

ANÁLISIS ECONOMETRICO TEÓRICO SOBRE LA HIPÓTESIS DE CONVERGENCIA

Ricardo Alan Solórzano Mendoza¹

RESUMEN

El presente trabajo tiene el objetivo de exponer desde una perspectiva teórica econométrica la naturaleza dinámica del llamado fenómeno de convergencia, para con ello exponer una modelización adecuada haciendo énfasis en el análisis de datos panel desde una perspectiva estocástica ortodoxa, pero que pueda abarcar aspectos argumentativos; como el del comercio internacional, que la propia teoría del crecimiento de la cual emana la hipótesis de convergencia carece. Así mismo, también se mencionan algunos otros enfoques estocásticos heterodoxos que estudian dicho fenómeno como el de las cadenas de Markov.

PALABRAS CLAVE: Econometría, Convergencia, Neoclásicos, Modelos de crecimiento económico.

CLASIFICACIÓN JEL: C01, O47, E13, O41.

INTRODUCCIÓN

El modelo de Solow tiene como corolario que siempre existirá lo que se ha llamado teóricamente convergencia económica. La anterior situación es predicha como un fenómeno invariante en el sentido de que siempre se generará una reducción de brechas en el ingreso per cápita entre las economías con similares características. En efecto, si las economías consideradas tienen parámetros similares, existirá un estado estacionario común y habrá convergencia. Tomando en cuenta la modelización de Solow, la primera propuesta econométrica para la medición de convergencia se presenta en el artículo de Mankiw, Romer y Weil (1992), de este trabajo han emanado análisis más

complejos y robustos como los de datos panel, pero para entender su esencia es necesario revisar cómo se presenta esa primera modelización.

Conforme se va haciendo el análisis más complejo existe la posibilidad de ir incorporando más variables al análisis, principalmente con el objetivo de ir dejando atrás varias de las críticas al modelaje ortodoxo del crecimiento, en especial aquellas sobre su mecanicismo y reduccionismo como lo argumenta Katz (2012); que señala que el enfoque de la teoría neoclásica que tradicionalmente se ocupa, se caracteriza por no señalar las interrelaciones que se crean para dar una explicación estructural de los procesos económicos, que im-

¹ Maestro y doctorante en Ciencias Económicas por la Universidad Autónoma Metropolitana. [lic.alan1193@gmail.com]

plicarían no solo tomar en cuenta los determinantes del estado estacionario de los modelos neoclásicos, sino también hacer un análisis integral que involucre otras variables de carácter no estrictamente económicas, como lo son la distribución geográfica de la población, amplitud territorial de las economías involucradas, distinción entre población rural y urbana, incluso en última instancia el idioma (Díaz, 2014; Jiménez & Narbona, 2007). De lo anterior, *mutatis mutandis*, el análisis se puede centrar en el terreno del comercio internacional, y esto por sí mismo ya es superar una de las más grandes barreras de simplicidad del análisis. En este sentido, se propone adicionar a la formalización una modelización gravitacional propuesta por Tinbergen (1962), que procure extender explicativamente otros aspectos.

Posteriormente se argumenta desde la perspectiva de datos panel, por que las modelizaciones de efectos fijos y aleatorios no son los adecuados para el análisis de la convergencia, y con ello aducir que será siempre necesario recurrir al tipo de modelaje de Arellano y Bond (1991). Lo anterior se vuelve necesario porque existen una gran variedad de trabajos empíricos sobre convergencia que utilizan la metodología de datos agrupados, de efectos fijos y aleatorios, por ejemplo el referente para el análisis estatal de México; Cermeño (2001), siendo que estos presentan problemas más graves que una subespecificación, ya que en realidad presentan endogeneidad, si bien estos muchas veces se muestran para fines comparativos o para reafirmar los resultados obtenidos con una estimación de paneles dinámicos, en realidad por construcción estos modelos estáticos no serían válidos.

Por último, se mencionan algunos estudios en el campo empírico más heterodoxo como soluciones distintas para el análisis de convergencia, que tiene sus propias fortalezas y debilidades, pero que merecen resaltar para también tenerlas en consideración, y con ello continuar la generación de nuevas alternativas que permitan una mejor cuantificación del fenómeno.

1. LA CONVERGENCIA DEL INGRESO PER CÁPITA

Para la teoría del crecimiento neoclásico el modelo de referencia es el de Solow, que surge en su versión original en 1956 como contraposición para lo que planteaba un modelo previo, el de Harrod Domar, este último proponía formalizar el crecimiento utilizando una función de producción con “la imposibilidad de sustituir el factor trabajo por capital” (Solow, 1956; 65). Por su parte Solow propuso sustituir lo anterior, considerando que el producto debe ser “producido por trabajo y capital bajo las condiciones neoclásicas estándar” (Solow, 1956; 66).

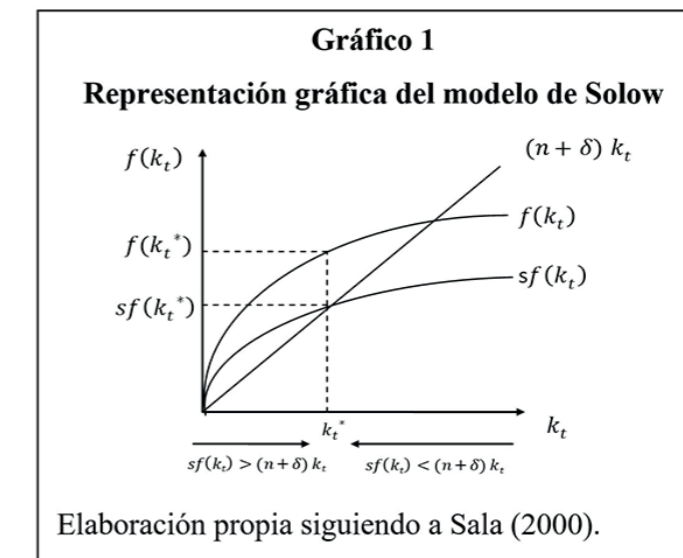
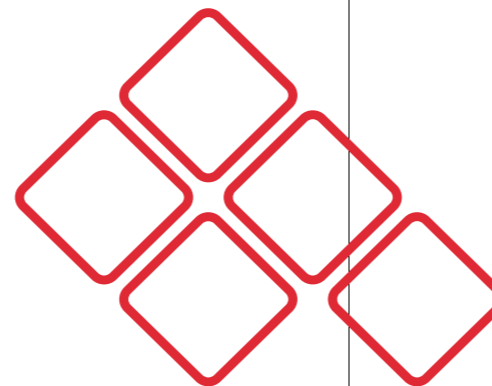
Posterior al planteamiento original de Solow surgieron formalizaciones que utilizaron de manera generalizada una función de producción Cobb-Douglas, destacándose los trabajos de Sala I Martin (2000), Gregory Mankiw (2000) y David Romer (2002). Sin embargo, estos tienen sus propias particularidades, como lo son el caso de incluir a la tecnología como una variable en el modelo o utilizar para el análisis unidades efectivas de trabajo.

En forma condensada (sin incluir a la tecnología), el modelo se especifica con cinco ecuaciones estructurales:

Ecuaciones estructurales del modelo de Solow		
Función de producción ²	[1]	$Y_t = K_t^\alpha L_t^{1-\alpha}$
Producto Interno Bruto (sin considerar exportaciones netas ni gasto de gobierno)	[2]	$Y_t = C_t + I_t$
Consumo agregado (con una tasa de ahorro constante “s”)	[3]	$C_t = (1 - s)Y_t$
Evolución del capital en el tiempo (a una tasa de depreciación constante “δ”)	[4]	$\dot{K}_t = I_t - \delta K_t$
Evolución de la población en el tiempo	[5]	$\dot{L}_t = nL_t$
Variables:	Y_t, K_t, L_t, C_t, I_t	
Parámetros:	s, α, δ, n	
Elaboración propia, a partir de lo expuesto por Sala (2000).		

Despejando la inversión bruta de la ecuación [4] y sustituyéndola, al igual que las ecuaciones [1] y [3] en la ecuación [2], considerando que la función de producción es homogénea de grado uno y que la definición de capital per cápita es $k_t = \frac{K_t}{L_t}$, entonces puede expresarse la evolución del mismo como $\dot{k}_t = sk_t^\alpha - (n + \delta)k_t$ a esta expresión también se le llama “ecuación fundamental del modelo de Solow”, misma que se compone de la resta de la inversión y de la inversión de reposición per cápita. Cuando se tiene la igualdad $sk_t^\alpha = (n + \delta)k_t$ es cuando la economía ha alcanzado el “estado estacionario”, lo cual implica que el cambio del capital en el tiempo es nulo $\dot{k}_t = 0$.

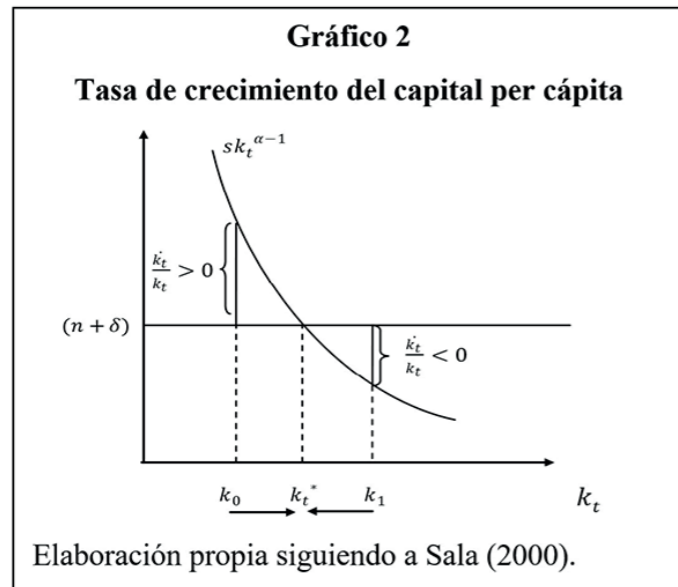
Con lo anterior, es posible obtener los niveles de capital y de la producción en el estado estacionario, sin embargo, en el presente análisis se referirá a los mismos como k^* y $y^* = f(k^*)$.



² Cumple las condiciones Inada, las cuales indican que a mayor cantidad de los factores de producción, entonces la productividad marginal de los mismos tendera a cero, y viceversa.

Al dividir la ecuación fundamental del modelo de Solow entre el capital per cápita, se obtiene la tasa de crecimiento del mismo $\frac{\dot{k}_t}{k_t} = s k_t^{\alpha-1} - (n + \delta)$. Para analizar su evolución hacia el estado estacionario, es preciso comparar lo que significa estar en dos situaciones: por encima y por debajo de k^* .

Cuando la economía se sitúa por debajo de k^* , el nivel de la inversión está por encima del de la inversión de reposición, por lo que $\frac{\dot{k}_t}{k_t} > 0$. En el caso contrario ocurrirá que $\frac{\dot{k}_t}{k_t} < 0$ (gráfico 2).



Utilizando la ecuación [1] puede demostrarse que la tasa de crecimiento del ingreso per cápita es proporcional al de la tasa del capital $\frac{\dot{y}_t}{y_t} = \alpha \frac{\dot{k}_t}{k_t}$, por lo que al describir lo que sucede con la tasa de crecimiento del capital, *mutatis mutandis*, ocurrirá lo mismo con la del ingreso, no importa el nivel inicial del mismo, siempre tenderá al nivel correspondiente y^* .

Todo lo anterior tiene una connotación teórica muy importante, ya que si el análisis no se limita al de una sola economía, sino que se generaliza para varias, entonces se implica que aquellas con un estado estacionario similar tenderán a un nivel de ingreso en común (y^*), y dada la dinámica antes descrita, economías con un nivel inicial de ingreso per cápita elevado tendrán una tasa de crecimiento menor en comparación con las que tienen un nivel inicial pequeño, lo que da la entrada al concepto de convergencia económica.

Cabría hacer mención que el modelo de Solow no es el único modelo de crecimiento que aduce convergencia, otro importante referente teórico es el de Ramsey que describe una economía cerrada, teniendo como principal característica un análisis de optimización del consumo por parte de los hogares, y que a diferencia de Solow no considera una tasa de ahorro constante. Debido a la extensión del modelo, en este trabajo no se desarrolla la formalización de Ramsey³, pero en lo que respecta a la convergencia llega a las mismas conclusiones que Solow, es decir:

³ Para la formalización en extenso véase el trabajo completo de Sulca, M. y Toribio C. (2014).

“La tendencia de las tasas de ahorro a aumentar o disminuir con el desarrollo económico afecta a la dinámica de transición, por ejemplo, a la velocidad de convergencia hacia el estado estacionario. Aunque la tasa de ahorro aumente en el modelo de Ramsey, la propiedad de convergencia se sigue cumpliendo con las condiciones generales. Es decir, una economía tiende a crecer más de prisa en términos per cápita cuanto más lejos esté de su propio estado estacionario.” (Sulca y Toribio, 2014; 60).

Oded Galor (1996) presenta dos clasificaciones en torno al fenómeno de convergencia, y lo que en cada caso se ha de entender en el campo de la investigación empírica bajo un enfoque ortodoxo:

a) Hipótesis de convergencia absoluta: “los ingresos per cápita de los países convergen entre sí en el largo plazo, independientemente de sus condiciones de inicio” (Galor, 1996; 1056).

En investigación empírica, para un periodo determinado se refiere que al encontrar evidencia estadística de dicho fenómeno, éste se verifica sin tomar en cuenta el grado de similitud de los parámetros entre las economías analizadas, pues implícitamente estas tienden a un mismo estado estacionario.

La base teórica para plantear una regresión consistente con la predicción de convergencia del modelo de Solow se encuentra en el texto de Mankiw, Romer y Weil (1992), en el cual al considerar una economía sin crecimiento tecnológico y el nivel de ingreso esta expresado en términos per cápita, muestran que a partir de la aproximación del ingreso alrededor de su estado estacionario; expresado en lo que en realidad es una ecuación diferencial $\frac{d \ln(y(t))}{dt} = \lambda [\ln(y^*) - \ln(y(t))]$ con $\lambda = -\frac{\partial(\frac{\dot{k}_t}{k_t})}{\partial \log(k)}$, y que en esta última se utiliza una transformación de la tasa de crecimiento del capital $\frac{\dot{k}_t}{k_t} = s e^{(\alpha-1) \log(k)} - (n + \delta)$ debiéndose cumplir que han de igualarse la inversión *per se* y la de reposición para tener $\lambda = (1 - \alpha)(n + \delta)$, finalmente puede obtenerse, al resolver dicha ecuación diferencial y restar a ambos lados de la ecuación resultante el término $\ln(y(0))$, lo que los autores denominaron la ecuación que debe utilizarse para estimar la convergencia.

$$\ln(y(t)) - \ln(y(0)) = (1 - e^{-\lambda t}) \ln(y^*) - (1 - e^{-\lambda t}) \ln(y(0))$$

Misma que muestra que “en el modelo de Solow el crecimiento del ingreso es una función de los determinantes del estado estacionario y del nivel inicial del ingreso” (*Ibidem* ; 423). Para estimar esta regresión se hace menester utilizar Mínimos Cuadrados Ordinarios no lineales, pero para evitar utilizar esta técnica (por simplicidad empírica) generalmente se hace una transformación equivalente que permite observar si existe o no convergencia bajo los supuestos clásicos del modelo de regresión lineal simple, buscando verificar el teorema de Gauss-Márkov. En este respecto, se utiliza la regresión de corte transversal:

$$\dot{y}_i = k + \beta \log(y_{i,0}) + u_i \quad [1.A]$$

Dónde:

\dot{y}_i : Tasa de crecimiento del ingreso per cápita.

β : Coeficiente de convergencia.

$\log(y_{i,0})$: Logaritmo del ingreso en el periodo inicial de análisis.

u_i : Es un término de error estocástico con $u_i \sim N(0, \sigma_u^2)$.

Si $\beta < 0$ y es estadísticamente significativo, entonces puede afirmarse que hay evidencia estadística de convergencia absoluta.

b) Teóricamente se ha llamado hipótesis de convergencia condicional cuando “los ingresos per cápita de los países que son similares en sus características, convergen entre sí en el largo plazo independiente de sus condiciones iniciales” (Ídem).

Para el análisis empírico, para realizar la métrica correspondiente es necesario tomar en cuenta las posibles diferencias entre las economías que sean determinantes del estado estacionario, solo así podrá determinarse la existencia del fenómeno. En este caso se utiliza una regresión como la utilizada para verificar la convergencia absoluta, solo que se agregan más parámetros en la parte sistemática, como la tasa de crecimiento de la población o de la depreciación, etc. Es decir, se tendrá que:

$$\dot{y}_i = k + \beta \log(y_{i,0}) + \sum_{j=1}^k \beta_j x_j + u_i \quad [1.A.1]$$

Si $\beta < 0$ y es estadísticamente significativo, entonces puede afirmarse que existe convergencia condicional.

En resumen, si se generaliza que las economías llegan a un mismo y^* , implícitamente se está demostrando matemáticamente la existencia de convergencia, y esto puede medirse estocásticamente mediante las regresiones anteriormente expuestas. Sin embargo, algunos autores recurren a una explicación de convergencia desde la perspectiva de una economía abierta a pesar de utilizar un modelo de economía cerrada (ecuación [2]); ejemplo de lo anterior es cuando Romer (2002) presenta dos argumentos inmediatamente después de realizar la formalización del modelo de Solow:

i) Bajo las condiciones de Inada, “el modelo de Solow implica que la tasa de rendimiento del capital es menor en aquellos países que disponen de un mayor capital ... el capital tendría un incentivo para desplazarse de los países más ricos a los más pobres, lo que de nuevo hace más verosímil la convergencia” (Romer, 2002; 33).

ii) “...es posible que las diferencias internacionales en los niveles de renta se deban a que algunos países no están explotando las mejores técnicas disponibles. Estas diferencias tenderían a reducirse a medida que los países más pobres tuvieran acceso a las técnicas más innovadoras” (Ídem), obviamente mediante el comercio internacional.

Argumentos aventurados que no son producto del análisis del modelo, pues como se mencionó, el mismo en ningún momento muestra las interrelaciones económicas del comercio internacional, pero mismas que si pueden ser visibles mediante un modelo de regresión.

En el caso de datos panel, para obtener mayor robustez en el modelo generalmente se elige por analizar el caso dinámico, esto es optar por una especificación que agrega una variable dependiente rezagada en el modelo, considerando que para el caso del análisis de convergencia esa variable debe ser la tasa de crecimiento del ingreso per cápita, ya que se busca encontrar la tendencia de la tasa de crecimiento contemporánea de cada economía con respecto a sus anteriores.

Análogamente para el caso de datos panel existen *a priori* dos consideraciones posibles, una que corresponde al análisis de convergencia absoluta [2.A] y otra a la de convergencia condicional [2.B].

$$\dot{y} = \alpha \zeta + \delta y_{-1} + u \quad [2.A]$$

$$\dot{y} = \alpha \zeta + \delta y_{-1} + X\beta + u \quad [2.B]$$

Dónde:

α : Corresponde al intercepto común entre las diferentes economías.

δ : Es el parámetro para contrastar la hipótesis de convergencia, si este es negativo y estadísticamente significativo, entonces se dice que hay evidencia estocástica de convergencia en el ingreso per cápita.

ζ : Es un vector de unos.

\dot{y} : Vector de tasas de crecimiento per cápita contemporánea.

y_{-1} : Vector de tasas de crecimiento per cápita rezagadas un periodo.

X : Matriz de las variables independientes.

β : Vector de parámetros a estimar.

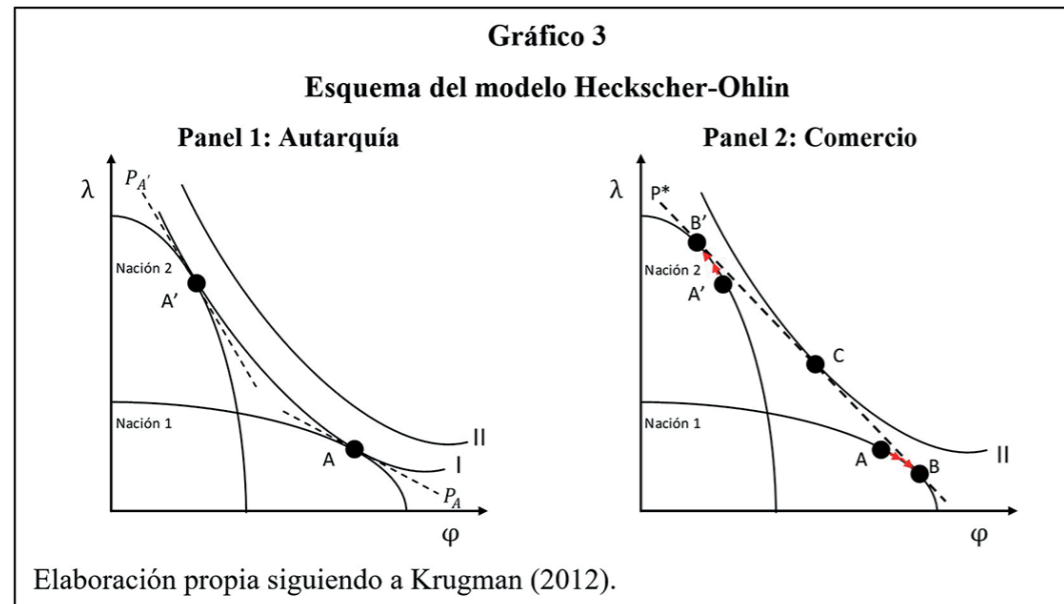
u : Vector de términos de errores compuestos, es decir que debe de considerar efectos inobservados.

La estimación de la regresión [2.B] permite un análisis más completo desde el punto de vista empírico, ya que por las propias limitaciones explicativas que pueda tener el modelo de crecimiento teórico que se utilice; puede que no incluya variables que sean estadísticamente significativas y que influyan en la dinámica de la convergencia. En este sentido, para el caso específico del modelo de Solow, el análisis empírico de convergencia puede enriquecerse mediante una incorporación paramétrica de un modelo gravitacional.

Los modelos gravitacionales son muy útiles para incorporar un análisis de economía abierta, de hecho tal modelaje surge a partir de que Gottfried Harberler adecuó el concepto original de la ley de la ventaja comparativa para expresarla en términos de costos de oportunidad, considerando que “El costo de una mercancía es la cantidad de una segunda mercancía a la que se debe renunciar para liberar los recursos estrictamente necesarios para producir una unidad adicional de la primera mercancía” Aquino (2012; 94), y en ese caso; “Un país tiene ventaja comparativa en la producción de un bien si el coste de oportunidad en la producción de este bien en términos de otros bienes es inferior en este país de lo que lo es en otros países.” (Krugman, 2012; 27).

Es a partir del concepto reformado Harberler que se representan las fronteras de posibilidades de producción de los diferentes países en un ámbito de comercio internacional, en este sentido y para hacer el análisis de forma esquemática del modelo H-O son necesarios los siguientes supuestos: se consideran solo dos factores de producción, trabajo y capital; existen dos economías con preferencias similares y cada una tiene abundancia de un solo factor de producción; los países utilizan la misma tecnología; únicamente dos mercancías son comercializadas: λ , φ , que

para su producción es necesario el uso intensivo de uno de los factores de producción; para la mercancía λ es trabajo, y para ϕ capital. De esta forma, considerando costos de oportunidad crecientes, se tendrán fronteras de posibilidades de producción cóncavas al origen (gráfico 3).



El panel 1 muestra que en una situación de autarquía, a los precios relativos internos P_A y $P_{A'}$, las naciones producirán y consumirán las combinaciones de mercancías correspondientes a los puntos A y A'. Es decir, donde haga tangencia la frontera de posibilidades de producción de cada economía y la curva de indiferencia en común (por el supuesto de preferencias similares).

El panel 2 introduce el análisis del comercio e indica que a un precio internacional (P^*), si la nación 1 es abundante en capital, siguiendo la ley de la ventaja comparativa, se especializará en la producción de ϕ ; mientras que la nación 2 que se asume abundante en trabajo se centrará en producir λ . De tal forma que cada nación se ha de beneficiar del comercio, pues con él producirán las combinaciones de bienes indicadas por los puntos B y B', y mediante el intercambio consumirán la miscelánea correspondiente al punto C, es decir que su consumo estará en una curva de indiferencia más alta que en autarquía.

De lo anterior se desprende el teorema H-O: "Un país exportará el bien que hace uso intensivo del factor de producción relativamente abundante en ese país, e importará el bien que hace uso intensivo del factor de producción relativamente escaso." (Díaz, 2014; 180).

Para comprobar empíricamente el cumplimiento del teorema anterior se ocupan modelos gravitacionales, los cuales están basados en la ley de gravitación universal propuesta por Isaac Newton. Tinbergen (1962) fue el primero en adaptar las fórmulas de Newton para estudiar los flujos del comercio internacional que por lo general incluyen como variable dependiente las exportaciones netas, y como variables explicativas al PIB del país exportador, del importador, los factores de producción, la distancia entre el país exportador hasta el país importador, el número de costas, y el idioma de los países (Díaz, 2014; Jiménez & Narbona, 2007). En ese sentido, en la especificación de [2.B] y en específico

en la matriz \mathbf{X} , es posible teóricamente la incorporación del análisis de una economía abierta con las anteriores variables mencionadas, introduciendo un análisis más completo empíricamente que el únicamente considerado en el modelo de Solow.

A continuación se explica porque se deben de agregar también los llamados Efectos Inobservados (EI) y porque los métodos estáticos más utilizados no son compatibles con el análisis de la hipótesis de convergencia en datos panel, pero que sin embargo erróneamente estos tienden a reportarse, y en muchos casos, compararse con las estimaciones dinámicas.

2. EL MODELAJE DE ARELLANO Y BOND COMO EJE DE ANÁLISIS EN DATOS PANEL PARA LA HIPÓTESIS DE CONVERGENCIA

El modelo base de datos panel (conocido como de datos agrupados o "pooled") sin la inclusión de especificidades, es decir sin incluir variables dicotómicas en las pendientes ni en el intercepto, es el siguiente:

$$\mathbf{y}_{[NT \times 1]} = \mathbf{X}_{[NT \times K]} \boldsymbol{\beta}_{[K \times 1]} + \mathbf{u}_{[NT \times 1]} \quad [1.B]$$

Dónde:

$\mathbf{y}_{[NT \times 1]}$: Vector de las observaciones de la variable dependiente.

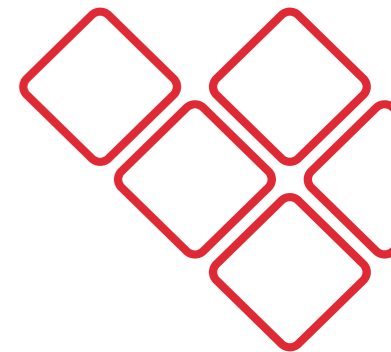
$\mathbf{X}_{[NT \times K]}$: Matriz de regresoras, con una columna de unos para considerar el intercepto.

$\mathbf{u}_{[NT \times 1]}$: Es un el vector de términos de errores estocásticos, cuya especificación de datos agrupados no considera EI, sin embargo, se rige bajo el supuesto de que sus elementos $u_{it} \sim N(0, \sigma_u^2)$.

$\boldsymbol{\beta}_{[K \times 1]}$: es el vector de parámetros a estimar que también tiene al elemento correspondiente al intercepto.

Sin embargo, cuando se trabaja empíricamente en el análisis de una economía, en primera instancia, debe de considerarse la existencia de los EI, ya que estos son parte esencial del estudio econométrico macroeconómico, esto es así porque el no incluirlos genera una subespecificación del modelaje repercutiendo en las estimaciones de los parámetros. Propiamente los EI son todos aquellos elementos que influyen en una variable macroeconómica, que en efecto, no pueden ser cuantificados o que simplemente influyen de manera agregada en el modelo y por algún motivo no es conveniente la desagregación; pero sabemos que existen, ejemplo de lo anterior es la propia cultura, raza, religión, tipo de gobierno, etc. de una economía en particular. Existen trabajos econométricos que plantean que en el modelaje de [1.B] si son considerados los EI⁴, pero en realidad si se mantiene dicha especificación, entonces se está suponiendo que tales EI no existen, o bien, que su agregación es idéntica a cero, no que todas las unidades de corte transversal compartan una misma constante con un mismo EI. Esta cuestión de especificación no es un asunto menor, pues si existen tales EI entonces existirá sesgo en las estimaciones emanadas de [1.B].

⁴ Considerando erróneamente que el intercepto captura tales efectos.



Por lo anterior, se hace menester ocupar una regresión más específica pero que contenga los mismos elementos básicos:

$$\mathbf{y}_{[NT \times 1]} = \alpha \boldsymbol{\zeta}_{[NT \times 1]} + \mathbf{X}_{[NT \times K]} \boldsymbol{\beta}_{[K \times 1]} + \mathbf{Z}_\mu \boldsymbol{\mu}_{[N \times 1]} + \mathbf{v}_{[NT \times 1]} \quad [1.C]$$

Dónde:

$\mathbf{Z}_\mu = \mathbf{I}_N \otimes \boldsymbol{\zeta}_{[T \times 1]}$,⁵ con $\boldsymbol{\zeta}_{[T \times 1]}$ definido como un vector de unos y \mathbf{I}_N una matriz identidad. $\mathbf{u}_{[NT \times 1]} = \mathbf{Z}_\mu \boldsymbol{\mu}_{[N \times 1]} + \mathbf{v}_{[NT \times 1]}$ cuyos elementos son $u_{it} = \mu_i + v_{it}$ un término de error compuesto, que se define como la suma de un componente de EI de corte transversal (que no cambian con el tiempo) pero estimables μ_i , y un término de error estocástico $v_{it} \sim N(0, \sigma_v^2)$.

Para obtener los parámetros estimados de [1.C] existen dos opciones que dependen de la suposición sobre el comportamiento de los EI, es decir, los estimadores serán distintos si estos efectos se consideran como fijos (parámetros), o bien, si se les supone como una variable estocástica $\mu_i \sim N(0, \sigma_\mu^2)$. La primera estimación es conocida como el Modelo de Efectos Fijos (MEF) y la segunda como Modelo de Efectos Aleatorios (MEA), aunque en realidad en ambos casos se trate de una misma ecuación de estimación.

Siguiendo a Baltagi (2008), tanto el MEF como el MEA tienen un tratamiento econométrico similar para la estimación de sus respectivos parámetros. Ya que en ambos casos se hace uso de dos matrices idempotentes simétricas, mismas que posteriormente se utilizan como ponderadores de [1.C]. Estas son:

- $\mathbf{P} = \mathbf{I}_N \otimes \left(\frac{1}{T} \boldsymbol{\zeta}_{[T \times T]}\right)$ donde $\boldsymbol{\zeta}_{[T \times T]}$ es una matriz de unos
- $\mathbf{Q} = \mathbf{I}_{NT} - \mathbf{P}$

Para el caso del MEF únicamente es necesario premultiplicar [1.C] por la matriz \mathbf{Q} . De tal forma que ahora se tendrá:

$$\tilde{\mathbf{y}} = \tilde{\mathbf{X}} \boldsymbol{\beta} + \mathbf{Q} \mathbf{v} \quad [1.D]$$

Dado que $\mathbf{Q} \mathbf{Z}_\mu = \mathbf{Q} \boldsymbol{\zeta}_{[NT \times 1]} = \mathbf{0}$ y definiendo $\tilde{\mathbf{y}} = \mathbf{Q} \mathbf{y}$, $\tilde{\mathbf{X}} = \mathbf{Q} \mathbf{X}$. Es decir, se han eliminado los EI, pero nótese que en realidad los parámetros a estimar no han sufrido ningún cambio, en efecto, el vector $\boldsymbol{\beta}$ de [1.C] es igual al de [1.D], es por ese motivo que puede utilizarse dicho tratamiento. Bastará con aplicar Mínimos Cuadrados Ordinarios (MCO) a [1.D] para obtener el estimador del MEF.

Para la estimación del MEA se parte de la matriz de varianzas covarianzas de los errores compuestos, es decir, de $\mathbf{u}_{[NT \times 1]} = \mathbf{Z}_\mu \boldsymbol{\mu}_{[N \times 1]} + \mathbf{v}_{[NT \times 1]}$ que bajo la consideración de que $\mu_i \sim N(0, \sigma_\mu^2)$, entonces estará determinada por:

⁵ \otimes denota un producto Kronecker.

$$\begin{aligned} \boldsymbol{\Omega} &= E(\mathbf{u} \mathbf{u}') = E(\mathbf{Z}_\mu \boldsymbol{\mu} \boldsymbol{\mu}' \mathbf{Z}_\mu' + \mathbf{Z}_\mu \boldsymbol{\mu} \mathbf{v}' + \mathbf{v} \boldsymbol{\mu}' \mathbf{Z}_\mu' + \mathbf{v} \mathbf{v}') = \mathbf{Z}_\mu E(\boldsymbol{\mu} \boldsymbol{\mu}') \mathbf{Z}_\mu' + E(\mathbf{v} \mathbf{v}') \\ &= \sigma_\mu^2 (\mathbf{I}_N \otimes \boldsymbol{\zeta}_{[T \times T]}) + \sigma_v^2 (\mathbf{I}_N \otimes \mathbf{I}_T) = (T \sigma_\mu^2 + \sigma_v^2) \mathbf{P} + \sigma_v^2 \mathbf{Q} \end{aligned}$$

Esta matriz se vuelve importante, pues es necesaria para obtener el estimador de Mínimos Cuadrados Generalizados (MCG), sin embargo, en realidad la técnica que utilizan los programas estadísticos no es la de MCG, sino es una técnica de Mínimos Cuadrados Ponderados, la cual consiste en recurrir a la matriz propuesta por Fuller y Battese (1974) $\boldsymbol{\theta} = \sigma_v \boldsymbol{\Omega}^{-1/2} = \frac{\sigma_v}{(T \sigma_\mu^2 + \sigma_v^2)^{1/2}} \mathbf{P} + \mathbf{Q}$.

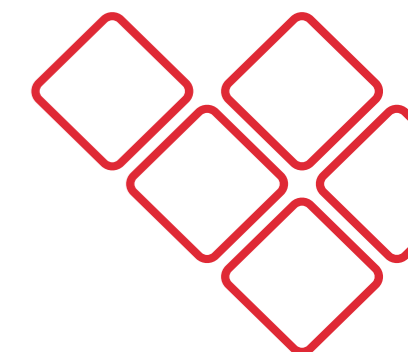
Premultiplicando [1.C] por $\boldsymbol{\theta}$, entonces:

$$\boldsymbol{\theta} \mathbf{y} = \alpha \boldsymbol{\theta} \boldsymbol{\zeta} + \boldsymbol{\theta} \mathbf{X} \boldsymbol{\beta} + \boldsymbol{\theta} \mathbf{Z}_\mu \boldsymbol{\mu} + \boldsymbol{\theta} \mathbf{v} \quad [1.E]$$

Para encontrar los estimadores correspondientes al MEA únicamente deben de aplicarse MCO a [1.E].

Matemáticamente los modelos de efectos fijos y aleatorios pueden ser muy atractivos, pero la realidad es que la principal deficiencia se encuentra en su propia especificación, ya que no se conoce absolutamente nada acerca de los EI, incluso si estos quieren ser estimados por variables indicadoras mediante un modelo Least Square Dummy Variables (LSDV), entonces matemáticamente debe cumplirse que $\sum_{i=1}^n \mu_i = 0$, pues estos son obtenidos a partir del intercepto estimado y de las medias en el tiempo de cada observación de corte transversal, para ello revítese a Baltagi (2008;13). Generalmente lo que los programas estadísticos reportarán de manera predeterminada serán los parámetros estimados de los intercepto de forma compuesta $\hat{\alpha} + \mu_i$.

Sin embargo, el principal problema para el análisis de convergencia es que estos modelos no son adecuados para el modelaje dinámico, esto es así formalmente porque en primera instancia, un modelo de datos agrupados como [1.B] tendría el problema de que los elementos y_{it} estarían correlacionados con todos los elementos de \mathbf{u}_i siempre que exista un $\mu_i \in \mathbf{u}_i$ (no importando que no exista correlación serial en \mathbf{v}). Ahora, un MEF tiene la particularidad de la transformación expresada en la ecuación [1.D] que elimina los componentes μ_i de \mathbf{u} , pero también genera que los elementos de los vectores se conviertan en desviaciones de las observaciones respecto de la media en el tiempo, esto también aplica para \mathbf{v} (ya que también sufre la pre-multiplicación por \mathbf{Q}) cuyos elementos serían definidos como $(v_{it} - \bar{v}_i)$, si se define a \mathbf{y}_{-1} como el vector de la variable dependiente con un rezago (cuyos elementos serían y_{it-1}) que se incluiría como regresor en el modelo, entonces existiría correlación entre los elementos de $\mathbf{Q} \mathbf{v}_i$ y $\mathbf{Q} \mathbf{y}_{-1i}$ (en específico con y_{it-1}), pues v_{it-1} está considerado en $\bar{v}_i = \frac{\sum_{t=1}^T v_{it}}{T}$ (es decir, v_{it-1} forma parte del promedio) y la $Cov(y_{it-1} v_{it-1}) \neq 0$. Por último, la transformación para el estimador del MEA tampoco soluciona el problema, ya que si se considera a [1.E] incluyendo $\boldsymbol{\theta} \mathbf{y}_{-1}$, entonces sus elementos estarían correlacionados con $\boldsymbol{\theta} \mathbf{v}$, esto por su componente que incluye a v_{it-1} .



Para eliminar dichos problemas, y utilizar la formalización que corresponde para analizar la convergencia en el ingreso per cápita, y que a su vez pasa a ser el modelo base a considerar si se requiere hacer alguna comparación con otros estimadores, se procede a utilizar el modelaje propuesto por Arellano y Bond (1991) que en la literatura también se conoce como el Método Generalizado de Momentos (por sus siglas en inglés GMM), que también se denominar como el modelo base de paneles dinámicos. Cabría mencionar que a pesar de su importancia, incluso la bibliografía especializada no le da un tratamiento adecuado, ejemplos de lo anterior son Wooldridge (2002) que solo menciona la existencia de dicho método sin entrar en detalles y Panchanan (2019) que a pesar de que su objetivo es entender lo que hace un programa estadístico, su explicación solo se enfoca en el caso de que no existan variables explicativas, es decir, en el caso de la hipótesis de convergencia sea especificada como [2.A], y también no explica la forma particular que adquiere la matriz de varianzas covarianzas de los errores.

Por lo anterior, y para entender como Arellano y Bond (1991) sortean las dificultades anteriormente externadas, se es menester desagregar (identificar la agrupación vectorial) y con ello plantear las estimaciones matriciales, ya que al no hacerlo de esta forma se genera una debilidad explicativa como en Baltagi (2008), en este respecto a continuación se presenta el modelaje correspondiente.

Arellano y Bond (1991) parten su explicación en dos, la primera sin incluir variables explicativas (además de \mathbf{y}_{-1}) y la segunda incluyéndolas, sin embargo, a continuación se presenta el modelo de manera conjunta:

$$\mathbf{y} = \alpha \zeta_{[NT \times 1]} + \delta \mathbf{y}_{-1} + \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}_\mu \boldsymbol{\mu} + \mathbf{v} = \alpha \zeta_{[NT \times 1]} + \mathbf{X}^* \boldsymbol{\beta}^* + \mathbf{Z}_\mu \boldsymbol{\mu} + \mathbf{v} \quad [3.F]$$

Dónde:

$$\mathbf{X}^* = [\mathbf{y}_{-1}, \mathbf{X}], \boldsymbol{\beta}^* = \begin{bmatrix} \delta \\ \boldsymbol{\beta} \end{bmatrix}, v_{it} \sim IID(0, \sigma_v^2).$$

El primer paso para estimar es aplicar primeras diferencias a [3.F] con el objetivo de eliminar los $\boldsymbol{\mu}$, aunque también se eliminará el intercepto, por lo que se obtiene.

$$\Delta \mathbf{y} = \Delta \mathbf{X}^* \boldsymbol{\beta}^* + \Delta \mathbf{v} \quad [3.F.1]$$

Δ indica primeras diferencias.

Agrupando vectorialmente por $i = 1, 2, \dots, N$. Entonces es posible observar a detalle $\Delta \mathbf{X}^*$

$$\begin{bmatrix} \Delta \mathbf{y}_1 \\ \Delta \mathbf{y}_2 \\ \vdots \\ \Delta \mathbf{y}_N \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{y}_{-11} & \Delta \mathbf{x}_{11} & \Delta \mathbf{x}_{21} & \cdots & \Delta \mathbf{x}_{K1} \\ \Delta \mathbf{y}_{-12} & \Delta \mathbf{x}_{12} & \Delta \mathbf{x}_{22} & \cdots & \Delta \mathbf{x}_{K2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \Delta \mathbf{y}_{-1N} & \Delta \mathbf{x}_{1N} & \Delta \mathbf{x}_{2N} & \cdots & \Delta \mathbf{x}_{KN} \end{bmatrix} \boldsymbol{\beta}^* + \begin{bmatrix} \Delta \mathbf{v}_1 \\ \Delta \mathbf{v}_2 \\ \vdots \\ \Delta \mathbf{v}_N \end{bmatrix} = \begin{bmatrix} \Delta \mathbf{X}_1^* \\ \Delta \mathbf{X}_2^* \\ \vdots \\ \Delta \mathbf{X}_N^* \end{bmatrix} \boldsymbol{\beta}^* + \begin{bmatrix} \Delta \mathbf{v}_1 \\ \Delta \mathbf{v}_2 \\ \vdots \\ \Delta \mathbf{v}_N \end{bmatrix}$$

Nótese que la estructura de $\Delta \mathbf{v}$ implica correlación serial de primer orden, es por eso que posteriormente Arellano y Bover (1995) plantearon un modelo con una transformación alternativa llamada Forward Orthogonal Deviation (FOD), que también eliminaba los efectos inobservados pero sin causar correlación serial.

Lo siguiente es especificar la estructura de la matriz de varianzas covarianzas de los errores, para ello se utiliza:

$$E(\Delta \mathbf{v}_i \Delta \mathbf{v}_i') = \sigma_v^2 \begin{bmatrix} 2 & -1 & 0 & \dots & \dots & 0 \\ -1 & 2 & -1 & \dots & \dots & 0 \\ 0 & -1 & 2 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & -1 & 0 \\ 0 & 0 & 0 & \dots & 2 & -1 \\ 0 & 0 & 0 & \dots & -1 & 2 \end{bmatrix} = \sigma_v^2 \boldsymbol{\Omega}_i_{[(T-2) \times (T-2)]}$$

Con $\boldsymbol{\Omega}_i$ ⁶ se busca aplicar el método de MCG pero tomando en cuenta una condición de ortogonalidad con una matriz de instrumentos, matriz que se utilizará como ponderadora a la regresión en primeras diferencias, a este procedimiento es al que se le llama método generalizado de momentos.

Bajo la inexistencia de regresores (además de \mathbf{y}_{-1}), Arellano y Bond (1991; 278) proponen usar como instrumentos a los rezagos de la variable dependiente, lo anterior bajo la lógica de que estos no estarán correlacionados con el término de error $\Delta \mathbf{v}_i$. Por lo que la matriz de instrumentos conjunta (para todo el sistema) será $\mathbf{z} = \begin{bmatrix} z_i \\ \vdots \\ z_n \end{bmatrix}$, cuyos elementos evidentemente son matrices de instrumentos para cada i .

$$\mathbf{Z}_i = \text{diag}([y_{i1}], [y_{i1}, y_{i2}], [y_{i1}, y_{i2}, y_{i3}], \dots, [y_{i1}, \dots, y_{iT-2}])$$

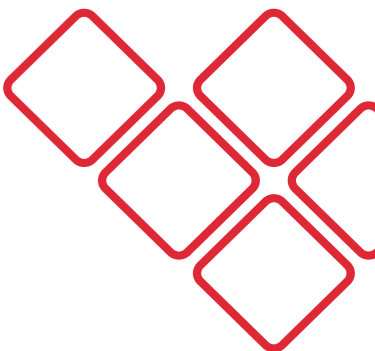
Nótese que se cumple ortogonalidad con el vector del termino de error en primeras diferencias, es decir que $E(\mathbf{Z}_i' \Delta \mathbf{v}_i) = 0$, ha esto se le llama condición de momentos. Ahora, cuando se utilizan otras variables en la parte sistemática del modelo, entonces estas pueden ser incluidas como instrumentos en la matriz \mathbf{Z}_i , simplemente se agregan columnas a ésta matriz con elementos de \mathbf{X} para cada i .

Una vez especificadas estas matrices el procedimiento es el siguiente: se premultiplica [3.F.1] por la matriz \mathbf{Z}' , es decir, se tendrá ahora $\mathbf{Z}' \Delta \mathbf{y} = \mathbf{Z}' \Delta \mathbf{X}^* \boldsymbol{\beta}^* + \mathbf{Z}' \Delta \mathbf{v}$. Posteriormente se aplica MCG utilizando la estructura de $\boldsymbol{\Omega}_i$ para con ello obtener el estimador GMM.

$$\hat{\boldsymbol{\beta}}^* = [(\mathbf{Z}' \Delta \mathbf{X}^*)' (\mathbf{Z}' \boldsymbol{\Omega}_i \mathbf{Z})^{-1} (\mathbf{Z}' \Delta \mathbf{X}^*)]^{-1} [(\mathbf{Z}' \Delta \mathbf{X}^*)' (\mathbf{Z}' \boldsymbol{\Omega}_i \mathbf{Z})^{-1} (\mathbf{Z}' \Delta \mathbf{y})]$$

Este modelo al ser la base explicativa dinámica, entonces los estimadores emanados de este modelaje pueden compararse con otros paneles dinámicos como los

⁶ Los programas estadísticos pueden utilizar variantes de ésta matriz, por ejemplo STATA pondera la misma por 0.5.



SYS-GMM o de Arellano y Bover (1995), considerando que con ello se ha podido eliminar el problema de endogeneidad (por la condición de momentos) que aparecía en los modelos de efectos fijos y aleatorios al agregar una variable dependiente como regresora, y con la ventaja de aun considerar la existencia de los EI. Así pues, esta es la base que debe de considerarse para la estimación de [2.A] y [2.B].

3. OTROS ANÁLISIS ESTOCÁSTICOS NO ORTODOXOS SOBRE CONVERGENCIA

Cabría resaltar que en términos de análisis econométricos más heterodoxos y de técnicas no paramétricas existen otras alternativas para el análisis de convergencia, cómo aquellas que usan Cadenas de Markov (CM), un ejemplo que se basa en varios trabajos del reconocido econometrista Quah y que utiliza los datos de México se encuentra en Rodríguez (2007), *grosso modo*, esta metodología consiste en agrupar a las unidades de corte transversal según el comportamiento promedio de la tasa de crecimiento del ingreso per cápita y de su nivel inicial, para posteriormente presentar matrices de transición probabilísticas $A_{[i,x,j]}$ cuyos elementos a_{ij} muestran la probabilidad de pasar del comportamiento inicial "i" al "j", y con ello después se puede presentar un análisis para identificar los factores que contribuyen a que las regiones pasen de un comportamiento a otro. Modelaje que es complementado con técnicas multinomiales de estimación de datos panel.

Otra forma de análisis de convergencia es a través de productividad, un referente para este tipo de análisis también usando los datos de México se encuentra en Blomström y Wolff (1977), este trabajo se centra en la premisa de que es posible que exista la derrama de las técnicas de producción que emplean las empresas multinacionales (asumidas más avanzadas) en las empresas nacionales, lo anterior teóricamente se explica por tres canales: la inversión extranjera directa, la competencia de mercado que fuerza la eficiencia, el entrenamiento de mano de obra y el nivel de dirección gerencial, que se asumen se diseminan principalmente por la rotación de per-

sonal entre las firmas nacionales y multinacionales y por medio de la producción subsidiaria. El análisis propone dos regresiones con dos variables dependientes: La tasa de crecimiento de la productividad de las empresas locales dentro de la industria y la tasa de convergencia en los niveles de productividad laboral entre empresas locales y extranjeras dentro de un sector, cuyos regresores más relevantes son el grado de propiedad extranjera de la industria y la brecha en la productividad laboral entre empresas locales y extranjeras en el periodo inicial.

Cabe resaltar que estos análisis heterodoxos pueden ser de un más alto poder estadístico inferencial que los correspondientes a los MEA y MEF, pues en términos generales dinámicos no tienen intrínsecamente sesgo asintótico.

4. CONCLUSIONES

En primera instancia, se hizo una exposición de donde emana por primera vez el concepto de convergencia económica, concatenándolo con su primera medición econométrica de la forma de datos de corte transversal según lo propuesto por Mankiw, Romer y Weil (1992), haciendo énfasis en dos casos posibles, el de convergencia absoluta y la condicionada. Posteriormente, centrándose el análisis en el caso de datos panel, se hizo un recuento de que en un modelo teórico como el de Solow, existen variables que no son consideradas y pueden ocultar relaciones intrínsecas al sistema, en ese sentido, empíricamente pueden ser introducidas y superar varios reduccionismos. En el ejemplo del comercio internacional, esto puede ser teóricamente posible gracias a los modelajes gravitacionales, que tienen la particularidad de que son análisis lineales y pueden ser introducidos fácilmente en la ecuación de regresión.

En la investigación teórica econométrica, se explicó porque los modelajes que no incluyen los EI tienen subespecificación, y por tanto, presentan problemas de sesgo. Así mismo, de manera formal se ha explicado la endogeneidad que se provoca por la inclusión de una variable dependiente en la parte sistemática del modelo al hacer estimacio-

nes de los MEF y MEA, inclusión que se hace necesario porque el análisis de convergencia implica investigar la relación de las tasas de crecimiento del ingreso contemporáneas con sus rezagadas. En ese sentido y desde la óptica empírica más ortodoxa, el modelo base que debe emplearse en este análisis es el de Arellano y Bond (1991), ya que este permite sortear esa endogeneidad mediante la condición de momentos, pero que por construcción prosigue con un problema de sesgo muestral (no asintótico); por la correlación serial de primer orden provocada por la corrección de primeras diferencias que eliminan en un primer paso los EI, por lo que para este modelo base será un buen complemento el análisis de un modelo tipo Arellano y Bover (1995), mismo que considera una corrección FOD.

Una alternativa para sortear varios de los obstáculos que se presentan en estos es recurrir a métodos más heterodoxos, aunque en muchas ocasiones esto se puede hacer solo si se sacrifican especificidades (como las estructurales). Pero aun así, estos métodos pueden arrojar respuestas a otras interrogantes, como las probabilidades de transición de un estado a otro como permiten identificarse utilizando las CM, o un análisis más microeconómico; como es el caso del trabajo pionero de Blomström y Wolff (1977), que aunque no tan moderno, lo cierto es que estos estudios se han ido relegando para ser reemplazados para un análisis desde el ámbito puramente macroeconómico.

BIBLIOGRAFÍA

Aquino, J. (2012): "Economía internacional". Red tercer milenio. Edo. de México, primera edición.
 Arellano M. y Bond S. (1991): "Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations". The Review of Economic Studies, Vol. 58, No. 2 (Apr., 1991), pp. 277-297.
 Baltagi, B. (2008). "Econometric analysis of panel data". John Wiley & Sons. New York, USA.
 Blomström, M., y Wolff, E. N. (1977). "Growth in a dual economy". World Development, 25(10), 1627-1637.

Cermeño, R. (2001). "Decrecimiento y convergencia de los estados mexicanos: Un análisis de panel". El Trimestre Económico, 603-629.
 Díaz, V. (2014): "¿Se cumple la teoría neoclásica del comercio internacional?: el caso de la economía colombiana entre 1980 y 2007". Universidad del Valle, Cali-Colombia.
 Fuller, W., y Battese, G. (1974): "Estimation of linear models with crossed-error structure". Journal of Econometrics, 2(1), 67-78.
 Galor, O. (1996): "Convergence?: Inferences from Theoretical Models". Economic Journal, Vol. 106, núm. 437.
 Jiménez, J. y Narbona, A. (2007): "El idioma común como determinante del comercio internacional: el caso del español". X Encuentro de Economía Aplicada, España.
 Krugman, P. (2012): "Economía internacional". Pearson education. Madrid-España, novena edición.
 Martín, S. (2000): "Apuntes de crecimiento económico". Antoni Bosch. Barcelona, primera edición.
 Mankiw, G. (2000): "Macroeconomía". Antoni Bosch. Barcelona, cuarta edición.
 Mankiw, G., Romer, D. y Weil D. (1992): "A contribution to the empirics of economic growth". The Quarterly Journal of Economics, Vol. 107, No. 2, pp.407-437.
 Panchanan, D. (2019): "Econometrics in Theory and Practice Analysis of Cross Section, Time Series and Panel Data with Stata 15.1". Springer Nature Singapore Pte Ltd.
 Romer, D. (2002): "Macroeconomía Avanzada". McGraw Hill. Madrid-España, segunda edición.
 Rodríguez, E. (2007). "Winners and Losers of Regional Growth in Mexico and their Dynamics". Investigaciones Económicas, vol. LXVI, 259, enero-marzo, 2007, pp.43-62.
 Solow, R. (1956): "A contribution to the theory of economic growth". The Quarterly Journal of Economics, Vol. 70, No. 1.
 Sulca y Toribio, (2014): "Estudio del modelo de Ramsey usando control óptimo". CONCYTEC, Peru.
 Tinbergen, J. (1962): "Shaping the World Economy; Suggestions for an International Economic Policy". Books (Jan Tinbergen). Twentieth Century Fund, New York.
 Wooldridge J. (2002): "Econometric analysis of cross section and panel data". The MIT Press Cambridge, Massachusetts London, England.