

# Modelos de datos de panel

## Aplicaciones en ordenador

Román Salmerón Gómez

Para ilustrar cómo abordar el análisis de **Modelos de datos de panel** con el software econométrico **Gretl** vamos a considerar el archivo de excel *Endeudamiento.xls* (<http://www.ugr.es/~romansg/material/WebEco/Eco2-Endeudamiento.xls>) donde se tiene información de 1204 pequeñas y medianas empresas (PYMEs) españolas en 4 periodos de tiempo sobre las siguientes variables:

**RTD:** nivel de endeudamiento de las empresas (medido como el ratio total de deuda, es decir, como el cociente entre el total de deuda de la empresa y el total de recursos de la misma),

**TAM:** tamaño de las empresas (que se estima como el logaritmo de los activos totales de la empresa),

**ESTA:** estructura de los activos (que se define como el cociente entre los activos fijos netos de la empresa y los activos totales),

**ROA:** rentabilidad de las empresas (calculada como la ratio entre el beneficio antes de intereses e impuestos y el activo total de la empresa),

**CRE:** crecimiento de las empresas (definido como la variación porcentual de los activos totales de una empresa con respecto al ejercicio anterior), y

**lnEDAD:** edad de las empresas (definida como el logaritmo del número de años que lleva funcionando la empresa).

El objetivo del análisis es estudiar qué factores influyen en el endeudamiento de una PYME y cómo lo hacen.

Al importar los datos desde excel hay que indicar que la estructura del conjunto de datos es de panel, más concretamente, de *Series temporales apiladas* con 1204 unidades de sección cruzada y 4 periodos temporales.

En primer lugar estimaremos el modelo por MCO, ignorando la estructura de panel de los datos. En tal caso se obtienen los siguientes resultados:

Modelo 1: MCO combinados, utilizando 4816 observaciones

Se han incluido 1204 unidades de sección cruzada

Largura de la serie temporal = 4

Variable dependiente: RTD

	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	0.933283	0.0374125	24.9458	0.0000
TAM	0.0156061	0.00414705	3.7632	0.0002
ESTA	-0.186070	0.0132441	-14.0492	0.0000
ROA	-1.00852	0.0366792	-27.4956	0.0000
CRE	0.114440	0.00983585	11.6350	0.0000
ln_EDAD_	-0.109826	0.00494316	-22.2179	0.0000
Media de la vble. dep.	0.655321	D.T. de la vble. dep.		0.210526
Suma de cuad. residuos	156.7541	D.T. de la regresión		0.180525
$R^2$	0.265465	$R^2$ corregido		0.264701
$F(5, 4810)$	347.6717	Valor p (de $F$ )		0.000000
Log-verosimilitud	1413.841	Criterio de Akaike		-2815.683
Criterio de Schwarz	-2776.805	Hannan-Quinn		-2802.031
$\hat{\rho}$	0.932376	Durbin-Watson		0.129608

Para tener en cuenta la naturaleza de los datos con los que se cuenta, datos de panel, hay estimar el modelo anterior bajo el supuesto de efectos fijos o aleatorios. Para ello hay que seguir la ruta *Modelo* → *Panel* → *Efectos fijos o efectos aleatorios*.

En la nueva ventana que aparece hay que especificar la variable dependiente y las independientes (como siempre), especificar si la estimación es bajo efectos fijos o aleatorios, si se desean obtener estimaciones con desviaciones típicas robustas (evitando problemas de heteroscedasticidad y/o autocorrelación) o incluir las variables ficticias temporales.

Considerando desviaciones típicas robustas y efectos fijos se obtienen los siguientes resultados:

Modelo 2: Efectos fijos, utilizando 4816 observaciones

Se han incluido 1204 unidades de sección cruzada

Largura de la serie temporal = 4

Variable dependiente: RTD

Desviaciones típicas robustas (HAC)

	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	0.493671	0.0740504	6.6667	0.0000
TAM	0.0909579	0.0108863	8.3552	0.0000
ESTA	-0.0234481	0.0282061	-0.8313	0.4059
ROA	-0.283984	0.0642849	-4.4176	0.0000
CRE	0.0402151	0.00592719	6.7849	0.0000
ln_EDAD_	-0.218768	0.0192028	-11.3926	0.0000
Media de la vble. dep.	0.655321	D.T. de la vble. dep.		0.210526
Suma de cuad. residuos	9.762124	D.T. de la regresión		0.052023
$R^2$	0.954256	$R^2$ corregido		0.938936
$F(1208, 3607)$	62.28822	Valor p (de $F$ )		0.000000
Log-verosimilitud	8098.855	Criterio de Akaike		-13779.71
Criterio de Schwarz	-5945.754	Hannan-Quinn		-11028.81
$\hat{\rho}$	0.071740	Durbin-Watson		1.170712

Contraste de diferentes interceptos por grupos –

Hipótesis nula: Los grupos tienen un intercepto común

Estadístico de contraste:  $F(1203, 3607) = 45.1471$

con valor p =  $P(F(1203, 3607) > 45.1471) = 0$

Considerando efectos aleatorios (no es posible seleccionar desviaciones típicas robustas) se obtienen los siguientes resultados:

Modelo 3: Efectos aleatorios (MCG), utilizando 4816 observaciones

Se han incluido 1204 unidades de sección cruzada

Largura de la serie temporal = 4

Variable dependiente: RTD

	Coefficiente	Desv. Típica	Estadístico $t$	Valor p
const	0.606417	0.0340803	17.7938	0.0000
TAM	0.0598154	0.00418277	14.3004	0.0000
ESTA	-0.0626182	0.0129809	-4.8239	0.0000
ROA	-0.358541	0.0214137	-16.7435	0.0000
CRE	0.0512767	0.00357761	14.3326	0.0000
ln_EDAD_	-0.156701	0.00700473	-22.3708	0.0000
Media de la vble. dep.	0.655321	D.T. de la vble. dep.		0.210526
Suma de cuad. residuos	174.8210	D.T. de la regresión		0.190625
Log-verosimilitud	1151.166	Criterio de Akaike		-2290.333
Criterio de Schwarz	-2251.455	Hannan-Quinn		-2276.681

$$\begin{aligned}\hat{\sigma}_\varepsilon^2 &= 0.00270644 \\ \hat{\sigma}_u^2 &= 0.0288907 \\ \theta &= 0.846965\end{aligned}$$

Contraste de Breusch-Pagan –

Hipótesis nula: Varianza del error específico a la unidad = 0

Estadístico de contraste asintótico:  $\chi^2(1) = 5573.55$

con valor p = 0

Contraste de Hausman –

Hipótesis nula: Los estimadores de MCO son consistentes

Estadístico de contraste asintótico:  $\chi^2(5) = 291.179$

con valor p = 7.88647e-061

Una vez estimado el modelo bajo los tres supuestos habrá que decidir cuál de ellos es el idóneo:

- Así, al estimar bajo efectos fijos, al final de los resultados, se tiene un contraste de hipótesis cuya hipótesis nula es si el modelo debe ser estimado mediante MCO agrupados frente a la alternativa de efectos fijos. En este caso, atendiendo al p-valor, se rechaza la hipótesis nula.
- Al estimar por efectos aleatorios, en el contraste de Breusch-Pagan (cuya hipótesis nula es estimar por MCO agrupados frente a la alternativa de efectos aleatorios) se tiene que se vuelve a rechazar la hipótesis nula.
- Por tanto, con los dos contrastes anteriores queda claro que se rechaza la hipótesis nula de estimación por MCO agrupados, por lo que habrá que decidir entre efectos fijos o aleatorios. El test de Hausman tiene como hipótesis nula que el método adecuado de estimación es el de efectos aleatorios frente a la alternativa de efectos fijos. Puesto que el p-valor es prácticamente cero, se rechaza la hipótesis nula.

Una vez decidido que el método idóneo para estimar el modelo son los efectos fijos (modelo 2), procedamos a interpretar los resultados:

- El p-valor asociado a la prueba de significación conjunta indica que se rechaza la hipótesis nula, por lo que la variabilidad explicada en el modelo no se debe al azar.
- El coeficiente de determinación indica que el ajuste explica el 95'42% de la variabilidad del endeudamiento.
- Atendiendo a los contrastes de significación individual:
  - El tamaño de la empresa influye positivamente en el endeudamiento, es decir, a mayor tamaño mayor endeudamiento.
  - La rentabilidad de la empresa influye negativamente en el endeudamiento, es decir, a mayor rentabilidad menor endeudamiento.
  - El crecimiento de la empresa influye positivamente en el endeudamiento, es decir, a mayor crecimiento mayor endeudamiento.
  - La edad de las empresas influye negativamente en el endeudamiento, es decir, a mayor edad menor endeudamiento.

Dejo una interpretación más profunda para los economistas...