
Datos de Panel
Problem Set 2
Modelos de Datos de Panel Lineales

1. Utilice nuevamente la base de datos “*cornwell.dta*” provista para el Problem Set 1. Considere el siguiente modelo de regresión:

$$\ln crmrte_{it} = \beta_0 + \beta_1 \ln prbarr_{it} + \beta_2 \ln prbconv_{it} + \beta_3 \ln prbpris_{it} + \beta_4 \ln avgse_{it} \\ + \beta_5 \ln polpc_{it} + \sum_{\tau=82}^{87} \beta_{\tau} \cdot I\{t = \tau\} + \mu_i + \varepsilon_{it}$$

- Utilizando el comando *egen* de STATA, construya las medias individuales de las variables del modelo.
 - Aplique la transformación *within* al modelo. Luego, estime el modelo transformado por POLS.
 - Comente sobre la validez de los errores estándar del inciso previo.
 - Utilice el comando *xtreg* para estimar nuevamente el modelo usando efectos fijos.
 - Estime el modelo usando diferencias finitas de primer orden.
2. Utilice la base de datos provista “*murder.dta*”. La base de datos es una muestra longitudinal de estados de EE.UU., para los años 1987, 1990 y 1993.

- Estime por OLS el efecto de las ejecuciones (x) sobre la tasa de homicidios (*murder rates*, m) controlando por desempleo (u) y año:

$$m_{i,t} = \alpha + \beta_x x_{i,t} + \beta_u u_{i,t} + \beta_{90} d_{90,t} + \beta_{93} d_{93,t} + \nu_{i,t}$$

Note que se omitió la dummy temporal para el año 1987. Interprete los resultados.

- ¿Por qué podría ser importante tener en consideración los efectos temporales agregados en el modelo?
- Ahora, considere la siguiente modificación en el modelo:

$$m_{i,t} = \alpha + \beta_x x_{i,t} + \beta_u u_{i,t} + \beta_{90} d_{90,t} + \beta_{93} d_{93,t} + c_i + e_{i,t}$$

donde c_i es un efecto individual por estado. Estime la ecuación usando efectos fijos.

- Repita la estimación del inciso previo usando diferencias finitas de primer orden.
- Brinde un ejemplo bajo el cual la variable de ejecuciones no sería estrictamente exógena (condicional en c_i). **Observación.** Para obtener estimaciones consistentes, el modelo de efectos fijos asume exogeneidad estricta de las variables explicativas condicionadas en c_i .
- Repita la estimación del inciso c) usando el estimador de GLS para diferencias finitas de primer orden. Compruebe que los coeficientes estimados son iguales a los obtenidos por FE.
- Reestimar el modelo del inciso c) usando efectos aleatorios. Implementar el test de Hausman. ¿Cuál es el mejor estimador?

3. Considere el siguiente modelo:

$$y_{it} = x_{it}\beta + \mu_i + \nu_{it}, \quad i = 1, 2, \dots, N, \quad t = 1, 2, \dots, T$$

donde $x_{it} \stackrel{iid}{\sim} \mathcal{N}(0, 1)$, $\mu_i \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\mu^2)$, $\nu_{it} \stackrel{iid}{\sim} \mathcal{N}(0, \sigma_\nu^2)$ y $\mu_i \perp \nu_{it}$ para todo i, t . Suponga que $\beta = \sigma_\mu^2 = \sigma_\nu^2 = 1$ y $T = 10$. La idea es realizar experimentos de Monte Carlo para evaluar la eficiencia de distintos estimadores de β .

- a) Caso 1: $N = 5$. Realice un experimento de Monte Carlo con 1000 simulaciones. Reporte media, desvío estándar y RMSE de la estimación de β usando: POLS, RE y FE.
 - b) Repita el punto anterior con $N = 10, 30, 50, 100$ y 500.
 - c) Comente los resultados obtenidos y su conclusión de qué estimador debiera utilizarse en la práctica.
4. Basado en el Ejercicio 10.18 de Wooldridge (2010). Utilice la base de datos *wagepan.dta* para responder las preguntas a continuación.
- a) Utilizando *lwage* como variable dependiente, estimar un modelo que contenga un intercepto y las variables *dummy* de año *d81* a *d87*. Estime el modelo por POLS, RE, FE y FD. ¿Qué puede concluir acerca de los coeficientes de las variables *dummy*?
 - b) Añada las variables constantes en el tiempo *educ*, *black* e *hisp* al modelo, y estímelo por POLS y RE. ¿Cómo se comparan los coeficientes? ¿Qué ocurre si se estima la ecuación por FE?
 - c) ¿Son iguales los errores estándar de POLS y RE del inciso b)? ¿Cuáles son probablemente más fiables?
 - d) Obtenga los errores estándar robustos para POLS. ¿Prefiere estos o los errores estándar habituales de RE?
 - e) Obtenga los errores estándar robustos de RE. ¿Cómo se comparan con los errores estándar robustos de POLS, y por qué?