

Burdisso, Tamara; Sangiácomo, Máximo

Working Paper

Series de tiempo en panel: Una reseña de la evolución metodológica

Working Paper, No. 2015/68

Provided in Cooperation with:

Economic Research Department (ie), Central Bank of Argentina

Suggested Citation: Burdisso, Tamara; Sangiácomo, Máximo (2015) : Series de tiempo en panel: Una reseña de la evolución metodológica, Working Paper, No. 2015/68, Banco Central de la República Argentina (BCRA), Investigaciones Económicas (ie), Buenos Aires

This Version is available at:

<https://hdl.handle.net/10419/126249>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

ESTUDIOS BCRA
Documentos de trabajo 2015 / 68

**Series de tiempo en panel.
Una reseña de la evolución metodológica**

Tamara Burdisso / Máximo Sangiácomo
Banco Central de la República Argentina

Noviembre, 2015



ie | BCRA
INVESTIGACIONES ECONÓMICAS

Banco Central de la República Argentina
ie | Investigaciones Económicas

Noviembre, 2015
ISSN 1850-3977
Edición Electrónica

Reconquista 266, C1003ABF
C.A. de Buenos Aires, Argentina
Tel: (5411) 4348-3582
Fax: (5411) 4348-3794
Email: investig@bcra.gov.ar
Pág. Web: www.bcra.gov.ar

Las opiniones vertidas en este trabajo son exclusiva responsabilidad de los autores y no reflejan necesariamente la posición del Banco Central de la República Argentina. La serie ESTUDIOS BCRA *Documentos de Trabajo* está compuesta por material preliminar que se hace circular con el propósito de estimular el debate académico y recibir comentarios. Toda referencia que desee efectuarse a estos Documentos deberá contar con la autorización del o los autores.

Series de tiempo en panel

Una reseña de la evolución metodológica

Tamara Burdisso^a y Máximo Sangiacomo^b

Resumen

El trabajo se focaliza en el tratamiento econométrico de los paneles macro, los que la literatura conoce con el nombre de *panel time series*. Este nuevo enfoque rechaza el supuesto de homogeneidad de las pendientes y se ocupa de la no estacionariedad. Además, reconoce que la presencia *cross-section dependence* (CSD), i.e. cierta estructura de correlación en el término de error entre las unidades debido a la presencia de factores comunes no observables, dilapida las ganancias de eficiencia por operar con un panel. Esto condujo a un nuevo conjunto de estimadores referidos en la literatura como *Common Correlated Effect* (CCE), que esencialmente consiste en aumentar el modelo que se desea estimar mediante la incorporación de los promedios de los individuos en cada t del tiempo, tanto de la variable dependiente como de los regresores específicos de cada individuo. Finalmente, se presentan dos códigos desarrollados en *Stata* para la evaluación y tratamiento de la *cross-section dependence*.

Abstract

The document focuses on the econometric treatment of macro panels, known in literature as panel time series. This new approach rejects the assumption of slopes' homogeneity and handles nonstationarity. It also recognizes that the presence of cross-section dependence (CSD), i.e. some correlation structure in the error term between units due to the presence of unobservable common factors, squanders efficiency gains by operating with a panel. This led to a new set of estimators known in literature as Common Correlated Effect (CCE), which essentially consists of increasing the model to be estimated by adding the averages of the individuals in each time t , of both the dependent variable and the specific regressors of each individual. Finally, two *Stata* codes developed for the evaluation and treatment of the cross-section dependence are presented.

Junio 2015

^a BCRA-UBA. tburdisso@bcra.gov.ar

^b BCRA-UNLP. maximo.sangiaco@bcra.gov.ar

Indice

| | |
|--|----|
| 1. Introducción | 3 |
| 2. Series de tiempo en panel..... | 4 |
| 3. Evaluación de la cross-section dependence: XTCSI | 12 |
| 4. El test de IPS en presencia de la cross-section dependence: XTCIPS..... | 15 |
| Bibliografía..... | 18 |

1. Introducción

Los modelos de datos de panel se han vuelto muy populares en la econometría empírica de finales del siglo XX y comienzos del XXI, fundamentalmente por su capacidad para capturar la heterogeneidad del comportamiento socio-económico de los agentes, en comparación con los modelos *cross-section* o de series de tiempo.

El término datos de panel¹ es utilizado para describir una amplia variedad de situaciones en econometría. Básicamente un panel de datos consiste en una muestra de individuos² en el tiempo, brindando de esta forma múltiples observaciones para cada individuo, mediante encuestas a familias /empresas, que se realizan con cierta periodicidad. Este tipo de paneles reciben el nombre de *paneles micro* ya que consisten en un gran número de N individuos –cientos o miles- sobre un período corto de tiempo – $T=2$ observaciones por individuo a un máximo de $T=10/20$ -. En contraste, los *paneles macro*, generalmente involucran a un número N de países –que pueden ir desde unos pocos países como un G7 hasta todos los países de la *Penn World Table* o de la *World Development Indicators*-, cuya información se muestra por lo general con frecuencia trimestral o anual, con rangos que van de 20 a 60 años (Arellano, 2003; Hsiao, 2003).

Paneles micro y macro requieren un tratamiento econométrico diferente (Baltagi, 2008). Por ejemplo, el estudio asintótico en paneles micro deber realizarse para N grande y T fijo, mientras que el análisis asintótico en los paneles macro se realiza permitiendo que tanto N como T tiendan a infinito (Phillips and Moon, 1999). Asimismo, un T grande en un panel macro debe lidiar con cuestiones de no estacionariedad propias del análisis de series de tiempo.

Los primeros desarrollos teóricos en el estudio de datos de panel surgieron para el tratamiento de los paneles micro. La literatura de segunda mitad de los 80 y gran parte de los 90 se focalizó en la estructura de los paneles micro: N grande y T chico. El estimador de efectos fijos, el estimador de Anderson y Hsiao, el estimador de Arellano y Bond o el *System GMM* fueron estimadores pensados para lidiar con el diseño del panel micro (ver Arellano, 2003; Hsiao, 2003; Baltagi, 2008). Por el contrario, es hacia finales de los años 90 cuando aparecen los primeros trabajos que advierten que la elección del estimador depende de manera crucial del diseño del panel, i.e del tamaño relativo de N y T (Pesaran and Smith, 1995; Im, Pesaran and Shin, 2003).

En este trabajo nos focalizamos en el tratamiento econométrico de los paneles macro, los que la literatura conoce con el nombre de *panel time series* y que aquí llamamos series de tiempo en panel. A continuación presentamos una breve descripción de las principales cualidades que caracterizan a la literatura de

¹ Datos de panel o datos longitudinales. Los nombres pueden variar dependiendo de la disciplina que analiza los datos.

² Por individuos se entiende, trabajadores, familias, empresas, industrias, regiones, países, etc.

series de tiempo en panel y dos códigos que desarrollamos en Stata para la evaluación y tratamiento de la *cross-section dependence*.

2. Series de tiempo en panel

La investigación empírica como la paridad del poder de compra (PPP), la relación ahorro-inversión o el problema de la convergencia en la teoría del crecimiento, entre otras, se han visto beneficiadas con el advenimiento de los paneles conformados por países, con T grande. El hecho de que T pueda tender a infinito contribuyó al desarrollo de dos grandes líneas de trabajo en la literatura de datos de panel. La primera de estas líneas rechazó el supuesto de homogeneidad de las pendientes, asumido en el uso de los estimadores *pooled* estándar (efectos fijos, diferencia o *System GMM*), y propuso pendientes heterogéneas, i.e. una regresión por individuo/unidad (Pesaran y Smith (1995), Pesaran, Shin y Smith (1999), Im, Pesaran y Shin (2003), entre otros). Esta literatura se basa en un T lo suficientemente grande como para estimar cada regresión separadamente, i.e una regresión por país.

La otra línea de trabajo se concentró en los métodos de series de tiempo aplicados a panel, ocupándose de la no estacionariedad, las regresiones espurias y las relaciones de cointegración. Incorporar la dimensión *cross-section* a la dimensión temporal, ofrece importantes ventajas en la evaluación de la no estacionariedad y la cointegración. En este sentido, la confianza en la econometría de paneles no estacionarios radica en combinar lo mejor de ambos mundos: el tratamiento de la no estacionariedad según los modelos de series de tiempo y al mismo tiempo incrementar datos y potencia de los tests a partir de la dimensión *cross-section*. Particularmente, añadir la dimensión *cross-section* bajo determinados supuestos puede interpretarse como diferentes extracciones de una misma distribución poblacional. De ahí que, al combinar la dimensión temporal con la dimensión *cross-section* se incrementa la potencia de los test estadísticos y los estimadores pueden converger en distribución a variables aleatorias normales (Baltagi y Kao, 2003).

Al igual que en el análisis empírico de los modelos de series de tiempo, los test de raíces unitarias se han vuelto en la actualidad una práctica frecuente en los modelos de panel. Es así que, hacia finales de los años 90 comienzan a desarrollarse los primeros tests de raíces unitarias en panel.³ Desde un punto de vista teórico, estos tests requieren que tanto N como T tiendan a infinito, aunque la manera en que el número de individuos/países N y T tienden a infinito es crucial para determinar las propiedades asintóticas de los estimadores (Phillips y Moon, 1999 y 2000, Levin y Lin, (LL), 1992). Uno de los test más utilizados es el test de Im, Pesaran y Shin (IPS, 1997, 2003), ya que

³ El primero de estos test fue el de Levin & Lin (1992). Posteriormente aparecen Im, Pesaran & Shin (2003), Maddala & Wu (1999), Choi (2001), Hadri (2000).

resulta ser menos restrictivo que el test de Levin, Lin y Chu (LLC, 2002). De acuerdo con el trabajo de Im, Pesaran y Shin (2003), se considera una muestra de N unidades/países a lo largo de T períodos. Supone además que el proceso estocástico y_{it} es generado por un proceso autorregresivo de primer orden

$$y_{it} = (1 - \phi_i)\mu_i + \phi_i y_{i,t-1} + \varepsilon_{it}, \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (1)$$

con valores iniciales y_{i0} . Al igual que en el test de Dickey-Fuller (DF), el interés radica en testear la hipótesis nula de raíz unitaria de $\phi_i = 1$ para todo i . La ecuación anterior puede ser reescrita como

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \varepsilon_{it} \quad \text{donde} \quad \alpha_i = (1 - \phi_i)\mu_i, \quad \beta_i = -(1 - \phi_i) \quad \text{y} \quad \Delta y_{it} = y_{it} - y_{i,t-1}.$$

El test de raíz unitaria es entonces

$$H_0 : \beta_i = 0 \quad \text{para todo } i, \quad \text{vs. las alternativas}$$

$$H_1 : \beta_i < 0, \quad i = 1, \dots, N_1, \quad \beta_i = 0, \quad i = N_1 + 1, N_1 + 2, \dots, N \quad 0 < N_1 \leq N$$

Nótese que la hipótesis alternativa permite que los β_i difieran entre unidades, a diferencia de la alternativa homogénea del test de LLC, que plantea $\beta_i = \beta < 0$ para todo i . Asimismo la fracción de unidades que siguen un proceso estacionario es distinta de cero, i.e. $\lim_{N \rightarrow \infty} (N_1/N) = \delta, \quad 0 < \delta \leq 1$. Esta condición es necesaria para la consistencia del test de IPS. El estadístico propuesto por IPS, al que llaman \bar{t}_{IPS} , es definido como el promedio de los estadísticos t individuales de la regresión (aumentada) DF.

Por su parte Maddala y Wu (MW, 1999) coinciden en la ventaja de la alternativa heterogénea de IPS pero resaltan que promediar los estadísticos DF no es la manera más eficiente de usar la información. Siguiendo a Fisher (1932) proponen un test estadístico que es un promedio de los logaritmos de los p -values asociados al estadístico t de cada unidad. De acuerdo con sus simulaciones, existen una variedad de situaciones donde el test de MW es superior tanto en tamaño como en potencia respecto del IPS, que a su vez es más potente que LL (Smith y Fuertes, 2010). Por otro lado, Breitung (2000) realiza simulaciones de Montecarlo para estudiar la potencia de los test de LL e IPS, y encuentra una dramática pérdida de potencia de ambos tests cuando se incorporan términos determinísticos.

Breitung y Pesaran (2008) proveen una revisión de la evolución de los tests de raíces unitarias para panel. Ellos destacan que uno de los objetivos primarios de la aplicación de estos tests en panel era mejorar el pobre desempeño de los test de raíces unitarias en series de tiempo. Por ejemplo, el test aumentado de DF por lo general no rechaza la hipótesis nula de que el tipo de cambio real es no estacionario. Por el contrario, los tests de raíz unitaria para panel aplicados a un conjunto de países industrializados, generalmente rechazan la hipótesis de una raíz unitaria, i.e. el tipo de cambio real muestra un comportamiento estacionario, respaldando empíricamente a la PPP (Coakley y Fuertes, 1997).

Si bien los tests de raíces unitarias para panel venían a solucionar la falta de potencia que mostraban dichos tests en series de tiempo, los mismos trajeron aparejados varios inconvenientes. En primer lugar, uno de los supuestos claves en los test de raíces unitarias en panel es **la independencia de las unidades**, condición necesaria para que el estadístico promedio de los DF individuales, \bar{t}_{IPS} , converja en distribución a la normal.⁴ En segundo lugar, si se rechaza la hipótesis nula de raíz unitaria, la interpretación de este resultado se dificulta, ya que la mejor conclusión que se puede extraer, es que existe una fracción de unidades que son estacionarias. Pero nada se puede decir respecto de cuantas y cuales son las unidades/países con comportamiento estacionario.

Una de las razones por la cual la investigación empírica se preocupa por la presencia de raíces unitarias en modelos de series de tiempo es para evitar el problema de correlación espuria. Como es sabido, cointegración es la condición requerida entre variables $I(1)$, para que la regresión no resulte espuria y el estimador de interés sea consistente. Es decir, si las variables están cointegradas entonces las mismas comparten una tendencia estocástica común que se cancela en la combinación lineal de estas. Pesaran y Smith (1995) señalaron que el problema de regresión espuria no se origina en la regresiones *cross-section* que surgen al colapsar la dimensión temporal, aún cuando la serie de tiempo de cada unidad/país contenga una raíz unitaria. Esta observación fue la que condujo a que el problema de correlación espuria se vea bastante mitigado al promediar las unidades. Phillips y Moon (1999, 2000), Pedroni (1996, 1997^a y 1997^b), y Kao & Chiang (2000) muestran que el *mean group estimator* que ellos proponen es más eficiente que el estimador que arroja una regresión *cross-section*.

Los modelos de cointegración en paneles se preocupan por estudiar relaciones económicas de largo plazo, típicas en el análisis de datos macroeconómicos y financieros. Estas relaciones de largo plazo son a menudo predichas por la teoría económica. Por ende, el interés de la investigación empírica radica en la estimación de los coeficientes de regresión para luego evaluar si las restricciones teóricas se satisfacen. Kao y Chen (1995) mostraron que el estimador *OLS* (*ordinary least squares*) en modelos de paneles cointegrados resulta asintóticamente normal pero **sesgado**. Chen, McCoskey y Kao (1999) encuentran que el estimador *OLS* corregido por sesgo no mejora respecto del estimador *OLS* en general. Los autores sugieren el uso del estimador *fully modified OLS* (FMOLS) o el *dynamic OLS* (DOLS). Phillips y Moon (1999) y Pedroni (1996) proponen el estimador FM como una generalización del de Phillips y Hansen (1990). Kao y Chiang (2000) estudian la distribución límite en una regresión de cointegración de acuerdo al estimador FM y muestran que es asintóticamente normal.

⁴ El supuesto de independencia de las unidades es crucial para satisfacer los requerimientos de Lindeberg-Levy del teorema central del límite en la elaboración del estadístico de raíz unitaria y de los estimadores y tests que son promedios de las relaciones individuales (Baltagi y Kao, 2003).

Asimismo, Pedroni (1996) y Phillips et al. (1999) también obtienen resultados similares para la distribución límite del estimador FM. Particularmente, el trabajo de Phillips et al. (1999) realiza un análisis pormenorizado de los diferentes tipos de relaciones que pueden verificarse en paneles no estacionarios.⁵ Los autores requieren que $N/T \rightarrow 0$, de ahí que los resultados sean válidos para paneles con N moderado y T grande, i.e. un panel macro, y no al revés (el típico panel micro). Dentro de los estimadores posibles, el *mean group estimator* (MG) propuesto en Pesaran y Smith (1995) o como lo llaman Phillips et al. (1999, 2000) *the average long run estimator* consiste básicamente en estimar el modelo de serie de tiempo $y_{it} = \eta_i + \lambda_i x_{it} + u_{it}$ para cada país y luego obtener el $\hat{\lambda}$, como $\hat{\lambda} = \sum_i \frac{\hat{\lambda}_i}{N}$. La $E(\hat{\lambda}_i) = \lambda$ representa el comportamiento promedio de los países. El estimador MG es un estimador consistente, aún cuando los $\hat{\lambda}_i$ no lo sean.

Al igual que en los tests de raíces unitarias presentados anteriormente, el supuesto clave para obtener la consistencia de los estimadores radica en la **independencia de los individuos cross-section**, de modo de añadir información al promediar los parámetros estimados que surgen del análisis de series de tiempo individual, atemperando la virtual correlación espuria. En este sentido, Phillips et al. (1999, 2000) proponen una variante sobre el estimador MG, al que denominan FMOLS. La idea detrás del estimador FMOLS es, además de su habilidad para contemplar la heterogeneidad entre las unidades del panel, la capacidad para controlar por el sesgo inducido por la potencial endogeneidad de los regresores y la correlación serial y heteroscedasticidad de los residuos (Pedroni, 2000, 2001, 2007).⁶

Uno de los resultados atractivo de los paneles no estacionarios, es que varios de los tests estadísticos y estimadores convergen en distribución a la normal. Tal es el caso de la distribución de estadístico de IPS como se mencionó anteriormente y de los estimadores FM y dynamic OLS (DOLS) (Kao y Chiang, 2000). Esta convergencia asintótica es un fuerte contraste con el comportamiento de los test de raíces unitarias y con los problemas de correlación espuria de los modelos de series de tiempo.

La teoría econométrica hasta aquí desarrollada para los tests de raíces unitarias en panel y la convergencia asintótica a la distribución normal de los

⁵ Los autores permiten que las series bajo análisis cointegren o no, y presentan un marco para el estudio secuencial y conjunto de la teoría asintótica en paneles no estacionarios. El modelo de panel considera cuatro casos (i) regresión de panel espuria, i.e. no hay cointegración entre las series, (ii) cointegración en panel heterogénea, i.e. cada individuo tiene su propia relación de cointegración, (iii) cointegración homogénea en panel, (iv) cointegración en panel cercana a la homogénea.

⁶ Mientras el estimador MG (*mean group estimator*) tiene dinámica de corto plazo paramétrica, el estimador FM se basa en métodos no paramétricos para remover los efectos de la dinámica y cualquier tipo de endogeneidad de los residuos sobre los coeficientes de largo plazo (Smith y Fuertes, 2010). El estimador es superconsistente bajo cointegración y es robusto a la omisión de variables que no formen parte de la relación de cointegración (Pedroni, 2007).

estimadores propuestos que suponen pendientes heterogéneas, se fundaron en la independencia de las unidades (países) del panel, situación que rara vez se verifica en el estudio empírico de los paneles macro. La ausencia de independencia entre las unidades se conoce en la literatura como *cross-section dependence* y su presencia es natural en el estudio de este tipo de datos, e.g. el ciclo económico-financiero mundial a través de la globalización de la actividad económica, las áreas comunes de comercio, el avance tecnológico y los efectos derrame, etc. Desconocer la *cross-section dependence* (CSD), i.e. cierta estructura de correlación en el término de error entre las unidades debido a la presencia de factores comunes no observables, dilapida las ganancias de eficiencia por operar con un panel y conduce a estimadores inconsistentes de los parámetros, invalidando la inferencia teórica en los modelos de datos de panel (Kapetanios, Pesaran y Yamagata (2011) y Banerjee y Carrion-i-Silvestre (2011)). Nótese que la presencia de factores comunes no observables no es otra cosa que una variable omitida en la especificación del modelo a estimar.

Las primeras armas de los investigadores empíricos para lidiar con la ausencia de independencia de las unidades y remover los factores comunes no observables, consistió en la incorporación de las *dummies* temporales. Sin embargo esta solución asume homogeneidad de las pendientes, i.e. $\lambda_i = \lambda$. Otra de las propuestas consistió en restar a cada variable la media correspondiente,

i.e. $\tilde{y}_{it} = y_{it} - \bar{y}_t$, donde $\bar{y}_t = \sum_{i=1}^N y_{it} / N$, y similarmente a los regresores x_{it} . Esto se

conoce con el nombre de *de-meaning*, pero una vez más se puede probar que esta potencial solución a la estimación de los factores no observables impone el supuesto de homogeneidad en el impacto de los no observables sobre las unidades.

A fin de modelar explícitamente la *cross-section dependence* se apela a la estimación de factores comunes no observables mediante el uso de las técnicas de componentes principales (Coakley, Fuertes y Smith, 2002; Bai, 2004; Bai y Ng, 2004). Pesaran (2006) objeta la propuesta de componentes principales para la estimación de la *cross-section dependence* de Coakley et. al (2002), y muestra que una combinación lineal de factores comunes no observables puede ser bien aproximada por los promedios sobre los individuos de los regresores del modelo como de la variable dependiente. Esto condujo a un nuevo conjunto de estimadores referidos en la literatura como *Common Correlated Effect* (CCE), que esencialmente consiste en aumentar el modelo que se desea estimar mediante la incorporación de los promedios de los individuos en cada t del tiempo, tanto de la variable dependiente como de los regresores específicos de cada individuo.

Diferentes intensidades en los tipos de manifestaciones de la CSD pueden ilustrarse, v.gr. efectos vecindario, efectos de red, la influencia de una unidad dominante o simplemente factores comunes no observables.⁷ De acuerdo con Pesaran (2006) el modelo econométrico puede representarse como,

⁷ Se puede distinguir entre dos tipos de CSD: débil y fuerte. La CSD débil, implica que las dependencias son locales y declinan con N . Este podría ser el caso de las correlaciones

$$\begin{aligned}
y_{it} &= \eta_i' z_t + \lambda_i' x_{it} + e_{it} & i=1,\dots,N \quad t=1,\dots,T \\
e_{it} &= \gamma_i' f_t + \varepsilon_{it}
\end{aligned} \tag{2}$$

donde y_{it} es la observación de la i -ésima unidad en el momento t ; z_t es un vector de $k_z \times 1$ de variables que no difieren sobre las unidades, i.e. ordenada al origen, tendencia y/o *dummies* estacionales; x_{it} es un vector de $k_x \times 1$ de regresores observables específicos de cada unidad al momento t ; f_t es un vector $r \times 1$ de factores no observables, que pueden influenciar a cada unidad de manera diferencial y además estar correlacionados con los x_{it} ; y ε_{it} es el disturbio no observable con $E(\varepsilon_{it})=0$, $E(\varepsilon_{it}^2)=\sigma_i^2$, el cual se distribuye independientemente a través de i y t . La covarianza entre los errores e_{it} está determinada por el *loading factor* γ_i . Puede observarse además que si f_t está correlacionado con x_{it} , como generalmente ocurren en muchas de las aplicaciones empíricas, v.gr. ciclos globales, desconocer la CSD al omitir el factor f_t , resulta en estimadores de λ_i sesgados e inconsistentes.

La idea de Pesaran (2006) es tratar al(los) factor(es) no observable(s) como parámetro(s) molesto(s) que uno desea controlar con el objetivo de tener una mejor estimación de los λ_i , a diferencia de quienes intentan estimarlos. El estimador propuesto, *common correlated effect* (CCE), consiste, como se mencionó anteriormente, en enriquecer el modelo a estimar mediante la incorporación de los promedios *cross-section* en cada t del tiempo, a fin de controlar por los no observables. Esto involucra tanto a las variables independientes como dependientes, como se muestra a continuación.

$$y_{it} = \eta_i' z_t + \lambda_i' x_{it} + \delta_{oi} \bar{y}_t + \delta_i' \bar{x}_t + u_{it} \tag{3}$$

Para entender la motivación de este procedimiento, supongamos un solo factor y promediamos la ecuación (2) sobre las unidades

$$\bar{y}_t = \bar{\eta}' z_t + \bar{\lambda}' \bar{x}_t + \bar{f}_t + \bar{\varepsilon}_t + \frac{1}{N} \sum (\lambda_i - \bar{\lambda})' x_{it} \tag{4}$$

$$f_t = \bar{\gamma}^{-1} \left\{ \bar{y}_t - \left(\bar{\eta}' z_t + \bar{\lambda}' \bar{x}_t + \bar{\varepsilon}_t + \frac{1}{N} \sum (\lambda_i - \bar{\lambda})' x_{it} \right) \right\} \tag{5}$$

Luego \bar{y}_t y \bar{x}_t funcionan como *proxy* del factor no observable. Notar que la covarianza entre \bar{y}_t y ε_{it} tiende a cero con N , entonces para N grande no hay problemas de endogeneidad. Si bien esta formulación supone coeficientes heterogéneos, existen versiones homogéneas de los mismos (ver Eberhardt y Teal, 2011). Esta misma idea de incorporar los promedios a las regresiones por país, es la que también utiliza Pesaran (2007) para inmunizar al test de raíz unitaria IPS contra la presencia de no observables. Estos tests de raíz unitaria

espaciales, donde cada unidad está correlacionado sólo con sus vecinos, mientras que CSD fuerte implica que la dependencia afecta a todas las unidades.

que controlan por la *cross-section dependence*, se los conoce como tests de segunda generación.

En resumen, Pesaran (2006) permite además de la heterogeneidad sobre los regresores observables (i) que los efectos comunes no observables tengan impacto diferencial sobre las unidades, (ii) que los errores por individuo muestren correlación serial y heteroscedasticidad, (iii) y que no sea necesario que los regresores por individuo sean idénticos y/o se distribuyan independientemente a través de los individuos, cualidad relevante en el caso de análisis de paneles de países. Sin embargo, Pesaran (2006) asume que tanto los regresores específicos por individuo como los factores comunes no observables son estacionarios y exógenos. Kapetanios et al. (2011) extienden el análisis de Pesaran (2006) permitiendo procesos $I(1)$ de los regresores por individuo como de los efectos no observables. La extensión de los resultados por parte de Kapetanios et al. (2011) dista de ser trivial y apela a resultados intermedios bien distintos para derivar la distribución asintótica de los estimadores cuando los datos son $I(1)$ que cuando son $I(0)$. Pero sorprendentemente, simulaciones de Montecarlo sugieren que el método CCE propuesto por Pesaran (2006) para abordar la CSD es robusto a una amplia variedad de procesos generadores de datos. Este resultado contrasta respecto a las diferencias substanciales que existen en los modelos de series de tiempo para las distribuciones de procesos $I(1)$ vs. la de procesos $I(0)$.

Si bien esta segunda generación de tests de raíz unitaria tuvo en cuenta la ausencia de independencia de las unidades al admitir la presencia de factores comunes no observables, los mismos introdujeron nuevos desafíos tanto en la interpretación de los test de raíz unitaria como en la interpretación de los test de cointegración (Breitung et. al, 2008). Estos factores comunes no observables pueden mostrar un comportamiento estacionario, v.gr. los ciclos económicos globales, o un comportamiento no estacionario, v.gr. el avance tecnológico global. En caso de que el factor no observable presente un comportamiento $I(1)$, i.e revele una raíz unitaria, habría que considerar la posibilidad de que dicho factor cointegre al interior de cada unidad como así también entre unidades. De ahí que la interpretación de los tests de raíz unitaria de segunda generación difiera respecto a la interpretación estándar de un test de raíz unitaria. Retomemos la ecuación (1), pero ahora conscientes de la presencia de un factor no observable

$$y_{it} = (1 - \phi_i)\mu_i + \phi_i y_{i,t-1} + \gamma_i f_t + \varepsilon_{it}, \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (6)$$

Rechazar la hipótesis nula $\phi_i = 1$, en favor de la alternativa $\phi_i < 1$, podría encubrir razones bien distintas. Esto podría deberse a que (i) tanto y_{it} como f_t son procesos estacionarios, o bien que (ii) y_{it} y f_t son $I(1)$ y cointegrados. Y esto es independiente del método que se utilice para dar cuenta de la CSD.

Retomando la solución propuesta por Pesaran (2006), i.e *common correlated effect*, para controlar por la presencia de la CSD, la misma fue extendida a los test de raíz unitaria por Pesaran (2007) y Pesaran, Smith y Yamagata (2008). La ecuación correspondiente para evaluar la presencia de una raíz unitaria es

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \delta_{0i} \overline{\Delta y_t} + \delta_{1i} \bar{y}_t + \varepsilon_{it}, \quad i = 1, \dots, N \quad t = 1, \dots, T \quad (7)$$

i.e. la ecuación convencional aumentada por los promedios de las unidades, tanto del regresor y_{it} como de la variable dependiente Δy_{it} . La hipótesis sería evaluar $\beta_i = 0$, utilizando algún test de panel. El propuesto por Pesaran (2007) consiste, al igual que el *IPS* (2003), en promediar los estadísticos t_i correspondiente al β_i de la ecuación (7). El nuevo estadístico, al que Pesaran denomina *CIPS* (*cross-sectional Im Pesaran y Shin*), tiene una distribución no estándar, aún con N grande. Esto se diferencia del resultado obtenido por Im et. al (2003), que bajo el supuesto de independencia de las unidades, el estadístico *IPS* se distribuye de acuerdo a una normal para N suficientemente grande.

Nótese que la ecuación (7) podría pensarse como un modelo de corrección al equilibrio, donde \bar{y}_t y y_{it} podrían ser $I(1)$, a pesar de que $\beta_i < 0$, simplemente porque cointegran. Esto último desalienta el uso de los test de raíz unitaria en panel ya que se dificulta la interpretación de los mismos. Tanto bajo H_0 como bajo la alternativa se estaría frente a hipótesis conjuntas: bajo H_0 se evalúa en simultáneo que todas las unidades son $I(1)$ y no cointegran mientras que en la alternativa se tiene $\beta_i < 0$ con la posibilidad de que $y_{it} \sim I(1)$ cointegre con el factor no observable.

Sintetizando lo anteriormente expuesto, se puede afirmar que este nuevo enfoque de la econometría de paneles de series de tiempo logra reunir las dos líneas de trabajo postuladas hacia fines de los años 90. Por un lado, el cuestionamiento de la homogeneidad de los parámetros de un modelo de panel macro puede provenir tanto del impacto de los observables -regresores- como de los factores no observables -*factor loadings*-. Desconocer la potencial heterogeneidad de los regresores observables como la de los factores no observables, tiene implicancias aún más serias si las variables observables y/o los no observables son no estacionarias. Este punto puede ilustrarse con el siguiente ejemplo: una ecuación en niveles estimada mediante algún estimador *pooled estándar* impone parámetros comunes para todos los países e implanta al mismo tiempo errores no estacionarios si los verdaderos parámetros del modelo son heterogéneos y las variables son no estacionarias (Eberhardt and Teal, 2011). Concretamente, no dar cuenta de la heterogeneidad de los parámetros de los regresores observables del modelo puede conducir a la ruptura de la relación de cointegración entre los regresores y la variable dependiente y producir potenciales resultados espurios (Smith y Fuertes, 2010). Similarmente una ecuación en niveles estimada mediante algún estimador *pooled estándar* y aumentada con $T-1$ variables *dummies* impone una evolución común de los no observables a todos los países, creando errores no estacionarios si los verdaderos no observables presentan un comportamiento no estacionario.

3. Evaluación de la cross-section dependence: XTCSI⁸

Como expusimos en la sección anterior, si bien la *CSD* es más un hecho que una rareza en los paneles macro, varios tests se han desarrollado en la literatura a fin de evaluarla. El uso del test del multiplicador de Lagrange (LM) propuesto por Breusch-Pagan (1980) podría ser una de las herramientas. Este test consiste en promediar los coeficientes de correlación al cuadrado de los residuos de a pares. Este test fue diseñado en el contexto de ecuaciones aparentemente no relacionadas (SURE, Zellner, 1962), con N fijo y $T \rightarrow \infty$, i.e. N pequeño relativo a T . Pesaran (2004) muestra que el test LM presenta serias distorsiones de tamaño cuando N es grande relativo a T , situación que se verifica en muchas de las aplicaciones empíricas. A fin de combatir el sesgo del test LM, Pesaran (2004) propone otro test al que llama *CD*, que consiste en promediar las correlaciones de los residuos de a pares. Bajo la hipótesis nula, para T suficientemente grande, el estadístico *CD* converge en distribución a la $N(0,1)$, cuando $N \rightarrow \infty$. Sin embargo, como resalta Pesaran (2004) el test *CD* puede resultar inconsistente en varias alternativas relevantes.

Pesaran, Ullah y Yamagata (2008), partiendo del test LM de Breusch-Pagan, proponen un nuevo test LM que corrige el sesgo del anterior, en paneles con regresores estrictamente exógenos y errores normales. Simulaciones de Montecarlo analizan la potencia y el tamaño de los tres estadísticos disponibles. Los autores concluyen que el test LM ajustado por sesgo controla de manera exitosa el tamaño del test manteniendo una potencia razonable.

Siguiendo el documento de Pesaran et. al (2008), diseñamos el código **XTCSI**, que computa los tres estadísticos antes mencionados de la siguiente manera:

Considerando el siguiente modelo de datos en panel

$$y_{it} = \lambda_i' x_{it} + u_{it} \quad \text{para } i = 1, 2, \dots, N; t = 1, 2, \dots, T \quad (8)$$

Donde la matriz $X_i = (x_{i1}, \dots, x_{iT})'$ de regresores puede contener en su primera columna el vector unitario para la constante y en la segunda una tendencia.⁹ Para cada i , $u_{it} \sim IIDN(0, \sigma_{ii}^2)$, para todo t , sin embargo podrían estar *cross-section correlated*

⁸ Disponible en: <https://ideas.repec.org/c/boc/bocode/s457848.html>

⁹ La incorporación de la constante y la tendencia en la matriz X_i a diferencia de lo presentado en la ecuación (2), donde ambas aparecían explicitadas en la matriz Z_i , es a los fines de simplificar el cálculo matricial que se muestra para la obtención del estadístico LM_{adj} .

Breusch y Pagan (1980) propusieron el siguiente estadístico LM para testear bajo la hipótesis nula correlación cero de los *cross equation error*

$$LM = T \sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij}^2 \quad (9)$$

Donde $\hat{\rho}_{ij}$ es la estimación muestral de la correlación de los errores que surgen de la estimación MCO de la regresión para cada individuo del panel.

Se puede demostrar que bajo la hipótesis nula:

$$H_0 : Cov(u_{it}, u_{jt}) = 0, \text{ para todo } t \text{ y } i \neq j \quad (10)$$

el estadístico LM se distribuye asintóticamente como una chi-cuadrado con $N(N-1)/2$ grados de libertad. Sin embargo, es probable que el test LM exhiba distorsiones sustanciales de tamaño para N grande y T chico, una situación usual en aplicaciones empíricas.

Ullah (2004) provee técnicas unificadas para obtener los momentos exactos y aproximados de los estimadores econométricos y los test estadísticos. Pesaran, Ullah y Yamagata (2008) utilizan este enfoque para corregir el sesgo en muestras chicas del estadístico LM.

Asumiendo:

Supuesto 1: Para cada i , los errores, u_{it} , son independientes con media 0 y varianza $0 < \sigma_i^2 < \infty$.

Supuesto 2: Bajo la hipótesis nula definida por $H_0 : u_{it} = \sigma_i \varepsilon_{it}$, $\varepsilon_{it} \sim IIDN(0,1)$ para todo i y t .

Supuesto 3: Los regresores, x_{it} , son estrictamente exógenos tal que $E(u_{it} / X_i) = 0$ para todo i y t , donde $X_i = (x_{i1}, \dots, x_{iT})'$ y $X_i' X_i$ es una matriz definida positiva.

Los autores introducen la siguiente matriz idempotente de rango $T - k$

$$M_i = I_T - H_i ; H_i = X_i (X_i' X_i)^{-1} X_i' \quad (11)$$

Considerando el modelo de datos en panel (8) y bajo los supuestos 1 a 3 la media exacta y la varianza de $(T - k) \hat{\rho}_{ij}^2$ están respectivamente dadas por:

$$\mu_{Tij} = E[(T-k)\hat{\rho}_{ij}^2] = \frac{1}{T-k} Tr[E(M_i M_j)] \quad (12)$$

y

$$v_{Tij}^2 = Var[(T-k)\hat{\rho}_{ij}^2] = \{Tr[E(M_i M_j)]\}^2 a_{1T} + 2Tr\{E[(M_i M_j)^2]\} a_{2T}, \quad (13)$$

donde

$$a_{1T} = a_{2T} - \frac{1}{(T-k)^2}, \quad a_{2T} = 3 \left[\frac{(T-k-8)(T-k+2)+24}{(T-k+2)(T-k-2)(T-k-4)} \right]^2 \quad (14)$$

Utilizando (12) y (13) el test estadístico LM ajustado por sesgo se define como:

$$LM_{adj} = \sqrt{\frac{2}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^N \frac{(T-k)\hat{\rho}_{ij}^2 - \mu_{Tij}}{v_{Tij}} \quad (15)$$

Bajo los supuestos 1 a 3, asumiendo que H_0 definida por (10), primero $T \rightarrow \infty$ después $N \rightarrow \infty$ se tiene:

$$LM_{adj} \rightarrow_d N(0,1) \quad (16)$$

Para lidiar con el sesgo de N grande del test LM , Pesaran (2004) sugiere utilizar el estadístico CD definido por:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \left(\sum_{i=1}^{N-1} \sum_{j=i+1}^N \hat{\rho}_{ij} \right) \quad (17)$$

y muestra que bajo H_0 y para un T suficientemente grande $CD \rightarrow_d N(0,1)$ a medida que $N \rightarrow \infty$.

Sintaxis

`xtcsi depvar indepvars [if] [in] [, trend]`

Descripción

`xtcsi` implementa en paneles heterogéneos (balanceados) distintos test de *error cross-section independence*: i) el test LM Breusch y Pagan (1980); ii) el test LM ajustado por sesgo de Pesaran, Ullah y Yamagata (2008); iii) el test CD Pesaran (2004).

Opciones

trend especifica una tendencia lineal a ser incluida en cada modelo de regresión individual.

Resultados

Escalares

r(N_g) Número de individuos del panel
r(lm) estadístico LM de Breusch y Pagan (1980)
r(p_lm) *p-value* de chi -cuadrado con $N(N-1)/2$ grados de libertad
r(lm_adj) estadístico LM ajustado por sesgo de Pesaran, Ullah y Yamagata (2008)
r(p_lm_adj) *p-value* de Normal(0,1) a dos colas
r(lm_cd) estadístico CD Pesaran (2004)
r(p_lm_cd) *p-value* de Normal(0,1) a dos colas

4. El test de IPS en presencia de la cross-section dependence: XTCIPS¹⁰

Siguiendo a Pesaran (2007), implementamos en Stata el test CIPS y el CIPS* (variante truncada del CIPS) al que llamamos XTCIPS. Tal cual lo puntualizamos anteriormente, la distribución límite del estadístico CIPS no es normal, y los correspondientes valores críticos están tabulados en Pesaran (2007). El código está diseñado para paneles balanceados, pero podría adaptarse para paneles desbalanceados, aunque eso implicaría realizar una simulación de Montecarlo para la obtención de los valores críticos acorde a la estructura del panel. Eso fue lo que se hizo en Bebczuk, Burdisso y Sangiácomo (2012).

Sintaxis

```
xtcips varname [if] [in] , maxlags(#) bglags(numlist) [ q trend noc ]
```

Descripción

xtcips estima el test CIPS de raíces unitarias en paneles heterogéneos (balanceados) desarrollado por Pesaran (2007; Sección 4, p. 275-279) .

Hay tres especificaciones posibles:

- Caso I: modelos sin intercepto o tendencia (ver opción **noc**)
- Caso II: modelos con intercepto individual específico (*default*)
- Caso III: modelos con tendencia lineal (ver opción **trend**)

¹⁰ Disponible en: <https://ideas.repec.org/c/boc/bocode/s457850.html>

Permite establecer especificaciones dinámicas individuales en cada regresión basada en dos criterios alternativos (ver opción **maxlags**(#)):

- i) Test de Wald de la hipótesis lineal compuesta sobre los parámetros del modelo (*default*)
- ii) Test Portmanteau (Q) de ruido blanco (ver opción **q**)

Informa el valor p del test Multiplicador de Lagrange de correlación serial de Breusch-Godfrey de cada regresión individual (ver opción **bglags**(*numlist*))

La hipótesis nula es (homogéneo no estacionario):

$$H_0 : \beta_i = 0 \text{ para todo } i$$

vs. las alternativas:

$$H_1 : \beta_i < 0, \quad i = 1, \dots, N_1, \quad \beta_i = 0, \quad i = N_1 + 1, N_1 + 2, \dots, N$$

en la siguiente regresión de Dickey-Fuller *cross-section augmented* (CADF):

$$\Delta y_{it} = \alpha_i + \beta_i y_{i,t-1} + \delta_{0i} \Delta \bar{y}_t + \delta_{1i} \bar{y}_{t-1} + \varepsilon_{it}, \quad i = 1, \dots, N \quad t = 1, \dots, T$$

Opciones

maxlags(#) requerida. Entero positivo. Establece la especificación dinámica individual. Indica el número máximo de rezagos que se incluirán en el modelo a estimar para cada individuo. Entonces, *xtcips* determina el número de rezagos a incluir en cada regresión individual con un proceso iterativo de 0 hasta **maxlags**, basado en el nivel de significación del test establecido para seleccionar la dinámica, donde la estrategia sugerida es seleccionar el mayor rezago significativo ya sea, -i) al rechazar H_0 (al 5 % o inferior) en el test de Wald¹¹ o; ii) al no rechazar H_0 (al 95 % o superior) en el test Portmanteau (Q) de ruido blanco- o **maxlags**, lo que suceda primero.

bglags(*numlist*) requerida. Establece el orden de correlación serial a ser probado en el test de multiplicador de Lagrange de Breusch-Godfrey en cada regresión individual. Si se proporciona un solo valor (entero positivo), ese orden se utiliza

¹¹ Se estima el modelo con el número de rezagos especificado de acuerdo a $\text{maxlag}(L)$, y el planteo del test de Wald es el siguiente: primero testea la hipótesis nula del test $H_0 : \delta_{0i}^1 = \delta_{0i}^2 = \delta_{0i}^3 = \dots = \delta_{0i}^L = \delta_{1i}^1 = \delta_{1i}^2 = \delta_{1i}^3 = \dots = \delta_{1i}^L = 0$. Si no se rechaza H_0 , la especificación sugerida sería el DF estándar sin aumentar. En caso de rechazar H_0 , se testea entonces $H_0 : \delta_{0i}^2 = \delta_{0i}^3 = \dots = \delta_{0i}^L = \delta_{1i}^2 = \delta_{1i}^3 = \dots = \delta_{1i}^L = 0$. Si no se rechaza H_0 , la especificación sugerida sería el DF aumentado en un rezago. En caso de rechazar H_0 , se procede de igual manera hasta determinar el rezago máximo para el $\text{maxlag}(L)$ especificado.

para todos los individuos. Si se proporciona una lista de números, su longitud debe coincidir con la cantidad de individuos en el panel.

trend incluye una tendencia temporal en la ecuación estimada (Caso III)

q establece el test Portmanteau (Q) de ruido blanco como el criterio para determinar la especificación dinámica

noc suprime el término constante (Caso I)

Resultados

`xtcips` guarda los siguientes resultados en `r()`:

Escalares

| | |
|------------------------|----------------------|
| <code>r(cvag1)</code> | valor crítico al 1% |
| <code>r(cvag5)</code> | valor crítico al 5% |
| <code>r(cvag10)</code> | valor crítico al 10% |
| <code>r(cips)</code> | estadístico CIPS |

Matrices

| | |
|-------------------|--|
| <code>r(W)</code> | Diagnósticos de las regresiones individuales |
|-------------------|--|

Bibliografía

- Arellano M (2003). *Panel Data Econometrics (Advanced Texts in Econometrics)*. Oxford University Press
- Bai J. (2004). "Estimating cross-section common stochastic trends in non-stationary panel data." *Journal of Econometrics*, 122, 137-183.
- Bai J. and S. Ng (2004). "A Panic attack on unit roots and cointegration." *Econometrica*, 72(4) 1127-1178.
- Baltagi B. H (2008). *Econometric Analysis of Panel Data*, 4th Edition. John Wiley & Sons, Ltd.
- Baltagi B. H. and C. Kao (2003). "Nonstationary Panels, Cointegration in Panels and Dynamic Panels: A survey. In *Nonstationary panels, cointegration in panels and dynamic panels.*" Edited by B. H. Baltagi, *Advances in Econometrics*, Vol 15.
- Banerjee A. and J. Carrion-i-Silvestre (2011). "Testing for Panel Cointegration using Common Correlated Effects Estimators." Mimeo. Department of Economics, University of Birmingham.
- Bebczuk R., T. Burdisso and M. Sangiácomo (2012). "Credit vs. Payment Services: Financial Development and Economic Activity Revisited." Working Paper 56. BCRA.
- Breitung J. and M.H. Pesaran (2008). *Unit Roots and Cointegration in Panels*, in L. Matyas and P. Sevestre, *The Econometrics of Panel Data (Third Edition)*, Kluwer Academic Publishers.
- Breusch, T. S. and A. R. Pagan (1980). "The Lagrange multiplier test and its applications to model specification tests in econometrics." *Review of Economic Studies* 47, 239-53.
- Chen B., S. McCoskey and C. Kao (1999). "Estimation and Inference of a Cointegrated Regression in Panel Data: A Montecarlo Study." *American Journal of Mathematical and Management Sciences*, 19, 75-114.
- Choi I. (2001). "Unit Root Test for Panel Data." *Journal of International Money and Finance*, Vol 20, Issue 2: 249-272
- Coakley J and Fuertes A. M. (1997). "New panel unit root tests of PPP." *Economics Letters*. Volume 57, Issue 1, Pages 17-22
- Coakley J., A. M. Fuertes and R. P. Smith (2002). "A Principal Component Approach to Cross-Section Independence in Panels." Mimeo. Birkbeck College, University of London.
- Eberhardt M. and F. Teal (2011). "Econometric for Gumblers: A New Look at the Literature on Cross Country Growth Empirics." *Journal of Economic Surveys*, Vol 25, N°1: 109-155.

- Fisher, R.A. 1932. *Statistical Methods for Research Workers*. Oliver and Boyd, Edinburgh.
- Hadri K. (2000). "Testing for Stationary in heterogeneous Panel Data." *The Econometric Journal*, Vol 3, Issue 2, 148-161
- Hsiao C. (2003). *Analysis of Panel Data*. Cambridge University Press.
- Im K. S., M. H. Pesaran and Y. Shin (2003). "Testing for Unit Roots in Heterogeneous Panels." *Journal of Econometrics* 115(1): 53-74.
- Kao C. and B. Chen (1995). "On the Estimation and Inference for Cointegration in Panel Data when the Cross-Section and Time Series Dimension." Manuscript, Center for Policy Research, Syracuse University.
- Kao C. and M. Chiang (2000). "On the estimation and inference of a cointegrated regression in panel data. In *Nonstationary panels, cointegration in panels and dynamic panels*." Edited by B.H. Baltagi, *Advances in Econometrics*, Vol 15.
- Kapetanios G., M. H. Pesaran and T. Yamagata (2011). "Panels with non-stationary multifactor error structures." *Journal of Econometrics* 160: 326-348.
- Levin A. and C. F. Lin (1992). "Unit Root Test in Panel Data: Asymptotic and Finite Sample Properties." Unpublished manuscript, University of California, San Diego.
- Levin A., C. F. Lin, and C. J. Chu (2002). "Unit root tests in panel data: asymptotic and finite- sample properties." *Journal of Econometrics*, Vol. 108, pp. 1-24
- Maddala G. S. and S. Wu (1999). "A Comparative Study of Unit Root Tests with Panel data and A New Simple Test." *Oxford Bulletin of Economics and Statistics*, 61:631-652.
- Pedroni P. (1996). "Fully Modified OLS for Heterogeneous Cointegrated Panels and the case of Purchasing Power Parity." W.P. Department of Economics, Indiana University.
- Pedroni P. (1997^a). "Panel Cointegration: Asymptotic and Finite Sample Properties of Pooled Time Series Tests with and Application to the PPP Hypothesis." W.P. Department of Economics, Indiana University.
- Pedroni P. (1997^b). "Cross Sectional Dependence in Cointegration tests of Purchasing Power Parity in Panels." W.P. Department of Economics, Indiana University.
- Pedroni P. (2000). "Fully Modified OLS for Heterogeneous Cointegrated Panels. In *Nonstationary panels, cointegration in panels and dynamic panels*." Edited by B.H. Baltagi, *Advances in Econometrics*, Vol 15.
- Pedroni P. (2001). "Purchasing Power Parity Tests in Cointegrated Panels." *Review of Economic and Statistic*, 83: 1371-1375

- Pedroni P. (2007). "Social capital, barriers to production and Capital shares: Implications for the Importance of Parameter Heterogeneity from a Nonstationary Panel Approach." *Journal of Applied Econometrics*, Vol 22, N°2: 429-451
- Pesaran M. H., Y. Shin and R. Smith (1999). "Pooled mean group estimation of dynamic heterogeneous panels." *Journal of the American Statistical Association* 94 (446): 289-326.
- Pesaran M. H. (2006). "Estimation and Inference in large heterogeneous panels with a multifactor error structure." *Econometrica* 74(4): 967-1012
- Pesaran M. H. (2007). "A simple panel unit root test in the presence of cross-section dependence." *Journal of Applied Econometrics* 22: 265-312.
- Pesaran M. H. and R. Smith (1995). "Estimating long-run relationships from dynamic heterogeneous panels." *Journal of Econometrics* 68(1): 79-113.
- Pesaran M. H., A. Ullah and T. Yamagata (2008). "A bias-adjusted LM test of error cross-section independence." *Econometrics Journal* (2008), volume 11, pp. 105-127.
- Pesaran, M. H. (2004). General diagnostic tests for cross section dependence in panels. CESifo Working Papers No.1233, 255-60.
- Phillips P. C. B. and B.E. Hansen (1990). "Statistical inference in instrumental variables regression with I(1) processes." *Review of Economic Studies* 57(1): 99-125
- Phillips P. C. B. and H. Moon (1999). "Linear regression Limit Theory for Nonstationary Panel Data." *Econometrica*, 67 1057-1111.
- Phillips P. C. B. and H. Moon (2000). "Nonstationary Panel Data Analysis: an Overview of Some Recent Developments." *Econometric Reviews* 19(3):263-286.
- Smith R. P. and A. M. Fuertes (2010). "Panel Time series." Mimeo, April, Centre for Microdata Methods and Practice.
- Ullah, A. (2004). "Finite Sample Econometrics." New York: Oxford University Press.
- Zellner, A. (1962). "An efficiency method of estimating seemingly unrelated regression equations and tests for aggregation bias." *Journal of the American Statistical Association* 57, 348-68.