Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 0
Repaso OLS, GLS & FGLS
Lectura y resumen de datos de panel en Stata

---

1. Considere el siguiente modelo de regresión:

$$\begin{aligned} ltotexp_i = & \beta_0 + \beta_1 suppins_i + \beta_2 phylim_i + \beta_3 actlim_i + \beta_4 totchr_i + \beta_5 age_i + \\ & \beta_6 female_i + \beta_7 income_i + u_i, \quad i = 1, \ldots, N \end{aligned}$$

   a) Use la base de datos "mus03data.dta", la cual contiene datos de corte transversal de gastos médicos, para estimar la ecuación por OLS usando comandos de matrices en Stata. Adicionalmente, reporte los errores estándar usuales de OLS y los estadísticos t asociados.
   b) Utilice el comando regress para verificar los resultados obtenidos.
   c) Implemente un test de significatividad individual para totchr.
   d) Implemente un test de significatividad conjunta para todas las variables del modelo, excluyendo el intercepto.

Solution: *Inciso (a)* Tenemos que reconstruir las expresiones usuales de OLS:

$$\hat{\beta} = (X'X)^{-1}X'y$$
$$\hat{\Sigma}_{\hat{\beta}} = s^2(X'X)^{-1}$$
$$s.e.(\hat{\beta}_j) = \left(\hat{\Sigma}_{\hat{\beta}}(j,j)\right)^{\frac{1}{2}}$$
$$t_j = \hat{\beta}_j / s.e.(\hat{\beta}_j)$$

Opción 1: Usar el comando matrix accum

Opción 2: Usar el comando mkmat

2. En este ejercicio vamos a aprender cómo setear los datos como panel en Stata y cómo generar estadísticas descriptivas del panel. Adicionalmente, veremos cómo convertir los datos de *wide form* a *long form* y cómo generar un panel para simulaciones.

   a) Utilice la base mus08psidextract.dta y describa la base de datos de la manera usual y como un panel.
   b) Utilice la base pigweights.dta. Los datos se encuentran en formato *wide*. Utilice el comando reshape para llevarlos a formato *long*. Luego, describa la base de la misma forma que en el inciso (a).
   c) Genere un panel de 5000 observaciones con 10 períodos temporales y 500 unidades en el corte transversal. El panel debe estar en formato *long*. Genere observaciones de $x_{it} \sim \mathcal{N}(0,1)$, $u_{it} \sim \mathcal{N}(0,1)$ y además $y_{it} = 1 + x_{it} + u_{it}$. Estime por POLS.

Solution: Ver: *"PS 0 - Ex3.do"*.

3. Considere el siguiente modelo:

$$\begin{aligned} y_i &= \beta_1 + \beta_2 x_{2i} + \beta_3 x_{3i} + u_i, \quad i = 1, \ldots, N \\ u_i &= \sqrt{\exp\left(-1 + 0{,}2 \cdot x_{2i}\right)} \cdot \varepsilon_i, \quad i = 1, \ldots, N \end{aligned}$$

con $\beta_1 = 1$, $\beta_2 = 1$, $\beta_3 = 1$, $x_2 \sim \mathcal{N}(0, 25)$, $x_3 \sim \mathcal{N}(0, 25)$ y $\varepsilon \sim \mathcal{N}(0, 25)$ Luego, el error u es heterocedástico con una varianza condicional igual a $25 \cdot \exp(-1 + 0{,}2 \cdot x2)$.

a) Genere 1000 muestras de N=10 observaciones a partir del modelo presentado. Para cada muestra estime por OLS, GLS y FGLS los parámetros del modelo y realice un test de hipótesis para contrastar que $H_0 : \beta_3 =\ 1$. Reporte tamaño del test al 1 %. Adicionalmente, reporte la media, mediana y desvío estándar de las estimaciones de β_1, β_2 y β_3.

b) Repita el punto anterior con N igual a 20, 30, 100, 200 y 500.

c) Describa detalladamente las propiedades de muestra finita de FGLS de acuerdo a lo que observó de los puntos anteriores.

Solution: Sea $E\left[uu' \mid X\right] = \sigma_u^2 \Omega$, donde Ω es definida positiva, por lo que su inversa también lo es. Por lo tanto, es posible encontrar una matriz no singular P tal que:

$$\Omega^{-1} = P'P$$

En este caso particular, tenemos que:

$$\Omega^{-1} = \begin{bmatrix} \frac{1}{\exp(-1+0{,}2\cdot x_{21})} & 0 & \cdots & 0 \\ \vdots & \frac{1}{\exp(-1+0{,}2\cdot x_{22})} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & & \frac{1}{\exp(-1+0{,}2\cdot x_{2N})} \end{bmatrix}$$

Luego,

$$P = \begin{bmatrix} \frac{1}{\sqrt{\exp(-1+0{,}2\cdot x_{21})}} & 0 & \cdots & 0 \\ \vdots & \frac{1}{\sqrt{\exp(-1+0{,}2\cdot x_{22})}} & \cdots & \vdots \\ \vdots & \cdots & \ddots & \vdots \\ 0 & \cdots & & \frac{1}{\sqrt{\exp(-1+0{,}2\cdot x_{2N})}} \end{bmatrix}$$

Premultiplicamos al modelo por P obtenemos que:

$$y_* = X_* \beta + u_*$$

donde $y_* = Py$, $X_* = PX$ y $u_* = Pu$. Luego, el estimador de MCG consiste en estimar el modelo transformado por OLS:

$$\begin{aligned} \hat{\beta} &= \left(X_*' X_*\right)^{-1} X_*' y_* \\ &= \left(X' \Omega^{-1} X\right)^{-1} X' \Omega^{-1} y \end{aligned}$$

En este caso, Ω es conocida. Ahora, supongamos que no la conocemos. Entonces, podemos utilizar el método de FGLS que utiliza un estimador consistente para Ω. FGLS es asintóticamente equivalente a MCG. En este caso, el procedimiento consiste en:

a) Estimar el modelo por OLS. Guardarse los residuos.

b) Generar el logaritmo de los residuos al cuadrado, $\ln(\hat{u}^2)$. Regresarlos contra la variable x_2:

$$\ln(\hat{u}_i^2) = \gamma_1 + \gamma_2 x_{2i} + \nu_i$$

Esto asegura que la varianza estimada sea positiva *c*) Generar la predicción del logaritmo de la varianza σ_i^2. Transformarlos tomándole la exponencial para obtener $\widehat{\sigma}_i^2 = exp(\widehat{\gamma}_1 + \widehat{\gamma}_2 x_{2i})$

d) Regresar $y_i/\widehat{\sigma}_i$ sobre $x_i/\widehat{\sigma}_i$.

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

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:

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

   a) Utilizando el comando egen de STATA, construya las medias individuales de las variables del modelo.

   b) Aplique la transformación within al modelo. Luego, estime el modelo transformado por POLS.

   c) Comente sobre la validez de los errores estándar del inciso previo.

Solution: La estimación de la varianza de los errores en el inciso previo es

$$\widehat{\sigma}_\varepsilon^2 = RSS/(NT-K).$$

Sin embargo, sabemos que una estimación consistente para σ_ε^2 es

$$\hat{\sigma}_\varepsilon^2 = \frac{1}{N(T-1)-K}\sum_{i=1}^{N}\sum_{t=1}^{T}\hat{\hat{\varepsilon}}_{it}^2.$$

Por lo tanto, los errores estándar reportados tienden a ser pequeños comparados a los verdaderos. El problema se encuentra en que los grados de libertad de aplicar OLS al modelo transformado no coinciden con el denominador del estimador consistente para σ_ε^2. Por consiguiente, salvo que T sea lo suficientemente grande, necesitamos corregir este denominador.

   d) Utilice el comando *xtreg* para estimar nuevamente el modelo usando efectos fijos.

Solution: Estimamos utilizando los comandos para datos de panel, en particular, la opción para la estimación por efectos fijos.

```
xtset county year
xtreg lcrmrte lprbarr lprbconv lprbpris lavgsen lpolpc d82-d87, fe
```

   e) Estime el modelo usando diferencias finitas de primer orden.

Solution: Al no haber un comando nativo del tipo "fd y x" necesitamos construir las variables. Esto puede hacerse creando las variables

```
gen dlcrmrte = lcrmrte−L1.lcrmrte
for any $xlist : gen dX= X−L1.X
reg dlcrmrte dlprbarr dlprbconv dlprbpris dlavgsen dlpolpc dd82−dd87, nocons
```

Solution: Las estimaciones de este ejercicio se presentan en la siguiente tabla:

	(1) POLS	(2) FE	(3) FD
lprbarr	-0.720***	-0.360***	-0.327***
	(0.110)	(0.0324)	(0.0300)
lprbconv	-0.546***	-0.286***	-0.238***
	(0.0704)	(0.0212)	(0.0182)
lprbpris	0.248**	-0.183***	-0.165***
	(0.109)	(0.0325)	(0.0260)
lavgsen	-0.0868	-0.00449	-0.0218
	(0.113)	(0.0264)	(0.0221)
lpolpc	0.366***	0.424***	0.398***
	(0.121)	(0.0264)	(0.0269)
d82	0.00514	0.0126	0.00771
	(0.0367)	(0.0215)	(0.0171)
d83	-0.0435	-0.0793***	-0.0844***
	(0.0336)	(0.0213)	(0.0235)
d84	-0.109***	-0.118***	-0.125***
	(0.0392)	(0.0216)	(0.0287)
d85	-0.0780**	-0.112***	-0.122***
	(0.0386)	(0.0218)	(0.0331)
d86	-0.0421	-0.0818***	-0.0863**
	(0.0429)	(0.0214)	(0.0367)
d87	-0.0270	-0.0405*	-0.0378
	(0.0381)	(0.0210)	(0.0400)
_cons	-2.082**	-1.604***	
	(0.865)	(0.169)	
N	630	630	540

Standard errors in parentheses

* $p < 0{,}10$, ** $p < 0{,}05$, *** $p < 0{,}01$

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.

   a) 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.

Solution: A continuación, se muestran los resultados de la estimación por OLS de la ecuación anterior

	Tasa de Homicidios
Ejecuciones	0.163
	(0.84)
Desempleo	1.391**
	(3.08)
Dummy año 1990	2.675
	(1.47)
Dummy año 1993	1.607
	(0.91)
Constante	-1.864
	(-0.61)
R^2	0.08
N	153

* $p < 0.05$; ** $p < 0.01$; t-statistics en paréntesis

En la tabla anterior se puede observar que el coeficiente de la cantidad de ejecuciones ($\hat{\beta}_x$) no es estadísticamente significativo para explicar la cantidad de asesinatos. Si la ecuación anterior estuviera bien estimada (es decir, si todos los supuestos necesarios para que OLS sea consistente se cumpliesen), los resultados indicarían evidencia en contra de la pena de muerte como forma de prevenir los homicidios. En otras palabras, una mayor cantidad de ejecuciones no tiene efecto disuasorio sobre posibles homicidas, ya que estadisticamente la variable no resulta significativa en la regresión corrida. Es más, no solo no resulta siginificativa, sino que el signo es el opuesto al que, a priori, los defensores de la pena de muerte afirman que debería ser.

Dentro de las variables incluidas como controles, la única que resulta significativa es el desempleo correspondiente a cada estado ($\hat{\beta}_u$). De acuerdo a los resultados de la estimación, la tasa de desempleo se relaciona de forma positiva con la tasa de homicidios, es decir, $\hat{\beta}_u > 0$, lo cual estaría indicando evidencia a favor de que el malestar socioeconómico (tomando como posible *proxy* la tasa de desempleo) debiera ser un factor a tener en cuenta al momento de analizar temas relacionados con la inseguridad y la criminalidad.

Por otra parte, podemos observar que ninguna de las dummies temporales ($\hat{\beta}_{90}$ y $\hat{\beta}_{93}$) resultan estadísticamente siginificativas, lo cual indica evidencia a favor de la ausencia de efectos fijos temporales.

b) ¿Por qué podría ser importante tener en consideración los efectos temporales agregados en el modelo?

Solution: Podría ser importante si la tasa de homicidios es afectada por factores macroeconómicos externos que afectan a todos los estados de EE.UU. de la misma manera. Por lo tanto, si no incluimos estas variables, debemos suponer que cualquier cambio en la media de la tasa de homicidios en el tiempo se debe a las ejecuciones o a la tasa de desempleo y no a factores externos. Por otra parte, controlar por estas variables hace más factible que se cumpla el supuesto de ausencia de autocorrelación serial.

c) 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.

d) Repita la estimación del inciso previo usando diferencias finitas de primer orden.

Solution: Para hallar los estimadores de Efectos Fijos, estimamos la siguiente ecuación:

$$m_{i,t} - \bar{m}_i = \beta_x(x_{i,t} - \bar{x}_i) + \beta_u(u_{i,t} - \bar{u}_i) + \beta_{90}(d_{90} - \bar{d}_{90,i}) + \beta_{93}(d_{93} - \bar{d}_{93,i}) + (e_{i,t} - \bar{e}_i)$$

donde $\bar{y}_i = T^{-1}\sum_{t=1}^{T} y_{it}$.

Mientras que para obtener los estimadores de Primeras Diferencias, estimamos la siguiente ecuación:

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

donde Δ indica que se han aplicado primeras diferencias a los datos correspondientes.

A continuación, se muestra una tabla con los resultados de las estimaciones de las ecuaciones anteriores:

	LSDV	FE	FD
Ejecuciones	-0.138	-0.138	
	(-0.78)	(-0.78)	
Desempleo	0.221	0.221	
	(0.75)	(0.75)	
Dummy año 1990	1.556*	1.556*	
	(2.09)	(2.09)	
Dummy año 1993	1.733*	1.733*	
	(2.47)	(2.47)	
Δ Ejecuciones			-0.115
			(-0.78)
Δ Desempleo			0.163
			(0.53)
Δ Dummy año 1990			1.511*
			(2.29)
Δ Dummy año 1993			1.725*
			(2.02)
Constante	5.904	5.822**	
	(1.78)	(3.04)	
R^2	0.91	0.07	0.06
N	153	153	102

* $p < 0.05$; ** $p < 0.01$; t-statistics en paréntesis

En la tabla reportamos tres estimaciones. Las dos primeras columnas son por efectos fijos: una de ellas utilizando el comando especifico que provee Stata y otra utilizando un set de dummies para cada individuo. Observamos que los coeficientes en ambas estimaciones (a excepción de la constante) son iguales. La tercera columna es la estimación de primeras diferencias, donde el coeficiente relevante viene dado por el de la variable Δ Ejecuciones, es decir, la primera diferencia de la variable ejecuciones.

En la estimación del modelo de Efectos Fijos se observa que el coeficiente $\hat{\beta}_x$ no es significativo, es decir que las ejecuciones no tienen impacto sobre la tasa de homicidios, al igual que en el resultado hallando en el inciso anterior. Sin embargo, el signo del coeficiente es opuesto al signo del coeficiente del modelo OLS, siendo ahora negativo, en concordancia con lo que afirman los defensores de la pena capital, aunque sin ser significativo.

Por otra parte, ahora tenemos que el coeficiente que acompaña a la tasa de desempleo no es significativo, por lo que en la estimación de este modelo la tasa de desempleo no tiene impacto sobre la tasa de homicidios. Adicionalmente, las dummies para el año 1990 y el año 1993 en este caso dan significativos, lo cual no ocurría en el inciso anterior.

e) 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.

Solution: La variable explicativa de cantidad de ejecuciones (x_{it}) podría fallar en cuanto a la exogeneidad estricta si los estados aumentan las ejecuciones futuras en respuesta a los shocks positivos actuales de la tasa de homicidios. Dado el tramo de tiempo relativamente corto de la base de datos, la retroalimentación de las tasas de homicidio a las ejecuciones futuras puede no ser muy preocupante, ya que el proceso judicial en los casos de pena capital tiende a moverse lentamente. (Por supuesto, si se acelerara debido a un aumento de las tasas de homicidio, eso podría violar la exogeneidad estricta). Con una serie temporal más larga podríamos añadir $x_{i,t+1}$ (e incluso valores de un futuro más lejano) y estimar la ecuación por FE, comprobando la significación estadística de la variable $x_{i,t+1}$. En el caso de que se encuentre que esta variable es estadísticamente significativa tendríamos evidencia en favor de que no se cumple el supuesto de exogeneidad estricta.

f) 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.

Solution: Computamos en Stata:

$$\begin{aligned}
\widehat{\beta}^{FDGLS} &= \left(\sum_{i=1}^{N} X_i' D' \left(DD'\right)^{-1} D X_i\right)^{-1} \left(\sum_{i=1}^{N} X_i' D' \left(DD'\right)^{-1} D y_i\right) \\
&= \left(X'(I_N \otimes D' \left(DD'\right)^{-1} D) X\right)^{-1} \left(X'(I_N \otimes D' \left(DD'\right)^{-1} D) y\right)
\end{aligned}$$

g) Reestimar el modelo del inciso c) usando efectos aleatorios. Implementar el test de Hausman. ¿Cuál es el mejor estimador?

Solution: A continuación se presentan los resultados obtenidos de Stata de la estimación del modelo por Efectos Aleatorios:

	RE
Ejecuciones	-0.0543
	(-0.34)
Desempleo	0.395
	(1.39)
Dummy año 1990	1.733**
	(2.32)
Dummy año 1993	1.700**
	(2.41)
Constante	4.635**
	(2.13)
N	153
Number of id	51

*** p<0.01, ** p<0.05, * p<0.1

En la estimación del modelo de Efectos Aleatorios se vuelve a observar que las ejecuciones no tienen efecto sobre la tasa de homicidios dado que el coeficiente que acompaña a esta variable no es significativo. Adicionalmente, el signo de este coeficiente es negativo, alineado a la teoría de quienes defienden la pena de muerte, aunque, nuevamente, repetimos que el efecto sigue sin ser significativo.

Por otra parte, al igual que en el inciso anterior, se vuelve a observar que el desempleo no tiene efecto alguno sobre la tasa de homicidios, mientras que las dummies para el año 1990 y el año 1993 son significativas, lo cual es evidencia a favor de la presencia de efectos fijos temporales.

Test de Hausman

Naturalmente surge la cuestión sobre cuál es el mejor estimador. Para ello se plantea el test de Hausman. De forma resumida, el test se basa en la diferencia de las estimaciones de efectos fijos y efectos aleatorios. Bajo la hipótesis nula de $cov(X, \mu_i) = 0$ para todas las covariables y para todo i, ambos estimadores son consistentes, por lo que $\hat{\beta}_{FE} - \hat{\beta}_{RE}$ converge en probabilidad a cero. Además, como bajo H_0 RE es eficiente, la matriz de varianzas y covarianzas de esa diferencia es la resta de las matrices de varianzas y covarianzas. El estadístico en el que se basa el test es

$$H = (\hat{\beta}_{FE} - \hat{\beta}_{RE})'[\hat{V}(\hat{\beta}_{FE}) - \hat{V}(\hat{\beta}_{RE})]^{-1}(\hat{\beta}_{FE} - \hat{\beta}_{RE})$$

Donde $(\hat{\beta}_{FE} - \hat{\beta}_{RE})$ es un vector columna de dimensión k, donde k es la cantidad de coeficientes estimados (no incluimos la fila correspondiente a la constante porque el test de Hausman por defecto no la usa, lo cual es logico porque en FE no está identificada. Se la puede identificar poniendo la restricción de que $\sum_i^I \mu_i = 0$ como hace Stata, pero de todas maneras no tiene sentido incluir la constante ya que es exógena por definición), $\hat{V}(\hat{\beta}_{FE})$ es la estimación de la matriz de varianzas y covarianzas del estimador de FE y $\hat{V}(\hat{\beta}_{RE})$ es la estimación de la matriz de varianzas y covarianzas del estimador de RE. La distribución asintotica del estimador es χ^2 con k grados de libertad.

Computando el test en Stata notamos en este caso que el programa nos reporta que la matriz de varianzas y covarianzas que no es positiva definida. Uno podría pensar que esta es una manifestación del problema de muestra finita que puede tener el estimador (la diferenecia de las matrices puede no ser definida positiva en muestra finita, aunque asintóticamente bajo H_0 tenga que serlo). Calculando los autovalores de la matriz, vemos que algunos son negativos, confirmando el problema de que la matriz $[\hat{V}(\hat{\beta}_{FE}) - \hat{V}(\hat{\beta}_{RE})]$ no es positiva definida.

Ahora bien, ¿realmente tenemos realmente un problema de muestra finita? La respuesta en este caso es negativa. En el libro de Wooldridge (2010, p. 239) citado en el programa, el autor nos aclara que, en relación al test de Hausman para comparar FE y RE:

"A third caveat concerns the set of parameters that we can compare. Because the FE approach only identifies coeficients on time-varying explanatory variables, we clearly cannot compare FE and RE coeficients on time-constant variables. But there is a more subtle issue: we cannot include in our comparison coeficients on aggregate time efect, that is, variables that change only across time. (*Nota*: como las dummies temporales que incluimos). *As with the case of comparing FE and FD estimates, the problem with comparing coeficients on aggregate time efects is not one of identification; we know RE and FE both allow inclusion of a full set of time period dummies. The problem is one of singularity in the asymptotic variance matrix of the diference between* $\hat{\beta}_{FE}$ *and* $\hat{\beta}_{RE}$. *[...] To summarize, we can estimate models that include aggregate time effects, time constant variables, and regressors that change across both i and t, by RE and FE estimation. But no matter how we compute a test statistic, we can only compare the coefficients on the regressors that change across both i and t."*

Un posible curso de acción en este caso es usar el test para evaluar la diferencia solo de estas dos variables (*exec* u *unem*). Notar que esto tiene cierto sentido teórico: lo que estamos excluyendo es algo que varía solo a través del tiempo, por lo que uno intuiría que $cov(d_t, \mu_i) = 0$, ya que una de las variables solo varía en una dimensión (la temporal) y la otra solo difiere entre individuos. En otras palabras, dado un i, μ_i es constante a traves del tiempo, por lo que la covarianza de dicha variable aleatoria debería ser cero.

Para hacer este test de Hausman solo sobre dos coeficientes, lo que hacemos es tomar la matriz de varianzas original y solo quedarnos con la submatriz de 2×2 que corresponde a las varianzas y covarianzas de la diferencia de los estimadores de *exec* y *unem*. Chequeando los autovalores de esta submatriz, vemos que son ambos positivos, por lo la matriz es definida positiva. Luego computamos el estadístico de Hausman con la formula provista. Obtenemos un valor del estadístico de Hausman de 5,7757, que corresponde a un *p-value* de 0,06 (este valor lo sacamos de la tabla de una distribución χ^2 con dos grados de libertad). En este caso, no rechazamos la hipotesis nula al 5 % de confianza, pero si al 10 %.

En conclusión, el test de Hausman rechaza la hipótesis nula de no correlación entre los regresores y los μ_i a un nivel de significatividad de 10 %. Por lo tanto, por un criterio puramente estadístico, si el nivel adoptado fuese de 10 %, el unico estimador consistente sería efectos fijos, por lo que este sería el mejor.

3. Considere el siguiente modelo:

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

donde $x_{it} \overset{iid}{\sim} \mathcal{N}(0, 1)$, $u_i \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\mu^2\right)$, $\nu_{it} \overset{iid}{\sim} \mathcal{N}\left(0, \sigma_\nu^2\right)$ 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.

Solution:

N	OLS			RE			FE		
	Media	SD	RMSE	Media	SD	RMSE	Media	SD	RMSE
5	0.9948	0.1986	0.1986	0.9959	0.1565	0.1565	0.9963	0.1527	0.1527
10	0.9987	0.1549	0.1549	0.9992	0.1090	0.1090	0.9992	0.1078	0.1078
30	1.0009	0.0828	0.0828	1.0014	0.0617	0.0617	1.0015	0.0619	0.0619
50	1.0022	0.0625	0.0625	1.0019	0.0464	0.0464	1.0019	0.0467	0.0467
100	0.9985	0.0438	0.0438	0.9999	0.0331	0.0331	1.0000	0.0333	0.0333
500	0.9996	0.0446	0.0446	0.9988	0.0332	0.0332	0.9987	0.0334	0.0334

Cuadro 1: Resultados de las simulaciones

En primer lugar, es importante destacar que dados los supuestos del modelo, los tres estimadores en consideración son consistentes. Por lo tanto, deberíamos esperar que a medida que el tamaño muestral aumente, la media de las estimaciones de β con los diferentes estimadores estén cerca del valor 1. Ahora bien, para $N < 10$ el estimador FE es el que mejor funciona en términos de sesgo y de eficiencia. Luego, a partir de un tamaño de muestra de $N = 30$ ya se observa como el estimador RE es el más eficiente de todos, es decir, es el que presenta un menor desvío estándar, lo cual se vincula a que, dados los supuestos del modelo, es el estimador con la menor varianza asintótica. En resumen, si en la práctica trabajaramos con un modelo donde suponemos que se cumplen los supuestos del modelo del inciso, entonces, para N muy pequeños uno optaría por utilizar el estimador FE, mientras que ya a partir de $N = 30$ uno optaría por el estimador de RE por su eficiencia.

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*?

   Solution: Se puede apreciar que las estimaciones de los coeficientes son numéricamente idénticas.

   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?

   Solution: Las estimaciones de POLS y RE son numéricamente idénticas. Este es un resultado general: si el modelo incluye sólo efectos temporales agregados y covariables específicas del individuo que no tienen variación temporal, entonces, POLS = RE.

   Por otra parte, cuando se utiliza FE, por supuesto, no se pueden estimar los coeficientes asociados a las variables constantes en el tiempo. Las estimaciones de las variables *dummy* de año son las mismas que las de POLS y RE. Sin embargo, cuando POLS y RE incluyen variables constantes en el tiempo, la estimación de la "constante" de FE no es igual a la estimación del intercepto en POLS/RE.

   c) ¿Son iguales los errores estándar de POLS y RE del inciso b)? ¿Cuáles son probablemente más fiables?

   Solution: Los errores estándar reportados para POLS y RE no son los mismos. Los errores estándar de POLS suponen, además de homocedasticidad, que no hay correlación serial en el error compuesto, es decir, que no considera la posible presencia de una heterogeneidad no observada. Al menos, los errores estándar de RE permiten en su estructura estándar la presencia de correlación serial, en particular, la cual es igual para todos los pares de períodos (t, s). Esto puede ser demasiado restrictivo, pero es menos restrictivo que los habituales errores estándar OLS.

   d) Obtenga los errores estándar robustos para POLS. ¿Prefiere estos o los errores estándar habituales de RE?

   Solution: Estos errores estándar robustos permiten cualquier tipo de correlación serial y de heterocedasticidad de los disturbios que varíen en el tiempo. Preferimos estos a los errores estándar habituales de RE ya que estos últimos imponen un tipo especial de correlación serial, y, además, asumen homocedasticidad.

   e) Obtenga los errores estándar robustos de RE. ¿Cómo se comparan con los errores estándar robustos de POLS, y por qué?

   Solution: Son numéricamente idénticos a los errores estándar robustos de POLS porque tenemos un solo estimador (POLS = RE en esta configuración) y, por lo tanto, hay una sóla varianza robusta.

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 3
Modelos de Datos de Panel Dinámicos

---

1. Considere la base de datos *mod_abdata.dta* que fue utilizada por Arellano y Bond en su famoso paper de 1991. Se trata de un panel de 140 empresas británicas encuestadas anualmente entre 1976 y 1984. El panel original no es balanceado, pero la versión para este ejercicio se trata de un panel balanceado de empresas con observaciones para exactamente 6 años entre 1977 y 1982. La variable que identifica la empresa es *id* y la variable que identifica el tiempo es *year*. La variable n es el empleo de la empresa. Luego, considere un modelo muy simplificado del siguiente tipo:

$$\ln n_{it} = \rho \ln n_{it-1} + \varepsilon_{it}$$
$$\varepsilon_{it} = c_i + \nu_{it}$$
$$E[c_i] = E[\nu_{it}] = E[c_i \nu_{it}] = 0$$

donde n_{it} es el empleo de la empresa i en el año t.

a) Estime el modelo por OLS. ¿Qué sesgo esperaría encontrar y por qué?

> **Solution:** Un problema inmediato que encontramos es que $\ln n_{it-1}$ está correlacionado con los efectos fijos que se encuentran en el término del error, lo cual da lugar al sesgo de paneles dinámicos. En general, bajo muchos supuestos, OLS sobrestima el valor real del parámetro ρ (ver Hsiao, p. 85).

b) Estime el modelo usando efectos fijos (FE). ¿Permite la transformación *within* eliminar el sesgo de paneles dinámicos?

Solution: Ahora, el problema se encuentra en que $\widetilde{\ln n}_{it-1} = \ln n_{it-1} - \overline{\ln n}_{i-1}$ está correlacionado con $\tilde{\nu}_{it} = \nu_{it} - \bar{\nu}_{i-1}$ aún cuando ν_{it} no tiene correlación serial. En particular, el término $\ln n_{it-1}$ correlaciona negativamente con $-(1/T-1)\nu_{it-1}$ que se encuentra dentro de $\bar{\nu}_{i-1}$, mientras que, simétricamente, $-(1/T-1)\ln n_{it}$ y ν_{it} también se encuentran correlacionados negativamente. Adicionalmente, hay otros pares de términos que correlacionan, pero su impacto es de segundo orden. Por lo tanto, notemos que en este caso el sesgo no surge por los c_i sino por el hecho de que, al ser dinámico el panel, estamos poniendo los ν_{it} en varios lados distintos, y eso es lo que genera el sesgo. Por último, cabe mencionar que Nickell mostró que este sesgo es siempre negativo si $\rho > 0$.

c) Considere una transformación de diferencias finitas de primer orden del modelo. ¿Continúa siendo la variable dependiente rezagada potencialmente endógena?

Solution: A pesar de que los efectos fijos se van, la variable dependiente rezagada aún es potencialmente endógena ya que el término $\ln n_{it-1}$ en $\Delta \ln n_{it-1} = \ln n_{it-1} - \ln n_{it-2}$ está correlacionado con ν_{it-1} en $\Delta\nu_{it} = \nu_{it} - \nu_{it-1}$.

d) Implemente el estimador de Anderson-Hsiao a partir del comando *ivregress* en Stata.

Solution: En este contexto, $\ln n_{it-2}$ surge como un instrumento candidato natural para $\Delta \ln n_{it-1}$ ya que matemáticamente están relacionados, pero $\ln n_{it-2}$ no está relacionado con el término de error $\Delta\nu_{it}$ siempre y cuando los v_{it} no presenten correlación serial. Esto nos conduce al estimador propuesto por Anderson y Hsiao (1982). De esta forma, esta es la primera estimación consistente del modelo dados nuestros supuestos. Ahora bien, recordemos que 2SLS es eficiente bajo el supuesto de errores i.i.d. esféricos. Sin embargo, luego de tomar diferencias de primer orden, los errores $\Delta\nu_{it}$ pueden estar lejos de ser independientes. $\Delta\nu_{it}$ puede estar correlacionado con $\Delta\nu_{it-1}$ dado que comparten el término ν_{it-1}.

e) Ahora, obtenga la estimación GMM de ρ utilizando todos los instrumentos posibles en niveles para el modelo en primeras diferencias. Para ello utilice el comando *xtabond2*.

> **Solution:** El estimador GMM de Arellano-Bond (1991) aborda directamente el problema del inciso previo, modelando la estructura del error de forma más realista, lo que hace que sea asintóticamente más preciso y se comporte mejor en la práctica. Adicionalmente, al estar utilizando más condiciones de momentos, esto nos provee de mayor información, lo que nos permite obtener un estimador más eficiente que el del inciso previo.

f) Obtenga la estimación de GMM de ρ utilizando todos los instrumentos posibles en niveles para el modelo en primeras diferencias e Δy_{it-1} como instrumento para el modelo en niveles.

> **Solution:** Aquí lo que estamos haciendo es implementar el estimador GMM de Blundell-Bond (1998).

g) Repita las estimaciones de los incisos e) y f) incluyendo efectos fijos de tiempo.

> **Solution:** En la estimaciones que vimos hasta aquí, se rechaza la hipótesis nula del test de Sargan, el cual sabemos que se puede pensar como un test de especificación del modelo. Esto implica que el *Data Generating Process* no es el modelo simple que vimos hasta aquí. Por lo tanto, la idea al incluir efectos fijos de tiempo es ir "mejorando" el modelo de a poco.

2. En este ejercicio se ilustrará el hecho de que los estimadores de Arellano-Bond y de Blundell-Bond pueden extenderse en forma directa a modelos que incluyan regresores estríctamente exógenos y regresores secuencialmente exógenos.

   En su paper original, Arellano y Bond modelaron el empleo de las empresas (n) utilizando un modelo de ajuste parcial para reflejar los costos de contratación y despido, incluyendo dos rezagos de la variable empleo. Otras variables incluidas fueron el nivel salarial actual y el rezagado (w), el stock de capital actual, rezagado una y dos veces (k) y la producción agregada actual, rezagada una y dos veces en el sector de la empresa (ys). Todas las variables se expresan en logaritmos. También se incluye un conjunto de variables *dummy* de tiempo.

a) Estime el modelo por OLS. Compute los errores estándar robustos a heterocedasticidad y correlación serial.

b) Estime el modelo por FE. Compute los errores estándar robustos a heterocedasticidad y correlación serial.

Solution: Sabemos que bajo ciertos supuestos, en el modelo de regresión la variable dependiente rezagada está correlacionada positivamente con el error, lo que hace que la estimación por OLS de su coeficiente esté sesgada hacia arriba. Por otra parte, en la estimación de efectos fijos, se subestima el verdadero valor del coeficiente que acompaña al primer rezago de n debido al signo negativo en el período $t-1$ en el error transformado. Por lo tanto, dadas las direcciones opuestas del sesgo presente en estas estimaciones, una estimación consistente para el verdadero parámetro debería estar entre estos valores.

c) Implemente el estimador de Anderson-Hsiao usando n_{it-2} como instrumento.

Solution: Aunque esta estimación debería ser consistente, los resultados obtenidos no son los esperados. El coeficiente del rezago de n está fuera de los límites de sus correspondientes estimaciones de OLS y FE, y, es mucho mayor que la unidad, una condición de estabilidad. Además, vemos que la estimación del error estándar es muy alta, y el coeficiente no es estadísticamente significativo a los niveles convencionales.

d) Estime la ecuación de empleo usando el estimador de Arellano-Bond. Asuma que la única endogeneidad presente es en el rezago de la variable dependiente.

Solution: Notamos que el coeficiente de la variable dependiente rezagas se encuentra ahora dentro del rango de estabilidad. Sin embargo, en ambas estimaciones encontramos que no es estadísticamente significativo.

e) Ahora, considere como hicieron Blundell y Bond (1998) que los salarios y el stock de capital no deben tomarse como estrictamente exógenos en este contexto (como se hizo en los modelos anteriores). Reestime el modelo usando el estimador de A-B y considerando a los salarios y el stock de capital como regresores secuencialmente exógenos.

Solution: Ahora, los resultados de la estimación en un paso y en dos pasos parecen razonables.

f) Adicionalmente, Blundell y Bond (1998) eliminan de su modelo los rezagos más largos (de dos períodos) del empleo y el capital, y prescinden del nivel de producto agregado sectorial. Considerando esta cuestión, compute el estimador de Blundell-Bond.

Solution: El resultado de la estimación está dentro de los valores esperados.

3. Cuando hay muchos instrumentos, surgen dos problemas principales:

   - Sobreestimación (*overfitting*) de la variable endógena
   - Mala estimación de la matriz de pesos W

   En estos casos, se proponen las siguientes soluciones:

   a) Probar diferentes especificaciones de IV recortando el número de rezagos en la matriz de instrumentos $\mathbf{Z}$.

   b) Colapsar/combinar instrumentos. Se modifica la matriz de instrumentos para el individuo i:

$$\mathbf{Z}_i = \begin{bmatrix} y_{i1} & 0 & 0 & \cdots \\ y_{i1} & y_{i2} & 0 & \cdots \\ y_{i3} & y_{i2} & y_{i1} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

Si el modelo funciona debería dar resultados similares con distintos instrumentos.

Retome el ejercicio 2.e) para ver una aplicación de esta cuestión. Estime el modelo de empleo restringiendo el máximo rezago a 3 y 4 períodos. Por último, estime el modelo colapsando instrumentos. Analice si los resultados obtenidos son robustos.

4. Considere nuevamente el modelo del primer ejercicio. Obtenga el estimador LSDVC propuesto por Kiviet (1995) a partir del comando *xtlsdvc*. Luego, estime la matriz de varianzas y covarianzas de los coeficientes de Kiviet siguiendo el procedimiento explicado en clase.

Universidad Torcuato Di Tella

Maestría en Economía y Econometría

2022

Datos de Panel

Problem Set 4

Modelos Lineales en Paneles Desbalanceados

---

1. Utilice la base de datos *"keane.dta"* la cual contiene el historial de empleo y escolaridad de una muestra de hombres para los años 1981 a 1987. Luego, considere la siguiente ecuación de salarios:

$$\ln(wage_{it}) = \beta_0 + \beta_1 exper_{it} + \beta_2 educ_{it} + c_i + u_{it}, \quad t = 1, 2, \ldots, T \tag{1}$$

donde $\ln(wage_{it})$ es el logaritmo del salario por hora, $exper_{it}$ son los años de experiencia en el mercado laboral y $educ_{it}$ son los años de escolaridad. Responda las siguientes preguntas:

a) Estime la ecuación usando efectos fijos. ¿Cuál es el sesgo potencial en este contexto?

Solution: En este contexto de efectos fijos la selección muestral por truncamiento incidental es un problema si la selección está relacionada con los errores idiosincráticos de la ecuación de interés, u_{it}. En este sentido, si pensamos que efectivamente lo anterior se cumple y que estamos observando los salarios "más altos" (los mejores salarios que se ofrecieron) entonces el truncamiento tendría como consecuencia una sobreestimación de los retornos a la educación.

b) Implemente el contraste de sesgo de selección propuesto por Wooldridge (1995) bajo el enfoque de Mundlak (1978).

c) Implemente el contraste de sesgo de selección propuesto por Wooldridge (1995) bajo el enfoque de Chamberlain (1980).

Solution: La hipótesis nula en ambos contrastes de hipótesis establece que la inversa del cociente de Mills estimada no debería ser estadísticamente significativa en nuestro modelo si no hay problemas de selección de la muestra. Sin embargo, en ambos casos encontramos evidencia en favor de rechazar la hipótesis nula.

2. Considerando nuevamente la ecuación de salarios del ejercicio previo, realice los siguientes procedimientos:

   a) Estime el modelo por Wooldridge (1995) bajo el enfoque de Chamberlain (1980).

   b) Estime el modelo por Wooldridge (1995) bajo el enfoque de Mundlak (1978).

   c) Comente sobre los errores estándar de las estimaciones anteriores.

   d) Estime los errores estándar vía bootstrapping.

   e) Estime los errores estándar analíticos (varianza asintótica).

Universidad Torcuato Di Tella
Maestría en Economía y Econometría
2022

Datos de Panel
Problem Set 5
Modelos de Variable Dependiente Discreta

---

1. El archivo *wagepan.dta* contiene los datos utilizados por Vella y Verbeek (1998). Estos datos contienen información para 545 hombres que trabajaron cada año de 1980 a 1987. Utilice los datos para analizar el impacto de la escolaridad ($educ_{it}$) en la probabilidad de estar afiliado a un sindicato ($union_{it}$). Las variables se describen en el conjunto de datos. Observe que la educación no cambia con el tiempo.

   a) Use Pooled OLS para estimar el modelo:

   $$P\left(union_{it} = 1 \mid educ_{it}\right) = \beta_0 + \beta_1 educ_{it} \tag{1}$$

   De acuerdo a los resultados obtenidos, ¿tiene impacto un año más de escolaridad sobre la probabilidad de estar afiliado a un sindicato?

   Solution: Si bien el coeficiente de interés estimado es negativo, $\widehat{\beta_1} = -0{,}0015$, el mismo no es estadísticamente significativo. Por lo tanto, no hay evidencia suficiente para rechazar la hipótesis nula que un año más de escolaridad no modifica la probabilidad de estar afiliado a un sindicato.

   Por otra parte, respecto a la metodología, utilizar un modelo de probabilidad lineal cuando la variable dependiente es lineal tiene los mismos problemas que en el caso de una muestra de corte transversal: 1) Nada garantiza que el valor predicho esté entre 0 y 1, y 2) el modelo tiene heterocedasticidad por construcción.

   b) Use Pooled Probit para estimar el modelo:

   $$P\left(union_{it} = 1 \mid educ_{it}\right) = \Phi\left(\beta_0 + \beta_1 educ_{it}\right) \tag{2}$$

   Comente sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato.

   Solution: Dado que la variable años de educación se trata de una variable continua, entonces el efecto parcial de la escolaridad sobre la probabilidad de estar afiliado a un sindicato es:

   $$\frac{\partial P\left(union_{it} = 1 \mid educ_{it}\right)}{\partial educ_{it}} = \beta_1 \phi\left(\beta_0 + \beta_1 educ_{it}\right)$$

Claramente el signo del efecto marginal depende del signo de β_1. Sin embargo, la magnitud cambiará con respecto a los valores de $educ_{it}$. Un consenso en la práctica es evaluar en el promedio de las x. Cuando hacemos esto en Stata obtenemos que un año más de educación reduce la probabilidad de estar afiliado en $-0{,}0016$, aunque esta caída no resulta estadísticamente significativa.

c) Use Pooled Logit para estimar el modelo:

$$P\left(union_{it}=1 \mid educ_{it}\right)=\Lambda\left(\beta_0+\beta_1 educ_{it}\right)=\frac{e^{\beta_0+\beta_1 educ_{it}}}{1+e^{\beta_0+\beta_1 educ_{it}}} \tag{3}$$

Comente sobre el impacto de un año más de educación en la probabilidad de estar afiliado a un sindicato. Compute el error estándar para esta estimación.

Solution: El efecto parcial de la escolaridad sobre la probabilidad de estar afiliado a un sindicato es:

$$\frac{\partial P\left(union_{it}=1 \mid educ_{it}\right)}{\partial educ_{it}}=g(\beta_0,\beta_1)=\beta_1\Lambda\left(\beta_0+\beta_1 educ_{it}\right)\left(1-\Lambda\left(\beta_0+\beta_1 educ_{it}\right)\right)$$

Al computar el efecto marginal en el promedio de la variable escolaridad, encontramos que un año más de escolaridad reduce la probabilidad de estar afiliado a un sindicato en $-0{,}0015$, aunque nuevamente se encuentra que esta caída no es estadísticamente significativa.

Para computar el error estándar de la estimación debemos utilizar el Método Delta:

$$V\left(\widehat{\frac{\partial P\left(union_{it}=1 \mid educ_{it}\right)}{\partial educ_{it}}}\right)=\frac{\widehat{\partial g}}{\beta'}\cdot\widehat{V(\beta)}\cdot\frac{\widehat{\partial g}}{\beta'}'$$

La estimación puntual $\widehat{V(\beta)}$ la recuperamos de los resultados de la estimación del modelo al utilizar el comando *logit*. Lo que es importante es calcular las derivadas parciales de la función g con respecto a los parámetros β. Lo hacemos a continuación:

$$\begin{aligned}\frac{g(\beta_0,\beta_1)}{\beta_1}&=\Lambda\left(1-\Lambda\right)+\beta_1\left[\Lambda'\left(1-\Lambda\right)educ_{it}-\Lambda'\Lambda educ_{it}\right]\\&=\Lambda\left(1-\Lambda\right)+\beta_1\left[1-2\Lambda\right]\Lambda' educ_{it}\end{aligned}$$

Luego, tenemos que:

$$\frac{\widehat{g(\beta_0,\beta_1)}}{\beta_1}=\widehat{\Lambda}\left(1-\widehat{\Lambda}\right)+\widehat{\beta}_1\left[1-2\widehat{\Lambda}\right]\widehat{\Lambda}'\overline{educ}_{it}$$

Bajo un procedimiento análogo se puede obtener que:

$$\frac{\widehat{g(\beta_0,\beta_1)}}{\beta_0}=\widehat{\beta}_1\left[1-2\widehat{\Lambda}\right]\widehat{\Lambda}'$$

Por último, se recuerda que $\Lambda'=\Lambda\left(1-\Lambda\right)$.

d) Estime la siguiente extensión del modelo (2):

$$P\left(union_{it}=1 \mid educ_{it}, c_i\right)=\Phi\left(\beta_0+\beta_1 educ_{it}+c_i\right) \tag{4}$$

donde c_i son efectos no observables individuales. Use el modelo Probit de efectos aleatorios. ¿Cuál es el problema que surge al momento de estimar el efecto parcial de interés?

> **Solution:** El efecto parcial de interés es:
>
> $$\frac{\partial P\left(union_{it}=1 \mid educ_{it}, c_i\right)}{\partial educ_{it}}=\beta_1 \phi\left(\beta_0+\beta_1 educ_{it}+c_i\right)$$
>
> El efecto parcial depende de c_i, lo cual no es estimado. Por lo tanto, no se puede estimar la magnitud del efecto parcial a menos que insertemos el valor de c_i. Una posibilidad que permite Stata es calcular el efecto marginal en $c_i=0$ lo que tiene sentido ya que estamos asumiendo que $c_i \mid X_i \sim Normal(0, \sigma_c^2)$. Haciendo esto último encontramos que un año más de educación reduce la probabilidad de estar afiliado a un sindicato en $-0{,}0076$. Sin embargo, este efecto no es estadísticamente significativo.

e) Estime la siguiente extensión del modelo (3):

$$P\left(union_{it}=1 \mid educ_{it}, c_i\right)=\Lambda\left(\beta_0+\beta_1 educ_{it}+c_i\right) \tag{5}$$

donde c_i son efectos no observables individuales. Use el modelo Logit de efectos aleatorios. ¿Surge el mismo problema que en el inciso anterior al momento de estimar el efecto parcial de interés?

> **Solution:** El efecto parcial de interés es:
>
> $$\frac{\partial P\left(union_{it}=1 \mid educ_{it}, c_i\right)}{\partial educ_{it}}=\beta_1 \frac{\exp\left(\beta_0+\beta_1 educ_{it}+c_i\right)}{\left(1+\exp\left(\beta_0+\beta_1 educ_{it}+c_i\right)\right)^2}$$
>
> Nuevamente, no se puede estimar la magnitud del efecto parcial a menos que insertemos el valor de c_i. Al calcular el efecto marginal en $c_i=0$ encontramos que un año más de educación reduce la probabilidad de estar afiliado a un sindicato en $-0{,}0059$, pero la caída no es estadísticamente significativa.

f) Compute el denominado estimador Logit de efectos fijos para el modelo (5). ¿Se puede computar el efecto de un año más de educación sobre la probabilidad de estar afiliado a un sindicato? Explique.

> **Solution:** El efecto de interés no puede ser computado porque en este método no se identifican los coeficientes de los regresores que no varían en el tiempo como es el caso de la variable educación en este ejercicio. Por otra parte, aún si el coeficiente estuviera

identificado (en el caso de que la variable varíe en el tiempo), correríamos la misma suerte que en los incisos previos de que deberíamos inservar el valor de c_i sumado a que en este caso es difícil saber que valor insertar ya que no estamos asumiendo que conocemos la distribución de c_i.

g) Considere la siguiente extensión del modelo (4):

$$
\begin{aligned}
&P\left(union_{it}=1 \mid educ_{it}, black_{it}, married_{it}, c_i\right)= \\
&\Phi\left(\beta_0+\beta_1 educ_{it}+\beta_2 black_{it}+\beta_3 married_{it}+c_i\right)
\end{aligned} \tag{6}
$$

donde $black_{it}$ es una variable binaria que toma valor 1 si la persona es afroamericana y $married_{it}$ es una variable binaria que toma valor 1 si la persona es casada. Asuma la siguiente versión de Mundlak (1978) del supuesto de Chamberlain (1980):

$$
c_i \mid X_i \sim Normal\left(\psi+\xi \cdot \overline{married}_i, \sigma_a^2\right) \tag{7}
$$

El modelo dado por (6) y (7) es un caso de lo que en la literatura se denomina modelo Probit de efectos aleatorios de Chamberlain. Al asumir solamente (6) y (7) se tiene que:

$$
\begin{aligned}
&P\left(union_{it}=1 \mid educ_{it}, black_{it}, married_{it}\right)= \\
&\Phi\left[\left(\beta_0+\beta_1 educ_{it}+\beta_2 black_{it}+\beta_3 married_{it}+\psi+\xi \overline{married}_i\right)\left(1+\sigma_a^2\right)^{-1/2}\right] \equiv \\
&\Phi\left[\beta_{0,a}+\beta_{1,a} educ_{it}+\beta_{2,a} black_{it}+\beta_{3,a} married_{it}+\xi_a \overline{married}_i\right]
\end{aligned}
$$

Use Pooled Probit para estimar el modelo. Estime el efecto de la escolaridad sobre la probabilidad de estar sindicalizado para una persona afroamericana casada.

Solution: Se obtiene que para una persona afroamericana y casada un año más de educación reduce la probabilidad de estar afiliado en un sindicato en $-0{,}0012$. Sin embargo, el efecto no es estadísticamente significativo.

La siguiente tabla del libro de Wooldridge (2010) resume de manera clara gran parte de lo que vimos hasta aquí:

Table 15.4
Summary of Features of Models and Estimation Methods for Unobserved Effects Binary Response Models

Model, Estimation Method	$P(y_{it}=1 \mid \mathbf{x}_{it}, c_i)$ Bounded in (0, 1)?	Restricts $D(c_i \mid \mathbf{x}_i)$?	Idiosyncratic Serial Dependence?	Partial Effects at $E(c_i)$?	APEs?
RE probit, MLE	Yes	Yes (independence, normal)	No	Yes	Yes
RE probit, pooled MLE	Yes	Yes (independence, normal)	Yes	No	Yes
RE probit, GEE	Yes	Yes (independence, normal)	Yes	No	Yes
Chamberlain's RE probit, MLE	Yes	Yes (linear mean, normal)	No	Yes	Yes
Chamberlain's RE probit, pooled MLE	Yes	Yes (linear mean, normal)	Yes	No	Yes
Chamberlain's RE probit, GEE	Yes	Yes (linear mean, normal)	Yes	No	Yes
LPM, within	No	No	Yes	Yes	Yes
FE logit, MLE	Yes	No	No	No	No

2. Considere los datos del ejercicio previo para analizar la probabilidad de estar afiliado a un sindicato según la situación de afiliación sindical del año previo.

   a) Use Pooled Probit para estimar el modelo:

   $$P\left(union_{it} = 1 \mid union_{it-1}\right) = \Phi\left(\psi + \rho \cdot union_{it-1}\right) \tag{8}$$

   A continuación, obtenga una estimación para $P\left(union_{it} = 1 \mid union_{it-1} = 1\right)$ y para $P\left(union_{it} = 1 \mid u$
   Comente sobre el efecto marginal de estar afiliado a un sindicato en el año $t-1$ en la probabilidad de estar afiliado a un sindicato en el año t.

   Solution: Se estima que:

   - La probabilidad estar afiliado a un sindicato dado que estaba afiliado en el período anterior es 0.727
   - La probabilidad estar afiliado a un sindicato dado que no estaba afiliado en el período anterior es 0.089

   Luego, el cambio marginal de estar afiliado a un sindicato en el año $t-1$ resulta en un aumento en la probabilidad de estar afiliado a un sindicato en el año t de $0{,}727 - 0{,}089 = 0{,}638$.

   b) Adicione al modelo el conjunto completo de variables binarias temporales. Vuelva a estimar las probabilidades solicitadas para cada año de la muestra.

   Solution: Se estima que:

   - La probabilidad estar afiliado a un sindicato en el año 1982 dado que estaba afiliado en el período anterior es 0.748
   - La probabilidad estar afiliado a un sindicato en el año 1982 dado que no estaba afiliado en el período anterior es 0.097
   - La probabilidad estar afiliado a un sindicato en el año 1983 dado que estaba afiliado en el período anterior es 0.717
   - La probabilidad estar afiliado a un sindicato en el año 1983 dado que no estaba afiliado en el período anterior es 0.082
   - La probabilidad estar afiliado a un sindicato en el año 1984 dado que estaba afiliado en el período anterior es 0.736
   - La probabilidad estar afiliado a un sindicato en el año 1984 dado que no estaba afiliado en el período anterior es 0.090

- La probabilidad estar afiliado a un sindicato en el año 1985 dado que estaba afiliado en el período anterior es 0.680
- La probabilidad estar afiliado a un sindicato en el año 1985 dado que no estaba afiliado en el período anterior es 0.067
- La probabilidad estar afiliado a un sindicato en el año 1986 dado que estaba afiliado en el período anterior es 0.687
- La probabilidad estar afiliado a un sindicato en el año 1986 dado que no estaba afiliado en el período anterior es 0.069
- La probabilidad estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el período anterior es 0.788
- La probabilidad estar afiliado a un sindicato en el año 1987 dado que no estaba afiliado en el período anterior es 0.121

c) Estime un modelo de efectos no observables dinámico. Use el modelo Probit de efectos aleatorios incluyendo $union_{i,80}$ como una variable explicativa adicional. Luego, promedie las probabilidades estimadas a lo largo de $union_{i,80}$ para obtener la probabilidad promedio de estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el período anterior.

Solution: Se nos solicita computar:

$$N^{-1} \sum_{i=1}^{N} \left\{ \Phi \left[\frac{\left(\hat{\psi} + \hat{\rho} + \hat{\xi} union_{i,80} + \hat{\delta}_{87} \right)}{(1 + \hat{\sigma}_a^2)^{1/2}} \right] \right\}$$

La probabilidad promedio de estar afiliado a un sindicato en el año 1987 dado que estaba afiliado en el año previo es igual a 0.397.