

APLICACIÓN DE LA COINTEGRACIÓN AL ESTUDIO DEL VOLUMEN PASAJEROS NACIONALES EN EL AEROPUERTO “JORGE CHAVEZ”

Zegarra Román Levy¹ & Ysela Agüero Palacios²

ABSTRACT. *Se demuestra que, a través de la Cointegración, es posible evaluar la relación real existente entre el volumen de pasajeros nacionales en el Aeropuerto Internacional Jorge Chávez y el producto bruto interno del país, aún cuando ambas variables corresponden a procesos no estacionarios. Se identifica un equilibrio a largo plazo entre las variables antes mencionadas y se obtiene estimadores eficientes y proyecciones a través de un modelo multivariante sin incurrir en violaciones a los supuestos del Modelo de Regresión Lineal Múltiple.*

1. INTRODUCCIÓN

Cuando se estudia la posibilidad de explicar el comportamiento de una variable (dependiente) en función de otras variables registradas a través del tiempo - muchas de ellas no necesariamente estacionarias -, en raras ocasiones los modelos formulados pueden ser estimados por el método de mínimos cuadrados, debido a que, puede encontrarse resultados alentadores como coeficientes de determinación (R^2) próximos a uno, pero también una notoria autocorrelación de residuos, lo cuál es una violación directa del supuesto de incorrelación de errores, establecido al formular el modelo de regresión lineal. La violación de este supuesto conduce a estimadores ineficientes, y al establecimiento de relaciones espúreas.

El concepto de Cointegración relaciona los conceptos de estacionariedad e integrabilidad, con los avances de la modelización econométrica dinámica de los mecanismos de corrección del error y la estructura de las relaciones de equilibrio establecidas por la teoría económica, estableciendo relaciones a largo plazo “reales” entre variables económicas, resolviendo así, los problemas que pueden acarrear las relaciones espúreas. El objetivo de este artículo es presentar una aplicación de la técnica de cointegración en el estudio de la relación entre el volumen de pasajeros nacionales en

¹G.G. - Poder Judicial- e-mail:levy-zegarra@yahoo.com

²Univ. Nac. Mayor de San Marcos. Fac. de Ciencias Matemáticas - e-mail:yaguero@terra.com.pe

el Aeropuerto Jorge Chávez y el producto bruto interno peruano, con el fin de obtener proyecciones confiables sin incurrir en violaciones a los supuestos de la regresión lineal múltiple.

2. TEORÍA DE LA COINTEGRACIÓN

El hecho de que dos o más variables estén asociadas, puede explicarse por una relación de causalidad, pero también puede ocurrir que estas relaciones sean casuales, las cuales son conocidas como relaciones sin sentido o espúreas.

Granger y Newbold (1974), estudiaron las consecuencias que involucran relaciones sin sentido, o espúreas simulando un par de variables, Y_t y X_t , generadas por paseos aleatorios independientes, luego, formularon el modelo de regresión relacionando ambas variables.

$$Y_t = \alpha + \beta X_t + u_t \quad t = 1, 2, \dots, T \quad (2.1)$$

Al ajustar el modelo mediante el método de mínimos cuadrados se obtuvieron coeficientes de determinación (R^2) próximos de 1, pero los residuos estaban fuertemente correlacionados. Así, al ajustar el modelo (2.1) por el método de mínimos cuadrados ordinarios, es de esperarse que se concluya erradamente que el modelo es adecuado para expresar la relación entre las variables involucradas cuando en realidad lo que se tiene es una relación espúrea. Granger y Newbold llegaron a la conclusión de que los contrastes convencionales que confirman la hipótesis nula de independencia entre dos variables no son potentes cuando las variables son generadas por procesos integrados de orden 1 o mayor, independientes estadísticamente. Esto llevó a los autores a ajustar un modelo de regresión relacionando las primeras diferencias finitas de ambas variables, lo cual dio buenos resultados. Este fue el punto de partida de la teoría de cointegración.

A continuación se presentan algunos conceptos básicos de la teoría de cointegración.

Definición 2.1.- Un proceso es integrado de orden d si su diferencia finita de orden " d " admite una representación ARMA estacionaria, y se denota como,

$$X_t \sim I(d) \quad (2.2)$$

Por ejemplo, un proceso con una raíz unitaria será denotado como $I(1)$ y uno estacionario como $I(0)$.

Dos propiedades importantes de los procesos integrados son las siguientes:

a) Si $X_t \sim I(d)$, $\Rightarrow aX_t + b \sim I(d)$; $(a, b) \in \mathcal{R}^2$

b) Sean $Y_t \sim I(d')$ y $X_t \sim I(d)$ tal que $d < d' \Rightarrow (X_t + Y_t) \sim I(d')$ (Prop. dominante).

En la metodología de Box y Jenkins, la identificación del orden de integrabilidad de una serie de tiempo, usualmente se realiza mediante el uso de gráficos exploratorios, y contrastes de hipótesis formales tales como los de Box y Pierce, y Box y Ljung (Box y Jenkins (1976)). Otros contrastes de hipótesis ampliamente utilizados en el contexto de la econometría son los de Durbin-Watson (DW) (Sargan y Barghava (1983)), de Dickey-Fuller (DF), (Dickey y Fuller (1979)), Schmit-Phillips (SP) (Schmit & Phillips (1992)), entre otros.

2.1 Cointegración

De acuerdo con la propiedad dominante de los procesos integrados, al realizar una combinación lineal de dos procesos del mismo orden de integrabilidad, esto es, dados X_t e Y_t integrados de orden d , entonces la combinación lineal $Z_t = Y_t - \alpha X_t$ es integrada de orden d .

La Cointegración constituye una excepción al caso general, esto es, si dos procesos, $X_t \sim I(e)$ e $Y_t \sim I(d)$ están cointegrados, la combinación lineal Z_t dada por $Z_t = Y_t - \alpha X_t$ puede ser de un orden de integrabilidad menor, esto es, $Z_t \sim I(d - e)$, donde $d > e > 0$.

A continuación, se presenta una definición formal de Cointegración, propuesta por Engle y Granger.

Definición 2.2.- “Las componentes de un vector Y_t , m dimensional, se dicen Cointegrados de órdenes $d > 0$ y $e > 0$, denotado por $Y_t \sim CI(d, e)$, si:

- i. Todos los componentes de Y_t son del mismo orden d , ($I(d)$).
- ii. Existe un vector \mathbf{a} no nulo, tal que $\mathbf{a}'Y_t = Z_t \sim I(d - e)$, con $e > 0$.”, donde el vector \mathbf{a} es denominado “vector de cointegración”.

Ejemplo.- Sean el vector $Y_t' = (Y_t, X_t)$ cuyos elementos son procesos $I(1)$, y $\mathbf{a}' = (1, -\alpha)$ un vector de constantes. La combinación lineal $\mathbf{a}'Y_t = Y_t - \alpha X_t = Z_t$ es un proceso estacionario ($I(0)$).

En otras palabras, esto quiere decir que, aunque los procesos X_t e Y_t tienen varianzas infinitas, existe una relación de equilibrio a largo plazo entre ellas, de modo que las situaciones de desequilibrio son de carácter estacionario, y por tanto transitorias.

Luego, la existencia de una relación de cointegración entre dos variables puede interpretarse como la existencia de una relación de equilibrio lineal entre ellas, cuantificada por el vector de cointegración.

2.2 Identificación de la Cointegración

Antes de identificar un modelo para la serie de tiempo estudiada es importante establecer si esta presenta raíces unitarias y establecer si dos o más variables están cointegradas. Con este fin podemos utilizar métodos exploratorios y posteriormente podemos utilizar contrastes de hipótesis para confirmar la presencia de raíces unitarias.

2.2.1 Análisis Exploratorio

El gráfico de la serie de tiempo, el de nivel - dispersión, el correlograma (FAS), ayudan a visualizar si los datos han sido generados por un proceso estacionario, tanto en la parte ordinaria como estacional, a determinar el orden de diferenciación necesario para transformar una serie que presenta raíces unitarias en estacionaria, si dos o más variables están cointegradas, etc.

2.2.2 Análisis Confirmatorio

Consistirá en la aplicación de los contrastes de raíces unitarias de Durbin-Watson (DW) y de Dickey-Fuller (DF) a los residuos de la Cointegración.

A. Contraste de Cointegración de Durbin-Watson (DW)

Propuesto por Sargan y Bhargava (1983), se basa en el Estadístico de Durbin Watson (DW) del análisis de regresión. Este contraste considera una hipótesis nula de "No Cointegración", es decir la existencia de raíz unitaria en la componente aleatoria del modelo propuesto. El contraste se realiza mediante la delimitación de cotas superior e inferior - como en el contraste DW de autocorrelación- para su comparación, pero estas cotas tienden a separarse entre sí en cuanto aumenta el número de variables explicativas. Una gran ventaja de este contraste es su invariancia frente a la posible inclusión de constantes y tendencias en el modelo, sin que por ello varíen sus valores críticos. La regla de decisión es, rechazar la hipótesis nula si el estadístico DW es significativamente mayor que cero.

B. Contraste de Cointegración de Dickey-Fuller (DF)

Propuesto por Dickey y Fuller (1979), asume un modelo de la forma,

$$\Delta \hat{u}_t = \gamma_1 \hat{u}_{t-1} + \sum_1^p \gamma_{2i} \Delta \hat{u}_{t-1} + a_t$$

donde \hat{u}_t son los errores estimados de la cointegración obtenidos por MCO, p es el número de retardos, γ_1 y γ_{2i} son los parámetros del modelo (autocorrelaciones), a_t un término del proceso de ruido blanco. Considerando un modelo AR(1) en los residuos, la hipótesis nula de no cointegración, es decir la existencia de raíz unitaria en los residuos.

$$H_0 : \gamma_1 = 0 \text{ Vs. } H_1 : \gamma_1 < 0$$

El estadístico de prueba y la regla de decisión serán los mismos que se utilizan en la prueba habitual de raíces unitarias de Dickey-Fuller (DF). Las tablas de los valores críticos con los que comparan la estadística t del coeficiente γ_1 aparecen en Engle y Granger (1987), y corrigen a los de Fuller (1976), debido a que son generados mediante un procedimiento que minimiza la varianza residual.

3. ESTIMACIÓN DEL VOLUMEN DE PASAJEROS NACIONALES EN EL AEROPUERTO INTERNACIONAL JORGE CHÁVEZ

La necesidad de obtener predicciones confiables del volumen de pasajeros nacionales en el Aeropuerto Jorge Chávez (VPN), condujo a considerar un conjunto de variables que tuvieran relación con el entorno económico del país. Se formularon y estimaron diversos modelos, encontrándose estimaciones ineficientes y problemas en la interpretación econométrica de las estimaciones de los parámetros, causados directamente por la propia naturaleza no estacionaria de la serie histórica del volumen de pasajeros nacionales.

Una variable que encontramos que está fuertemente relacionada con la variación del volumen de pasajeros a nivel nacional (VPN) es el poder adquisitivo de los agentes económicos dentro del país (explicada macroeconómicamente por el producto bruto interno peruano (PBI)).

El ajuste de un modelo de regresión lineal relacionando ambas variables dio como resultado la ecuación:

$$VPN_t = -1.243088 + 389.459(PBI)_t$$

El coeficiente de determinación indicó que el 70.6% de la variación del VPN es explicada por las variación del PBI ($R^2 = 0.7059$). Sin embargo, el valor de la estadística de Durbin Watson ($DW = 1.072$) indica una marcada autocorrelación

de residuos, y por ende, inconsistencia en los estimadores de los parámetros. Ante este hecho, existe evidencia para sospechar que el modelo ajustado no será adecuado, mientras no se resuelva el problema de la autocorrelación de los errores.

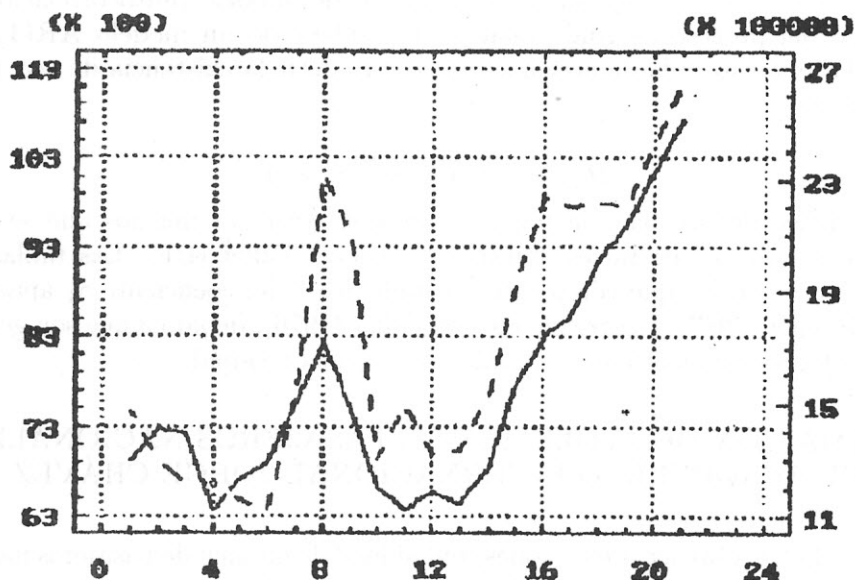


Figura N°1 - Evolución del Producto Bruto Interno Peruano (PBI) y del Volumen de Pasajeros Nacionales (VPN) en el Aeropuerto Jorge Chávez. Periodo : desde 1980 hasta el 2000.

La imposibilidad de ajustar un modelo de regresión relacionando estas dos variables sin violentar el supuesto de incorrelación de los errores, exhortaba hacia el uso de la técnica de cointegración, en un intento por evitar los inconvenientes causados por la introducción de variables no estacionarias en un modelo de regresión.

El análisis exploratorio del VPN y PBI (Fig. N° 1) reveló que ambas variables son no estacionarias, además aún cuando no tienen una tendencia clara ni obedecen a un patrón lineal, es evidente que existe una relación de equilibrio a largo plazo entre ellas durante el periodo de observación (1980 - 2000).

3.1 Contrastes de raíces unitarias

Para confirmar los hallazgos del análisis exploratorio se procedió a realizar dos contrastes de raíces unitarias.

a) Contraste de Raíces Unitarias de Durbin Watson

Bajo la hipótesis nula de no estacionariedad, se consideraron los modelos

$$PBI_t = C + u_t \quad VPN_t = C + u_t,$$

donde C y C'son constantes. Las estadísticas de prueba de DW para cada modelo fueron 0.334 y 0.543 respectivamente. Ambos valores están próximos de 0, y son mayores que los valores críticos de la prueba para un nivel de significación del 1% (0.386 y 0.511, respectivamente). Finalmente, para ambas variables, existe evidencia estadísticamente significativa para rechazar la hipótesis nula de estacionariedad, y aceptar la existencia de raíces unitarias, en ambas series de tiempo.

b) Contraste de Raíces Unitarias de Dickey-Fuller

Bajo la hipótesis nula de existencia de una raíz unitaria, se consideraron los modelos

$$\begin{aligned} \Delta PBI_t &= v(PBI_{t-1}) + a_t \\ \Delta VPN_t &= v'(VPN_{t-1}) + a_t. \end{aligned}$$

El siguiente paso será estimar los parámetros v y v' de los modelos formulados y verificar si son estadísticamente significativos.

Los modelos ajustados son;

$$\begin{aligned} \Delta PBI_t &= 0.0053(PBI_{t-1}) + a_t \\ \Delta VPN_t &= 0.0058(VPN_{t-1}) + a_t. \end{aligned}$$

Observamos que ambas estimaciones están próximas de cero y no hay evidencia estadísticamente significativa la afirmar que los parámetros sean diferentes de cero ($T = 0.91$ y $T = 0.79$ para el PBI y VPN, respectivamente, $\alpha = 0.05$), entonces concluimos que ambas series son integradas de orden 1.

3.2 Estimación del Modelo de Cointegración

Hasta ahora, las pruebas indican que tanto el PBI como el Volumen de pasajeros nacionales tienen un comportamiento $I(1)$. Según la técnica de Cointegración, si estas dos variables están cointegradas, al realizar una regresión entre ambas, se esperaría que los residuos resultantes correspondan a una realización de un proceso $I(0)$. El modelo de cointegración formulado fue

$$\Delta VPN_t = a + b(\Delta PBI_t).$$

Estimado los parámetros del modelo a partir de las series mensuales observadas del PBI y volumen de pasajeros nacionales durante el periodo de 1980 al 2000, se obtuvo la ecuación de regresión siguiente:

$$\Delta VPN_t = -1.92468e^{-4} + 412.19498(\Delta PBI_t) \quad (3.1)$$

El modelo ajustado explica el 72% de la variación en el volumen de pasajeros nacionales, las estimaciones son consistentes, los residuos están incorrelacionados (DW = 2.162) y tienen varianza mínima. Así este modelo será el adecuado para explicar la variación de VPN y realizar pronósticos muy próximos de la realidad. Asimismo, cabe recalcar que los coeficientes estimados fueron significativos a niveles de confianza superiores al 95%.

3.3 Verificación del Modelo

La estacionariedad de los residuos de la cointegración obtenidos en el modelo (3.1) garantizaría un equilibrio a largo plazo entre el Volumen de Pasajeros Nacionales (VPN) y el Producto Bruto Interno (PBI). A continuación se muestran los resultados de las pruebas de hipótesis de Durbin-Watson y de Dickey-Fuller aplicadas a los residuos del ajuste del modelo B.

a) Contraste de Cointegración de Durbin-Watson (DW).

El valor obtenido del DW fue de 1.80, el cual supera a los valores críticos (0.386 y 0.511 a un nivel de significación de 5% y 1%, respectivamente). Este hecho implicaría la existencia de estacionariedad en los residuos ($t \sim I(0)$) y el rechazo de la hipótesis nula de no cointegración, propuesta por este contraste.

b) Contraste de Cointegración de Dickey-Fuller (DF)

Se realizó el contraste sobre un modelo sin componentes determinísticas, obteniéndose $\Delta \hat{u}_t = 0.88342(\hat{u}_{t-1}) + e_t$, donde $\tau(\hat{\nu}) = 11.3456$.

La estimación puntual de ν arroja un resultado cuyo valor numérico es próximo a 1, el cual provee de evidencias sobre la no existencia de raíces unitarias. El valor de $\tau(\hat{\nu})$ supera ampliamente al valor de las tablas (-3.37 a un nivel de significancia de 5%), concluyendo que los residuos de la cointegración son estacionarios ($\hat{u}_t \sim I(0)$), y rechazando la hipótesis nula de no cointegración.

4. CONCLUSIONES

- a) El concepto de Cointegración conjuga tanto los conceptos de estacionariedad y orden de integrabilidad implícitos en la metodología de análisis de serie de tiempo propuesta por Box y Jenkins como la modelización econométrica dinámica (MCE), con las ideas de relaciones estructurales de equilibrio establecidas por la teoría económica y que se manifiestan en las variables.
- b) Tras los resultados obtenidos, llegamos a la conclusión que el Volumen de Pasajeros Nacional (VPN) y el Producto Bruto Interno (PBI) están cointegrados de la forma $CI(1,1)$, es decir, aún cuando estas dos variables son no estacionarias, podemos inferir que en el largo plazo, la relación entre estas dos variables se mantiene constante, aún cuando el comportamiento en el corto plazo de cada uno de estos se muestra inestable.
- c) Es necesario prestar mayor atención al peligro que supone establecer relaciones que en realidad son espúreas, es decir relaciones de casualidad y no de causalidad, cuando las series de tiempo son integradas, dado que el análisis de regresión puede producir resultados erróneos, cuando no se verifican las suposiciones establecidas en la formulación del modelo.

BIBLIOGRAFÍA

- [1] G.E.P. BOX - G.M. JENKINS. *Time Series Analysis : Forecasting and Control*. University of Wisconsin- U.S.A, University of Lancaster, U.K.(1976).
- [2] D.A. DICKEY - W.A. FULLER . *Distribution of the Estimators for Autorregresive Time Series With a Unit Root*. Journal of the American Statistical Association. Vol.74. (1979). pp.427-431.
- [3] D.A. DICKEY - W.A. FULLER . *Likelihood Ratio Statistics for Autoregressive Time Series with a Unit Root*. Econometrica. Vol. 49,(1981). pp.1057-1072.
- [4] J. ENGLE - C. GRANGER. *Cointegration and Error Correction: Representation, Estimation and Testing*. Econometrica. Vol. 50,(1987). pp.1057-1072.
- [5] C. GRANGER - P. NEWBOLD. *Spurious Regressions in Econometrics*. Journal of Econometrics. Vol.2, (1974). pp.111-120.
- [6] T. MILLS *The Econometric Modelling of Financial Time Series*. Department of Economics. University of Hull. Cambridge University Press.(1993).
- [7] J. SARGAN - A. BHARGAVA *Testing Residuals from Least Squares Regression for Being Generated by the Gaussian Random Walk*. Econometrica. Vol.51, (1983). pp.153-174.
- [8] P. SCHMIDT - P. C. B. PHILLIPS *LM Test for a Unit Root in the Presence of Deterministic Trends*. Oxford Bulletin of Economics and Statistics. Vol. 54, (1992). pp.257-288.
- [9] J. SURIÑACH - M. ARTIS - E. LÓPEZ - A. SANSÓ *Análisis Económico Regional: Nociones Básicas de la Teoría de la Cointegración*. Universidad de Barcelona-España.(1995).