

Estrategias Metodológicas para Datos de Panel. El caso de los bancos típicos en Argentina

Margarita Díaz, José Miguel Vargas¹

¹Instituto de Demografía y Estadística. Facultad de Ciencias Económicas.UNC

Congreso Interamericano de Estadística
Octubre 2017

Objetivos

- Evaluar la performance de diversas estrategias para modelar datos de panel
- Analizar la información financiera de los bancos típicos en Argentina para el período 2005-2015
- Cuantificar los efectos de indicadores específicos de en la rentabilidad de los bancos clasificados como *Típicos*.

Metodología

Información longitudinal \Rightarrow dependencia temporal y/dependencia *Cross-section*

- **Dependencia temporal:** Las variables se miden en el mismo individuo en distintos momentos del tiempo.
 - Nivel 1: Ocasión
 - Nivel 2: Bancos
- El valor de una observación asociada a una Ocasión está relacionado con el valor de otras observaciones del mismo Banco
- **Dependencia Cross-section:** los datos de panel a menudo presentan dependencias transversales y temporales (Feng and Seasholes (2004) y Hirshleifer y Teoch (2003) entre otros).

Algunos Estimadores de Datos de Panel

- **Modelo Pooled OLS**

Es el modelo más restrictivo, ya que especifica que la ordenada y los coeficientes son constantes

$$y_{it} = \alpha + x'_{it}\beta + \varepsilon_{it}$$
$$i = 1, 2, \dots, n \quad t = 1, 2, \dots, T$$

- La regresión pooled OLS subestima los errores standard

- **Modelos Sujeto Específico**

Incluyen una ordenada para cada individuo, lo que capta la heterogeneidad no observada que induce la dependencia temporal

$$y_{it} = \alpha + x'_{it}\beta + \mu_i + \varepsilon_{it}$$

- Las μ_i son variables aleatorias (Cameron y Trivedi (2009))

Efectos Fijos y Efectos Aleatorios

Modelo de Efectos Fijos

- μ_i es una variable aleatoria no observada, potencialmente correlacionada con las regresoras observadas

Modelo de Efectos Aleatorios

- μ_i es una variable aleatoria no observada independiente de regresoras observadas

Modelo de Efectos Aleatorios

- Residuos a nivel dos

$$\mu_i \sim N(0, \sigma_\mu^2)$$

- Residuos a nivel uno

$$\varepsilon_{it} \sim N(0, \sigma_\varepsilon^2)$$

- Varianza total

$$\sigma^2 = \sigma_\mu^2 + \sigma_\varepsilon^2$$

- Coeficiente de correlación intraclase (ICC)

$$ICC = \frac{\sigma_\mu^2}{\sigma_\mu^2 + \sigma_\varepsilon^2}$$

Modelo de Efectos Fijos: Estimador Dentro

- Se estima un modelo OLS al siguiente modelo de regresión transformado:

$$(y_{it} - \bar{y}_i + \bar{\bar{y}}) = \alpha + (x_{it} - \bar{x}_i + \bar{\bar{x}})' \beta + (\varepsilon_{it} - \bar{\varepsilon}_i + \bar{\bar{\varepsilon}})$$

- Medias de grupos:

$$\bar{y}_i \quad \bar{x}_i \quad \bar{\varepsilon}_i$$

- Medias de todas las observaciones:

$$\bar{\bar{y}} \quad \bar{\bar{x}} \quad \bar{\bar{\varepsilon}}$$

Propiedades de OLS, EF y EA

- En el modelo de **EA** se pueden calcular los valores predichos $E[y_{it} | x_{it}] = \alpha + x'_{it}\beta$; mientras que en el modelo de **EF** $E[\mu_i | x_{it}]$ no es estimable consistentemente para paneles cortos.
- El modelo **EF** permite identificar consistentemente sólo el efecto marginal para regresoras que varían en el tiempo.
- Si el verdadero modelo es **EF**, tanto **Pooled OLS** como **EA** arrojan **estimadores inconsistentes**.
- Si el modelo de **EA** es correcto, **Pooled OLS** subestima varianzas \Rightarrow hay estimadores robustos a heterocedasticidad y autocorrelación

Estimadores robustos de las varianzas

- Para **Pooled OLS**, modelo **EF** y modelo **EA**, los estimadores robustos están basados en la corrección de White, conocidos como estimadores sándwich y permiten obtener estimadores de los errores estándar consistentes en presencia de heterocedasticidad y autocorrelación, siempre que se cumpla la independencia transversal (Baltagi (2008)).

Prueba de Dependencia Cross-section de Pesaran (CD)

- Numerosos estudios indican que los paneles microeconómicos exhiben dependencia *cross-section*.
- Es una prueba aplicable a modelos de panel, que se basa en el promedio de coeficientes de correlación de parejas de los residuos OLS de las regresiones individuales en el panel (Pesaran (2004))

$$y_{it} = \alpha_i + x'_{it}\beta_i + \varepsilon_{it}$$

$$\hat{\rho}_{ij} = \frac{\sum_{t=1}^T e_{it}e_{jt}}{\sqrt{\sum_{t=1}^T e_{it}^2}\sqrt{\sum_{t=1}^T e_{jt}^2}} \quad e_{it} = y_{it} - \hat{\alpha}_i - x'_{it}\beta_i$$

Prueba de dependencia cross-section de Pesaran

$$CD = \sqrt{\frac{2T}{n(n-1)}} \left(\sum_{i=1}^{n-1} \sum_{j=i+1}^n \hat{\rho}_{ij} \right)$$

- Bajo $H_0 : Cov(\varepsilon_{it}, \varepsilon_{jt}) = 0 \quad \forall t, \quad i \neq j,$
 $\sqrt{T} \hat{\rho}_{ij} \Rightarrow N(0, 1)$
 $CD \Rightarrow N(0, 1)$

Incorporando la dependencia *Cross-section*

- Driscoll and Kraay (1998) proponen un estimador no paramétrico de la matriz de covarianza consistente y “robusto” a varias formas de dependencia transversal y temporal.
- Ese estimador es implementado en Stata a través del programa xtsc (D. Hoechle, Stata Journal 2007, v.7, n.3.) para *pooled OLS* y regresión de efectos fijos.

- El grupo objeto de análisis constituido por 25 "Bancos Típicos" se identificó en un estudio previo, a través de una clasificación de los bancos mediante la aplicación del algoritmo K Medias robusto.
- Para cada uno de ellos se tomaron los estados financieros de los últimos once años. La fuente de datos son las publicaciones realizadas por el Banco Central de la República Argentina, para los meses de diciembre de los años 2005 al 2015.

Modelo de rentabilidad

- **Variable respuesta**

La variable respuesta es el ROA, que relaciona los beneficios con el activo

- **Covariables:** similares a las utilizadas en el trabajo de Athanasoglou et al (2006)

- Capital (Patrimonio Neto/Activo)
- Riesgo Crediticio (Previsiones/Préstamos)
- Productividad (Ingresos Brutos/Gastos Administrativos),
- Gestión de Gastos (Gastos Administrativos/Activo)
- Tamaño (logaritmo del Activo).

Resultados de Modelos, cont.

Variable	PoolRob	EA	EF	EFRob	EFDris
Capital	36.116146	32.232528	30.31337	30.31337	30.31337
	2.431413	2.1480232	2.7322293	4.3219412	1.4015979
	14.85	15.01	11.09	7.01	21.63
	0.0000	0.0000	0.0000	0.0000	0.0000
Riesgo Crediticio	-21.032378	-17.657172	-12.210044	-12.210044	-12.210044
	5.6701896	4.5406705	4.7922412	6.7649724	3.6117592
	3.71	-3.89	-2.55	-1.80	-3.38
	0.0003	0.0001	0.0115	0.0837	0.0070
Productividad	.85422618	1.0128743	.78802075	.78802075	.78802075
	.16107808	.17714311	.21015708	.28066434	.24576423
	5.30	5.72	3.75	2.81	3.21
	0.0000	0.0000	0.0002	0.0098	0.0094
Gestión de Gastos	14.83654	14.643283	5.4465569	5.4465569	5.4465569
	3.0846133	4.0734656	5.7035351	6.5588249	4.4641559
	4.81	3.59	0.95	0.83	1.22
	0.0000	0.0003	0.3406	0.4145	0.2504
Tamaño	.10122693	.2308353	.56900885	.56900885	.56900885
	.04986016	.08923801	.14430852	.15897208	.14813406
	2.03	2.59	3.94	3.58	3.84
	0.0433	0.0097	0.0001	0.0015	0.0033
cons	-4.6087751	-5.8020812	-7.6443662	-7.6443662	-7.6443662
	.61758105	.81200462	1.0517878	1.3676926	1.2315498
	7.46	-7.15	-7.27	-5.59	-6.21
	0.0000	0.0000	0.0000	0.0000	0.0001

legend: b/se/t/p

Resultados de Modelos, cont.

- En primer lugar se concluye que el panel presenta efectos individuales:
EF versus Pooled Ols:

$$F \text{ test } u_i = 0 : F(24, 237) = 6,90 \text{ Prob} > F = 0,0000$$

- En segundo lugar, a través de la prueba de Hausman, se comparan las estimaciones de EF y EA

$$\text{chi2}(5) = (b - B)'[(V_b - V_B)^{-1}](b - B) = 24,35$$

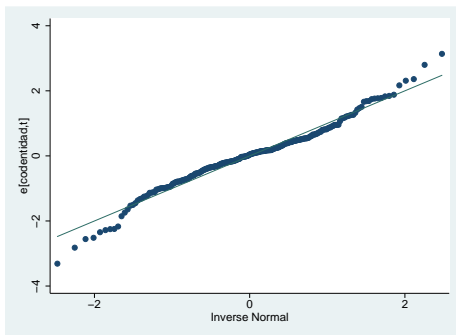
$$\text{Prob} > \text{chi2} = 0,0002$$

- En tercer lugar, se prueba la Dependencia Cross-section a través de la prueba de Pesaran

$$\text{Pesaran's test of cross sectional independence} = 6,072, \text{ Pr} = 0,0000$$

$$\text{Average absolute value of the off - diagonal elements} = 0,336$$

Normalidad de los residuos



Prueba de simetría y kurtosis para una normal

Variable	Obs	Pr(Skewness)	Pr(Kurtosis)	adj chi2(2)	Prob>chi2
error	267	0.3550	0.0009	10.55	0.0051

Odenamiento de variables significativas

- Capital: positivo y altamente significativo
- Tamaño: el tamaño del banco (activo en log) es positivo
- Riesgo Crediticio: negativo. En línea con la teoría, que sostiene que la mayor exposición al riesgo de crédito se asocia normalmente con la disminución de la rentabilidad de la empresa
- Productividad: efecto positivo
- Gestión de Gasto: Es la variable predictora cuyo coeficiente es muy distinto en los modelos Pooled y EA. Se espera que sea negativo, no resulta significativa.

Bibliografía

- Athanasoglou P. et al (2006) Bank specific, industry-specific and macroeconomic determinants of bank profitability . Journal of International Financial Markets Institutions and Money
- Baltagi, B. (2008). Econometric analysis of panel data. John Wiley Sons.
- Driscoll J.C. y Kraay A.C. (1998) Consistent Covariance Matrix Estimation with Spatially Dependent Panel Data. The MIT Press
- Hoechle D. (2007). Robust Standard Errors for Panel Regressions with Cross-Sectional Dependence. The Stata Journal, v. 7, n. 3.
- Feng, L., y M. S. Seasholes. 2004. Correlated Trading and Location. Journal of Finance 59(5): 2117–2144.

- Hirshleifer, D., y S. H. Teoh. 2003. Herd Behaviour and Cascading in Capital Markets: A Review and Synthesis. *European Financial Management* 9(1): 25–66.
- Pesaran M.H (2004) General Diagnostic Tests for Cross Section Dependence in Panels. CESifo Working Paper No. 1229.