

ECONOMETRÍA

APLICADA UTILIZANDO R.

PAPIME PE302513 LIBRO ELECTRÓNICO Y COMPLEMENTOS DIDÁCTICOS EN MEDIOS COMPUTACIONALES, PARA EL FORTALECIMIENTO DE LA ENSEÑANZA DE LA ECONOMETRÍA

Capítulo 14

Modelos Panel: Aplicaciones en Software R

Miguel Ángel Mendoza González



RESULTADOS DE LOS MODELOS ECONOMÉTRICOS PANEL CON EL PAQUETE PLM DE R.

La curva de Philips para las ciudades de México

- El modelo de la curva de Phillips tradicionalmente explica la inflación con base a las expectativas, a los factores de demanda y de oferta (Varela y Torres, 2009). La especificación de un modelo estático general de la curva de Phillips para las i ciudades de México, se puede escribir como en la ecuación 4.

$$\pi_{i,t} = \mu_i + \beta_i(u_{i,t} - u_{i,t}^*) + \gamma_i z_{i,t} + \varepsilon_{i,t} \quad (4)$$

- Donde la inflación para cada ciudad $\pi_{i,t}$ se explica por una constante individual μ_i , por el exceso de demanda que se deriva al observar que la tasa de desempleo actual se encuentra por arriba de la tasa de desempleo natural o potencial $u_{i,t}^*$, que en la literatura se le conoce como el componente del desempleo actual que no está correlacionado con la inflación de largo plazo (NAIRU), y por variables de oferta y/o de política monetaria $z_{i,t}$.

Aplicación de los modelos panel con R

DATOS

- La base de datos fue construida por Mendoza, M.A. (2013) y contiene la inflación (INF), medida por la tasa de crecimiento del índice nacional de precios al consumidor por ciudad, la tasa de desocupación (U) por ciudad, la tasa de desocupación natural (UN) por ciudad, estimada con el filtro *Hodrick-Prescot*, y la tasa de interés medida con CETES a 28 días (CETES28).
- El índice nacional de precios al consumidor y los CETES28 se construyeron con base al Banco de México (BANXICO) y la tasa de desocupación al INEGI



Análisis Exploratorio

- Con el objetivo de llevar a cabo el análisis exploratorio, desde la consola de *R*, se escribe el siguiente comando:

```
#Mantener activa la base de datos, guardando con formato de R (Rdata)  
> load("C:/R/InflacionDesempleoCiudades.RData")
```

- La instrucción *load* tiene que incluir el lugar donde se localiza el archivo (C:/R/) y su nombre (InflacionDesempleoCiudades.RData).

- Para obtener el resumen de los estadísticos de la base de datos, escribimos:

```
#Llamar la base de datos
> summary(Datos)
```

- Se utiliza el comando *summary* con el nombre de la tabla en la base de datos entre paréntesis, que en este caso se le asignó el nombre de Datos. El resultado es:

```
#Llamar la base de datos
> summary(Datos)
```

Ciudad	Periodo	Nom_Ciudad	INF
Min. : 1.00	Min. :1995	LEON : 17	Min. : 2.30
1st Qu.: 7.75	1st Qu.:1999	ACAPULCO : 16	1st Qu.: 4.40
Median :14.50	Median :2002	AGUASCALIENTES: 16	Median : 5.50
Mean :14.50	Mean :2002	CAMPECHE : 16	Mean :11.17
3rd Qu.:21.25	3rd Qu.:2006	CD DE MEXICO : 16	3rd Qu.:16.30
Max. :28.00	Max. :2010	CHIHUAHUA : 16	Max. :41.70
(Other)	:351		

U	UN	CETES28	U_UN
Min. :0.700	Min. :1.012	Min. : 4.300	Min. :-1.565e+00
1st Qu.:2.300	1st Qu.:2.829	1st Qu.: 6.732	1st Qu.: -5.513e-01
Median :3.350	Median :3.470	Median : 8.120	Median :-6.304e-02
Mean :3.531	Mean :3.531	Mean :14.319	Mean :-6.691e-14
3rd Qu.:4.500	3rd Qu.:4.167	3rd Qu.:17.500	3rd Qu.: 4.226e-01
Max. :8.000	Max. :6.989	Max. :48.620	Max. : 2.794e+00

- En la tabla se incluyen en las dos primeras columnas las variables que identifican las ciudades (sección cruzada) y el periodo (tiempo) del formato panel, donde los valores mínimos y máximos indican 28 ciudades y una serie anual de 1995 a 2010. También se incluye una columna con el nombre de las 28 ciudades y adicionalmente a las variables INF, UN y CETES28 que se describieron anteriormente, se incluye la variable U_UN que es la diferencia de la tasa de desocupación observada y la tasa de desocupación natural (NAIRU).
- Con el objetivo de analizar la relación entre inflación, la diferencia de la tasa de desocupación observada y la tasa natural, y la tasa de interés de Cetes, se construyeron las matrices de diagramas de dispersión para los periodos 1995-2010.

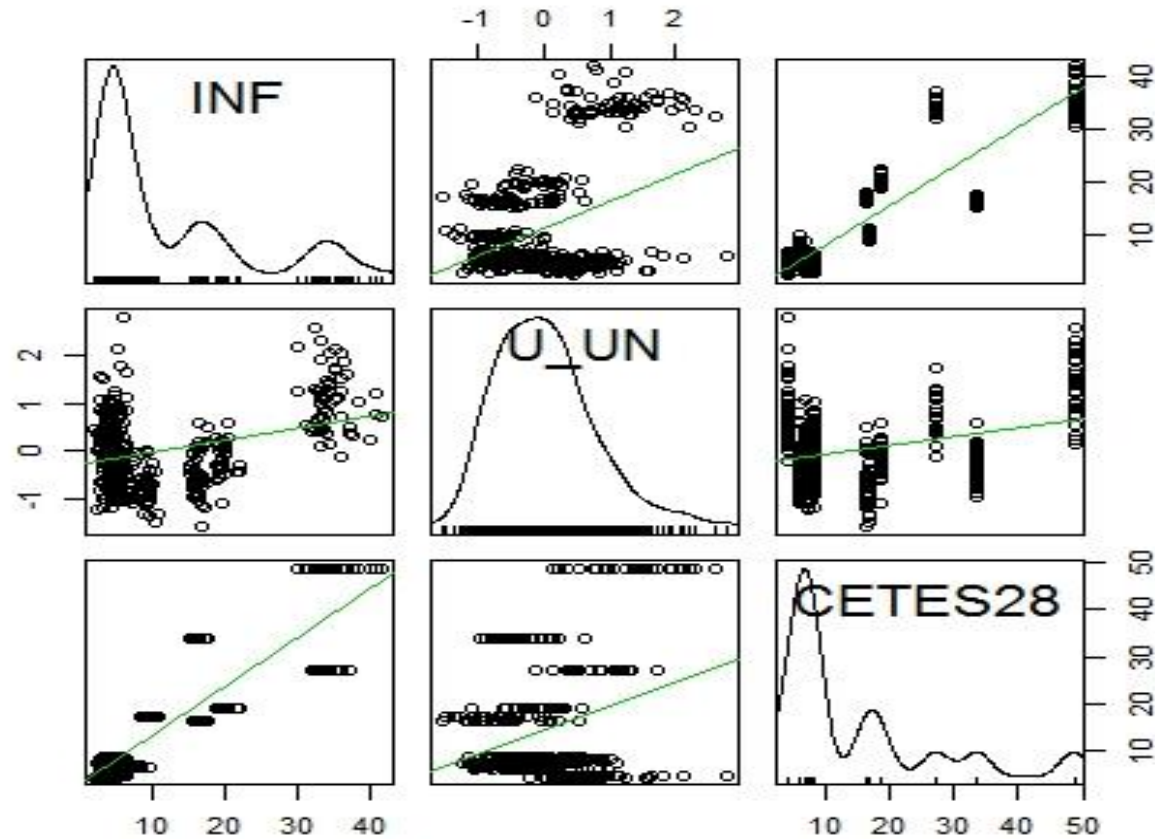
Matrices de diagramas de dispersión 1995-2010

- El comando para construir matriz en *R* incluye la especificación de las variables que se analizarán en los diagramas de dispersión, el tipo de línea de relación que en este caso es por medio de una línea de regresión, no se incluye una línea suavizada ni su correspondiente amplitud, en la diagonal se especifica la función de densidad para cada variable de la matriz.

```
#Construir una matriz
```

```
> scatterplotMatrix(~INF+U_UN+CETES28, reg.line=lm, smooth=FALSE,  
spread=FALSE,span=0.5, diagonal = 'density', data=Datos, Periodo < 2001)
```

Figura 1: Matriz de diagramas de dispersión 1995-2010 entre inflación, desempleo y tasa de interés



- En la diagonal de la matriz, se observa que la inflación y la tasa de interés presentan más de una moda y el desempleo solo una, pero es interesante como en los tres casos las modas con las mayores concentraciones se localizan en los niveles más bajos. Con respecto al diagrama de dispersión entre inflación y desempleo (segundo diagrama de la primera fila de la matriz) la relación es claramente positiva y se identifican tres grandes concentraciones relacionadas con las modas de la inflación. En cuanto a la relación inflación y tasa de interés (tercer diagrama de la primera fila), parece que la relación es positiva y muy fuerte.

Análisis confirmatorio

- Se estiman los tres tipos de modelos panel (pool, efectos fijos y aleatorios) y se aplican las pruebas de pooling y Hausman para tomar la decisión del mejor modelo en términos de consistencia.

Estimación Pool

- El comando para estimar el modelo pool incluye de derecha a izquierda, la especificación del modelo (*pooling*), la fuente de información que en nuestro caso se encuentra en Datos que está definida dentro de la base de datos de *InflacionDesempleoCiudades.RData*, la especificación de la función de inflación con respecto al diferencial del desempleo y los CETES a 28 días ($INF \sim U_UN + CETES28$), el comando que llama a la programación de los modelos panel (*plm*) y la instrucción para asignar el resultado (\leftarrow) en un objeto llamado modelo.pool.

```
#Estimación del modelo Pool  
  
> modelo.pool <- plm(INF~U_UN+CETES28, data = Datos, model = "pooling")
```

- Una vez, que se estimó el modelo panel tipo *pool* se puede observar los resultados con:

```
#Traer resultados del modelo Pool  
  
> summary(modelo.pool)
```

- Se obtiene:

Oneway (individual) effect Pooling Model

Call:

```
plm(formula = INF ~ U_UN + CETES28, data = Datos, model = "pooling")
```

Balanced Panel: n=28, T=16, N=448

Residuals :

Min.	1st Qu.	Median	3rd Qu.	Max.
-9.530	-2.500	-0.958	1.540	16.100

Coefficients :

	Estimate	Std. Error	t-value	Pr(> t)	
(Intercept)	0.876198	0.356340	2.4589	0.01432	*
U_UN	1.445834	0.319420	4.5264	7.71e-06	***
CETES28	0.718582	0.019378	37.0820	< 2.2e-16	***

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

Total Sum of Squares: 46897

Residual Sum of Squares: 9964.2

R-Squared : 0.78753

Adj. R-Squared : 0.78226

F-statistic: 824.704 on 2 and 445 DF, p-value: < 2.22e-16



Modelo de Efectos Aleatorios

- Finalmente para estimar el modelo de efectos aleatorios, se especifica el modelo con *"random"* y el mecanismo para estimar la varianza, que en este caso se utiliza el método de amemiya. Como en los casos anteriores se utiliza el comando *summary*, para ver los resultados de la estimación.

```
#Modelo de efectos aleatorios  
  
> modelo.ea <- plm(INF ~ U_UN+CETES28, data = Datos, model =  
"random", random.method="amemiya")  
  
> summary(modelo.ea)
```

- Se obtiene:

```
Oneway (individual) effect Random Effect Model
(Amemiya's transformation)

Call:
plm(formula = INF ~ U_UN + CETES28, data = Datos,
model = "random",
random.method = "amemiya")

Balanced Panel: n=28, T=16, N=448

Effects:
                var std.dev share
idiosyncratic 23.72    4.87     1
individual     0.00    0.00     0
theta: 0

Residuals :
  Min. 1st Qu.  Median 3rd Qu.  Max.
-9.530 -2.500  -0.958   1.540  16.100
```

```
Coefficients :
                Estimate Std. Error t-value
Pr(>|t|)
(Intercept) 0.876198    0.356340  2.4589
0.01432 *
U_UN         1.445834    0.319420  4.5264
7.71e-06 ***
CETES28      0.718582    0.019378 37.0820 <
2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01
'*' 0.05 '.' 0.1 ' ' 1

Total Sum of Squares:    46897
Residual Sum of Squares: 9964.2
R-Squared      : 0.78753
Adj. R-Squared : 0.78226
F-statistic: 824.704 on 2 and 445 DF, p-
value: < 2.22e-16
```

Elección de modelos alternativos

- Como se mencionó previamente, el procedimiento para elegir los modelos es el modelo pool versus el de efectos fijos y en segundo lugar el modelo de efectos fijos versus el modelo de efectos aleatorios.

```
#Modelo de efectos fijos (ef) versus modelo Pool (pool)

> pooltest(modelo.pool, modelo.ef)

      F statistic

data:  INF ~ U_UN + CETES28
F = 0.0763, df1 = 27, df2 = 418, p-value = 1
alternative hypothesis: unstability
```

Los resultados muestran que el mejor modelo es de tipo *pool*, por lo que se puede concluir que es *consistente* con respecto al modelo de *efectos fijos*.

Modelo de efectos aleatorios (ea) versus el modelo de efectos fijos (ef)

- En sentido estricto, como el modelo *pool* fue mejor que el de *efectos fijos* no es necesario hacer la prueba de *Hausman* para elegir entre el modelo de efectos fijos y aleatorios. Sin embargo, para fines ilustrativos se muestran los comandos y los resultados de la prueba de *Hausman*.

```
#Modelo de efectos aleatorios (ea) versus modelo de efectos fijos (ef)

> phtest(modelo.ea, modelo.ef)

      Hausman Test

data:  INF ~ U_UN + CETES28
chisq = 0, df = 2, p-value = 1
alternative hypothesis: one model is inconsistent
```

Aunque la prueba de *Hausman* indica que el mejor modelo es el de efectos aleatorios, no tiene sentido debido a que el modelo *pool* es el adecuado.

Consideraciones finales sobre los resultados

- Con el objetivo de probar en el nivel más general la hipótesis sobre desempleo e inflación, en este trabajo se estimaron modelos panel con el paquete *plm* del software *R* en sus tres especificaciones alternativas (*pool*, *efectos fijos* o *aleatorios*), Con las pruebas *pooling* y de *Hausman* se encontró que el modelo *pool* es consistente con respecto a los modelos de efectos fijos y aleatorio y de acuerdo a los resultados econométricos se concluyó que tanto la tasa de desempleo bajo el mecanismo NAIRU como la tasa de interés, tienen un efecto positivo y homogéneo sobre el proceso inflacionario de las ciudades para el periodo 1995-2010.



- **ARCHIVOS DE DATOS ASOCIADO AL CAPÍTULO**
InflacionDesempleoCiudades.RData



Bibliografía

- Cameron, C. y Trivedi, P. (2005) *Microeconometrics, Methods and Applications*, primera edición, Cambridge University Press.
- Croissant y Millo (2008), Panel Data Econometrics in R: The *plm* Package, en Journal of Statistical Software.
- Frees, E. (2004) *Longitudinal and Panel Data, Analysis and Applications in the Social Sciences*, Cambridge University Press.
- Greene, W. (1998) *Análisis Económico*, Prentice Hall, Tercera edición.
- Hsiao, C. 2003. *Analysis of Panel Data*. Cambridge University Press: segunda edición.
- Mendoza, M.A. (2013) *Inflación y desempleo en ciudades de México: una evaluación con modelos panel y aplicaciones en software R*,
- Varela, R. y Torres, V. (2009) Estimación de la tasa de desempleo no aceleradora de la inflación en México, *Análisis Económico*, Núm. 57, vol. XXIV.
- Wooldridge, J. (2002) *Econometric Analysis of Cross Section and Panel Data*, Massachusetts Institute of Technology.

