Lección 3.
Regresión con constantes diferentes.

Teodosio Pérez Amaral,
Universidad Complutense
Telefónica, mayo 2005

1. Introducción

- Si se rechaza la homogeneidad, una forma simple y útil de recoger la heterogeneidad es usar un modelo con constantes diferentes para cada individuo.

- Constantes diferentes para cada individuo recogen (condicionales en los valores de las explicativas):
  - A. Variables individuales que no varían en el tiempo, (raza, origen)
  - B. Variables temporales que no varían por individuos, (tipos interés, precios)
  - C. Variables que varían por individuos y tiempo, (ventas, beneficios).

- Los modelos de constantes diferentes recogen la heterogeneidad con diferentes constantes que pueden representar:
- A. Las numerosas variables omitidas que varían a lo largo del tiempo que son individualmente poco importantes, pero colectivamente significativas.

- B. El efecto de las variables omitidas que no varían en el tiempo y
- C. El efecto de otras variables que no varían en el corte transversal,

2. Modelos de efectos fijos: mínimos cuadrados con variables ficticias.

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

- α_i^* es un escalar
- β es un vector $1 \times k$ de constantes desconocidas que no varían en i y t.

- $t=1, \ldots, T$; tiempo
- $i=1, \ldots, N$; individuos
- $k=1, \ldots, K$; variables explicativas

- Se supone que no hay correlación entre las x_i y los errores u_{it}.
- $u_{it} \sim niid(0, \sigma^2)$

- Modelo de análisis de covarianza: analiza los factores cuantitativos y cualitativos.

- Usando una ficticia de constante para cada individuo, podemos estimar por mínimos cuadrados ordinarios, MCO, que es ELIO, estimador insesgado y óptimo.

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right]^{-1} \left[\sum_{i=1}^{N} \sum_{t=1}^{T} (x_{it} - \bar{x}_i)(y_{it} - \bar{y}_i)' \right]$$

$$\hat{\alpha}_i^* = \bar{y}_i - \beta' \, \bar{x}, \qquad i = 1, \ldots, N.$$

- Sea la matriz idempotente de transformación de covarianzas Q

$$Q = I_T - \frac{1}{T} ee'$$

- Que elimina los efectos individuales, midiendo las variables en desviaciones con respecto a su media muestral.

- Aplicando MCO al modelo en desviaciones con respecto a las medias individuales.

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^{N} X_i' Q X_i\right]^{-1} \left[\sum_{i=1}^{N} X_i' Q y_i\right]$$

Propiedades estimador covarianzas

- Existencia.
- Unicidad.
- Insesgadez $E(\hat{\beta}_{CV}) = \beta$
- Consistencia, $\hat{\beta}_{CV} \to \beta$ cuando N, T o ambos tienden a infinito.
- Eficiencia. Teorema de Gauss-Markov.

+Propiedades

- Matriz de varianzas y covarianzas

$$Var(\hat{\beta}_{CV}) = \sigma_u^{\ 2} \left[\sum_{i=1}^{N} X_i' Q X_i \right]^{-1}$$

- Estimador efectos fijos, insesgado

$$E(\hat{\alpha}^*_{\ i}) = \alpha_i$$

- Estimador efectos fijos, solo consistente si T→∞. $$p\lim(\hat{\alpha}^*_{\ i}) = \alpha_i$$

3. Modelo de efectos aleatorios. Modelos de componentes del error.

- Suponemos que el error ν_{it} tiene tres componentes

$$\nu_{it} = \alpha_i + \lambda_t + u_{it}$$

$$E(\alpha_i) = E(\lambda_t) = E(u_{it}) = 0$$

$$E(\alpha_i \lambda_t) = E(\alpha_i u_{it}) = E(\lambda_t u_{it}) = 0$$

- Covarianzas de los componentes del error iguales a cero.
- Y ausencia de correlación entre los regresores y los componentes del error.
- $$E(\alpha_i x'_{it}) = E(\lambda_t x'_{it}) = E(u_{it} x'_{it}) = 0'$$

- La varianza de y_{it} condicional en las x_{it}.

$$\sigma^2{}_y = \sigma^2{}_\alpha + \sigma^2{}_\lambda + \sigma^2{}_u$$

- Y la ecuación

$$\underset{T\times 1}{y_i} = \underset{T\times(k+1)}{\widetilde{X}_i} \; \underset{(k+1)\times 1}{\delta} + \underset{T\times 1}{\nu_i}$$

- V es la matriz de varianzas y covarianzas de los errores v_i

$$Ev_iv_i = \sigma_u^{\;2} I_T + \sigma_\alpha^{\;2} ee' = V$$

$$V^{-1} = \frac{1}{\sigma_u^{\;2}}\left[I_T - \frac{\sigma_\alpha^{\;2}}{\sigma_u^{\;2} + T\sigma_\alpha^{\;2}} ee' \right]$$

3.1 Estimación por covarianzas

- El α_i produce correlación serial entre los errores del mismo individuo.
- Pero se puede eliminar con la transformación Q.
- Sean fijos o aleatorios los efectos, los β se pueden estimar por CV y el estimador es insesgado y consistente con N, T o ambos tendiendo a infinito.

- Pero $\hat{\beta}_{CV}$ es ELIO bajo efectos fijos, pero no lo es bajo efectos aleatorios.
- En caso de efectos aleatorios sería óptimo (más eficiente, más preciso) el estimador de Mínimos Cuadrados Generalizados.

3.2 Estimador entre grupos (between, corte transversal)

- Un estimador alternativo que usa solamente la variación entre grupos

$$\hat{\beta}_b = \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1} \left[\sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y}) \right]$$

- 1. Se calcula la **media** muestral de cada una de las variables, endógena y explicativas de **cada individuo** del panel.
- 2. Se estima un modelo de regresión por MCO con datos de corte transversal. Con una observación por individuo.
- 3. El resultado es el estimador between, entregrupos. Usa solo la variación entre grupos.

3.1. Estimación por generalizados

$$\hat{\beta}_{GLS} = \left[\tfrac{1}{T}\sum_{i=1}^{N} X_i' Q X_i + \Psi \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})' \right]^{-1}$$

$$\left[\tfrac{1}{T}\sum_{i=1}^{N} X_i' Q y_i + \Psi \sum_{i=1}^{N} (\bar{x}_i - \bar{x})(\bar{y}_i - \bar{y}) \right]$$

$$\Psi = \frac{\sigma_u^2}{\sigma_u^2 + T\sigma_\alpha^2}$$

CARACTERISTICAS DE MC GENERALIZADOS.

- Aprovecha estructura matriz covarianzas errores.
- Combina la variación intragrupos y entregrupos de las variables.
- El estimador es una combinación de los estimadores intragrupos y entregrupos.

$$\hat{\beta}_{\mathrm{GLS}} = \Delta \hat{\beta}_{\mathrm{b}} + (I_k - \Delta)\hat{\beta}_{\mathrm{cv}}$$

- Δ son ponderaciones que dependen de los datos (Hsiao, p. 37).
- Si $\Psi \to 1,$ MCG converge a MCO, del modelo conjunto.
- Si $\Psi \to 0,$ MCG de β converge al intragrupos.

- Ψ es el peso que se da a la variación entre grupos.

- Efectos aleatorios es un intermedio entre tratar a todos los efectos iguales o todos diferentes.

$$Var(\hat{\beta}_{GLS}) = \sigma_u^{\,2}\left[\sum_{i=1}^{N} X_i' Q X_i + T\,\Psi \sum_{i=1}^{N}(\bar{x}_i - \bar{x})(\bar{x}_i - \bar{x})'\right]^{-1}$$

- La diferencia entre las matrices de varianzas y covarianzas de $\hat{\beta}_{CV}$ y $\hat{\beta}_{GLS}$ es positiva semidefinida.

- $\hat{\beta}_{GLS}$ más eficiente que $\hat{\beta}_{CV}$

- Estimación: por mínimos cuadrados generalizados en dos etapas.
- 1. Pseudo diferencias (1- $\psi^{1/2}$),
- 2. MCO.
- Como no se conocen, hay que estimar
- $$\hat{\sigma}_u^2 = \frac{\sum_{i=1}^{N}\sum_{t=1}^{T}\left[(y_{it} - \bar{y}_i) - \hat{\beta}_{CV}'(x_{it} - \bar{x}_i)\right]^2}{N(T-1) - K}$$

- Y usando el estimador entregrupos, between, podemos estimar

$$\hat{\sigma}_\alpha^2 = \frac{\sum_{i=1}^{N}(\bar{y}_i - \tilde{\mu} - \tilde{\beta}'\bar{x}_i)^2}{N-(K+1)} - \frac{1}{T}\hat{\sigma}_u^2$$

3.2 Estimación por máxima verosimilitud del modelo de efectos aleatorios.

- Forma de estimar tradicional.
- Especificar distribución.
- Si suponemos normalidad de α_i y $u_{it.}$

$$\log L = -\frac{NT}{2}\log 2\pi - \frac{N}{2}\log|V| - \frac{1}{2}\sum_{i=1}^{N}(y_i - e\mu\ X_i\beta)'\,V^{-1}(y_i - e\mu - X_i\beta)$$

$$= -\frac{NT}{2}\log 2\pi - \frac{N(T-1)}{2}\log\sigma_u^2 - \frac{N}{2}\log(\sigma_u^2 + T\sigma_\alpha^2)$$

$$-\frac{1}{2\sigma_u^2}\sum_{i=1}^{N}(y_i - e\mu - X_i\beta)'\,Q\,(y_i - e\mu - X_i\beta)$$

$$-\frac{T}{2(\sigma_u^2 + T\sigma_\alpha^2)}\sum_{i=1}^{N}(\bar{y}_i - \mu - \beta'\bar{x}_i)'^2$$

- Y derivando LogL con respecto a: β, μ, σ^2_u y σ^2_α

- Obtenemos las ecuaciones normales, que generan las condiciones de ortogonalidad que definen el estimador de Máxima verosimilitud.

- Estimación simultánea es complicada.
- Se puede usar procedimiento iterativo de Newton-Raphson.

$$\tilde{\delta}^{(j)} = \tilde{\delta}^{(j-1)} - \left[\frac{\partial^2 \log L}{\partial \tilde{\delta}\, \partial \tilde{\delta}'}\right]^{-1}_{\tilde{\delta} = \tilde{\delta}^{(j-1)}} \left.\frac{\partial \log L}{\partial \tilde{\delta}}\right|_{\tilde{\delta} = \tilde{\delta}^{(j-1)}}$$

- Que se repite hasta la convergencia.
- Estimador de máxima verosimilitud es consistente y asintóticamente normal,

$$\mathrm{Var}(\sqrt{N}\tilde{\delta}_{MLE}) = \quad NE\left[-\frac{\partial^2 \log L}{\partial \tilde{\delta}\, \partial \tilde{\delta}'}\right]^{-1}$$

$$= \begin{bmatrix} \dfrac{T}{\sigma^2} & \dfrac{T}{\sigma^2}\dfrac{1}{N}\sum_{i=1}^{N}\bar{x}_i' & 0 & 0 \\ & \dfrac{1}{\sigma^2_u}\dfrac{1}{N}\sum_{i=1}^{N} X_i'\left(I_t - \dfrac{\sigma_\alpha^2}{\sigma^2}ee'\right)X_i & 0 & 0 \\ & & \dfrac{T-1}{2\sigma_u^2}+\dfrac{1}{2\sigma^4} & \dfrac{T}{2\sigma^4} \\ & & & \dfrac{T^2}{2\sigma^4} \end{bmatrix}$$

3.4 Efectos fijos o aleatorios

- **Modelo de efectos fijos:** inferencia condicional en los efectos presentes en la muestra.
- **Modelos de efectos aleatorios:** inferencia incondicional.
- El investigador **decide** qué es lo más apropiado.

- Si T grande, estimación por efectos fijos y aleatorios, da resultados similares.

- Con **T finito y N grande**, la elección entre modelos no es fácil, y suele ser **grande la diferencia** entre efectos fijos y aleatorios.

Cuando es más apropiado?

- **Modelo de efectos fijos:** la muestra coincide con población, individuos importantes, queremos decir cosas de algunos individuos o grupos. Interés de la inferencia se ciñe solo a los efectos presentes en la muestra.
- **Modelo de efectos aleatorios:** muestra de de una población más amplia. La inferencia se quiere hacer con respecto a una población de efectos de los que los datos son solamente una muestra aleatoria.

- Una vez elegido el modelo, en la práctica suele haber **importantes diferencias** entre los dos modelos, de efectos fijos y aleatorios.
- Posible causa: error especificación.
- Ej: **Correlación entre x_{it} y los** α_i.

$$\hat{\beta}_{MCO} = \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\bar{x})(x_{it}-\bar{x})'\right]^{-1}\left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\bar{x})(y_{it}-\bar{y})\right]^{-1}$$

$$= \beta + \left[\sum_{i=1}^{N}\sum_{t=1}^{T}(x_{it}-\bar{x})(x_{it}-\bar{x})'\right]^{-1}\left\{T\sum_{i=1}^{N}(x_i-\bar{x})(\alpha_i^{*}-\bar{\alpha})\right\}$$

3.4.2.a Formulación de Mundlak

- **Critica** el uso del modelo de **efectos aleatorios** porque, con datos económicos, suele haber correlación entre regresores y efectos individuales.
- Y dice que la diferencia aparente entre efectos fijos y aleatorios se debe a un **error de especificación**.

- **Mundlak** afirma que si se corrige la correlación entre efectos y regresores, el estimador apropiado sería el de efectos fijos.
- **Comentario:** esto es cierto solamente para el modelo lineal, pero no para modelos dinámicos, coeficientes variables o elección discreta.

3.5 Contrastes de especificación.

- Hausman (1978)
- Ayuda a decidir entre **ESTIMAR** por MCO o por MCG un modelo de efectos aleatorios.
- H_0: No hay correlación entre ✂$_i$ y las x_i
- H_1: Hay correlación entre ✂$_i$ y las x_i

- H_0: $E(\alpha_i|X_i) = 0$,
  - MCG es consistente y eficiente.
  - El estimador de CV es consistente pero ineficiente.
- H_1: $E(\alpha_i|X_i) \neq 0$,
  - MCG es inconsistente.
  - El estimador de CV es consistente.

- Procedimiento práctico: mide la distancia entre las dos estimaciones: por MCO con variables ficticias(CV) y por MCG.

$$\hat{q} = \hat{\beta}_{CV} - \hat{\beta}_{GLS}$$

- Usando el famoso lema de Hausman.

$$Var(\hat{q}) = Var(\hat{\beta}_{CV}) - Var(\hat{\beta}_{GLS})$$

- Hausman propone el estadístico

$$m = \hat{q}' \, Var(\hat{q}')^{-1} \hat{q}$$

- Bajo la H_0 de no correlación, el estadístico se distribuye como una Ji-cuadrado con el mismo número de grados de libertad que parámetros que se contrastan.

- Bajo la alternativa, H_1 de correlación, el estadístico m se comporta asintóticamente como una Ji-cuadrado no centrada, con parámetro de no-centralidad, $\bar{q}'\ Var(\hat{q}')^{-1}\bar{q}$, donde

$$\bar{q} = p\lim(\hat{\beta}_{CV} - \hat{\beta}_{GLS})$$

3.6 Modelos con variables y efectos específicos individuales y temporales.

- Caso 1. Variables individuales especificas

$$y_i = e\mu + Z_i\gamma + X_{it}\beta + e\alpha_i + u_i$$

  - i=1, ..., N.
  - Donde Z_i = e z_i´
  - Si los efectos son fijos, tenemos multicolinealidad perfecta.

- No se pueden estimar conjuntamente γ,
- μ, y α_i.
- Pero si podemos estimar β por CV
- Y para estimar γ podemos escribir

$$\bar{y}_i - \bar{x}_i'\beta = \mu + z_i' \gamma + \alpha_i + \bar{u}_i$$

- Estimamos γ por MCO.

$$\hat{\gamma} = \left[\sum_{i=1}^{N} (z_i - \bar{z})(z_i - \bar{z})' \right]^{-1}$$

$$\left[\sum_{i=1}^{N} (z_i - \bar{z})\left[(\bar{y}_i - \bar{y}) - (\bar{x}_i - \bar{x})'\beta\right] \right]$$

$$\hat{\mu} = \bar{y} - \bar{x}'\beta - \bar{z}'\hat{\gamma}_i$$

- Estimación.
- Estimamos β por CV.
- Lo sustituimos en las dos fórmulas siguientes.
- Estimamos γ y μ.
- Propiedades.
- β y γ son consistentes cuando $N \rightarrow \infty$.

- Cuando los efectos α_i son aleatorios e incorrelacionados con X_i y Z_i, el estimador CV no es ELIO.
- El estimador ELIO sería el de MCG. Ver Hsiao, p 53.

- Una interpretación de este modelo es que se incluyen variables que no varían a lo largo del tiempo, Z_i para "explicar" los efectos individuales, y esto permite disminuir o eliminar las correlaciones entre los efectos individuales y los regresores.

Caso 2. Modelos con efectos individuales y temporales.

$$y_{it} = \mu + z_i' \gamma + r_t' \rho + x_{it} \beta + \alpha_i + \lambda_t + u_{it}$$

- Si efectos fijos, no se pueden estimar todos los parámetros.
- Si aleatorios, se usa un estimador en varios pasos como el anterior.
- Si aleatorios e incorrelacionados con r, z y x, se usarían MCG, que es ELIO.

3.7 Heteroscedasticidad

- Muchos paneles tienen unidades de tamaño muy diverso. Puede dar lugar a heteroscedasticidad.

- En el caso de componentes del error, la heteroscedasticidad se puede dar porque la varianza, $\sigma^2_{\alpha i}$ de los α_i varía con i, o bien porque la varianza, σ^2_{ui} de los u_{it} varía con i, o ambos.

- Problema complejo por el número de parámetros. Haciendo simplificaciones.

$$\hat{\sigma}^2_{ui} = \frac{1}{T-1}\sum_{i=1}^{T}(\hat{v}_{it} - \bar{\hat{v}}_{i})^2$$

El estimador de mínimos cuadrados generalizados factibles

$$\hat{\delta}_{FGLS} = \left[\sum_{i=1}^{N} \tilde{X}_i' \hat{\Omega}_i^{-1} \tilde{X}_i\right]^{-1} \left[\sum_{i=1}^{N} \tilde{X}_i' \hat{\Omega}_i^{-1} y_i\right]$$

- Otra alternativa en caso de que haya heteroscedasticidad a lo largo de las unidades y de los errores comunes sería tratar los efectos como fijos y luego aplicar mínimos cuadrados generalizados factibles.
- Primero se divide cada observación por su desviación típica estimada y luego

- Se estima por covarianzas usando datos transformados.

$$\hat{\beta}_{CV} = \left[\sum_{i=1}^{N} X_i^{*\prime} Q X_i^{*}\right]^{-1}\left[\sum_{i=1}^{N} X_i^{*\prime} Q y_i^{*}\right]$$

- Los estimadores serían ELIO