Paneles heterogéneos

Gabriel Montes-Rojas

Paneles heterogéneos

- En los modelos previos, toda la heterogeneidad se modelaba con un intercepto que era específico de cada individuo, y autocorrelación y/o heteroscedasticidad de los errores.
- Paneles heterogéneos trata de modelos con pendientes heterogéneas (slope heterogeneity).
- En general se necesita que $T \to \infty$ para poder estimar cada ecuación por separado.

Paneles heterogéneos

Dos formas de modelar la heterogeneidad, coeficientes fijos y coeficientes aleatorios.

- Podriamos considerar que cada individuo tiene una estructura de coeficientes única, β_i.

  $$y_{it} = \beta_i x_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T.$$

  Esto se puede modelar usando interacciones de las Xs con las dummies por individuo o estimando un modelo diferente para cada i.
  También podríamos tener que algunos coeficientes son específicos de cada individuo i y otros son comunes:

  $$y_{it} = \beta_i x_{it} + \gamma z_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T.$$

  Si todos los coeficientes fueran específicos para cada individuo, el modelo es equivalente a MCO para cada i.

- Otra forma más simple de modelar heterogeneidad es con un **modelo de coeficientes aleatorios**. En este caso, los coeficientes mismos con variables aleatorias, que se asumen estacionarios (misma media y varianza).

Paneles heterogéneos

Supongamos el modelo de Swamy (1970)

$$y_{it} = \beta_i x_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T,$$

$$\beta_i = \beta + \eta_i,\ E(\eta_i) = 0,\ E(\eta_i x'_{it}) = 0,$$

$$E(\eta_i \eta'_j) = \begin{cases} \Omega_\eta, & si \quad i = j, \\ 0, & si \quad i \neq j. \end{cases}$$

- η es una variable aleatoria de dimensión K (la cantidad de variables explicativas).
- En este modelo la aleatoriedad es invariante en el tiempo.
- Nos interesa la media β y la funcin de varianzas-covarianzas de los parmetros.

Paneles heterogéneos

Supongamos el modelo de Hsiao (1974,1975)

$$y_{it} = \beta_{it} x_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T$$

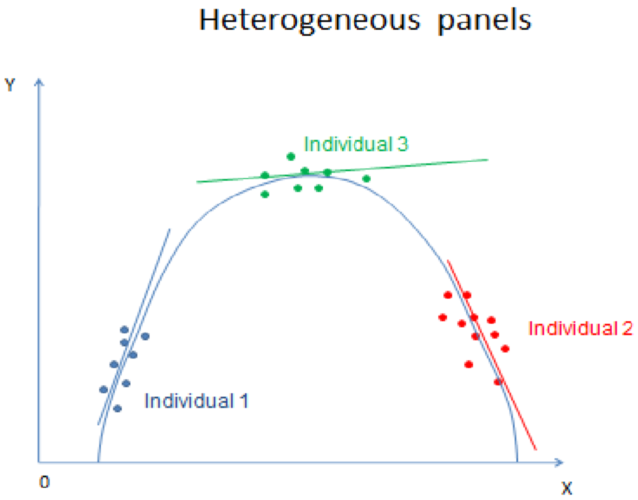
$$\beta_{it} = \beta + \eta_i + \lambda_t$$

$$E(\eta_i) = 0, E(\lambda_t) = 0, E(eta_i' \lambda_t) = 0$$

$$E(x_{it}' \eta_i) = 0, E(x_{it}' \lambda_t) = 0$$

$$E(\eta_i \eta_j') = \begin{cases} \Omega_\eta, & si \quad i = j, \\ 0, & si \quad i \neq j. \end{cases}$$

$$E(\lambda_t \lambda_h') = \begin{cases} \Omega_\lambda, & si \quad t = h, \\ 0, & si \quad t \neq h. \end{cases}$$


Heterogeneous panels
Y
Individual 3
Individual 2
Individual 1
0
X

Consecuencias de obviar heterogeneidad

- Supongamos el siguiente modelo generador de los datos:

$$y_{it} = \mu_i + \beta_i x_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T,$$

$$u_{it} \sim iid(0, \sigma_i^2),$$

$$\beta_i = \beta + \eta_i,$$

- Ahora supongamos que se estima el modelo

$$y_{it} = \alpha_i + \delta_x x_{it} + \delta_z z_{it} + \nu_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T,$$

  donde z es un regresor adicional que el investigador cree que es importante, cuando no lo es.

- ν_{it} es iid, distribuido independientemente de $w_{it} = (x_{it}, z_{it})$ para todo i, t. w sigue un proceso estacionario en covarianza de $\Omega_i = \begin{bmatrix} \omega_{ixx} & \omega_{ixz} \\ \omega_{izx} & \omega_{izz} \end{bmatrix}$.
- Consideremos el estimador FE

$$plim_{N,T \to} \hat{\hat{\delta}}_{FE} = \left[plim \frac{1}{NT} \sum_{i=1}^{N} W_i' Q_T W_i \right]^{-1} \left[plim \frac{1}{NT} \sum_{i=1}^{N} W_i' Q_T (X_i \beta_i) \right],$$

  donde $Q_T = I_T - \iota_T () \iota_T$ (¿qué transformación hace esta matriz?)

Consecuencias de obviar heterogeneidad

- Si $\beta_i = \beta$ entonces $plim_{N,T\to\infty}\hat{\delta}_{FE} = (\beta', 0')'$
- Si los paneles son heterogéneos

$$plim \frac{1}{NT} \sum_{i=1}^{N} W_i' Q(X_i \eta_i) = plim \begin{pmatrix} \frac{1}{NT} \sum_{i=1}^{N} X_i' Q(X_i \eta_i) \\ \frac{1}{NT} \sum_{i=1}^{N} X_i' Q(Z_i \eta_i) \end{pmatrix}$$

- Entonces necesitamos que $\frac{1}{N} \sum_{i=1}^{N} \omega_{ixx} \eta_i \xrightarrow{p} 0$ y $\frac{1}{N} \sum_{i=1}^{N} \omega_{ixz} \eta_i \xrightarrow{p} 0$. En otras palabras, no tiene que haber correlación sistemática entre los β_i y los **momentos segundos** de los regresores.
-
$$plim\hat{\delta}_{x,FE} - \beta = \frac{cov(\omega_{ixx}, \eta_i)E(\omega_{izz}) - E(\omega_{ixz})cov(\omega_{ixz}, \eta_i)}{E(\omega_{ixx})E(\omega_{izz}) - (E(\omega_{ixz}))^2}$$
$$plim\hat{\delta}_{z,FE} = \frac{cov(\omega_{ixz}, \eta_i)E(\omega_{ixx}) - E(\omega_{ixz})cov(\omega_{ixx}, \eta_i)}{E(\omega_{ixx})E(\omega_{izz}) - (E(\omega_{ixz}))^2}$$
- Supongamos el modelo $y_{it} = \alpha_i + \delta_x x_{it} + \delta_z x_{it}^2 + \nu_{it}$. En este caso en general no podemos obtener que FE estima bien a menos que β_i sea proporcional a x_{it}.

Estimador de Swamy (1970)

- En el modelo de Swamy, la matriz de varianzas-covarianzas es

$$\Sigma = E(\nu\nu') = \begin{bmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_N \end{bmatrix}$$

donde $\nu_{it} = u_{it} + \eta_i x_{it}$ y $\Sigma_i = \sigma_i^2 I_T + X_i \Omega_\eta X_i'$ para $i = 1, 2, ..., N$. [Notas que $X_i \Omega_\eta X_i$ modela las correlaciones entre η y X.]

- Se propone entonces el estimador

$$\hat{\beta}_{SW} = (X'\Sigma X)^{-1}(X'\Sigma y) = [\sum_{i=1}^{N} X_i' \Sigma_i^{-1} X_i]^{-1} \sum_{i=1}^{N} X_i' \Sigma_i^{-1} y_i$$

Estimador de Swamy (1970)

- Si Ω_η es no singular,

$$\Sigma_i^{-1} = \frac{I_T}{\sigma_i^2} - \frac{X_i}{\sigma_i^2}\left(\frac{X_i'X_i}{\sigma_i^2} + \Omega_\eta^{-1}\right)^{-1}\frac{X_i'}{\sigma_i^2} = \frac{I_T}{\sigma_i^2} - \frac{X_i\Omega_\eta}{\sigma_i^2}\left(I_K + \frac{X_i'X_i}{\sigma_i^2}\Omega_\eta\right)^{-1}\frac{X_i'}{\sigma_i^2}$$

(para la última igualdad no necesitamos que sea no singular)

Estimador de Swamy (1970)

- El estimador se puede escribir como:

$$\overline{\hat{\overline{\beta}}}_{SW} = \sum_{i=1}^{N} R_i \hat{\beta}_i$$

donde $R_i = [\sum_{i=1}^{N}(\Omega_\eta + \Sigma_{\hat{\beta}_i})^{-1}]^{-1}(\Omega_\eta + \Sigma_{\hat{\beta}_i})^{-1}$, donde $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i$ y $\Sigma_{\hat{\beta}_i} = \sigma_i^2(X_i'X_i)^{-1}$ son los estimadores y sus varianzas, respectivamente, para cada i.

- $var(\overline{\hat{\overline{\beta}}}_{SW}) = [\sum_{i=1}^{N}(\Omega_\eta + \Sigma_{\hat{\beta}_i})^{-1}]^{-1}$.
- Para estimar la varianza necesitamos un estimador de Ω_η. Se propone:

$$\hat{\Omega}_\eta = \frac{1}{N-1}\sum_{i=1}^{N}(\hat{\beta}_i - \hat{\beta}_{MG})(\hat{\beta}_i - \hat{\beta}_{MG})'$$

donde $(\hat{\beta}_i - \hat{\beta}_{MG})$ se define abajo como el estimador de *mean-group*.

Estimador media de grupos (MG estimator)

- Una alternativa es el mean-group (MG)

$$\hat{\beta}_{MG} = \frac{1}{N}\sum_{i=1}^{N}\hat{\beta}_i$$

  donde $\hat{\beta}_i = (X_i'X_i)^{-1}X_i'y_i$ son los estimadores OLS de cada i.

- Notar que $\hat{\beta}_i = \beta + \eta_i + \zeta_i$, donde $\zeta_i = (X_i'X_i)^{-1}X_i'u_i$, entonces $\hat{\beta}_{MG} = \beta + \bar{\eta} + \bar{\zeta}$ con $\bar{\eta} = \frac{1}{N}\sum_{i=1}^{N}\eta_i$ y $\bar{\zeta} = \frac{1}{N}\sum_{i=1}^{N}\zeta_i$.
- $var(\hat{\beta}_{MG}) = var(\bar{\eta}) + var(\bar{\zeta}) = \frac{1}{N}\Omega_\eta + \frac{1}{N^2}\sum_{i=1}^{N}\sigma_i^2 E\left[T(X_i'X_i)^{-1}\right]$

Contrastes para homogeneidad vs. heterogeneidad

- Dadas las potenciales consecuencias negativas de obviar la heterogeneidad es importante chequear por la misma.
- Los test de heterogeneidad se pueden interpretar como de *poolability*, o sea, si se pueden "juntar" los modelos en una misma función lineal.
- Supongamos el siguiente modelo generador de los datos:

$$y_{it} = \mu_i + \beta_i x_{it} + u_{it}, i = 1, 2, ..., N; t = 1, 2, ..., T$$

$$u_{it} \sim iid(0, \sigma_i^2)$$

- La hipótesis nula de interés es

$$H_0 : \beta_i = \beta, \forall i$$

  y la alternativa

$$H_A : \beta_i \neq \beta_j, i \neq j$$

- Notar que los contrastes de homogeneidad se refieren a las pendientes β, y en general no se contrasta por el intercepto.

Contrastes para homogeneidad vs. heterogeneidad

- F-test de $K(N-1)$ restricciones. Este contraste funciona para N fijo y $T \to \infty$ y $\sigma_i^2 = \sigma^2$ (homoscedasticidad):

$$F = \frac{N(T-K-1)}{K(N-1)} \frac{RSSR - USSR}{USSR}$$

  Este test corresponde a correr una regresión de efectos fijos con las interacciones de cada i con cada variable de control X, excepto una, más la X sin interactuar, y luego chequear que todas las interacciones son cero. En general tiene un resultado pobre si N es grande y/o la varianza es heteroscedástica.

- **Contraste de Hausman.** Comparar $\hat{\beta}_{MG}$ y $\hat{\beta}_{FE}$. Para poder usar este contraste se necesita:
  (a) Bajo la hipótesis nula, FE es más eficiente. ¿Por qué? Porque estima sólo un parámetro en ves de N.
  (b) Bajo la alternativa $\hat{\beta}_{MG} - \hat{\beta}_{FE}$ debería tender a un vector no cero.

Contrastes para homogeneidad vs. heterogeneidad

- **Contraste de Swamy (1970).**

$$\hat{S} = \sum_{i=1}^{N} (\hat{\beta}_i - \hat{\beta}_{WFE})' \frac{X_i' Q_T X_i}{\hat{\sigma}_i^2} (\hat{\beta}_i - \hat{\beta}_{WFE})$$

donde $\hat{\sigma}_i^2 = \frac{1}{T-K-1}(y_i - X_i\hat{\beta}_{WFE})' Q_T (y_i - X_i\hat{\beta}_{WFE})$ y
$\hat{\beta}_{WFE} = (\sum_{i=1}^{N}(X_i' Q_T X_i)/\hat{\sigma}_i^2)^{-1} \sum_{i=1}^{N}(X_i' Q_T y_i)/\hat{\sigma}_i^2$ donde $Q = I_T - \bar{J}_T$.
Cuando N está fijo y $T \to \infty$, bajo la hipótesis nula, $\hat{S} \xrightarrow{d} \chi^2_{K(N-1)}$.
La intuición de este contraste es que nosotros no conocemos el valor de β bajo la hipótesis nula. Sin embargo, si este fuera uno se podría estimar con FE. Entonces podríamos compara todas las estimaciones de β_i con esta referencia. Si esa diferencia no es muy grande, entonces podríamos aceptar la nula.

- **Contraste de Pesaran y Yamagata (2008).**
Este contraste parte del de Swamy, pero estandariza el estadístico $\hat{S}$ usando su distribución chi cuadrado.
Reemplazar $\hat{\sigma}_i^2$ por $\tilde{\sigma}_i^2 = \frac{1}{T-1}(y_i - X_i\hat{\beta}_{FE})' Q_T (y_i - X_i\hat{\beta}_{FE})$ y usar

$$\Delta = \sqrt{N}\left(\frac{N^{-1}S - K}{\sqrt{2K}}\right)$$

Este contraste funciona para $N, T \to \infty$.