

Datos en panel I

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 las dos anteriores.

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

1. **Cortes transversales independientes**
2. **Muestra longitudinal**

Cortes transversales independientes

- La estructura de datos en panel que es más frecuente son los cortes transversales independientes.
- Ej: Muchas encuestas de hogares no siguen los mismo hogares a lo largo del tiempo, estos cambian cada vez. La EPH contiene hogares que se renuevan cada vez y otros a los que se los sigue repetidamente.
- Supongamos dos periodos $t = 0, 1$, con n_0 individuos en $t = 0$ y n_1 en $t = 1$ (entonces tenemos un total de $n = n_0 + n_1$). Los datos son $\{y_{i0}, x_{i0}\}_{i=1}^{n_0}$, $\{y_{j1}, x_{j1}\}_{j=1}^{n_1}$

Supongamos que corremos una regresión simple:

$$y_{ht} = \beta_0 + \beta_1 x_{ht} + u_{ht}, \quad h = 1, 2, \dots, (n_0 + n_1), \quad t = 0, 1$$

¿Qué mide β_1 ?

Cortes transversales independientes

Supongamos que tanto y como x son mayores en $t = 1$ que en $t = 0$. ¡Entonces con una regresión simple estaríamos capturando el efecto del tiempo como un efecto de x en y !

Para ver esto definamos

$$u_{ht} = a_t + e_{ht}$$

donde a_t mide el efecto del tiempo mientras que e es un verdadero error. Entonces, $E(xu) = E(xa) + E(xe) = E(xa) \neq 0$.

⇒ ¿Cómo controlaríamos por el efecto de tiempo? Con variables dummy.

Cortes transversales independientes

Consideremos el siguiente modelo

$$y_{ht} = \beta_0 + \beta_1 x_{ht} + \delta_0 a_t + \delta_1 x_{ht} \times a_t + e_{ht},$$
$$h = 1, 2, \dots, (n_0 + n_1), t = 0, 1$$

donde a_t es una variable dummy que toma valor 0 en $t=0$ y 1 en $t=1$.

- En este modelo no solo tenemos un efecto del tiempo sobre y , sino también de tiempo en la relación misma de x con y .
- El intercepto para $t = 0$ es β_0 pero para $t = 1$ es $\beta_0 + \delta_0$.
- La pendiente (efecto de un cambio de 1 unidad de x en y) para $t = 0$ es β_1 pero para $t = 1$ es $\beta_1 + \delta_1$.
- ¿Cómo contrastaría que la pendiente en $t = 0$ es la misma que en $t = 1$?

Estimador de diferencia en diferencias (d-en-d)

Esta metodología es muy común para establecer relaciones causales.

Tratamiento: Lo que queremos evaluar.

Grupo de control: Grupo de individuos que no tienen el tratamiento.

Grupo de tratamiento: Grupo de individuos que se ve afectado por el tratamiento.

$$y_{it} = \beta_0 + \beta_1 TREAT_i + \delta_0 AFTER_t + \delta_1 TREAT_i \times AFTER_t + u_{it}$$

	Antes	Después	Antes-Después
Control	β_0	$\beta_0 + \delta_0$	δ_0
Tratamiento	$\beta_0 + \beta_1$	$\beta_0 + \delta_0 + \beta_1 + \delta_1$	$\delta_0 + \delta_1$
Tratamiento-Control	β_1	$\beta_1 + \delta_1$	δ_1

δ_1 es el estimador **d-en-d**.

Efecto de la ubicación de un incinerador de basura en los precios de las casas

- Ejemplo tomado de Kiel y McClain (1995) sobre el efecto de un incinerador de basura en los precios de las casas en North Andover, Massachusetts, para los años 1978 y 1981.
- La base de datos se puede acceder en:
use <http://fmwww.bc.edu/ec-p/data/wooldridge/KIELMC>, clear
- Las variables son
rprice: precio de las casas en términos reales.
nearinc: variable dummy para cercanía al incinerador.
- Los rumores de que el incinerador se iba a construir empezaron **después** de 1978, su construcción empezó en 1981, y empezó a operar en 1985.
- Un estimador simple usaría sólo datos de 1981:

$$\begin{aligned}
 rprice &= \gamma_0 + \gamma_1 \text{nearinc} + u \\
 \widehat{rprice} &= 101308 - 30688 \text{nearinc} \\
 &\quad (3093) \quad (5828)
 \end{aligned}$$

- ¿Es correcto afirmar que vivir cerca del incinerador causa que los precios disminuyan 30 mil dólares?

Efecto de la ubicación de un incinerador de basura en los precios de las casas

- El estimador previo no implica que el incinerador **cause** una disminución del precio de las casas. De hecho la misma regresión para 1978 da:

$$rprice = 82517 - 18824 \text{ nearinc}$$

$$(2654) \quad (4744)$$

- ¿Por qué? Porque el incinerador se construyó donde el precio de las casas ya era bajo.
- Estimador d-en-d:**

$$\hat{\delta}_1 = (\widehat{rprice}_{81,nr} - \widehat{rprice}_{81,fr}) - (\widehat{rprice}_{78,nr} - \widehat{rprice}_{78,fr})$$

donde *nr* significa cerca del incinerador y *fr* sin (free) incinerador.

- Usando el modelo general:

$$\widehat{rprice} = \beta_0 + \beta_1 \text{ nearinc} + \delta_0 \text{ y81} + \delta_1 \text{ nearinc*y81} + u$$

$$\widehat{rprice} = 82517 - 18824 \text{ nearinc} + 18790 \text{ y81} - 11864$$

$$(2727) \quad (4875) \quad (4050) \quad (7466)$$

Card y Krueger (1994)

Estudio de cambios en el salario mínimo sobre empleo

David Card y Alan B. Krueger (1994), "Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey," The American Economic Review 84(4), pp. 772-793. Ver una discusión avanzada en <http://econ.lse.ac.uk/staff/spischke/ec524/evaluation3.pdf>

Card y Krueger (1994)

- En noviembre de 1989 se aprobó una ley para incrementar el salario mínimo a nivel federal de \$3.35 por hora a \$3.80 (en abril 1990), \$4.25 por hora (en abril 1991), y al principio de 1990 se decidió incrementarlo a \$5.05 por hora (en abril de 1992).
- Sin embargo había diferencias a nivel estadual en cuanto al salario mínimo vigente (Pennsylvania y New Jersey)
- ¿Qué tipo de firmas iban a ser más afectadas? Aquellas que usaban trabajo no calificado (*unskilled labor*).
- Entonces estudiaron: Burger King, KFC, Wendy's, Roy Rogers
 - 1 son los empleadores líderes para trabajadores de bajo salario
 - 2 cumplen con la regulación de salarios mínimos
 - 3 los trabajos son relativamente homogéneos
 - 4 es relativamente fácil construir una muestra de franquicias de esas cadenas

Card y Krueger (1994)

- **Tratamiento:** incremento del salario mínimo efectivo
- **Grupo de control:** restaurantes en Pennsylvania
- **Grupo de tratamiento:** restaurantes en New Jersey
- Mapa of New Jersey
- **Antes:** Wave 1, February 15 - March 4, 1992
- **Después:** Wave 2, November 5 - December 31, 1992
- Efecto sobre el empleo: Los autores no encuentran ningún efecto significativo sobre el empleo.
- Impacto sobre otros beneficios no salariales: Podría ser que no hay efecto sobre el empleo porque los restaurantes reducen otros beneficios, o reducen training y las promociones. No hay evidencia.
- Impacto sobre precios: En un modelo competitivo incrementar el salario implica incrementar los precios. Tampoco hay evidencia de esto.