

Modelos y datos de panel

Econometría Aplicada y Ciencia de Datos - CIDE Fall 2026

Carlos Brito

- Consideremos el siguiente modelo para unidades indexadas como i observadas en varios periodos indexados con t

$$y_{it} = \alpha_{it} + x'_{it}\beta_{it} + u_{it}$$

con $i = 1, \dots, N$ y $t = 1, \dots, T$

- Este modelo así escrito es muy general pues tiene un intercepto y una pendiente para cada i y t

- Consideremos el siguiente modelo para unidades indexadas como i observadas en varios periodos indexados con t

$$y_{it} = \alpha_{it} + x'_{it}\beta_{it} + u_{it}$$

con $i = 1, \dots, N$ y $t = 1, \dots, T$

- Este modelo así escrito es muy general pues tiene un intercepto y una pendiente para cada i y t
- Sin embargo, este modelo es imposible de estimar:

Estructura de datos en panel

- Una forma de resolver el problema es formulando un *modelo de efectos específicos por individuo*

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}$$

donde ε_{it} es iid sobre i y t

- Las α_i las pensamos como variables aleatorias que capturan la heterogeneidad
- Para propósitos de estimación, recurriremos a un supuesto clave:

Exogeneidad fuerte o estricta: asumimos que el error tiene media cero condicional cuando condicionamos en los valores contemporáneos, pasados y futuros

- La exogeneidad fuerte descarta modelos donde y_{it} depende de sus rezagos o donde alguna de las x_{it} es endógena
- La forma en que tratemos a α resultará en los dos tipos de modelos principales de esta clase

- Si asumimos que los efectos individuales están potencialmente correlacionados con x_{it} , nuestro modelo se conoce como de **efectos fijos**
- α_i es la heterogeneidad no observada
- Con paneles cortos no podemos estimar de forma consistente las α_i
- Recurriremos a métodos para deshacernos de las α_i

Modelo de efectos aleatorios

- Por otro lado, si asumimos que la heterogeneidad no observada es una variable aleatoria independiente de x_{it} , podemos hacer algunos supuestos sobre su distribución y especificar un modelo de **efectos aleatorios**
- Suponemos que tanto la heterogeneidad no observada como el error son iid
- Podemos ver al modelo como una transformación del original con un error compuesto $u_{it} = \alpha_i + \varepsilon_{it}$:

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

- Podemos mostrar que la correlación de los errores compuestos de dos observaciones en distintos momentos del tiempo es $cor(u_{it}, u_{is}) = \frac{\sigma_{\alpha}^2}{\sigma_{\alpha}^2 + \sigma_{\varepsilon}^2}$
- También se conoce como **modelo equicorrelacionado**

- Quizás un mejor nombre para el modelo de efectos aleatorios es el a veces usado modelo de **intercepto aleatorio** o de **componentes aleatorios**
- Para el modelo de efectos fijos, Lee (2002) usa el nombre de **efecto relacionado** y para el de efectos aleatorios el de **efecto no relacionado**
- Lo más importante es tener en cuenta que α_i es una variable aleatoria, lo que cambia es lo que asumimos sobre su correlación con los regresores

Estimadores para el modelo de efectos fijos

- Una desventaja: no se pueden identificar los parámetros sobre regresores que son fijos en el tiempo (sexo, IQ, lugar de nacimiento, raza, ...)

- Una desventaja: no se pueden identificar los parámetros sobre regresores que son fijos en el tiempo (sexo, IQ, lugar de nacimiento, raza, ...)
- Frecuentemente no estamos interesados en estimar los efectos fijos, es decir, los consideramos **parámetros incómodos**

- Una desventaja: no se pueden identificar los parámetros sobre regresores que son fijos en el tiempo (sexo, IQ, lugar de nacimiento, raza, ...)
- Frecuentemente no estamos interesados en estimar los efectos fijos, es decir, los consideramos **parámetros incómodos**
- Si quisiéramos estimarlos directamente, podríamos incluir dummies individuales
- Se requieren paneles largos para que los efectos estimados sean consistentes
- En el resto del curso nos enfocaremos en casos donde podemos ignorar estos parámetros incómodos

- Consideremos el modelo
- El estimador de primeras diferencias es el estimador de MCO de esta ecuación

- El estimador de primeras diferencias es consistente
- El estimador de primeras diferencias y el within son iguales cuando $T = 2$
- Con $T > 2$ el estimador de primeras diferencias es menos eficiente
- Los paquetes estadísticos usan un estimador within

Propiedades del estimador within

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}$$

$$\bar{y}_i = \alpha_i + \bar{x}_i\beta + \bar{\varepsilon}_i$$

- Y luego restando del modelo original estas medias (sobre i , $\bar{y}_i = \frac{1}{T} \sum_i y_{it}$)

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

- El estimador within puede escribirse como

- La consistencia del estimador depende de que

$$p \lim \frac{1}{NT} \sum_i \sum_t (x_{it} - \bar{x}_i)(\varepsilon_{it} - \bar{\varepsilon}_i) = 0 \text{ cuando } N \rightarrow \infty \text{ o } T \rightarrow \infty$$

Propiedades del estimador within

$$y_{it} = \alpha_i + x'_{it}\beta + \varepsilon_{it}$$

$$\bar{y}_i = \alpha_i + \bar{x}_i\beta + \bar{\varepsilon}_i$$

- Y luego restando del modelo original estas medias (sobre i , $\bar{y}_i = \frac{1}{T} \sum_i y_{it}$)

$$y_{it} - \bar{y}_i = (x_{it} - \bar{x}_i)'\beta + (\varepsilon_{it} - \bar{\varepsilon}_i)$$

- El estimador within puede escribirse como

- La consistencia del estimador depende de que $p \lim \frac{1}{NT} \sum_i \sum_t (x_{it} - \bar{x}_i)(\varepsilon_{it} - \bar{\varepsilon}_i) = 0$ cuando $N \rightarrow \infty$ o $T \rightarrow \infty$
- Frecuentemente tenemos pocos periodos (páneos cortos), por lo que esperamos que haya muchos individuos para probar consistencia
- Exogeneidad fuerte ($E(\varepsilon_{it}|\alpha_i, x_{i1}, \dots, x_{iT}) = 0$) es suficiente para consistencia

- Podemos asumir que los errores ε_{it} son iid y obtener

$$\begin{aligned}\hat{V}(\hat{\beta}_{W,h}) &= \hat{\sigma}_\varepsilon^2 \left(\sum_i \sum_t (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \\ &= \hat{\sigma}_\varepsilon^2 \left(\sum_i \sum_t \ddot{x}_{it} \ddot{x}_{it}' \right)^{-1}\end{aligned}$$

donde $\ddot{x}_{it} = x_{it} - \bar{x}_i$ y se puede estimar $\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1)-K} \sum_i \sum_t \hat{\varepsilon}_{it}^2$

- Podemos asumir que los errores ε_{it} son iid y obtener

$$\begin{aligned}\hat{V}(\hat{\beta}_{W,h}) &= \hat{\sigma}_\varepsilon^2 \left(\sum_i \sum_t (x_{it} - \bar{x}_i)(x_{it} - \bar{x}_i)' \right)^{-1} \\ &= \hat{\sigma}_\varepsilon^2 \left(\sum_i \sum_t \ddot{x}_{it} \ddot{x}_{it}' \right)^{-1}\end{aligned}$$

donde $\ddot{x}_{it} = x_{it} - \bar{x}_i$ y se puede estimar $\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1)-K} \sum_i \sum_t \hat{\varepsilon}_{it}^2$

- Sin embargo, en panel esto es un supuesto fuerte pues esperamos que, aunque las observaciones sean independientes entre individuos, haya correlación serial

- Generalmente usamos una versión robusta, que considera la correlación serial y permite heterocedasticidad

$$\hat{V}(\hat{\beta}_W) = \left(\sum_i \sum_t \ddot{x}_{it} \ddot{x}'_{it} \right)^{-1} \left(\sum_i \sum_t \sum_s \ddot{x}_{it} \ddot{x}'_{is} \hat{e}_{it} \hat{e}_{is} \right) \left(\sum_i \sum_t \ddot{x}_{it} \ddot{x}'_{it} \right)^{-1}$$

- Esta versión con la estructura de la matriz de White es robusta a la heterocedasticidad y a la correlación serial
- Es lo que se conoce como matriz de varianzas agrupada

Estimador between

Estimador between

- Aunque es de poca utilidad práctica, el estimador *between* nos servirá de auxiliar más adelante
 - Explota la variación entre unidades
 - Al promediar sobre el **tiempo** recordemos que obtenemos
-
- El estimador between es MCO a la ecuación transformada (con intercepto)
 - Solo es consistente si $\bar{x}_i$ es independiente del error compuesto $\alpha_i - \alpha + \bar{\varepsilon}_i$
 - Es equivalente a una regresión de sección cruzada con $T = 1$

Mínimos cuadrados generalizados

Breve desviación a mínimos cuadrados generalizados

- En ECNI aprendimos que el estimador de MCO es el estimador de varianza mínima entre los estimadores lineales insesgados cuando los errores son iid
- Ahora consideremos el modelo $y = X\beta + u$
- Asumamos que la varianza de los errores es $\Omega \neq \sigma^2 I$
- Asumamos que conocemos Ω , una matriz no singular
- Si premultiplicamos el modelo lineal por $\Omega^{-1/2}$

se puede mostrar con algo de álgebra que

Breve desviación a mínimos cuadrados generalizados

- Es decir, los errores del modelo transformado son iid
- Así que podemos estimar β eficientemente por MCO al modelo transformado
- El estimador **mínimos cuadrados generalizados** (MCG) es

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

Breve desviación a mínimos cuadrados generalizados

- Es decir, los errores del modelo transformado son iid
- Así que podemos estimar β eficientemente por MCO al modelo transformado
- El estimador **mínimos cuadrados generalizados** (MCG) es

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

- En la práctica, no conocemos Ω
- Podemos proponer un modelo para $V(u|x)$ y obtener un estimador consistente para Ω
- Esto da lugar al estimador de **mínimos cuadrados generalizados factibles**

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

Estimador para modelo de efectos aleatorios

- Podemos escribir el modelo general añadiendo un intercepto no aleatorio μ para que los efectos aleatorios puedan tener media cero:

$$y_{it} = \mu + x'_{it}\beta + \alpha_i + \varepsilon_{it} = w'_{it}\delta + \alpha_i + \varepsilon_{it}$$

donde $w_{it} = [1 \ x_{it}]$ y $\delta = [\mu \ \beta']'$

- Asumimos que α_i y ε_i son iid
 - $\alpha_i \sim (\alpha, \sigma_\alpha^2)$
 - $\varepsilon_{it} \sim (0, \sigma_\varepsilon^2)$
- μ simplemente garantiza que los errores tengan media cero.

- El estimador de efectos aleatorios se logra premultiplicando el modelo $y_{it} = w'_{it}\delta + \alpha_i + \varepsilon_{it}$ por $\Omega^{-1/2}$ y con algo de álgebra
- Se asume que α_i y ε_{it} son independientes
- Hacemos Ω_i con entradas $\sigma_\alpha^2 + \sigma_\varepsilon^2$ en la diagonal principal
- Fuera de la diagonal tiene σ_α^2

$$\Omega = \sigma_\varepsilon^2 I_T + \sigma_\alpha^2 ee'$$

donde $e = (1, 1, \dots, 1)'$ es un vector de unos de dimensión $T \times 1$

- Premultiplicamos entonces el modelo con efectos aleatorios para obtener uno que tenga como matriz de varianzas $\sigma_\varepsilon^2 I_T$, es decir, que sea homocedástico

Con algunas cuentas se puede demostrar que:

$$\Omega^{-1/2} y_{it} = y_{it} - \lambda \bar{y}_i$$
$$\lambda = 1 - \frac{\sigma_\varepsilon}{\sqrt{T\sigma_\alpha^2 + \sigma_\varepsilon^2}}$$

- El estimador de MCG de μ y β puede implementarse como MCO a la ecuación transformada

$$y_{it} - \lambda \bar{y}_i = (1 - \lambda)\mu + (x_{it} - \lambda \bar{x}_i)' \beta + \nu_{it}$$

donde $\nu_{it} = (1 - \lambda)\alpha_i + (\varepsilon_{it} - \lambda \bar{\varepsilon}_i)$

- Construimos un estimador consistente de λ usando estimadores consistentes de σ_ε^2 y σ_α^2

$$\hat{\lambda} = 1 - \frac{\hat{\sigma}_\varepsilon}{\sqrt{T\hat{\sigma}_\alpha^2 + \hat{\sigma}_\varepsilon^2}}$$

Estimación de $\hat{\lambda}$

- El software estadístico hace estas estimaciones basadas en:
- Para σ_ε^2 usamos el estimador within $\hat{\beta}_W$ para obtener:

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1) - k} \sum_i \sum_t \left((y_{it} - \bar{y}_i) - (x_{it} - \bar{x}_i)' \hat{\beta}_W \right)^2$$

- Para σ_α^2 , usamos el estimador between $\hat{\beta}_B$ para obtener

$$\hat{\sigma}_\alpha^2 = \frac{1}{N - (k-1)} \sum_i \left(\bar{y}_i - \hat{\alpha}_B - \bar{x}_i' \hat{\beta}_B \right)^2 - \frac{1}{T} \hat{\sigma}_\varepsilon^2$$

Propiedades del estimador de efectos aleatorios

- El estimador de efectos es consistente si $NT \rightarrow \infty$, lo cual requiere que $N \rightarrow \infty$ o $T \rightarrow \infty$ (o ambos)
- Con paneles cortos, es natural requerir suficientes individuos en la muestra
- Por otro lado, CT proveen dos expresiones para las matrices de varianza en los casos de homocedasticidad (es decir, si ε_{it} y α_i son iid) y para el caso general donde se permite heterocedasticidad y autocorrelación de ε_{it}

Modelo agrupado o *pooled*

Modelo agrupado o *pooled*

- Si asumimos coeficientes constantes, como en sección cruzada

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

Tenemos lo que se conoce como modelo pooled o de **coeficientes constantes**

- Este modelo no explota la estructura del panel
- El modelo es correcto si los regresores no están correlacionados con el error
- Podemos estimar consistentemente β por MCO. Pero la matriz de varianza habitual de MCO se basa en errores i.i.d.
- Los errores estándar deben tomar en cuenta la correlación serial: las observaciones correlacionadas NT contienen menos información que las observaciones independientes NT.
- Si estamos dispuestos a asumir un modelo pooled, debemos al menos estimar un modelo de MCO con errores robustos agrupados **a nivel individual**

¿Efectos fijos o efectos aleatorios?

¿Qué estimador usar?

- Si el modelo asumido es de efectos fijos:
 - Solo el estimador within o de primeras diferencias es consistente
 - Estimadores de efectos aleatorios y pooled son inconsistentes

¿Qué estimador usar?

- Si el modelo asumido es de efectos fijos:
 - Solo el estimador within o de primeras diferencias es consistente
 - Estimadores de efectos aleatorios y pooled son inconsistentes
- Si el modelo asumido es de efectos aleatorios
 - El estimador de efectos aleatorios y el estimador pooled son consistentes, pero el estimador de efectos aleatorios es más eficiente

Prueba de Hausman

Prueba de Hausman

- Una prueba de Hausman compara dos estimadores, uno consistente bajo la H_0 y otro que no lo es
- Si los efectos individuales son efectos fijos correlacionados con el error, el estimador within es consistente y el de efectos aleatorios no lo es
- La H_0 es que los efectos fijos no están correlacionados con el error
- Si se rechaza la H_0 , existe evidencia en favor de usar un estimador de efectos fijos

H_0 : RE es totalmente eficiente.

$$H = \left(\hat{\beta}_{1,RE} - \hat{\beta}_{1,W} \right)' \left[\hat{V}[\hat{\beta}_{1,W}] - \hat{V}[\hat{\beta}_{1,RE}] \right]^{-1} \left(\hat{\beta}_{1,RE} - \hat{\beta}_{1,W} \right)$$

Donde β_1 denota el subvector de coeficientes de los regresores que varían en el tiempo. Este estadístico se distribuye asintóticamente $\chi^2(\dim[\beta_1])$ bajo la hipótesis nula.