

Modelos básicos de datos en panel

Gabriel Montes-Rojas

3 tipos de estructuras de datos

- **Corte transversal (Cross section)** Muestra de individuos, hogares, firmas, países, etc. que se toman en un momento dado del tiempo.

$\{y_i, x_i\}_{i=1}^N$, donde i representa individuos.

Ej.: EPH de 2013, datos de PBI en 2013 para muchos países.

- **Serie de tiempo** Muestra por varios periodos del mismo individuo, país, firma, etc.

$\{y_t, x_t\}_{t=1}^T$, donde t es tiempo.

Ej.: Inflación en la Argentina.

- **Datos en panel** Combinación de los datos anteriores.

$\{y_{it}, x_{it}\}_{i=1, t=1}^{N, T}$, donde i representa individuos y t tiempo.

1. **Cortes transversales independientes**
2. **Muestras longitudinales**

One-way error components model

En un panel longitudinal el mismo individuo es observado a lo largo del tiempo.

$$y_{it} = \alpha + X'_{it}\beta + u_{it}$$

$i = 1, 2, \dots, N$ es el índice de individuos (firmas, familias, hogares, países),
 $t = 1, 2, \dots, T$ es el índice de tiempo,
 α es un escalar, β es $K \times 1$, X_{it} es la observación it de las K variables explicativas.
El error tiene esta estructura:

$$u_{it} = \mu_i + v_{it}$$

es el **error compuesto**. En inglés: **one-way error components model**.

- μ_i : efecto individual no observado que captura todos los factores constantes a lo largo del tiempo en y_{it} ,
- v_{it} : errores idiosincráticos o shocks.

One-way error components model

En notación matricial

$$y = \alpha \iota_{NT} + X\beta + u = Z\delta + u,$$

donde y y u son vectores $NT \times 1$, X es una matriz $NT \times K$, $Z = [\iota'_{NT}, X']'$, $\delta' = [\alpha, \beta']$ y ι_{NT} es un vector de 1s con dimensión $NT \times 1$

$$u = Z_{\mu}\mu + v$$

donde $Z_{\mu} = I_N \otimes \iota_T$ es una matriz $NT \times N$ de 1s y 0s, \otimes es el producto de Kronecker, ι_T es un vector de 1s de dimensión T , $\mu = [\mu_1, \dots, \mu_N]'$ es un vector $N \times 1$ que contiene los efectos individuales, y v es un vector $NT \times 1$ con los errores. De esta manera $Z_{\mu}\mu$ es un vector también $NT \times 1$.

Two-way error components model

Consideremos el modelo

$$y_{it} = \alpha + X'_{it}\beta + u_{it}$$

$i = 1, 2, \dots, N$ es el índice de individuos (firmas, familias, hogares, países),

$t = 1, 2, \dots, T$ es el índice de tiempo,

α es un escalar, β es $K \times 1$, X_{it} es la observación it de las K variables explicativas.

$$u_{it} = \mu_i + \lambda_t + v_{it}$$

es el **error compuesto**. En inglés: **two-way error components model**.

- μ_i : efecto individual no observado que captura todos los factores constantes a lo largo del tiempo en y_{it} ,
- λ_t : efecto temporal no observado que captura todos los factores constantes a lo de los individuos en y_{it} ,
- v_{it} : errores idiosincráticos o shocks.

Two-way error components model

En notación matricial

$$u = Z_{\mu}\mu + Z_{\lambda}\lambda + v,$$

donde $Z_{\lambda} = \iota_N \otimes I_T$ es una matriz $NT \times T$ de 1s y 0s, ι_N es un vector de 1s de dimensión N , $\lambda = [\lambda_1, \dots, \lambda_T]'$ es un vector $T \times 1$ que contiene los efectos temporales.

Paneles multi-dimensionales

- Los modelos anteriores se pueden extender a una mayor cantidad de dimensiones.
- Ejemplos de 3 dimensiones (i, j, t):
 - 1 Modelos de comercio exterior podrían tener efectos por exportador, importador, tiempo, etc.
 - 2 Modelos de inmigración podrían tener efectos por país de origen, país de destino, año de inmigración, etc.

En ambos casos podríamos tener efectos para cada elemento,

$$u_{ijt} = \mu_i + \alpha_j + \lambda_t + v_{ijt},$$

o efectos bilaterales, donde nos importan los efectos de a pares,

$$u_{ijt} = \mu_{ij} + \lambda_t + v_{ijt}$$

- Un caso particular es el de los modelos anidados. Por ejemplo, si evaluamos políticas educativas, tendríamos un efecto por alumno (i), que pertenece a una determinada clase (j), de una escuela (k) a lo largo del tiempo (t). Así podríamos pensar en una estructura de errores:

$$u_{ijkt} = \mu_i + \alpha_j + \gamma_k + \lambda_t + v_{ijkt}$$

- Ver el libro de Laszlo Matyas (<http://www.metrixmdp.eu>)

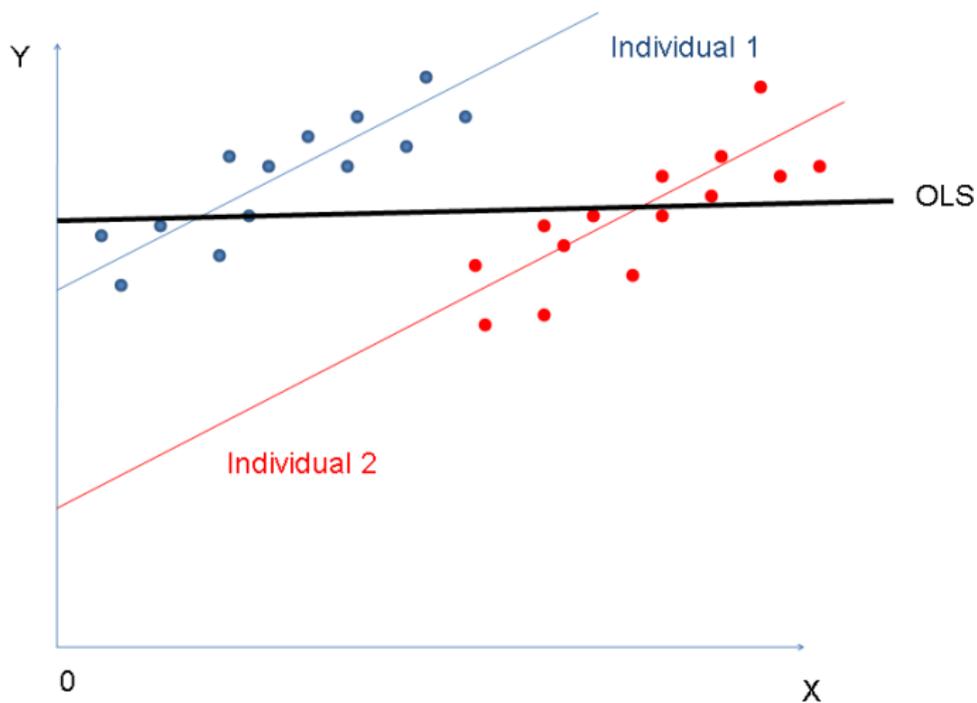
Una primera distinción

Si $\text{cov}(x_{it}, u_{it}) \neq 0 \Rightarrow$ OLS es sesgado. ¿Por qué?

Los siguientes estimadores se proponen como soluciones a este problema:

- **Estimador en primeras diferencias**
- **Efectos fijos**

Panel data



Estimador en diferencias (first differences, FD)

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + v_{it}$$

$$y_{it-1} = \beta_0 + \beta_1 x_{it-1} + \mu_i + v_{it-1}$$

$$\Rightarrow \Delta y_{it} = \beta_1 \Delta x_{it} + \Delta v_{it}$$

donde Δ es el operador de diferencias, $\Delta y_{it} = y_{it} - y_{it-1}$

¡Lo que pasa es que μ_i desaparece, entonces se acaba el problema!

Efectos fijos (fixed-effects, FE)

$$y_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + v_{it}$$

$$\bar{y}_i = \beta_0 + \beta_1 \bar{x}_i + \bar{\mu}_i + \bar{v}_i$$

$$\Rightarrow \tilde{y}_{it} = \beta_1 \tilde{x}_{it} + \tilde{v}_{i,t}$$

donde $\tilde{\cdot}$ es una transformación que se aplica a cada individuo (se usa en inglés **within** transformation), $\tilde{y}_{it} = y_{it} - \bar{y}_i$

¡Otra vez desaparece μ_i ! (porque $\bar{\mu}_i = T^{-1} \sum_{t=1}^T \mu_i = T^{-1} T \mu_i = \mu_i$)

Efectos fijos (fixed-effects, FE)

En los estimadores de efectos fijos μ_i se asumen como **parámetros fijos a ser estimados** y $v_{it} \sim IID(0, \sigma_v^2)$. $x_{it} \perp v_{it}, \forall i, t$.

Otro modo de ver los modelos FE es $E(v_{it}|x_{it}) \neq 0$ pero que $E(v_{it}|x_{it}, \mu_i) = 0$.
Entonces necesitamos estimar μ_i .

Efectos fijos (fixed-effects, FE)

- En notación matricial

$$y = \alpha \iota_{NT} + X\beta + Z_{\mu}\mu + v = Z\delta + Z_{\mu}\mu + v$$

- Definamos $P_{\mu} = Z_{\mu}(Z'_{\mu}Z_{\mu})^{-1}Z'_{\mu}$ como la matriz de proyección en Z_{μ} , el subespacio de dummies por individuo. $Z_{\mu}Z'_{\mu} = I_N \otimes J_T$. Entonces $P_{\mu} = Z_{\mu}(Z'_{\mu}Z_{\mu})^{-1}Z'_{\mu} = I_N \otimes \overline{J_T}$ donde $\overline{J_T} = J_T / T$.
- P_{μ} es una matriz que promedia las observaciones para los individuos. Así $P_{\mu}y$ tiene elementos $\bar{y}_i = 1/T \sum_{t=1}^T y_{it}$ repetido T veces para cada i .
- Probar que P_{μ} es simétrica ($P'_{\mu} = P_{\mu}$) e idempotente ($P_{\mu} \times P_{\mu} = P_{\mu}$) por lo que $rank(P_{\mu}) = tr(P_{\mu}) = N$.

Efectos fijos (fixed-effects, FE)

- Definamos también el complemento $Q_\mu = I_{NT} - P_\mu$. Q_μ se define entonces como las desviaciones con respecto a las medias (en inglés within-group operator): $\tilde{y}_{it} = y_{it} - \bar{y}_i$.
- Probar que Q_μ es simétrica e idempotente, y que $P_\mu Q_\mu = 0_{NT}$, por lo que $\text{rank}(Q_\mu) = \text{tr}(Q_\mu) = N(T - 1)$.
- Entonces,

$$Q_\mu y = Q_\mu X\beta + Q_\mu v, \Rightarrow \hat{\beta}_{FE} = (X' Q_\mu X)^{-1} X' Q_\mu y$$

- Notar que $\hat{\beta}_{FE} = (\sum_{i=1}^N (X_i' (I_T - \bar{J}_T) X_i))^{-1} \sum_{i=1}^N (X_i' (I_T - \bar{J}_T) y_i)$. ¿Qué significa?
- ¿Qué tipo de variables quedan excluidas?

Efectos fijos (fixed-effects, FE)

Notar que $\hat{\beta}_{FE}$ es equivalente a un modelo OLS con una dummy para cada individuo i , $\hat{\beta}_{LSDV}$.

La prueba es una aplicación del Teorema de **Frisch-Waugh-Lovell**

$$y = X_1\beta_1 + X_2\beta_2 + u$$

con estimadores OLS $\hat{\beta} = [\hat{\beta}_1 \hat{\beta}_2]$.

Usemos $M_1y = M_1X_2\beta_2 + M_1u$ donde M_1 es la proyección residual de X_1 , el teorema muestra que

$$\hat{\beta}_2 = (X_2' M_1 X_2)^{-1} X_2' M_1 y$$

Prueba: Consideremos $y = P_X y + M_X y = X_1 \hat{\beta}_1 + X_2 \hat{\beta}_2 + M_X y$. Multiplicar ambos lados por $X_2' M_1 y$ obtenemos $X_2' M_1 y = X_2' M_1 X_2 \hat{\beta}_2$. (usando $M_1 X_1 = 0$ y $M_X M_1 X_2 = 0$). Resolver.

Efectos fijos (fixed-effects, FE)

- Notar que $\text{var}(\hat{\beta}_{FE}) = \sigma_v^2 (X' Q_{\mu} X)^{-1}$.
- Podemos plantear $\hat{\mu}_i = \hat{y}_i - \hat{\beta}_{FE} \bar{x}_i$ o $\hat{\mu}_{FE} = Q_{\mu} y - \hat{\beta}_{FE} Q_{\mu} X$ como el estimador de los “efectos fijos”.
- Si $T \rightarrow \infty$, $(\hat{\mu}_{FE}, \hat{\beta}_{FE})$ son estimadores consistentes e insesgados.
- Pero si T está fijo y $N \rightarrow \infty$, sólo $\hat{\beta}_{FE}$ es consistente, aunque ambos son insesgados. El problema es conocido como el problema de parámetros incidentales de Neyman y Scott (1948).
- Contraste de efectos fijos: correr el modelo LSDV, y contrastar conjuntamente por la significatividad de las dummies.

Efectos fijos (fixed-effects, FE)-two-way

Para el modelo two-way tenemos $Z_\lambda Z'_\lambda = J_N \otimes I_T$. Entonces

$P_\lambda = Z_\lambda (Z'_\lambda Z_\lambda)^{-1} Z'_\lambda = \overline{J_N} \otimes I_T$ donde $\overline{J_N} = J_N / N$.

Definamos la proyección residual

$$Q_{\mu\lambda} = E_N \otimes E_T = I_N \otimes I_T - I_N \otimes \overline{J_T} + \overline{J_N} \otimes I_T + \overline{J_N} \otimes \overline{J_T}$$

donde $E_N = I_N - \overline{J_N}$ y $E_T = I_T - \overline{J_T}$. Esta transformación elimina los efectos de μ_i y λ_t simultáneamente. La matriz $Q_{\mu\lambda}$ hace la siguiente transformación:

$y_{it} - \bar{y}_i - \bar{y}_t + \bar{y}$. Entonces,

$$Q_{\mu\lambda} y = Q_{\mu\lambda} X \beta + Q_{\mu\lambda} v, \Rightarrow \hat{\beta}_{FE-2} = (X' Q_{\mu\lambda} X)^{-1} X' Q_{\mu\lambda} y$$

¿Qué tipo de variables quedan excluidas?

Efectos fijos (fixed-effects, FE)

Usar efectos fijos tiene algunas desventajas:

- Hay un costo en comparación con OLS sin controlar por las dummies por individuo. La transformación within es equivalente a estimar una dummy para cada individuo. O sea estimar $N - 1$ parámetros adicionales. Más parámetros significa menos precisión.
- Entonces eso afecta los grados de libertad y por ende la precisión de lo que estimamos. Los grados de libertad son $NT - N - K$, comparado con $NT - 1 - K$
- También la transformación within (y first-differences) elimina TODO aquello que esta fijo para cada individuo. Entonces no se puede medir por ejemplo el efecto de SEXO, para individuos, o CONTINENTE para países.

Efectos fijos (fixed-effects, FE)

Comparando FE con FD (Baltagi, p.17):

- FE es más eficiente que FD cuando $v_{it} \sim iid(0, \sigma_v^2)$.
- FD es más eficiente que FE cuando v_{it} es un paseo aleatorio (random walk).

Efectos aleatorios (random-effects, RE)

Un modelo alternativo es el de efectos aleatorios. Tiene un supuesto MUY importante y restrictivo: $cov(X, \mu) = 0$.

Consideremos el modelo

$$y_{it} = \beta_0 + \beta_1 x_{it} + u_{it} = \beta_0 + \beta_1 x_{it} + \mu_i + v_{it}$$

En este caso hay **correlación serial**:

$Cov(u_{it}, u_{is}) = Cov(\mu_i + v_{it}, \mu_i + v_{is}) = Var(\mu_i) \neq 0$ para $t \neq s$. El estimador de efectos aleatorios (RE) tiene en cuenta esta particularidad y produce un estimador eficiente GLS:

$$\hat{\beta}_{RE} = (X' \Omega^{-1} X)^{-1} (X' \Omega^{-1} y),$$

donde $\Omega = E(uu')$ es la matriz de varianzas-covarianzas de los errores compuestos.

Una de las ventajas de RE es que se pueden reincorporar variables fijas por individuos:

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \mu_i + v_{it}$$

donde z captura todas las variables que están fijas para cada individuo y que se pueden observar.

Efectos aleatorios (random-effects, RE)

En el modelo RE $\mu_i \sim iid(0, \sigma_\mu^2)$, $v_{it} \sim iid(0, \sigma_v^2)$, $\mu_i \perp v_{it}$. Además tenemos $X_{it} \perp \mu_i$ y $X_{it} \perp v_{it}$ para todo i y t .

$$\Omega = E(uu') = Z_\mu E(\mu\mu')Z_\mu' + E(vv') = \sigma_\mu^2(I_N \otimes J_T) + \sigma_v^2(I_N \otimes I_T).$$

Ω tiene esta estructura:

$$\begin{aligned} cov(u_{it}, u_{js}) &= \sigma_\mu^2 + \sigma_v^2 && \text{para } i = j, t = s \\ &= \sigma_\mu^2 && \text{para } i = j, t \neq s \\ &= 0 && \text{para } i \neq j \end{aligned}$$

Otra forma de verlo es como un modelo de **equicorrelación** intra cluster:

$$\begin{aligned} correl(u_{it}, u_{js}) &= 1 && \text{para } i = j, t = s \\ &= \sigma_\mu^2 / (\sigma_\mu^2 + \sigma_v^2) := \rho && \text{para } i = j, t \neq s \\ &= 0 && \text{para } i \neq j, t \neq s \end{aligned}$$

Efectos aleatorios (random-effects, RE)

¿Cuáles serían las consecuencias de ignorar la correlación serial?

Supongamos el siguiente modelo de regresión univariada: $y_{it} = \beta x_{it} + \mu_i + v_{it}$, donde $\mu_i \sim iid(0, \sigma_\mu^2)$,

$v_{it} \sim iid(0, \sigma_v^2)$, $\mu_i \perp v_{it}$. Además tenemos $x_{it} \perp \mu_i$ y $X_{it} \perp v_{it}$ para todo i y t .

- Por un lado se puede probar que el estimador OLS $\hat{\beta}_{OLS} = \frac{\sum_{i=1}^N \sum_{t=1}^T y_{it} x_{it}}{\sum_{i=1}^N \sum_{t=1}^T x_{it}^2}$ es insesgado y consistente.
- Sin embargo la varianza de OLS

$$\text{var}(\hat{\beta}_{OLS}) \neq \sigma_{\mu+v}^2 \frac{1}{\sum_{i=1}^N \sum_{t=1}^T x_{it}^2}$$

que es lo que estimaríamos de ignorar la correlación entre observaciones del mismo individuo. En particular, tendríamos

$$\text{var}(\hat{\beta}_{OLS}) = \sigma_{\mu+v}^2 \frac{1}{\sum_{i=1}^N \sum_{t=1}^T x_{it}^2} + \sigma_\mu^2 \frac{\sum_{i=1}^N \sum_{t=1}^T \sum_{s=1, t \neq s}^T x_{it} x_{is}}{(\sum_{i=1}^N \sum_{t=1}^T x_{it}^2)^2}.$$

- El resultado es que los errores estándar que estima STATA son incorrectos, por lo que toda la inferencia es incorrecta.
- Esta diferencia se llama el "factor de Moulton". Notar que depende de la correlación intra-individuo de las X_s . En particular si las X_s no están relacionadas entre sí para un mismo individuo, o sea $\text{Cov}(x_{it}, x_{is}) = 0$, $t \neq s$, entonces la varianza estándar es correcta.

Efectos aleatorios (random-effects, RE)

Para estimarlo vamos a tratar el problema como GLS (generalized least squares) donde se estima Ω .

- Siguiendo a Wansbeek y Kapteyn (1982,1983), Baltagi p.18, reemplacemos J_T por $T\overline{J_T}$, y I_T por $(E_T + \overline{J_T})$, donde por definición $E_T = I_T - \overline{J_T}$. Entonces reagrupando términos tenemos

$$\begin{aligned}\Omega &= T\sigma_\mu^2(I_N \otimes \overline{J_T}) + \sigma_v^2(I_N \otimes E_T) + \sigma_v^2(I_N \otimes \overline{J_T}) \\ &= (T\sigma_\mu^2 + \sigma_v^2)(I_N \otimes \overline{J_T}) + \sigma_v^2(I_N \otimes E_T) = \sigma_1^2 P_\mu + \sigma_v^2 Q_\mu\end{aligned}$$

donde $\sigma_1^2 = (T\sigma_\mu^2 + \sigma_v^2)$.

- Esta es la descomposición espectral de Ω donde σ_1^2 (de multiplicidad N) y σ_v^2 (de multiplicidad $N(T-1)$) son las raíces características de Ω .
- Esto permite hallar $\Omega^r = \sigma_1^{2r} P + \sigma_v^{2r} Q$ para todo r (en particular $r = -1$). [¿Por qué? Probar que es cierto.]

Efectos aleatorios (random-effects, RE)

- Entonces, surgen los siguientes estimadores

$$\hat{\sigma}_1^2 = \frac{u' P_\mu u}{\text{tr}(P_\mu)} = T \sum_{i=1}^N \bar{u}_i^2 / N$$

$$\hat{\sigma}_v^2 = \frac{u' Q_\mu u}{\text{tr}(Q_\mu)} = \frac{\sum_{t=1}^T \sum_{i=1}^N (u_{it} - \bar{u}_i)^2}{N(T-1)}$$

- Notar que u no es observado... hay diferentes alternativas, todas ellas consistentes: (i) usar los residuos OLS, (ii) los residuos de FE, y (iii) otras.

Efectos aleatorios (random-effects, RE)

El estimador RE es entonces

$$\begin{aligned}\hat{\beta}_{RE} &= [(X'Q_{\mu}X/\sigma_v^2) + (X'(P_{\mu} - \bar{J}_{NT})X/\sigma_1^2)]^{-1}[(X'Q_{\mu}y/\sigma_v^2) + (X'(P_{\mu} - \bar{J}_{NT})y/\sigma_1^2)] \\ &= [W_{XX} + \phi^2 B_{XX}]^{-1}[W_{Xy} + \phi^2 B_{Xy}]\end{aligned}$$

donde $W_{XX} = X'Q_{\mu}X$, $B_{XX} = X'(P_{\mu} - \bar{J}_{NT})X$, $\phi^2 = \sigma_v^2/\sigma_1^2$. También $\text{var}(\hat{\beta}_{RE}) = \sigma_v^2[W_{XX} + \phi^2 B_{XX}]^{-1}$.

Por otro lado $\hat{\beta}_{FE} = W_{XX}^{-1}W_{Xy}$ y $\hat{\beta}_{BE} = B_{XX}^{-1}B_{Xy}$, FE: fixed-effects-within; BE: between. Entonces, $\hat{\beta}_{RE} = W_1\hat{\beta}_{FE} + W_2\hat{\beta}_{BE}$ donde $W_1 = [W_{XX} + \phi^2 B_{XX}]^{-1}W_{XX}$ y $W_2 = [W_{XX} + \phi^2 B_{XX}]^{-1}\phi^2 B_{XX} = I - W_1$.

Nota: Análisis cuando $T \rightarrow \infty$, $N \rightarrow \infty$, $(\sigma_v^2, \sigma_{\mu}^2) \rightarrow 0, \infty$.

Efectos aleatorios (random-effects, RE)-two-way

En el modelo RE two-way $\mu_i \sim iid(0, \sigma_\mu^2)$, $\lambda_t \sim iid(0, \sigma_\lambda^2)$, $v_{it} \sim iid(0, \sigma_v^2)$, $\mu_i \perp v_{it}$, $\lambda_i \perp v_{it}$. Además tenemos $X_{it} \perp \mu_i$, $X_{it} \perp \lambda_t$ y $X_{it} \perp v_{it}$ para todo i y t .

$$\begin{aligned}\Omega = E(uu') &= Z_\mu E(\mu\mu')Z_\mu' + Z_\mu E(\lambda\lambda')Z_\mu' + E(vv') = \\ &\sigma_\mu^2(I_N \otimes J_T) + \sigma_\lambda^2(J_N \otimes I_T) + \sigma_v^2(I_N \otimes I_T).\end{aligned}$$

Ω tiene esta estructura:

$$\begin{aligned}\text{cov}(u_{it}, u_{js}) &= \sigma_\mu^2 + \sigma_\lambda^2 + \sigma_v^2 && \text{para } i = j, t = s \\ &= \sigma_\mu^2 && \text{para } i = j, t \neq s \\ &= \sigma_\lambda^2 && \text{para } i \neq j, t = s \\ &= 0 && \text{para } i \neq j\end{aligned}$$

Efectos aleatorios (random-effects, RE)-two-way

Para estimarlo vamos a tratar el problema como GLS (generalized least squares) donde se estima Ω .

- Baltagi p.37, reemplazar J_N por $N\overline{J}_N$, I_N por $E_N + \overline{J}_N$, J_T por $T\overline{J}_T$, I_T por $E_T + \overline{J}_T$, entonces

$$\Omega = \sum_{j=1}^4 \eta_j Q_j$$

donde $\eta_1 = \sigma_v^2$, $\eta_2 = T\sigma_\mu^2 + \sigma_v^2$, $\eta_3 = N\sigma_\lambda^2 + \sigma_v^2$, y $\eta_4 = T\sigma_\mu^2 + N\sigma_v^2 + \sigma_v^2$, $Q_1 = E_N \otimes E_T$, $Q_2 = E_N \otimes \overline{J}_T$, $Q_3 = \overline{J}_N \otimes E_T$, y $Q_4 = \overline{J}_N \otimes \overline{J}_T$. Esta es la descomposición espectral de Ω donde η_j son las raíces características. Cada Q_j es simétrica e idempotente.

- Entonces, surgen los siguientes estimadores

$$\hat{\eta}_j = \frac{u' Q_j u}{\text{tr}(Q_j)}, j = 1, 2, 3, 4$$

Efectos aleatorios (random-effects, RE)-two-way

El estimador RE es entonces

$$\begin{aligned}\hat{\beta}_{RE} &= [(X'Q_1X/\sigma_v^2) + (X'Q_2X/\eta_2) + (X'Q_3X/\eta_3)]^{-1} \\ &\quad [(X'Q_1y/\sigma_v^2) + (X'Q_2y/\eta_2) + (X'Q_3y/\eta_3)] \\ &= [W_{XX} + \phi_2^2 B_{XX} + \phi_3^2 C_{XX}]^{-1} [W_{Xy} + \phi_2^2 B_{Xy} + \phi_3^2 C_{Xy}]\end{aligned}$$

donde $W_{XX} = X'Q_1X$, $B_{XX} = X'Q_2X$, $C_{XX} = X'Q_3X$, $\phi_j^2 = \sigma_v^2/\eta_j$, $j = 2, 3$. También $\text{var}(\hat{\beta}_{RE}) = \sigma_v^2 [W_{XX} + \phi_2^2 B_{XX} + \phi_3^2 C_{XX}]^{-1}$.

Entonces, $\hat{\beta}_{RE} = W_1 \hat{\beta}_{FE} + W_2 \hat{\beta}_{BE1} + W_3 \hat{\beta}_{BE2}$ donde
 $W_1 = [W_{XX} + \phi_2^2 B_{XX} + \phi_3^2 C_{XX}]^{-1} W_{XX}$, $W_2 = [W_{XX} + \phi_2^2 B_{XX} + \phi_3^2 C_{XX}]^{-1} \phi_2^2 B_{XX}$, y
 $W_3 = [W_{XX} + \phi_2^2 B_{XX} + \phi_3^2 C_{XX}]^{-1} \phi_3^2 C_{XX}$.

[Probar que en este caso también el estimador es un promedio ponderado de tres estimadores. Análisis cuando $T \rightarrow \infty$, $N \rightarrow \infty$, $(\sigma_v^2, \sigma_\mu^2, \sigma_\lambda^2) \rightarrow 0, \infty$.]

Contraste de Hausman

El estimador de efectos fijos es siempre **consistente**. Sin embargo, el de efectos aleatorios es válido si las Xs no están correlacionadas con los efectos individuales (μ_i). Entonces, un contraste de la validez de efectos aleatorios es

$$H_0 : \hat{\beta}_{FE} - \hat{\beta}_{RE} = 0$$

$$H_A : \hat{\beta}_{FE} - \hat{\beta}_{RE} \neq 0$$

Este es el llamado contraste de Hausman.

Contraste de Hausman

- Notemos que el estimador RE es GLS, entonces es el de menor varianza.
- Por lo tanto $var(\hat{\beta}_{FE}) - var(\hat{\beta}_{RE})$ es una matriz definida positiva. Para demostrarlo notemos que tanto W_{XX} como B_{XX} son matrices definidas positivas (formas cuadráticas), mientras que $\phi^2 > 0$. Entonces $[W_{XX} + \phi^2 B_{XX}] - W_{XX}$ es definida positiva, por lo que $[W_{XX} + \phi^2 B_{XX}]^{-1} - W_{XX}^{-1}$ es definida negativa, o $var(\hat{\beta}_{RE}) - var(\hat{\beta}_{FE})$ es definida negativa.
- Se puede probar que $var(\hat{\beta}_{FE} - \hat{\beta}_{RE}) = var(\hat{\beta}_{FE}) - var(\hat{\beta}_{RE})$.
- Entonces, definamos

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})' [var(\hat{\beta}_{FE}) - var(\hat{\beta}_{RE})]^{-1} (\hat{\beta}_{FE} - \hat{\beta}_{RE})$$

tiene una distribución χ_K^2 bajo la nula.

Contraste de Mundlak

(Hsiao, 2003, sec. 3.2; Wooldridge, 2012, sec. 10.7.3)

- Mundlak (1978) propone usar un modelo de regresión que contenga los promedios de individuos i de las variables que varían en i y t :

$$y_{it} = \beta_0 + \beta_1 x_{it} + \beta_2 z_i + \beta_3 \bar{x}_i + \mu_i + v_{it}$$

- Entonces un contraste de $H_0 : \beta_3 = 0$ es un contraste por la validez de RE.

¿Cómo implementar paneles en STATA?

- Los datos hay que organizarlos:

id	tiempo	YVAR	XVAR
1	1	y ₁₁	x ₁₁
1	2	y ₁₂	x ₁₂
1	3	y ₁₃	x ₁₃
2	1	y ₂₁	x ₂₁
2	2	y ₂₂	x ₂₂
2	3	y ₂₃	x ₂₃

Es muy importante que no haya valores repetidos. En el ejemplo se representa un panel balanceado. Podría ser desbalanceado si tuviera gaps (ej., la obs. $i = 1, t = 2$ no está).

- Primero STATA tiene que identificar que se trata de datos en paneles. Para eso se necesita una variable numérica, ej. `id`, que identifica el individuo. Luego, `iis id`
- Pero si se tiene una muestra longitudinal con una estructura de series de tiempo, con una variable `tiempo`, `tsset id tiempo`
- Nota: la variable de tiempo tiene que ser en números discretos consecutivos. O sea, $t = -2, -1, 0, 1, \dots$ o $t = 1980, 1981, 1982, \dots$

¿Cómo implementar paneles en STATA?

- Entonces estamos listos para usar datos en paneles:
- `reg D.y D.x1 D.x2 D.x3` (modelo en diferencias, sólo con `tsset`)
- `xtreg y x1 x2 x3, fe` (modelo de efectos fijos)
- `xi: reg y x1 x2 x3 i.id` (lo mismo pero implementado “a mano” con dummies para id) [Nota: comparar con una regresión de las variables transformadas within. ¿Cuál sería el problema con este modelo?]
- `xtreg y x1 x2 x3, re` (modelo de efectos aleatorios)
- `xtreg y x1 x2 x3, be` (modelo between)
- `xi: xtreg y x1 x2 x3 i.tiempo, fe` (modelo de efectos fijos, two-way)

¿Cómo implementar paneles en STATA?

- Contraste de Hausman test
xtreg y x1 x2 x3, fe
est store fe
xtreg y x1 x2 x3, re
est store re
hausman fe re

¿Cómo implementar paneles en STATA?

- <http://www.ats.ucla.edu/stat/stata/examples/eacspd/chapter10.htm>
- <http://www.wiley.com/legacy/wileychi/baltagi/datasets.html>
- `webuse grunfeld`