

Ejercicio 3

Curso: Econometría II

Profesor: Mauricio Tejada

Estudiantes: Dania Bustamante, Rosana Cardona, José Casanova

15 mayo 2024

Cornwell y Trumbull (1994) utilizaron datos de 90 condados en Carolina del Norte, de los años 1981 a 1987, para estimar un modelo de efectos inobservables de la delincuencia. Diversos factores, entre los que se cuentan la ubicación geográfica, las actitudes hacia la delincuencia, los registros históricos y las normas de denuncia del crimen, podrían estar contenidos en el efecto no observable a_i . El modelo propuesto para explicar el efecto disuasivo de la justicia penal sobre la tasa de criminalidad es el siguiente:

donde i es el condado y t es el año.

Las variables de salario se incluyen en el modelo para controlar por la situación económica y por la importancia de los sectores económicos a nivel de condado. El número de policías (por habitante) se incluye para controlar por el efecto del control policial.

Pregunta 1

Estime el modelo por el método de mínimos cuadrados agrupados (pooled MCO). Estamos interesados en las tres variables de probabilidad (arresto, condena y tiempo cumplido en prisión) que son las variables asociadas a la justicia penal. ¿Por qué es probable que los estimadores obtenidos sean sesgados?

Pooling Model

Call:

```
plm(formula = log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) +  
      log(avgsen) + log(polpc) + log(wcon) + log(wtuc) + log(wtrd) +  
      log(wfir) + log(wser) + log(wmfg) + log(wfed) + log(wsta) +  
      log(wloc), data = panel, model = "pooling")
```

Balanced Panel: $n = 90$, $T = 7$, $N = 630$

Residuals:

Min.	1st Qu.	Median	3rd Qu.	Max.
-1.825262	-0.192315	0.015372	0.224594	1.305420

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)	
(Intercept)	-5.647995	0.727636	-7.7621	0.00000000000003498	***
log(prbarr)	-0.677261	0.036368	-18.6226	< 0.0000000000000022	***
log(prbconv)	-0.513053	0.026102	-19.6555	< 0.0000000000000022	***
log(prbpris)	0.123758	0.064823	1.9092	0.0567052	.
log(avgsen)	-0.063690	0.053146	-1.1984	0.2312252	
log(polpc)	0.334638	0.030089	11.1217	< 0.0000000000000022	***
log(wcon)	-0.031310	0.075292	-0.4158	0.6776698	
log(wtuc)	-0.068225	0.040288	-1.6934	0.0908801	.
log(wtrd)	0.201556	0.084834	2.3759	0.0178118	*
log(wfir)	0.033159	0.062842	0.5277	0.5979298	
log(wser)	-0.022411	0.043355	-0.5169	0.6053971	
log(wmfg)	-0.160275	0.072882	-2.1991	0.0282420	*
log(wfed)	0.870285	0.141457	6.1523	0.00000000137643680	***
log(wsta)	-0.420231	0.109188	-3.8487	0.0001312	***
log(wloc)	0.138523	0.159416	0.8689	0.3852208	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 206.38

Residual Sum of Squares: 78.749

R-Squared: 0.61843

Adj. R-Squared: 0.60974

F-statistic: 71.1972 on 14 and 615 DF, p-value: < 0.00000000000000222

- $\hat{\beta}_1 \log(prbarr)$: Un aumento de 1% en la probabilidad de arresto en disminuye en promedio 0.677% la tasa de criminalidad, manteniendo todo lo demás constante.
- $\hat{\beta}_2 \log(prbconv)$: Un aumento de 1% en la probabilidad de condena disminuye en promedio un 0.513% la tasa de criminalidad, manteniendo todo lo demás constante.
- $\hat{\beta}_3 \log(prbconv)$: Un aumento de 1% en la probabilidad de sentencia de prisión aumenta en promedio 0.124% la tasa de criminalidad, manteniendo todo lo demás constante.

Los estimadores obtenidos a través del minimos cuadrados agrupados podrían tener sesgo debido a varios factores:

- **Endogeneidad:** Las variables de la probabilidad de arresto, condena y tiempo cumplido en prisión pueden tener correlación con el término de error. Los condados con altas tasas de criminalidad pueden tener más recursos u opciones para la aplicación de la ley, lo que aumentaría la probabilidad de ser arrestado o tener algún tipo de condena. Lo anterior podría generar una correlación positiva entre la probabilidad de arresto y el término de error, generando un sesgo en el estimador para la probabilidad de arresto.

- **Efectos fijos no observados:** El modelo tiene un término de características o cualidades no observadas (a_i), que puede capturar elementos inobservables y constantes en el tiempo de cada lugar. Si estas cualidades tienen correlación con las variables de probabilidad usadas, los estimadores estarán sesgados. Por ejemplo, si los condados con leyes más estrictas hacia el crimen tienen una mayor probabilidad de ser arrestado, el parámetro de $\log(\text{prbarr})$ y el término a_i estarán correlacionados.
- **Variables omitidas:** Si existen variables relevantes que no se incluyen en el modelo, esto genera sesgo en los estimadores. Por ejemplo, si la tasa de desempleo afecta potencialmente a la tasa de criminalidad, pero no se incluye en el modelo, esto puede sesgar los estimadores de las variables incluidas en el modelo.

Pregunta 2:

Estime ahora el modelo usando el método de primeras diferencias. ¿Existe un cambio notable (respecto a lo hallado en 1) en el signo o en la magnitud de los coeficientes de las variables asociadas a la justicia penal? ¿Tiene un efecto disuasivo la justicia penal? ¿Son los efectos estadísticamente significativos individualmente? Interprete los resultados.

Oneway (individual) effect First-Difference Model

Call:

```
plm(formula = log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) +
    log(avgsen) + log(polpc) + log(wcon) + log(wtuc) + log(wtrd) +
    log(wfir) + log(wser) + log(wmfg) + log(wfed) + log(wsta) +
    log(wloc), data = panel, model = "fd")
```

Balanced Panel: n = 90, T = 7, N = 630

Observations used in estimation: 540

Residuals:

	Min.	1st Qu.	Median	3rd Qu.	Max.
	-0.6660817	-0.0807718	0.0043541	0.0825722	0.6814652

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)
(Intercept)	0.0013449	0.0137444	0.0978	0.92209
$\log(\text{prbarr})$	-0.3450694	0.0306502	-11.2583	< 0.00000000000000022 ***
$\log(\text{prbconv})$	-0.2524166	0.0185458	-13.6105	< 0.00000000000000022 ***
$\log(\text{prbpris})$	-0.1769305	0.0265786	-6.6569	0.00000000007059 ***
$\log(\text{avgsen})$	-0.0080575	0.0220809	-0.3649	0.71533
$\log(\text{polpc})$	0.3901473	0.0277500	14.0594	< 0.00000000000000022 ***
$\log(\text{wcon})$	-0.0411658	0.0314129	-1.3105	0.19061
$\log(\text{wtuc})$	0.0111858	0.0134466	0.8319	0.40586

log(wtrd)	-0.0411234	0.0317844	-1.2938	0.19630
log(wfir)	0.0032492	0.0218791	0.1485	0.88200
log(wser)	0.0165830	0.0147935	1.1210	0.26282
log(wmfg)	-0.2297353	0.1020216	-2.2518	0.02475 *
log(wfed)	-0.1797402	0.1708226	-1.0522	0.29319
log(wsta)	0.1238491	0.0931275	1.3299	0.18413
log(wloc)	0.0954568	0.1022299	0.9337	0.35086

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 22.197

Residual Sum of Squares: 13.408

R-Squared: 0.39596

Adj. R-Squared: 0.37986

F-statistic: 24.5825 on 14 and 525 DF, p-value: < 0.000000000000000222

- $\hat{\beta}_1 \log(prbarr)$: Un aumento del 1% en la probabilidad de arresto disminuye, en promedio, un 0.345% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_2 \log(prbconv)$: Un aumento del 1% en la probabilidad de condena disminuye, en promedio, un 0.252% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_3 \log(prbpris)$: Un aumento del 1% en la probabilidad de sentencia de prisión disminuye, en promedio, 0.177% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.

Comparando los resultados entre el modelo de mínimos cuadrados agrupados y el modelo de primeras diferencias, podemos observar algunos cambios en los coeficientes de las variables asociadas a la justicia penal y en su significancia estadística.

- $\log(prbarr)$: En el modelo de pooling, tiene un coeficiente negativo (-0.677) y significativo a un nivel de confianza muy alto con un valor p cercano a cero, lo que sugiere que un aumento en la probabilidad de arresto se relaciona con una disminución en la tasa de criminalidad. Sin embargo, en el modelo de primeras diferencias, el signo es también negativo (-0.345), pero con una magnitud menor y sigue siendo significativo a un nivel muy alto con un valor p cercano a cero. Esto indica que el efecto disuasivo del arresto sobre la tasa de criminalidad se mantiene, pero quizás no es tan fuerte como en el modelo de pooling. La variable de “probabilidad de arresto” ($\log(prbarr)$) tiene un coeficiente negativo significativo en ambos modelos. Esto sugiere que un aumento en la probabilidad de arresto está asociado con una disminución en la tasa de criminalidad.
- $\log(prbconv)$: Tanto en el modelo de pooling (-0.513) como en el de primeras diferencias (-0.252), el estimador tiene un coeficiente negativo y significativo, lo que sugiere que

un aumento en la probabilidad de condena se asocia con una disminución en la tasa de criminalidad para ambos modelos. Sin embargo, el coeficiente en el modelo de primeras diferencias es menor en magnitud que en el modelo de pooling. En ambos modelos, estos parametros indican que un aumento en la probabilidad de condena está asociado con una disminución en la tasa de criminalidad.

- $\log(\text{prbpris})$: En el modelo de pooling, la variable tiene un coeficiente positivo (0.124) y significativo a un 1%, lo que indica que un aumento en la probabilidad de sentencia de prisión podría estar relacionado con un aumento en la tasa de criminalidad. En el modelo de primeras diferencias, el coeficiente es negativo (-0.177) y significativo con un valor p menor al nivel de significancia. Esto sugiere que un aumento en la probabilidad de sentencia de prisión está asociado con una disminución en la tasa de criminalidad.

En resumen, aunque los coeficientes de las variables asociadas a la justicia penal tienen algunas diferencias en la magnitud y la significancia estadística, muestran un efecto disuasivo sobre la tasa de criminalidad en ambos modelos. Estas diferencias podrían generarse debido a la forma en que se controlan los efectos individuales en cada modelo.

Pregunta 3:

Vuelva a estimar el modelo pero ahora use efectos fijos en vez de diferenciación. ¿Existe un cambio notable (respecto a lo hallado en 2) en el signo o en la magnitud de los coeficientes de las variables asociadas a la justicia penal? ¿Qué sucede con la significancia estadística? Explique porqué existirían (o no) dichas diferencias.

Oneway (individual) effect Within Model

Call:

```
plm(formula = log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) +
      log(avgsen) + log(polpc) + log(wcon) + log(wtuc) + log(wtrd) +
      log(wfir) + log(wser) + log(wmfg) + log(wfed) + log(wsta) +
      log(wloc), data = panel, model = "within")
```

Balanced Panel: n = 90, T = 7, N = 630

Residuals:

Min.	1st Qu.	Median	3rd Qu.	Max.
-0.58697823	-0.07292538	-0.00046989	0.07755788	0.51724500

Coefficients:

	Estimate	Std. Error	t-value	Pr(> t)
$\log(\text{prbarr})$	-0.3859180	0.0324015	-11.9105	< 0.00000000000000022 ***
$\log(\text{prbconv})$	-0.3036104	0.0211703	-14.3413	< 0.00000000000000022 ***
$\log(\text{prbpris})$	-0.1919037	0.0325306	-5.8992	0.000000006544 ***

log(avgsen)	0.0240314	0.0252588	0.9514	0.3418364
log(polpc)	0.4300293	0.0268068	16.0418	< 0.00000000000000022 ***
log(wcon)	-0.0330217	0.0391582	-0.8433	0.3994497
log(wtuc)	0.0287136	0.0177099	1.6213	0.1055464
log(wtrd)	-0.0393415	0.0412091	-0.9547	0.3401785
log(wfir)	-0.0129121	0.0286998	-0.4499	0.6529656
log(wser)	0.0040193	0.0194285	0.2069	0.8361861
log(wmfg)	-0.3577230	0.0995598	-3.5930	0.0003575 ***
log(wfed)	-0.5765660	0.1620628	-3.5577	0.0004080 ***
log(wsta)	0.1802300	0.0954239	1.8887	0.0594780 .
log(wloc)	0.3282463	0.1019427	3.2199	0.0013615 **

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 17.991

Residual Sum of Squares: 10.411

R-Squared: 0.42132

Adj. R-Squared: 0.308

F-statistic: 27.3541 on 14 and 526 DF, p-value: < 0.000000000000000222

- $\hat{\beta}_1 \log(\ddot{prbarr})$: Un aumento del 1% en la probabilidad de arresto disminuye, en promedio, un 0.386% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_2 \log(\ddot{prbconv})$: Un aumento del 1% en la probabilidad de condena disminuye, en promedio, un 0.304% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_3 \log(\ddot{prbpris})$: Un aumento del 1% en la probabilidad de sentencia de prisión disminuye, en promedio, 0.192% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.

Al realizar la comparación de los resultados entre el modelo de primeras diferencias y el modelo de efectos fijos, se observan algunos cambios en los coeficientes de las variables de probabilidad asociadas a la justicia penal y en su significancia estadística.

- $\log(\ddot{prbarr})$: En el modelo de primeras diferencias, tiene un coeficiente negativo (-0.345) y significativo con un p valor muy cercano a cero, indicando un efecto disuasivo de la probabilidad de arresto sobre la tasa de criminalidad. Por otra parte, en el modelo de efectos fijos el coeficiente es también negativo y significativo (-0.386), pero la magnitud es ligeramente menor que en el modelo de primeras diferencias.
- $\log(\ddot{prbconv})$: En ambos modelos, tiene un coeficiente negativo y significativo. Para modelo de diferencias (-0.252), indicando un efecto disuasivo de la condena sobre la tasa de criminalidad. Sin embargo, la magnitud del coeficiente es menor en el

modelo de efectos fijos (-0.304). Aunque la magnitud es menor que en el modelo de primeras diferencias los estimadores de ambos modelos sugieren que, un aumento en la probabilidad de tener condena luego de ser arrestado disminuye la tasa de criminalidad.

- $\log(\text{prbpris})$: En el modelo de primeras diferencias el coeficiente es negativo y significativo (-0.177). Por otra parte, en el modelo de efectos fijos el coeficiente es negativo y significativo (-0.192). Aunque la magnitud es menor que en el modelo de primeras diferencias los estimadores de ambos modelos sugieren que, un aumento en la probabilidad de ir a prisión disminuye la tasa de criminalidad.

Por otra parte, en ambos modelos, la mayoría de los coeficientes de las variables asociadas a la justicia penal son estadísticamente significativos a un nivel de confianza muy alto, dado que el valor p es menor al nivel de significancia ($p\text{-valor} < 0.001$).

Pregunta 4:

Realice un test de significancia estadística de los efectos fijos no observables a_i . Para ello use el comando `pFtest(modelo_efecto_fijo, modelo_pooling)`, el cual realiza el test F discutido en clase (requiere como input los modelos estimados en 1 y en 3).

$$\text{Hipotesis nula : } H_0 : a_i = 0$$

$$\text{Hipotesis alternativa : } H_1 : \text{Al menos uno es significativo}$$

Si el valor p es menor que el nivel de significancia, se rechaza la hipótesis nula.

F test for individual effects

```
data: log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) + log(avgsen) + ...  
F = 38.794, df1 = 89, df2 = 526, p-value < 0.000000000000000022  
alternative hypothesis: significant effects
```

El resultado arroja un valor-p muy pequeño de 0.000000000000000022 , lo que indica que existe una fuerte evidencia estadística para rechazar la hipótesis nula. Finalmente se concluye que al menos uno de los efectos individuales es significativo con un α de 0.05 , es decir que el modelo en el cual se utiliza el método de los efectos fijos mejora significativamente el modelo en comparación con el modelo de mínimos cuadrados agrupados.

Pregunta 5:

Vuelva a estimar el modelo pero ahora use efectos aleatorios. ¿Existe un cambio notable (respecto a lo hallado en 3) en el signo o en la magnitud de los coeficientes de las variables asociadas a la justicia penal? Que conclusión podría sacar de la comparación de los modelos de efectos fijos y efectos aleatorios?

Oneway (individual) effect Random Effect Model
(Swamy-Arora's transformation)

Call:

```
plm(formula = log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) +
      log(avgsen) + log(polpc) + log(wcon) + log(wtuc) + log(wtrd) +
      log(wfir) + log(wser) + log(wmfg) + log(wfed) + log(wsta) +
      log(wloc), data = panel, model = "random")
```

Balanced Panel: n = 90, T = 7, N = 630

Effects:

```
              var std.dev share
idiosyncratic 0.01979 0.14069 0.193
individual    0.08277 0.28770 0.807
theta: 0.8182
```

Residuals:

```
      Min.    1st Qu.      Median    3rd Qu.      Max.
-0.898669 -0.078049  0.016534  0.094230  0.532297
```

Coefficients:

	Estimate	Std. Error	z-value	Pr(> z)
(Intercept)	-0.8675192	0.5899323	-1.4705	0.141415
log(prbarr)	-0.4613921	0.0326700	-14.1228	< 0.00000000000000022 ***
log(prbconv)	-0.3559045	0.0214099	-16.6233	< 0.00000000000000022 ***
log(prbpris)	-0.1954782	0.0350384	-5.5790	0.0000000242 ***
log(avgsen)	0.0246280	0.0274570	0.8970	0.369737
log(polpc)	0.4322964	0.0272689	15.8531	< 0.00000000000000022 ***
log(wcon)	-0.0243095	0.0422322	-0.5756	0.564875
log(wtuc)	0.0227704	0.0193170	1.1788	0.238487
log(wtrd)	-0.0152818	0.0447416	-0.3416	0.732684
log(wfir)	-0.0105343	0.0312808	-0.3368	0.736293
log(wser)	-0.0037376	0.0211780	-0.1765	0.859913
log(wmfg)	-0.2574320	0.0915891	-2.8107	0.004943 **
log(wfed)	-0.0239451	0.1556607	-0.1538	0.877745
log(wsta)	-0.0643540	0.0950754	-0.6769	0.498486
log(wloc)	0.2014455	0.1066653	1.8886	0.058949 .

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares: 24.214

Residual Sum of Squares: 14.649

R-Squared: 0.39501

Adj. R-Squared: 0.38124

Chisq: 401.544 on 14 DF, p-value: < 0.000000000000000222

- $\hat{\beta}_1 \log(\tilde{prbarr})$: Un aumento del 1% en la probabilidad de arresto disminuye, en promedio, un 0.461% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_2 \log(\tilde{prbconv})$: Un aumento del 1% en la probabilidad de condena disminuye, en promedio, un 0.355% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.
- $\hat{\beta}_3 \log(\tilde{prbpris})$: Un aumento del 1% en la probabilidad de sentencia de prisión disminuye, en promedio, 0.195% la tasa de criminalidad, manteniendo constantes las demás variables en el modelo.

Al comparar los resultados del modelo de efectos aleatorios con los modelos de efectos fijos y de primeras diferencias, podemos observar algunas diferencias en los coeficientes de las variables asociadas a la justicia penal.

$\log(prbarr)$: En el modelo de efectos aleatorios el coeficiente estimado es (-0.461) . Por otra parte, en el modelo de efectos fijos el coeficiente estimado es (-0.226) . Ambos modelos

$\log(prbconv)$: En el modelo de efectos aleatorios el coeficiente estimado es (-0.355) . Por otra parte, el modelo de efectos fijos el coeficiente estimado es (-0.137) .

$\log(prbpris)$: En el modelo de efectos aleatorios el coeficiente estimado es (-0.195) . Por otra parte, en el modelo de efectos fijos el coeficiente estimado es (-0.133) .

En el modelo de efectos aleatorios, los coeficientes tienen magnitudes y significancias estadísticas similares a los del modelo de efectos fijos, lo que indica que los efectos aleatorios no alteran sustancialmente las conclusiones sobre la relación entre las variables asociadas a la justicia penal y la tasa de criminalidad. Además, en ambos modelos, un aumento en la probabilidad de arresto, condena y sentencia de prisión se asocia con una disminución en la tasa de criminalidad.????

Pregunta 6:

Realice el test de Hausman. Para ello use el comando `phtest(modelo_efecto_fijo, modelo_efecto_aleatorio)`, el cual realiza el test de Hausman discutido en clase (requiere como input los modelos estimados en 3 y en 5). ¿Cuál es el mejor estimador, el de efectos fijos o el de efectos aleatorios?

El test de Hausman se utiliza para determinar si es más apropiado utilizar un modelo de efectos fijos o un modelo de efectos aleatorios.

$$\text{Hipotesis nula} : H_0 : Cov(x_{kit}, a_i) = 0, \forall k$$

$$\text{Hipotesis alternativa} : H_1 : Cov(x_{kit}, a_i) \neq 0, \text{ para algún } k$$

Hausman Test

```
data:  log(crmrte) ~ log(prbarr) + log(prbconv) + log(prbpris) + log(avgsen) + ...  
chisq = 143.58, df = 14, p-value < 0.000000000000000022  
alternative hypothesis: one model is inconsistent
```

Rechazamos H_0 . solo EF es consistente.