**Trabajo Práctico N° 2:**

**Modelos de Datos de Panel Lineales.**

**Ejercicio 1.**

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

*= + + + + + + + + .*

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

Stata.

**(b)** *Aplicar la transformación within al modelo. Luego, estimar el modelo transformado por POLS.*

POLS:

Source | SS df MS Number of obs = 630

-------------+---------------------------------- F(11, 619) = 43.19

Model | 7.81221835 11 .710201668 Prob > F = 0.0000

Residual | 10.1785214 619 .016443492 R-squared = 0.4342

-------------+---------------------------------- Adj R-squared = 0.4242

Total | 17.9907397 630 .02855673 Root MSE = .12823

---------------------------------------------------------------------------------

within\_lcrmrte | Coefficient Std. err. t P>|t| [95% conf. interval]

----------------+----------------------------------------------------------------

within\_lprbarr | -.3597944 .0299699 -12.01 0.000 -.4186493 -.3009395

within\_lprbconv | -.2858733 .0196143 -14.57 0.000 -.3243919 -.2473547

within\_lprbpris | -.1827812 .0300086 -6.09 0.000 -.2417122 -.1238502

within\_lavgsen | -.0044879 .024449 -0.18 0.854 -.0525009 .043525

within\_lpolpc | .4241142 .0243741 17.40 0.000 .3762483 .4719802

within\_d82 | .0125802 .0199141 0.63 0.528 -.0265271 .0516875

within\_d83 | -.0792813 .0197277 -4.02 0.000 -.1180225 -.04054

within\_d84 | -.1177281 .0199815 -5.89 0.000 -.1569678 -.0784884

within\_d85 | -.1119561 .0201954 -5.54 0.000 -.151616 -.0722962

within\_d86 | -.0818268 .0198078 -4.13 0.000 -.1207254 -.0429282

within\_d87 | -.0404704 .0194497 -2.08 0.038 -.0786657 -.0022751

---------------------------------------------------------------------------------

**(c)** *Comentar sobre la validez de los errores estándar del inciso previo.*

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 . Por consiguiente, excepto que T sea lo suficientemente grande, se necesita corregir este denominador.

**(d)** *Utilizar el comando xtreg para estimar, nuevamente, el modelo usando efectos fijos.*

FE:

Fixed-effects (within) regression Number of obs = 630

Group variable: county Number of groups = 90

R-squared: Obs per group:

Within = 0.4342 min = 7

Between = 0.4066 avg = 7.0

Overall = 0.4042 max = 7

F(11,89) = 11.49

corr(u\_i, Xb) = 0.2068 Prob > F = 0.0000

(Std. err. adjusted for 90 clusters in county)

------------------------------------------------------------------------------

| Robust

lcrmrte | Coefficient std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

lprbarr | -.3597944 .0594678 -6.05 0.000 -.4779557 -.2416332

lprbconv | -.2858733 .051522 -5.55 0.000 -.3882464 -.1835001

lprbpris | -.1827812 .0452811 -4.04 0.000 -.2727538 -.0928085

lavgsen | -.0044879 .0333499 -0.13 0.893 -.0707535 .0617777

lpolpc | .4241142 .0849052 5.00 0.000 .2554095 .592819

d82 | .0125802 .0160066 0.79 0.434 -.0192246 .044385

d83 | -.0792813 .0195639 -4.05 0.000 -.1181544 -.0404081

d84 | -.1177281 .0217118 -5.42 0.000 -.160869 -.0745872

d85 | -.1119561 .0256583 -4.36 0.000 -.1629386 -.0609736

d86 | -.0818268 .0236276 -3.46 0.001 -.1287745 -.0348792

d87 | -.0404704 .0241765 -1.67 0.098 -.0885087 .0075678

\_cons | -1.604135 .5102062 -3.14 0.002 -2.617904 -.5903664

-------------+----------------------------------------------------------------

sigma\_u | .43487416

sigma\_e | .13871215

rho | .90765322 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

**(e)** *Estimar el modelo usando diferencias finitas de primer orden.*

FD:

Source | SS df MS Number of obs = 540

-------------+---------------------------------- F(11, 529) = 36.66

Model | 9.60258308 11 .872962098 Prob > F = 0.0000

Residual | 12.5963755 529 .023811674 R-squared = 0.4326

-------------+---------------------------------- Adj R-squared = 0.4208

Total | 22.1989586 540 .041109183 Root MSE = .15431

------------------------------------------------------------------------------

D.lcrmrte | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

lprbarr |

D1. | -.3274942 .0299801 -10.92 0.000 -.3863889 -.2685995

|

lprbconv |

D1. | -.2381066 .0182341 -13.06 0.000 -.2739268 -.2022864

|

lprbpris |

D1. | -.1650462 .025969 -6.36 0.000 -.2160613 -.1140312

|

lavgsen |

D1. | -.0217607 .0220909 -0.99 0.325 -.0651574 .021636

|

lpolpc |

D1. | .3984264 .026882 14.82 0.000 .3456177 .451235

|

d82 |

D1. | .0077134 .0170579 0.45 0.651 -.0257961 .0412229

|

d83 |

D1. | -.0844391 .0234564 -3.60 0.000 -.1305182 -.03836

|

d84 |

D1. | -.1246632 .0287464 -4.34 0.000 -.1811344 -.068192

|

d85 |

D1. | -.121561 .03315 -3.67 0.000 -.1866827 -.0564392

|

d86 |

D1. | -.0863333 .0366763 -2.35 0.019 -.1583823 -.0142842

|

d87 |

D1. | -.0377932 .0399728 -0.95 0.345 -.116318 .0407316

------------------------------------------------------------------------------

**Ejercicio 2.**

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

*= + + + + + ,*

*Notar que se omitió la dummy temporal para el año 1987. Interpretar los resultados.*

POLS:

Source | SS df MS Number of obs = 153

-------------+---------------------------------- F(4, 148) = 3.05

Model | 977.390644 4 244.347661 Prob > F = 0.0190

Residual | 11867.9475 148 80.1888343 R-squared = 0.0761

-------------+---------------------------------- Adj R-squared = 0.0511

Total | 12845.3381 152 84.5088034 Root MSE = 8.9548

------------------------------------------------------------------------------

mrdrte | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

exec | .1627547 .1939295 0.84 0.403 -.2204738 .5459832

unem | 1.390786 .4508653 3.08 0.002 .4998207 2.281751

d90 | 2.675335 1.816934 1.47 0.143 -.91515 6.26582

d93 | 1.607317 1.774768 0.91 0.367 -1.899842 5.114476

\_cons | -1.864393 3.069517 -0.61 0.545 -7.930134 4.201349

------------------------------------------------------------------------------

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

Tener en consideración los efectos temporales agregados en el modelo 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 se incluyen estas variables, se debe 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, considerar la siguiente modificación en el modelo:*

*= + + + + + + ,*

*donde es un efecto individual por estado. Estimar la ecuación usando efectos fijos.*

FE:

Fixed-effects (within) regression Number of obs = 153

Group variable: id Number of groups = 51

R-squared: Obs per group:

Within = 0.0734 min = 3

Between = 0.0037 avg = 3.0

Overall = 0.0108 max = 3

F(4,98) = 1.94

corr(u\_i, Xb) = 0.0010 Prob > F = 0.1098

------------------------------------------------------------------------------

mrdrte | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

exec | -.1383231 .1770059 -0.78 0.436 -.4895856 .2129395

unem | .2213158 .2963756 0.75 0.457 -.366832 .8094636

d90 | 1.556215 .7453273 2.09 0.039 .0771369 3.035293

d93 | 1.733242 .7004381 2.47 0.015 .3432454 3.123239

\_cons | 5.822104 1.915611 3.04 0.003 2.020636 9.623572

-------------+----------------------------------------------------------------

sigma\_u | 8.7527226

sigma\_e | 3.5214244

rho | .86068589 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

F test that all u\_i=0: F(50, 98) = 17.18 Prob > F = 0.0000

**(d)** *Repetir la estimación del inciso previo usando diferencias finitas de primer orden.*

FD:

Source | SS df MS Number of obs = 102

-------------+---------------------------------- F(4, 98) = 1.61

Model | 119.103298 4 29.7758244 Prob > F = 0.1778

Residual | 1812.28656 98 18.49272 R-squared = 0.0617

-------------+---------------------------------- Adj R-squared = 0.0234

Total | 1931.38986 102 18.9351947 Root MSE = 4.3003

------------------------------------------------------------------------------

cmrdrte | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

cexec | -.1150682 .1473871 -0.78 0.437 -.407553 .1774166

cunem | .1630854 .3079049 0.53 0.598 -.4479419 .7741126

cd90 | 1.51099 .6608967 2.29 0.024 .1994623 2.822518

cd93 | 1.725263 .8533453 2.02 0.046 .0318275 3.418699

------------------------------------------------------------------------------

**(e)** *Brindar un ejemplo bajo el cual la variable de ejecuciones no sería, estrictamente, exógena (condicional en ). Observación: Para obtener estimaciones consistentes, el modelo de efectos fijos asume exogeneidad estricta de las variables explicativas condicionadas en .*

Un ejemplo bajo el cual la variable de ejecuciones () no sería estrictamente exógena (condicional en ) podría ser 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 la tasa 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 la tasa de homicidios, eso podría violar la exogeneidad estricta. Con una serie temporal más larga, se podría añadir (e, incluso, valores de un futuro más lejano) y estimar la ecuación por FE, comprobando la significatividad estadística de la variable . En el caso de que se encuentre que esta variable es estadísticamente significativa, se tendría evidencia a favor de que no se cumple el supuesto de exogeneidad estricta.

**(f)** *Repetir la estimación del inciso (c) usando el estimador de GLS para diferencias finitas de primer orden. Comprobar que los coeficientes estimados son iguales a los obtenidos por FE.*

bfdgls[4,1]

mrdrte

exec -.13832306

unem .22131582

d90 1.5562147

d93 1.7332421

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

RE:

Random-effects GLS regression Number of obs = 153

Group variable: id Number of groups = 51

R-squared: Obs per group:

Within = 0.0680 min = 3

Between = 0.0731 avg = 3.0

Overall = 0.0426 max = 3

Wald chi2(4) = 8.52

corr(u\_i, X) = 0 (assumed) Prob > chi2 = 0.0743

------------------------------------------------------------------------------

mrdrte | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

exec | -.0543375 .1595008 -0.34 0.733 -.3669533 .2582784

unem | .3947507 .2848133 1.39 0.166 -.1634732 .9529745

d90 | 1.732981 .7478556 2.32 0.020 .2672106 3.19875

d93 | 1.699913 .7065606 2.41 0.016 .3150796 3.084746

\_cons | 4.635132 2.179451 2.13 0.033 .3634863 8.906778

-------------+----------------------------------------------------------------

sigma\_u | 8.2056677

sigma\_e | 3.5214244

rho | .84447636 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

---- Coefficients ----

| (b) (B) (b-B) sqrt(diag(V\_b-V\_B))

| est\_fe est\_re Difference Std. err.

-------------+----------------------------------------------------------------

exec | -.1383231 -.0543375 -.0839856 .0767503

unem | .2213158 .3947507 -.1734349 .0819749

d90 | 1.556215 1.732981 -.1767658 .

d93 | 1.733242 1.699913 .0333292 .

------------------------------------------------------------------------------

b = Consistent under H0 and Ha; obtained from xtreg.

B = Inconsistent under Ha, efficient under H0; obtained from xtreg.

Test of H0: Difference in coefficients not systematic

chi2(4) = (b-B)'[(V\_b-V\_B)^(-1)](b-B)

= 5.78

Prob > chi2 = 0.2165

(V\_b-V\_B is not positive definite)

Por lo tanto, se puede observar que, considerando un nivel de significación del 10%, el mejor estimador es el de efectos fijos, ya que se rechaza la hipótesis nula de no correlación entre los regresores y los efectos fijos, por lo que el estimador de efectos aleatorios no es consistente.

**Ejercicio 3.**

*Considerar el siguiente modelo:*

*= + + , i= 1, 2, … , N; t= 1, 2, … , T,*

*donde (0, 1), (0, ), (0, ) y para todo i, t. Suponer = = = 1 y T= 10. La idea es realizar experimentos de Monte Carlo para evaluar la eficiencia de distintos estimadores de .*

**(a)** *Caso 1: N= 5. Realizar un experimento de Monte Carlo con 1000 simulaciones. Reportar media, desvío estándar y RMSE de la estimación de usando: POLS, RE y FE.*

**(b)** *Repetir el punto anterior con N= 10, 30, 50, 100 y 500.*

N\_5 N\_10 N\_30 N\_50 N\_100 N\_500

media\_beta~s .99804484 .99234775 1.0019568 1.0039147 1.0005582 .99963949

de\_beta\_pols .19410442 .14170935 .07988537 .06148401 .04521915 .02023065

rmse\_beta\_~s .1940172 .14184504 .0798694 .06157782 .04519998 .02022375

media\_beta~e .99311721 .99036578 1.0021895 1.0027847 1.0005854 .99967618

de\_beta\_fe .15215654 .10712643 .06006592 .04603343 .03388416 .01449802

rmse\_beta\_fe .15223612 .10750542 .06007579 .0460946 .03387227 .01449439

media\_beta~e .99239395 .99075022 1.0022435 1.0029077 1.0005371 .99967278

de\_beta\_re .15558624 .10656719 .05945567 .04579737 .0335924 .01453039

rmse\_beta\_re .15569433 .10691477 .05946827 .04586672 .03357989 .01452681

**(c)** *Comentar los resultados obtenidos y su conclusión de qué estimador debiera utilizarse en la práctica.*

En primer lugar, es importante destacar que, dados los supuestos del modelo, los tres estimadores en consideración son consistentes. Por lo tanto, se debería esperar que, a medida que el tamaño muestral aumenta, la media de las estimaciones de con los diferentes estimadores estén cerca del valor poblacional (= 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 cómo 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, se trabajara con un modelo donde se supone que se cumplen los supuestos del modelo del inciso, entonces, para N muy pequeño se optaría por utilizar el estimador FE, mientras que, a partir de N= 30, se optaría por el estimador RE.

**Ejercicio 4.**

*Basado en el Ejercicio 10.18 de Wooldridge (2010). Utilizar 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. Estimar el modelo por POLS, RE, FE y FD. ¿Qué se puede concluir acerca de los coeficientes de las variables dummy?*

POLS:

Source | SS df MS Number of obs = 4,360

-------------+---------------------------------- F(7, 4352) = 50.54

Model | 92.9668229 7 13.2809747 Prob > F = 0.0000

Residual | 1143.56282 4,352 .262767192 R-squared = 0.0752

-------------+---------------------------------- Adj R-squared = 0.0737

Total | 1236.52964 4,359 .283672779 Root MSE = .51261

------------------------------------------------------------------------------

lwage | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .0310529 3.84 0.000 .0585107 .1802697

d82 | .1781901 .0310529 5.74 0.000 .1173106 .2390696

d83 | .2257865 .0310529 7.27 0.000 .1649069 .286666

d84 | .2968181 .0310529 9.56 0.000 .2359386 .3576976

d85 | .3459333 .0310529 11.14 0.000 .2850538 .4068128

d86 | .4062418 .0310529 13.08 0.000 .3453623 .4671213

d87 | .4730023 .0310529 15.23 0.000 .4121228 .5338818

\_cons | 1.393477 .0219577 63.46 0.000 1.350429 1.436525

------------------------------------------------------------------------------

RE:

Random-effects GLS regression Number of obs = 4,360

Group variable: nr Number of groups = 545

R-squared: Obs per group:

Within = 0.0000 min = 8

Between = 0.0000 avg = 8.0

Overall = 0.0752 max = 8

Wald chi2(7) = 738.94

corr(u\_i, X) = 0 (assumed) Prob > chi2 = 0.0000

------------------------------------------------------------------------------

lwage | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .021487 5.56 0.000 .0772765 .1615039

d82 | .1781901 .021487 8.29 0.000 .1360764 .2203038

d83 | .2257865 .021487 10.51 0.000 .1836728 .2679001

d84 | .2968181 .021487 13.81 0.000 .2547044 .3389318

d85 | .3459333 .021487 16.10 0.000 .3038196 .388047

d86 | .4062418 .021487 18.91 0.000 .3641281 .4483555

d87 | .4730023 .021487 22.01 0.000 .4308886 .515116

\_cons | 1.393477 .0219577 63.46 0.000 1.350441 1.436513

-------------+----------------------------------------------------------------

sigma\_u | .37007665

sigma\_e | .35469771

rho | .52120938 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

FE:

Fixed-effects (within) regression Number of obs = 4,360

Group variable: nr Number of groups = 545

R-squared: Obs per group:

Within = 0.1625 min = 8

Between = . avg = 8.0

Overall = 0.0752 max = 8

F(7,3808) = 105.56

corr(u\_i, Xb) = 0.0000 Prob > F = 0.0000

------------------------------------------------------------------------------

lwage | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .021487 5.56 0.000 .0772631 .1615173

d82 | .1781901 .021487 8.29 0.000 .136063 .2203172

d83 | .2257865 .021487 10.51 0.000 .1836594 .2679135

d84 | .2968181 .021487 13.81 0.000 .254691 .3389452

d85 | .3459333 .021487 16.10 0.000 .3038063 .3880604

d86 | .4062418 .021487 18.91 0.000 .3641147 .4483688

d87 | .4730023 .021487 22.01 0.000 .4308753 .5151294

\_cons | 1.393477 .0151936 91.71 0.000 1.363689 1.423265

-------------+----------------------------------------------------------------

sigma\_u | .39074676

sigma\_e | .35469771

rho | .54824631 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

F test that all u\_i=0: F(544, 3808) = 9.71 Prob > F = 0.0000

FD:

Source | SS df MS Number of obs = 3,815

-------------+---------------------------------- F(7, 3808) = 14.06

Model | 19.3631642 7 2.76616631 Prob > F = 0.0000

Residual | 749.249837 3,808 .196756785 R-squared = 0.0252

-------------+---------------------------------- Adj R-squared = 0.0234

Total | 768.613001 3,815 .201471298 Root MSE = .44357

------------------------------------------------------------------------------

D.lwage | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 |

D1. | .1193902 .0190006 6.28 0.000 .0821379 .1566425

|

d82 |

D1. | .1781901 .0268709 6.63 0.000 .1255074 .2308728

|

d83 |

D1. | .2257865 .03291 6.86 0.000 .1612636 .2903093

|

d84 |

D1. | .2968181 .0380011 7.81 0.000 .2223136 .3713226

|

d85 |

D1. | .3459333 .0424866 8.14 0.000 .2626347 .4292319

|

d86 |

D1. | .4062418 .0465417 8.73 0.000 .3149927 .4974908

|

d87 |

D1. | .4730023 .0502708 9.41 0.000 .3744421 .5715626

------------------------------------------------------------------------------

Tabla comparativa:

----------------------------------------------------------------------------

(1) (2) (3) (4)

POLS RE FE FD

----------------------------------------------------------------------------

d81 0.119\*\*\* 0.119\*\*\* 0.119\*\*\* 0.119\*\*\*

(0.0311) (0.0215) (0.0215) (0.0190)

d82 0.178\*\*\* 0.178\*\*\* 0.178\*\*\* 0.178\*\*\*

(0.0311) (0.0215) (0.0215) (0.0269)

d83 0.226\*\*\* 0.226\*\*\* 0.226\*\*\* 0.226\*\*\*

(0.0311) (0.0215) (0.0215) (0.0329)

d84 0.297\*\*\* 0.297\*\*\* 0.297\*\*\* 0.297\*\*\*

(0.0311) (0.0215) (0.0215) (0.0380)

d85 0.346\*\*\* 0.346\*\*\* 0.346\*\*\* 0.346\*\*\*

(0.0311) (0.0215) (0.0215) (0.0425)

d86 0.406\*\*\* 0.406\*\*\* 0.406\*\*\* 0.406\*\*\*

(0.0311) (0.0215) (0.0215) (0.0465)

d87 0.473\*\*\* 0.473\*\*\* 0.473\*\*\* 0.473\*\*\*

(0.0311) (0.0215) (0.0215) (0.0503)

\_cons 1.393\*\*\* 1.393\*\*\* 1.393\*\*\*

(0.0220) (0.0220) (0.0152)

----------------------------------------------------------------------------

N 4360 4360 4360 3815

r2 0.0752 0.163 0.0252

----------------------------------------------------------------------------

Standard errors in parentheses

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Por lo tanto, lo que se puede concluir acerca de los coeficientes de las variables *dummy* es que son numéricamente idénticos.

**(b)** *Añadir las variables constantes en el tiempo educ, black e hisp al modelo, y estimar por POLS y RE. ¿Cómo se comparan los coeficientes? ¿Qué ocurre si se estima la ecuación por FE?*

POLS:

Source | SS df MS Number of obs = 4,360

-------------+---------------------------------- F(10, 4349) = 73.66

Model | 179.091659 10 17.9091659 Prob > F = 0.0000

Residual | 1057.43798 4,349 .243145087 R-squared = 0.1448

-------------+---------------------------------- Adj R-squared = 0.1429

Total | 1236.52964 4,359 .283672779 Root MSE = .4931

------------------------------------------------------------------------------

lwage | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .029871 4.00 0.000 .0608279 .1779526

d82 | .1781901 .029871 5.97 0.000 .1196277 .2367524

d83 | .2257865 .029871 7.56 0.000 .1672241 .2843488

d84 | .2968181 .029871 9.94 0.000 .2382557 .3553804

d85 | .3459333 .029871 11.58 0.000 .287371 .4044957

d86 | .4062418 .029871 13.60 0.000 .3476794 .4648041

d87 | .4730023 .029871 15.83 0.000 .41444 .5315647

educ | .0770943 .0043766 17.62 0.000 .0685139 .0856747

black | -.1225637 .0237021 -5.17 0.000 -.1690319 -.0760955

hisp | .024623 .0213056 1.16 0.248 -.0171468 .0663928

\_cons | .4966384 .0566686 8.76 0.000 .3855391 .6077377

------------------------------------------------------------------------------

RE:

Random-effects GLS regression Number of obs = 4,360

Group variable: nr Number of groups = 545

R-squared: Obs per group:

Within = 0.1625 min = 8

Between = 0.1296 avg = 8.0

Overall = 0.1448 max = 8

Wald chi2(10) = 819.51

corr(u\_i, X) = 0 (assumed) Prob > chi2 = 0.0000

------------------------------------------------------------------------------

lwage | Coefficient Std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .021487 5.56 0.000 .0772765 .1615039

d82 | .1781901 .021487 8.29 0.000 .1360764 .2203038

d83 | .2257865 .021487 10.51 0.000 .1836728 .2679001

d84 | .2968181 .021487 13.81 0.000 .2547044 .3389318

d85 | .3459333 .021487 16.10 0.000 .3038196 .388047

d86 | .4062418 .021487 18.91 0.000 .3641281 .4483555

d87 | .4730023 .021487 22.01 0.000 .4308886 .515116

educ | .0770943 .009177 8.40 0.000 .0591076 .0950809

black | -.1225637 .0496994 -2.47 0.014 -.2199728 -.0251546

hisp | .024623 .0446744 0.55 0.582 -.0629371 .1121831

\_cons | .4966384 .1122718 4.42 0.000 .2765897 .7166871

-------------+----------------------------------------------------------------

sigma\_u | .34337144

sigma\_e | .35469771

rho | .48377912 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

FE:

Fixed-effects (within) regression Number of obs = 4,360

Group variable: nr Number of groups = 545

R-squared: Obs per group:

Within = 0.1625 min = 8

Between = . avg = 8.0

Overall = 0.0752 max = 8

F(7,3808) = 105.56

corr(u\_i, Xb) = 0.0000 Prob > F = 0.0000

------------------------------------------------------------------------------

lwage | Coefficient Std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .021487 5.56 0.000 .0772631 .1615173

d82 | .1781901 .021487 8.29 0.000 .136063 .2203172

d83 | .2257865 .021487 10.51 0.000 .1836594 .2679135

d84 | .2968181 .021487 13.81 0.000 .254691 .3389452

d85 | .3459333 .021487 16.10 0.000 .3038063 .3880604

d86 | .4062418 .021487 18.91 0.000 .3641147 .4483688

d87 | .4730023 .021487 22.01 0.000 .4308753 .5151294

educ | 0 (omitted)

black | 0 (omitted)

hisp | 0 (omitted)

\_cons | 1.393477 .0151936 91.71 0.000 1.363689 1.423265

-------------+----------------------------------------------------------------

sigma\_u | .39074676

sigma\_e | .35469771

rho | .54824631 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

F test that all u\_i=0: F(544, 3808) = 9.71 Prob > F = 0.0000

Tabla comparativa:

------------------------------------------------------------

(1) (2) (3)

POLS RE FE

------------------------------------------------------------

d81 0.119\*\*\* 0.119\*\*\* 0.119\*\*\*

(0.0299) (0.0215) (0.0215)

d82 0.178\*\*\* 0.178\*\*\* 0.178\*\*\*

(0.0299) (0.0215) (0.0215)

d83 0.226\*\*\* 0.226\*\*\* 0.226\*\*\*

(0.0299) (0.0215) (0.0215)

d84 0.297\*\*\* 0.297\*\*\* 0.297\*\*\*

(0.0299) (0.0215) (0.0215)

d85 0.346\*\*\* 0.346\*\*\* 0.346\*\*\*

(0.0299) (0.0215) (0.0215)

d86 0.406\*\*\* 0.406\*\*\* 0.406\*\*\*

(0.0299) (0.0215) (0.0215)

d87 0.473\*\*\* 0.473\*\*\* 0.473\*\*\*

(0.0299) (0.0215) (0.0215)

educ 0.0771\*\*\* 0.0771\*\*\* 0

(0.00438) (0.00918) (.)

black -0.123\*\*\* -0.123\*\* 0

(0.0237) (0.0497) (.)

hisp 0.0246 0.0246 0

(0.0213) (0.0447) (.)

\_cons 0.497\*\*\* 0.497\*\*\* 1.393\*\*\*

(0.0567) (0.112) (0.0152)

------------------------------------------------------------

N 4360 4360 4360

r2 0.145 0.163

------------------------------------------------------------

Standard errors in parentheses

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Por un lado, se puede observar que las estimaciones de POLS y RE son numéricamente idénticas, ya que, si el modelo incluye sólo efectos temporales agregados y covariables específicas del individuo que no tienen variación temporal, entonces, los coeficientes de POLS son iguales a los de RE.

Por otra parte, lo que ocurre si se estima la ecuación por FE es que los coeficientes asociados a las variables constantes en el tiempo no se pueden estimar y, en consecuencia, cuando se incluyen variables constantes en el tiempo, la estimación de la constante en FE no es igual a la estimación de la constante en POLS/RE.

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

Los errores estándar de POLS y RE del inciso (b) no son iguales. 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 heterogeneidad individual no observable. Los errores estándar de RE, al menos, en su estructura estándar, permiten 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 de POLS.

**(d)** *Obtener los errores estándar robustos para POLS. ¿Son preferibles estos o los errores estándar habituales de RE?*

POLS (con errores estándar robustos):

Linear regression Number of obs = 4,360

F(10, 544) = 49.41

Prob > F = 0.0000

R-squared = 0.1448

Root MSE = .4931

(Std. err. adjusted for 545 clusters in nr)

------------------------------------------------------------------------------

| Robust

lwage | Coefficient std. err. t P>|t| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .0244086 4.89 0.000 .0714435 .1673369

d82 | .1781901 .0241987 7.36 0.000 .1306558 .2257243

d83 | .2257865 .0243796 9.26 0.000 .1778968 .2736761

d84 | .2968181 .0271485 10.93 0.000 .2434894 .3501468

d85 | .3459333 .0263181 13.14 0.000 .2942358 .3976309

d86 | .4062418 .0273064 14.88 0.000 .3526029 .4598807

d87 | .4730023 .025996 18.20 0.000 .4219374 .5240672

educ | .0770943 .0090198 8.55 0.000 .0593763 .0948122

black | -.1225637 .0532662 -2.30 0.022 -.2271964 -.017931

hisp | .024623 .0411235 0.60 0.550 -.0561573 .1054033

\_cons | .4966384 .1097474 4.53 0.000 .2810579 .7122189

------------------------------------------------------------------------------

Estos errores estándar robustos son preferibles a los errores estándar habituales de RE, ya que estos errores estándar robustos permiten cualquier tipo de correlación serial y de heterocedasticidad de los disturbios que varían en el tiempo.

**(e)** *Obtener los errores estándar robustos de RE. ¿Cómo se comparan con los errores estándar robustos de POLS y por qué?*

RE (con errores estándar robustos):

Random-effects GLS regression Number of obs = 4,360

Group variable: nr Number of groups = 545

R-squared: Obs per group:

Within = 0.1625 min = 8

Between = 0.1296 avg = 8.0

Overall = 0.1448 max = 8

Wald chi2(10) = 494.13

corr(u\_i, X) = 0 (assumed) Prob > chi2 = 0.0000

(Std. err. adjusted for 545 clusters in nr)

------------------------------------------------------------------------------

| Robust

lwage | Coefficient std. err. z P>|z| [95% conf. interval]

-------------+----------------------------------------------------------------

d81 | .1193902 .0244086 4.89 0.000 .0715502 .1672302

d82 | .1781901 .0241987 7.36 0.000 .1307616 .2256186

d83 | .2257865 .0243796 9.26 0.000 .1780033 .2735696

d84 | .2968181 .0271485 10.93 0.000 .2436081 .3500281

d85 | .3459333 .0263181 13.14 0.000 .2943508 .3975159

d86 | .4062418 .0273064 14.88 0.000 .3527222 .4597613

d87 | .4730023 .025996 18.20 0.000 .422051 .5239536

educ | .0770943 .0090198 8.55 0.000 .0594157 .0947728

black | -.1225637 .0532662 -2.30 0.021 -.2269636 -.0181638

hisp | .024623 .0411235 0.60 0.549 -.0559775 .1052236

\_cons | .4966384 .1097474 4.53 0.000 .2815375 .7117392

-------------+----------------------------------------------------------------

sigma\_u | .34337144

sigma\_e | .35469771

rho | .48377912 (fraction of variance due to u\_i)

------------------------------------------------------------------------------

Tabla comparativa:

--------------------------------------------

(1) (2)

POLS (robu~) RE (robust)

--------------------------------------------

d81 0.119\*\*\* 0.119\*\*\*

(0.0244) (0.0244)

d82 0.178\*\*\* 0.178\*\*\*

(0.0242) (0.0242)

d83 0.226\*\*\* 0.226\*\*\*

(0.0244) (0.0244)

d84 0.297\*\*\* 0.297\*\*\*

(0.0271) (0.0271)

d85 0.346\*\*\* 0.346\*\*\*

(0.0263) (0.0263)

d86 0.406\*\*\* 0.406\*\*\*

(0.0273) (0.0273)

d87 0.473\*\*\* 0.473\*\*\*

(0.0260) (0.0260)

educ 0.0771\*\*\* 0.0771\*\*\*

(0.00902) (0.00902)

black -0.123\*\* -0.123\*\*

(0.0533) (0.0533)

hisp 0.0246 0.0246

(0.0411) (0.0411)

\_cons 0.497\*\*\* 0.497\*\*\*

(0.110) (0.110)

--------------------------------------------

N 4360 4360

r2 0.145

--------------------------------------------

Standard errors in parentheses

\* p<0.10, \*\* p<0.05, \*\*\* p<0.01

Por lo tanto, se puede observar que estos errores estándar son numéricamente idénticos a los errores estándar robustos de POLS porque se tiene un solo estimador y, entonces, hay una sola varianza robusta.