

Principios básicos de microeconometría y del uso de variables instrumentales para la inferencia causal

*Principles of microeconometrics and instrumental
variables for causal inference*

Primer envío: 13/09/2018; segundo envío: 20/02/2019; aceptado: 03/04/2019

*Owen Eli Ceballos Mina**

Resumen

La microeconometría por medio de su lógica de identificación causal y del uso de microdatos, conduce a un conjunto de técnicas de medición y modelaje muy particular en el marco de la economía aplicada. El presente artículo constituye una presentación de los rasgos distintivos de esta rama considerando que su auge va de la mano de la creciente disponibilidad de información de corte transversal y de datos tipo panel. Se identifican tres elementos como principios característicos de los estudios microeconómicos: *i*) las estrategias de identificación causal por medio del aislamiento de la correlación que resulta de las variables omitidas o de la selección muestral; *ii*) el modelaje no lineal para variables discretas y censuradas como

* Profesor Investigador del Departamento de Economía de la Universidad Autónoma Metropolitana Unidad Azcapotzalco. (Correo electrónico: oecm@azc.uam.mx). El autor agradece a los tres dictaminadores anónimos que contribuyeron al desarrollo de este ensayo con sus invaluable comentarios; también agradece a Fabian Uvaldo Mendoza y a Luis Francisco Rodríguez, ayudantes del proyecto “Laboratorio Cuanti” de la UAM-Azcapotzalco, por su asistencia de investigación.

modelos *logit*, *probit* o *tobit*; y *iii*) la reducción de supuestos acerca de las distribuciones de las variables y la flexibilización de las formas funcionales. Asimismo, se hace una breve exposición del problema frecuente de endogeneidad y del uso de variables instrumentales como una sencilla pero potente estrategia de solución; y finalmente, se recuerda que las técnicas microeconómicas necesitan del resto de elementos que dan forma al conjunto de investigación para alcanzar el máximo de su potencial.

Palabras Clave: microeconometría, inferencia causal, variables instrumentales.

Clasificación JEL: C01, C10, C26, D00

Abstract

Microeconometrics with its causal identification logic and the microdata use, leads to a set of specific measurement and modeling techniques in applied economics. This paper shows basic elements of microeconometrics, taking account that the peak of its application goes together with availability rise of cross-sectional and panel-type data. It describes three representative features of microeconometry: i) strategies of causal identification for isolating the correlation from omitted variables or from sample selection; ii) non-linear modeling for discrete and censored variables such as Logit, Probit or Tobit models; and iii) reduction of assumptions about the distributions and the flexibilization of functional forms. In addition, a brief exposition is made of the frequent problem of endogeneity and the use of instrumental variables as a simple but powerful solution strategy; and finally, it is recalled that the microeconomic techniques without the rest of the elements that shape the research set, loses its potential.

Keywords: Microeconometry, causal inference, instrumental variables.

Clasificación JEL: C01, C10, C26, D00

Introducción

Una de las ramas más activas de la ciencia económica es aquella que se ocupa de la aplicación de métodos cuantitativos para el análisis del comportamiento de las variables y las relaciones que observamos en nuestro acontecer diario. Comúnmente llamamos economía aplicada al uso de datos para analizar y resolver problemas prácticos relacionados con la conducta de los agentes y la política económica (Pulido y Pérez, 2006; Vicéns, 1998; Ibarra, 1998). Más allá de discusiones de si la economía aplicada es una división legítima o si sencillamente es un complemento indivisible de la forma de hacer ciencia en nuestra disciplina, la mayoría de economistas coin-

cidimos en lo limitado del análisis de datos sin el adecuado fundamento teórico; y probablemente también estemos de acuerdo en que, el creciente volumen de estudios empíricos que se apoyan en la observación de la realidad son un acierto para el fin de una ciencia social como la economía (Greene, 1999; Angrist y Pischke, 2009).

En este gran universo de estudios empíricos, los que se centran en la aplicación de herramientas econométricas usando datos de individuos, familias, empresas o unidades territoriales como municipios, estados o regiones, pertenecen al ámbito microeconómico. En particular, la microeconometría se centra en el uso de métodos de regresión con datos de corte transversal o tipo panel para el análisis del comportamiento económico (Cameron y Trivedi, 2005; Wooldridge 2002, Hausman, 2001). El reciente auge del análisis microeconómico es resultado de los constantes desarrollos teóricos en la materia y de los desarrollos tecnológicos que permiten un incremento en la disponibilidad de microdatos y de la capacidad de cómputo; en la actualidad se puede encontrar un sinnúmero de artículos que contrastan sus relaciones teóricas con técnicas estadísticas que utilizan datos a nivel individual de corte transversal o longitudinales (tipo panel).

En el conjunto de la econometría aplicada, el modelaje microeconómico con su rasgo distintivo de inferencia causal, se ha convertido en una herramienta recurrente en ramas como la economía laboral, economía de la salud, economía regional, organización industrial, entre otras. Una visible consecuencia de este auge del uso de la técnica es la gran lista de métodos microeconómicos que, como sugieren Angrist y Pischke (2009), es un “intrincado menú” del que se deben seleccionar las estrategias adecuadas para atender la problemática de interés. Entre los errores más frecuentes y con peores consecuencias en el ámbito aplicado, destaca el atribuir dones mágicos a la técnica y centrarse en ella sin reconocer su verdadero alcance; en ocasiones se olvida que la medición microeconómica es sólo una parte del análisis y que lo realmente importante es el problema en su conjunto: causas, hipótesis, medición y alternativas de solución.

La finalidad de este ensayo es realizar una introducción a los principios característicos de la microeconometría, asentando los elementos básicos para un uso familiar del menú de estrategias disponibles en esta rama. Se debe tener en mente que de la gran lista de técnicas existentes no todas son igual de importantes; hay un núcleo de herramientas fundamentales para poder iniciar con la aplicación y comprensión del análisis microeconómico; pero no debemos olvidar que ni este núcleo básico, ni las técnicas más sofisticadas son efectivas sin una visión estratégica del conjunto de análisis. Con tales premisas, esta investigación apunta en tres dimensiones específicas: *i*) la naturaleza y los rasgos distintivos del análisis microeconómico en el marco de estudio del comportamiento económico; *ii*) la

relevancia de los supuestos de identificación y el uso de variables instrumentales como estrategia básica frente a la problemática de endogeneidad; y *iii*) los alcances y límites del uso de las técnicas de medición y modelaje microeconómico en el análisis económico y las ciencias sociales en general.

Este artículo se desarrolla en tres apartados incluyendo esta introducción y sus conclusiones. En el primero sobre el papel de la microeconometría en la economía aplicada, se conceptualizan los elementos que hereda la rama particular de la econometría en general; se enuncian los rasgos distintivos del análisis, la importancia de su lógica de identificación causal, los conjuntos de datos disponibles en México, y se expone brevemente el modelo de regresión como punto de partida para el análisis. La segunda sección se adentra en el problema de endogeneidad como un incidente frecuente de identificación causal, y se hace una exposición sintética del método de variables instrumentales como una potente alternativa de solución. En la tercera sección, sobre los usos y alcances de la microeconometría, se menciona el potencial de su aplicación en la evaluación de política pública, los modelos más usados y se refieren algunos estudios microeconómicos en México. En las conclusiones de forma adicional a los principios sentados a lo largo del ensayo, se hace un llamado al uso cauto de la técnica microeconómica en el marco de la economía aplicada.

1. Microeconometría en la economía aplicada

El objeto de la econometría ha sido el mismo desde que la sociedad econométrica a principios de los años 30 en el siglo XX estableció que la estadística, la economía matemática y la teoría económica son condiciones necesarias, pero no suficientes para la comprensión de las relaciones cuantitativas en el mundo económico (Frisch, 1933; Haavelmo, 1944; Greene, 1999).¹ El consenso apunta a que la econometría es ese campo en que la estadística matemática y la inferencia estadística son aplicadas para probar las relaciones empíricas sugeridas por la teoría económica (Wooldridge, 2000; Greene, 1999; Theil, 1971; Goldberger, 1964).² Algunos autores definen la econometría en su tipología teórica como el uso de las herramientas de la estadística matemática para apoyar el ejercicio empírico de la economía (Gujarati y Porter,

¹ Según Fernández y Adalid (2000) el concepto de econometría fue introducido por primera vez en 1926 por el noruego Ragnar Frisch para referir a los estudios económicos que hacían uso de métodos estadísticos.

² La lista de autores que documentan la historia y evolución de la econometría es extensa, pero no es el objetivo de este ensayo detallar dicha secuencia; revisiones desde la perspectiva histórica puede ser vistas en Goldberger (1964), Fernández y Adalid (2000); entre otros.

2010; Fernández y Adalid, 2000); a pesar de esta línea teórica, dada la naturaleza social de la ciencia económica, el papel esencial de la econometría en la economía aplicada es proveer un conjunto de herramientas cuantitativas para el dialogo entre la información que describe el comportamiento de los agentes y las predicciones de la teoría económica.

Con el paso del tiempo y según la atención que la economía ha dedicado a los fenómenos en sus dos principales niveles de análisis, la econometría ha marcado al menos dos vertientes bien definidas en su ejercicio empírico; *i*) una vinculada con el uso de series de tiempo para dar seguimiento a los importantes agregados macroeconómicos; y *ii*) otra relacionada con el uso de datos de individuos, empresas, territorios u otro tipo de información pertinente para el análisis a nivel microeconómico (Ruiz, 2016; Gujarati y Porter, 2010; Cameron y Trivedi, 2005; Enders, 2004; Ibarra, 1998). Este ensayo se centra en la vertiente de la aplicación econométrica relacionada con el uso de datos a nivel microeconómico. La microeconometría como rama específica de la econometría hereda su naturaleza; la acción conjunta de la teoría económica, las matemáticas y la inferencia estadística, en el particular mundo de los datos microeconómicos, dan origen al análisis microeconométrico.

Entre los rasgos procedentes de la econometría, el énfasis sobre la estrategia de identificación para la inferencia causal es una de las características más deseables en los estudios microeconométricos (Wooldridge, 2000; Wooldridge, 2002; Cameron y Trivedi, 2005). Los manuales en la materia dedican buena parte de su introducción a recordar la importancia de identificar relaciones causa-efecto para poder hacer predicciones sobre el cambio de estado de una variable y lograr control sobre ciertos escenarios (Angrist y Pischke, 2009). Sin restar importancia a la parte descriptiva del análisis, en economía comúnmente se prioriza la identificación de relaciones causales que van más allá de la simple asociación o correlación entre variables; este camino para identificar el parámetro causal de x sobre el resultado de interés y , depende tanto del conjunto de herramientas estadísticas como del tipo de datos disponibles para el análisis.

1.1 Microdatos y rasgos distintivos de la microeconometría

Si bien la economía experimental se ha encargado de la producción de datos en ambientes controlados donde un grupo de tratamiento, un grupo de control y la aleatorización permiten identificar relaciones causales sin grandes sofisticaciones estadísticas, también es cierto que es una rama de poca expansión dado su alto costo de desarrollo (Cameron y Trivedi, 2005; Smith, 2005). Por otro lado, la automatización y el uso de la tecnología en los procesos de recolección, registro y almace-

namiento de censos, encuestas e información administrativa, ha traído consigo una gran disponibilidad de datos no experimentales. En México contamos con un gran acervo de información de libre acceso como resultado del ejercicio del Instituto Nacional de Estadística y Geografía (INEGI), instituciones académicas, centros de investigación y algunos sectores de la administración pública.

Del conjunto de información microeconómica que se produce en el país destacan los censos y conteos de población y vivienda que se realizan cada cinco años, los censos económicos y agropecuarios, las encuestas nacionales agropecuarias (ENA), las encuestas anuales y mensuales en establecimientos económicos de comercio, industria y otros sectores, las encuestas nacionales de ingresos y gastos de los hogares (ENIGH), las encuestas nacionales de inclusión financiera (ENIF), las encuestas de movilidad social (EMOVI), las encuestas nacionales de salud y nutrición (ENSANUT), las encuestas tipo panel rotativo de ocupación y empleo (ENOE), las encuestas panel nacional de niveles de vida (ENNVIIH), las encuesta panel nacional de salud y envejecimiento en México (ENASEM), por sólo mencionar algunas en la gran variedad de fuentes de información disponibles a nivel nacional, como se muestra en el cuadro 1 sobre la disponibilidad de fuentes de información y microdatos en México.³

Una característica visible del conjunto de información que se produce a nivel nacional es la gran proporción de encuestas que se levantan para un único momento en el tiempo (datos de corte transversal); esta abundancia de datos de sección cruzada genera gran diversidad y especialización en los temas disponibles, desde el nivel de detalle en la descripción de los ingresos y gastos de los hogares que logra la ENIGH hasta la especialización en aspectos tecnológicos (ENDUTIH), adicciones (ENADIC), juventud (ENJU), violencia y seguridad (ENVIF, ENVIM, ENVIPE), finanzas (ENAFIN), entre otras. Pese a que los datos longitudinales brindan mejores posibilidades para la identificación causal, el seguimiento del mismo individuo en varios periodos de tiempo es un objetivo que muy pocas encuestas se proponen dados los altos costos que implica este tipo de levantamiento.⁴

³ El sitio www.bdsocial.inmujeres.gob.mx del Instituto Nacional de las Mujeres hace una buena presentación de una larga lista de encuestas disponibles en el país, la descripción temática de cada una y el vínculo de descarga para los microdatos en la mayoría de los casos.

⁴ Véase Ruiz (2016) para una revisión histórica del uso de datos panel en México. El autor confirma que el uso de paneles de datos frente al conjunto de estudios aplicados es relativamente reducido. En su clasificación general del uso de datos panel incluye ejercicios macroeconómicos, sin dejar de reconocer que el mayor volumen de producción procede del ámbito microeconómico.

Cuadro 1
Disponibilidad de fuentes de información y tipos de microdatos en México

<i>Fuente de información</i>	<i>Tipo de micro-dato</i>	<i>Temática</i>	<i>Frecuencia y últimos levantamientos</i>	<i>Dato informativo</i>	<i>Responsable</i>
Censo de Población y Vivienda	Censal	Contabiliza la población en sus viviendas, sus características demográficas y socioeconómicas.	Decenal 2000 y 2010	En 1895 inician los censos contemporáneos	INEGI
Conteo de Población y Vivienda	Censal	Da continuidad a los censos, a la contabilidad de la población en sus viviendas y sus características.	Decenal 1995 y 2005	A partir de 1995 incrementa la serie histórica del censo de población.	INEGI
Censo económico	Censal	Información sobre todos los establecimientos productores, comercializadores de bienes y prestadores de servicios.	Quinquenal 2009 y 2014	Comenzó en 1930 como un censo manufacturero, y se extendió a toda la actividad económica.	INEGI
Encuestas anuales de: comercio (EAC), manufacturas (EAIM), construcción (EAEC).	Encuesta de establecimientos	Muestran datos sobre comercio, manufacturas y construcción.	Anuales	Dieron inicio: EAC en 1994 EAIM en 1963 EAEC en 2000	INEGI
Encuestas mensuales: Comercio (EMEC), manufacturas (EMIM).	Encuesta de establecimientos	Muestran datos sobre comercio y manufacturas.	Mensuales	Dieron inicio: EMEC en 1983 EMIM en 1976	INEGI
Encuesta nacional agropecuaria (ENA)	Encuesta de hogares y establecimientos	Muestra datos de los principales productos agrícolas, ganaderos y forestales.	2012, 2014 y 2017	Adicional: características de productores y sus hogares.	INEGI
Encuesta nacional de ingresos y gastos de los hogares (ENIGH)	Encuesta de hogares	Proporciona un panorama detallado de la distribución de los ingresos y los gastos de los hogares.	Bienal 2014 y 2016	Inició en 1984. A partir de 2008, se publica nueva serie nueva construcción de variables.	INEGI
Encuesta Nacional de Inclusión Financiera (ENIF)	Encuesta de individuos en hogares	Información estadística e indicadores oficiales sobre inclusión y educación financiera.	Trienal 2015 y 2018	Inició en 2009 para medir acceso y uso de servicios financieros.	CNBV - INEGI

<i>Fuente de información</i>	<i>Tipo de micro-dato</i>	<i>Temática</i>	<i>Frecuencia y últimos levantamientos</i>	<i>Dato informativo</i>	<i>Responsable</i>
Encuesta de Movilidad Social (EMOVI)	Encuesta de hogares	Información sobre factores laborales, educativos y familiares que determinan y promueven la movilidad social.	2006, 2011 y 2015	Ofrece información sobre hijos (encuestados) como de los padres cuando el encuestado era menor.	Centro de Estudios Espinosa Yglesias (CEEY)
Encuesta Nacional de Salud y Nutrición (ENSANUT)	Encuesta de individuos en hogares	Información sobre el estado de salud y las condiciones nutricionales de la población.	Sexenal 2006, 2012 y 2016-Medio Camino	Antecedentes: Encuestas Nacionales de Nutrición (ENN) 1988 y 1999. Inicio: ENSANUT 2006	Instituto Nacional de Salud Pública (INSP)
Encuesta Nacional de Ocupación y Empleo (ENOE)	Encuesta panel rotativo-individuos en hogares	Información sobre las características ocupacionales y económicas para el análisis de aspectos laborales.	Trimestral	Antecedentes: ENEU y ENE. Inicio: ENOE 2005	INEGI
Encuesta Nacional de Niveles de Vida en los Hogares (ENNVIH)	Encuesta panel de hogares	Información sobre distintas dimensiones del bienestar de la población y su cambio a través del tiempo.	2002, 2005 y 2009	Encuesta longitudinal con línea basal en 2002 seguimientos en 2005 y 2009	UIA, CIDE, INSP, UCLA
Encuesta Nacional sobre Salud y Envejecimiento en México (ENASEM)	Encuesta panel de individuos en hogares	Información sobre el proceso de envejecimiento, el impacto de las enfermedades y la discapacidad en la población de 50 años y más.	2001, 2003, 2012 y 2015	Línea basal en 2001	INEGI, UTMB, INSP, INGER

Fuente: elaboración propia con información del INEGI, CEEY, CNBV, INSP, ENNVIH y ENASEM.

A diferencia de lo que ocurre con los datos experimentales, en el caso de los observacionales la identificación de relaciones causales en el análisis conlleva un mayor esfuerzo estadístico (Cameron y Trivedi, 2005; Wooldridge, 2002), en tanto que, la interrelación de las variables observadas y no observables desencadenan un conjunto de problemas comunes y rasgos distintivos en la rama microeconómica. De la parte observada se pueden desprender complicaciones como los errores de medición, la simultaneidad, el truncamiento de variables y la presencia de datos discretos; del lado de lo no observable brotan dificultades como la endogeneidad por variables omitidas, los sesgos de selección muestral, entre otros asociados con

la heterogeneidad no observada. El conjunto de estas incidencias estadísticas, junto con el establecimiento de las hipótesis de distribución de las variables y relaciones, dan origen al reto de identificación causal.

Dado el problema de identificación, la búsqueda de soluciones en los estudios microeconómicos habitualmente parte del análisis de regresión -al igual que en toda la rama econométrica-, adicionando elementos que la hacen muy particular; entre las características del análisis microeconómico destaca el uso de métodos específicos para aislar efectos causales, el modelaje con métodos no lineales, el modelaje con la disminución de hipótesis paramétricas de distribución, el uso de más inferencia estadística con base en errores estándar robustos y la teoría asintótica.

De las herramientas para aislar los efectos causales sobresalen los modelos de variables instrumentales, los modelos de ecuaciones simultáneas, los de corrección de sesgo de selección, los de corrección de error de medida, los modelos panel de efectos fijos y los modelos de diferencias en diferencias. En el uso de métodos no lineales es imposible dejar de mencionar los modelos *Probit* y *Logit* para el tratamiento de variables dependientes discretas; así como los modelos *Tobit* en el caso de variables dependientes censuradas y truncadas. Cuando se piensa en la disminución de hipótesis de distribución paramétrica y la flexibilidad de la forma funcional, aparecen herramientas como los modelos semiparamétricos y no paramétricos.

1.2 Un punto de partida: análisis de regresión

Aunque el análisis de regresión en su concepción moderna consiste en la estimación o predicción del valor esperado de una variable dependiente y en términos de los valores fijos de la variable explicativa x_{k-1} , esta relación estadística en sí misma, no implica causalidad sin algunas consideraciones *a priori* (Angrist y Pischke, 2009; Gujarati y Porter, 2010; Greene, 1999); en especial suponemos un comportamiento ortogonal entre la matriz x de variables explicativas y el vector de perturbaciones aleatorias ε (o errores estocásticos), que son las desviaciones del valor esperado de la variable dependiente condicional en x , con respecto a su valor observado en la muestra⁵:

$$\begin{aligned} y_i &= E[y|\mathbf{x}_i] + \varepsilon_i & (1) \text{ Modelo de regresión paramétrico} \\ E[\varepsilon|\mathbf{x}] &= 0 & (S.1) \text{ Perturbaciones aleatorias} \end{aligned}$$

⁵ A lo largo de este texto las variables son denotadas con letras minúsculas en cursivas como x o y ; los vectores con letras minúsculas o símbolos en negrita como, β o ε ; una matriz se representa con una letra en negrita como \mathbf{x} donde x_{ik} se refiere a la i -ésima observación de la k -ésima variable de la matriz.

Es posible resumir la regresión paramétrica en (1) y su supuesto de identificación (S.1) en tres enunciados: *i*) dados los valores fijos de \mathbf{x} se quiere predecir el valor esperado de la variable y para el individuo i en una muestra de tamaño n ; *ii*) dado que la relación entre las $k-1$ variables explicativas y la dependiente no es exacta más que en términos teóricos, existe un conjunto de desviaciones $\boldsymbol{\varepsilon}_i = \mathbf{y}_i - E[y|\mathbf{x}_i]$ llamadas perturbaciones aleatorias que representan el conjunto de variables no observadas por el modelo; y *iii*) para poder hablar del efecto causal de una variable de x_{k-1} en y , se necesita que x_{k-1} no esté vinculada con aquellas que no se incluyeron en el modelo.

En el conocido caso del modelo de regresión lineal, tomando en cuenta que $E[y|\mathbf{x}_i]$ es una función de \mathbf{x} , de manera conveniente asignamos una forma funcional paramétrica que relaciona a \mathbf{x} y a y por medio de un conjunto de parámetros lineales $\boldsymbol{\beta}$, tal y como se presenta en la ecuación (2). El modelo de regresión lineal múltiple hereda el supuesto de que el valor esperado de las perturbaciones aleatorias debe ser cero (S.1); este último implica que la covarianza entre \mathbf{x} y $\boldsymbol{\varepsilon}$ es cero (S.2), dado que $\text{Cov}[\mathbf{x}, E[\boldsymbol{\varepsilon}|\mathbf{x}]] = 0$; asimismo, se hacen consideraciones adicionales acerca de la varianza y distribución de la perturbación aleatoria en (S.3) para una estimación adecuada del vector parámetros $\boldsymbol{\beta}$.

$$\begin{array}{ll}
 y = \mathbf{x}\boldsymbol{\beta} + \boldsymbol{\varepsilon} & (2) \textit{ Modelo de regresión lineal} \\
 \text{Cov}[\mathbf{x}, \boldsymbol{\varepsilon}] = 0 & (S.2) \textit{ Exogeneidad o independencia} \\
 \boldsymbol{\varepsilon}|\mathbf{x} \sim N[0, \sigma^2 \mathbf{I}] & (S.3) \textit{ Normalidad y homocedasticidad}
 \end{array}$$

En este punto el objetivo estadístico es calcular un valor de $\boldsymbol{\beta}$ tal que sea lo más parecido posible al valor del parámetro poblacional de la relación de interés. Para estimar los parámetros del modelo de regresión lineal bajo los supuestos establecidos, la minimización del cuadrado de las perturbaciones aleatorias, conocidos como estimadores mínimos cuadrados ordinarios (MCO), generan valores de $\boldsymbol{\beta}$ con las características que se buscan. Tras el proceso de optimización, los estimadores MCO se definen:

$$\hat{\boldsymbol{\beta}}_{\text{MCO}} = (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\mathbf{y} \quad (3) \textit{ Estimador MCO}$$

Los estimadores $\hat{\boldsymbol{\beta}}_{\text{MCO}}$ poseen propiedades deseables para la inferencia en el caso de muestras finitas tales como la insesgadez (el valor esperado del estimador muestral $\hat{\boldsymbol{\beta}}_{\text{MCO}}$ es igual al valor del parámetro poblacional $\boldsymbol{\beta}_{\text{Pob}}$) y la eficiencia (la varianza del estimador es mínima); y en el caso de muestras grandes, la propiedad asintótica de consistencia (el límite de la probabilidad de $\hat{\boldsymbol{\beta}}_{\text{MCO}}$ es igual a $\boldsymbol{\beta}_{\text{Pob}}$ a

medida que el tamaño de la muestra n crece). Particularmente la insesgidez y la consistencia del vector de estimadores $\hat{\beta}_{MCO}$ dependen invariablemente de los supuestos de independencia (S.2) y del valor esperado de las perturbaciones aleatorias (S.1).⁶

$$E[\hat{\beta}_{MCO}] = \beta_{Pob} \quad (P.1) \text{ Insesgidez mco}$$

$$\text{plim}_{n \rightarrow \infty} \hat{\beta}_{MCO} = \beta_{Pob} \quad (P.2) \text{ Consistencia MCO}$$

Por lo que, con las mencionadas consideraciones se tiene una buena estimación del efecto causal de x_{k-1} sobre y . La pregunta inmediata es ¿qué ocurre cuando el regresor del parámetro de interés causal no cumple con el supuesto de independencia?; es decir, ¿qué ocurre cuando la variable independiente está correlacionada con las perturbaciones?

$$\text{Cov}[x, \varepsilon] \neq 0 \quad (\text{RS.2}) \text{ Problema de endogeneidad}$$

2. El problema de endogeneidad

El problema de endogeneidad es uno de los más recurrentes y debe ser considerado entre los más importantes del análisis microeconómico. Básicamente, el problema consiste en que no se cumple el supuesto de independencia ante la correlación entre la variable de interés y la perturbación aleatoria; esta correlación puede resultar de la omisión de variables relevantes no observadas, de un error de medición en la variable explicativa de interés, de la simultaneidad de ecuaciones o del caso de autoselección muestral. Un ejemplo clásico en la literatura microeconómica es el ejercicio de Thomas Mroz (1987), sobre la oferta de horas de trabajo de las mujeres casadas en Estados Unidos; el autor enfrenta dos problemas de endogeneidad: *i*) las mujeres se autoseleccionan para participar en el mercado laboral mientras el promedio de horas ofrecidas depende de dicha decisión (se usa una solución tipo *Tobit*); y *ii*) dada la no observación de variables relevantes que afectan tanto al salario como a las horas ofrecidas, el salario de las mujeres casadas es una variable endógena en la explicación de la oferta laboral (se soluciona por variables instrumentales).

⁶ Dado que en estricto sentido la propiedad de consistencia sólo requiere de exogeneidad débil vista como $E[\varepsilon|x] = E[\varepsilon] = 0$ entonces S.1 y S.2 son condiciones suficientes para su cumplimiento. Una revisión de las propiedades de los estimadores MCO a nivel básico se puede encontrar en Gujarati y Porter (2010, Cap. 7) y Wooldridge (2001, Cap. 3); una revisión más formal se puede encontrar en los respectivos apéndices de los dos libros y en Greene (1999).

Otro ejemplo que aparece como una aplicación frecuente en los libros de texto de microeconometría para abordar el problema de endogeneidad y el aislamiento de efectos causales, es la estimación de los retornos de la escolaridad en los salarios; dada la imposibilidad de observar el nivel de habilidad del individuo que afecta tanto los salarios como a la educación, muchos estudios sobre el tema asumen el carácter endógeno de la variable años de educación como una dificultad econométrica preponderante (Wooldridge, 2002, 2000; Angrist, Imbens y Rubin, 1996; Card 1993; Angrist y Krueger, 1991).

El problema estadístico se resume en que dado $\hat{\beta}_{MCO} = (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\mathbf{y}$, si sustituimos el modelo poblacional $\mathbf{y} = \mathbf{x} \beta_{Pob} + \varepsilon$, se obtiene:

$$\hat{\beta}_{MCO} = \beta_{Pob} + (\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\varepsilon \quad (4) \textit{ Estimador MCO}$$

Reescribiendo (4) como la diferencia de estimadores y tomando el valor esperado en ambos lados de la ecuación, se tiene que $E[\hat{\beta}_{MCO} - \beta_{Pob}] = \mathbf{E}[(\mathbf{x}'\mathbf{x})^{-1} \mathbf{x}'\varepsilon]$. Esto implica que en el caso $\text{Cov}[\mathbf{x}, \varepsilon] \neq 0$ de la diferencia entre el valor estimado del estimador muestral y el valor poblacional no es cero; es decir, el estimador $\hat{\beta}_{MCO}$ posee un sesgo asociado a la correlación entre la variable omitida y la variable explicativa.

Entre las soluciones comunes al problema de endogeneidad se puede optar por extraer la variable omitida del error estocástico, adicionando en la especificación del modelo una variable proxy de la variable omitida; sin embargo, la corrección del sesgo será tan buena como la aproximación observada de la variable omitida. En el ejercicio de los retornos de la escolaridad sobre el salario y la no observación de la habilidad, Wooldridge (2000) sugiere que la medición del coeficiente intelectual de los individuos "IQ" puede ser buena variable proxy de la habilidad; sin embargo, la experiencia empírica ha demostrado que no siempre es fácil encontrar variables observadas que sean buenas para sustituir las variables omitidas relevantes como ocurre en el caso del efecto de los salarios en la oferta laboral (Mroz, 1987). Otra opción frente al problema es adicionar información de la variable explicativa endógena que no tenga relación con la variable omitida; es decir, instrumentar el regresor endógeno de interés (Angrist y Krueger, 1991, Angrist et al., 1996, Hausman 2001).

2.1 Variables instrumentales para la inferencia causal

El método de variables instrumentales (VI) es sin duda la técnica más popular en la corrección del sesgo de endogeneidad en el modelaje microeconómico; es usado en muchos artículos científicos en busca de solucionar el problema de inferencia

causal. Para remontarnos a sus antecedentes históricos Goldberger (1972) provee referencias que conducen a los trabajos de Wright (1928, 1934) y Haavelmo (1943). El uso de VI ha sido una respuesta común frente a problemas como las variables relevantes omitidas, los errores de medición en las variables explicativas, el sesgo de autoselección muestral y el problema de estimación de ecuaciones simultáneas (Wooldridge, 2000; Cameron y Trivedi, 2005).

La técnica consiste en proponer información adicional por medio de variables conocidas como instrumentos para aislar la correlación entre la variable explicativa y la perturbación aleatoria; un instrumento debe tener la propiedad de explicar el regresor endógeno, pero sin estar directamente vinculado con la explicación de la variable dependiente o la variable omitida del modelo en su forma estructural.⁷ En términos formales, sea el modelo lineal

$$y = \alpha + \beta x + \varepsilon \quad \text{tal que } \text{Cov}[x, \varepsilon] \neq 0;$$

entonces una variable z es un instrumento o variable instrumental del regresor endógeno x si cumple con dos condiciones:

$$\text{Cov}[z, x] \neq 0 \quad (\text{S.5}) \textit{ Condición de relevancia}$$

$$\text{Cov}[z, \varepsilon] = 0 \quad (\text{S.6}) \textit{ Condición de exclusión o exogeneidad}$$

Tal como propone Mroz (1987), una variable instrumental de los salarios para su identificación sobre la oferta laboral de mujeres casadas es la variable “experiencia”; dicha variable tiene efectos directos sobre el salario percibido, pero no tan claros sobre el número de horas ofrecidas, por lo que puede considerarse que cumple con las condiciones necesarias de un buen instrumento. De igual forma, en el caso de la identificación de la decisión de participación de las mujeres casadas, el autor propone que la misma experiencia y el número de hijos menores de 6 años tienen efectos sobre la decisión de trabajar, pero no afectan directamente las horas de trabajo ofrecidas. Mroz (1987) prueba la existencia del sesgo de autoselección sobre la oferta laboral, así como la sensibilidad de las estimaciones del salario al conjunto de instrumentos; el autor muestra la relevancia de la selección de los instrumentos y el cumplimiento de sus condiciones, para la corrección del sesgo de endogeneidad.

⁷ En econometría se hace referencia a la forma estructural del modelo cuando la variable dependiente se escribe en función de variables explicativas endógenas y sus alternativas de identificación; en contraste, en una forma reducida la variable dependiente es escrita exclusivamente en términos de variables exógenas, es decir, variables que no están correlacionadas con la perturbación aleatoria.

En el caso de los retornos de la escolaridad sobre los salarios y la no observación de la habilidad de los individuos, Card (1993) propuso como variable instrumental “la proximidad geográfica” de la residencia del individuo a la escuela, argumentando que la gente que vive más lejos de las universidades tiene menos probabilidad de asistir a ella. Parece que este instrumento al igual que en el caso de la experiencia en los salarios, cumple con la condición de relevancia, pero puede discutirse si cumple con la condición de exogeneidad; la gente que vive lejos de las universidades puede también estar más lejos del mercado de trabajo por lo que podrían tener salarios más bajos; quizás este instrumento necesite de controles adicionales -como indicadores de zonas metropolitanas- para lograr satisfacer plenamente las dos condiciones de una buena variable instrumental y corregir el carácter endógeno de la educación sobre los salarios.

Otra variable instrumental para medir los retornos de la educación en los salarios es “el mes de nacimiento del individuo” propuesto por Angrist y Krueger (1991); es fácil identificar que el mes de nacimiento del individuo no está correlacionado con su habilidad; es decir, la variable cumple con la condición de exogeneidad; más complicado es imaginar que el mes de nacimiento puede explicar el nivel de educación; sin embargo, en Estados Unidos el mes de nacimiento determina la edad de primer ingreso a la escuela, por lo que la variable puede afectar los años de escolaridad. Entre las variables propuestas para instrumentar la educación en otros estudios destaca el nivel de educación de los padres cuando se tiene información retrospectiva, como la de la EMOVI en México, o el número de hermanos del entrevistado; Card (2001) brinda una extensa revisión de los retornos de la escolaridad y los instrumentos asociados a su medición.

La discusión acerca de los posibles instrumentos y la satisfacción de sus condiciones para una variable endógena es una parada inevitable; si bien la condición de relevancia de la variable instrumental es verificable dada la significancia del instrumento z en la explicación del regresor x , la condición de exclusión no se puede probar de forma directa, puesto que involucra la no correlación de z con la parte no observada de y . En términos estadísticos se han desarrollado herramientas de apoyo para la verificación de estas condiciones. Por ejemplo, la prueba de identificación de instrumentos débiles para la consideración de relevancia y la prueba de sobreidentificación en referencia a la condición de exclusión; sin embargo, en el caso de la prueba de la condición de exclusión se necesita que el número de variables instrumentales sea mayor al número de regresores endógenos.⁸

⁸ En referencias a estas pruebas y el tema VI en general se pueden consultar Cameron y Trivedi (2005, Cap. 4), Wooldridge (2002, Cap. 5) para un desarrollo intermedio y Wooldridge (2000, Cap. 15) para una forma introductoria del tema.

El estimador $\hat{\beta}_{VI}$ de variables instrumentales que identifica el efecto causal del regresor endógeno x sobre la variable dependiente y por medio del instrumento z , es producto de la relación entre el cambio marginal en la variable dependiente ante el cambio de la variable instrumental y el cambio marginal del regresor endógeno ante el cambio de la variable instrumental, así:

$$\hat{\beta}_{VI} = \frac{dy/dz}{dx/dz} \quad (5) \text{ Forma del estimador de VI}$$

Si usamos la definición en (3) de estimadores MCO bajo los supuestos del modelo de regresión lineal, para calcular los efectos marginales dy/dz , dx/dz , se puede escribir el estimador de variables instrumentales de la siguiente manera:

$$\hat{\beta}_{VI} = \frac{(z'z)^{-1}z'y}{(z'z)^{-1}z'x} \quad (6) \text{ Derivación estimador VI}$$

$$\hat{\beta}_{VI} = (z'x)^{-1}z'y \quad (7) \text{ Estimador VI}$$

Una manera común de obtener una estimación $\hat{\beta}_{VI}$ es usar mínimos cuadrados en dos etapas (MC2E); este procedimiento consiste en estimar por MCO una regresión de la variable endógena x sobre su instrumento z en una primera etapa; y en la segunda etapa, se estima una regresión de la variable dependiente y sobre el valor predicho de x proveniente de la primera etapa. Sea el modelo $y = \alpha + \beta x + \varepsilon$ con x endógeno y z variable instrumental de x entonces:

$$x = \theta + \delta z + u \quad (8) \text{ Primera etapa MC2E}$$

$$y = \alpha + \beta_{VI} \hat{x} + \varepsilon \quad (9) \text{ Segunda etapa MC2E}$$

Donde \hat{x} de la ecuación (9) es la predicción que resulta de estimar la primera etapa descrita en (8) y no está correlacionada con ε . Así el parámetro β_{VI} de la ecuación (9) es estimador de variables instrumentales definido en (7) que recoge el efecto causal de x sobre y .

El método de variables instrumentales para la identificación causal es ampliamente usado en microeconometría ya que su lógica principal se puede reproducir más allá de la estimación de modelos lineales (Cameron y Trivedi, 2005; Angrist y Krueger, 2001; Angrist *et al.*, 1996; Maddala, 1983); en el caso de variables dependientes discretas se llevan a cabo estimaciones por variables instrumentales mediante los modelos *IVPROBIT* o en el caso de variables dependientes censuradas vía los modelos *IVTOBIT*. El tratamiento de variables instrumentales también

es adoptado en el caso de los modelos de selección muestral; es decir, cuando el comportamiento de la variable dependiente de interés está precondicionado por una decisión o circunstancia no observada en los datos, la lógica de VI permite la corrección del sesgo.

Supóngase, por ejemplo, que se quieren identificar los determinantes del gasto de los individuos que tienen tarjetas de crédito; rápidamente se puede ver que la tenencia de tarjetas de crédito no es un evento aleatorio: es una decisión endógena asociada con factores no observables como las expectativas económicas y otras características de los individuos. En este contexto, la medición de los determinantes del gasto con tarjetas de crédito tendría un sesgo asociado con la decisión de tener o no el servicio financiero y con la omisión de aquella parte de la muestra de individuos que no tienen tarjeta de crédito. La estimación de forma directa de los determinantes del gasto sin tomar en cuenta la endogeneidad de la decisión de tener el servicio financiero, distorsiona la inferencia causal.

Para corregir este sesgo de selección muestral, Heckman (1979) sugiere incluir en la estimación de los determinantes del gasto un factor de corrección conocido como el inverso de la razón de Mills. La solución al problema proviene de modelar en una primera etapa la decisión de tener el servicio financiero usando una variable instrumental; la variable debe ser buena explicando la decisión de tener tarjeta de crédito (*condición de relevancia*), pero no puede afectar de forma directa el monto del gasto realizado con el servicio financiero (*condición de exclusión*); por ejemplo, la proximidad geográfica a los bancos o el acceso a educación financiera. Con la predicción que resulta de la estimación de la primera etapa por medio de un modelo *Probit*, se calcula el inverso de la razón de Mills (factor de corrección) para incluirlo como variable explicativa en la estimación de los determinantes del gasto con tarjeta de crédito en la segunda etapa.

No son pocos los fenómenos económicos que tienen un comportamiento similar al que se representó con el ejemplo de selección o autoselección muestral del gasto con tarjetas de crédito; variables como el monto de ahorro, el monto de deuda, el consumo de ciertos bienes, la oferta laboral, los salarios, entre otras, al igual que el gasto con tarjeta de crédito están asociados con decisiones endógenas o precondiciones no observadas que generan sesgos en la estimación de sus efectos determinantes. Como se mostró en el ejemplo, la lógica de variables instrumentales para estimar modelos de selección, entre otros modelos, resulta fundamental en la corrección del sesgo para la inferencia causal.

3. Usos y alcances del análisis microeconómico

Una de las aplicaciones económicas en gran apogeo es el uso de las técnicas de medición y modelaje microeconómico para la evaluación de políticas públicas (Angrist y Pischke, 2009). Los actuales ejercicios de gobierno anclados en ejes de transparencia y presupuestos basados en resultados requieren gran apoyo de la economía y de técnicas de medición específicas que permitan observar los impactos de sus programas y políticas desde los microdatos (Moreno-Torres, 2013; Heckman, 2000). La evaluación de programas públicos con fines de mejora en el diseño y ajustes presupuestales es el mejor ejemplo de la imperiosa necesidad de identificar efectos causales para incidir de manera controlada sobre las variables de interés.

La relevancia y particularidad de la información disponible para el análisis microeconómico no es excepción en el ámbito de evaluación de política pública. A pesar de que las evaluaciones de impacto con diseños experimentales conducen de manera más rápida y segura a la identificación de efectos causales del tratamiento sobre el resultado de interés, dichos estudios que garantizan la aleatoriedad en la definición de grupos de tratamiento y control son costosos en los procesos de campo y pueden conducir a problemas éticos que limitan su implementación (Moreno-Torres, 2013; Heckman, 2000). En este marco, los microdatos observacionales y las técnicas microeconómicas toman relevancia en la búsqueda de identificación causal a menores costos y responsabilidades éticas.

Los modelos de regresión discontinua, de diferencias en diferencias, panel de efectos fijos, técnicas de pareo por puntaje (*propensity score matching*), entre otros, ofrecen una gran variedad de opciones ante la disponibilidad de datos observacionales y evaluaciones con diseños *ex post* a la implementación del programa público a evaluar. La combinación de técnicas y modelos para atender las problemáticas específicas en el ámbito de evaluación es parte de la estructura y rasgos de muchos estudios en esta rama (Moreno-Torres, 2013).

3.1 Referencias al análisis microeconómico en México

México no ha sido la excepción en dar muestras del auge de las aplicaciones microeconómicas, tanto en la evaluación de política pública como en el análisis general de relaciones económicas a partir del uso de microdatos; en la producción científica del país se observa desde el modelaje paramétrico con datos de corte transversal o panel para la identificación y aislamiento de relaciones causales, así como el uso del modelaje no lineal, de variables discretas, censuradas, hasta el

modelaje no-paramétrico y semi-paramétrico de la teoría asintótica. Si bien, como mencionan Díaz y Martínez (2006) y Moreno-Torres (2013) el apogeo del uso de las técnicas microeconómicas a nivel mundial se vive desde la década de los años ochenta, su aplicación masiva a nivel nacional está asociada con la disponibilidad y el libre acceso a las fuentes de información adecuadas como se mencionó en el primer apartado de este trabajo.

De forma particular el Cuadro 2, resume algunos ejemplos recientes de las aplicaciones microeconómicas en el contexto mexicano. No han sido pocos los esfuerzos científicos en los que se ocupa alguna técnica microeconómica para analizar temas como la evaluación del sistema de pensiones, del mercado de vivienda, el estudio de los determinantes de la pobreza y la desigualdad, de la discriminación salarial y el mercado laboral, o para evaluación de políticas públicas particulares como el Seguro Popular o el Programa Oportunidades entre otros. El Cuadro 2 agrupa estos ejercicios empíricos por el tipo de modelaje usado en la identificación: *i*) evaluación paramétrica y uso de datos panel; *ii*) modelos no lineales, de variable censurada y autoselección; y *iii*) modelos no-paramétricos y semiparamétricos. Asimismo, el cuadro hace referencia a la técnica usada en cada caso dando un panorama de la gran variedad metodológica ocupada según las necesidades de análisis y pasando de forma transversal por el objetivo de evaluación de política pública y de análisis relaciones económicas y del comportamiento de los agentes.

Cuadro 2
Ejemplos de algunas aplicaciones microeconómicas en México

<i>Título</i>	<i>Año</i>	<i>Técnica</i>	<i>Objetivo del estudio</i>	<i>Publicación: Autor(es)</i>
<i>Modelos de evaluación paramétrica y con datos panel</i>				
México: impacto de la educación en la pobreza rural.	2009	Propensity Score Matching	Se evalúa el impacto de la educación primaria y secundaria sobre la pobreza alimentaria, de capacidades, y patrimonial del sector rural.	<i>Serie estudios y perspectivas:</i> Ordaz, J.
Desarrollo e igualdad de oportunidades en México: una evaluación por cuantiles del Programa Oportunidades	2015	Regresión Cuantílica	Se evalúa el impacto del Programa Oportunidades en la reducción de la desigualdad.	<i>R. legislativa de estudios sociales:</i> Barrón, K., Gómez, C., Orozco, P.

Continúa...

<i>Título</i>	<i>Año</i>	<i>Técnica</i>	<i>Objetivo del estudio</i>	<i>Publicación: Autor(es)</i>
Efectos de las becas educativas del programa Oportunidades sobre la asistencia escolar. El caso de la zona urbana del noreste de México.	2012	Propensity Score Matching	Se evalúa el impacto del Programa Oportunidades en la disminución de la inasistencia escolar.	<i>Revista Desarrollo y Sociedad:</i> Martínez, O.
Impacto del Seguro Popular en el gasto catastrófico y de bolsillo en el México rural y urbano, 2005-2008.	2011	VARIABLES INSTRUMENTALES	Estimar el efecto del Seguro Popular sobre la incidencia del gasto catastrófico en salud y sobre el gasto de bolsillo en salud en el mediano plazo.	<i>Salud Pública de México:</i> Sosa, S., Salinas, A., Galaraga, O.
Determinantes del gasto estatal en México.	2006	Modelos de efectos fijos y aleatorios	Estimar los determinantes del gasto público viendo los impactos de las variables políticas y participaciones federales.	<i>Gestión y Política Pública:</i> Velázquez, C.
Remesas, crecimiento y convergencia regional en México: aproximación con un modelo panel-espacial.	2016	Modelos de efectos fijos y aleatorios	Estimar el efecto de las remesas sobre el crecimiento regional	<i>Estudios Económico:</i> Mendoza, M., Valdivia, M.
<i>Modelos no lineales y de variables censuradas</i>				
Estimación de la demanda de uso de autos particulares en la zona metropolitana del Valle de México: un análisis Tobit	2017	Modelo Tobit	Analizar el efecto del ingreso, de variables socio-económicas y geográficas sobre la demanda de uso de autos particulares.	<i>Panorama económico:</i> Perez, R.
El mercado de vivienda en México: un análisis de toma de decisiones basado en la microeconomía.	2006	Modelos Logit y Probit	Estimar los determinantes de la toma de decisiones en el mercado de vivienda.	<i>Denarius:</i> Díaz, J., Martínez, G.
Cobertura de los sistemas de pensiones y factores asociados al acceso a una pensión de jubilación en México.	2011	Modelo Logit	Identificar la cobertura y los determinantes y de los sistemas de pensiones.	<i>Papeles de población:</i> Murillo, S., Venegas, F.
La importancia de corregir por el sesgo de selección en el análisis de las brechas salariales por género: un estudio para Argentina, Brasil y México.	2012	Modelo Heckit	Estimar determinantes de las brechas salariales por género controlando por la decisión de participación laboral.	<i>Revista de Economía:</i> Rubil, A.

Continúa...

<i>Título</i>	<i>Año</i>	<i>Técnica</i>	<i>Objetivo del estudio</i>	<i>Publicación: Autor(es)</i>
<i>Modelos No-paramétricos y Semi-paramétricos</i>				
Salarios y calificación laboral en México	2008	Modelo no-paramétrico	Analizar los salarios en el mercado laboral de las regiones en México por trabajo calificado y no calificado entre los años 2000 y 2004.	<i>Problemas del Desarrollo</i> : Huesca, L., Rodríguez, E.
Evolución de la brecha salarial de género en México	2014	Modelo Semi-paramétrico	Analizar la brecha salarial de género en México entre 1990 y 2010 utilizando los censos de población	<i>El Trimestre Económico</i> : Arceo, E., Campos, R.
El uso de internet y su relación con los salarios en México: un análisis no paramétrico	2017	Modelo no-paramétrico	Caracterizar la relación entre los ingresos salariales y el uso de internet en el trabajo	<i>Los retos del cambio económico actual</i> : Torres, J., Félix, G.
Análisis de transición dinámica: un enfoque no paramétrico aplicado a la región centro de México (1988-2003).	2011	Modelo no-paramétrico	Analizar la dinámica de la distribución del ingreso de la región centro del país a escala municipal.	<i>Problemas del Desarrollo</i> : Borrayo, R., Castañeda, J.

Fuente: elaboración propia con información de las revistas enunciadas en el cuadro.

En este ejercicio, por otro lado se debe reconocer que, las técnicas de medición poseen importantes límites asociados con el contexto de aplicación, la disponibilidad de información y el objetivo mismo para el que se hayan elegido; el reconocimiento de tales alcances es forzoso para un uso responsable. Desde LaLonde (1986) se ha mostrado que la aplicación de diferentes técnicas viables en un mismo problema y conjunto de datos puede conducir a distintos resultados, por lo que argumentar la elección del método es un ejercicio primordial en los estudios aplicados. En muchos casos las diferencias en los resultados de evaluación de una misma política son atribuibles a diferencias en la especificación de variables dados los supuestos propios de la herramienta seleccionada; en general, las particularidades de cada técnica conducen a una lógica de resultados, pero lo realmente importante es la justificación de su elección dada las particularidades de la problemática que se quiere cuantificar.

Un punto adicional a tener en cuenta en los alcances del análisis micro-econométrico es la visión objetiva e imparcial que debería poseer el investigador social que aplica la técnica e interpreta los resultados. En las ciencias sociales la

visión del observador es un tema muy discutido; en este caso se debe considerar que los números por sí solos no hablan, más bien sugieren circunstancias, relaciones, efectos, etc. Sin la adecuada interpretación o lectura de la evidencia sugerida por los resultados, el uso de la técnica por más sofisticada que sea es inefectiva. Consecuentemente, los resultados del modelaje microeconómico se interpretan y lucen mejor a la luz de la evidencia que se genera con la estadística y el análisis descriptivo.

Conclusiones

Muchas son las posibilidades que resultan de las variaciones y combinaciones de las técnicas de medición para la inferencia causal ante la necesidad de cada problema que se enfrenta en el ámbito de los microdatos de corte transversal y tipo panel. No fue objeto de este ensayo realizar una lista exhaustiva de las técnicas de medición y de modelaje; sin embargo, se identificaron los rasgos distintivos del análisis microeconómico, se expusieron los elementos básicos de un punto de partida asentado en el modelo de regresión, la relevancia de los supuestos de identificación, el uso de variables instrumentales y, finalmente, se hizo una breve mención de usos y alcances de la técnica microeconómica en la economía aplicada y su ejercicio en México.

Se caracterizó la microeconometría como aquella rama de la econometría aplicada que se apoya en técnicas específicas para la identificación causal -como variables instrumentales, modelos de selección, entre otros-, en el uso de técnicas del modelaje no lineal -como modelos de variables dependientes discretas y censuradas-, en la reducción y flexibilización de hipótesis distribucionales -como los modelos semi y no paramétricos-, entre otros rasgos muy particulares para la medición del comportamiento microeconómico usando datos de corte transversal y datos panel. Se identificó el importante papel de la disponibilidad de las fuentes de información para el análisis y se destacó la gran cantidad de datos observacionales de libre acceso en el país.

Este ensayo expuso la relevancia de los supuestos de identificación. Específicamente en el caso de la violación del supuesto de no correlación endógena entre la perturbación aleatoria y la variable explicativa de interés, se revisó la lógica de variables instrumentales como una solución muy recurrente en los estudios microeconómicos; se mostró que, a pesar de que VI es una técnica bastante antigua (surge poco después de los inicios de la econometría misma), es un método de estimación que continúa ocupando un lugar especial en los problemas de inferencia causal. Asimismo, se mostró que toda técnica -incluyendo VI y sus condiciones para

un buen instrumento- posee un espacio abierto para la discusión de su adecuada implementación, uso de supuestos y propiedades.

Este trabajo también hizo referencia al ejercicio de la técnica microeconómica en la evaluación de la política pública. En un marco de administración pública más controlado, las mediciones causales de esta rama son la mejor opción para la toma de decisiones que logren incidencia y control mismo sobre la política social. Se destacó que se deben medir los alcances de la técnica para no hacer un uso indiscriminado de ella; asimismo se hace referencia al importante acompañamiento del análisis descriptivo que motiva y complementa la lectura de resultados del análisis de inferencia causal.

Finalmente, no quisiera cerrar este ensayo sin mencionar que, en el gran reto de enseñanza y acercamiento al uso de métodos cuantitativos en nuestro país, estoy convencido de que la mejor estrategia una vez identificado el objetivo correcto, es lograr el dominio y la familiaridad con las técnicas más básicas para que cada usuario pueda adaptarlas, combinarlas e incluso mejorarlas según sus necesidades. Por ello, este escrito espera contribuir al acercamiento de las técnicas de medición y ejercicio microeconómico, mediante la caracterización de sus rasgos distintivos asociados con el objetivo de inferencia causal, el establecimiento de los elementos de partida para la aplicación de técnicas, la descripción de sus alcances y el énfasis sobre el hecho de que la estrategia de medición es sólo una parte del problema de investigación.

Referencias

- Angrist, J., G. Imbens and D. Rubin (1996), "Identification of Causal Effects Using Instrumental Variables", *Journal of the American Statistical Association*, 434 (91), pp. 444-455. DOI: 10.1080/01621459.1996.10476902
- Angrist, J. and A. Krueger (2001), "Instrumental Variables and the Search for Identification: From Supply and Demand to Natural Experiments", *Journal of Economic Perspectives*, 15(4), pp. 69-85. DOI: 10.1257/jep.15.4.69
- Angrist, J. and A. Krueger (1991), "Does Compulsory School Attendance Affect Schooling and Earnings?", *The Quarterly Journal of Economics*, 106(4), pp. 979-1014. <https://doi.org/10.2307/2937954>
- Angrist J. and J. Pischke (2009), *Mostly Harmless Econometrics. An Empiricist's Companion*, Princeton, N. J.: Princeton University Press.
- Arceo, E. y R. Campos (2014), "Evolución de la brecha salarial de género en México", *El trimestre económico*, 81 (323), pp. 619-653. DOI: <http://dx.doi.org/10.20430/ete.v81i323.125>

- Borraro, R. y J. Castañeda (2011), “Análisis de transición dinámica: un enfoque no paramétrico aplicado a la región centro de México (1988-2003)”, *Problemas del desarrollo*, 42(167), pp. 7-44. DOI: 10.22201/iiec.20078951e.2011.167.27718
- Barrón, K., Gómez, C., y P. Orozco (2015), “Desarrollo e igualdad de oportunidades en México: una evaluación por cuantiles del Programa Oportunidades”, *Revista legislativa de estudios sociales y de opinión pública*, 8(15), pp. 107-146.
- Cameron C. and K. Trivedi (2005), *Microeconometrics. Methods and Applications*, Cambridge: Cambridge University Press.
- Card, D. (1993), “Using Geographic Variation in College Proximity to Estimate the Return to Schooling”, *NBER Working Papers 4483*, National Bureau of Economic Research.
- Card, D. (2001), “Estimating the Return to Schooling: Progress on Some Persistent Econometric Problems”, *Econometrica*, 5(69), pp. 1127-1160. <https://doi.org/10.1111/1468-0262.00237>
- Díaz, J. y G. Martínez (2006), “El mercado de vivienda en México: un análisis de toma de decisiones basado en la microeconomía”, *Denarius*, 10 (1), pp. 155-176. Recuperado de <http://tesiuami.uam.mx/revistasuam/denarius/include/getdoc.php?id=321&article=120&mode=pdf>.
- Enders, W. (2004), *Applied Econometric Time Series*, 2nd Edition, Wiley.
- Fernández, J. y C. Adalid (2000), “Para una breve historia de la econometría”, *Política y Cultura*, 13, pp. 7-32. Recuperado de <https://www.redalyc.org/articulo.oa?id=26701302>
- Frisch, R. (1933), “Editorial”, *Econometrica*, 1, pp. 1-4.
- Goldberger, A. (1972), “Structural Equation Methods in the Social Sciences”, *Econometrica*, 40, pp. 979-1001. [https://doi.org/0012-9682\(197211\)40:6<979:SEMITS>2.0.CO;2-9](https://doi.org/0012-9682(197211)40:6<979:SEMITS>2.0.CO;2-9)
- Goldberger, A. (1964), *Econometric theory*, Wiley.
- Greene, W. (1999), *Análisis econométrico*, Primera edición, Madrid: Prentice Hall.
- Gujarati, D. y D. Porter (2010). *Econometría*, Quinta edición, México: Mc Graw Hill.
- Haavelmo, T. (1944), “The Probability Approach in Econometrics”, *Supplement to Econometrica*, vol 12, pp. 1-115. <http://www.jstor.org/stable/1906935>
- Haavelmo, T. (1943), “The Statistical Implications of a System of Simultaneous Equations”, *Econometrica*, 11, pp. 1-12. DOI: 10.2307/1905714
- Hausman, J. (2001), “Microeconometrics”, *Journal of Econometrics*, 1(100), pp. 33-35. [https://doi.org/10.1016/S0304-4076\(00\)00050-6](https://doi.org/10.1016/S0304-4076(00)00050-6)
- Heckman, J. (1979), “Sample Selection Bias as a Specification Error”, *Econometrica*, 47(1), pp. 153-161. DOI: 10.2307/1912352 <https://www.jstor.org/stable/1912352>

- Heckman, J. (2000), “Datos Microeconómicos, Heterogeneidad y la Evaluación de Políticas Públicas”. *Fundación Nobel, Revista Asturiana de Economía* (2003), 26, pp. 211-295.
- Huesca L. y E. Rodríguez (2008), “Salarios y calificación laboral en México”, *Problemas del desarrollo*. 39 (154), pp. 61-86. DOI: 10.22201/iiec.20078951e.2008.154.7727
- Ibarra, A. (1998), “Historia cuantitativa, serial y cliometría: una apreciación general y de su impacto en la historia mexicana reciente”, *Investigación Económica*, 224(58), pp. 119-135.
- LaLonde, R. (1986), “Evaluating the Econometric Evaluations of Training Programs with Experimental Data”, *American Economic Review*, 76(4), pp. 604-20. Recuperado en <http://links.jstor.org/sici?sici=0002-8282%28198609%2976%3A4%3C604%3AETEEOT%3E2.0.CO%3B2-P>
- Maddala, G. (1983), *Limited dependent qualitative variables in econometrics*, Cambridge: Cambridge University Press.
- Martínez, O. (2012), “Efectos de las becas educativas del programa Oportunidades sobre la asistencia escolar. El caso de la zona urbana del noreste de México”, *Revista Desarrollo y Sociedad*, (69), pp. 99-131. <https://doi.org/10.13043/dys.69.4>
- Mendoza, M. y M. Valdivia (2016), “Remesas, crecimiento y convergencia regional en México: aproximación con un modelo panel-espacial”, *Estudios Económicos*, 31(1), pp. 125-167. Recuperado de <https://EconPapers.repec.org/RePEc:emx:esteco:v:31:y:2016:i:1:p:125-167>
- Moreno-Torres, A. (2013), “Técnicas microeconómicas para la evaluación de políticas públicas”, *Economía Industrial*, 390, pp. 33-54. Recuperado de: <https://www.mincotur.gob.es/Publicaciones/Publicacionesperiodicas/EconomiaIndustrial/RevistaEconomiaIndustrial/390/ANTONIO%20MORENO%20TORRES.pdf>
- Mroz, T. (1987), “The Sensitivity of an Empirical Model of Married Women’s Hours of Work to Economic and Statistical Assumptions”, *Econometrica*, 4(55), pp. 765-799. DOI: 10.2307/1911029 <https://www.jstor.org/stable/1911029>.
- Murillo, S. y F. Venegas (2011), “Cobertura de los sistemas de pensiones y factores asociados al acceso a una pensión de jubilación en México”, *Papeles de población*, (67), pp. 209-250. Recuperado de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S1405-74252011000100008&lng=es&tlng=es.
- Ordaz, J. (2009), “México: impacto de la educación en la pobreza rural”, *Serie estudios y perspectivas*, 1(105), pp.5-40.
- Pérez, R. (2017), “Estimación de la demanda de uso de autos particulares en la zona metropolitana del valle de México: un análisis *tobit*”, *Panorama económico*, 11 (21), pp. 1-37.

- Pulido, A. y J. Pérez (2006), “Lawrence R. Klein y la economía aplicada”, *Estudios de Economía Aplicada*, 1(24), pp. 43-94. Recuperado en <http://www.revista-eea.net/documentos/24120.pdf>.
- Rubil, A. (2012), “La importancia de corregir por el sesgo de selección en el análisis de las brechas salariales por género: un estudio para Argentina, Brasil y México”, *Ensayos Revista de Economía*, XXXI (2), pp. 1- 36.
- Ruiz, A. (2016), “La investigación econométrica mediante paneles de datos: historia, modelos y usos en México”, *Economía y Política*, 24(12), pp.12-32. Recuperado en [Dialnet-LaInvestigacionEconometricaMediantePanelesDeDatos-6166187.pdf](http://dialnet-lainvestigacioneconometricamediantepanelesdedatos-6166187.pdf)
- Smith, V. (2005), “¿Qué es la economía experimental?”, *Revista Apuntes del CENES*, 39(25), pp. 7-16. Recuperado de <https://www.redalyc.org/pdf/4795/479548746002.pdf>
- Sosa, S., Salinas, A. y O. Galarraga (2011), “Impacto del Seguro Popular en el gasto catastrófico y de bolsillo en el México rural y urbano, 2005-2008”, *Salud Pública de México*, 53, pp. 425-435. Recuperado de http://www.scielo.org.mx/scielo.php?script=sci_arttext&pid=S0036-36342011001000006&lng=es&tlng=es.
- Theil, H. (1971), *Principles of Econometrics*, New York: Wiley.
- Torres, J. y G. Félix (2017), “El uso de internet y su relación con los salarios en México: un análisis no paramétrico”, *Los retos del cambio económico actual: revisión y aplicaciones para el caso mexicano*, pp. 33-50.
- Velázquez, C. (2006), “Determinantes del gasto estatal en México”, *Gestión y Política Pública*, XV (1), pp. 83-109.
- Vicéns, J. (1998), “Econometría y contrastación empírica. Concepto e historia”, *Documento de trabajo No. 98/1*, Universidad Autónoma de Madrid.
- Wooldrige, J. (2000), *Introducción a la econometría. Un enfoque moderno*. México: Thomson Learning.
- Wooldrige, J. (2002), *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT press.
- Wright, S. (1928), “A New Design for Randomized Clinical Trial”, *New England Journal of Medicine*, 300, pp. 1242-1245.
- Wright, S. (1934), “The Method of Path Coefficients”, *Annals of Mathematical Statistics*, 5, pp. 161-215. DOI: 10.1214/aoms/1177732676