sigma_\mu^2 + \sigma_\nu^2$

(2.19) $\Omega^{-1} = \frac{1}{\sigma_1^2} P + \frac{1}{\sigma_v^2} Q$

(2.20) $\Omega^{-1/2} = \frac{1}{\sigma_1} P + \frac{1}{\sigma_v} Q$

The best quadratic unbiased (BQU) estimators of the variance components arise naturally from the spectral decomposition of Ω.

In fact, $Pu \sim (0, \sigma_1^2 P)$ and $Qu \sim (0, \sigma_v^2 Q)$

(2.21) $\hat{\sigma}_1^2 = \frac{u'Pu}{tr(P)} = T \sum_{i=1}^{N} \bar{u}_{i\cdot}^2 / N$

(2.22) $\hat{\sigma}_v^2 = \frac{u'Qu}{tr(Q)} = \frac{\sum_{i=1}^{N} \sum_{t=1}^{T} (u_{it} - \bar{u}_{i\cdot})^2}{N(T-1)}$

(2.21) and (2.22) provide the BQU Estimators for σ_1^2 and σ_v^2 respectively!

The true disturbances are not known and therefore (2.21) and (2.22) are not feasible? Wallace and Hussain (1969) suggest using $\hat{u}_{OLS}$ instead of the true u, because under random effect model, OLS is unbiased and consistent, but no longer efficient.

Amemiya (1971) uses LSDV residual: $\tilde{u} = y - \tilde{\alpha} \iota_{NT} - X\tilde{\beta}$ where $\tilde{\alpha} = \bar{y}_{\cdot\cdot} - \bar{X}_{\cdot\cdot}\tilde{\beta}$

(2.23) $$\begin{pmatrix} \sqrt{NT}(\hat{\sigma}_v^2 - \sigma_v^2) \\ \sqrt{N}(\hat{\sigma}_\mu^2 - \sigma_\mu^2) \end{pmatrix} \sim N\left(0, \begin{pmatrix} 2\sigma_v^4 & 0 \\ 0 & 2\sigma_\mu^4 \end{pmatrix}\right)$$

where $\hat{\sigma}_\mu^2 = (\hat{\sigma}_1^2 - \hat{\sigma}_v^2)/T$

Swamy and Arora (1972) run 2 regressions to get variance component estimates

$$\begin{pmatrix} Qy \\ Py \end{pmatrix} = \begin{pmatrix} QZ \\ PZ \end{pmatrix} \delta + \begin{pmatrix} Qu \\ Pu \end{pmatrix} \qquad \Omega = \begin{pmatrix} \sigma_v^2 Q & 0 \\ 0 & \sigma_1^2 P \end{pmatrix}$$

(2.24) $\hat{\hat{\sigma}}_v^2 = [y'Qy - y'QX(X'QX)^{-1}X'Qy]/[N(T-1)-K]$

(2.27) $\hat{\hat{\sigma}}_1^2 = [y'Py - y'PZ(Z'PZ)^{-1}Z'Py]/(N-K-1)$

(2.29) $$\begin{pmatrix} Qy \\ (P - \bar{J}_{NT})y \end{pmatrix} = \begin{pmatrix} QX \\ (P - \bar{J}_{NT})X \end{pmatrix} \beta + \begin{pmatrix} Qu \\ (P - \bar{J}_{NT})u \end{pmatrix} \qquad \Omega = \begin{pmatrix} \sigma_v^2 Q & 0 \\ 0 & \sigma_1^2 (P - \bar{J}_{NT}) \end{pmatrix}$$

(2.30) $\hat{\beta}_{GLS} = [(X'QX/\sigma_v^2) + X'(P-\bar{J}_{NT})X/\sigma_1^2]^{-1}[(X'Qy/\sigma_v^2) + X'(P-\bar{J}_{NT})y/\sigma_1^2]$

$= [W_{XX} + \phi^2 B_{XX}]^{-1}[W_{Xy} + \phi^2 B_{Xy}]$

with $\mathrm{var}(\hat{\beta}_{GLS}) = \sigma_v^2[W_{XX} + \phi^2 B_{XX}]^{-1}$

Note that: $W_{XX} = X'QX$, $B_{XX} = X'(P-\bar{J}_{NT})X$

$\phi^2 = \sigma_v^2/\sigma_1^2 = \sigma_v^2/(T\sigma_\mu^2 + \sigma_v^2)$

$\tilde{\beta}_{within} = W_{XX}^{-1} W_{Xy}$

$\hat{\beta}_{between} = B_{XX}^{-1} B_{Xy}$

(2.31) $\hat{\beta}_{GLS} = W_1 \tilde{\beta}_{within} + W_2 \hat{\beta}_{between}$

where $W_1 = [W_{XX} + \phi^2 B_{XX}]^{-1} W_{XX}$

$W_2 = [W_{XX} + \phi^2 B_{XX}]^{-1}(\phi^2 B_{XX}) = I - W_1$

(1) $\sigma_\mu^2 = 0$, $\phi^2 = 1$, $\hat{\beta}_{GLS} = \hat{\beta}_{OLS}$

(2) $T \to \infty$, $\phi^2 \to 0$, $\hat{\beta}_{GLS} \to \tilde{\beta}_{within}$

(3) B_{XX} dominates W_{XX}, $\hat{\beta}_{GLS} \to \hat{\beta}_{between}$

× "Feasible GLS is more efficient than LSDV for all but the fewest degrees of freedom"

× Nerlove (1971a)

2.4 Maximum Likelihood Estimation

Under normality of the disturbances,

(2.32) $L(\alpha, \beta, \phi^2, \sigma_v^2) = \text{constant} - \frac{NT}{2}\log\sigma_v^2 + \frac{N}{2}\log\phi^2 - \frac{1}{2\sigma_v^2}u'\Sigma^{-1}u$

where $\Omega = \sigma_v^2\Sigma$, $\phi^2 = \sigma_v^2/\sigma_1^2$

$\Sigma = Q + \phi^{-2}P$

(2.33) $\hat{\sigma}^2_{v,mle} = d'[Q + \phi^2(P - \bar{J}_{NT})]d/NT$

where $d = y - X\hat{\beta}_{mle}$

(2.34) $L_c(\beta, \phi^2) = \text{constant} - \frac{NT}{2}\log\{d'[Q + \phi^2(P - \bar{J}_{NT})]d\} + \frac{N}{2}\log\phi^2$

$\downarrow$

concentrated likelihood

$\Rightarrow$ (2.35) $\hat{\phi}^2 = \frac{d'Qd}{(T-1)d'(P - \bar{J}_{NT})d} = \frac{\sum\sum(d_{it} - \bar{d}_{i.})^2}{T(T-1)\sum(\bar{d}_{i.} - \bar{d}_{..})^2}$

$\Rightarrow$ (2.36) $\hat{\beta}_{mle} = [X'(Q + \phi^2(P - \bar{J}_{NT}))X]^{-1}X'[Q + \phi^2(P - \bar{J}_{NT})]y$

2.5 Prediction

Best linear unbiased predictor (BLUP) of $y_{i,T+S}$ under GLS

(2.37) $\hat{y}_{i,T+S} = Z'_{i,T+S}\hat{\delta}_{GLS} + w'\Omega^{-1}\hat{u}_{GLS}$ for $s \geq 1$

where $\hat{u}_{GLS} = y - Z\hat{\delta}_{GLS}$, $w = E(u_{i,T+S} \cdot u)$

(2.38) $u_{i,T+S} = \mu_i + \nu_{i,T+S}$

and $w = \sigma_\mu^2(l_i \otimes \iota_T)$ where l_i is the i^{th} column of I_N, l_i is a vector that has 1 in the ith position and 0 elsewhere

(2.39) $w'\Omega^{-1} = \sigma_\mu^2 (\iota_i' \otimes \iota_T') \left[\frac{1}{\sigma_1^2} P + \frac{1}{\sigma_\nu^2} Q \right] = \frac{\sigma_\mu^2}{\sigma_1^2} (\iota_i' \otimes \iota_T')$

since $(\iota_i' \otimes \iota_T') P = (\iota_i' \otimes \iota_T')$

$(\iota_i' \otimes \iota_T') Q = 0$