

Trabajo Práctico N° 4: **Modelos Lineales en Paneles Desbalanceados.**

Ejercicio 1.

Utilizar 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, considerar la siguiente ecuación de salarios:

$$\ln(\text{wage}_{it}) = \beta_0 + \beta_1 \text{exper}_{it} + \beta_2 \text{educ}_{it} + c_i + u_{it}, t = 1, 2, \dots, T, \quad (1)$$

donde $\ln(\text{wage}_{it})$ 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. Responder las siguientes preguntas:

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

FE:

```
Fixed-effects (within) regression      Number of obs   =      5,837
Group variable: id                   Number of groups =      1,531

R-squared:                            Obs per group:
    Within = 0.2373                     min =          1
    Between = 0.1857                    avg =         3.8
    Overall = 0.1767                    max =          7

corr(u_i, Xb) = -0.3197                F(2,4304)       =      669.73
                                      Prob > F          =      0.0000
```

```
-----+-----
      lwage | Coefficient  Std. err.      t    P>|t|    [95% conf. interval]
-----+-----
      exper |   .0964067   .0027721    34.78  0.000   .0909721   .1018414
       educ |   .1697764   .0243797     6.96  0.000   .1219795   .2175732
       _cons |   7.270616   .295975    24.56  0.000   6.690352   7.850879
-----+-----
      sigma_u |   .45083563
      sigma_e |   .31573611
       rho    |   .67092951   (fraction of variance due to u_i)
-----+-----
F test that all u_i=0: F(1530, 4304) = 5.46                Prob > F = 0.0000
```

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. Por lo tanto, si se piensa que, efectivamente, lo anterior se cumple y que se están 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) Implementar el contraste de sesgo de selección propuesto por Wooldridge (1995) bajo el enfoque de Mundlak (1978).

Se rechaza la hipótesis nula, por lo que existe evidencia suficiente de que hay sesgo de selección.

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

Se rechaza la hipótesis nula, por lo que existe evidencia suficiente de que hay sesgo de selección.

Ejercicio 2.

Considerando, nuevamente, la ecuación de salarios del ejercicio previo, realizar los siguientes procedimientos:

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

POLS (Chamberlain):

Source	SS	df	MS	Number of obs	=	5,837
				F(15, 5821)	=	111.85
Model	357.875358	15	23.8583572	Prob > F	=	0.0000
Residual	1241.62941	5,821	.213301737	R-squared	=	0.2237
				Adj R-squared	=	0.2217
Total	1599.50477	5,836	.274075526	Root MSE	=	.46185

lwage	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
exper	.0566567	.0130844	4.33	0.000	.0310064	.0823071
educ	.101945	.0187839	5.43	0.000	.0651215	.1387684
exper81	.1106423	.0330149	3.35	0.001	.0459208	.1753637
educ81	.0986618	.0204867	4.82	0.000	.0585003	.1388233
exper82	-.0102639	.0284707	-0.36	0.718	-.0660771	.0455492
educ82	-.2174353	.0366788	-5.93	0.000	-.2893393	-.1455313
exper83	-.0704014	.0284495	-2.47	0.013	-.1261731	-.0146298
educ83	.0856036	.0363551	2.35	0.019	.0143341	.1568731
year#c.lambdac_chamberlain						
81	-.2683315	.0873923	-3.07	0.002	-.4396528	-.0970101
82	-.3211014	.0890068	-3.61	0.000	-.4955878	-.1466151
83	-.3500805	.0833289	-4.20	0.000	-.5134361	-.1867249
84	-.3390861	.0853207	-3.97	0.000	-.5063464	-.1718258
85	-.3585597	.0887959	-4.04	0.000	-.5326328	-.1844867
86	-.3195615	.09277	-3.44	0.001	-.5014252	-.1376977
87	-.3590845	.0978012	-3.67	0.000	-.5508112	-.1673579
_cons	8.87783	.1514153	58.63	0.000	8.580999	9.17466

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

POLS (Mundlak):

Source	SS	df	MS	Number of obs	=	5,837
Model	354.236597	11	32.203327	F(11, 5825)	=	150.64
Residual	1245.26817	5,825	.213779943	Prob > F	=	0.0000
				R-squared	=	0.2215
				Adj R-squared	=	0.2200
Total	1599.50477	5,836	.274075526	Root MSE	=	.46236

lwage	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
exper	.0678465	.007649	8.87	0.000	.0528517	.0828413
educ	.0881093	.0296264	2.97	0.003	.0300305	.146188
mean_exper	-.0275146	.0156551	-1.76	0.079	-.0582044	.0031753
mean_educ	-.0167755	.0291599	-0.58	0.565	-.0739398	.0403887
year#c.lambdas_mundlak						
81	-.2756566	.0778676	-3.54	0.000	-.4283059	-.1230073
82	-.3239156	.0765798	-4.23	0.000	-.4740404	-.1737908
83	-.3637717	.0737034	-4.94	0.000	-.5082579	-.2192856
84	-.3154769	.0727725	-4.34	0.000	-.4581381	-.1728158
85	-.3101603	.073666	-4.21	0.000	-.454573	-.1657477
86	-.2515261	.074834	-3.36	0.001	-.3982286	-.1048236
87	-.2525642	.0770484	-3.28	0.001	-.4036077	-.1015207
_cons	8.781247	.1284501	68.36	0.000	8.529437	9.033056

(c) *Comentar sobre los errores estándar de las estimaciones anteriores.*

La varianza asintótica de los estimadores de la segunda etapa necesita ser corregida por heterocedasticidad y correlación serial arbitraria, así como, además, por la estimación de la primera etapa.

(d) *Estimar los errores estándar vía bootstrapping.*

Stata.

(e) *Estimar los errores estándar analíticos (varianza asintótica).*

Stata.