

¿Ciencia económica o el arte de hacer economía? Metodología científica y replicaciones en economía

Verónica De Jesús Romo^a

Resumen

El objetivo de este artículo es argumentar que para que la economía se fundamente plenamente como una ciencia, y no simplemente como un arte basado en la intuición, se deben seguir ciertos principios científicos. El artículo contribuye al debate sobre la materia, específicamente en dos aspectos metodológicos: 1) el adecuado análisis de datos para la verificación de las hipótesis de una investigación y 2) la replicación de los hallazgos científicos con la finalidad de evaluar a los mismos. El primer aspecto, con base en el *enfoque probabilístico de la econometría* de Haavelmo (1944), implica que los economistas aseguremos que los supuestos probabilísticos de los modelos estadísticos que utilizamos en nuestras investigaciones se cumplan. El segundo aspecto plantea que las *replicaciones* son un medio para establecer la fragilidad (derivada por errores del manejo de bases de datos o por la inadecuada especificación de los modelos estadísticos), o la solidez (derivada de la robustez de los resultados ante cambios en el periodo de análisis, en la economía analizada, en la especificación de las variables, en el método de estimación, etc.) de los resultados reportados en una investigación publicada. El artículo contribuye con una ilustración de nuestra propuesta metodológica: se trata de la replicación del artículo de Kakkar (2001), “Long run real exchange rates: evidence from Mexico”, publicado en la revista *Economics Letters*.

Palabras clave: replicación, metodología de la ciencia, *Economic Letters*, Kakkar (2001).

Clasificación JEL: B4, C5.

Abstract

The aim of this article is to argue that for the economics to be fully settled as a science, and not simply as an art based on intuition, it has to follow certain scientific principles. The article contributes to the debate on the matter, specifically on two methodological aspects: 1) the proper analysis of data to confirm the hypothesis of a research, and 2) the replication of scientific findings in order to evaluate them. The first aspect, based on Haavelmo's monograph of 1944 entitled “The probabilistic approach to econometrics”,

Manuscrito recibido el 10 de octubre de 2015; aceptado el 12 de abril de 2016.

a Posgrado en Economía de la Facultad de Economía de la Universidad Nacional Autónoma de México (UNAM, México). Correspondencia: veronica.dejesus@ciencias.unam.mx. La autora agradece los comentarios y sugerencias del Dr. Julio López Gallardo.

implies that economists should make sure that the probabilistic assumptions of the statistical models used in their research are met. The second aspect suggests that the replications are a means to show the fragility (arose by errors of mishandling databases or inadequate specification of statistical models), or the strength (arose from the robustness of the results from changes in time period, the economy under analysis, the specification of the variables, the estimation methods, etc.) of the results reported in a published research. The article contributes with an illustration of our own methodological proposal: it is a replication of an article by Kakkar (2001), entitled “Long run actual exchange rates: Evidence from Mexico”, published in the journal *Economics Letters*.

Key words: replication, methodology of science, *Economic Letters*, Kakkar (2001).

INTRODUCCIÓN

El estatuto de la economía como *una disciplina científica*, y no simplemente como un *arte basado en la intuición*, es algo que se viene discutiendo prácticamente desde su nacimiento. El objetivo de este artículo es contribuir al debate sobre la materia, específicamente en dos aspectos metodológicos: 1) el adecuado análisis de datos para la verificación de las hipótesis de una investigación y 2) la replicación de los hallazgos científicos con la finalidad de evaluar a los mismos. La discusión abstracta podría extenderse hasta el infinito, y por eso consideramos que es más útil seguir un camino diferente. Así, primeramente, se hará referencia en esta introducción y en la segunda sección a las cuestiones metodológicas generales relativas al análisis de datos en economía que se desprenden del artículo pionero “The probability approach in econometrics” (“El enfoque probabilístico de la econometría”) de Haavelmo (1944). Inmediatamente después se presentará un reporte sobre las replications a investigaciones empíricas en economía, que describe la escasa cultura entre los economistas de replicar los resultados de investigaciones publicadas y, posteriormente, publicar sus hallazgos. Finalmente, mediante una replicación del artículo de Kakkar (2001), titulado “Long run real exchange rates: Evidence from Mexico”, publicado en la revista *Economics Letters*, se ilustrará cómo ambos aspectos deben ser debidamente garantizados para evitar que las conclusiones de una investigación sean inválidas, o poco confiables.

Por lo común sólo aquellas disciplinas que tienen que ver con las ciencias naturales se reconocen como propiamente científicas, fundamentalmente porque en ellas es posible someter las hipótesis a experimentación y, mediante el uso de métodos estadísticos, validarlas o rechazarlas. Por supuesto, las investigaciones en economía aplicada son distintas, porque no se pueden medir características de

un gran número de elementos de una muestra bajo un mismo experimento. No se trata de que las estadísticas públicas del producto interno bruto (PIB) en un año únicamente se reunieron una vez, sino que el PIB de ese año en cuestión existió sólo una vez; la historia exclusivamente nos proporciona ese “experimento”.¹

Sin embargo, el carácter no experimental de los datos económicos, que proceden de “la corriente de experimentos que la naturaleza está constantemente realizando en su enorme laboratorio, de la que meramente somos observadores pasivos” (Haavelmo, 1944, p. 14), no debería ser un obstáculo para que en economía se apliquen metodologías estadísticas basadas en la teoría de la probabilidad para estimar y probar relaciones a cerca de fenómenos reales con todo rigor científico.

La econometría, o al menos la corriente de la econometría con la que me identifico, adopta el fundamento metodológico del enfoque probabilístico de Haavelmo (1944). En este enfoque se supone que todas y cada una de las variables que conforman un cierto fenómeno son variables aleatorias y que, por tanto, se caracterizan por poseer una cierta función de densidad que, *grosso modo*, indica la probabilidad de ocurrencia de los distintos valores que podrían tomar cada una de esas variables. A la vez, se considera el conjunto de datos disponibles respecto de cada una de esas variables (\mathbf{x}) como una muestra, fruto de la realización de un *proceso generador de datos* (PGD).² Este proceso es desconocido, y de hecho quizá no lleguemos nunca a conocerlo con precisión. Pero al caracterizarlo mediante una determinada *distribución de probabilidad*³ resulta posible realizar pruebas de hipótesis o cualquier otro tipo de inferencia estadística. En palabras de Haavelmo (1944, p. iii):

It is sufficient to assume that the *whole set* of, say n , observations may be considered as one observation of n variables (or a “sample point”) following an n -dimensional *joint* probability law, the “existence” of which may be purely hypothetical. Then, one can test hypotheses regarding this joint probability law, and draw inference as to its possible form, by means of *one* sample point (in n dimensions).

1 Aunque el desarrollo de la economía experimental está abordado de esta manera los fenómenos económicos por medio de la simulación de eventos.

2 En el artículo se usan letras mayúsculas para denotar variables aleatorias, los valores observados de éstas se denotan con las letras minúsculas correspondientes. Las letras en formato negrita denotan arreglos de números o variables; las mayúsculas denotan matrices y las minúsculas vectores.

3 Dicha distribución fue llamada ley de probabilidad por Haavelmo.

Ahora bien, consideramos que en economía al igual que en las ciencias naturales (y también la medicina, por ejemplo) es importante, antes que nada, modelar de manera formal, y general, nuestra hipótesis, por ejemplo, cuáles son los determinantes del tipo de cambio real en un periodo de tiempo dado. Pero si estudiamos un fenómeno real, para ello no basta una formulación general, o teórica. Es preciso que el modelo en cuestión tome en cuenta debidamente la forma en que el fenómeno se manifiesta en la realidad concreta; esto es, que recoja adecuadamente los datos que proporciona la realidad.

Pero, ¿qué quiere decir en este contexto recoger adecuadamente los datos? En la econometría inspirada en el enfoque de Haavelmo, esto quiere decir ni más ni menos que tomamos esos datos como si ellos provinieran de una determinada función de probabilidad, y que verificamos que, al estimar parámetros o contrastar hipótesis (teorías económicas), cumplimos los supuestos probabilísticos del modelo estadístico. Esta es una cuestión fundamental puesto que los métodos estadísticos no tienen sentido sin la teoría de la probabilidad que los fundamenta. “For *no tool developed in the theory of statistics has any meaning—except, perhaps, for descriptive purposes—without being referred to some stochastic scheme*” (Haavelmo 1944, p. iii).

En la siguiente sección se exponen algunos aspectos del enfoque probabilístico de la econometría de Haavelmo, que representan un conjunto de principios metodológicos para la economía aplicada.

EL ENFOQUE PROBABILÍSTICO DE LA ECONOMETRÍA

El objetivo de esta sección es destacar los fundamentos metodológicos que los planteamientos del economista noruego Trygve Haavelmo dieron a la econometría, al adaptar las nuevas perspectivas probabilísticas y estadísticas de la época a los datos económicos, y que hoy en día permiten el uso de métodos estadísticos para evaluar teorías económicas, o hipótesis económicas, contra los datos, con todo rigor científico. Dichas ideas se publicaron en 1944, en la revista *Econométrica*, bajo el título “The probability approach to econometrics”.⁴

El enfoque probabilístico que Haavelmo dio a la econometría fue influenciado por los axiomas de la teoría de la probabilidad (Kolmogorov, 1933), por

4 Trabajo por el cual Haavelmo recibió el grado de doctor en 1946, y junto a su análisis de estructuras económicas simultáneas lo hicieron acreedor al Premio Nobel de Economía en 1989.

los planteamientos de inferencia estadística basada en la teoría de muestras finitas, la teoría de la estimación y de pruebas de hipótesis de Fisher (1922), así como por los métodos estadísticos basados en probabilidad de Jerzy Neyman y Abraham Wald.⁵

Previo a los planteamientos de Haavelmo, los economistas de la época se habían mostrado renuentes a adoptar modelos probabilísticos para modelar fenómenos económicos por la falta de independencia y homogeneidad de los datos. Se creía que, basados en el concepto de muestra aleatoria,⁶ los modelos de probabilidad estaban reservados a experimentos cuyas realizaciones, con objetos (o individuos) de una misma “población” y bajo las mismas condiciones, generen una serie de observaciones independientes con idéntica distribución, por ejemplo en juegos de azar (véase Spanos, 2015).

Así pues, había mucho escepticismo en la mayoría de los economistas de la época, ya que se creía que el análisis econométrico basado en la teoría de la probabilidad no era aplicable en la mayoría de las situaciones. No obstante, Haavelmo insistió que la teoría de probabilidad era lo suficientemente general para hacer frente a las particularidades de los datos económicos.

What we want are theories that, without involve us in direct logical contradictions, state that the observations will *as a rule* cluster in a limited subset of the set of all conceivable observations, while it is still consistent with the theory that an observation falls outside this subset ‘now and then’.

As far as is known, the scheme of probability and random variables is, at least for the time being, the only scheme suitable for formulating such theories (Haavelmo, 1944, p. 40).

El primer aspecto metodológico del enfoque probabilístico de la econometría para comparar la teoría contra los datos es puntualizar bajo qué condiciones se espera que la teoría se cumpla. Es decir, para que un modelo tenga significado económico deberá tener asociado un diseño de experimentos que describa e indique cómo medir un sistema de variables verdadera⁷ (X_1, X_2, \dots, X_n) asociadas a las

5 Haavelmo estudió la entonces nueva estadística frecuentista con Neyman en 1936, en Londres, y en 1939, en Berkeley, y con Wald estudió inferencia estadística con muestras heterogéneas y dependientes durante sus estudios en Estados Unidos entre 1939-1944 (véase Spanos, 2015).

6 La sucesión de n variables aleatorias independientes e idénticamente distribuidas, (X_1, X_2, \dots, X_n) , se define como muestra aleatoria de tamaño n .

7 Las variables “verdaderas” (o funciones de tiempo) representan nuestro ideal en cuanto a las medidas exactas de la realidad “como lo es en realidad” (Haavelmo, 1944, p. 5).

correspondientes variables teóricas.⁸ De esta forma, el modelo se convierte en una hipótesis del fenómeno real, donde todo sistema de valores que podríamos observar⁹ de las variables “verdaderas” será uno que pertenece al conjunto de sistemas de valores admisibles dentro del modelo. En pocas palabras, la filosofía del planteamiento es la siguiente:

(...) that Nature has a way of selecting joint value-systems of the ‘true’ variables such that these systems are as if the selection had been made by the rule defining our theoretical model. Hypotheses in the above sense are thus the joint implications—and the only testable implications, as far as observations are concerned—of a theory and a design of experiments (Haavelmo, 1944, p. 9).

Haavelmo distinguió entre dos tipos de experimentos, los que se pueden aislar artificialmente, y los de la naturaleza:

(1) experiments that we should like to make to see if certain real phenomena—when artificially isolated from “other influences”—would verify certain hypotheses, and (2) the stream of experiments that Nature is steadily turning out from her own enormous laboratory, and which we merely watch as passive observers (Haavelmo, 1944, p. 14).

En la mayoría de los casos, los fenómenos económicos no se pueden realizar un gran número de veces bajo las mismas condiciones. Los economistas, por tanto, deben modelar con resultados que la naturaleza ha realizado, y de los cuales son observadores pasivos. En este caso, Haavelmo propone utilizar la teoría para explicar los resultados que la naturaleza está produciendo:

In order to test a theory against facts, or to use it for predictions, either the statistical observations available have to be “corrected”, or the theory itself has to be adjusted, so as to make the facts we consider the “true” variables relevant to the theory (...). The economist (...) is presented with some results which, so to speak, Nature has produced in all their complexity, his task being to build models that explain what has been observed (p. 7) (...) try to adjust our theories to reality as it appears before us (Haavelmo, 1944, p. 14).

8 Las variables teóricas son las medidas verdaderas que debemos hacer si la realidad estuviera realmente de acuerdo con nuestro modelo teórico (Haavelmo, 1944, p. 5).

9 Los valores observados, cuando contradicen la teoría, dejan la posibilidad de que estuviésemos probando la teoría en hechos para los que la teoría no estaba destinada a cumplirse (Haavelmo, 1944, p. 7).

Y los datos disponibles van a ser tratados como si hubiesen sido generados por el diseño de experimento y la teoría elegida: “We try to choose a theory and a design of experiments to go with it, in such a way that the resulting data *would be* those which we get by passive observation of reality” (Haavelmo, 1944, p. 14).

El fundamento probabilístico de los modelos económicos consiste en considerar el conjunto de los n valores observados, (x_1, x_2, \dots, x_n) , de las n variables (X_1, X_2, \dots, X_n) involucradas en alguna relación de teoría económica de interés, $X_1 = f(X_2, X_3, \dots, X_n)$, como un punto muestral, extraído del universo n -dimensional, \mathbb{R}^n , obedeciendo alguna función de distribución de probabilidad conjunta n -dimensional desconocida, denotada por:

$$P_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = P_{\mathbf{X}}(\mathbf{x}), (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

Dicha función contiene toda la información probabilística del vector aleatorio, de manera que esta función adquiere central importancia puesto que ella constituye la base del modelo estadístico que captura toda la información contenida en las series de datos observados.¹⁰ De hecho, esta información expresada en la función de densidad conjunta de todas las variables bajo análisis:

$$D_{X_1, X_2, \dots, X_n}(x_1, x_2, \dots, x_n) = D_{\mathbf{X}}(\mathbf{x}), (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$$

constituye el *proceso generador de datos*. Por lo general, la función de densidad conjunta del proceso observable tiene una estructura probabilística complicada debido a su alta dimensionalidad y a la heterogeneidad de los datos económicos.

A partir de lo anterior, podemos dar un nuevo paso. En el marco del enfoque probabilístico se ha desarrollado *la teoría de la reducción probabilística* (Hendry y Richard, 1983; Spanos 1986). La anterior consiste en evaluar rigurosamente supuestos sobre el vector observable, para llegar a una estructura probabilística simplificada y aceptable del mismo. Los supuestos en cuestión se refieren a: 1) la distribución (normal, t de Student, logística, etc.), 2) la dependencia (propiedad de Markov (p), martingala, ergodicidad, etc.) y 3) la heterogeneidad (idéntica distribución, estacionariedad estricta, estacionariedad de orden k , etc), de dicho vector.

¹⁰ En el caso de las variables aleatorias con valores reales, la distribución de una variable aleatoria X queda determinada por su función de distribución $P_X(x)$.

Respecto al supuesto de distribución Haavelmo señaló:

Since the assignment of a certain probability law to a system of observable variables is a trick of our own, invented for analytical purposes, and since the same observable results may be produced under a great variety of different probability schemes, the question arises as to which probability law should be chosen, in any given case, to represent the ‘true’ mechanism under which the data considered are being produced. To make this a rational problem of statistical inference we have to start out by an axiom, postulating that every set of observable variables has associated with it one particular ‘true’, but unknown, probability law (Haavelmo, 1944, p. 49).

Haavelmo fue consciente de que las características de dependencia y heterogeneidad exhibidas por las variables económicas (principalmente series de tiempo de macroeconomía) eran fuente de correlaciones (regresiones) espurias, tal como lo evidenció Yule (1926).¹¹ Sin embargo, él estaba completamente convencido de que tales características de las variables económicas, sólo podían ser captadas mediante un modelo estadístico rigurosamente formulado en la teoría de la probabilidad, específicamente en la distribución conjunta del proceso estocástico observable mediante los datos (véase Haavelmo, 1943a y 1944).

Siguiendo el enfoque probabilístico, la adopción de tales supuestos se convierte en un problema de inferencia estadística que debe ser contrastado estadísticamente con los datos observados: “We make hypothetical statements before we draw the sample, and we are only concerned with whether the sample rejects or does not reject an a priori hypothesis” (Haavelmo, 1944, p. 70).

Así, el diseño de experimentos de Haavelmo vía la distribución conjunta del vector observable coincide con una primera aproximación del proceso generador de datos, PGD, cuando los datos son observaciones pasivas de la realidad.¹² No obstante, en la práctica este proceso es desconocido y debe ser determinado a partir de la información disponible, es decir, del conjunto de variables observadas para

11 El coeficiente de correlación muestral (contemporáneo) entre dos series diferentes $\{(x_t, y_t), t = 1, 2, \dots, T\}$ es una medida de dependencia de primer orden significativa sólo cuando las medias de ambos procesos observables son constantes, $E(X_t) = \mu_x$, $E(Y_t) = \mu_y$, para todo $t \in T$, de otra forma la medida no tendrá sentido.

12 Hendry acuñó el concepto de proceso generador de datos local (LDGP) que ofrece un “diseño de experimentos” apropiado, mediante el cual se describen los datos observados de forma pasiva dentro de los más pequeños errores posibles, dada la elección del conjunto de datos (véase Hendry, 2016).

el análisis de acuerdo a alguna teoría económica, de todos los términos determinísticos asociados, del conocimiento institucional y de la evidencia anterior. Por lo tanto, el éxito de los modelos empíricos para recoger toda la información del fenómeno de interés depende de la *especificación del modelo estadístico*, que consiste en reconocer ciertas regularidades en los datos y la adopción de supuestos probabilísticos apropiados para captar dichas regularidades.

Para hacer más concretas las observaciones anteriores, considérese el siguiente modelo estadístico:

$$M_0(\mathbf{x}) = \{D(\mathbf{x}|\boldsymbol{\theta}), \boldsymbol{\theta} \in \Theta\}, \mathbf{x} \in \mathbb{R}^n, \Theta \subset \mathbb{R}^m, m < n \quad [1]$$

donde Θ denota el espacio paramétrico; \mathbb{R}^n , el espacio muestral; $D(\mathbf{x}|\boldsymbol{\theta})$, la densidad conjunta. El modelo $M_0(\mathbf{x})$ es la parametrización del proceso $\{X_t, t \in \mathbb{N}\}$ conociendo los valores de la muestra \mathbf{x} , con la estructura de probabilidad elegida. Por ejemplo, para el modelo estadístico conocido como regresión lineal, $y_t = \alpha + \boldsymbol{\beta}'\mathbf{x}_t + u_t, t \in T$, su estructura probabilística se presenta en el cuadro 1 usando dos enfoques distintos.

Ahora bien, antes de contrastar las hipótesis acerca del fenómeno de interés con el modelo [1], es necesario *verificar* que el modelo estadístico satisfaga la lista completa de supuestos probabilísticos subyacentes a la especificación elegida con los datos de la muestra \mathbf{x} . Dicha verificación garantiza la confiabilidad de cualquier inferencia basada en el modelo. Nótese que cuando el modelo es incorrectamente especificado, en el sentido que alguno de los supuestos del modelo fue rechazado, entonces la distribución $P(\cdot; \boldsymbol{\theta})$ estará erróneamente especificada para la muestra \mathbf{x} e invalidará la distribución de los estimadores, de los estadísticos de prueba y de cualquier estadístico T_n obtenido a partir de $D(\mathbf{x}|\boldsymbol{\theta})$.

Cuadro 1
Supuestos del modelo de regresión lineal

	<i>Enfoque probabilístico</i>	<i>Enfoque tradicional</i>
1 Normalidad	$(y_t \mathbf{X}_t = \mathbf{x}_t) \sim N(\cdot, \cdot)$	$u_t \sim N(\cdot, \cdot)$
2 Linealidad	$E(y_t \mathbf{X}_t = \mathbf{x}_t) = \alpha + \boldsymbol{\beta}'\mathbf{x}_t$	$E(u_t \mathbf{X}_t = \mathbf{x}_t) = 0$
3 Homocedasticidad	$Var(y_t \mathbf{X}_t = \mathbf{x}_t) = \sigma^2$	$Var(u_t \mathbf{X}_t = \mathbf{x}_t) = \sigma^2$
4 t-invarianza de los parámetros	$\alpha, \boldsymbol{\beta}, \sigma^2$	$\alpha, \boldsymbol{\beta}, \sigma^2$
5 Independencia	$\{(y_t \mathbf{X}_t = \mathbf{x}_t), t \in T\}$	$\{(u_t, u_s \mathbf{X}_t = \mathbf{x}_t) = 0, t \neq s, t \in T\}$
$\alpha = E(y_t) - \boldsymbol{\beta}'E(\mathbf{X}_t), \boldsymbol{\beta} = Cov(\mathbf{X}_t)^{-1}Cov(\mathbf{X}_t, y_t)$ y $\sigma^2 = Var(y_t) - \boldsymbol{\beta}'Cov(\mathbf{X}_t, y_t)$		

Supóngase que mediante alguna prueba (test) estadística se encuentra que el modelo [1] está incorrectamente especificado. Pues bien, el modelo se debe de *reespecificar*, elegir una nueva especificación que considere las regularidades en los datos no explicadas por el modelo incorrectamente especificado, y posteriormente evaluar que los datos, \mathbf{x} , no rechazan los supuestos de esta nueva especificación. Este procedimiento se debe de repetir hasta determinar una especificación estadísticamente válida para los datos, a partir de la cual se puedan hacer inferencias confiables.

Debe reconocerse que, a partir del enfoque probabilístico de la econometría, los economistas cuentan con un método analítico formal basado en la teoría de la probabilidad que permite el uso de los métodos estadísticos con el mismo rigor científico que en las ciencias experimentales.

Purely empirical investigations have taught us that certain things in the real world happen only very rarely, they are “miracles” while others are “usual events”. The probability calculus has developed out of a desire to have a formal logical apparatus for dealing with such phenomena of real life. The question is not whether probabilities exist or not, but whether —if we proceed *as if* they existed— we are able to make statements about real phenomena that are “correct for practical purposes” (Haavelmo 1944, p. 43).

Dicho método consiste en: 1) definir el diseño de experimento que relacione el modelo de teoría económica a los datos en un esquema probabilístico mediante la especificación del modelo estadístico; 2) la verificación de los supuestos estadísticos subyacentes a la especificación, y 3) la reespecificación del modelo con el objeto de establecer un modelo correctamente especificado con el cual contrastar las hipótesis, a fin de establecer conclusiones estadísticamente confiables a la luz de los datos.

A continuación se desarrolla el enfoque probabilístico del modelo de regresión lineal múltiple. En la exposición se enfatiza la razón por la cual cada supuesto del cuadro 1 es necesario en la construcción del modelo y las implicaciones de no cumplirlos.

ENFOQUE PROBABILÍSTICO DEL MODELO DE REGRESIÓN LINEAL MÚLTIPLE

Supongamos que nos interesa modelar los determinantes del PIB en un periodo determinado. Sea la siguiente relación económica de interés:

$$Y_t = f(X_{1,t}, X_{2,t}, \dots, X_{m,t}), \quad t = 1, 2, \dots, n \quad [2]$$

Sea $\{\mathbf{Z}_t\}_{t=1}^n$ el proceso estocástico que considera todas las variables que suponemos están involucradas en la relación [2], donde $\mathbf{Z}_t := (Y_t, X_{1,t}, X_{2,t}, \dots, X_{m,t})$, Y_t denota la variable dependiente, y $X_{i,t}$, $i = 1, 2, \dots, m$ denota la i -ésima variable independiente al tiempo t . Enseguida se muestra como el modelo estadístico de regresión lineal múltiple:

$$M_\theta(\mathbf{z}) : y_t = \alpha + \boldsymbol{\beta}' \mathbf{x}_t + u_t, \quad u_t \sim N(0, \sigma^2)$$

para $\mathbf{z} \in \mathbb{R}^{n(m+1)}$ con vector de parámetros $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}, \sigma^2)$ y bajo los supuestos del cuadro 1, es una parametrización de la densidad conjunta de todas las variables observables del proceso $\{\mathbf{Z}_t\}_{t=1}^n$ bajo los siguientes supuestos: distribución normal, independencia e idéntica distribuido, es decir: $\mathbf{Z}_t \sim NIID(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, $t = 1, 2, \dots, n$.

Por el supuesto de independencia e idéntica distribución de \mathbf{Z}_t , la función de densidad conjunta del proceso es igual a $D_Z(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n; \Phi) = \prod_{t=1}^n D(\mathbf{Z}_t; \Psi)$ aplicando la definición de probabilidad conjunta se obtiene:

$$D_Z(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n; \Phi) = \prod_{t=1}^n D(Y | \mathbf{X}_t; \boldsymbol{\psi}_1) D(\mathbf{X}_t | \boldsymbol{\psi}_2)$$

Por el supuesto de distribución normal de \mathbf{Z}_t , y suponiendo exogeneidad de las variables $X_{i,t}$ para $i = 1, 2, \dots, m$, es posible ignorar la densidad marginal de $\mathbf{X}_t = (X_{1,t}, X_{2,t}, \dots, X_{m,t})$, denotada por $D(\mathbf{X}_t | \boldsymbol{\psi}_2)$, para hacer inferencia en el espacio paramétrico $\boldsymbol{\psi}_1$ (véanse Spanos, 1986 y Hendry y Nielsen, 2007, cap. 10), entonces:

$$D_Z(\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n; \Phi) \cong \prod_{t=1}^n D(Y | \mathbf{X}_t; \boldsymbol{\psi}_1)$$

Por el teorema 1 (en Apéndice 1) sabemos que Y_t , dados los valores del vector aleatorio $\mathbf{X}_t = \mathbf{x}_t$, se distribuye como una variable aleatoria normal:

$$Y_t | \mathbf{X}_t = \mathbf{x}_t \sim N(\alpha + \boldsymbol{\beta}' \mathbf{x}_t, \sigma^2) \quad [3]$$

donde $\alpha = E(y_t) - \boldsymbol{\beta}' E(\mathbf{X}_t)$, $\boldsymbol{\beta} = Cov(\mathbf{X}_t)^{-1} Cov(\mathbf{X}_t, y_t)$ y $\sigma^2 = Var(y_t) - \boldsymbol{\beta}' Cov(\mathbf{X}_t, y_t)$.

Este resultado muestra que existe una relación lineal entre Y_t y \mathbf{X}_t del siguiente tipo:

$$Y_t = E(y_t | \mathbf{X}_t = \mathbf{x}_t) + u_t, \quad t = 1, 2, \dots, n$$

donde el término de error $u_t = y_t - E(y_t | \mathbf{X}_t = \mathbf{x}_t)$ no es autónomo, su estructura probabilística está completamente determinada por [3]. De hecho, los supuestos del modelo estadístico se pueden expresar en términos de u_t , como en el cuadro 1.

Como es habitual, para determinar los valores más probables de los parámetros del modelo estadístico $\boldsymbol{\theta} = (\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\sigma}^2)$, cuando el proceso aleatorio $\{\mathbf{Z}_t\}_{t=1}^n$ ha sido observado, maximizamos el logaritmo de la función de verosimilitud respecto a $\boldsymbol{\theta}$. Pero, como la función de verosimilitud es la densidad conjunta del proceso observado $\mathbf{z} = (z_1, z_2, \dots, z_n)$ condicionada por $\boldsymbol{\theta}$, entonces:

$$L(\mathbf{z} | \boldsymbol{\theta}) = \prod_{t=1}^n D(y_t | \mathbf{x}_t; \boldsymbol{\psi}_1)$$

donde $D(y_t | \mathbf{x}_t; \boldsymbol{\psi}_1)$ es la densidad normal multivariada dada por [3]. Por lo tanto, las propiedades probabilísticas de los estimadores de máxima verosimilitud, de cualquier estadístico de prueba, y medida de bondad de ajuste, serán completamente determinadas por [3].

Supongamos ahora que $M_0(\mathbf{z})$, el modelo de regresión lineal, no cumple con alguno de los supuestos probabilísticos del cuadro 1 *vis-à-vis* los datos, entonces $D(y_t | \mathbf{x}_t; \boldsymbol{\psi}_1)$ será erróneamente especificada e invalidará las propiedades probabilísticas de cualquier estadístico derivado de ésta. Pues bien, ello implica no sólo que la inferencia estadística, las medidas de bondad de ajuste y los pronósticos realizados a partir del modelo estadístico no sean confiables, sino que el modelo completo estará en tela de juicio como PGD.

El supuesto de la distribución normal de \mathbf{Z}_t , $t = 1, 2, \dots, n$, es muy importante en este modelo. Nótese que los supuestos 1-3 en el cuadro 1 dependen de éste, sin embargo, también es uno de los supuestos más difícil de cumplir. En Hoover, Johansen y Juselius (2008) y Hoover y Juselius (2012) se argumenta que la hipótesis de normalidad multivariada de los datos económicos no es una característica que esperamos que se cumpla, más bien es una hipótesis que nos permite asegurar que hemos considerado tanto los eventos *usuales*, que son adecuadamente des-

critos por la distribución normal, como los *milagros*,¹³ que tienden a caer fuera del rango de la distribución normal. Tales eventos extraordinarios, frecuentemente son la causa de asimetría o exceso de curtosis y, por tanto, del rechazo del supuesto de normalidad de u_t . Es decir, la falta de control adecuada de este tipo de eventos puede ser la causa de autocorrelación entre los errores, de sesgo en los estimadores y de imprecisión en las inferencias.¹⁴

Del supuesto de normalidad también depende la linealidad, en variables y parámetros, del modelo estadístico. Por ejemplo, en un modelo donde se supone que el PIB es una función lineal de un conjunto de variables del tipo $PIB = ae_r + b\omega$ (donde a y b son parámetros, e_r es el tipo de cambio real y ω es la participación salarial), si encontramos que el supuesto de *normalidad* no se cumple, la forma funcional del modelo también será cuestionada.

Para Haavelmo, el término de error debe captar todos los factores que influyen en el fenómeno que no fueron considerados por el modelo empírico, y dada la complejidad de la economía real, estos factores debían de ser muchos respecto a los considerados. De modo que si tales factores son independientes, los errores podrían distribuirse aproximadamente como variables aleatorias normales por el teorema del límite central:

if we consider a set of related economic variables, it is, in general, not possible to express any one of the variables as an exact function of the other variables only. There will be an “unexplained rest,” and, for statistical purposes, certain stochastic properties must be ascribed to this rest, a priori. (...) For the necessity of introducing “error terms” in economic relations is not merely a result of statistical errors of measurement. It is as much a result of the very nature of economic behavior, its dependence upon an enormous number of factors, as compared with those which we can account for, explicitly, in our theories. We need a stochastic formulation to make simplified relations elastic enough for applications (Haavelmo, 1943b, p. 1).

El supuesto de independencia de Z_t , $t = 1, 2, \dots, n$, en el proceso observable de variables económicas, también ha sido cuestionado (supuesto 4 del cuadro 1);

13 Entre los eventos extraordinarios caben las reformas económicas y, en general, las intervenciones políticas.

14 Por ejemplo, la autocorrelación en los errores generada al no modelar los cambios en las tasas medias de crecimiento de las variables, puede sugerir el uso innecesario de mayores rezagos en un modelo de vectores autorregresivos (VAR).

de hecho, es muy frecuente observar que el proceso $\{\mathbf{Z}_t\}_{t=1}^n$ muestra algún tipo de dependencia. Por ejemplo, si $\{\mathbf{Z}_t\}_{t=1}^n$ cumple la propiedad de Markov, $P(\mathbf{z}_t | \mathbf{z}_{t-1}, \mathbf{z}_{t-2}, \dots, \mathbf{z}_0) = P(\mathbf{z}_t | \mathbf{z}_{t-1})$ es necesario incluir a \mathbf{z}_{t-1} como regresor en la especificación del modelo para dar cuenta de este tipo de dependencia.

El supuesto de idéntica distribución de \mathbf{Z}_t , $t = 1, 2, \dots, n$, es difícil de cumplir, principalmente en variables macroeconómicas, donde la heterogeneidad de estas variables induce que tanto la esperanza como la matriz de varianzas y covarianzas del proceso observable sean función de t . Es decir, el proceso $\{\mathbf{Z}_t\}_{t=1}^n$ no es estacionario. Por ejemplo, los modelos integrados y cointegrados surgen cuando se sustituye el supuesto de variables idénticamente distribuidas por el de variables estacionarias de segundo orden.

Así pues, la única estrategia correcta para conseguir inferencias válidas y confiables es adoptar un modelo estadístico cuyos supuestos probabilísticos sean válidos *vis-à-vis* los datos antes de realizar cualquier inferencia. Esta es la razón por la cual todo artículo de investigación de economía aplicada, cuyo propósito sea explicar la realidad tal como se manifiesta frente a nosotros, por medio de métodos estadísticos, debe presentar de manera clara y precisa: 1) el diseño de experimento, en el sentido de Haavelmo, propuesto como PGD del fenómeno objeto de la investigación, y sobre el cual se probaron las hipótesis (teóricas), y 2) los resultados de las pruebas de incorrecta especificación, es decir, de las pruebas que evalúan el cumplimiento de los supuestos probabilísticos de las metodologías estadísticas empleadas *vis-à-vis* los datos.

En la siguiente sección se describe y analiza el segundo requisito importante para que exista rigor científico en nuestra disciplina. Nos referimos a la posibilidad de que otro investigador pueda reproducir (*repetir* exactamente los procedimientos con los mismos datos) los números reportados en artículos publicados. Lo cual es importante, porque es lo único que un investigador puede garantizar de un estudio, ya que no puede garantizar que las afirmaciones hechas son correctas, a menos que sean puramente descriptivas. En última instancia, el tiempo nos dirá si las conclusiones de una investigación fueron correctas o incorrectas, pero lo que sí podemos saber en este momento es si los números sobre los cuales se realizaron dichas conclusiones se pueden recalcular o no.

REPLICACIONES EN LA ECONOMÍA

Aún no hay un consenso en la literatura sobre la definición de replicación. Algunas propuestas y debates se pueden consultar en Collins (1991), Cartwright

(1991), Lindsay and Ehrenberg (1993), Hamermesh (2007), Clemens (2015), Zimmermann *et al.* (2015) entre otros.

Hamermesh (2007) define *replicación pura* como hacer algo otra vez exactamente en la misma forma. En investigaciones económicas implica examinar la misma pregunta, modelo y base de datos que la investigación original. Cuando se utiliza diferente muestra, diferente población y, quizá, un modelo similar, pero no idéntico, se define como *replicación científica*.

La definición de Zimmermann *et al.* (2015) comprende tanto el concepto de replicación pura como el de replicación científica. Una *replicación* es repetir un estudio previo para verificar su validez (verificación). Lo cual podría hacerse con los datos y el código computacional originales, pero también podría realizarse con otra muestra de los datos originales (reproducción) o con nuevos datos o con un código computacional similar al que se usó originalmente (robustez). Una definición más amplia implica usar nuevos datos (diferente periodo, geografía); una solución diferente o método de aproximación o un nuevo enfoque estadístico (extensión).

Para el *Journal of Applied Econometrics* (JAE) (Pesaran, 2003) la *replicación en un sentido estricto* implicaría la comprobación de los datos presentados contra las fuentes primarias (cuando se aplique) para la consistencia y exactitud, la comprobación de la validez de los cálculos, ya sea directamente o mediante la realización de la estimación (incluyendo el cálculo de los errores estándar) usando otro software. La *replicación en un sentido amplio* supone que la replicación en sentido estricto ha sido un éxito, y cuestiona si el hallazgo empírico de fondo del artículo puede ser replicado con datos de otros periodos, países, regiones u otras entidades, según proceda.

De estas definiciones se reconocen dos tipos de replications, aquellas cuyo objetivo es *reproducir* los números reportados en un estudio previo, usando los mismos datos y procedimientos, y aquellas que buscan establecer la solidez (*robustness*) de los resultados. Ambos tipos de replications son muy importantes en economía, las primeras constituyen la mejor manera de establecer la validez de los cálculos, mientras que las segundas prueban la generalización de resultados y/o regularidades empíricas documentadas en los artículos originales.

Recientemente *The Replication Network*,¹⁵ organizada por Maren Duvendack y W. Robert Reed, definió un estudio de replicación como: cualquier estudio, cuyo

15 *The Replication Network*. Furthering the Practice of Replication in Economics. Disponible en: <<http://replicationnetwork.com/>>.

objetivo principal sea establecer la exactitud de un estudio previo. Esta definición más general, y por lo tanto más incluyente, parece más adecuada porque reconoce como replicaciones a los ejercicios donde podría haber subjetividad. Es decir, replicaciones donde por falta de acceso a bases de datos y códigos de estimación originales (lo cual sucede frecuentemente en economía, véase la siguiente sección), no es posible realizar la replicación estricta. Sin embargo, el autor replicante recurre a las fuentes originales referidas y siguiendo las indicaciones metodológicas publicadas en el artículo, crea la base de datos y códigos de estimación con alguna discrepancia respecto a los originales (periodo, uso de alguna variable proxi, distinto método de estimación) con el objetivo de replicar los resultados originales publicados y obtener un resultado coherente.

En economía, como en cualquier otra ciencia, se deben realizar replications para establecer la fragilidad o solidez de los resultados de investigaciones publicadas que sirvan de guía en el proceso de generación de conocimiento. Además, las replications podrían desincentivar malas prácticas científicas, como señalan Crocker y Cooper (2011) en la Editorial del número especial de la revista *Science*, titulado *Data Replication & Reproducibility*: “Scientists generally trust that fabrication will be uncovered when other scientists cannot replicate (and therefore fail to validate) findings” (p. 1182). En este sentido, las replications constituyen un medio que nos provee la información para discriminar entre conclusiones establecidas sobre procedimientos que la propia comunidad económica es capaz de juzgar y verificar, y conclusiones generadas en la opacidad de bases de datos y métodos ocultos.

Un beneficio adicional de realizar replications es la mejor comprensión que se adquiere de las ideas de los autores del artículo original cuando se realizan replications. A continuación se describe el actual estado de las replications en economía.

REPLICACIONES DE ESTUDIOS PUBLICADOS EN LAS MÁS IMPORTANTES REVISTAS DE ECONOMÍA

Según el reporte sobre las replications en economía de Duvendack, Palmer-Jones y Reed (2015), de 333 revistas de economía (*Economics journals*) que los autores encontraron listadas en el Thomson Reuters *Journal Citation Report* en 2013, únicamente:

- En veintisiete revistas se puede acceder regularmente (en al menos 50% de los artículos) en línea a los datos y códigos de los artículos empíricos publicados.¹⁶
- En diez revistas se menciona explícitamente en su sitio web que publican estudios de replicación.¹⁷

Estos datos evidencian la escasez de insumos para realizar estudios de replicación de artículos publicados, y la falta de espacios dedicados a la divulgación de los hallazgos de las replications en las principales revistas de economía a nivel internacional

En Duvendack, Palmer-Jones y Reed (2015) también se describe el estado del arte de las replications publicadas en economía, a partir de una muestra de 162 estudios que cumplieron con dos características definidas por los autores para considerar un estudio como un “estudio de replicación”: 1) haberse publicado en una revista *peer-reviewed* y 2) tener como principal objetivo la replicación de un artículo de investigación.¹⁸ De esta muestra, los autores reportan que:

16 Información del cuadro 1 de Duvendack, Palmer-Jones y Reed (2015): 1. *Agricultural Economics*; 2. *American Economic Journal: Applied Economics*; 3. *American Economic Journal: Economic Policy*; 4. *American Economic Journal: Macroeconomics*; 5. *American Economic Journal: Microeconomics*; 6. *American Economic Review*; 7. *Brookings Papers on Economic Activity*; 8. *Econometrica*; 9. *Economic Journal*; 10. *Econometrics Journal*; 11. *Economics: The Open-Access, Open-Assessment E-Journal*; 12. *European Economic Review*; 13. *Explorations in Economic History*; 14. *International Journal of Forecasting*; 15. *Jahrbücher für Nationalökonomie und Statistik/ Journal of Economics and Statistics*; 16. *Journal of Applied Econometrics*; 17. *Journal of Labor Economics*; 18. *Journal of Money, Credit, and Banking*; 19. *Journal of Political Economy*; 20. *Journal of the European Economic Association*; 21. *Quarterly Journal of Economics*; 22. *Review of Economic Dynamics*; 23. *Review of Economic Studies*; 24. *Review of Economics and Statistics*; 25. *Review of International Organizations*; 26. *Studies in Nonlinear Dynamics and Econometrics*, y 27. *World Bank Economic Review*.

17 Información del cuadro 2 de Duvendack, Palmer-Jones y Reed (2015): 1. *Econ Journal Watch*; 2. *Economic Development and Cultural Change*; 3. *Economics of Education Review*; 4. *Empirical Economics*; 5. *Experimental Economics*; 6. *Explorations in Economic History*; 7. *International Journal of Forecasting*; 8. *Jahrbücher für Nationalökonomie und Statistik/ Journal of Economics and Statistics*; 9. *Journal of Applied Econometrics*, y 10. *Review of International Organizations*

18 Es importante señalar que el tamaño de muestra se redujo debido a que los artículos no cumplieron con las características señaladas. Inicialmente, los autores seleccionaron una muestra aleatoria de 1 601 artículos de un total de 13 261 identificados mediante la búsqueda de los términos “replicat*” y “Replicate*” en distintas fuentes que, podríamos decir, abarcaron las 50 revistas de economía más importantes respecto al factor de impacto. La búsqueda se realizó en las siguientes fuentes: Google Scholar y Web of Science, “Replication in Economics” wiki (es un proyecto de la Universidad de Gottinger financiado por el Institute for New Economic Thinking que ha reunido en un sitio web (wiki) un gran número de estudios de replicación publicados en revistas de economía), sugerencias de editores de revistas y colecciones de replications de autores.

- El 59.4% de los estudios de replicación se publicaron en seis revistas: *Journal of Applied Econometrics* (19.1%), *Journal of Human Resources* (11.7%), *American Economic Review* (9.3%), *Econ Journal Watch* (6.8%), *Journal of Development Studies* (6.2%) y *Experimental Economics* (5.6%)
- En el 64% de los estudios, la replicación utilizó exactamente los mismos datos, especificación y procedimiento de estimación que el estudio original.
- En el 50.5% de los estudios, la replicación incluyó una extensión de los hallazgos originales, por ejemplo mediante la validación de los resultados en distintas economías o periodos de análisis.

Respecto a la confirmación de resultados originales:

- En el 66% de los estudios se encontró alguna diferencia significativa respecto al estudio original.
- En el 12.3% de los estudios se obtuvo alguna confirmación junto con, al menos, una diferencia significativa de los principales hallazgos del estudio original.
- En el 21.3% de los estudios se confirmaron todas las conclusiones principales del estudio original.

Otros importantes estudios que mostraron que en la economía muchas investigaciones empíricas no son replicables fueron el de Dewald, Thursby y Anderson (1986), donde los autores sólo fueron capaces de replicar dos de 54 artículos, y el de McCullough, McGeary y Harrison (2006), donde únicamente se pudo replicar 14 de 62 artículos.

Esta información es evidencia de dos hechos que cuestionan seriamente la transparencia y credibilidad de las investigaciones económicas:

1. La inusual publicación de estudios de replicación entre las principales revistas de economía de acuerdo a su factor de impacto.
2. La alta proporción de estudios de replicación publicados, el 78.3% en el reporte de Duvendack, Palmer-Jones y Reed (2015), donde no se valida al menos alguno de los principales hallazgos de los artículos originales.¹⁹

¹⁹ Es posible que esta proporción esté sobrestimada porque algunas revistas únicamente publican replications fallidas. Para evitar este sesgo, en Duvendack, Palmer-Jones y Reed (2015) también se calculó esta proporción considerando únicamente artículos publicados en el *Journal of Applied Econometrics*, donde —según Drukker y Guan (2003)— no existe esta discriminación. No obstante, la proporción de replications que no confirman por lo menos alguno de los principales resultados fue del 65%, es decir, aún muy alta.

A la luz de esta información vale la pena analizar cuáles son las razones de la escasez de replicaciones en materia económica. Sin lugar a dudas, no es la falta de resultados y conclusiones de relevancia en la economía, de hecho existen suficientes investigaciones dignas de ser replicadas. La evidencia en favor de ello está disponible en el sitio *Economics Replication Wiki*,²⁰ por medio de un listado de artículos publicados en las principales revistas de economía, para los cuales no existe replicación. Tampoco es porque los estudios de replicación no generen beneficios para la economía como una ciencia, como ya se apuntó en la sección anterior; las replicaciones son una guía para construir conocimiento con base en conclusiones juzgadas y verificadas por la propia comunidad de economistas. En seguida se describen dos de los principales factores que frenan la realización de replicaciones.

ACCESO A BASES DE DATOS Y CÓDIGOS COMPUTACIONALES

A pesar de que el interés por proveer las bases de datos que permitieran replicaciones de investigaciones en el campo de la economía data desde 1933, y quedó manifestada en la editorial de Ragnar Frisch del primer número de la revista *Econometrica*:

In statistical and other numerical work presented in *Econometrica* the original raw data will, as a rule, be published, unless their volume is excessive. This is important in order to stimulate criticism, control, and further studies. The aim will be to present this kind of paper in a condensed form. Brief, precise descriptions of (1) the theoretical setting, (2) the data, (3) the method, and (4) the results, are the essentials (Frisch, 1933, p. 3),

dicha práctica no ha sido seriamente atendida. McCullough, McGeary y Harrison (2006), en el artículo “Lessons from the JMCB Archive”, señalan que en no pocos casos los artículos no pueden ser replicados ni siquiera en principio, porque los autores generalmente no están obligados a hacer disponibles sus datos y el código para su verificación. Incluso cuando tal requisito existe, puede ser (y es) ignorado impunemente.

²⁰ Disponible en: <http://replication.uni-goettingen.de/wiki/index.php/Category:Study_lacking_replication>.

Previo a la década de los noventa muy pocas revistas contaban con alguna política de acceso a bases de datos y/o publicaron estudios de replicación, tal es el caso de *Journal of Human Resources*, *Journal of Political Economy* y *Journal of Money, Credit and Banking*.

Sin embargo, fue hasta el 2004 cuando, a causa del intento fallido de McCullough y Vinod (2003) por replicar algunos artículos de *American Economic Review* (AER), el entonces editor de esta revista, Ben Bernanke, fortaleció la política de acceso a bases de datos y códigos computacionales. El cambio consistió en requerir las bases de datos y códigos para correr el modelo final, una descripción de cómo fueron usadas las bases de datos primarias y códigos intermedios para crear la base de datos final de artículos econométricos y de simulación. Desde entonces *Econometría*, *Journal of Applied Econometrics*, *Journal of Political Economy* y *Review of Economic Studies* han ido adoptando la política de requerir tanto las bases de datos como los códigos de estimación de los artículos publicados para mantenerlos disponibles en su sitio web.

Es cierto que las políticas de acceso a datos y códigos de algunas de las revistas en economía han mejorado considerablemente en las últimas décadas. No obstante, los autores pueden argumentar que sus bases de datos no pueden ser liberadas por motivos de confidencialidad.

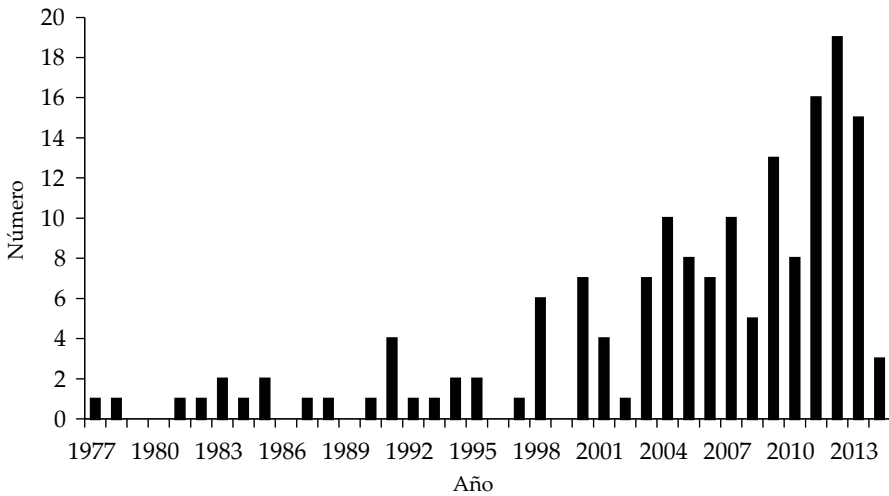
PUBLICACIÓN DE ESTUDIOS DE REPLICACIÓN

Según el reporte de Duvendack, Palmer-Jones y Reed (2015), a partir de las políticas editoriales de acceso a datos de la AER, y desde la creación de la sección de replications del JAE en enero de 2003, ha habido una mayor disposición a la publicación de estudios de replicación que en años anteriores (véase la gráfica 1). No obstante, a partir de la evidencia mostrada por estos autores aún son muy pocas revistas que explícitamente mencionan en su sitio web que publican replications (véase Duvendack, Palmer-Jones y Reed, 2015, cuadro 2, p. 175), entre las cuales destacan: *Journal of Applied Econometrics*, que se limita a publicar réplicas de artículos empíricos propios y publicados en nueve de las más prestigias revistas de economía;²¹ *Econ Journal Watch*, una revista en línea de acceso

21 *Econometría*, *American Economic Review*, *Journal of Political Economy*, *Quarterly Journal of Economics*, *Review of Economics and Statistics*, *Review of Economic Studies*, *Journal of Econometrics*, *Journal of Business* y *Economic Statistics Economic Journal*.

libre y que pone gran énfasis en las replicaciones fallidas. En áreas específicas destacan: *Experimental Economics* (en simulación) y *Empirical Economics*. Otras revistas que no figuran en esta lista, pero que también publican replicación son: *Journal of Human Resources*, *American Economic Review*, *Journal of Development Studies*, *Applied Economics* y *Journal of Economic and Social Measurement*.

Gráfica 1
Frecuencia anual de los estudios de replicación publicados
en las principales revistas económicas (por factor de impacto)



Fuente: Duvendack, Palmer-Jones y Reed (2015).

A partir de 2016, *Economics: The Open Access, Open Assessment e-Journal*, que desde 2009 tiene una política de acceso a datos, comenzó a aceptar replicaciones de cualquier revista de economía *peer-reviewed* y de otras fuentes originales (libros y publicaciones gubernamentales) de reconocido impacto para su publicación.

En el caso de las revistas mexicanas de economía, hasta hace menos de un año, ninguna recibía investigaciones de replicación para su publicación. Fue en el segundo trimestre de 2015 cuando la revista *Investigación Económica*, editada por la Facultad de Economía de la UNAM, comenzó a aceptar estudios de replicación que confirmen o no los hallazgos de artículos publicados.

La divulgación de los resultados de los estudios de replicación debe dejar de ser infravalorada en la economía. Es cierto que una replicación no amplía la

frontera de conocimiento como lo podría hacer un artículo original, pero su importancia radica en evidenciar si las conclusiones establecidas en una investigación lo fueron sobre resultados que no se pueden replicar, o bien, asegurar que la contribución al cuerpo de conocimiento de un artículo está basada en un análisis de datos exitosamente replicable.

A continuación se analizan los incentivos de los editores de revistas para publicar replicaciones, de los autores originales para compartir bases de datos y códigos de estimación y de los autores de estudios de replicación para realizar este tipo de trabajos, que han conducido a la relativa escasez de las replicaciones en la literatura económica.

ACCIONES DE LAS PARTES DE LA PROFESIÓN ECONÓMICA QUE INFLUYEN EN LA REALIZACIÓN Y PUBLICACIÓN DE REPLICACIONES

El *editor de una revista* tiene como objetivo publicar artículos que por su novedad y relevancia académica mantengan el interés de actuales y nuevos lectores. Con la finalidad de asegurar la credibilidad de la revista, el editor debería tener incentivos para mantener una política de acceso a datos y códigos que faciliten la replicación de los artículos que publica.

Sin embargo, la evidencia muestra que la gran mayoría de las revistas (editores) no están interesadas en que las investigaciones que publican sean replicables. Zimmermann *et al.* (2015) define el problema como “un dilema del prisionero”. En la carrera por factores de impacto altos, las revistas que dejen de publicar artículos originales por publicar replicaciones podrían devaluar su contenido respecto a las que no publican réplicas, si acaso las replicaciones son menos citadas que los artículos originales.²² Además, siempre correrán el riesgo de publicar replicaciones que contradigan artículos originales previamente publicados. Por estos motivos, muchas revistas incluso son reacias a incluir una sección de replicaciones.

Los *autores originales*, a sabiendas de que su investigación podría ser objeto de replicaciones y de que les sean solicitados los códigos de estimación de sus modelos econométricos, los archivos de datos originales y de datos subyacentes

22 Aunque también es posible que las revistas que publican replicaciones mejoren sus factores de impacto cuando los estudios de replicación sean mejor valorados por la comunidad de economistas. Es decir, cuando los economistas prefieran investigaciones para las cuales exista algún trabajo donde se haya verificado los cálculos sobre los cuales se fundamentan las conclusiones de los autores originales.

(transformaciones), deberían tener mayores incentivos por realizar rigurosos análisis de datos en sus investigaciones. Lo anterior debido a que una replicación fallida podría implicar desde la pérdida de financiamiento o nombramiento hasta el rechazo académico por falta de credibilidad de su investigación. Por lo tanto, la confianza en las bases de datos y códigos que garanticen la replicación exitosa de una investigación debería ser el principal incentivo para hacer públicos estos archivos.

Empero, una práctica común entre los investigadores es obtener mayor provecho posible de una base de datos desarrollada, por lo que con frecuencia éstas no se hacen públicas. Lo mismo ocurre con el código de estimación, ya que el investigador se abre a la crítica acerca de los procedimientos. Es mucho más fácil ocultar todo. Como ha ocurrido hasta ahora, no afecta mucho a la credibilidad porque la mayoría de los investigadores procede de la misma manera (Dewald, Thursby y Anderson, 1986; McCullough, McGeary y Harrison, 2006) debido a que son pocas las revistas que exigen las bases de datos y código para que una investigación sea publicada.

Los *replicantes* encuentran como mayor desincentivo para realizar trabajos de replicación el costo de oportunidad que significa construir las bases de datos y/o el código de estimación del artículo original a replicar, respecto a las investigaciones propias, estudios de posgrado, docencia u otras actividades.

¿Por qué gastar el tiempo en replications si uno tiene ideas propias y datos para probarlas? Considerando que actualmente el reconocimiento académico, los nombramientos y las plazas de investigación están relacionados con el grado académico, el número de investigaciones publicadas, el prestigio de la revista donde se publican los artículos propios y la cantidad de citas que tengan los artículos de un investigador (véase Liebowitz, 2014), y dado que muy pocas revistas publican replications, es más probable que una investigación de conocimiento nuevo sea publicada que una replicación. Por lo tanto, el beneficio académico de realizar estudios de replicación no es muy grande.

A menos que exista la voluntad de todas las partes de la profesión económica (revistas, investigadores, profesores y estudiantes) por acoger el principio de la replicación, las replications seguirán siendo inusuales en la literatura económica. Si nuestro interés es hacer que las investigaciones sean replicables, lo primero que podemos hacer es dar acceso a los datos y códigos de estimación de nuestras investigaciones. Pero esto no es lo único que como comunidad académica podemos hacer. Por ejemplo:

- Si usted es árbitro de un artículo en una revista con política de acceso a datos, cuando le pidan dictaminar un artículo de investigación debería verificar los datos y códigos de estimación. Incluso en los casos donde el acceso a éstos le sea negado, cabría la posibilidad de negarse a arbitrar dicha investigación.
- Si usted es jurado de un concurso de oposición o de un comité que otorga reconocimientos a la investigación, becas u otras posiciones u honores, antes de emitir su voto sería recomendable que consulte el sitio web de los candidatos, y el de las revistas donde éstos hayan publicado, para verificar el acceso a datos y códigos de estimación de sus investigaciones. Lo anterior, con la finalidad de reconocer en tales promociones u honores a aquellos candidatos que no temen a la replicación de sus investigaciones.
- Si usted es investigador, podría dejar de citar artículos no replicables o bien hacer explícito que sus conclusiones podrían no ser confiables debido a que los códigos y datos de cierta investigación no están disponibles.

Ahora bien, con el objetivo de incrementar los incentivos para realizar estudios de replicación:

- Si usted es editor de una revista, podría iniciar la publicación de estudios de replicaciones.
- Si usted es organizador de conferencias o seminarios, podría incluir exposiciones o alguna sesión de replicaciones.
- Si usted es jurado o forma parte de un comité que otorga posiciones o reconocimientos, valore los estudios de replicación que los candidatos hayan realizado.
- Si usted es profesor, podría promover la realización de replicaciones entre sus alumnos.

A fin de contribuir con la realización de replicaciones en economía, en la siguiente sección se desarrolla una replicación del artículo “Long run real exchange rates: evidence from Mexico” de Kakkar (2001) publicado en la revista *Economic Letters*.

UNA REPLICACIÓN DEL ARTÍCULO “LONG RUN REAL EXCHANGE RATES: EVIDENCE FROM MEXICO” (ECONOMICS LETTERS, 2001)

La replicación que a continuación se presenta no es una replicación exacta debido a que la base de datos original no es pública,²³ y a que se va a emplear un

23 En el artículo de Kakkar (2001) no se menciona que las bases de datos o el código de estimación serán enviados (abiertos) a los lectores que lo soliciten, y la revista *Economic Letters* no da acceso a ellos.

método de estimación distinto al que usó el autor. Sin embargo, dado que se utiliza el mismo periodo de estudio, fuente de datos y especificación de variables que en el artículo original, la replicación podría ser considerada como un estudio que busca confirmar la solidez (robustez) de los hallazgos publicados empleando un método de estimación distinto.

El objetivo de la investigación de Kakkar (2001) fue determinar si los cambios en los precios relativos de las mercancías no comerciables pueden explicar los cambios persistentes del tipo de cambio real bilateral entre México y Estados Unidos a lo largo del periodo 1955-1996. La conclusión de la investigación es que existe una relación de cointegración entre el tipo de cambio real y los precios relativos de mercancías no comerciables a comerciables de México y de Estados Unidos que confirma su hipótesis.

La modelación de Kakkar (2001) parte del hecho estilizado de que el tipo de cambio real es no estacionario, así como de los siguientes supuestos:

1. El tipo de cambio real se define con índices de precios generales (deflactor del PIB) mediante la siguiente ecuación:

$$e_r = \frac{p^{GDP}}{Ep^{GDP^*}}$$

donde E denota el tipo de cambio nominal (pesos por un dólar), y el índice general de precios en cada país se define como el promedio geométrico de los precios de mercancías comerciables (p^T) y no comerciables (p^N) ponderados por un factor α , según la siguiente ecuación: $p^{GDP} = c^{GDP} [p^T]^\alpha [p^N]^{(1-\alpha)}$. Las variables de Estados Unidos se denotan con el superíndice *.

- De acuerdo a la definición de Kakkar (2001), un incremento (decremento) de e_r equivale a una apreciación (depreciación) real del peso respecto al dólar.
2. Se supone que en el largo plazo se cumple la paridad de poder de compra (PPP por sus siglas en inglés, *Purchasing Power Parity*) en mercancías comerciables, es decir:

$$\ln(p^T) = \ln(E) + \ln(p^{T*}) + u$$

donde u es una variable aleatoria estacionaria con media cero, lo cual implica que un choque en el precio de mercancías comerciables en un país que desvíe al tipo de cambio real de la condición de PPP se desvanecerá con el tiempo.

De los supuestos 1 y 2, el tipo de cambio real (TCR) se puede escribir de la siguiente manera:

$$\ln(e_r) = \alpha \ln\left(\frac{p^N}{p^T}\right) - \alpha^* \ln\left(\frac{p^{N^*}}{p^{T^*}}\right) + \varepsilon \quad [4]$$

donde $\varepsilon = \ln(c^{GDP}) - \ln(c^{GDP^*}) + u$. En esta ecuación se observa que la dinámica del TCR depende de la dinámica de los precios relativos (no comerciables a comerciables) internos y externos. En la cual un incremento de los precios internos de los bienes no comerciables respecto a los comerciables conduce a una apreciación real, y un incremento de los precios externos de los bienes no comerciables respecto a los comerciables conduciría a una depreciación real.

Si $\ln(p^N/p^T)$ y $\ln(p^{N^*}/p^{T^*})$ son integrados de orden uno y se cumple la PPP en las mercancías comerciables en el largo plazo, entonces las variables $\ln(e_r)$, $\ln(p^N/p^T)$ y $\ln(p^{N^*}/p^{T^*})$ deberían de cointegrar con vector de cointegración igual a $(1, -\alpha, \alpha^*)$. Por lo tanto, para estimar los coeficientes de la ecuación [4] es necesario que se satisfaga la hipótesis de raíz unitaria, lo que permitirá, posteriormente, probar la hipótesis de cointegración de las variables.

Para estimar la ecuación [4] en el periodo de análisis 1955-1996, el autor utilizó el índice de precios al mayoreo (IPW) y el índice de precios al consumidor (IPC) como variables proxy de los precios de mercancías comerciables y no comerciables respectivamente.²⁴ La validación empírica se realizó con datos anuales del Fondo Monetario Internacional (FMI), específicamente del CD-ROM de las *International Financial Statistics* (IFS) [*Estadísticas Financieras Internacionales*, EFI] (IMF, 2015). Según lo indicado en Kakkar (2001), las tres variables involucradas en la ecuación [4] se especifican como en el cuadro 2, pero no se indica el año base de los índices de precios.

Los resultados de las pruebas de raíces unitarias en Kakkar (2001) no rechazan la hipótesis de raíces unitarias de las variables $\ln(e_r)$, $\ln(q_{mx})$ y $\ln(q_{us})$ con las pruebas Dickey-Fuller aumentada (ADF) (Dickey y Fuller, 1979) y Philips-Perron (PP) (Phillips y Perron, 1988) al 90% de confianza.

24 El autor asume que el IPC se basa en una canasta fija de bienes y servicios consumidos por el hogar promedio, y es probable que tenga una gran participación de los bienes no transables en forma de servicios de venta al por menor, vivienda, transporte y otros servicios. Mientras que, el IPW se limita, generalmente, a los bienes del sector agrícola y manufacturero que son en gran medida comerciables.

Cuadro 2
Base de datos para estimar la ecuación [4], 1955-1996

E	Tipo de cambio nominal pesos / dólar, promedio del periodo
p^{GDP}	Deflactor del PIB
p^T	IPW proxi de precios de mercancías comerciables
p^N	IPC proxi de precios de mercancías no comerciables
$\ln(e_r)$	$= \ln\left(\frac{p^{GDP}}{E p^{GDP^*}}\right)$
$\ln(q_{mx})$	$= \ln\left(\frac{p^N}{p^T}\right)$
$\ln(q_{us})$	$= \ln\left(\frac{p^{N^*}}{p^{T^*}}\right)$

Nota: variables sin (con) asterisco corresponde a México (Estados Unidos).

La hipótesis de cointegración fue probada mediante una regresión de cointegración canónica (véanse Park, 1992, Ogaki, 1993 y Ogaki y Park, 1997) estimada en el paquete CCR de Gauss desarrollado por Ogaki (1993). Kakkar (2001, p. 83), con base en sus estimaciones publicadas en el cuadro 2 (reproducida en el Apéndice 2), concluye que a los niveles de significancia convencionales no se rechaza la hipótesis de cointegración estocástica con los estadísticos $H(0,2)$ y $H(0,3)$, y tampoco se rechaza la hipótesis de cointegración determinística con la estadística $H(0,1)$. Asimismo, el autor señala que los coeficientes estimados de la ecuación [4], $\alpha = 3.731$ y $\alpha^* = -1.163$, están de acuerdo a lo esperado en la teoría, en signo y magnitud, y son estadísticamente significativos.

Sin embargo, el autor no reporta los estadísticos o valores p (p-values) de las pruebas de significancia ni la distribución de los estadísticos de prueba ni los niveles de confianza empleados en las pruebas de significancia. Tampoco reporta ninguna prueba de incorrecta especificación asociada al modelo.

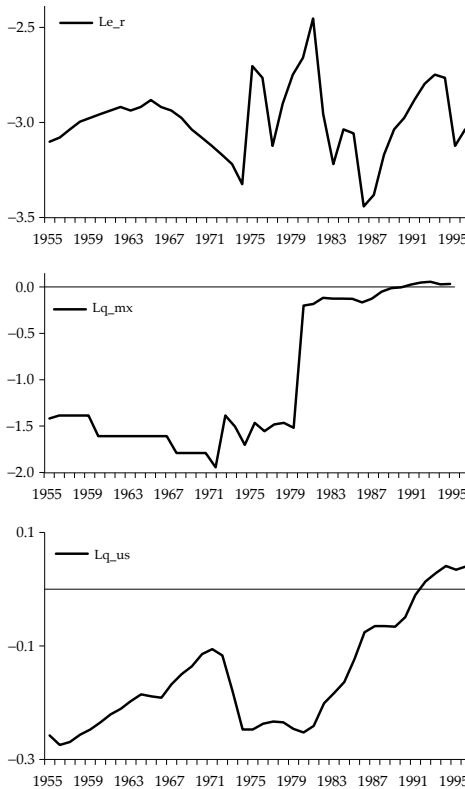
RESULTADOS DE LA REPLICACIÓN

La replicación de la investigación de Kakkar (2001) consiste en realizar pruebas de raíces unitarias y de traza que permitan determinar el rango de cointegración del sistema $\mathbf{x}_t = (\ln(e_{r_t}), \ln(q_{mx_t}), \ln(q_{us_t}))$ en el mismo periodo de estudio del artículo de Kakkar, 1955-1996 ($t = 1, 2, \dots, 42$), a fin de estimar los coeficientes de la ecuación [4].

La base de datos fue construida usando la misma especificación de las variables que el autor uso (véase el cuadro 2). La única diferencia es que los datos provienen de la edición de 2015 de las IFS, donde los índices de precios usan como base al año 2000. En la gráfica 2 se muestran las series de tiempo de las variables $\ln(e_r)$, $\ln(q_{mx})$ y $\ln(q_{us})$.

Además de las pruebas de raíces unitarias reportadas en Kakkar (2001), en la replicación se muestran los resultados de las pruebas Kwiatkowski, Phillips, Schmidt y Shin (KPSS) (Kwiatkowski *et al.*, 1992) y la Zivot-Andrews Z-A (Zivot y Andrews, 2002). Ésta última debido a que la serie de tiempo de la variable $\ln(q_{mx})$ presenta un cambio permanente en el nivel de la serie para 1981.

Gráfica 2
Tipo de cambio real, precios relativos de mercancías no comerciables a comerciables en México y Estados Unidos, 1955-1996



Fuente: elaboración propia con datos de las IFS, IMF (2015).

A diferencia de la metodología econométrica utilizada en Kakkar (2001), en la replicación el análisis de cointegración fue realizado con la metodología de Johansen (1988), descrita en Juselius (2006), y las estimaciones se realizaron en PcGive. A partir de un VAR(k) correctamente especificado para el vector $\mathbf{x}_t = (\ln(e_r), \ln(q_{mx_t}), \ln(q_{us_t}))$, la hipótesis de cointegración se contrasta con la prueba de la traza, y la identificación del vector de cointegración (β) se realiza en el modelo corrector de errores que representa al VAR(k) restringido por el rango de cointegración:

$$\Delta \mathbf{x}_t = \pi + \sum_{l=1}^{k-1} \Phi_l \Delta \mathbf{x}_{t-l} + \alpha \beta' \mathbf{x}_{t-1} + \mathbf{u}_t, \mathbf{u}_t \sim IIDN(\mathbf{0}, \Omega)$$

En el cuadro 3 se muestran los resultados de las pruebas de raíces unitarias en el periodo 1955-1996.²⁵ A continuación se comentan los resultados para cada variable.

Cuadro 3
Pruebas de raíces unitarias 1955-1996
(modelo con constante)

	ADF	PP	KPSS	Z-A cambio estructural año	
$\ln(e_r)$	-3.446	-3.554	0.045	-3.306	1975
$\ln(q_{mx})$	-0.590	-0.483	0.633	-10.274	1981
$\ln(q_{us})$	-0.096	-0.211	0.533	-4.873	1973
Valores críticos (modelo con constante):			1%	5%	10%
ADF y PP			-3.60	-2.94	-2.61
KPSS			0.74	0.46	0.35
Z-A			-5.34	-4.93	-4.58

La variable $\ln(e_r)$ rechaza la hipótesis de raíz unitaria al 5% de significancia, pero no al 1% con las pruebas ADF y PP, con la prueba KPSS no existe evidencia estadística suficiente para rechazar la hipótesis de estacionariedad de $\ln(e_r)$. No obstante, la prueba Z-A no rechaza la hipótesis nula de raíz unitaria que excluye cambio estructural. De acuerdo a esta información, se concluye que hay

25 Las pruebas de raíces unitarias se realizaron con las rutinas de Eviews 8.

evidencia suficiente en favor de la no estacionariedad del TCR para el periodo 1955-1996.

La variable $\ln(q_{mx})$ no rechaza la hipótesis de raíz unitaria con las pruebas ADF y PP a los niveles de significancia usuales, con la prueba KPSS rechaza al 5% de significancia la hipótesis de estacionariedad. Sin embargo, la prueba Z-A rechaza al 99% de confianza la hipótesis nula de raíz unitaria, lo que implica que $\ln(q_{mx})$ es estacionaria con cambio estructural para 1981, lo cual parece evidente del comportamiento descrito en la gráfica 2. En resumen, a pesar que las pruebas ADF, PP y KPSS encuentran evidencia estadística a favor, $\ln(q_{mx}) \sim I(1)$, la conclusión de la prueba Z-A, $\ln(q_{mx}) \sim I(0)$, es la adecuada, puesto que, como Perron (1989) mostró, la potencia de las pruebas ADF y PP para rechazar una raíz unitaria disminuye cuando la hipótesis alternativa (la variable es estacionaria) es verdadera y se ignora un cambio estructural.

La variable $\ln(q_{us})$ presenta raíz unitaria, puesto que las pruebas ADF y PP no rechazan esta hipótesis, la prueba KPSS rechaza al 5% la hipótesis de estacionariedad y la prueba Z-A no rechaza al 5% de significancia la hipótesis de raíz unitaria sin cambio estructural.

La conclusión final acerca de la heterogeneidad de los datos en el periodo 1955-1996 es que: $\ln(e_r) \sim I(1)$, $\ln(q_{mx}) \sim I(0)$ y $\ln(q_{us}) \sim I(1)$. A consecuencia de ello, el análisis de cointegración del vector $\mathbf{x}_t = (\ln(e_r), \ln(q_{mx}), \ln(q_{us}))$ es inadecuado.

Sólo con la finalidad de observar el resultado del análisis de cointegración de la ecuación [4], se ignorará el resultado de la prueba Z-A para $\ln(q_{mx})$.

La especificación del modelo VAR(2) para el vector $\mathbf{x}_t = (\ln(e_r), \ln(q_{mx}), \ln(q_{us}))$, en el periodo 1955-1996 ($t = 1, 2, \dots, 42$), incluye término constante; una variable dummy de impulso en 1973, D_{1973} (toma el valor de uno en 1973, cero en otro caso), y una variable dummy de cambio en 1981, S_{1981} (toma el valor de uno en el periodo 1955-1980, cero en cualquier otro caso). El modelo cumple con los supuestos probabilísticos del modelo estadístico, según se observa en las pruebas de incorrecta especificación del cuadro 4.²⁶

26 Hendry y Juselius (2001) reconocen que la inferencia estadística de un modelo VAR es sensible a la validez de los supuestos de parámetros constantes, la no correlación serial y la simetría de la distribución de los errores, mientras que la inferencia se mantiene moderadamente robusta al exceso de curtosis (distribuciones de cola pesadas) y heterocedasticidad de los errores. Por lo que recomiendan asegurar, especialmente, la validez de los tres primeros supuestos.

La prueba de la traza (véase el cuadro 5) determinó el rango de cointegración del sistema en $r = 1$, lo cual implica que existe una combinación lineal estable, $\beta'x_t$, denominada también relación de largo plazo o de cointegración, entre las variables del sistema cuya solución para $\ln(e_t)$ se interpreta como la relación de equilibrio de largo plazo del TCR, y está dada por la siguiente ecuación:²⁷

$$\ln(e_t) = 1.70 \ln\left(\frac{p^N}{p^T}\right) + 2.23 * \ln\left(\frac{p^{N*}}{p^{T*}}\right) + 3.25S_{1981} - 3.04 \quad [4']$$

La principal diferencia respecto al resultado que Kakkar (2001, p. 83) reporta en el cuadro 2, es que en la relación de cointegración estimada por el método de Johansen la variable dummy de cambio (S_{1981})²⁸ es estadísticamente significativa (véase el cuadro 6) y el signo estimado de los precios relativos de las mercancías no comerciables respecto a las comerciables en Estados Unidos es incorrecto, puesto que éste debía ser negativo de acuerdo al marco teórico que sirve de fundamento al estudio del autor.

En la gráfica 3 se muestra la relación de cointegración dada por $\beta'x_t$ y el tipo de cambio real de largo plazo estimado mediante el análisis de cointegración dado por [4'].

La evidencia estadística muestra que los resultados reportados por Kakkar (2001) no son robustos al método de estimación. Más aún, las pruebas de raíces unitarias del cuadro 3 muestran que los precios relativos de los no comerciables en México, $\ln(q_{mx_t})$, para el periodo 1955-1996 son una variable estacionaria con cambio estructural. Esto implica que no se cumple el supuesto que exige el análisis de cointegración²⁹ requerido para el cumplimiento de la hipótesis principal de la investigación:

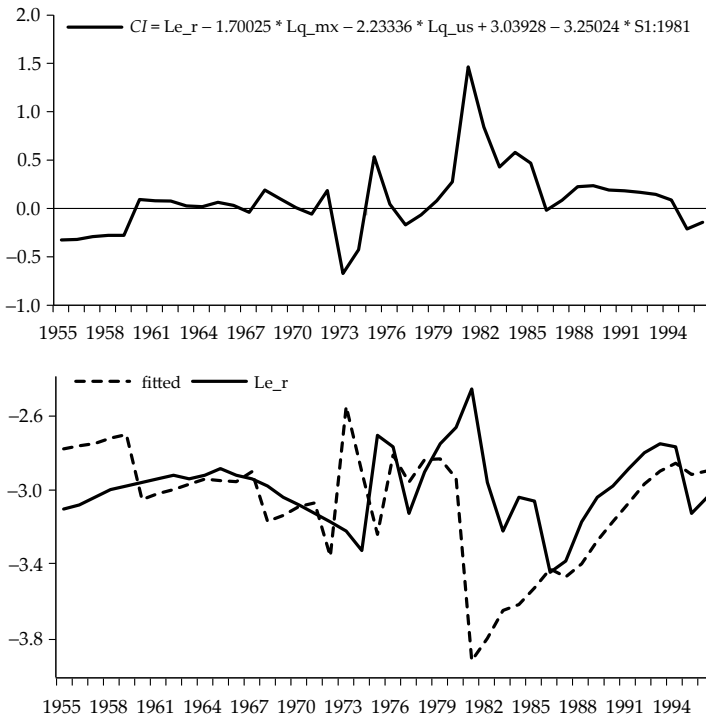
27 La relación de equilibrio del TCR de la ecuación [4'] rechaza al 5% de significancia la hipótesis nula de raíz unitaria con la prueba ADF con constante, estadístico de prueba $\tau = -3.591$, valores críticos 5% = -2.94 y 1% = -3.61 de significancia. Este resultado nos lleva a concluir que la relación es estacionaria.

28 Las pruebas de exclusión se realizaron condicionadas por el orden de cointegración del sistema, es decir, $r = 1$. El estadístico de la prueba del cociente de verosimilitudes LR (*Likelihood Ratio*) se distribuye asintóticamente como χ^2 , con un grado de libertad.

29 La cointegración del conjunto de datos, por ejemplo, $x_t = (\ln(e_t), \ln(q_{mx_t}), \ln(q_{us_t}))$ está basado en un supuesto, $x_t \sim I(1)$.

The main purpose of this paper is to study whether permanent (or persistent) changes in the relative price of nontradable goods can explain the persistent movements of Mexico's real exchange rate. In other words, we test whether the Mexican-US bilateral real exchange rate is cointegrated with the Mexican and US relative price of nontradables. The evidence supports the existence of this cointegrating relationship (Kakkar, 2001, p. 80).

Gráfica 3
Relación de cointegración y tipo de cambio real del peso respecto al dólar observado y de equilibrio de largo plazo, 1955-1996



Aunque no resulta indispensable para nuestro trabajo, anotemos aquí que, por ejemplo, Kakkar y Ogaki (1999) refieren argumentos por los cuales el TCR y los precios relativos internos y externos pudieran no cointegrar. En primer lugar, muchos bienes comerciables entre países no son idénticos. También es difícil encontrar bienes comerciables puros porque los bienes comerciables a menudo se combinan con los servicios no comerciables, como los servicios de venta al

por menor. Por lo tanto, la PPP no puede mantenerse para estos bienes comerciables en el largo plazo. En segundo lugar, los pesos colocados (en el índice general de precios) en bienes comerciables y no comerciables pueden no ser estables y moverse junto con el TCR. Tercero, las variables proxies, disponibles de los precios relativos de los bienes comerciables y no comerciables, pueden ser aproximaciones pobres de los precios relativos. Probablemente alguno de estos argumentos se aplique al caso del TCR del peso respecto al dólar en el periodo 1955-1996.

Cuadro 4
Pruebas de incorrecta especificación del VAR(2)
en el periodo 1955-1996

Pruebas individuales (valores p)						
	<i>Portmanteau</i> (5) $\chi^2_{(3)}$	<i>AR</i> 1-2 F(2,29)	<i>Normalidad</i> $\chi^2_{(2)}$	<i>Hetero</i> F(13,25)	<i>Hetero-X</i> F(28,10)	<i>ARCH</i> 1-1 F(1,38)
$\ln(e_t)$	2.0537 [0.5613]	0.32993 [0.7216]	4.317 [0.1155]	1.0208 [0.4628]	0.85257 [0.6511]	0.020593 [0.8867]
$\ln(q_{mx})$	1.0115 [0.7985]	0.20725 [0.8140]	1.0342 [0.5962]	2.4054 [0.0288]*	3.6547 [0.0179]*	0.47404 [0.4953]
$\ln(q_{us})$	1.5851 [0.6628]	0.82641 [0.4477]	0.93099 [0.6278]	3.2621 [0.0054]**	1.6914 [0.1927]	1.1217 [0.2962]
Pruebas conjuntas						
	<i>Portmanteau</i> (5) $\chi^2_{(27)}$	<i>AR</i> 1-2 F(18,65)	<i>Normalidad</i> $\chi^2_{(6)}$	<i>Hetero</i> F(78,116)	<i>Hetero-X</i> F(168,37)	<i>RESET</i> 23 F(18,65)
	44.335 [0.0191]*	1.4702 [0.1307]	5.1642 [0.5229]	1.4073 [0.0471]*	2.6278 [0.0005]**	2.0826 [0.0166]*

Cuadro 5
Pruebas de cointegración

Rango	<i>Trace test</i>	[Prob]	<i>Max test</i>	[Prob]	<i>Trace test</i>	(T-nm)	<i>Max test</i>	(T-nm)
0	114.15	[0.000]**	91.64	[0.000]**	97.03	[0.000]**	77.89	[0.000]**
1	22.51	[0.022]*	15.76	[0.051]	19.14	[0.070]	13.39	[0.121]
2	6.76	[0.144]	6.76	[0.144]	5.74	[0.219]	5.74	[0.219]

Cuadro 6
Coefficientes estimados del vector de cointegración
y pruebas de exclusión

	$\ln(e_r)$	$\ln(q_{mx})$	$\ln(q_{us})$	Constante	S_{1981}
Coefficiente	1	-1.70	-2.23	3.04	-3.25
$\chi^2_{(1)}$		53.233	13.485	13.034	74.283
[valor p]		[0.0000]**	[0.0002]**	[0.0003]**	[0.0000]**

REFLEXIONES FINALES

En toda investigación, el autor debe garantizar el análisis de datos sobre el cual establece las conclusiones de su investigación. En economía, los datos representan observaciones pasivas de la realidad, es decir, no hay posibilidad de repetir experimentos en la población objetivo bajo las mismas circunstancias para generar tantas observaciones como deseemos. El enfoque probabilístico de la econometría de Haavelmo establece un conjunto de principios metodológicos para la economía aplicada que permite el uso de métodos estadísticos para evaluar teorías económicas, o hipótesis económicas, contra los datos con todo rigor científico.

Bajo el enfoque de Haavelmo, si suponemos que los datos, de los cuales disponemos, de las variables que intervienen en la relación económica de estudio se pueden considerar como una realización de la distribución conjunta del proceso observable, podemos parametrizar un modelo estadístico haciendo algunos supuestos. Para que las estimaciones de los parámetros y las inferencias realizadas a partir de la distribución conjunta del proceso observable sean estadísticamente válidas, es necesario que los supuestos probabilísticos del modelo estadístico no sean rechazados por los datos, es decir, que el modelo estadístico esté correctamente especificado.

Ahora bien, en la economía, como en cualquier otra ciencia, la generación de conocimiento debe basarse en investigaciones replicadas y replicables. Replicadas, es decir, que un investigador independiente, haciendo uso de los datos originales y el código de estimación original, obtenga (o no) los mismos resultados que fueron reportados en el estudio original. Tales estudios establecen la solidez (o fragilidad) de la investigación en cuestión. Replicables, es decir, de acceso libre a datos y códigos de estimación, para mitigar los altos costos de

oportunidad que significa crear las bases datos y códigos de estimación para verificar los resultados y el análisis estadístico de un artículo original.

Me parece que, a menos que exista una fuerte demanda profesional por la transparencia y una voluntad académica por hacer las investigaciones empíricas replicables, las replicaciones en economía seguirán siendo inusuales. Por lo tanto, debemos cambiar algunos comportamientos en aras de hacer replicables los resultados económicos, comenzando por: 1) dar acceso a los datos y códigos de estimación de nuestras investigaciones y 2) crear espacios para la divulgación de tales estudios de replicación.

De no promover la realización y publicación de los estudios de replicación, estamos impidiendo que tales estudios: 1) mediante la verificación, garanticen que la generación de conocimiento en nuestra disciplina esté libre de errores y 2) mediante el uso de conjuntos de datos alternos (distinta definición de variables, distinto periodo de estudio o economía de análisis), métodos de estimación y modelo, doten de solidez y/o extiendan los hallazgos de una investigación publicada.

REFERENCIAS

- Cartwright, N., 1991. Replicability, reproducibility, and robustness: Comments on Harry Collins. *History of Political Economy*, 23(1), pp. 143-155.
- Clemens, M.A., 2015. The meaning of failed replications: A review and proposal. [Discussion Paper no. 9000]. *Institute for the Study of Labor (IZA)*, Bonn, Alemania.
- Collins, H.M., 1991. The meaning of replication and the science of economics. *History of Political Economy*, 23(1), pp. 123-142.
- Crocker, J. y Cooper, M.L., 2011. Addressing scientific fraud. *Science*, 334(6060) pp. 1182-1182.
- Dickey, D. A., and Fuller, W.A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, 74(366a), pp. 427-431.
- Dewald, W.G., Thursby, J.G. y Anderson, R.G., 1986. Replication in empirical economics: The journal of money, credit and banking project. *The American Economic Review*, 76(4), pp. 587-603.
- Drukker, D.M. y Guan, W., 2003. Replicating the results in 'on efficient estimation with panel data: an empirical comparison of instrumental variables estimators'. *Journal of Applied Econometrics*, 18(1), p. 119.

- Duvendack, M., Palmer-Jones, R.W. y Reed, W.R., 2015. Replications in Economics: A progress report. *Econ Journal Watch*, 12(2), pp. 164-191.
- Fisher, R.A., 1922. On the mathematical foundations of theoretical statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, 222, pp. 309-368.
- Frisch, R., 1933. Editorial. *Econometrica*, 1(1), pp. 1-4.
- Haavelmo, T., 1943a. Statistical testing of business-cycle theories. *The Review of Economic and Statistics*, 25(1), pp. 13-18.
- Haavelmo, T., 1943b. The statistical implications of a system of simultaneous equations. *Econometrica*, 11(1), pp. 1-12.
- Haavelmo, T., 1944. The probability approach in econometrics. *Econometrica* (Supplement), 12, pp. iii-115.
- Hamermesh, D.S., 2007. Viewpoint: Replication in economics. *Canadian Journal of Economics (Revue canadienne d'économique)*, 40(3), pp. 715-733.
- Hendry, D.F., 2016. Deciding between alternative approaches in macroeconomics [Department of Economics Discussion Paper Series no. 778]. *University of Oxford, Department of Economics*, Oxford, Reino Unido.
- Hendry, D.F. y Juselius, K., 2001. Explaining cointegration analysis: Part II. *The Energy Journal*, 22(2), pp. 75-120.
- Hendry, D.F. y Nielsen, B., 2007. *Econometric Modeling: A Likelihood Approach*. Princeton, NJ: Princeton University Press.
- Henry, D.F. y Richard, J.F., 1983. The econometric analysis of economics time series. *International Statistical Review*, 51(2), pp. 111-163.
- Hoover, K.D., Johansen, S. y Juselius, K., 2008. Allowing the data to speak freely: The macroeconometrics of the cointegrated vector autoregression. *The American Economic Review*, 98(2), pp. 251-255.
- Hoover, K.D. y Juselius, K., 2012. Experiments, passive observation and scenario analysis: Trygve Haavelmo and the cointegrated vector autoregression [Discussion Paper no. 12-16]. *University of Copenhagen, Department of Economics*, Copenhagen, Dinamarca.
- IMF (International Monetary Fund), 2015. *International Financial Statistics* (Edición: diciembre 2015). UK Data Service. DOI: <http://dx.doi.org/10.5257/imf/ifs/2015-12>
- Johansen, S., 1988. Statistical analysis of cointegration vectors. *Journal of Economic Dynamics and Control*, 12(2), pp. 231-254.
- Juselius, K., 2006. *The cointegrated VAR model: Methodology and applications*. Oxford: Oxford University Press.

- Kakkar, V., 2001. Long run real exchange rates: evidence from Mexico. *Economics Letters*, 72(1), pp. 79-85.
- Kakkar, V. y Ogaki, M., 1999. Real exchange rates and nontradables: A relative price approach. *Journal of Empirical Finance*, 6(2), pp. 193-215.
- Kolmogorov, A.N., 1933. *Foundations of the Theory of Probability*. 2da edición. Nueva York: Chelsea Publishing Co.
- Kwiatkowski, D., Phillips, P.C.B., Schmidt, P. y Shin, Y., 1992. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1), pp. 159-178.
- Liebowitz, S.J., 2014. Willful blindness: The inefficient reward structure in academic research. *Economic Inquiry*, 52(4), pp. 1267-1283.
- Lindsay, R.M. y Ehrenberg, A.S.C., 1993. The design of replicated studies. *The American Statistician*, 47(3), pp. 217-228.
- McCullough, B.D., McGeary, K. A. y Harrison, T.D., 2006. Lessons from the JMCB archive. *Journal of Money, Credit, and Banking*, 38(4), pp. 1093- 1107.
- McCullough, B.D. y Vinod, H.D., 2003. Verifying the solution from a nonlinear solver: A case study. *The American Economic Review*, 94(1), pp. 873- 892.
- Ogaki, M., 1993. GMM: A user guide [Rochester Center for Economic Research Working Paper no. 348]. *University of Rochester*, Rochester, Nueva York.
- Ogaki, M. y Park, J.Y., 1997. A cointegration approach to estimating preference parameters. *Journal of Econometrics*, 82(1), pp. 107-134.
- Park, J.Y., 1992. Canonical cointegrating regressions. *Econometrica*, 60(1), pp. 119-143.
- Perron, P., 1989. The great crash, the oil price shock and the unit root hypothesis, *Econometrica*, 57(6), pp. 1361-1401.
- Pesaran, H., 2003. Introducing a replication section. *Journal of Applied Econometrics*, 18(1), p. 111.
- Phillips, P. y Perron, P., 1988. Testing for a unit root in time series regression. *Biometrika*, 75(2), pp. 335-346.
- Spanos, A., 1986. *Statistical Foundations of Econometric Modelling*. Cambridge, MA: Cambridge University Press.
- Spanos, A., 2015. Revisiting Haavelmo's structural econometrics: Bridging the gap between theory and data. *Journal of Economic Methodology*, 22(2), pp. 171-196.
- Yule, G.U., 1926. Why do we sometimes get nonsense-correlations between Time-Series?-a study in sampling and the nature of time-series. *Journal of the royal statistical society*, 89(1), pp. 1-63.

Zimmermann, C. *et al.*, 2015. On the need for a replication journal [Working Paper Series no. 2015-016]. *Federal Reserve Bank of St. Louis*, St. Louis, MO.
 Zivot, E. y Andrews, D.W.K., 2002. Further evidence on the great crash, the oil-price shock, and the unit-root hypothesis. *Journal of Business & Economic Statistics*, 20(1), pp. 25-44.

APÉNDICE 1

Teorema. Sea:

$$\begin{pmatrix} Y \\ X \end{pmatrix} \sim N_{p+1} \left[\begin{pmatrix} \mu_Y \\ \mu_X \end{pmatrix}, \begin{pmatrix} \Sigma_{YY} & \Sigma_{YX} \\ \Sigma_{XY} & \Sigma_{XX} \end{pmatrix} \right]$$

donde Y es una variable aleatoria escalar y X es un vector aleatorio p -dimensional, $\mu_Y = E(Y)$, $\mu_X = E(X)$, $\Sigma_{YY} = Var(Y) = E[(Y - E(Y))(Y - E(Y))']$, $\Sigma_{XX} = Var(X) = E[(X - E(X))(X - E(X))']$, $\Sigma_{YX} = Cov(Y, X) = E[(Y - E(Y))(X - E(X))']$, $\Sigma_{XY} = Cov(X, Y) = E[(X - E(X))(Y - E(Y))'] = \Sigma'_{YX}$ y Σ_{XX} es una matriz no singular. Entonces, condicionando sobre $X = x$, la variable aleatoria $(Y|X = x)$ se distribuye normal con esperanza condicional $E(Y|X = x) = \alpha + \beta'X$ donde $\beta = \Sigma_{XX}^{-1} \Sigma_{XY}$ y $\alpha = \mu_Y - \beta' \mu_X$ y con varianza condicional $Var(Y|X = x) = \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY}$. Es decir:

$$Y|X = x \sim N(\alpha + \beta' X, \Sigma_{YY} - \Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY})$$

APÉNDICE 2

Regresión de cointegración canónica entre el tipo de cambio real y los precios relativos de bienes no comerciables, 1995-1996

$\alpha^a/$	$\alpha^{*a}/$	H(0,1) ^{b/}	H(0,2) ^{b/}	H(0,3) ^{b/}
3.731	-1.163	1.508	0.021	0.139
(-0.83)	(-0.428)	(0.219)	(0.885)	(0.933)

Notas: a/ Errores estándar en paréntesis. b/ p-values en paréntesis.
 Fuente: Kakkar (2001, cuadro 2, p. 83).