

Eficiencia de los indicadores tipo PMI publicados por INEGI e IMEF

Gerardo Leyva y Olinca Páez

Índice Ponderado de Consumo de Electrodomésticos: una propuesta de medición a partir de datos de encuestas en hogares en México

Ana Ruth Escoto Castillo y Landy Sánchez Peña

Determinantes subnacionales de la informalidad laboral en México

Benjamín Temkin y Jorge Cruz Ibarra

Obteniendo indicadores de actividad económica municipal basados en información representativa de los Censos Económicos

Francisco de Jesús Corona Villavicencio y Jesús López-Pérez

Vinculación longitudinal de los Censos Económicos 1994-2014 de México

Matías Busso, Óscar Eduardo Fentanes Téllez y Santiago Levy Algazi

Solo se puede actuar sobre aquello que se mide

Reseña

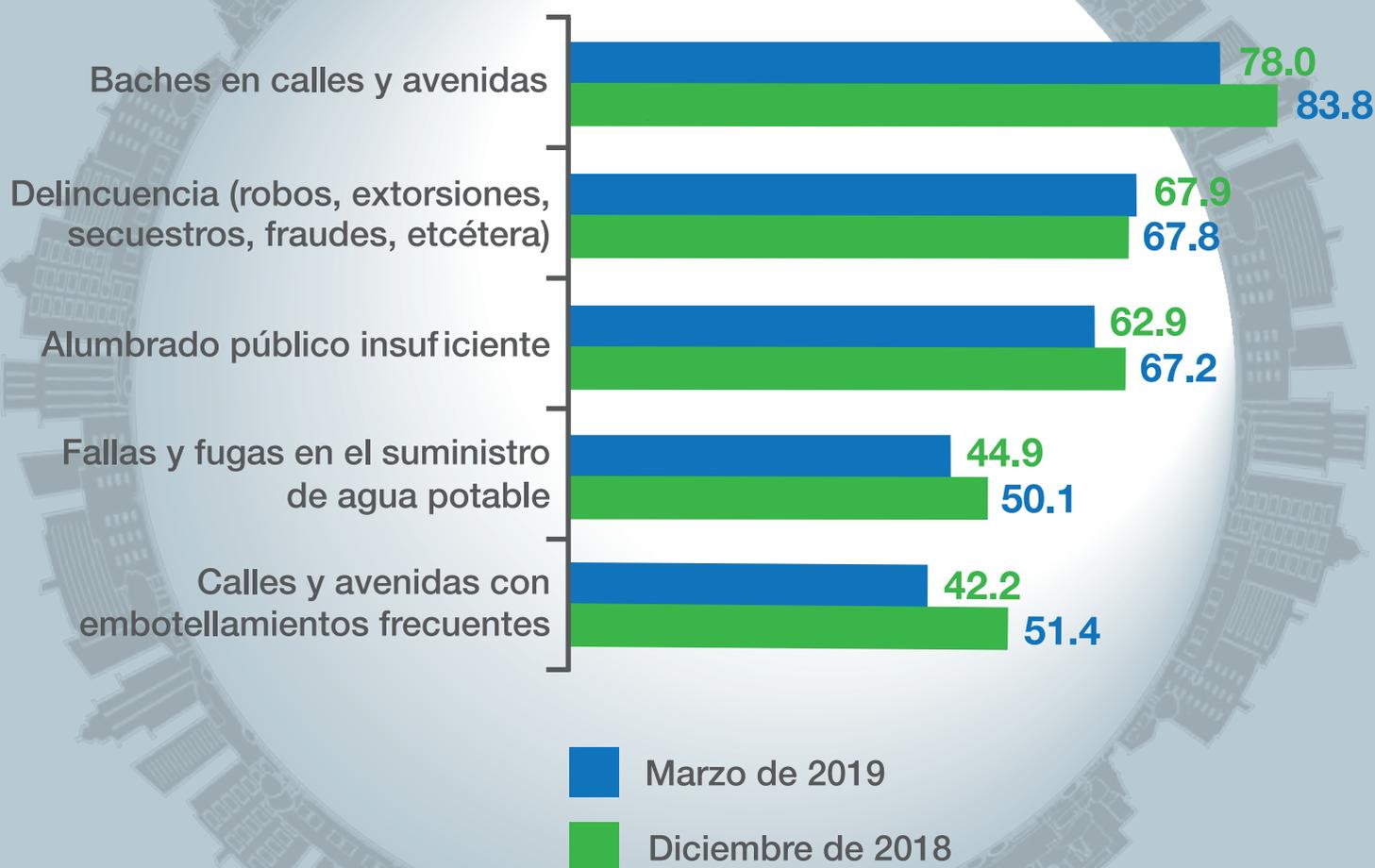
Adriana Pérez Amador

Por un logro sostenido en el tiempo de capacidades estadísticas

Reseña

Ximena A. Clark

Principales problemáticas en las zonas urbanas (porcentajes)



Nota: percepción de la población de 18 años y más de edad; excluye la respuesta *No sabe o no responde*.

Fuente: INEGI. *Encuesta Nacional de Seguridad Pública Urbana*.

Conociendo México

01 800 111 46 34

www.inegi.org.mx

atencion.usuarios@inegi.org.mx

f i t v INEGI Informa

INEGI

Contenido

Eficiencia de los indicadores tipo PMI publicados por INEGI e IMEF <i>Efficiency of PMI Indicators Published by INEGI and IMEF</i> Gerardo Leyva y Olinca Páez	4
Índice Ponderado de Consumo de Electrodomésticos: una propuesta de medición a partir de datos de encuestas en hogares en México <i>Weighted Index of Household Appliances Consumption: a Proposal Based on Household Survey Data in Mexico</i> Ana Ruth Escoto Castillo y Landy Sánchez Peña	26
Determinantes subnacionales de la informalidad laboral en México <i>Subnational Determinants of Labor Informality in Mexico</i> Benjamín Temkin y Jorge Cruz Ibarra	46
Obteniendo indicadores de actividad económica municipal basados en información representativa de los Censos Económicos <i>Obtaining Municipality-Level Economic Activity Indicators Based on Representative Economic Censuses Information</i> Francisco de Jesús Corona Villavicencio y Jesús López-Pérez	62
Vinculación longitudinal de los Censos Económicos 1994-2014 de México <i>The Longitudinal Linkage of Mexico's Economic Census 1994-2014</i> Matías Busso, Óscar Eduardo Fentanes Téllez y Santiago Levy Algazi	82
Solo se puede actuar sobre aquello que se mide <i>If you don't measure it, you can't act on it</i> Reseña Adriana Pérez Amador	100
Por un logro sostenido en el tiempo de capacidades estadísticas <i>For a sustained achievement of statistical capacities</i> Reseña Ximena A. Clark	104
Colaboran en este número	108

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Presidente del Instituto

Julio Alfonso Santaella Castell

Vicepresidentes

Enrique de Alba Guerra

Paloma Merodio Gómez

Enrique Jesús Ordaz López

Adrián Franco Barrios

Dirección General de Estadísticas Sociodemográficas

Edgar Vielma Orozco

Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia

Oscar Jaimes Bello

Dirección General de Estadísticas Económicas

José Arturo Blancas Espejo

Dirección General de Geografía y Medio Ambiente

María del Carmen Reyes Guerrero

Dirección General de Integración, Análisis e Investigación

Sergio Carrera Riva Palacio

Dirección General de Coordinación del Sistema Nacional de Información Estadística y Geográfica

María Isabel Monterrubio Gómez

Dirección General de Vinculación y Servicio Público de Información

Eduardo Javier Gracida Campos

Dirección General de Administración

Marcos Benerice González Tejeda

Contraloría Interna

Manuel Rodríguez Murillo

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Editor responsable

Enrique Jesús Ordaz López

Editor técnico

Gerardo Leyva Parra

Coordinación editorial

Virginia Abrín Batule y Mercedes Pedrosa Islas

Corrección de estilo

José Pablo Covarrubias Ordiales y Laura Elena López Ortiz

Corrección de textos en inglés

Gerardo Piña

Diseño y formación edición impresa

Juan Carlos Martínez Méndez y Eduardo Javier Ramírez Espino

Indizada en: Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal *Latindex Catálogo*; Citas Latinoamericanas en Ciencias Sociales y Humanidades (*CLASE*) y en la Red Iberoamericana de Innovación y Conocimiento (*REDIB*).

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, Vol. 10, Núm. 2, mayo-agosto, 2019, es una publicación cuatrimestral editada por el Instituto Nacional de Estadística y Geografía, Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276 Aguascalientes, Aguascalientes, entre la calle INEGI, Avenida del Lago y Avenida Paseo de las Garzas, México. Teléfono 55 52781069. Toda correspondencia deberá dirigirse al correo: rde@inegi.org.mx

Editor responsable: Enrique Jesús Ordaz López. Reserva de Derechos al Uso Exclusivo del Título Núm. 04-2012-121909394300-102, ISSN Núm. 2007-2961, ambos otorgados por el Instituto Nacional del Derecho de Autor. Certificado de Licitud de Título y Contenido Núm. 15099, otorgado por la Comisión Calificadora de Publicaciones y Revistas Ilustradas de la Secretaría de Gobernación. Domicilio de la publicación, imprenta y distribución: Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276 Aguascalientes, Aguascalientes, entre la calle INEGI, Avenida del Lago y Avenida Paseo de las Garzas, México.

El contenido de los artículos, así como sus títulos y, en su caso, fotografías y gráficos utilizados son responsabilidad del autor, lo cual no refleja necesariamente el criterio editorial institucional. Asimismo, la Revista se reserva el derecho de modificar los títulos de los artículos, previo acuerdo con los autores. La mención de empresas o productos específicos en las páginas de la Revista no implica el respaldo por el Instituto Nacional de Estadística y Geografía.

Se permite la reproducción total o parcial del material incluido en la Revista, sujeto a citar la fuente. Esta publicación consta de 400 ejemplares y se terminó de imprimir en enero del 2019.

Versión electrónica: <http://rde.inegi.org.mx>

ISSN 2395-8537

CONSEJO EDITORIAL

Enrique de Alba Guerra

Presidente del Consejo

Fernando Cortés Cáceres

Profesor Emérito de FLACSO

PUED de la UNAM

México

Gerardo Bocco Verdinelli

Universidad Nacional Autónoma de México

México

Juan Carlos Chávez Martín del Campo

Banco de México

México

Lidia Bratanova

UNECE Statistical Division

Switzerland

Tonatiuh Guillén López

Instituto Nacional de Migración

México

Víctor Manuel Guerrero Guzmán

Instituto Tecnológico Autónomo de México

México

En el primer artículo de esta edición, *Eficiencia de los indicadores tipo PMI publicados por INEGI e IMEF (Efficiency of PMI Indicators Published by INEGI and IMEF)*, los autores analizaron variables relacionadas con la producción a diferentes escalas y seleccionaron al Indicador Global de la Actividad Económica para evaluar, en su función de anticipar la actividad económica agregada en México, la eficiencia tanto del Indicador de Pedidos Manufactureros —generado por el Instituto Nacional de Estadística y Geografía (INEGI) y el Banco de México— como de los desarrollados por el Instituto Mexicano de Ejecutivos de Finanzas (IMEF) en sus versiones de ponderación igualitaria (IMEF-IG) y la ajustada por tamaño de empresa (IMEF-AT).

Desde la perspectiva de la promoción de consumo energético eficiente y responsable en los hogares, *Índice Ponderado de Consumo de Electrodomésticos: una propuesta de medición a partir de datos de encuestas en hogares en México (Weighted Index of Household Appliances Consumption: a Proposal Based on Household Survey Data in Mexico)* es un artículo donde se estudia la posesión de equipamiento de este tipo y sugiere la necesidad de incorporar medidas de estos bienes que den cuenta con dicho índice de sus usos diferenciados.

Posteriormente, *Determinantes subnacionales de la informalidad laboral en México (Subnational Determinants of Labor Informality in Mexico)* hace un análisis de las dinámicas del trabajo informal a nivel entidad federativa a través de una base de datos que considera los 32 estados del país en la que se incluyen como variables la tasa de informalidad, el Producto Interno Bruto per cápita, la Inversión Extranjera Directa, el porcentaje de afiliados al Seguro Popular, la cualificación de la mano de obra y la intensidad de la microempresarialidad, factores que impactan la prevalencia de este fenómeno de ocupación en la economía mexicana.

Más adelante, en el trabajo *Obteniendo indicadores de actividad económica municipal basados en información representativa de los Censos Económicos (Obtaining Municipality-Level Economic Activity Indicators Based on Representative Economic Censuses Information)* se presenta una metodología para desarrollarlos con una temporalidad y nivel sectorial pertinente, por lo cual se utilizaron datos de ese operativo censal del INEGI del

periodo 2003 a 2013; para verificar el error de estimación, se construyeron indicadores estatales con lo aquí propuesto y se compararon con los productos internos brutos de las entidades generados por el Sistema de Cuentas Nacionales de México.

Lo siguiente es la nota técnica *Vinculación longitudinal de los Censos Económicos 1994-2014 de México (The Longitudinal Linkage of Mexico's Economic Census 1994-2014)*, la cual describe la metodología para construir una base de datos longitudinal que se puede resumir en cinco pasos: estandarización, indexación (previnculación), comparación, clasificación de enlaces y validación; para su proceso, se definió un algoritmo con reglas de continuidad que consideran tres factores: ubicación, entidad legal y clase de actividad.

Solo se puede actuar sobre aquello que se mide (If you don't measure it, you can't act on it) es la reseña del libro *No todo el trabajo es empleo. Avances y desafíos en la conceptualización y medición del trabajo en México*, obra coordinada por Mauricio Padrón Innamorato, Luciana Gandini y Emma Liliana Navarrete en la que confluyen 14 investigadores de distintas instituciones nacionales e internacionales, la cual resulta muy pertinente porque logra plasmar diversas inquietudes alrededor de la nueva forma en que serán abordados conceptos como *trabajo, fuerza de trabajo, ocupación y subutilización*.

Finalmente, se presenta también la reseña del libro *¿A quién le importa saber? La economía política de la capacidad estadística en América Latina* bajo el título *Por un logro sostenido en el tiempo de capacidades estadísticas (For a sustained achievement of statistical capacities)*. La obra está basada en el estudio de casos para 10 países de América Latina que responde a una serie de interrogantes acerca del porqué es importante que las naciones produzcan datos de calidad, oportunos y relevantes para que la sociedad tome decisiones sustentadas o cuál es la causa de que algunos países desarrollen un sistema estadístico altamente técnico y coordinado por sus oficinas nacionales de estadística, mientras que otros de características similares no logren tal nivel, entre muchas otras preguntas relacionadas.

<http://rde.inegi.org.mx>

Eficiencia de los indicadores tipo PMI publicados por INEGI e IMEF

Efficiency of PMI Indicators Published by INEGI and IMEF

Gerardo Leyva* y Olinca Páez**

Con la finalidad de contribuir al mejor uso de los indicadores de difusión, en esta investigación evaluamos la capacidad que tienen tres de ellos para anticipar cambios en la actividad económica agregada en México: se trata del Indicador de Pedidos Manufactureros (IPM) que genera el INEGI en colaboración con el Banco de México y de los indicadores manufactureros del IMEF en sus versiones de ponderación igualitaria (IMEF-IG) y ajustada por tamaño de empresa (IMEF-AT).

La técnica de partida es un análisis de correlaciones que después se complementa con la evaluación del éxito con el que un umbral clasifica correctamente los resultados en la variable de interés. La clasificación puede ser exitosa bajo distintos criterios. En este caso, evaluamos según cuatro medidas distintas: la tasa de error, el

With the purpose of contributing to the best use of the diffusion indicators, we evaluate the capacity of three different Purchasing Manager's Indicators produced in Mexico to anticipate changes in the aggregated economy, one generated by INEGI and Banco de México, and two designed by the Mexican Institute of Finance Executives (IMEF by its Spanish acronym).

The starting technique is a correlation analysis that is then complemented by the evaluation of the success with which a threshold correctly classifies the results in the variable of interest. Since classification can be successful under several criteria, we define four different measures to broadly compare the results: error rate, negative predictive value, true negative rate, and Φ coefficient.

* Instituto Nacional de Estadística y Geografía (INEGI), gerardo.leyva@inegi.org.mx

** INEGI, olinca.paez@inegi.org.mx

valor predictivo negativo, la tasa de correctos negativos y el coeficiente Φ .

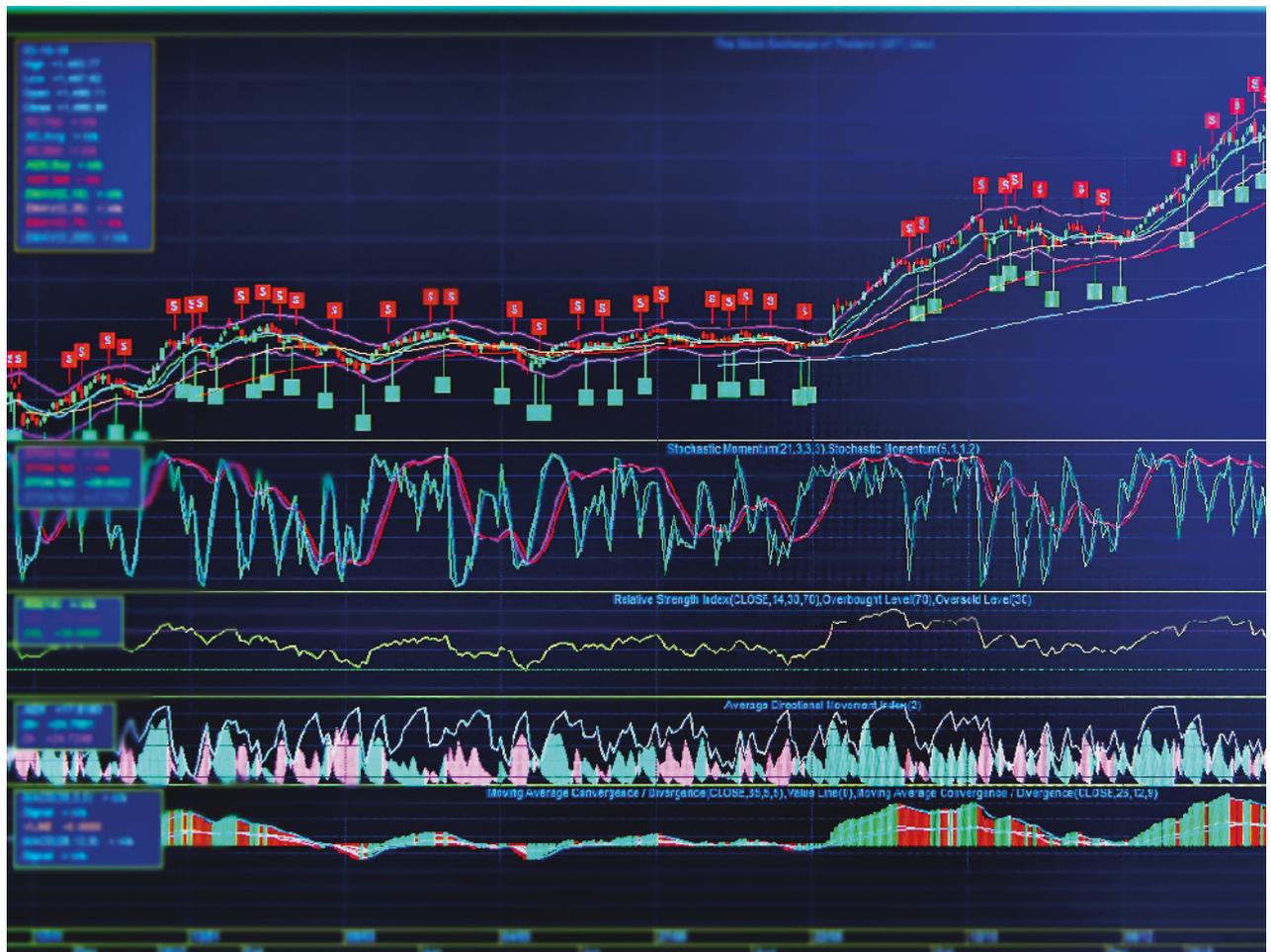
Concluimos que, aunque el indicador IMEF-IG es marginalmente superior en términos de correlaciones, para anticipar el signo de la variación trimestral del Indicador Global de la Actividad Económica (IGAE), el IPM y el IMEF-AT son, prácticamente, igual de eficientes y mejores comparados con el IMEF-IG. Además, demostramos que la capacidad de este último para adelantar variaciones del IGAE mejora con un umbral empírico de 48.5.

Palabras clave: Indicador de Pedidos Manufactureros; PMI; Encuesta Mensual de Opinión Empresarial; Indicador IMEF; IMEF; IGAE; indicadores de difusión; umbral empírico; clasificación; *logit*.

Recibido: 8 de agosto de 2018
Aceptado: 12 de diciembre de 2018

In order to anticipate the sign of the quarterly variation of the Global Indicator of Economic Activity (IGAE by its Spanish acronym), IPM and IMEF-AT are basically equally efficient and better compared to IMEF-IG. Besides, the capacity of the latter for antedating variations of the IGAE is better when an empirical threshold of 48.5 is used as reference.

Keywords: Purchasing Manager's Indicator; IPM; *Encuesta Mensual de Opinión Empresarial*; IMEF; IMEF Indicator; IGAE; diffusion indicators; empirical threshold; classification; *logit*.



Fotografía de stock/Kanok Sultaiman/Getty Images

Introducción

Dos instituciones producen y difunden indicadores de pedidos manufactureros para México: el Instituto Mexicano de Ejecutivos de Finanzas (IMEF) y el Instituto Nacional de Estadística y Geografía (INEGI).¹ El IMEF pone a disposición dos versiones de su Indicador del Entorno Empresarial Mexicano para las actividades —esencialmente— manufactureras, uno ajustado por tamaño de empresa y otro sin ajustar (en este trabajo nos referimos a ellos como IMEF-AT e IMEF-IG, respectivamente).² El INEGI publica el Indicador de Pedidos Manufactureros (IPM), cuyo diseño toma en consideración el tamaño de las empresas. En ambos casos, la referencia es el *Purchasing Managers' Index (PMI)* que el *Institute for Supply Management (ISM)* reporta desde 1931 para Estados Unidos de América (EE. UU.).³ Los especialistas emplean este tipo de indicador para anticipar el curso de la economía en distintos países y si bien es usado frecuentemente de manera directa y sin mediación de modelos, también ha sido objeto de análisis estadísticos y econométricos con el fin de mejorar su utilidad.

Para obtener la información que se requiere, el IMEF envía un cuestionario a sus miembros vía internet; los indicadores que de él se derivan (IMEF-IG, IMEF-AT y niveles de cada uno de sus componentes) son difundidos con suma oportunidad, pues se publican el primer día hábil del mes posterior al de referencia. El INEGI levanta la Encuesta Mensual

de Opinión Empresarial (EMOE) durante las tres primeras semanas de cada mes por entrevista directa o vía internet en empresas del sector manufacturero que emplean más de cien personas: sus resultados (entre ellos el IPM) también son difundidos durante los primeros días del mes siguiente, lo que la ubica como la encuesta en unidades económicas más oportuna que ofrece esta institución.⁴

Sin duda, la oportunidad es una de las ventajas de los indicadores de difusión como los que aquí se analizan. Estos permiten conocer la percepción que tienen los directivos de las empresas acerca de la situación económica y de las decisiones que serán tomadas en relación con la producción, el empleo y los precios. Por lo tanto, son un recurso valioso para anticipar lo que ocurrirá con variables económicas duras menos oportunas. En ese sentido, las encuestas de opinión son cualitativas y complementan la información cuantitativa menos oportuna, pero más sólida, que ofrecen otras fuentes. En general, cuando se tienen dos indicadores que hacen referencia al mismo tema, el menos oportuno suele ser más preciso. Sin embargo, la oportunidad es un atributo valioso, que para ciertos propósitos y circunstancias compensa el sacrificio en materia de precisión. Los usuarios de indicadores oportunos deben conocer la manera más adecuada de aprovecharlos a la luz de las diferencias y similitudes entre la serie menos oportuna, pero más completa, y la más oportuna, pero menos completa.

En síntesis, el objetivo de los indicadores de difusión de opinión empresarial (como el IPM, IMEF-IG e IMEF-AT) es ofrecer cada mes información altamente oportuna que sea útil para la toma de decisiones en los sectores público y privado en tanto que permitan anticipar el comportamiento del sector manufacturero (o del sector económico de referencia o de la economía en su conjunto) tal

1 El INEGI trabaja en colaboración con el Banco de México (BM), quien cofinancia la Encuesta Mensual de Opinión Empresarial (EMOE) y participó en el diseño de las preguntas que se utilizan en la construcción del Indicador de Pedidos Manufactureros (IPM). Incluso, el nombre del indicador se estableció de común acuerdo entre ambas instituciones. Sin dejar de reconocer lo anterior en ningún momento, por facilidad de exposición, a lo largo de este documento nos referiremos al IPM como el indicador del INEGI, pero en todos los casos esto se deberá leer como el indicador que genera el INEGI con el apoyo del BM o el indicador diseñado y financiado conjuntamente por el BM y el INEGI.

2 Aunque de manera formal se denomina Indicador IMEF del Entorno Empresarial Mexicano, por brevedad llamamos a cada una de sus dos versiones como IMEF-IG e IMEF-AT para referirnos a la versión del indicador IMEF del Entorno Empresarial Mexicano que pondera igual a cada empresa independientemente de su tamaño y a la que reconoce más peso a las más grandes, respectivamente.

3 Un tercer indicador de este tipo lo produce para México *IHS Markit* desde el 2012, pero se desempeña pobremente en términos de correlaciones con indicadores duros como el Indicador Global de la Actividad Económica (ver Heath, 2016). Además, la serie histórica del indicador de *IHS Markit* no está disponible de manera gratuita, por lo que éste no compete en términos de accesibilidad con los producidos por las otras dos instituciones y, por lo tanto, se descarta de nuestro análisis.

4 Vale la pena distinguir que las publicaciones quincenales del Índice Nacional de Precios al Consumidor (INPC) forman la serie más frecuente, pero no la más oportuna. Mientras que el tiempo que transcurre entre la conclusión del periodo de referencia y la publicación de la información sobre el INPC es de 10 días —y de nueve cuando se trata del Índice Nacional de Precios Productor (INPP); en el caso de la EMOE la diferencia es de máximo cuatro días—; por otra parte, la información del INPC no proviene de encuestas ni se obtiene mediante entrevistas en establecimientos, sino a través de cotizaciones directas.

como se reportará varias semanas después con los indicadores más duros que difunde el INEGI.

El pulso de la actividad económica manufacturera se puede seguir por medio de estos indicadores de difusión, observando los cambios ocurridos de un mes a otro en relación con su umbral teórico. Los usuarios de estas fuentes típicamente basan sus interpretaciones con referencia al umbral teórico de 50 puntos, pues se trata de la mediana en el rango de puntuación posible y representa la opinión de que los temas sobre los que se pregunta han permanecido y se mantendrán igual.

El uso directo del umbral de 50 apela al diseño de las preguntas en el cuestionario y, concretamente, a los pesos relativos que se dan a las opciones de respuesta. Las opciones asociadas con una situación sin cambios tienen una ponderación de 50.0 (en escala de 0.0 a 100.0),⁵ por lo que, si todos los informantes respondieran con esa opción, el indicador nos estaría sugiriendo un crecimiento muy cercano a 0 en la variable de interés. De ahí que, si el indicador de difusión está por encima del umbral de 50, se suponga una tasa de crecimiento positiva de la variable de interés, y viceversa; en este sentido, mientras mayor sea la diferencia entre el nivel del indicador y el umbral, mayor será la tasa de crecimiento o decrecimiento esperado en la variable de interés. En la práctica, sin embargo, es posible que los informantes incorporen sesgos sistemáticos en sus respuestas que lleven a que el umbral verdadero, que separa la zona de crecimiento de la de decrecimiento, sea distinto del teórico. Es, justamente, en estos casos cuando resulta pertinente valerse de modelos estadísticos para estimar el valor del *umbral empírico*.⁶

Este trabajo condensa la investigación llevada a cabo para evaluar la capacidad del IPM, IMEF-IG e IMEF-AT de anticipar los cambios en la actividad

económica agregada.⁷ Además de analizar correlaciones, ajustamos un modelo de regresión logística como técnica para distinguir el umbral que permite optimizar la predicción del cambio en la variable de interés.

Como punto de partida, se explican las semejanzas y diferencias en el diseño de los indicadores del INEGI y del IMEF con el fin de establecer los límites de la comparabilidad. También, previo al análisis de correlaciones y a la evaluación de umbrales, dedicamos un apartado para exponer las estrategias en la elección de la variable susceptible de ser anticipada.

Acercando a los usuarios los parámetros de referencia adecuados y mostrando las ventajas de cada uno de los indicadores de difusión, buscamos contribuir al mejor aprovechamiento de la información generada por el INEGI y el IMEF.

Metodología

Como se indicó antes, la evaluación de la eficiencia de los indicadores tipo PMI disponibles para México requiere de dos pasos previos: una revisión detallada del diseño de los indicadores de difusión a evaluar y de las fuentes de las que proceden, y la identificación de la variable susceptible de anticiparse. Podemos distinguir entonces tres etapas consecutivas en la investigación:

1. Comprender el diseño de los indicadores para establecer comparaciones válidas. Conocer las semejanzas y diferencias entre el IPM, IMEF-AT e IMEF-IG en cuanto a su diseño y a las características de las fuentes de las que proceden es indispensable para evitar sesgos en la evaluación comparativa; por ejemplo, dado que la serie del IPM fue objeto de un

⁵ Es decir, una ponderación de 0.50 en una escala de 0.00 a 1.00.

⁶ Por ejemplo, el Indicador de Confianza del Consumidor (ICC), que producen conjuntamente el INEGI y el BM, no tiende a oscilar en torno a su umbral teórico de 50, sino que lo hace sobre el empírico de 34.5. Cuando ambos son tan diferentes, el uso acríptico del teórico puede conducir a serios equívocos en la interpretación del indicador (ver Leyva, Páez y Sainz, 2016).

⁷ Nuestra investigación no explora otras combinaciones de variables de la EMOE para el diseño de un indicador de difusión más eficiente ni modelos de uso de la información distintos a los que por tradición se han empleado para obtener este tipo de indicadores. Reconocemos que hay oportunidades para avanzar técnica y analíticamente en la integración de la información de manera novedosa y, en ese sentido, consideramos que un primer paso en esa dirección es la evaluación de los indicadores con las características que en la actualidad tienen.

cambio de diseño conceptual en el 2007, a efecto de usar series intertemporalmente consistentes, el periodo por analizar en esta investigación inicia en el 2008 y considera años completos hasta el 2017.

- II. Seleccionar la variable dura de referencia. Para llevar a cabo la evaluación, hay que elegir la variable susceptible de anticipar, es decir, la variable dura de referencia. Para ello, revisamos el comportamiento de las series históricas desestacionalizadas de los indicadores de difusión y de algunas variables económicas de interés⁸ buscando sincronías y/o correlaciones reveladoras a través de la comparación gráfica de la serie histórica del indicador frente a la de cada una de las variables económicas *candidatas*, la medición de su grado de correlación⁹ (incluidos rezagos para el indicador de difusión) y el cálculo del porcentaje de coincidencia en el sentido de las variaciones.
- III. Precisar el enfoque de evaluación. Partimos del análisis de la asociación estadística entre cada uno de los indicadores de difusión y la variable dura de referencia, para centrarnos posteriormente en la evaluación del éxito de la clasificación que produce el umbral teórico en cada caso. Cada uno de los tres indicadores de difusión se evalúa considerando cuatro medidas distintas computables a partir de una tabla de contingencia de 2 x 2, que son: 1) la tasa de error, 2) el valor predictivo negativo, 3) la tasa de correctos negativos y 4) el coeficiente Φ .¹⁰ La tabla de contingencia despliega en sus filas la agrupación de los niveles que ha

tomado el IPM, IMEF-IG e IMEF-AT según dos categorías basadas en el umbral teórico (<50 y >=50) que anticipan respectivamente que la variación porcentual de la variable dura de referencia será negativa o positiva; en sus columnas, divide las variaciones negativas de las positivas, de manera que la diagonal principal presenta los casos para los que el IPM, IMEF-AT e IMEF-IG anticipan correctamente el signo de la variación porcentual de la variable dura de referencia.

En esta fase es de utilidad ajustar un modelo *logit* con el propósito de distinguir el umbral para el indicador de difusión que mejor separa las variaciones positivas y negativas de la variable dura de referencia; de esta manera se pone a prueba la pertinencia de la utilización del umbral teórico de 50 puntos. En tal modelo, la variable dependiente es el signo de la variación porcentual de la variable dura de referencia y la variable independiente es el nivel del indicador de difusión. Con los parámetros estimados es posible despejar el valor del umbral empírico. Aunque de especificación muy elemental, el modelo descrito estima un valor de referencia alrededor del cual se pueden revisar resultados en tablas de contingencia alternativas. Con los resultados de la evaluación de los umbrales teóricos y su comparación con alternativos umbrales empíricos, sugerimos en qué caso conviene cambiar el valor de referencia para interpretar mejor el indicador de difusión.

Diseño del IPM y de los indicadores del IMEF

Como ya se mencionó, a partir de su encuesta a miembros, el IMEF produce y publica los indicadores IMEF-IG e IMEF-AT.¹¹ Sus análisis mensuales ponen énfasis en el primero, pero el equivalente al IPM del INEGI es el segundo, en tanto que valora distinto la opinión de las empresas en función de su tamaño. Si bien ambas instituciones toman

8 Disponibles en las páginas institucionales <http://www.inegi.org.mx/sistemas/bie/> y <http://www.iiem.org.mx>

9 Medida con el coeficiente de correlación producto o momento de Pearson, que refiere el grado de relación lineal, positiva o negativa, entre dos variables: $r = \frac{\sum(x - \bar{x})(y - \bar{y})}{\sqrt{\sum(x - \bar{x})^2 \sum(y - \bar{y})^2}}$, [-1, 1].

10 El coeficiente Φ mide la asociación entre dos variables binarias y se interpreta a la manera del coeficiente de correlación de Pearson. En *machine learning* se conoce como coeficiente de Matthews y es indicativo de la calidad de una clasificación binaria (Matthews, 1975). A diferencia de otras medidas de la calidad o eficiencia de la clasificación, el coeficiente de Matthews es útil incluso cuando las clases son muy diferentes en tamaño (Boughorbel, Jarray y El-Anbari, 2017). Puede tomar valores desde -1 (tasa de error máxima) hasta +1 (predicción perfecta). Más adelante, en la figura 3, se especifica la fórmula para su cálculo con base en una tabla de contingencia.

11 Se difunden bajo el nombre de IMEF e IMATE, pero en este documento nos referimos a ellos como IMEF-IG e IMEF-AT, respectivamente.

como referencia la construcción del PMI de EE. UU., hay diferencias de consideración en el diseño estadístico de sus levantamientos de información (ver figura 1).

La muestra a la que se aplica el cuestionario para generar los indicadores del IMEF es pequeña, pero es posible que los informantes estén especialmente bien calificados para reportar sobre su empresa; también hay que notar que no es probabilística, pues se basa sobre todo en un directorio de miembros, de manera que la composición sectorial es más reflejo de la composición sectorial de la membresía del IMEF que de la estructura de

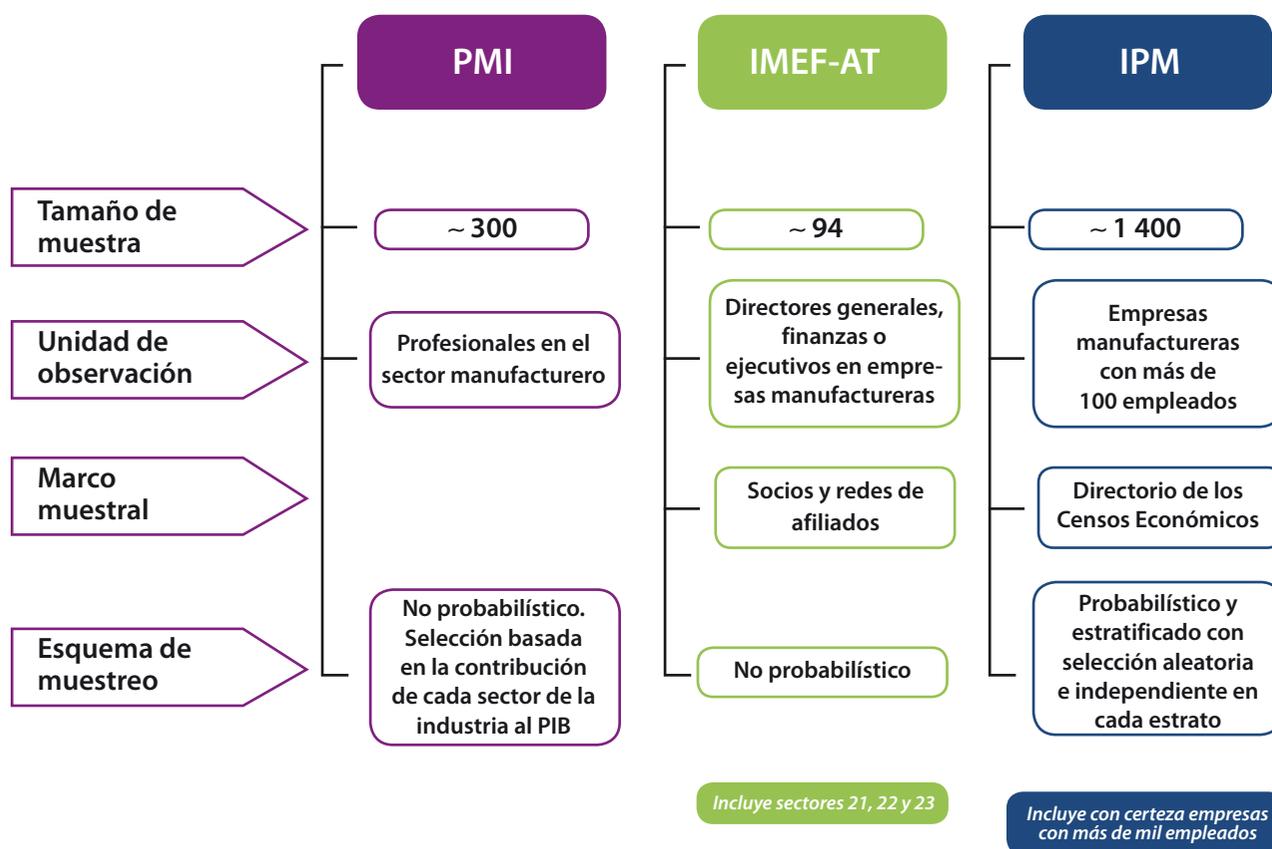
la economía que muestran los Censos Económicos o las cuentas nacionales.¹²

Más aún, cada indicador se construye de manera distinta. Si bien el IMEF pondera igual los cinco componentes de sus indicadores —siguiendo el ejemplo del PMI que cambió los ponderadores en el 2008—, los de *Oportunidad en la entrega...* e *Inventario de...* están referidos a los productos, no a los insumos. Por otra parte, la construcción del IPM está basada en preguntas conceptualmente

12 Otras diferencias respecto al IPM son que el IMEF considera para la construcción de sus indicadores manufactureros la opinión de empresas con menos de 100 empleados y también de algunas pocas que pertenecen al sector de la construcción.

Figura 1

Diseño estadístico comparado



Fuente: elaboración propia con información del ISM, IMEF e INEGI.

semejantes a las del PMI, pero sus ponderadores no son iguales, sino que pretenden reflejar la diferente importancia de las etapas en la cadena de producción, como se hacía antes en el caso del PMI (ver figura 2); es decir que, si bien tanto el indicador de difusión del INEGI como el del IMEF se inspiran en el PMI, ninguno de los dos es una réplica exacta de éste.

Además, es importante destacar cómo se calcula el nivel de cada componente de los indicadores de difusión que son nuestro objeto de estudio. En contraste con el PMI de EE. UU., los indicadores del IMEF y del INEGI están basados en preguntas con cinco opciones de respuesta,¹³ mientras que el PMI

usa únicamente tres. Las ponderaciones aplicadas en los indicadores mexicanos diluyen el peso de las opiniones negativas y positivas al incorporar escalas intermedias entre el 0.50 y el 1.00 (0.75) y entre el 0.50 y el 0 (0.25). Esto permite captar de manera diferenciada la intensidad de las apreciaciones y contribuye a reducir la volatilidad de las series.¹⁴

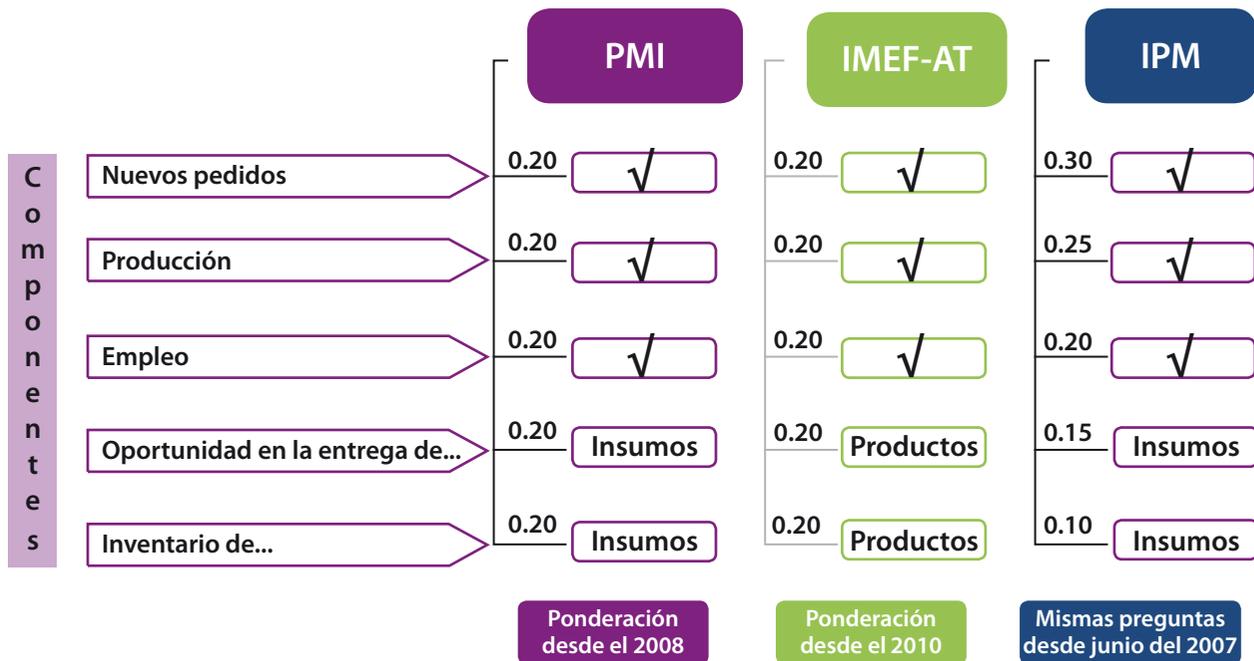
A pesar de las diferencias, las series históricas de los indicadores del INEGI y del IMEF tienen una notable asociación estadística: la correlación del IPM con los indicadores IMEF-IG e IMEF-AT es de 0.72 y 0.66, respectivamente, si se emplean datos mensuales, y de 0.79 y 0.73 si se calculan con promedios trimestrales. En general, muestran tendencias semejantes,

¹³ El número de opciones de respuesta usado en la EMOE, con la que se construye el IPM, responde a la intención por parte del INEGI de homologar estos reactivos con los de la Encuesta Nacional sobre Confianza del Consumidor (ENCO), que la precedió y en la que se utilizan cinco opciones de respuesta para las preguntas de las que se deriva el ICC: una opción neutral, dos de cambio moderado y dos de cambio extremo. Por su parte, los indicadores del IMEF heredan el número de opciones de respuesta del IPM.

¹⁴ Para conocer con mayor detalle las preguntas, opciones de respuesta, ponderadores y metodología para el cálculo de los indicadores de pedidos manufactureros del IMEF y del INEGI, se sugiere consultar el *Marco conceptual y metodológico del Indicador IMEF* en <http://www.iiem.org.mx/home/faq> y la *Síntesis metodológica de la Encuesta Mensual de Opinión Empresarial del Sector Manufacturero* en http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/metodologias/EMOE/emoe2012/sm_emoe_2012.pdf

Figura 2

Construcción de los indicadores



Fuente: elaboración propia con información del ISM, IMEF e INEGI.

aunque los indicadores del IMEF presentan mayor variabilidad (coeficiente de variación) en el tiempo (ver gráfica 1), lo cual seguramente se relaciona con el hecho de que su muestra es más pequeña y con que los informantes del IMEF son menos consistentes para responder a lo largo del tiempo.

Evaluando la eficiencia de los indicadores de difusión

La variable a anticipar: el IGAE

Los componentes de los indicadores de difusión que se muestran en la figura 2 pueden entenderse como parte de una cadena productiva. En ese sentido, un componente central por sus vínculos con cada uno de los otros componentes es la producción. Este componente es conceptualmente próximo a variables económicas de interés para los analistas, quienes podrían anticipar su evolución a partir de los niveles de los indicadores de difusión.

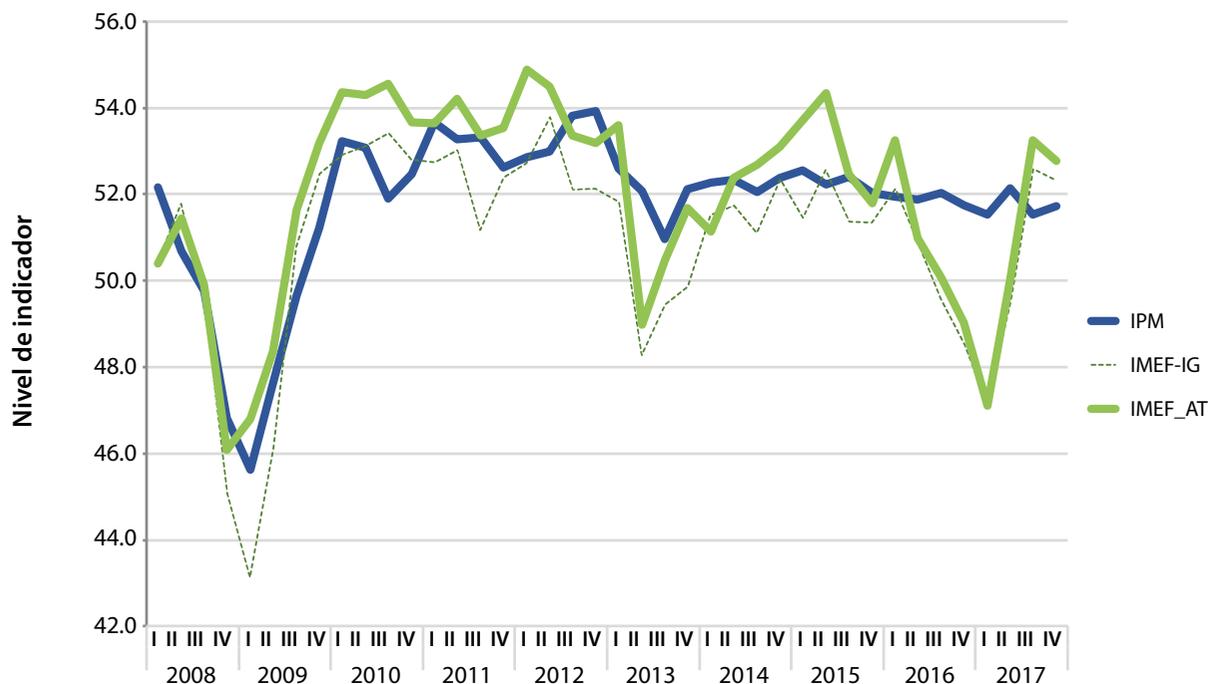
Para los propósitos de esta investigación fueron evaluadas las siguientes variables relacionadas con la producción a diferentes escalas: Producto Interno Bruto (PIB), PIB-Actividades Secundarias (PIB-AS), PIB-Industrias Manufactureras (PIB-IM), Indicador Global de la Actividad Económica (IGAE) e IGAE-Industrias Manufactureras (IGAE-IM). En el documento nos referimos a ellas como variables candidatas y también como variables duras.

Los elementos tomados en consideración para elegir cuál de estas variables intentaremos anticipar son: coincidencia en el sentido de las variaciones del indicador y de la variable candidata, análisis de correlación y correlación cruzada con un rezago para el indicador de difusión, análisis gráfico de las series históricas desestacionalizadas y análisis de diagramas de dispersión.¹⁵ El IGAE obtuvo los mejores resultados globales y, por lo tanto, fue la variable seleccionada para evaluar la efi-

¹⁵ En el *Anexo* se reportan los resultados para todas las variables candidatas.

Gráfica 1

Series históricas 2008-2017 del IPM, IMEF-IG e IMEF-AT trimestrales. Cifras desestacionalizadas



Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

ciencia del IPM, IMEF-IG e IMEF-AT en su función de anticipar la actividad económica agregada en México.¹⁶

Otros autores han reportado altas correlaciones entre indicadores de difusión manufactureros y el crecimiento del PIB o del IGAE, tal es el caso de Chien y Morris, quienes estimaron una correlación entre el PMI y el crecimiento del PIB de 0.75 para EE. UU. y de 0.73 para China.¹⁷ Replicando el ejercicio para México, Jonathan Heath reporta correlaciones con el IGAE de 0.81 y 0.73 para el IMEF-IG y el IPM, respectivamente.¹⁸

Chien y Morris promedian los PMI del trimestre y anualizan las tasas de crecimiento del PIB trimestral. La serie va del 2005 al 2015. Heath también *trimestraliza* las series desestacionalizadas y analiza el periodo del primer trimestre del 2006 al primero del 2016. En nuestra investigación, confirmamos que los niveles más altos de correlación se obtienen cuando las series de los datos desestacionalizados se convierten a trimestres calendario; en

este caso, el periodo de análisis es diferente, va del 2008 al 2017, y las series económicas empleadas tienen como año base el 2013 (ver cuadro 1).

El IPM tiene mejor correlación con el IGAE en comparación con los indicadores del IMEF si se emplean las cifras mensuales, aunque la correlación no es tan buena como la encontrada al *trimestralizar* las series. Tanto el IPM como las dos versiones del IMEF tienen correlaciones más altas con el IGAE cuando se emplean las series en trimestres calendario. Como Heath lo había observado, considerando exclusivamente esta medida de asociación, el IMEF-IG es un poco mejor que el IPM. No obstante, comparando entre comparables, se encuentra que el IPM es superior al IMEF-AT. En el siguiente apartado mostraremos el desempeño de los indicadores de difusión tomando en cuenta otras medidas que evalúan precisamente el éxito con el que anticipan el IGAE.

El cuadro 2 ilustra que las correlaciones con el PIB son, como se espera, muy parecidas a las que hay con el IGAE, aunque ligeramente menores. Por otra parte, también muestra que las correlaciones son más fuertes con las variables económicas a nivel más agregado a pesar de que los indicadores de difusión son extraídos del sector manufacturero. Se debe resaltar, además, que el ordenamiento del IPM, IMEF-IG e IMEF-AT a partir de las correlaciones es consistente independientemente de con cuál de estas variables se le asocie.

¹⁶ Para asegurar que el resultado del análisis pueda generalizarse hacia otros periodos no considerados, es importante verificar la estacionariedad de las series (ver Gujarati, 2004). Mediante la prueba Dickey Fuller Aumentada (ADF, por sus siglas en inglés) comprobamos que la serie del IGAE es estacionaria hasta con un rezago, independientemente de si la hipótesis nula considera un término de deriva o si la regresión para obtener el estadístico de prueba incluye una constante o un término para la tendencia en el tiempo. Las series del IPM y de los indicadores del IMEF son estacionarias sin rezagos, con constante y sin tendencia.

¹⁷ Chien y Morris, 2016.

¹⁸ Heath, 2016.

Cuadro 1

Coeficientes de correlación del IPM, IMEF-IG e IMEF-AT con el IGAE. Cifras desestacionalizadas 2008-2017

Nivel del indicador	Variación porcentual del IGAE		
	Mensual (n = 120)	Trimestre móvil (n = 120)	Trimestre calendario (n = 40)
IPM	0.3940	0.5889	0.6751
IMEF-IG	0.3824	0.6071	0.7168
IMEF-AT	0.3763	0.5339	0.5970

Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

Cuadro 2

Coefficientes de correlación del IPM, IMEF-IG e IMEF-AT con el IGAE, PIB, PIB-AS e IGAE-IM. Cifras desestacionalizadas 2008-2017

Nivel del indicador	Variación porcentual en el trimestre calendario			
	IGAE	PIB	PIB-AS	IGAE-IM
IPM	0.6751	0.6563	0.6019	0.5959
IMEF-IG	0.7168	0.7080	0.6542	0.6502
IMEF-AT	0.5970	0.5820	0.5440	0.5111

Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

En resumen, cualquiera de las variables incluidas en el cuadro podría ser anticipada a partir del nivel de los indicadores de difusión que aquí revisamos. También, debemos señalar que la lista de variables examinada no es exhaustiva, por lo cual el tipo de análisis que describimos en este documento puede ser replicado para otras variables de interés.

Funcionamiento de los umbrales teórico y empíricos

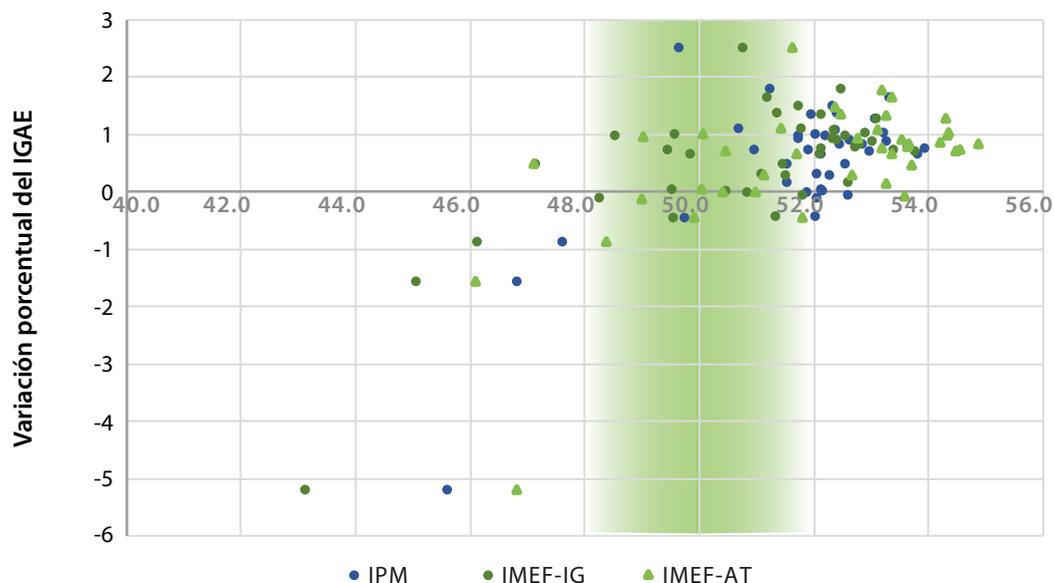
El supuesto para la lectura de los indicadores es que cuando se supera el umbral teórico se espe-

ran variaciones positivas del IGAE, y viceversa. El éxito con el que el indicador de difusión anticipa la actividad económica agregada es la esencia de su utilidad práctica. Como se ve en la gráfica 2, algunas veces el nivel de los indicadores de difusión ha sido menor a 50 y, sin embargo, el IGAE ha variado de manera positiva; también, ha ocurrido que el nivel ha sido mayor a 50 y el IGAE ha variado negativamente.

El análisis de correlaciones llevado a cabo en otras investigaciones, y replicado en ésta, es informativo pero insuficiente si se desea evaluar la

Gráfica 2

Relación de los indicadores de difusión con el IGAE. Series trimestrales 2008-2017



Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

eficiencia del umbral 50 de cada indicador para anticipar el cambio en el IGAE.¹⁹ Para llevar a cabo esta evaluación, organizamos la información en una tabla de contingencia en la que incluimos los valores esperados en las filas y los ocurridos en las columnas.²⁰ Los valores esperados dependen del nivel del indicador de difusión y se agrupan en dos clases: *por debajo del umbral de 50 y en el umbral o por encima de él*. En las columnas se agrupan las variaciones del IGAE también en dos clases: *negativas y nulas o positivas* (ver cuadro 3). Así, la diagonal principal (celdas A y D) presenta los casos en los que el indicador de difusión anticipa

correctamente el signo de la variación porcentual del IGAE. Cabe precisar que hay una diversidad de medidas que pueden calcularse a partir de una tabla de este tipo. La medida adecuada depende del propósito de la clasificación, de qué tipo de error es más tolerable e incluso de cómo están distribuidos los datos entre las clases definidas; en otras palabras, se debe definir qué se desea maximizar o minimizar.

En este caso, decidimos evaluar la eficiencia del umbral teórico usando cuatro medidas alternativas: 1) la tasa de error, 2) el valor predictivo negativo, 3) la tasa de correctos negativos y 4) el coeficiente Φ (ver figura 3).

A continuación, se describe brevemente cada una de ellas:

1. La **tasa de error** reporta el porcentaje del total de casos en los que la posición del indicador respecto al umbral no concuerda con el signo de la variación observada en la variable de interés. Esta medida no distingue entre los dos tipos de fallos (anticipar que la variación será negativa y que resulte positiva o anticipar que será positiva y resulte negativa), por lo cual no es la más adecuada en casos como el que nos ocupa porque la

19 Este tipo de análisis lo hemos llevado a cabo con anterioridad para el caso del ICC que producen conjuntamente el INEGI y el BM (Leyva, Páez y Sainz, 2016).

20 A esta clase particular de tabla de contingencia se le conoce como *matriz de confusión* en el campo de *machine learning*.

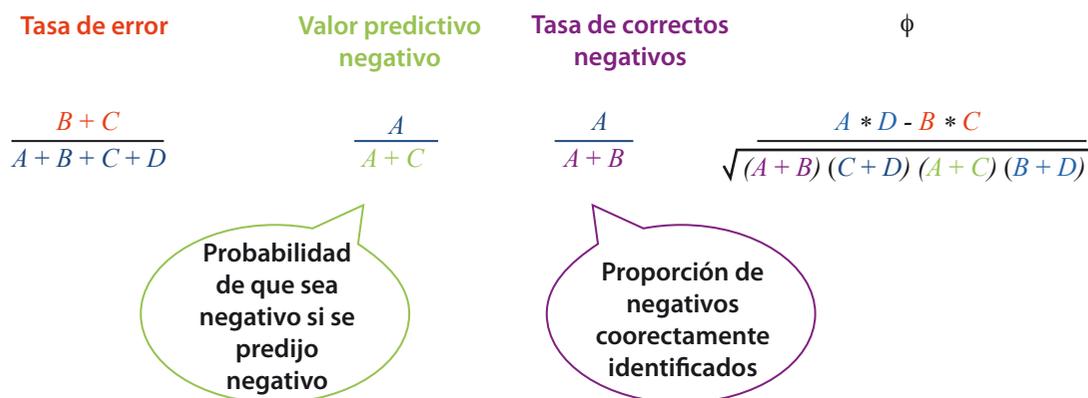
Cuadro 3
Tabla de contingencia para la evaluación de la eficiencia del umbral teórico

Variación porcentual del IGAE			
Indicador	Negativa	Positiva	Total
<50	A	C	A + C
>= 50	B	D	B + D
Total	A + B	C + D	A + B + C + D

Fuente: elaboración propia.

Figura 3

Medidas de eficiencia en una tabla de contingencia (a partir de la definición de las celdas del cuadro 3)



Fuente: elaboración propia con base en Boughorbel, S. et al. y James, G. et al.

ocurrencia de variaciones negativas es bastante menos frecuente que la de positivas, pero la incluimos por ser la medida básica de eficiencia de la clasificación.

2. El **valor predictivo negativo** indica la probabilidad de que la variación porcentual del IGAE sea negativa dado que se anticipó que lo sería. Es, entonces, una medida de la certeza de la predicción que resulta especialmente relevante. La incluimos porque al ser poco frecuente que ocurran las variaciones negativas en el IGAE, interesa que cuando se detecte esa posibilidad se haga con la mayor certeza posible.
3. La **tasa de correctos negativos** es la proporción de negativos correctamente identificados. En este caso particular, es relevante su cálculo porque las opiniones empresariales suelen tender al optimismo y es posible que por ello el indicador de difusión pierda parte de su potencial para detectar disminuciones del IGAE.
4. Finalmente, **el coeficiente Φ** resume en una cifra la asociación entre las dos variables. En

el mejor de los casos, cuando no hubiere discordancia entre ellas y la diagonal principal concentrare todas las observaciones, la tasa de error sería nula y, el valor predictivo negativo, la tasa de correctos negativos, así como Φ , serían igual a la unidad. Esta medida es mejor que la tasa de error porque toma en cuenta de manera simultánea toda la información de la matriz, pero de forma tal que es sensible a los desbalances entre las categorías de la clasificación, lo cual es importante en casos como el que nos interesa, en el que la frecuencia de variaciones positivas del IGAE es mayor que la de negativas.

Si bien evaluamos tanto las series mensuales como las trimestrales, por brevedad de exposición a continuación presentamos exclusivamente las clasificaciones en tablas de contingencia resultantes de las series trimestrales (ver cuadro 4), pues son las que mejores resultados producen. Enseguida, incluimos un comparativo de las medidas de eficiencia seleccionadas para nuestros propósitos (ver cuadro 5).

Cuadro 4

Tablas de contingencia para evaluar la eficiencia de los umbrales teóricos del IPM, IMEF-IG e IMEF-AT con series trimestrales 2008-2017

Variación porcentual del IGAE			
Umbral IPM	Negativa	Positiva	Total
<50	4	1	5
>= 50	5	30	35
Total	9	31	40
Umbral IMEF-IG	Negativa	Positiva	Total
<50	5	6	11
>= 50	4	25	29
Total	9	31	40
Umbral IMEF-AT	Negativa	Positiva	Total
<50	5	2	7
>= 50	4	29	33
Total	9	31	40

Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

Cuadro 5

Eficiencia de los indicadores de difusión según su capacidad de anticipar las variaciones del IGAE con el umbral teórico: cuatro medidas alternativas. Series trimestrales 2008-2017

Indicador	(Min) Tasa de error	(Max) Valor predictivo negativo	(Max) Tasa de correctos negativos	(Max) Φ
IPM	0.1500	0.8000	0.4444	0.5204
IMEF-IG	0.2500	0.4545	0.5555	0.3385
IMEF-AT	0.1500	0.7143	0.5555	0.5396

Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

Puede observarse que el IMEF-AT es marginalmente mejor que el IPM en dos de los cuatro criterios considerados: tasa de correctos negativos y Φ . Sin embargo, el IPM es marginalmente preferible con el criterio de valor predictivo negativo. Los analistas económicos pueden decidir observar el nivel del IPM o el del IMEF-AT dependiendo de lo que buscan: si lo que se quiere es maximizar la probabilidad de anticipar de manera acertada variaciones negativas, la balanza se inclina ligeramente hacia el IPM; si lo que interesa es tener un óptimo comportamiento global del indicador, como lo mide Φ , la balanza se inclina de forma leve hacia el IMEF-AT. Si ambos criterios importan, entonces tal vez lo procedente sería observar simultáneamente al IPM y al IMEF-AT. En todo caso, es conveniente tener presente que las diferencias en general son pequeñas y que, con series *trimestralizadas*, las señales que se pueden extraer tanto del IPM como del IMEF-AT son muy similares.

Para completar el análisis, ajustamos un modelo *logit* en cada caso para detectar si otro nivel de los indicadores de difusión evaluados anticipa mejor las variaciones positivas y negativas del IGAE (ese otro nivel se encontraría en la zona cálida ilustrada en la gráfica 2). Se trata de modelar el sentido de la variación trimestral del IGAE en función del indicador de difusión:

$$\log\left(\frac{p(\Delta Y \geq 0|X)}{p(\Delta Y < 0|X)}\right) = \beta_0 + \beta_1 X$$

ΔY : variación porcentual del IGAE

X : nivel del indicador de difusión

El umbral empírico $\frac{-\beta_0}{\beta_1}$ resultante para el IPM es 49.8. Su proximidad con el teórico revela que no se requiere de un umbral alternativo para mejorar la capacidad de anticipar la evolución del IGAE. Para el IMEF-IG, el umbral sugerido por el modelo es 48.5. Aunque con él se consigue una menor tasa de error, un mayor poder predictivo negativo y un mayor grado de correlación entre las dos variables, disminuye la tasa de variaciones negativas correctamente identificadas. En cuanto al IMEF-AT, con el umbral empírico de 49 mejora el valor predictivo negativo, pero empeora la tasa de correctos negativos y el coeficiente Φ . Por lo tanto, solo para el indicador IMEF-IG sería recomendable el uso de un umbral empírico (ver cuadro 6).

Conclusiones

Los analistas económicos profesionales que usan los indicadores de difusión para tener un panorama oportuno de la situación económica del país disponen de tres indicadores de diseño semejante, pero no idéntico, que pueden anticipar el signo de la variación del IGAE. El IPM y el IMEF-AT (las dos medidas ajustadas por tamaño de empresa) se comportan algo mejor que el IMEF-IG pues tienen coeficientes Φ superiores. Además, sus umbrales teóricos minimizan la tasa de error al anticipar el sentido de la variación del IGAE de un trimestre a otro, por lo que la utilización de umbrales empíricos distintos del teórico no ofrece ganancia. En el caso del IMEF-IG un umbral de 48.5 mejora su capacidad predictiva.

Aunque tanto el IPM como el IMEF-AT son buenos indicadores de la evolución del IGAE de un

Cuadro 6

Eficiencia de los indicadores de difusión según su capacidad de anticipar las variaciones del IGAE con umbrales teóricos y empíricos: cuatro medidas alternativas. Series trimestrales 2008-2017

Indicador	(Min) Tasa de error	(Max) Valor predictivo negativo	(Max) Tasa de correctos negativos	(Max) Φ
IPM				
50.0	0.1500	0.8000	0.4444	0.5204
49.8	0.1500	0.8000	0.4444	0.5204
IMEF-IG				
50.0	0.2500	0.4545	0.5555	0.3385
48.5	0.1500	0.8000	0.4444	0.5204
IMEF-AT				
50.0	0.1500	0.7143	0.5555	0.5396
49.0	0.1500	0.8000	0.4444	0.5204

Fuente: elaboración propia con datos del INEGI e IMEF consultados el 1 de marzo de 2018.

trimestre a otro, es importante distinguir para qué propósito es mejor cada uno de ellos. Las observaciones empíricas hasta el momento muestran que cuando la variación del IGAE ha sido negativa, el IMEF-AT ha estado más veces por debajo de 50, de manera que si lo que interesa en primera instancia es adelantar de forma adecuada el mayor porcentaje de variaciones negativas del IGAE, entonces el IMEF-AT es potencialmente de mayor utilidad. No obstante, si lo que interesa es estar lo más seguro posible de que cuando se adelante una variación negativa del IGAE ésta resulte en realidad negativa, entonces, en el margen, conviene usar el IPM.

La metodología desarrollada en estas páginas puede replicarse para analizar el vínculo entre estos indicadores de difusión y otras variables económicas, tales como: PIB, PIB de Actividades Secundarias e IGAE de Industrias Manufactureras.

La utilidad del IPM y de los indicadores del IMEF queda de manifiesto en términos de la oportunidad con la que se publican, pues con ellos es posible anticipar hasta en siete semanas si el IGAE variará positiva o negativamente respecto al trimestre previo. En consideración de los criterios

usados en este documento, las claras ventajas del IPM en cuanto a sofisticación en la operación, rigor conceptual en el fraseo de las preguntas y tamaño de muestra no se reflejan en ventajas prácticas notables con respecto a los indicadores del IMEF, llegando a ocurrir en algunos casos que la ventaja, si bien marginal, está del lado de éstos.

Fuentes

- Boughorbel, S., F. Jarray y M. El-Anbari. "Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric", en: *PLOS One*. Vol. 12, No. 6, 2017 (DE) <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0177678>
- Chien, Y. y P. Morris. "PMI and GDP: Do They Correlate for the United States? For China?", en: *Economic Synopses*. No. 6, 2016 (DE) <https://doi.org/10.20955/es.2016.6>
- Gallardo, M. y M. Pedersen. *Indicadores líderes compuestos. Resumen de metodologías de referencia para construir un indicador regional en América Latina. Serie Estudios estadísticos y prospectivos*. Núm. 49. Santiago de Chile, Naciones Unidas, CEPAL, División de Estadística y Proyecciones Económicas, 2007 (DE) https://repositorio.cepal.org/bitstream/handle/11362/4750/1/S0700234_es.pdf
- Gujarati, D. *Basic Econometrics*. Fourth Edition. New York, The McGraw-Hill Companies, 2004.

- Harris, E. "Tracking the Economy with the Purchasing Managers' Index", en: *Quarterly Review*. Vol. 16, No. 3. Federal Reserve Bank of New York, Autumn 1991, pp. 61-69 (DE) https://www.newyorkfed.org/medialibrary/media/research/quarterly_review/1991v16/v16n3article5.pdf
- Heath, J. *Los PMI de México*. 2016 (DE) <http://jonathanheath.net/2016/05/>
- _____. *Lo que indican los indicadores. Cómo utilizar la información estadística para entender la realidad económica de México*. México, INEGI, 2012 (DE) http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/estudios/indican_indi/indica_v25iv12.pdf
- Heath, J. y L. Domínguez. *Marco conceptual y metodológico del indicador IMEF*. 2016 (DE) <http://www.iiem.org.mx/home/faq>
- IMEF. *Indicador IMEF del Entorno Empresarial Mexicano. Información estadística*. México, IMEF, 2018 (DE) <http://www.iiem.org.mx/>
- INEGI. *Banco de Información Económica (BIE)*. México, INEGI, 2018 (DE) <http://www.inegi.org.mx/sistemas/bie/>
- _____. *Encuesta Mensual de Opinión Empresarial del Sector Manufacturero: Documento de resultados*. Aguascalientes, México, INEGI, 2015 (DE) <http://www.beta.inegi.org.mx/app/biblioteca/ficha.html?upc=702825051198>
- _____. *Encuesta Mensual de Opinión Empresarial (EMOE) 2015. Manufacturas. Documento metodológico sobre diseño muestral*. México, INEGI, 2014 (DE) www3.inegi.org.mx/rnm/index.php/catalog/133/download/4249
- _____. *Encuesta Mensual de Opinión Empresarial del Sector Manufacturero 2012. Síntesis metodológica. EMOE*. México, INEGI, 2012 (DE) http://internet.contenidos.inegi.org.mx/contenidos/productos/prod_serv/contenidos/espanol/bvinegi/productos/metodologias/EMOE/emoe2012/sm_emoe_2012.pdf
- ISM. *ROB Manufacturing Data*. Arizona, ISM, 2018 (DE) <https://www.instituteforsupplymanagement.org/news/content.cfm?ItemNumber=28859&SSO=1>
- James, G. et al. *An Introduction to Statistical Learning with Applications in R*. New York, Springer, 2015.
- Koenig E. "Using the Purchasing Managers' Index to Assess the Economy's Strength and the Likely Direction of Monetary Policy", en: *Economic & Financial Policy Review*. Vol. 1, No. 6. Federal Reserve Bank of Dallas, 2002 (DE) <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.199.5065&rep=rep1&type=pdf>
- Lahiri, K. y G. Monokroussos "Nowcasting US GDP: The role of ISM Business Surveys", en: *International Journal of Forecasting*. Vol. 29, No.4, 2013 (DE) http://www.researchgate.net/publication/228455397_Nowcasting_US_GDP_The_role_of_ISM_business_surveys
- Leyva, G., O. Páez y E. Sainz "Un umbral empírico y otras recomendaciones para el reporte de la confianza del consumidor en México", en: *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*. Vol. 7, Núm. 1. México, INEGI, 2016 (DE) <http://www.inegi.org.mx/rde/2016/01/05/un-umbral-empirico-y-otras-recomendaciones-para-el-reporte-de-la-confianza-del-consumidor-en-mexico/>
- Matthews, B. "Comparison of the predicted and observed secondary structure of T4 phage lysozyme", en: *Biochimica et Biophysica Acta (BBA)—Protein Structure*. Vol. 405, No. 2, 1975, pp. 442-45 (DE) https://www.researchgate.net/publication/21974618_Matthews_B_Comparison_of_the_predicted_and_observed_secondary_structure_of_T4_phage_lysozyme_Biochimica_et_Biophysica_Acta_BBA-Protein_Structure_4052_442-451
- StataCorp. *Stata: Release 15*. Statistical Software. College Station. TX New York, StataCorp LLC, 2017.

Anexo

Resultados de la evaluación estadística para seleccionar la variable de referencia

1

Coincidencia en el signo de las variaciones mensuales. Series desestacionalizadas 2008-2017

Nivel del indicador	IGAE	IGAE-IM
IPM	58%	58%
IMEF-IG	50%	43%
IMEF-AT	51%	58%

Fuente: elaboración propia con datos del INEGI e IMEF.

Coincidencia en el signo de las variaciones trimestrales. Series desestacionalizadas 2008-2017

Nivel del indicador	IGAE	IGAE-IM	PIB	PIB-AS	PIB-IM
IPM	65%	58%	63%	60%	63%
IMEF-IG	70%	63%	63%	60%	58%
IMEF-AT	60%	63%	58%	50%	53%

Fuente: elaboración propia con datos del INEGI e IMEF.

Correlación y correlación cruzada con series mensuales. Series desestacionalizadas 2008-2017

Nivel del indicador	IGAE	IGAE-IM
IPM	0.39	0.39
IMEF-IG	0.38	0.30
IMEF-AT	0.38	0.24
IPM ^{t-1}	0.28	0.26
IMEF-IG ^{t-1}	0.30	0.30
IMEF-AT ^{t-1}	0.23	0.26

Fuente: elaboración propia con datos del INEGI e IMEF.

Correlación y correlación cruzada con series trimestrales. Series desestacionalizadas 2008-2017

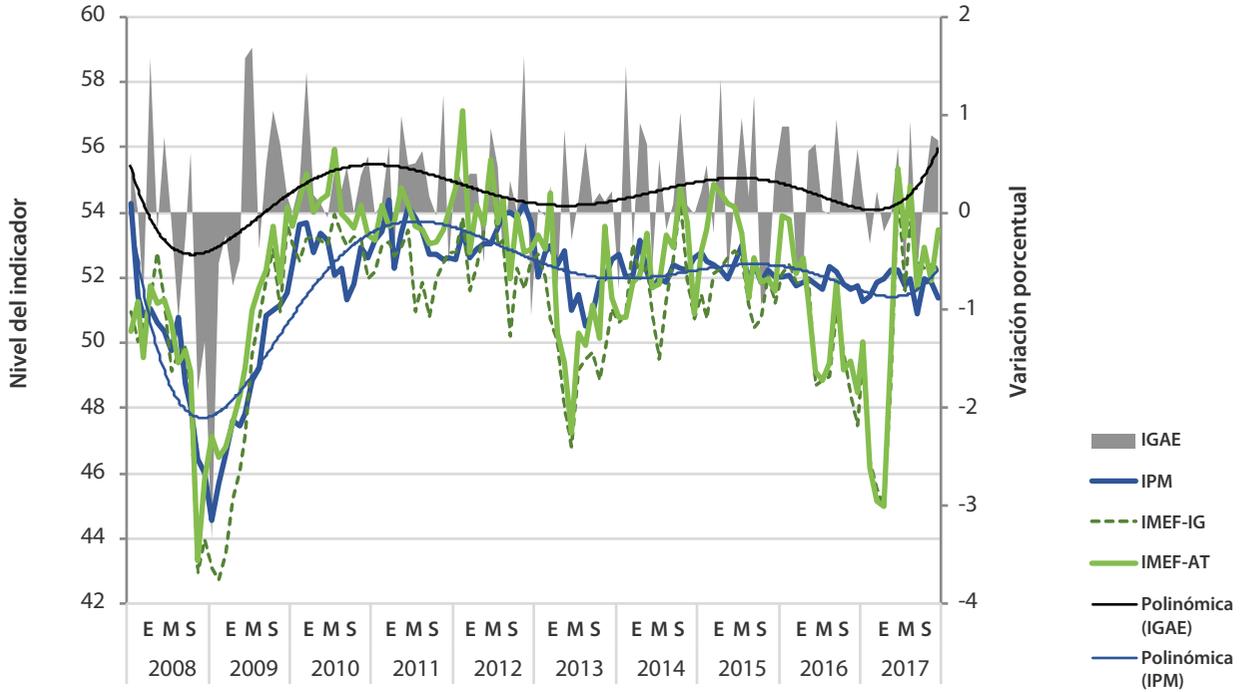
Nivel del indicador	IGAE	IGAE-IM	PIB	PIB-AS	PIB-IM
IPM	0.68	0.60	0.66	0.60	0.60
IMEF-IG	0.72	0.65	0.71	0.65	0.65
IMEF-AT	0.60	0.51	0.58	0.54	0.50
IPM ^{t-1}	0.44	0.35	0.43	0.37	0.33
IMEF-IG ^{t-1}	0.44	0.30	0.41	0.36	0.30
IMEF-AT ^{t-1}	0.45	0.34	0.42	0.42	0.34

Fuente: elaboración propia con datos del INEGI e IMEF.

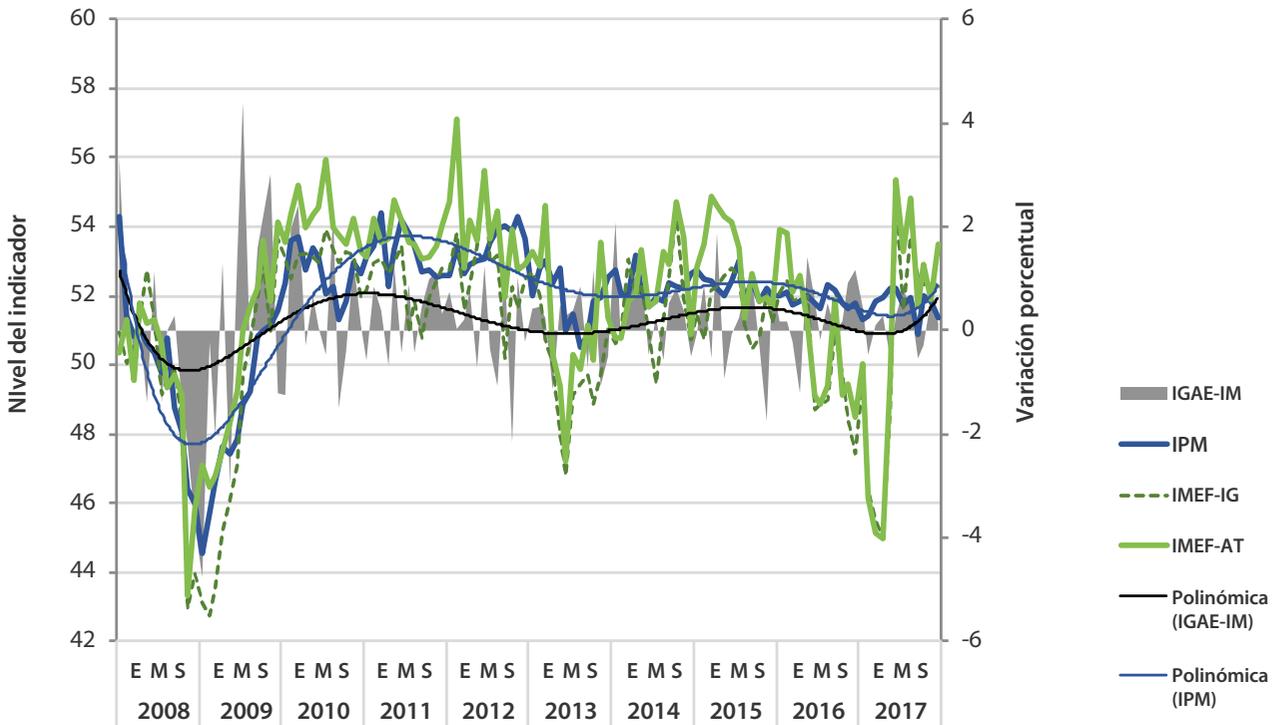
5

Gráficas de las series históricas desestacionalizadas 2008-2017

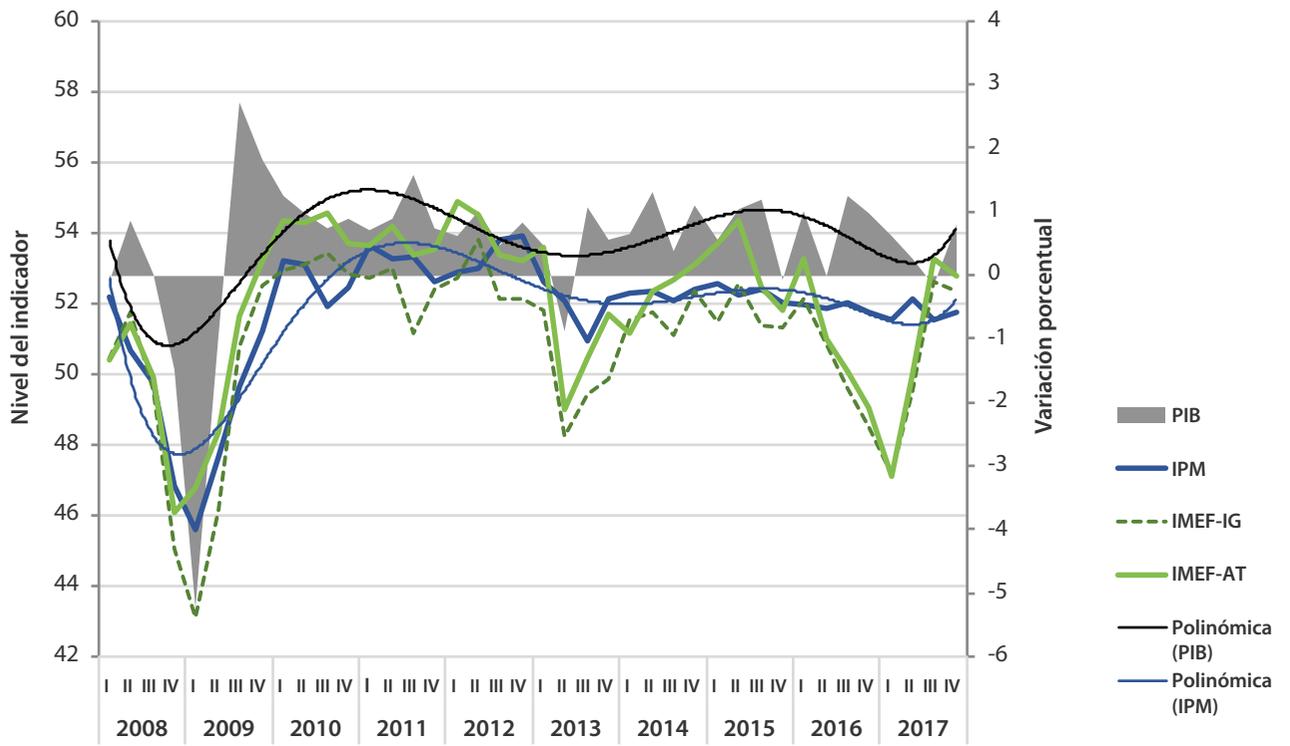
a



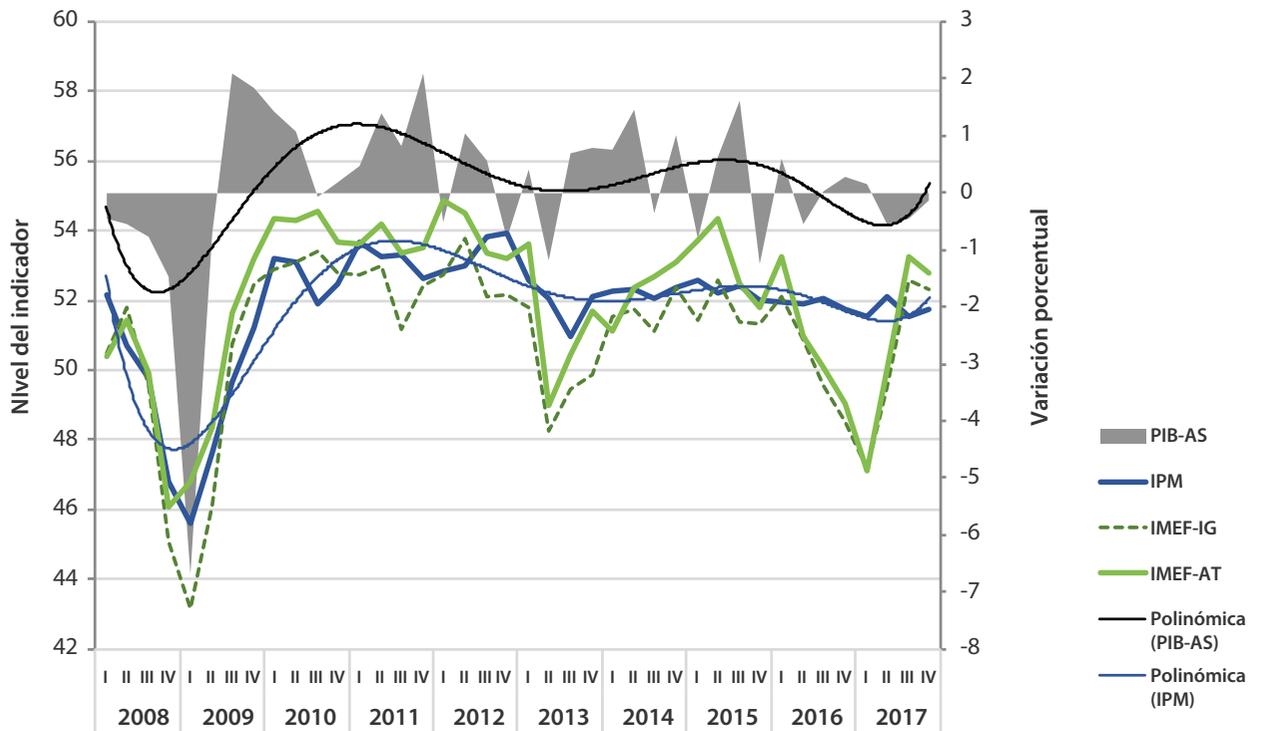
b

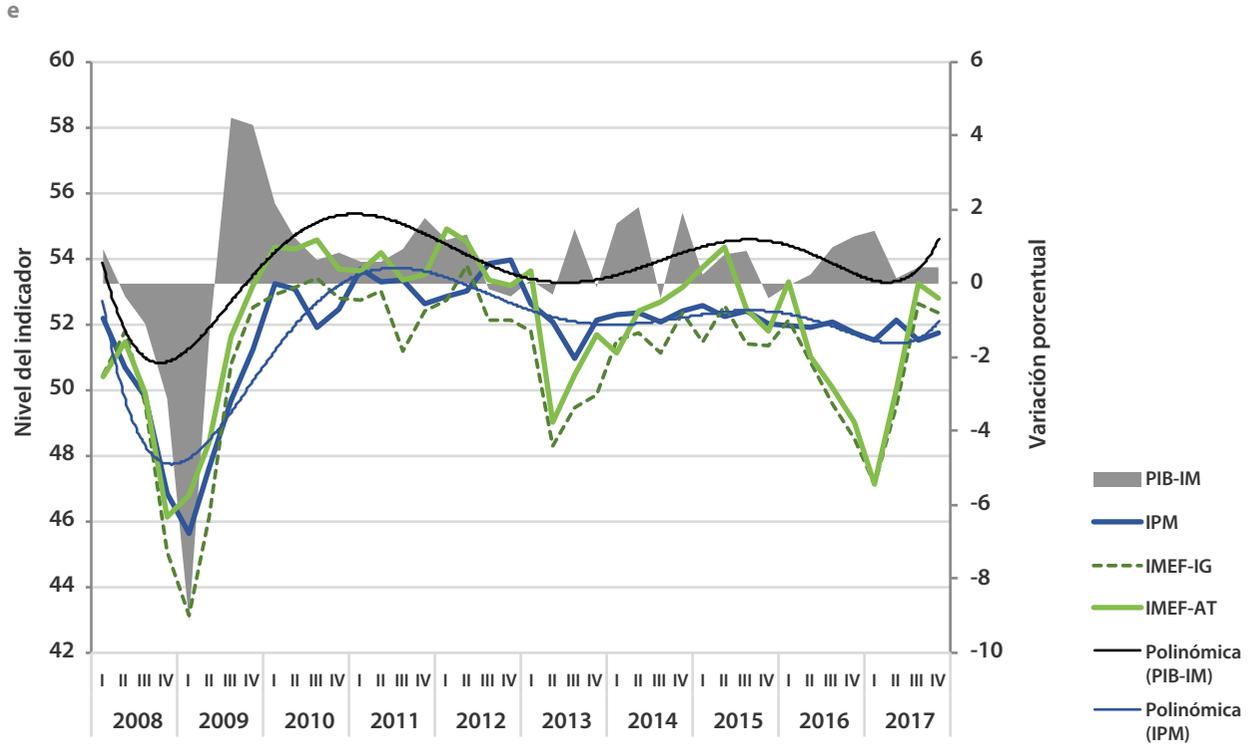


c



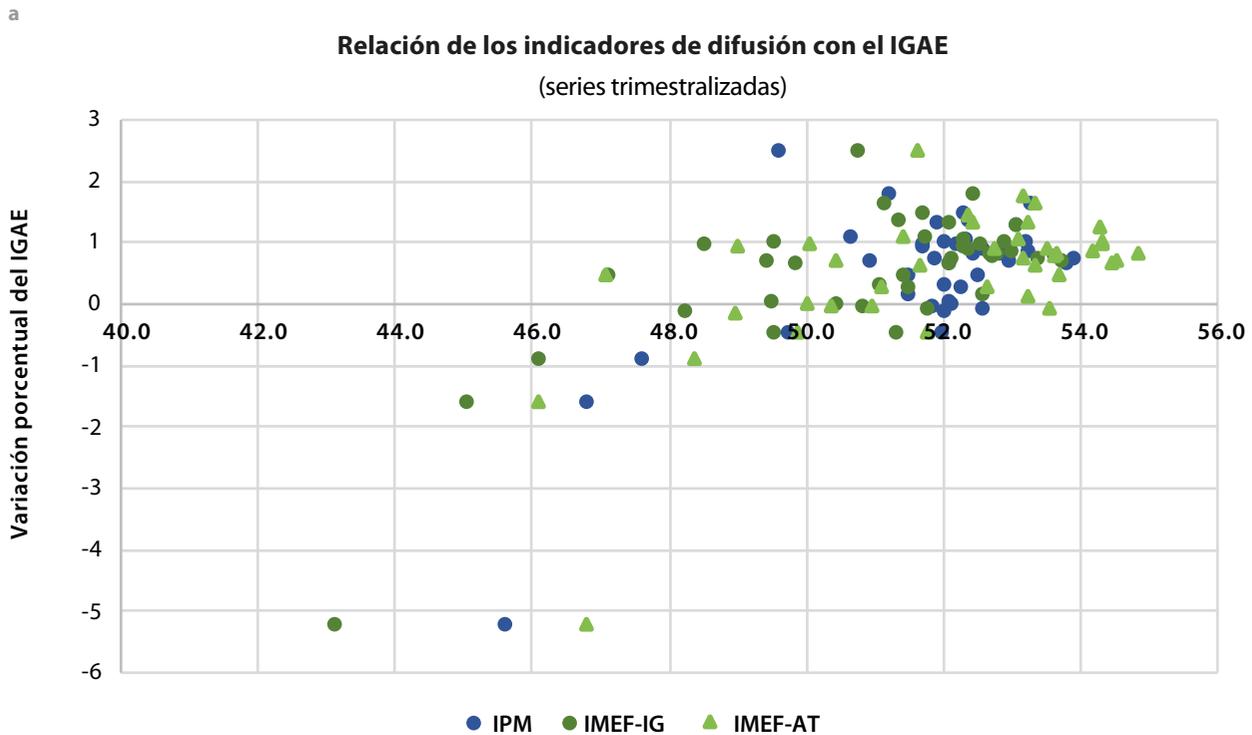
d



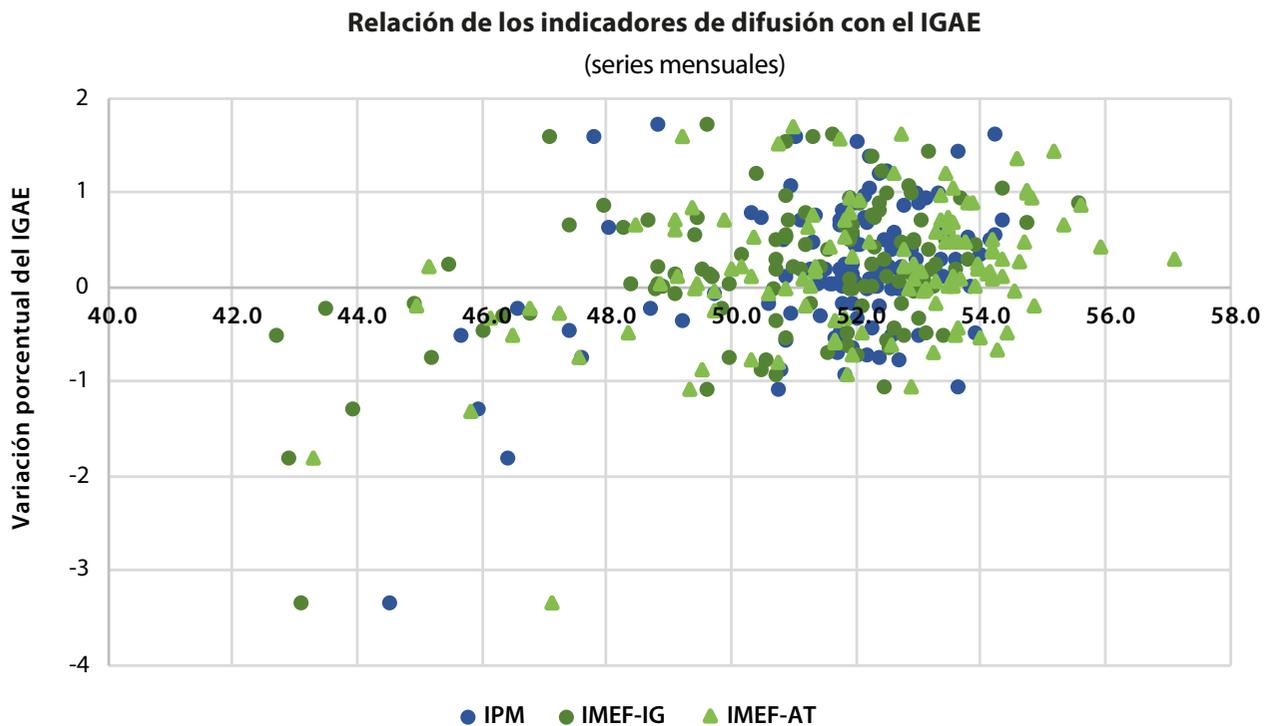


Fuente: elaboración propia con datos del INEGI e IMEF.

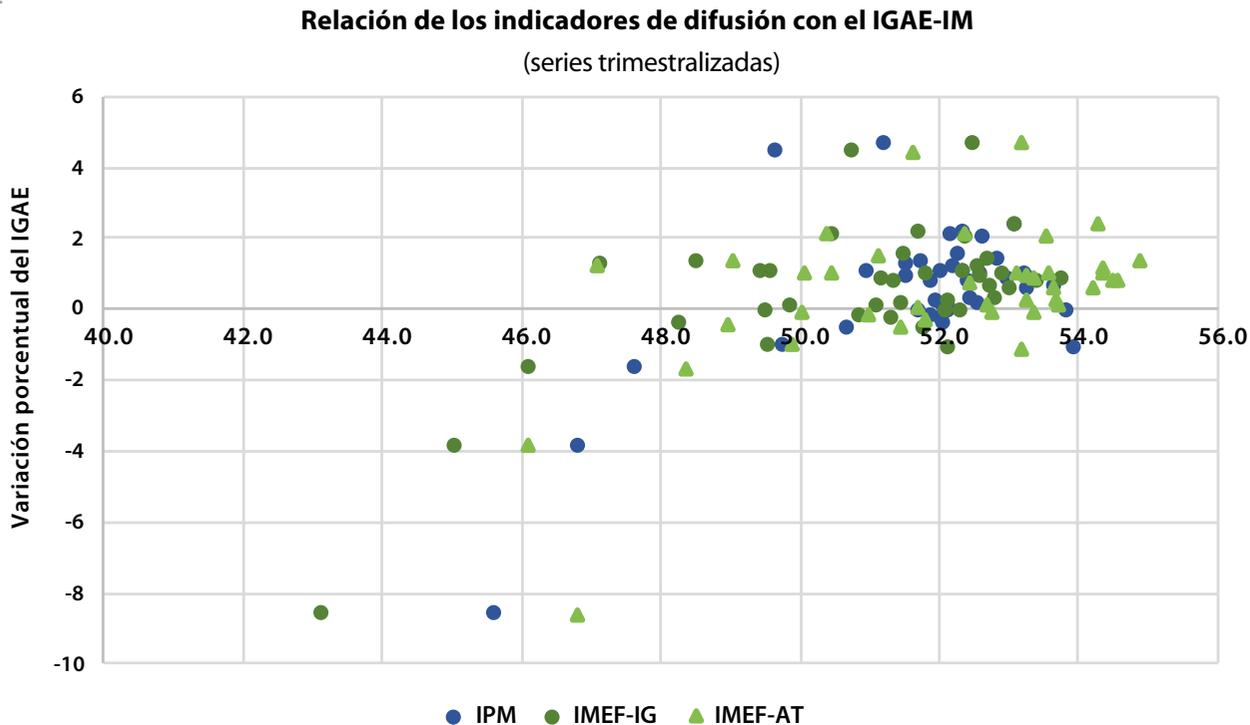
Diagramas de dispersión con series desestacionalizadas 2008-2017

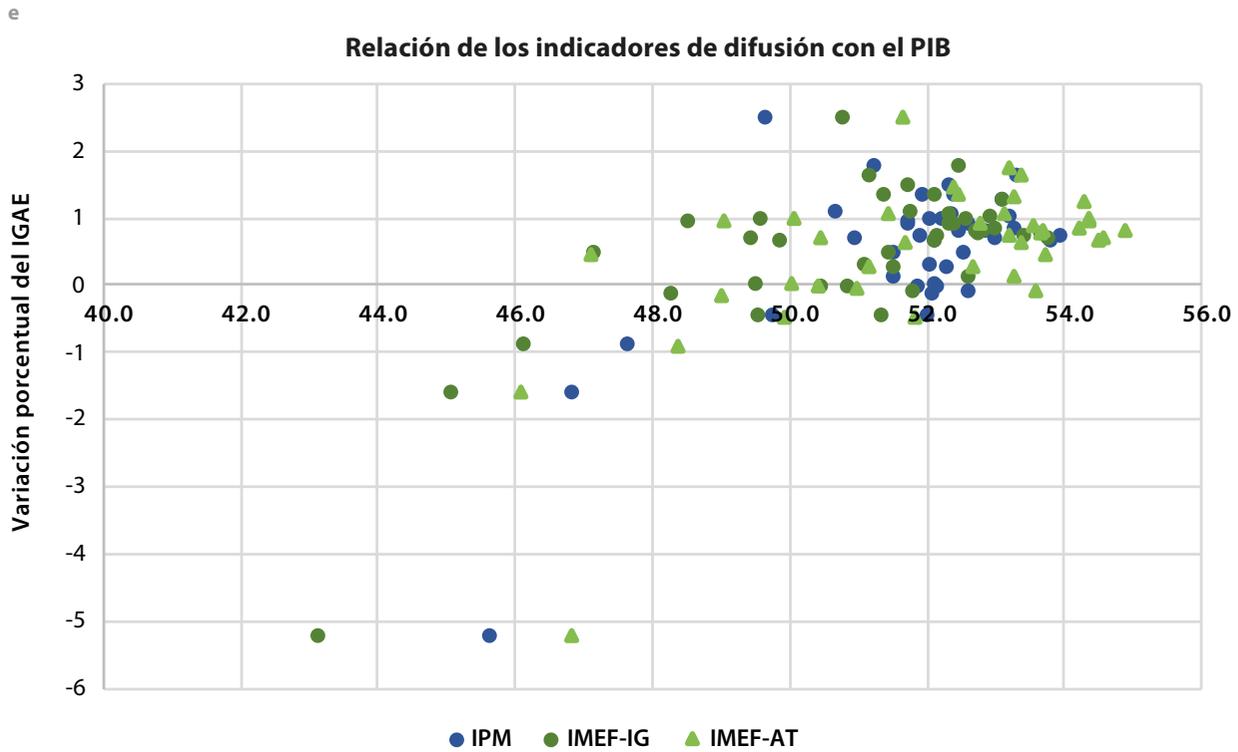
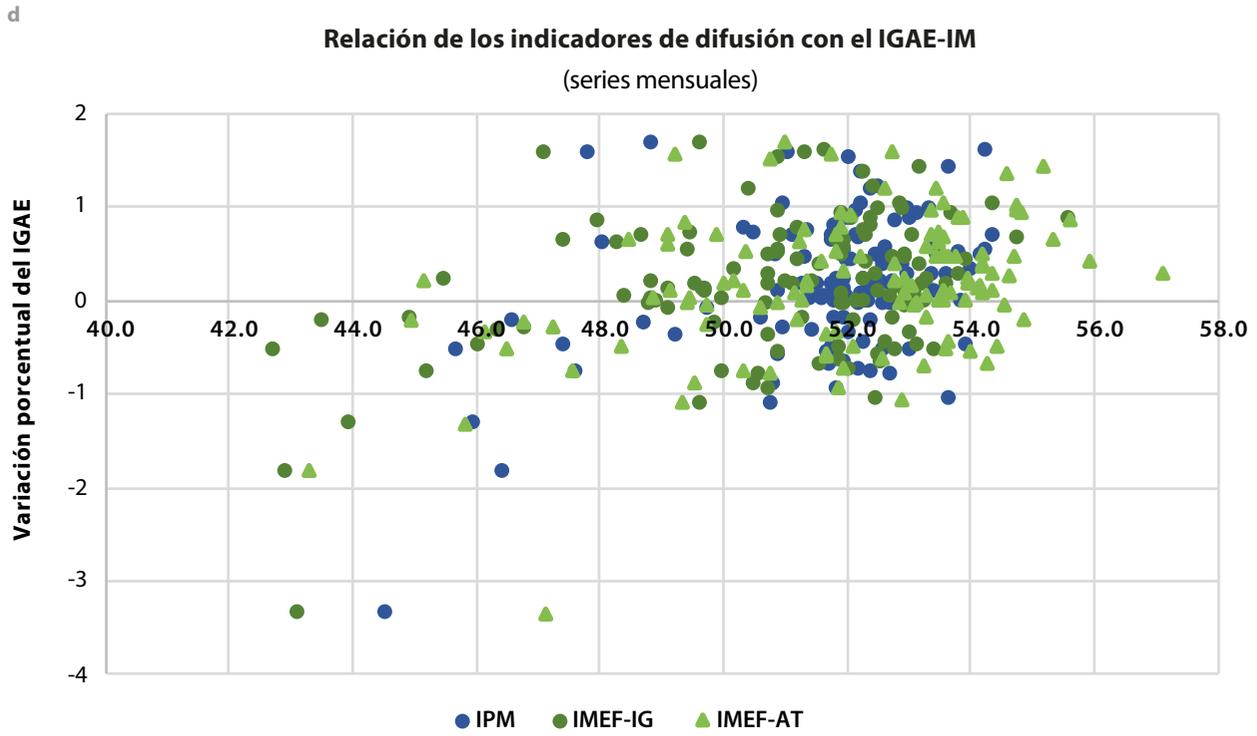


b

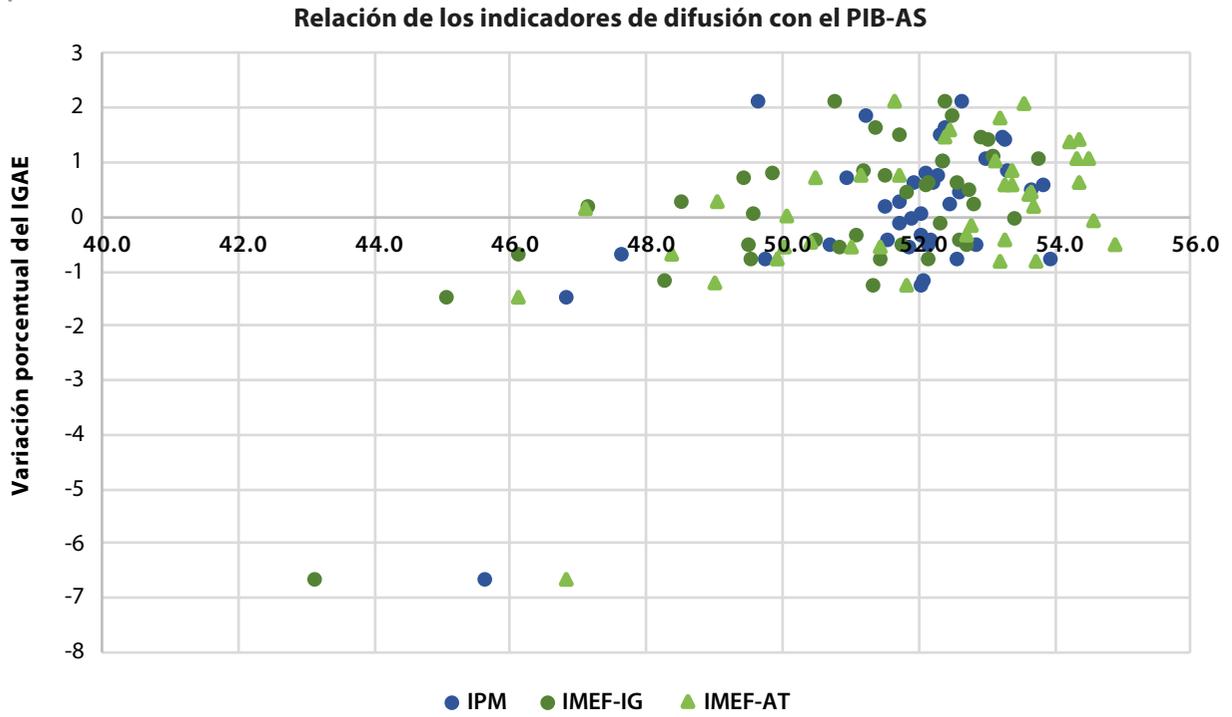


c

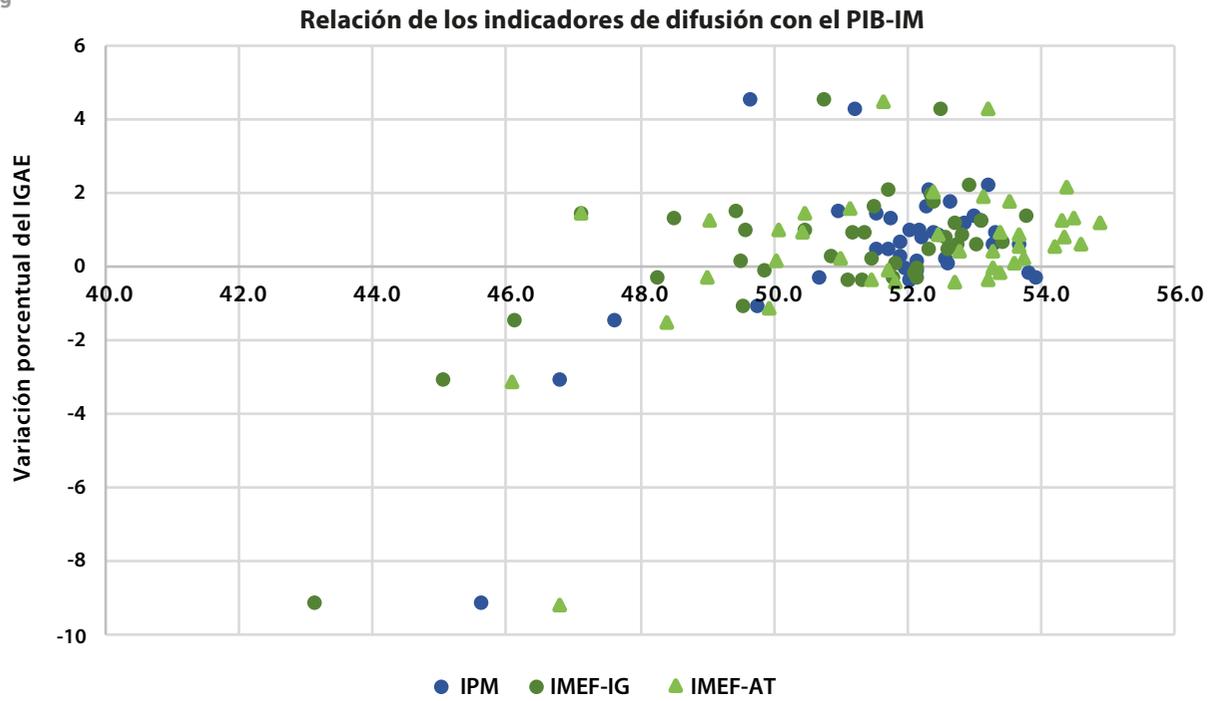




f



g



Fuente: elaboración propia con datos del INEGI e IMEF.

Índice Ponderado de Consumo de Electrodomésticos:

una propuesta de medición a partir de datos de encuestas en hogares en México

Weighted Index of Household Appliances Consumption:

a Proposal Based on Household Survey Data in Mexico

Ana Ruth Escoto Castillo* y Landy Sánchez Peña**

El consumo sustentable es parte de la nueva agenda del desarrollo; un elemento central es la promoción de consumo energético eficiente y responsable en los hogares. Se requiere tener indicadores que capturen las diferencias en el uso de energía residencial, donde los electrodomésticos que los hogares poseen juegan un papel central. Con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares ediciones de 1984 al 2014, analizamos la posesión de aparatos electrodomésticos y proponemos el Índice Ponderado de Consumo de Electrodomésticos, indicador homologado en el tiempo

Sustainable consumption is part of the new development agenda; a central element is the promotion of efficient and responsible energy consumption by households. It is necessary to have indicators that capture the differences in the use of residential energy, where household appliances play a central role. With data from the National Survey of Household Income and Expenses (1984-2014), we analyzed the ownership of household electrical appliances and proposed the Weighted Index of Household Appliance Consumption (IPCE, by its Spanish acronym); an indicator homolo-

* Universidad Nacional Autónoma de México, ana.escoto@politicas.unam.mx

** El Colegio de México, lsanchez@colmex.mx

Introducción

La sostenibilidad es parte de la agenda política de los países para avanzar en los indicadores de desarrollo. Más recientemente, los Objetivos de Desarrollo Sustentable (ODS) han planteado una serie de metas donde las condiciones ambientales son parte integral del desarrollo y de las políticas públicas (Sachs, 2012 y 2015; Sánchez Peña, 2015). En específico, el ODS 7 apunta a garantizar el acceso a una energía asequible, segura, sostenible y moderna para todos, mientras que el 12 señala la necesidad de promover un consumo energético sostenible. La búsqueda de ambos implica cambios en múltiples ejes, desde las fuentes de energía y los sistemas de provisión hasta impulsar el uso más eficiente de la energía y la reducción del desperdicio (Bizikova *et al.*, 2015, p. 9). En este último aspecto, es necesario entender mejor los usos de energía en los hogares con el fin de identificar tanto deficiencias en el acceso a la energía como posibilidades para reducir su consumo. Indicadores que permitan evaluar la evolución histórica del consumo energético y compararlo entre grupos de hogares y entre países son indispensables para el diseño de políticas medioambientales.

Sin embargo, los datos para estudiar el consumo energético a nivel de los hogares son limitados. En México, como en la mayoría de los países de la región, hay pocas fuentes especializadas en el tema. Para subsanar esta carencia, desarrollamos la propuesta del Índice Ponderado de Consumo de Electrodomésticos (IPCE) como una aproximación al consumo eléctrico de los hogares dado el alto acceso a la electricidad en la región y su creciente demanda observada.

Tomando en cuenta los antecedentes teóricos sobre elementos que inciden en el consumo energético y su intensidad, el IPCE busca ser útil para identificar las diferencias en los niveles de consumo residencial en México y dar cuenta de su cambio en el tiempo. Esto constituye un primer paso para construir un indicador de prácticas de consumo y un indicio de los cambios en los estilos de vida de los habitantes del país. Presentamos a de-

talle su implementación con datos de la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH),¹ ediciones de 1984 al 2014 en México, pero también señalamos que la información utilizada con frecuencia se recopila de otras fuentes similares en América Latina y el resto del mundo, por lo que puede extenderse su uso a otras naciones.

El IPCE parte de dos ideas centrales: 1) busca construir un indicador que no solo recoja el número de bienes, sino que también aproxime los usos y funciones de los electrodomésticos, incorporando datos de fuentes secundarias; esta propuesta no obvia la necesidad de contar con información más detallada sobre la manera en que los hogares usan la energía de forma cotidiana, pero sugiere una alternativa en la ausencia de estos datos y 2) señala que la propiedad de aparatos electrodomésticos delinea perfiles de consumo de electricidad en tanto éstos moldean la capacidad para la conservación de energía de los hogares (Escoto Castillo & Sánchez Peña, 2017; Fischer, 2008). De hecho, las políticas de eficiencia energética necesitan atender mejor la diversidad en la demanda energética de la población en razón de su nivel socioeconómico, características demográficas y habitacionales, así como ubicación geográfica (Sánchez Peña y Escoto Castillo, 2018).

Hemos dividido el presente texto en cuatro grandes secciones: en la primera se muestra la literatura concerniente a la dotación de electrodomésticos, su clasificación y usos, así como la manera en que pueden estudiarse empleando la ENIGH para el caso mexicano. La segunda expone la propuesta de la construcción de un índice de consumo de electrodomésticos, mientras la tercera analiza los resultados del IPCE y su comparación con otras mediciones, así como su asociación con el consumo de electricidad residencial. El cuarto apartado presenta la discusión de estos elementos, haciendo una breve revisión sobre las potencialidades y limitaciones del índice propuesto.

¹ Del Instituto Nacional de Estadística y Geografía (INEGI).

Electrodomésticos, su importancia y usos

Existe consenso acerca de que las fuentes modernas de energía son necesarias para el desarrollo social y económico de los países y que la falta de acceso a éstas son un buen indicador de la pobreza (Birol, 2007; Pachauri & Spreng, 2004 y 2011). En el contexto de la implementación de políticas orientadas al desarrollo y combate a la pobreza, la electricidad —fuente moderna de energía— ha aumentado su penetración en las últimas décadas en el mundo y es casi universal en Latinoamérica. Destacan naciones como Costa Rica, México y Chile, que han tenido una trayectoria más rápida que el resto de la región, pues ya para el inicio de la década de los 90 más de nueve de cada 10 personas tenía acceso a ella (World Bank, 2017). Aunque estos países tienen niveles de penetración de electricidad parecidos, mantienen fuentes en su generación distintas y, por lo tanto, diferentes implicaciones ambientales: alrededor de 90% de la electricidad de Costa Rica proviene de fuentes limpias (hidroeléctrica, renovables y nuclear); para el caso chileno, este porcentaje desciende a 40; y en México es de solo 20% (World Bank, 2017). Al mismo tiempo, su amplia penetración ha llevado consigo un mayor uso. El consumo per cápita ha aumentado, en especial en las economías emergentes que, como las latinoamericanas, han consolidado el acceso eléctrico, lo que explica su creciente demanda mundial, que se ha duplicado en dos décadas, de 9.2% en 1973 a 18% en el 2013, según datos de la Agencia Internacional de Energía (IEA, por sus siglas en inglés) (IEA, 2015).

Por lo anterior, es necesario aproximarnos a las prácticas de consumo eléctrico. Proponemos iniciar con el estudio de la dotación y uso de los electrodomésticos que permiten vincular los elementos físicos y conductuales del consumo. Estos aparatos tienen un componente *físico* —como lo mencionan Mansouri, Newborough & Probert (1996)— que proviene de una decisión pasada (la compra), pero también uno de *conducta* que se refiere al uso —medido en tiempo y frecuencia— que se les da. Este último no se encuentra en otro tipo de bienes

muebles o en las condiciones materiales de la vivienda. Esta dualidad los hace importantes de estudiar y analizar a la luz del consumo eléctrico que conllevan.

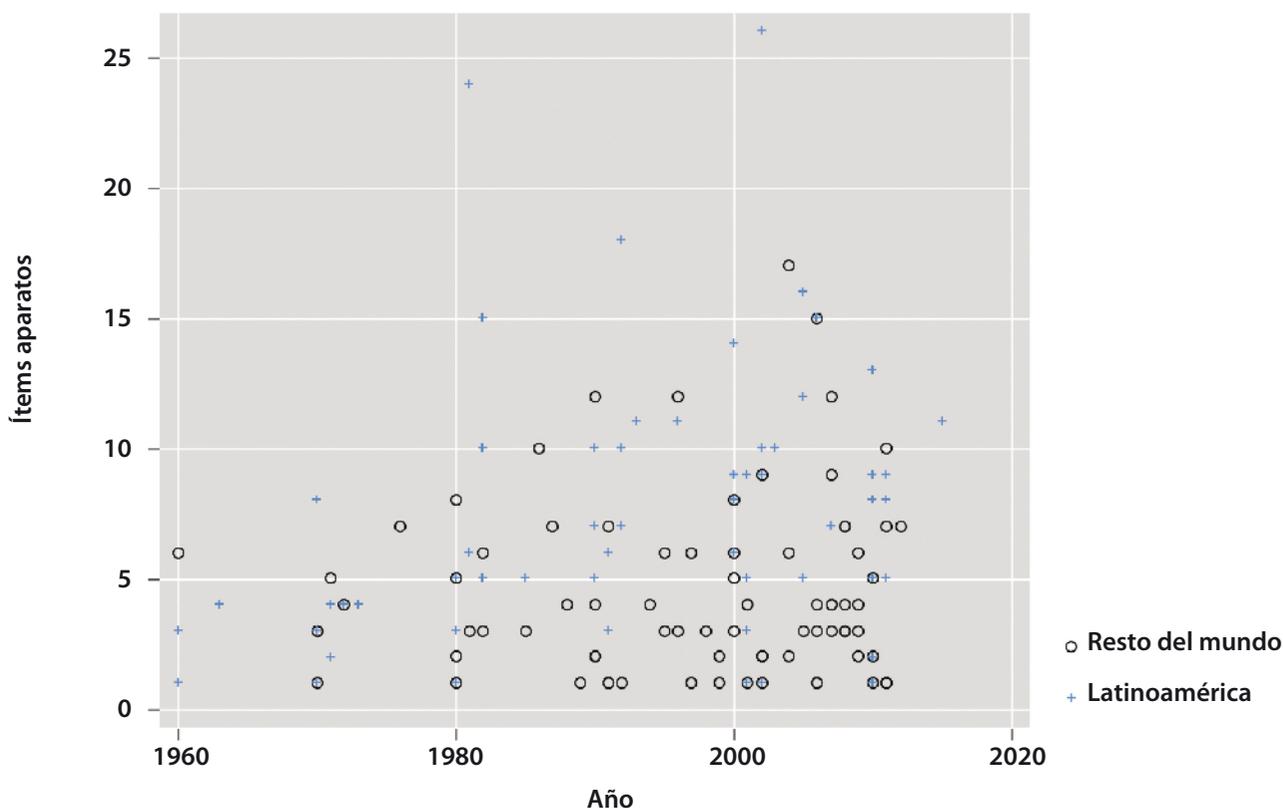
Para esta investigación, consideraremos un bien electrodoméstico como aquel aparato que desempeña una función mecánica o de transformación dentro del ámbito doméstico, concentrándonos en los que funcionan a través de la corriente eléctrica. La información disponible sobre estos bienes —cómo se estudian y por qué se han incorporado a algunas mediciones— no ha estado siempre promovida por el estudio del comportamiento energético; no obstante, existe un acervo histórico de datos sobre ellos que es útil para nuestro objetivo.

En primer lugar, los bienes del hogar —específicamente los electrodomésticos— constituyen también parte de sus acervos y han sido examinados como una porción en la definición del estrato socioeconómico, así como del monto de activos y condiciones de la vivienda. Es por ello que se presentan en distintas fuentes como los censos de población y vivienda y en las encuestas de hogares destinadas a medir tanto el bienestar como las condiciones de vida y las habitacionales. La gráfica 1 muestra cuántos ítems sobre electrodomésticos se han incorporado en 148 muestras de las 305 provenientes de los censos desde 1960, las cuales han sido recopiladas por el proyecto de Uso Público Integrado de Microdatos (IPUMS, por sus siglas en inglés) (Minnesota Population Center, 2017). Resalta la presencia de América Latina en este listado, donde países como Chile y México han registrado el uso de los equipos electrodomésticos de manera creciente a lo largo del tiempo en sus levantamientos censales.

Las encuestas en hogares también proveen información sobre los electrodomésticos, aunque las fuentes varían en los aparatos sobre los que se recaban datos y el detalle de los mismos (por ejemplo, número de dispositivos, año o modelo). Destacan los proyectos internacionales que permiten la comparabilidad entre países, como

Gráfica 1

Número de ítems sobre electrodomésticos en los censos de población y vivienda, 1960-2015



Fuente: Minnesota Populación Center. *Uso Integrado de Microdatos Público*. 2017.

las Encuestas para Medir los Estándares de Vida (LSMS, por sus siglas en inglés) promovidas por el Banco Mundial (Grosh & Glewwe, 1995), así como las Encuestas Demográficas y de Salud, impulsadas por la Agencia de Ayuda de Estados Unidos de América (USAID, por sus siglas en inglés), aunque estas últimas no son tan exhaustivas como las primeras. También, las que se realizan a nivel nacional proveen datos sobre los electrodomésticos. Para la región latinoamericana, diferentes operativos en hogares muestran información sobre su presencia, por ejemplo, las encuestas Multipropósito 2014 de Colombia, Nicaragüense de Demografía y Salud 2011 y Permanente de Hogares 2017 de Paraguay. Algunos sondeos detallan la cantidad de cada electrodoméstico, como la Encuesta de Hogares de Propósitos Múltiples 2015 de El Salvador, la Nacional de Hogares 2017 de Costa Rica y la Continua de Hogares 2017 de Uruguay. En tanto, otros ejercicios de recolección especifican

el modelo o año de compra, como la Encuesta de Condiciones de Vida 2014 de Guatemala y la ENIGH 2010 en adelante de México. En este sentido, nuestro país figura como uno donde la penetración de la electricidad es casi global, el consumo eléctrico tiene mayores implicaciones ambientales por el origen de la energía eléctrica y, además, tiene información disponible y de manera histórica sobre el uso de electrodomésticos, lo cual es propicio para la presente propuesta de análisis.²

En un segundo grupo de estudios, la información del equipamiento del hogar ha sido incorporada directa o indirectamente a las mediciones de la pobreza en México. De esta manera, se incluye el

² La estabilidad en el registro de los electrodomésticos en México está relacionada, también, con el hecho de que algunos de estos bienes emplean la definición de estratos socioeconómicos de las unidades primarias de muestreo usadas para el levantamiento de encuestas en hogares.

acceso a energía eléctrica y a las comunicaciones como una de las necesidades básicas, mientras que el Consejo Nacional de Evaluación de la Política Nacional de Desarrollo (CONEVAL) ha planteado el equipamiento del hogar como una de las medidas complementarias para esta medición multivariada³ (Boltvinik, 2012; CONEVAL, 2010; Freyermuth, Belismelis, Licona y Hernández, 2011). Otro conjunto de trabajos analiza los electrodomésticos desde una perspectiva histórica, pues han formado parte de la mecanización del trabajo doméstico. En este sentido, estos aparatos son analizados como factor que incide en el uso del tiempo al interior del hogar, contribuyendo a reducir la carga doméstica, facilitando la participación laboral femenina y la reducción de la pobreza de tiempo (Coen-Pirani, León & Lugauer, 2010; Damián, 2005; Greenwood, Seshadri & Yorukoglu, 2005; Hardymont, 1988).

Sin negar la importancia de su estudio en esos ámbitos, es notoria la poca atención que han recibido los electrodomésticos para examinar el consumo energético de los hogares en América Latina (Jones, Fuertes & Lomas, 2015). En el caso particular mexicano, tienen implicaciones ambientales con respecto a las emisiones de CO₂, pues la generación de electricidad representa una quinta parte de ellas (INECC, 2013). Además, otras fuentes generadoras de emisiones de efecto invernadero (como la leña y el carbón) han disminuido su participación en el consumo energético de los hogares, mientras que el acceso a la electricidad se ha generalizado en México conforme se ha transitado de servicios energéticos no modernos a modernos (Pachauri & Spreng, 2011; Sovacool, 2011). En este contexto, es importante cuestionarnos sobre las implicaciones medioambientales de la adquisición y uso de los electrodomésticos, así como su papel en la búsqueda del desarrollo sustentable.

En México, se ha analizado la tenencia de electrodomésticos en relación con las estimaciones nacionales de los balances de energía, así como la determinación de los posibles cambios en la de-

manda energética ante las políticas de eficiencia energética (Martínez-Montejo & Sheinbaum-Pardo, 2016; Rosas, Sheinbaum & Morillon, 2010; Rosas-Flores, Rosas-Flores & Gálvez, 2011; Sheinbaum & Dutt, 1996). No obstante, éstos no buscan examinar los determinantes del consumo a nivel hogar, lo cual es necesario para avanzar en la promoción de prácticas de reducción de la demanda energética, así como de mejoras en la eficiencia en su uso. Los hogares no solo inciden en el uso de electricidad —directa e indirectamente a través de los bienes y servicios que demandan— sino que también son espacios para la conformación de expectativas y consensos sobre las políticas energéticas a favorecerse (Sánchez, 2016), de ahí que es central entender cómo usan la energía.

Tipos y clasificaciones de los electrodomésticos

Estos aparatos pueden estudiarse por su finalidad, frecuencia e intensidad de uso. En el caso mexicano (como en otros países) hay información limitada sobre esto en las encuestas, pese a que son centrales para entender impactos ambientales y políticas energéticas para los hogares.⁴ Para subsanar el problema, en este trabajo proponemos un índice que utilice la información de dotación de electrodomésticos y emplee datos de fuentes secundarias sobre demanda energética en virtud de los fines para los que son utilizados.

Los electrodomésticos se pueden clasificar por sus condiciones físicas y uso o por combinación de ambas. La clasificación más utilizada es la de los almacenes que los venden (*gamas*), ordenada por colores, que refieren al tamaño y a la movilidad. En este sentido, la *línea blanca* necesita instalación para su uso o no se puede mover muy fácil dentro de la vivienda;⁵ el resto es considerado como pequeños electrodomésticos. En fecha reciente, con

3 En específico, el indicador sería *Porcentaje de hogares que no cuentan con televisión, refrigerador, lavadora ni estufa de gas o eléctrica*.

4 Recientemente, en noviembre del 2018, el INEGI publicó la Encuesta Nacional sobre Consumo de Energéticos en Viviendas Particulares (ENCEVI), primer instrumento especializado en consumo de energéticos residencial en México.

5 Por ejemplo, los refrigeradores no necesitan una instalación especial, pero no son fácilmente movibles y, por lo tanto, son aparatos electrodomésticos grandes.

el desarrollo de la tecnología, se hacen dos distinciones dentro de éstos: la *línea café* o de entretenimiento y los *pequeños aparatos electrodomésticos* (PAE). Estas gamas también estarían definiendo diferentes consumos por parte de los hogares y una diferencia en su uso (De-Juan-Vigaray y Vadell, 2007). Este ordenamiento empírico distingue en forma parcial a los electrodomésticos según su función y su carga eléctrica. Otros autores prefieren clasificarlos por su uso más común (Rosas *et al.*, 2010; Rosas-Flores & Gálvez, 2010; Rosas-Flores *et al.*, 2011) o bien, distinguirlos por su utilización, diferenciando a los usuarios únicos de los multiusuarios, es decir, quienes poseen más de un bien del mismo tipo en la unidad doméstica (Mansouri *et al.*, 1996).

No obstante, clasificaciones más adecuadas para capturar el consumo energético provienen de los estudios ingenieriles (*building stock models*) que reconstruyen desde abajo hacia arriba (*bottom-up*) la demanda energética de los hogares, en específico cuando se enfocan en el consumo eléctrico.⁶ Desde esta perspectiva, los electrodomésticos se distinguen por la generación, el tipo y la constancia del consumo eléctrico. Si bien no existe una forma de agrupación única, hay coincidencias importantes en los ordenamientos que diversos autores utilizan. Por un lado, algunos esfuerzos los clasifican en función de su interacción con el consumidor y la continuidad de su encendido (Firth, Lomas, Wright & Wall, 2008); esta clasificación, sin embargo, es muy demandante en información, pues la mayoría de las encuestas en hogares no incluye datos sobre las formas de uso de electricidad de los aparatos.⁷ Por su parte, otros autores proponen una basada en los *usos finales* y, por lo tanto, en el consumo eléctrico esperado de los aparatos. Widén & Wäckelgård (2010) señalan nueve usos: *electrodomésticos de enfriamiento, iluminación, para cocinar, lavar platos, lavar ropa, TV, computadora, estéreo y adicionales*. La versión de Yao y Steemers (2005)

propone un número menor de categorías, pero con la misma lógica de incluir los usos finales de los aparatos; además, estos autores retoman la lógica de comercialización de las gamas de electrodomésticos, lo cual la hace coincidir con muchas baterías de preguntas de las encuestas en hogares; para ellos, existen cinco clases: a) electrodomésticos de entretenimiento, b) para estabilizar la temperatura o enfriamiento, c) para cocinar, d) para lavar y e) una categoría miscelánea. Este trabajo retoma la clasificación de aparatos propuesta por Yao y Steemers por ser más parsimoniosa y puede ser fácilmente trasladada a la información histórica sobre equipamiento para México.

El caso mexicano desde la ENIGH

La información sobre consumo de energía es limitada. Con la publicación de los resultados de la ENCEVI se dieron a conocer datos acerca del consumo de energía en viviendas particulares, el equipamiento y las prácticas de manejo y ahorro de la energía. Esta encuesta es similar a las realizadas en otros países y ofrece posibilidades de adentrarse en cuánta energía usan los hogares y sobre sus prácticas de ahorro. Sin embargo, solo se cuenta con la edición 2018 y se desconoce si se repetirá en el futuro, lo que limita analizar la evolución en el tiempo del consumo de energéticos. Además, la ENCEVI ofrece una caracterización limitada del perfil sociodemográfico de los hogares, por lo que puede decirse poco acerca de los determinantes de dicho consumo. En México también se cuenta con el Módulo de Hogares y Medio Ambiente (MOHOMA) de la Encuesta Nacional de Hogares (ENH) del 2015 y 2017 también desarrollada por el INEGI, cuyo objetivo es conocer las prácticas o hábitos de los hogares relacionados con diversos aspectos del medio ambiente, incluyendo el manejo de residuos, así como el consumo y ahorro del agua y la energía (INEGI, 2016). El MOHOMA representa un avance en la indagación de las prácticas de consumo energético, sin embargo, no permite tener un panorama histórico por su reciente desarrollo y las preguntas incluidas no detallan lo suficiente para estimar el uso de los electrodomésticos.

6 Para una revisión más profunda de este tipo de modelos, también se pueden consultar las revisiones elaboradas por Kavgic *et al.* (2010) y Swan & Ugursal (2009).

7 No obstante, el tiempo en espera de los aparatos ha sido sujeto de investigación para las políticas públicas en el extranjero y ya ha sido legislado recientemente en México (Brown, 2001; de Almeida, Fonseca, Schlomann & Feilberg, 2011; Norma Oficial Mexicana NOM-032-ENER-2013, Límites máximos de potencia eléctrica para equipos y aparatos que demandan energía en espera. Métodos de prueba y etiquetado, 2014).

En México hay una gran cantidad de fuentes que incorporan información sobre la dotación de electrodomésticos, aunque se recoge de manera distinta entre ellas. La ENIGH es en particular útil por su amplia cobertura temporal, la extensión de los bienes encuestados y la posibilidad de vincular dichos bienes con el consumo energético de los hogares, así como un amplio número de variables sociodemográficas. No obstante, su mayor limitación es que no incluye la frecuencia de uso de los electrodomésticos;⁸ además, le falta información histórica sobre modelos o la antigüedad para poder estimar mejor el consumo energético según la tecnología empleada,⁹ que permitiría evaluar mejor la eficiencia energética de los aparatos, complementándola con fuentes adicionales, como las estimaciones de demanda energética de bienes de la Comisión Federal de Electricidad (CFE, 2014).

Los electrodomésticos como predictores del consumo y las prácticas en México

Extensos estudios a nivel mundial plantean el vínculo entre electrodomésticos y consumo eléctrico como estadísticamente significativo (Jones *et al.*, 2015). En México, algunos análisis han integrado un promedio de consumo estándar de los electrodomésticos a partir de un índice de saturación; éste es calculado como el número de aparatos más utilizados entre la cantidad de hogares. Esa medida se multiplica con una unidad promedio de consumo energético de los principales aparatos para obtener el consumo total de los hogares. Si bien esta cifra puede ser calculada por grupos poblacionales, es frecuente que el consumo sea asumido como constante a través de los grupos y, con frecuencia, también en el tiempo ante la falta de información del consumo energético específico (Rosas *et al.*, 2010; Rosas-Flores & Gálvez, 2010; Rosas-Flores *et al.*, 2011; Sheinbaum & Dutt, 1996). Por ende, si-

8 Algunos estudios sugieren que, además de la dotación y uso de electrodomésticos, se requiere información detallada sobre las prácticas y los estilos de vida de los miembros de los hogares (Sanquist, Orr, Shui & Bittner, 2012), así como de sus implicaciones para el consumo indirecto de energía (Bedir, Hasselaar & Itard, 2013).

9 Como ya se mencionó, en el 2010, el año de la compra o antigüedad del aparato fue introducido en la ENIGH.

guiendo esta metodología no se pueden establecer los cambios en los modos de vida y prácticas de los ciudadanos en México, pues son elementos endógenos al esquema del análisis.

La propuesta del índice establecerá que el equipamiento puede ser examinado en relación con otras características para evaluar las implicaciones ambientales del consumo eléctrico a nivel de hogar. Así, se podrían estudiar los otros determinantes de las prácticas de consumo, como la estructura familiar, información sobre los ocupantes dentro del hogar, el ingreso del hogar y la ubicación geográfica. Además, proponemos una homologación histórica de los bienes electrodomésticos que posibilita comparar en el tiempo su posesión y la demanda energética asociada. Ello hace posible tomar en cuenta efectos de periodo, como los de una crisis económica o la adopción de nuevas normativas y los cambios en las políticas de eficiencia energética.

Metodología: la construcción del IPCE, propuesta con base en la ENIGH

Si se busca entender cómo ha cambiado el consumo de electrodomésticos y cuáles son las implicaciones medioambientales del consumo eléctrico, se requiere construir una medida que solvete al menos dos problemas: a) cómo capturar su impacto sobre la demanda de energía cuando no se tiene información sobre frecuencia e intensidad del uso y b) su comparabilidad en el tiempo. En este artículo construimos un índice a partir de la información disponible de la ENIGH. En esta construcción, la clasificación y ponderación de los tipos de electrodomésticos da cuenta de las diversas demandas energéticas de ellos y, además, se ha llevado a cabo un proceso de armonización que permite examinar tendencias de cambio en el tiempo.

La clasificación de Yao y Steemers (2005) funciona de mejor manera para el caso mexicano al permitir una categorización que recupera las diferencias en la demanda energética de los elec-

trodomésticos, pero sin que ésta requiera información detallada sobre ellos. El cuadro 1 muestra esta clasificación con la información disponible de la ENIGH en el periodo de 1992 al 2014 para 14 aparatos que se pueden seguir a lo largo del tiempo, los cuales fueron agrupados según su uso funcional que implica prácticas y consumos distintos.¹⁰

Desde 1984 ha habido cambios metodológicos en la ENIGH, así como otros propios de las transformaciones históricas del uso y desarrollo de los electrodomésticos, aunque siempre ha considerado su cantidad en el hogar. A lo largo de los años no se han preguntado los mismos elementos ni de una manera homogénea. En los primeros levantamientos, la sección de aparatos del hogar se combinaba con otros elementos físicos del mismo, como la bomba de agua, el sistema de calefacción y otra infraestructura. En general, el cuestionario no distingue específicamente su fuente de energía, como es el caso de la estufa en algunos años, o si la máquina de coser es manual, lo cual afectaría sobre todo en las observaciones de mayor antigüedad. En la gráfica 2 se muestra el número de reactivos o aparatos que ha incluido la encuesta a lo largo de los años, según los grupos definidos en Yao y

¹⁰ En esta clasificación también se incluyen electrodomésticos para cocinar: estufas eléctricas y hornos. Éstos no pueden ser rastreados a lo largo del tiempo para el caso mexicano

Steeemers (2005) y ejemplificada con anterioridad en el cuadro 1. Se observa que existen dos categorías que se fueron ampliando desde 1984 hasta el 2006 (*entretenimiento y miscelánea*); a partir del 2008 hubo un recorte en los reactivos que afectaron a todas las categorías, menos a *lavar*, el cual no ha tenido cambios en el tiempo.

El que se haya cambiado el listado de electrodomésticos que la ENIGH captura cada año tiene una repercusión directa en los promedios de las dotaciones de los hogares, tal como se observa en la gráfica 3. Esto se aprecia sobre todo en el crecimiento importante en los promedios de dotación en los años en que se incrementaron los reactivos asociados con los electrodomésticos (2002-2006) y con su contracción en los años posteriores. La gráfica evidencia que el cambio en la posesión de electrodomésticos no puede distinguirse de la modificación en el instrumento si se emplean los conteos *crudos* que arroja la ENIGH. En contraste, la propuesta aquí desarrollada incluye una clasificación de los aparatos que refleja su uso energético, misma que parte de una homologación de los dispositivos a considerar.

En el proceso de armonización entre encuestas hay dos tipos de cambio que deben tomarse en cuenta: tecnológicos y metodológicos. Los primeros son, por ejemplo, el ítem de antena parabólica que

Cuadro 1

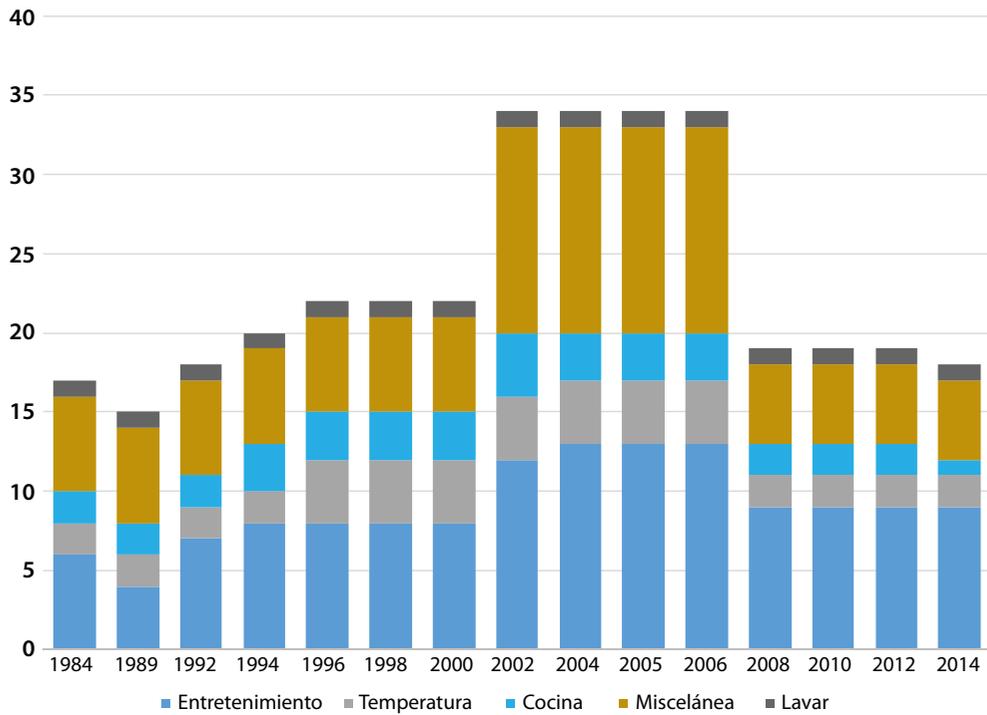
Clasificación de los electrodomésticos constantes en la ENIGH 1992-2002

Electrodomésticos electrónicos de entretenimiento	Electrodomésticos para estabilizar temperatura	Electrodomésticos para lavar	Electrodomésticos misceláneos
• Radio	• Refrigeradores	• Lavadora	• Máquina de coser
• Radiograbadora	• Ventiladores		• Licuadora
• Consola, estéreo			• Plancha
• Televisión			• Aspiradora
• Reproductor de películas (VCR o DVD)			
• Videojuegos			
• Computadoras			

Fuente: elaboración propia siguiendo a Yao y Steemers (2005).

Gráfica 2

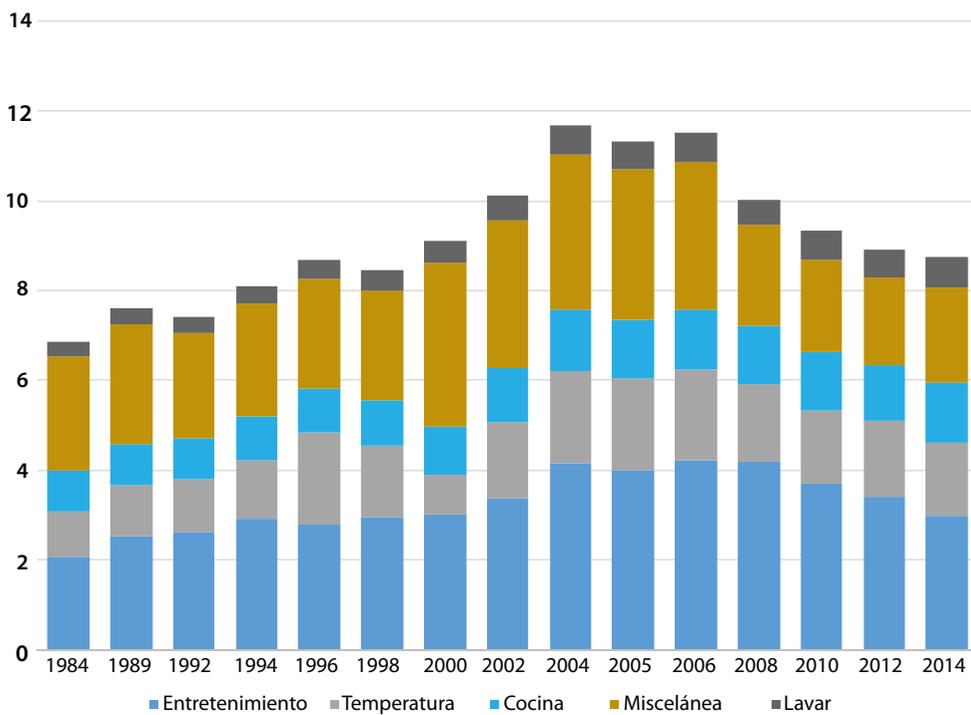
Reactivos en la ENIGH según tipo de electrodomésticos, México, 1984-2014



Fuente: elaboración propia con información de la ENIGH, 1984-2014

Gráfica 3

Promedio de aparatos en el hogar según tipo de electrodoméstico, México, 1984-2014



Fuente: elaboración propia con información de la ENIGH, 1984-2014.

se incluía en las décadas de los 80 y 90 desapareció por completo en la de los 2000, por su evidente desuso. Por otro lado, hay modificaciones por las nuevas tecnologías: en 1984 y 1989 no se preguntaba por la presencia de videojuegos, computadoras y reproductores de video o películas; bien podría aducirse que su penetración en el mercado se dio más tarde, a partir de la década de los 90; incluso, la presencia de videojuegos y computadoras en 1992 era bastante residual (menos de 2% en los hogares), aunque la de reproductores de películas alcanzaba alrededor de 30% de los hogares.¹¹ Del mismo modo, desde 1994, se puede seguir la información del horno de microondas y desde el 2002, la de impresoras y horno tostador. En

algunos casos, como se observa en el cuadro 2, se consolidaron categorías, o bien, se tomó en cuenta la evolución de la tecnología (por ejemplo, el reproductor de VHS se transformó en DVD, consola en minicomponente, etcétera). De esta manera, se puede dar una continuidad según los fines últimos de uso.

Por otro lado, los cambios metodológicos apuntan a modificaciones en el diseño y conceptualización de la fuente. Una transformación notoria fue la disminución de reactivos en la información sobre electrodomésticos que la ENIGH recolecta. Implementado desde el 2008, este cambio no responde a evoluciones tecnológicas sino a decisiones metodológicas en torno a cómo medir la dotación y bienestar de los hogares. En términos del análisis energético, estas decisiones no son menores.

¹¹ En este sentido, se habla de una homologación parcial desde 1984 si se considera la poca penetración de estas tecnologías de entretenimiento y una plena desde 1992.

Cuadro 2

Reactivos en la ENIGH y su seguimiento en el tiempo, México, ediciones de 1984 al 2014

Tipo de aparato	Ítems en la encuesta	Desde	Clasificación
Radio	Radios	1984	Entretenimiento
Música (de menos movilidad)	Estéreos, consolas, modulares, minicomponentes	1984	Entretenimiento
Radiograbadora (movible)	Radiograbadoras (con o sin CD)	1984	Entretenimiento
TV	Televisiones (blanco y negro o color), pantallas	1984	Entretenimiento
Ventiladores	Ventiladores	1984	Temperatura
Máquinas de coser	Máquinas de coser	1984	Miscelánea
Estufa ^a	Estufas de gas Estufas de otro combustible Estufa de gas o eléctrica	1984	Cocina
Refrigerador	Refrigerador	1984	Temperatura
Licadoras	Licadora	1984	Miscelánea
Plancha	Plancha	1984	Miscelánea
Lavadora	Lavadora, lavadora automática	1984	Lavar
Aspiradora	Aspiradora	1984	Miscelánea
Videograbadoras y/o DVD	Videocasetera, DVD	1992	Entretenimiento
Consolas videojuegos	Videojuegos	1992	Entretenimiento
Computadoras	Computadoras	1992	Entretenimiento
Horno de microondas	Horno de microondas	1994	Cocina
Tostador	Tostador	2002	Cocina
Impresora	Impresora	2002	Entretenimiento

^a Este ítem no fue incluido en los análisis posteriores, porque su principal fuente de energía es el gas y no la electricidad.

Fuente: elaboración propia con información de la ENIGH, 1984-2014.

Se aprecia que desde el 2008 se dejó de preguntar por aires acondicionados y sistemas de calefacción (ver cuadro 3); si bien solo un pequeño porcentaje de hogares declaraba contar con éstos en el 2006, su impacto en la demanda energética fue alto.¹² En este sentido, la construcción de un índice de electrodomésticos homologado sacrifica información

importante disponible en algunos años, en aras de la comparabilidad en el tiempo. En las secciones sucesivas de ejemplificación de la propuesta hemos utilizado la homologación desde 1992, donde creemos que tenemos datos históricos extensos y una comparabilidad sustantiva. Considerando las dificultades señaladas de forma previa, en la información de la ENIGH existen 15 aparatos que se pueden seguir a lo largo del tiempo, de los cuales 14 son electrodomésticos.

A partir de estos procedimientos de homologación y clasificación de electrodomésticos, en el siguiente apartado se analizan algunos ejemplos de índices y su relación con el consumo energético.

12 Algunos elementos específicos a considerar: 1) no hay comparabilidad en el tiempo del indicador de aire acondicionado o *cooler* por variaciones en el cuestionario; 2) en México, la cocina utiliza sobre todo gas; incluso, ésta fue registrada desde 1984 hasta el 2002 como *estufa de gas* o *estufa de otros combustibles*; fue en el 2004 cuando se especificó como *estufa de gas* o *eléctrica* y el registro de estufas de otros combustibles desapareció; la participación de estufas de otros combustibles desde 1992 es marginal, menos de 1% de la muestra; y 3) una nota en los documentos metodológicos de la ENIGH señala que la captación de las lavadoras no puede ser comparada en el 2008 con la serie histórica anterior ni con la posterior.

Cuadro 3

Continúa

Equipamiento en los ítems en la ENIGH, México, ediciones de 1984 al 2014

Aparato	1984	1989	1992	1994	1996	1998	2000	2002	2004	2005	2006	2008	2010	2012	2014
Aire acondicionado o cooler					1	1	1	1	1	1	1				
Abrelatas								1	1	1	1				
Antena parabólica			1	1	1	1	1	1							
Aparato calefactor (no sistema)					1	1	1	1	1	1	1				
Aspiradora	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Batidora								1	1	1	1				
Cafetera								1	1	1	1				
Computadora			1	1	1	1	1	1	1	1	1	1	1	1	1
Consola o modular	1														
DVD								1	1	1	1	1	1	1	1
Escáner, quemador, módem y otros								1	1	1	1				
Estéreo, modular, minicomponente o consola										1	1				

Equipamiento en los ítems en la ENIGH, México, ediciones de 1984 al 2014

Aparato	1984	1989	1992	1994	1996	1998	2000	2002	2004	2005	2006	2008	2010	2012	2014
Estéreo, modular o consola								1							
Estéreo, modular, minicomponente o consola									1						
Estéreo, modular, minicomponente												1	1	1	1
Estufa de gas	1	1	1	1	1	1	1	1							
Estufa de gas o eléctrica									1	1	1	1	1	1	1
Estufa de otro combustible	1	1	1	1	1	1	1	1							
Exprimidor de jugos								1	1	1	1				
Grabadora	1	1	1	1	1	1	1	1							
Horno eléctrico								1	1	1	1				
Impresora								1	1	1	1	1	1	1	1
Lavadora	1	1	1	1	1	1	1	1	1	1	1				
Lavadora automática												1	1	1	1
Licuadora	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Máquina de coser	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Microondas				1	1	1	1	1	1	1	1	1	1	1	1
Molino de mano	1	1	1	1	1	1	1	1	1	1	1				
Otros	1	1	1	1	1	1	1	1	1	1	1				
Plancha	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Radio	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Radiograbadora												1	1	1	1
Radiograbadora con reproductor de CD									1	1	1				
Radiograbadora sin reproductor de CD									1	1	1				
Refrigerador	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Reproductor de CD				1	1	1	1	1	1	1	1				
Sandwichera								1	1	1	1				
Tocadiscos	1														
Tocadiscos, consola o modular		1	1	1	1	1	1								

Equipamiento en los ítems en la ENIGH, México, ediciones de 1984 al 2014

Aparato	1984	1989	1992	1994	1996	1998	2000	2002	2004	2005	2006	2008	2010	2012	2014
Tostador								1	1	1	1	1	1	1	1
TV		1	1	1	1	1	1					1	1	1	1
TV B/N	1							1	1	1	1				
TV color	1							1	1	1	1				
Ventilador	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
Videocasetera			1	1	1	1	1	1	1	1	1	1	1	1	1
Videojuegos			1	1	1	1	1	1	1	1	1	1	1	1	1

Fuente: elaboración propia con información de la ENIGH, 1984-2014.

Resultados

Indicadores de electrodomésticos

Esta sección revisa cuatro formas de analizar el consumo de éstos: 1) un conteo crudo de los aparatos, 2) un índice sumatorio simple homologado, 3) un índice construido a partir del análisis por componentes principales y 4) el IPCE.

El índice sumatorio simple totaliza los electrodomésticos de la canasta homologada.¹³ Este tipo de índices se ha utilizado en algunos estudios como un determinante más del consumo eléctrico, con un efecto estadísticamente significativo (Jones *et al.*, 2015, p. 912). En la gráfica 4, podemos ver la diferencia entre la dotación que provee un índice homologado y el presentado en la gráfica 3. Solo en 1992 y en el 2008 las diferencias entre ambos no son estadísticamente significativas, pero para el resto de los años optar por uno u otro cambia de forma notoria lo que se señale sobre el consumo de los hogares: si bien el índice simple homologado subestima el consumo total de electrodomésticos, el análisis del cambio en el tiempo es más confiable, pues no refleja los cambios metodológicos y tecnológicos discutidos de manera previa.

¹³ Para los siguientes ejercicios, ocuparemos la homologación que deviene desde 1992, ver cuadro 2.

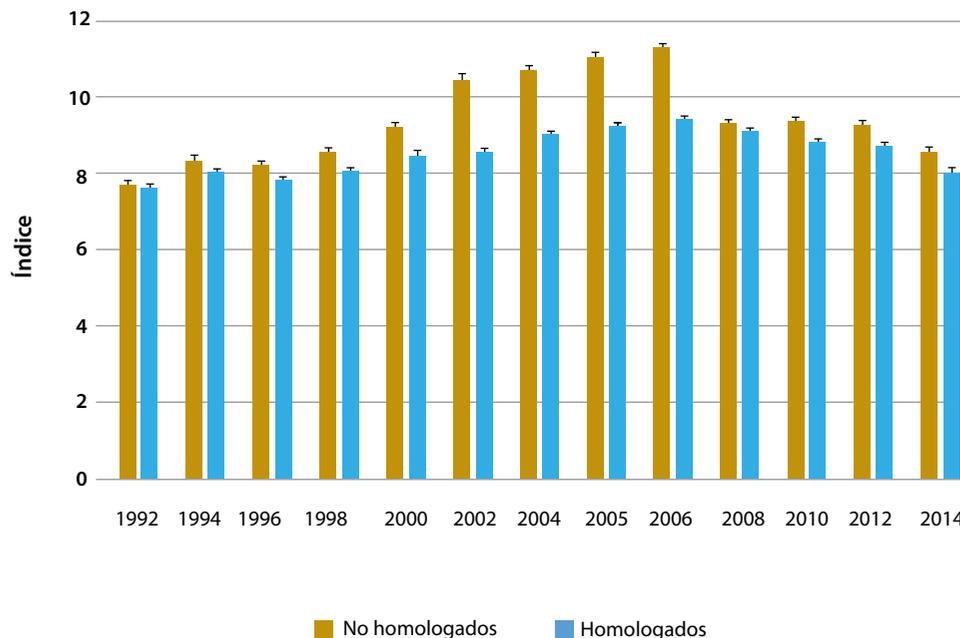
Sin embargo, este índice meramente aditivo no da cuenta de la diferenciación en los usos y demanda energética de los aparatos. Las clasificaciones de los electrodomésticos responden a la necesidad de entender sus usos, niveles de penetración o saturación de los mismos, así como distintos comportamientos a través de los grupos poblacionales; esto es útil en especial para establecer procesos de difusión de prácticas de consumo. Por ejemplo, si la clasificación propuesta se examina a través de quintiles de ingreso, muestra cómo los electrodomésticos de mayor consumo (*lavar y temperatura*) están menos presentes en los deciles más bajos, mientras que los de entretenimiento son los que tienen más presencia a lo largo de los estratos de ingreso.

Requerimos un indicador que dé cuenta de las diferencias en términos de la demanda energética esperada de los electrodomésticos. Una estrategia común en la construcción de índices para reducir la multidimensionalidad es un proceso a través de la estimación de componentes principales o un modelo factorial. Estas técnicas se basan en las varianzas de las variables que se estudian dentro de una población; de este modo, estarían muy cercanas a estimar la penetración de los electrodomésticos y sopesaría aún más, aquellos que son más comunes entre sí. Se estimó un índice por componentes principales utilizando la clasificación previa en usos, el cual explica 74% de la varianza de las do-

Gráfica 4

Promedio de electrodomésticos en el hogar, México, 1992-2014

(intervalos de confianza a 95%)



Fuente: elaboración propia con información de la ENIGH, 1992-2014.

taciones de electrodomésticos.¹⁴ Este método no da cuenta de las variaciones de consumo eléctrico, sino de las variaciones dentro de la población, pues sopesa por las categorías más frecuentes porque son las que comparten mayor varianza conjunta; por ejemplo, se asignan mayores pesos a los electrodomésticos de entretenimiento que no son necesariamente los más demandantes de energía, pero sí los que han aumentado su saturación. Sin duda, si se trata de medir la penetración de los aparatos, este tipo de índices funciona, pero no así para el consumo energético.

Otra alternativa —y la que este texto considera como la más adecuada— es recuperar las clasificaciones de electrodomésticos revisadas en tanto proveen información sobre la demanda energética esperada, pues cada grupo hace referencias a usos, frecuencias y prácticas distintas. En este sentido, Yao y Steemers (2005) señalan que los usos más intensivos en consumo eléctrico refieren a los utiliza-

dos para enfriar (*temperatura*) y lavar, mientras que los dedicados al entretenimiento y los misceláneos representan menos del consumo total. Este patrón de distinción entre los electrodomésticos coincide con la información recogida por la CFE (2014) para el consumo en México.

Como ya se mencionó, la ENIGH no proporciona datos sobre el modelo o antigüedad de los aparatos sino hasta fecha muy reciente (2010), por lo que no es posible estimar con precisión la demanda eléctrica de los aparatos y es difícil vincular la información específica proporcionada por la CFE (2014). Sin embargo, tenemos alguna información disponible sobre la demanda eléctrica de ellos a través de la clasificación y sus elementos físicos y de fines de uso de los aparatos. Por lo anterior, planteamos una ponderación alejada de los tradicionales métodos multivariados basados en la varianza. Nuestra propuesta consiste en que se deben crear índices ponderados, a manera de una variable *proxy*, a partir de sopesar más las categorías o clases de aparatos que demandan un mayor

¹⁴ Estimado a partir de la matriz de covarianza, con índice KMO de 0.93.

consumo por sus cualidades físicas y de uso finales. Se construye el IPCE con base en la demanda esperada:

$$IPCE = w_1clase_1 + w_2clase_2 + \dots + w_nclase_n \quad (1)$$

donde:

w_i refiere a los pesos de cada clase.
 $clase_i$, a las clases que se pueden establecer en la dotación de electrodomésticos.
 i , de 1 a n , número de clases.

En la medida en que se adapte la información disponible con clasificaciones más precisas sobre las cargas eléctricas que demandan los aparatos, el índice se puede ir ampliando en términos del número de clases e, incluso, estableciendo subclases. En este caso, hemos utilizado una clasificación de cuatro clases ($n = 4$), para el caso mexicano:

$$IPCE = w_1entretenimiento + w_2temperatura + w_3lavar + w_4misc \quad (2)$$

Los pesos pueden incluir información sobre la carga o demanda eléctrica y ello puede permitir recolectar mejor el consumo eléctrico. La manera en que éstos se calculen puede ir siendo más exacta en la medida que los datos disponibles lo sean, es decir, pueden estimarse de forma más precisa cuando se cuente con información acerca del año o modelo de los aparatos. Para ejemplificar la propuesta, se asumió la igualdad entre las categorías de más consumo; en esta clasificación serían *temperatura* y *lavar* y las de menor, *miscelánea* y *entretenimiento*, pero sabiendo que las de mayor gasto deben tener más peso, por lo tanto, se estimó el promedio ponderado con el conjunto de ponderaciones que se muestra en la ecuación 3, asumiendo que la suma de todos los pesos es 1. Por lo cual, la ecuación 2 se reescribe, tomando una de las combinaciones posibles de ponderaciones que cumple con las condiciones que provienen del

conocimiento del funcionamiento de los electrodomésticos según sus clases:

$$IPCE = 0.2entretenimiento + 0.3temperatura + 0.3lavar + 0.2misc \quad (3)$$

Esto nos permitiría una mejor aproximación a las implicaciones ambientales de los electrodomésticos que un índice simple que solo sume el número de aparatos por hogar, o bien, uno multivariado sopesado por la penetración de los electrodomésticos.

En este ejercicio se plantea una ponderación igual para todos los hogares y fija en el tiempo, pero podrían proponerse otras estimaciones, por ejemplo, podría variar a lo largo de diferentes dimensiones dentro del análisis, o bien, ajustarse los pesos para reflejar los cambios en la eficiencia energética de los electrodomésticos o en variaciones climáticas que incidan sobre la demanda.

Ilustración de la relación entre electrodomésticos y consumo

En este apartado ahondaremos en la asociación del índice con el consumo eléctrico del hogar. Este último se calculó de manera exógena a los electrodomésticos, derivado a partir del gasto energético reportado.¹⁵ Esta comparación también es relevante, no solo por el objetivo de analizar las consecuencias ambientales del consumo, sino también para la comparación más clara de los índices, pues al mantener distintas escalas, éstos no pueden ser estrictamente comparables si no es en relación con otro elemento. La gráfica 5 muestra la evolución de la correlación de Pearson de los cuatro indicadores y el consumo eléctrico. Cualquiera de sus formas

15 Para estimar éste, se transformó el gasto monetario en electricidad a su poder calórico (expresado en gigajoules), dividiendo el gasto entre el precio de uso residencial de electricidad de cada año. Estas unidades monetarias fueron transformadas a kilowatt-hora (KWh), que luego se estandarizaron a gigajoules. Cabe señalar que, desde el 2010, la forma de recolección del gasto en electricidad fue modificado en la Encuesta, pues hoy se lleva a cabo mediante la consulta de recibos; esto se refleja en una reducción del consumo eléctrico para todos los grupos en dicho año, lo cual afecta la comparabilidad en el tiempo de los niveles promedio de consumo eléctrico.

presenta una correlación positiva y relativamente alta del consumo eléctrico del hogar. Esta relación disminuye de forma notable en las últimas dos observaciones, para todos los indicadores.¹⁶

Los resultados para el total de electrodomésticos sin homologar mantienen más altas correlaciones que el resto, pero esta relación también aumenta en los años donde metodológicamente se amplió la canasta de equipamiento en la Encuesta, lo cual pone en duda la comparabilidad de esta relación a lo largo del tiempo; para estudios intertemporales, este indicador no sería el más idóneo. El índice sumatorio homologado y el ponderado mantienen un comportamiento similar; éste mantiene una correlación ligeramente más alta. Al introducir la información disponible sobre las cargas eléctricas, mejora la capacidad predictiva, lo cual podría in-

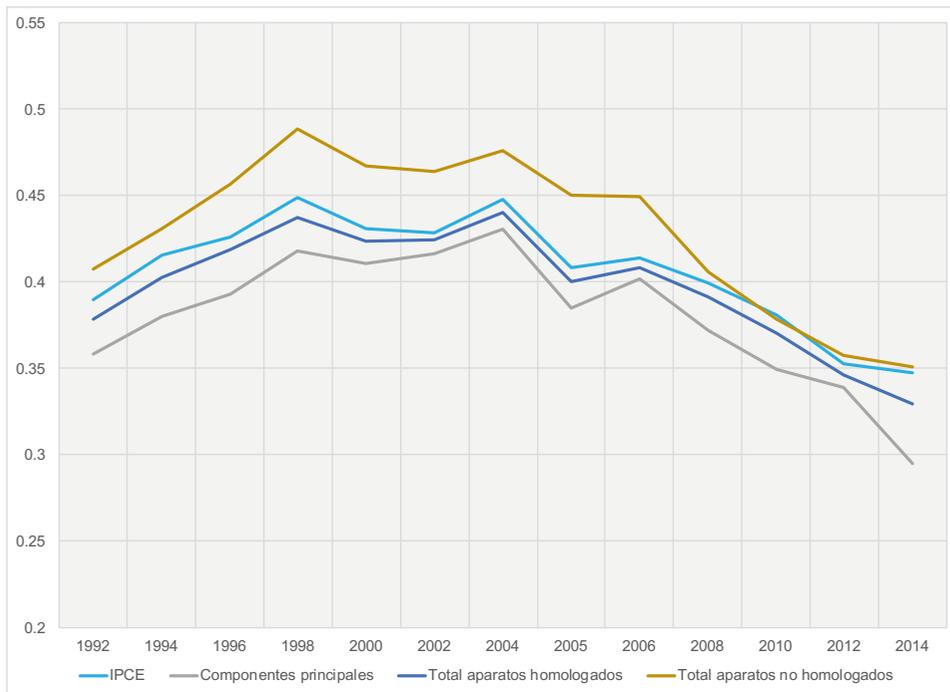
dicar que mientras más información se incorpore al registro de los electrodomésticos, mejoraría la estimación de su relación con respecto al consumo energético. Por su parte, el índice calculado con componentes principales se mantiene por debajo de todos alejándose en los primeros y en los últimos años, lo que puede estar explicado por la pérdida de información de la varianza y porque la estructura de pesos se aleja de los patrones de consumo energético.

Sin duda, la relación descrita establece el poder predictivo de los electrodomésticos sobre el consumo eléctrico y, dado que las emisiones están relacionadas con éste, los aparatos pueden dar respuesta de la contribución de los hogares a las emisiones de efecto invernadero. La ventaja del índice y la clasificación con la que se construye es que permite observar cuáles componentes de la dotación de electrodomésticos están permeando los cambios en el consumo energético.

¹⁶ Debe señalarse que, a partir del 2010, las estimaciones del consumo eléctrico se realizan mediante la consulta de recibos y no por la declaración del usuario, lo que ha disminuido las estimaciones tanto del gasto como del presente cálculo de energía en electricidad.

Gráfica 5

Correlación del consumo eléctrico por hogar en gigajoules (GJ) y diversos índices de aparatos electrodomésticos



Fuente: elaboración propia con información de la ENIGH, 1992-2014.

No obstante, aun cuando la relación es sustantiva y estable en el tiempo, no explica por completo el consumo energético. Existe una mejora al incorporar elementos de la carga eléctrica demandada por los aparatos; como ya se señaló, la dotación física debe ser complementada por el elemento conductual de los electrodomésticos. Por lo tanto, a este análisis deben incorporarse elementos que den cuenta de la organización de los hogares, de las condiciones físicas de las viviendas y de los aspectos socioespaciales, elementos que mejorarían su capacidad explicativa. Además, su inclusión se hace más importante debido a que estos elementos no se han mantenido exentos de variaciones en las últimas décadas.¹⁷

Discusión

En este documento establecemos la necesidad de incorporar medidas de los electrodomésticos que den cuenta de sus usos diferenciados y sus consumos. La propuesta se concentra en introducir una diferenciación de los bienes de acuerdo con su demanda eléctrica y uso. El ejercicio muestra cómo la incorporación de información sobre ellos a través de clasificaciones y ponderaciones mejora la estimación de la relación del consumo eléctrico, a la par que evidencia la necesidad de mejorar la información disponible sobre los aparatos que están siendo utilizados en las viviendas. En específico, el IPCE es consistente con la literatura que vincula a la tenencia de aparatos electrodomésticos con el consumo eléctrico. Aun cuando hubo un cambio en la medición del gasto en electricidad en el 2010, la correlación a lo largo de los años se mantuvo por encima de 0.4, lo que explicaría alrededor de 16% de la varianza, lo cual es consistente con la extensa revisión realizada por Jones *et al.* (2015) a nivel mundial. Asimismo, los resultados permiten avanzar en la investigación sobre la demanda de electricidad, tanto en términos de brindar una alternativa para estimarla de manera indirecta a partir de los electrodomésticos como de abrir la puerta para

considerar su crecimiento a futuro a partir de supuestos sobre la adopción de dichos aparatos. En este sentido, futuros trabajos pueden avanzar la propuesta aquí desarrollada al ajustar las ponderaciones con mediciones de la demanda efectiva de los aparatos o estimaciones de cambios en la eficiencia. Sin duda, estas propuestas pueden ser complementadas con otras estrategias analíticas y estadísticas, como el cálculo de perfiles latentes de uso de la energía (Escoto y Sánchez, 2017), o bien, ecuaciones estructurales que examinen los determinantes directos e indirectos del consumo eléctrico.

Esta línea de investigación es en particular relevante en relación con las políticas de eficiencia energética, en especial aquéllas orientadas a la adopción de nuevas tecnologías y de prácticas de ahorro. En México, la mayoría de los indicadores de eficiencia están basados en promedios per cápita nacionales de consumo residencial (Martínez-Montejo & Sheinbaum-Pardo, 2016; Rosas *et al.*, 2010; Rosas-Flores & Gálvez, 2010; Rosas-Flores *et al.*, 2011; Sheinbaum & Dutt, 1996). Éstos pueden mejorarse estimando perfiles de consumo a partir de la posesión de electrodomésticos y características sociodemográficas.

Dado que el IPCE distingue por el uso final de los electrodomésticos, puede servir para indagar sobre el potencial efecto que la difusión de la tenencia de ciertos aparatos puede tener en el consumo eléctrico o sobre el impacto de medidas de eficiencia energética; por ejemplo, puede considerar escenarios donde continúe el crecimiento de los bienes de entretenimiento y misceláneos a lo largo de los estratos socioeconómicos y estimar sus potenciales impactos en la demanda energética. Asimismo, los resultados del índice podrían combinarse con proyecciones sobre la adopción de medidas de eficiencia energética obtenidos en otras investigaciones para estimar el ahorro energético potencial.

Estas estimaciones indirectas del consumo energético son soluciones frente a la limitada información sobre los usos de la energía de los hogares,

¹⁷ Para considerar cuánto varía el índice por estrato socioeconómico y otras características sociodemográficas, ver Escoto y Sánchez (2017) y Escoto, Sánchez y Guadián (2017).

pero posibilitan tener indicadores a lo largo del tiempo y a distintas escalas geográficas, a la par de que pueden ser utilizadas con información que con frecuencia se incluye en encuestas en hogares. Ello hace posible que este tipo de índices pueda dar cuenta de los cambios del consumo eléctrico en relación con la evolución de la población, en específico, de las formas de organización familiar, la urbanización y los estilos de vida (Izazola, 2014; Medina-Ross, Mata-Sandoval & López-Pérez, 2005; Mercado García, 2014; Sánchez Peña, 2014). Los resultados permiten establecer diferencias entre los hogares en términos de su consumo energético, estableciendo que la tenencia de los electrodomésticos es un indicador importante para la estadística ambiental energética en México, sobre todo en términos de la necesidad de considerar la sustentabilidad de las prácticas de los hogares.

Fuentes

- Bedir, M.; E. Hasselaar & L. Itard. "Determinants of electricity consumption in Dutch dwellings", en: *Energy and Buildings*. 58, 2013, 194-207 (DE) <https://doi.org/10.1016/j.enbuild.2012.10.016>
- Biról, F. "Energy Economics: A Place for Energy Poverty in the Agenda?", en: *The Energy Journal*. 28(3), 2007, 1-6.
- Bizikova, L.; L. Pinter; G. Huppe; H. Schandl; C. Arden-Clarke; S. Averous; A. Mansion & C. O'Connor. *Sustainable Consumption and Production Indicators for the Future SDGs (UNEP Discussion Paper)*. Paris, France, UNEP, 2015 (DE) <http://www.iisd.org/library/sustainable-consumption-and-production-indicators-future-sdgs>
- Boltvinik, J. "Treinta años de medición de la pobreza en México. Una mirada desde Coplamar", en: *Estudios Sociológicos*. 30, 2012, 83-110.
- Brown, M. A. "Market failures and barriers as a basis for clean energy policies", en: *Energy Policy*. 29(14), 2001, 1197-1207 (DE) [https://doi.org/10.1016/S0301-4215\(01\)00067-2](https://doi.org/10.1016/S0301-4215(01)00067-2)
- CFE. *¿Cuál es el consumo de los principales aparatos electrodomésticos?* Sección Preguntas frecuentes. 10 de noviembre de 2014 (DE) <http://www.cfe.gob.mx/Lists/PreguntasFrecuentesTransp/DispForm.aspx?ID=16&ContentTypeld=0x01004DB3C7433B9CCA4CA541A3D199DE5CA4>, consultado el 20 de agosto de 2016.
- Coen-Pirani, D.; A. León & S. Lugauer. "The effect of household appliances on female labor force participation: Evidence from microdata", en: *Labour Economics*. 17(3), 2010, 503-513 (DE) <https://doi.org/10.1016/j.labeco.2009.04.008>
- CONEVAL. *Metodología para la medición de la pobreza multidimensional en México*. 2010 (DE) http://www.coneval.org.mx/Informes/Coordinacion/INFORMES_Y_PUBLICACIONES_PDF/Metodologia_Multidimensional_web.pdf
- Damián, A. "La pobreza de tiempo. El caso de México", en: *Estudios Sociológicos*. 23(69), 2005, 807-843.
- De Almeida, A.; P. Fonseca; B. Schlomann & N. Feilberg. "Characterization of the household electricity consumption in the EU, potential energy savings and specific policy recommendations", en: *Energy and Buildings*. 43(8), 2011, 1884-1894 (DE) <https://doi.org/10.1016/j.enbuild.2011.03.027>
- De-Juan-Vigaray, M. D. y J. B. G. Vadell. "Distribución de electrodomésticos", en: *Distribución y Consumo*. 17(91), 2007, 103-116.
- Escoto Castillo, A. & L. Sánchez Peña. "Diffusion of Electricity Consumption Practices in Mexico", en: *Social Sciences*. 6(4), 2017, 144 (DE) <https://doi.org/10.3390/socsci6040144>
- Firth, S.; K. Lomas; A. Wright & R. Wall. "Identifying trends in the use of domestic appliances from household electricity consumption measurements", en: *Energy and Buildings*. 40(5), 2008, 926-936 (DE) <https://doi.org/10.1016/j.enbuild.2007.07.005>
- Fischer, C. "Feedback on household electricity consumption: a tool for saving energy?", en: *Energy Efficiency*. 1(1), 2008, 79-104 (DE) <https://doi.org/10.1007/s12053-008-9009-7>
- Freyermuth, J. R. D.; G. T. Belismelis; G. H. Licona & V. P. Hernández. "Metodología para la medición multidimensional de la pobreza", en: *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*. 2(1), 2011, 36.
- Greenwood, J.; A. Seshadri & M. Yorukoglu. "Engines of liberation", en: *The Review of Economic Studies*. 72(1), 2005, 109-133.
- Grosh, M. E. & P. Glewwe. *A guide to living standards measurement study surveys and their data sets*. Vol. 120. World Bank Publications, 1995.
- Hardyment, C. *From mangle to microwave: the mechanization of household work*. Cambridge, UK; Oxford, UK; New York, NY, USA; Polity Press; Basil Blackwell; 1988.
- IEA. *World energy outlook 2015*. Paris, OECD, 2015 (DE) <http://www.worldenergyoutlook.org/weo2015/>
- INEGI. *Módulo de Hogares y Medio Ambiente 2015*. 2016.
- Instituto Nacional de Ecología y Cambio Climático (INECC). *Inventario Nacional de Emisiones de Gases de Efecto Invernadero 1990-2010*. 2013 (DE) http://www.inecc.gob.mx/descargas/cclimatico/inf_inegei_public_2010.pdf
- Izazola, H. "Hogares y medio ambiente. Reflexiones desde la investigación sociodemográfica", en: *La Estadística Ambiental en México*. El Colegio de México, 2014, pp. 261-286 (DE) <http://www.jstor.org/stable/j.ctt14jxnjz.15>
- Jones, R. V.; A. Fuertes & K. J. Lomas. "The socio-economic, dwelling and appliance related factors affecting electricity consumption in domestic

- buildings", en: *Renewable and Sustainable Energy Reviews*. 43, 2015, 901-917 (DE) <https://doi.org/10.1016/j.rser.2014.11.084>
- Kavgic, M.; A. Mavrogianni; D. Mumovic; A. Summerfield; Z. Stevanovic & M. Djurovic-Petrovic. "A review of bottom-up building stock models for energy consumption in the residential sector", en: *Building and Environment*. 45(7), 2010, 1683-1697 (DE) <https://doi.org/10.1016/j.buildenv.2010.01.021>
- Kavousian, A.; R. Rajagopal & M. Fischer. "Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior", en: *Energy*. 55, 2013, 184-194 (DE) <https://doi.org/10.1016/j.energy.2013.03.086>
- Mansouri, I.; M. Newborough & D. Probert. "Energy consumption in UK households: impact of domestic electrical appliances", en: *Applied Energy*. 54(3), 1996, 211-285.
- Martínez-Montejo, S. A. & C. Sheinbaum-Pardo. "The impact of energy efficiency standards on residential electricity consumption in Mexico", en: *Energy for Sustainable Development*. 32, 2016, 50-61 (DE) <https://doi.org/10.1016/j.esd.2016.02.010>
- Medina-Ross, J. A.; J. C. Mata-Sandoval & R. López-Pérez. "Indicators for sustainable energy development in Mexico", en: *Natural Resources Forum*. 29(4), 2005, 308-321 (DE) <https://doi.org/10.1111/j.1477-8947.2005.00143.x>
- Mercado García, A. "Datos ambientales de los hogares mexicanos", en: Mercado García, A. y C. R. López Pérez. *La estadística ambiental en México*. El Colegio de México, AC, 2014, pp. 247-260.
- Minnesota Population Center. *Integrated Public Use Microdata Series, International: Version 6.5 (dataset)*. Minneapolis, University of Minnesota, 2017 (DE) <http://doi.org/10.18128/D020.V6.5>
- Pachauri, S. & D. Spreng. "Energy Use and Energy Access in Relation to Poverty", en: *Economic and Political Weekly*. 39(3), 2004, 271-278.
- _____ "Measuring and monitoring energy poverty", en: *Energy Policy*. 39(12), 2011, 7497-7504 (DE) <https://doi.org/10.1016/j.enpol.2011.07.008>
- Rosas, J., C. Sheinbaum & D. Morillon. "The structure of household energy consumption and related CO₂ emissions by income group in Mexico", en: *Energy for Sustainable Development*. 14(2), 2010, 127-133 (DE) <https://doi.org/10.1016/j.esd.2010.04.002>
- Rosas-Flores, J. A. & D. M. Gálvez. "What goes up: Recent trends in Mexican residential energy use", en: *Energy*. 35(6), 2010, 2596-2602 (DE) <https://doi.org/10.1016/j.energy.2010.01.015>
- Rosas-Flores, J. A.; D. Rosas-Flores & D. M. Gálvez. "Saturation, energy consumption, CO₂ emission and energy efficiency from urban and rural households appliances in Mexico", en: *Energy and Buildings*. 43(1), 2011, 10-18 (DE) <https://doi.org/10.1016/j.enbuild.2010.08.020>
- Sachs, J. D. "From millennium development goals to sustainable development goals", en: *The Lancet*. 379(9832), 2012, 2206-2211.
- _____ "Goal-based development and the SDGs: implications for development finance", en: *Oxford Review of Economic Policy*. 31(3-4), 2015, 268-278 (DE) <https://doi.org/10.1093/oxrep/grv031>
- Sánchez Peña, L. "Una mirada demográfica a las metas de mitigación de gases de efecto invernadero en México", en: Saucedo, S. E. G. y V. Ugalde (eds.). *Gobierno, territorio y población: las políticas públicas en la mira*. El Colegio de México, 2014, pp. 611-644.
- _____ "Sustentabilidad y equidad: reflexiones en torno a la agenda latinoamericana post 2015", en: *Revista Brasileira de Estudos de População*. 32(3), 2015, 599-607 (DE) <https://doi.org/10.1590/S0102-30982015000000036>
- _____ "Ingresos y consumo energético de los hogares en el Distrito Federal" en: Aguilar, A.G. (ed.). *La Ciudad de México en el Siglo XXI*. DF, SECITI-Miguel Ángel Porrúa, 2016.
- Sánchez Peña, L. & A. Escoto Castillo. "Desigualdades en el consumo energético de los hogares en México", en: Lezama, J. L. (ed.). *Cambio climático, ciudad y gestión ambiental. Los ámbitos nacional e internacional*. Ciudad de México, El Colegio de México, AC, 2018, pp. 403-422.
- Sanquist, T. F.; H. Orr; B. Shui & A. C. Bittner. "Lifestyle factors in U.S. residential electricity consumption", en: *Energy Policy*. 42, 2012, 354-364 (DE) <https://doi.org/10.1016/j.enpol.2011.11.092>
- Secretaría de Gobernación. *Diario Oficial de la Federación. NOM-032-ENER-2013* (DE) http://www.dof.gob.mx/nota_detalle.php?codigo=5330530&fecha=23/01/2014
- Sheinbaum, C. & G. S. Dutt. "The structure of residential energy consumption in the Mexico City Metropolitan Area", en: *Energy for Sustainable Development*. 3(1), 1996, 43-48 (DE) [https://doi.org/10.1016/S0973-0826\(08\)60181-7](https://doi.org/10.1016/S0973-0826(08)60181-7)
- Sovacool, B. K. "Conceptualizing urban household energy use: Climbing the 'Energy Services Ladder'", en: *Energy Policy*. 39(3), 2011, 1659-1668 (DE) <https://doi.org/10.1016/j.enpol.2010.12.041>
- Swan, L. G. & V. I. Ugursal. "Modeling of end-use energy consumption in the residential sector: A review of modeling techniques", en: *Renewable and Sustainable Energy Reviews*. 13(8), 2009, 1819-1835 (DE) <https://doi.org/10.1016/j.rser.2008.09.033>
- Widén, J. & E. Wäckelgård. "A high-resolution stochastic model of domestic activity patterns and electricity demand", en: *Applied Energy*. 87(6), 2010, 1880-1892.
- World Bank. *World Development Indicators*. 2017 (DE) <http://wdi.worldbank.org/table/3.7>, consultado el 6 de junio de 2017.

Determinantes subnacionales *de la informalidad laboral en México*

Subnational Determinants *of Labor Informality in Mexico*

Benjamín Temkin* y Jorge Cruz Ibarra**

Escena de la calle de la Ciudad de México/Starcevic/Getty Images



* Facultad Latinoamericana de Ciencias Sociales (FLACSO) México, temkin@flacso.edu.mx

** Universidad Panamericana, Ciudad de México, jmcruz@up.edu.mx

El artículo pretende contribuir a la identificación de los principales factores que influyen sobre los niveles diferenciales de informalidad laboral en las entidades federativas en México. Encontramos que las variables que mejor explican las diferencias en los niveles de empleo informal entre los estados de la República son el Producto Interno Bruto (PIB) per cápita, el porcentaje de afiliados al Seguro Popular y la intensidad de la microempresarialidad.

Palabras clave: informalidad laboral; entidades federativas; PIB per cápita; Seguro Popular; microempresarialidad; México.

Recibido: 24 de agosto de 2018
Aceptado: 6 de febrero de 2019

Introducción

La importancia del estudio de la informalidad laboral, tanto para identificar sus causas como para comprender sus efectos, radica en la potente asociación que tiene con la productividad, el desempeño de la economía, el tamaño de las empresas y las condiciones de vida de los trabajadores. El crecimiento de la informalidad en el empleo se ha asociado de manera clara con efectos negativos sobre el crecimiento económico y la productividad laboral, así como con el aumento de la pobreza y la desigualdad (Dougherty y Escobar, 2013).

En México, de acuerdo con datos del Instituto Nacional de Estadística y Geografía (INEGI) del 2016, la tasa de informalidad laboral alcanza cerca de 60% de la población ocupada, aunque se observa una significativa variación entre entidades federativas y periodos en los últimos años.¹ En América Latina, este valor oscila entre 40% en Uruguay y 75% en Bolivia (OIT, 2014).

Controlando por características intrínsecas a cada país o región, la informalidad laboral afec-

¹ Para dar un ejemplo, en el último periodo que analiza este trabajo (2015), la tasa de informalidad laboral por entidad federativa va de 34% en Chihuahua a 77% en Guerrero, mientras que la desviación estándar entre los valores de todas las entidades fue de 11.6 por ciento.

This article intends to contribute to the identification of the main factors influencing the differential levels of labor informality at the subnational level in Mexico. We found that the variables that best explain the differences in the levels of informal employment between the different Mexican states are GDP per capita, *Seguro Popular* membership, and the prevalence of microenterprises.

Key words: labor informality; federal states; GDP per capita; *Seguro Popular*; microenterprises; Mexico.

ta de forma negativa el crecimiento económico (Loayza, Servén y Sugawara, 2009); asimismo, está muy ligada con la productividad empresarial debido a las condiciones de irregularidad de las empresas informales, las cuales, con el fin de permanecer fuera del registro de las autoridades, se mantienen pequeñas y sin capacidad de adoptar nuevas tecnologías, lo cual impacta de manera negativa su productividad (Dabla-Norris *et al.*, 2008). Por otro lado, aun cuando la actividad informal suele estar estrechamente asociada con las microempresas, en América Latina se detecta un importante porcentaje de trabajo informal dentro de las medianas y grandes empresas (Fajnzylber, 2007). En el caso de México, el porcentaje de los trabajadores informales que labora en empresas formales, instituciones y gobiernos alcanza 25% del total (Temkin y Bensusán, 2014).

Dado que los empleados informales —sin importar el sector en el que se desempeñan y el salario que reciben— se encuentran legalmente desprotegidos, tanto su lugar de trabajo como sus condiciones laborales, y sus ingresos son inestables. Diversos autores han evidenciado el impacto negativo de la informalidad en el empleo sobre el bienestar fisiológico y psicológico de los trabajadores, mostrando que los informales experimentan estrés, inseguridad e inestabilidad, factores que disminuyen sus niveles

de bienestar (Artazcoz, Benach, Borrel y Cortés, 2005; Benach, Vives *et al.*, 2014; Kuykendall y Tay, 2015; Temkin, 2016; Temkin y Cruz Ibarra, 2018).

Este artículo intenta identificar, a partir de datos recolectados para cada una de las entidades federativas de México entre el 2005 y el 2015, cuáles son las principales variables que explican las diferencias en los niveles de informalidad laboral al interior del país. Con información proveniente sobre todo de la Encuesta Nacional de Ocupación y Empleo (ENOE) y el Banco de Información Económica (BIE) del INEGI, desarrollamos modelos de regresión que nos permiten identificar el grado de asociación y probables efectos de nuestras variables explicativas, controlando por características intrínsecas a cada estado.

Dinámica de la informalidad laboral en México 2005-2015

La informalidad laboral es una de las situaciones específicas en que se desenvuelve un número importante —y en muchos casos creciente— de los participantes de la fuerza de trabajo en los distintos países y, en particular, en aquéllos con un menor nivel de desarrollo. De acuerdo con la Organización Internacional del Trabajo (OIT, 2002), el empleo informal, también conocido como *empleo atípico*, "...se compone de una gran diversidad de ocupaciones en una amplia gama de lugares de trabajo (que a su vez pueden ser formales o informales)...". Incluye, por un lado, a comerciantes ambulantes, autoempleados, todo tipo de cuentapropistas y familiares no remunerados y, por el otro, a empleados asalariados informales —esto es, que no cuentan con contratos formales, beneficios o protección social y de salud— que trabajan en empresas informales o formales. Este rubro incluye, también, a las empleadas domésticas y la mayoría de los trabajadores agrícolas (OIT, 2002).

Por su parte, el INEGI (2018) define la informalidad laboral como el porcentaje de los ocupados que son "...laboralmente vulnerables por la naturaleza de la unidad económica para la que traba-

jan..." (esto se puede deber a la carencia de constitución legal) o "...cuyo vínculo o dependencia laboral no es reconocido por su fuente de trabajo..." (por ejemplo, el trabajo sin contrato).

En los últimos años, hemos podido observar ligeras fluctuaciones en los indicadores de informalidad en México; los datos agregados de todo el país señalan una ligera reducción en ella para el periodo que abarca del 2005 al 2015 (ver gráfica 1); sin embargo, cuando se analizan los datos por entidad federativa, encontramos importantes diferencias con porcentajes que fluctuaron entre 32 y 76 por ciento.²

Este artículo se basa en la definición y los indicadores de informalidad del INEGI para el análisis de las dinámicas del trabajo informal entre el 2005 y el 2015. La gráfica 2 ilustra las diferencias en la tasa de informalidad entre entidades federativas para este periodo.

Aquellos estados con menor porcentaje de trabajadores informales son Chihuahua, Baja California, Baja California Sur, Coahuila de Zaragoza, Nuevo León y Aguascalientes, en su mayoría, del norte del país. Por el contrario, los que presentaron niveles más altos de informalidad son Chiapas, Guerrero, Hidalgo, Michoacán de Ocampo, Oaxaca, Puebla y Tlaxcala.

Observamos un crecimiento importante de la informalidad entre el 2005 y el 2015 sobre todo en Campeche, Colima, Guerrero y Tabasco; mientras que, en Coahuila de Zaragoza, la Ciudad de México, Jalisco y Nuevo León hubo importantes retrocesos. En total, son 15 las entidades con disminuciones en este periodo, 13 donde se registró un crecimiento y cinco (Baja California, Chihuahua, Hidalgo, México y Zacatecas) en los cuales se mantuvo casi sin cambios.

Los datos anteriores indican que, sin considerar la tendencia nacional en los porcentajes de informalidad, existen importantes diferencias en las tendencias a nivel subnacional, lo cual justifica un estudio que tome en cuenta las características de cada entidad federativa.

² Se revisaron datos de la ENOE del 2005 al 2015 para las 32 entidades federativas.

Gráfica 1

Cambios en la formalidad e informalidad

México 2005-2015

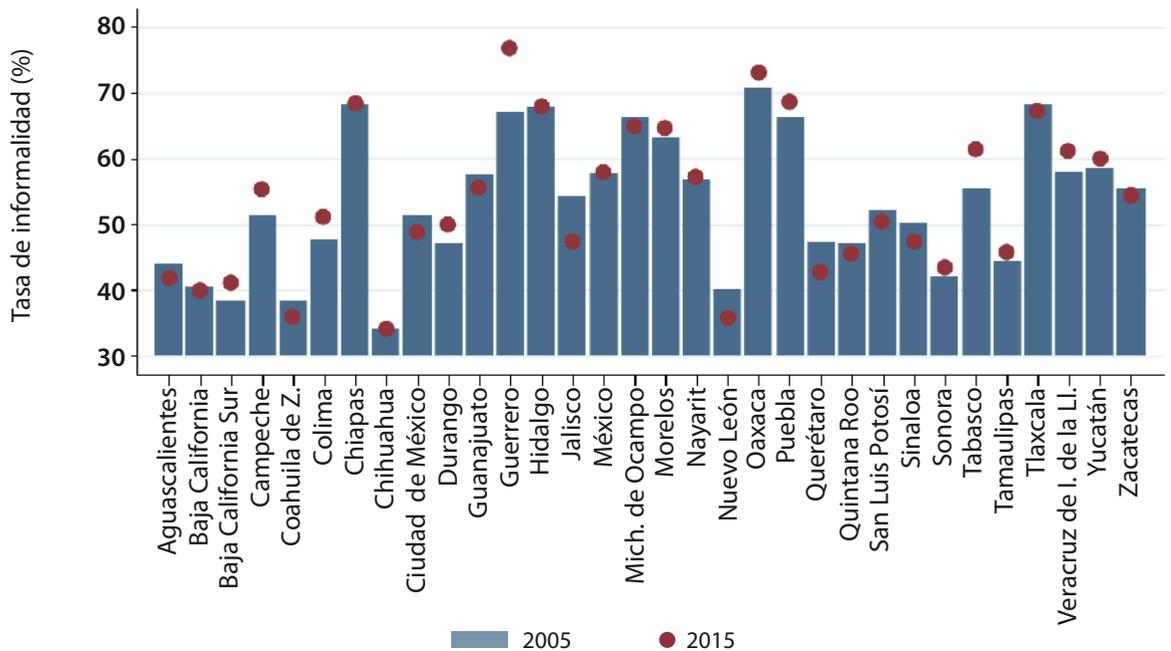


Fuente: elaboración propia con datos de INEGI. Banco de Información Económica (BIE).

Gráfica 2

Cambios en la tasa de informalidad

32 entidades federativas, 2005 y 2015



Fuente: INEGI. Encuesta Nacional de Ocupación y Empleo (ENOE). INEGI.

Determinantes de la informalidad laboral

Para indagar acerca de los factores asociados a la informalidad en el empleo, se construyó una base de datos que contempla las 32 entidades federativas mexicanas durante un periodo de 11 años: 2005 al 2015. Entre las variables³ analizadas (ver cuadro 1) están la tasa de informalidad, el PIB per cápita (a precios constantes del 2008), la Inversión Extranjera Directa (IED), el porcentaje de afiliados al

3 Es importante mencionar que los trabajos de Azuara y Marinescu (2011) y Dougherty y Escobar (2013) han guiado nuestra selección de variables.

Seguro Popular (SP), la cualificación de la mano de obra⁴ y la intensidad de la microempresarialidad.⁵

Desempeño económico

Un entorno económico favorable es fundamental para contribuir a la reducción de la informalidad. Las oportunidades de empleo formal son más abundantes en las economías en crecimiento que

4 Entendida como el porcentaje de los miembros de la Población Económicamente Activa (PEA) no agropecuaria con educación media superior o superior.

5 Definida como el porcentaje de la PEA no agropecuaria que labora en microempresas (aquellos negocios con menos de 10 empleados).

Cuadro 1

Variable	Definición	Unidad	Fuente
Entidad	Etiqueta numérica de cada entidad federativa	N/A	N/A
Año	Año	N/A	N/A
Abreviación	ISO por entidad	N/A	N/A
PIBpb (millones de pesos a precios constantes del 2008)	PIB per cápita (millones de pesos a precios constantes del 2008)	Millones de pesos	BIE (INEGI).
IED (millones de dólares)	IED (millones de dólares)	Millones de dólares	BIE (INEGI).
Logaritmo natural del PIBpc	Logaritmo natural del PIB per cápita	N/A	Cálculo con datos del BIE (INEGI).
Tasa de informalidad laboral	Porcentaje de trabajadores informales en el 4.º trimestre (no toma en cuenta el sector agropecuario)	Porcentaje	ENOE (INEGI).
(%) Afiliados al Seguro Popular	Afiliados al Seguro Popular como porcentaje de la población	Porcentaje	Cálculo con datos de la Secretaría de Salud.
Cualificación de la mano de obra (población no agropecuaria)	Porcentaje de los miembros de la PEA no agropecuaria con educación media superior o superior	Porcentaje	Cálculo con datos de la ENOE (INEGI).
Cualificación de la mano de obra al cuadrado (población no agropecuaria)	Porcentaje de los miembros de la PEA no agropecuaria con educación media superior o superior (valores elevados al cuadrado)	Porcentaje al cuadrado	Cálculo con datos de la ENOE (INEGI).
Intensidad de la microempresarialidad (población no agropecuaria)	Porcentaje de la PEA no agropecuaria que labora en microempresas (aquellos negocios con menos de 10 empleados)	Porcentaje	Cálculo con datos de la ENOE (INEGI).
Facilidad para abrir un negocio	Costo de comenzar un negocio (como % del PIB)	Porcentaje	<i>Doing Business Database</i> , World Bank.
Índice Nacional de Corrupción y Buen Gobierno	Índice Nacional de Corrupción y Buen Gobierno de Transparencia Mexicana. Valores más altos indican mayor corrupción	Índice	Transparencia Mexicana, vía base de datos del IMCO (original solo datos para el 2005, el 2007 y el 2010).

en aquellas en recesión (Loayza *et al.*, 2009); es así que una de las principales variables explicativas de los niveles de informalidad laboral es el PIB, obtenido a partir de datos del INEGI con valores constantes del 2008.

La literatura sobre la influencia de la IED y el desarrollo de los mercados locales coincide en el diagnóstico en cuanto a sus efectos positivos sobre la generación de empleos formales. La presencia de empresas multinacionales influencia de forma positiva el crecimiento de los mercados (Markusen y Venables, 1999; Haskel *et al.*, 2002). Además, dado que las empresas internacionales invierten su capital solo en compañías locales que cumplen con las reglamentaciones pertinentes, el resultado es el crecimiento del número de empleos formales (Dabla-Norris *et al.*, 2008).

Por otro lado, trabajos como los de Blomström & Kokko (2003) y Smarzynska (2004) concuerdan con la valoración positiva del impacto de la inversión

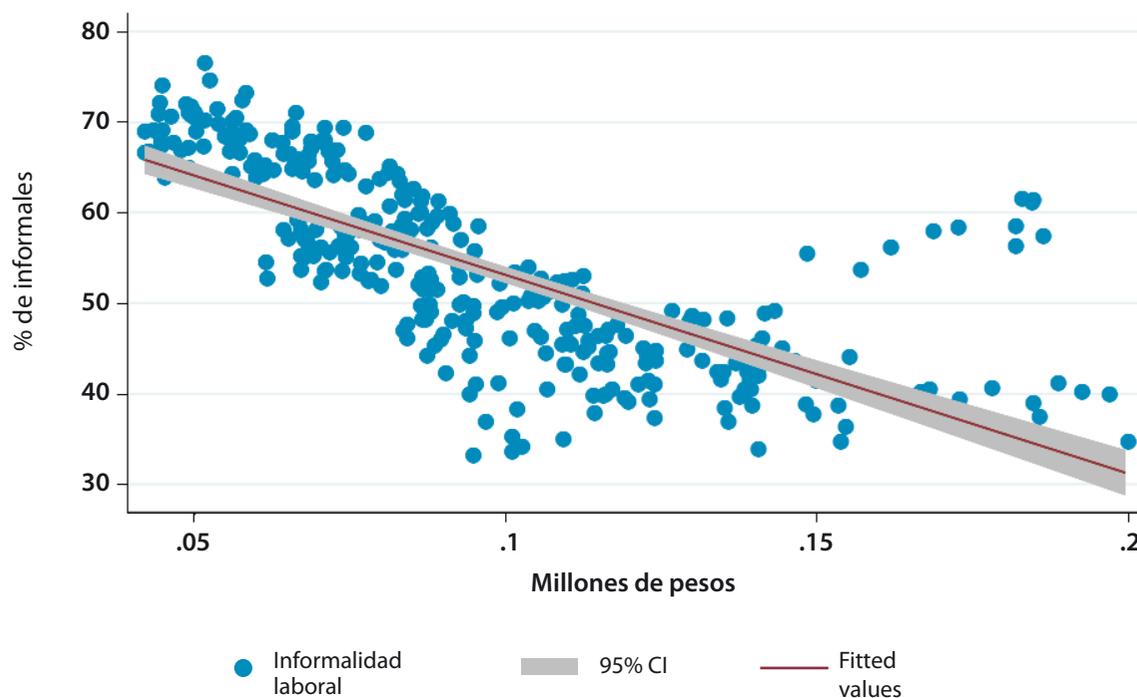
extranjera sobre los mercados laborales locales; sin embargo, consideran que para que estas influencias se manifiesten, en especial sobre las medianas y pequeñas empresas, son necesarias ciertas condiciones de asimilación de nuevas técnicas y tecnologías, así como políticas públicas (e.g. subsidios gubernamentales) que fomenten dichas condiciones. De acuerdo con lo anterior, analizamos los datos del INEGI acerca de la IED para medir el efecto de dicha inversión sobre la economía de los estados y, a su vez, su impacto en los niveles de informalidad.

El análisis de correlación entre el PIB, la IED y la tasa de informalidad indica una asociación negativa y significativa entre la variable dependiente y las dos variables explicativas, con coeficientes de -0.20 y -0.29, respectivamente. Las gráficas 3 y 4 ilustran dichas asociaciones.⁶

⁶ Es importante mencionar que en ambas gráficas se eliminaron los valores extremos para una mejor visualización de la relación.

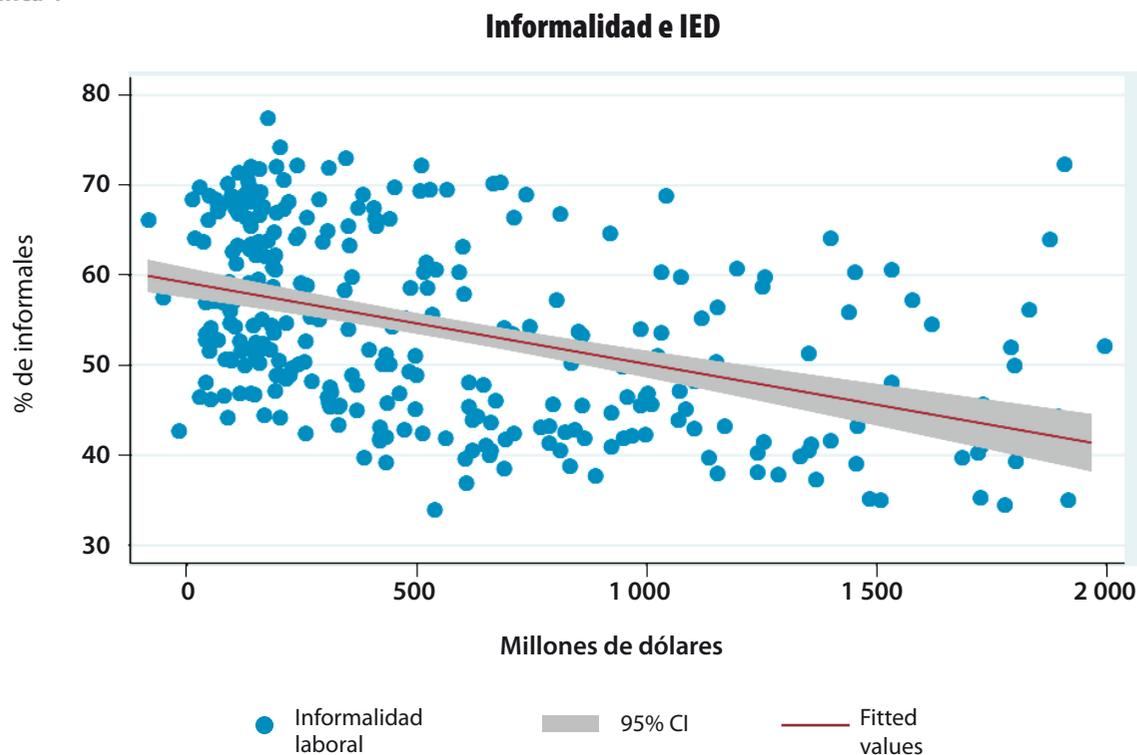
Gráfica 3

Informalidad y PIBpc



Nota: se eliminaron valores extremos (mayores a 0.2 mdp).

Gráfica 4



Nota: se eliminaron valores extremos (mayores a 2 mil mdd).

Afiliación al Seguro Popular

El SP forma parte del Sistema Nacional de Salud de México como complemento a los demás servicios de seguridad social públicos y privados. Es de inscripción voluntaria e independiente de las condiciones laborales, el cual entró en vigor desde el 2004 en 24 entidades federativas y que, hasta el 2015, tenía presencia en todos los estados de la República con una cobertura promedio de 48% de la población del país.⁷ También aquí, el porcentaje de afiliación al Seguro Popular varía para las distintas entidades federativas mexicanas y ha sufrido cambios durante el periodo analizado (ver gráfica 5).

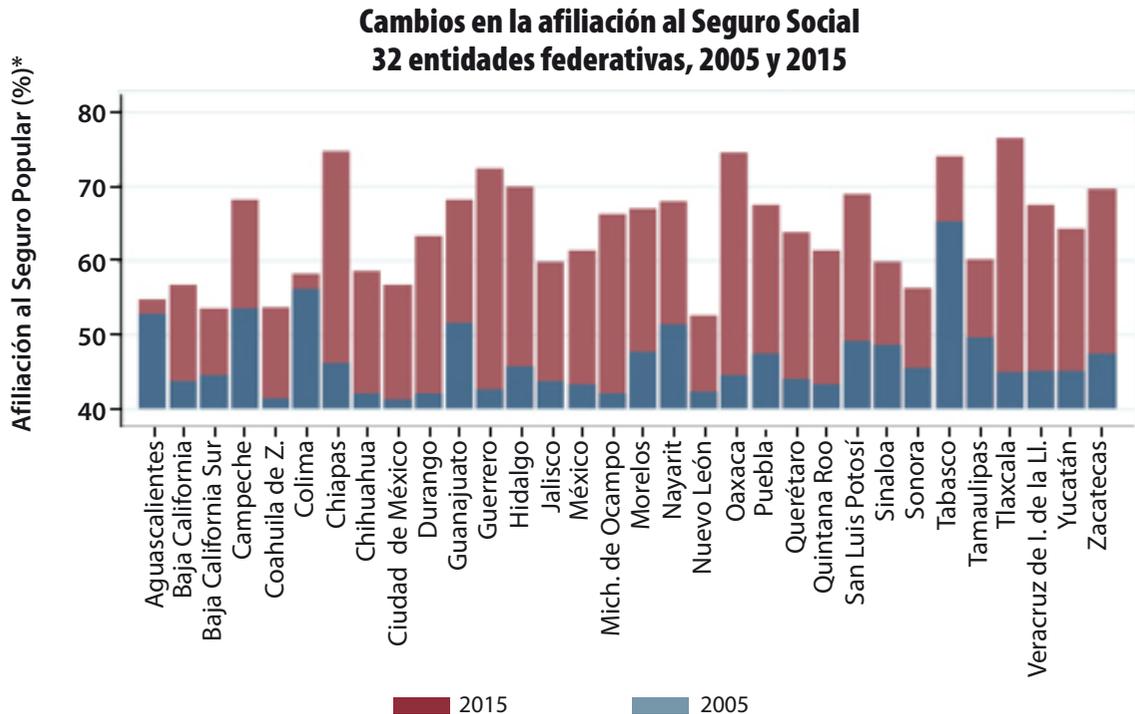
Una de las principales discusiones en la literatura concierne al impacto de la introducción del SP en México sobre los incentivos para operar en el sector formal, es decir, si este programa motiva a los

trabajadores formales a tomar empleos en el sector informal. En esta línea, distintas investigaciones sustentan que la relación entre la expansión del SP y el empleo formal es negativa (Azua y Marinescu, 2011; Aterido *et al.*, 2011; Duval Hernández y Smith Ramírez, 2011; Bosch *et al.*, 2012; Bosch y Campos-Vázquez, 2014; Arenas *et al.*, 2015).

Los efectos encontrados en la literatura no son homogéneos, sino que obedecen a las particularidades de las investigaciones: unidades de análisis elegidas, periodos estudiados, fuentes de los datos, metodología, etcétera. Bosch y Campos-Vázquez (2014) encuentran un efecto negativo sobre la creación de empleos formales, en específico entre las pequeñas y medianas empresas; su trabajo atribuye al SP una reducción de 300 mil empleos formales desde su instauración en el 2002 hasta el 2010. Por su parte, Arenas *et al.* (2015) registran un impacto negativo en el empleo formal de las mujeres de origen urbano y hombres de origen rural.

⁷ Cálculo elaborado con datos de la Secretaría de Salud para el 2015.

Gráfica 5



* Porcentaje de la población de la entidad federativa que está afiliada.

Fuente: Secretaría de Salud, fichas de información de afiliación.

Aterido *et al.* (2011) observan diferencias a nivel municipal en la penetración del programa y encuentran una reducción de 0.4 a 0.7 puntos porcentuales en la formalidad para el periodo que abarca del 2002 al 2011. Duval Henández y Smith Ramírez (2011) desarrollan un modelo que concluye que el SP, entre el 2002 y el 2011 ha generado desincentivos para la formalización de la fuerza laboral; aunque precisan que los efectos son pequeños si se les compara con el de las variables de control y que éstos son más grandes en áreas rurales que en urbanas.

Los efectos encontrados también varían en tamaño. De acuerdo con Azuara y Marinescu (2011), la afiliación al Seguro Popular tiene un efecto mínimo sobre la informalidad. Ellos hallaron que para cierto sector de la PEA (trabajadores mayores de 30 años de edad, casados y con hijos) la presencia del SP tiene un efecto significativo y positivo, pero mínimo (coeficiente = 0.013**).

A partir del cálculo del coeficiente de correlación entre el porcentaje de afiliados al SP y la tasa

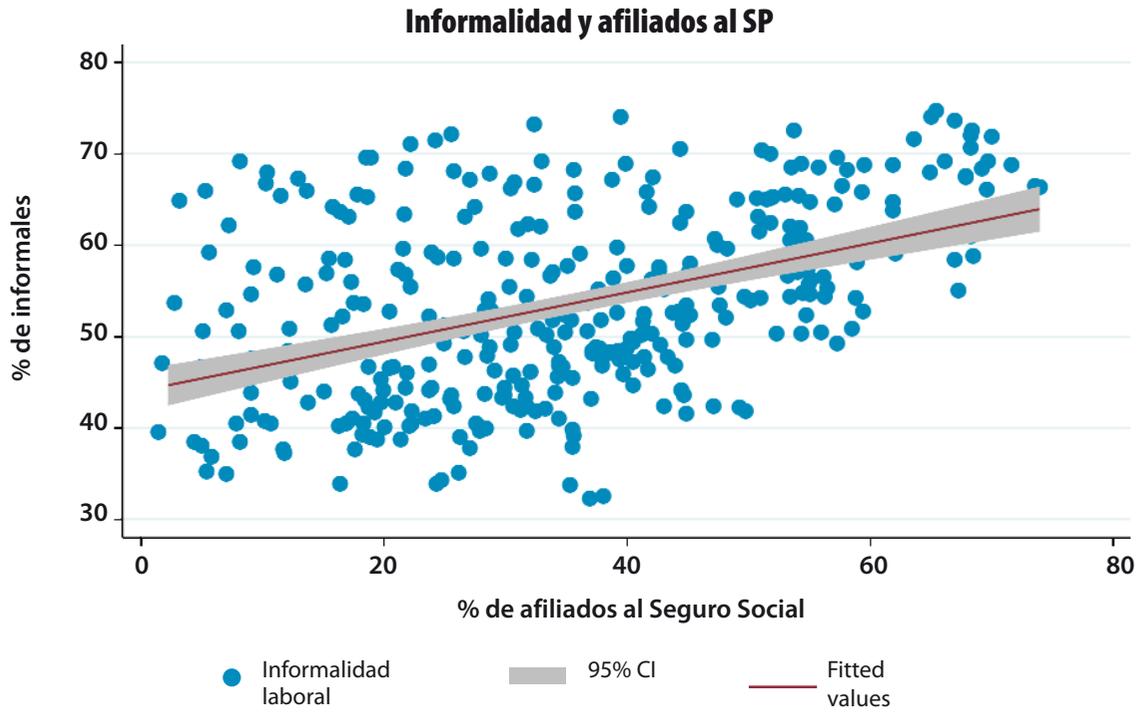
de informalidad por estados, encontramos que la asociación entre dichas variables para el periodo analizado es de $r = 0.45$, lo cual indica una relación positiva y significativa (ver gráfica 6).

Cualificación de la mano de obra

En la gráfica 7 observamos una relación negativa entre informalidad y cualificación de la mano de obra, es decir, aquellos estados con trabajadores más preparados son, también, los que presentan menos informalidad. Dado un coeficiente de correlación de Pearson de apenas -0.25 entre estas dos variables, se observa una línea de regresión mucho más plana en esta gráfica que en las anteriores.

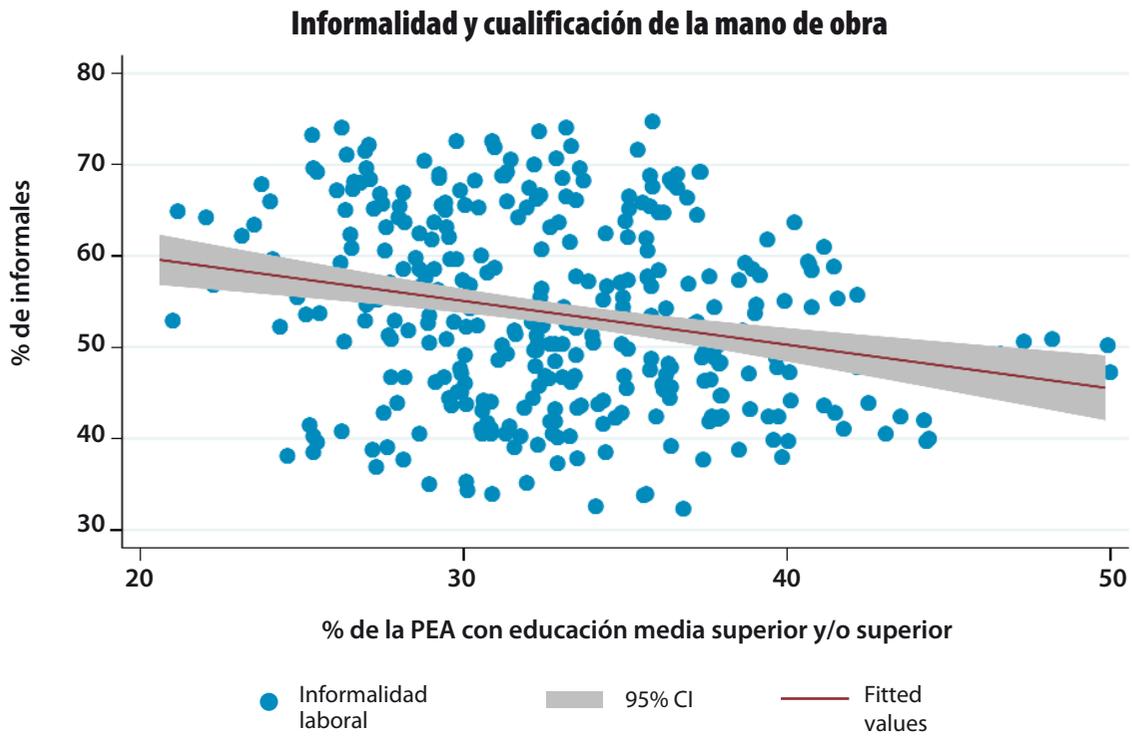
Nos atrevemos a pensar que la relación entre la informalidad y la cualificación de la mano de obra no es lineal; si se presta atención a la gráfica 7, identificamos que las observaciones con valores más altos en X poseen valores ligeramente más altos en Y que las que las preceden. Pondremos a

Gráfica 6



Nota: la variable afiliados al SP toma como base la población total de cada entidad federativa.

Gráfica 7



Nota: la variable cualificación de la mano de obra toma como base la PEA no agropecuaria.

prueba en el modelo de regresión la posibilidad de una relación no lineal entre informalidad y cualificación de la mano de obra.

Intensidad de la microempresarialidad

La asociación entre la microempresarialidad en los estados y su porcentaje de informales es muy fuerte ($r = 0.93$). Descartamos una coincidencia en la operacionalización de las variables, ya que ambas son medidas de manera diferente por el INEGI. La gráfica 8 comprueba la asociación teórica fuerte entre informalidad y microempresarialidad presente en la literatura, como la que exponen Dabla-Norris *et al.* (2008), quienes encuentran que la informalidad está muy ligada a las pequeñas empresas. Según los autores, éstas son, por lo general, menos productivas, ya sea por falta de desarrollo tecnológico o como estrategia para permanecer fuera de la vista de la autoridad.

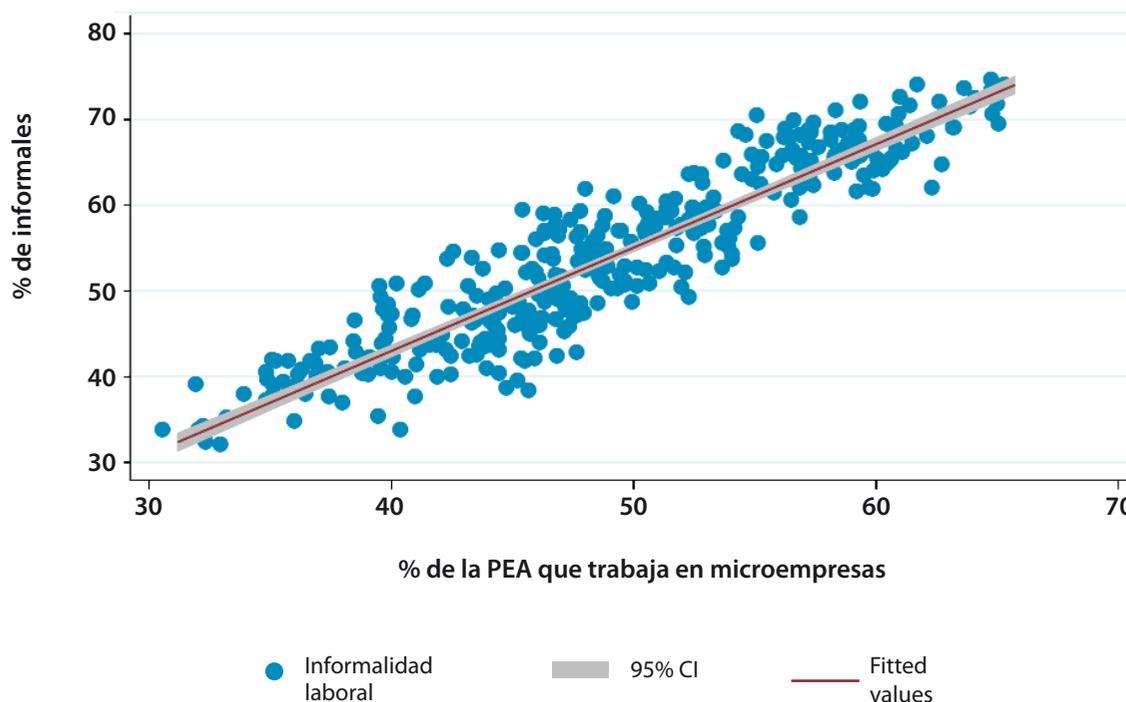
Corrupción y costo de iniciar un negocio

La literatura ha señalado como determinantes importantes de los niveles de informalidad a aquellas variables que inciden sobre la productividad y el desarrollo económico, como: el nivel de corrupción, el acceso a financiamiento y, en general, el nivel de dificultad para abrir un negocio. Distintos trabajos han encontrado una relación fuerte y positiva entre los niveles de corrupción y el crecimiento de la informalidad (Gatti y Honorati, 2008; Dougherty y Escobar, 2013; Mishra y Ray, 2013; Dutta, Kar y Roy, 2013).

Dougherty (2013) explora la relación entre la corrupción y el tamaño de las empresas, argumentando que niveles altos de corrupción entorpecen el desarrollo empresarial a la vez que fomentan el crecimiento de la informalidad. Por su parte, Mishra y Ray (2010) dicen que son complementarias y que ambas están determinadas tanto por la desigualdad como por la precariedad económica.

Gráfica 8

Informalidad y microempresarialidad



Notas:

- La variable microempresarial toma como base la PEA no agropecuaria.
- Microempresa definida como aquella con 10 empleados o menos, $n = 352$.

En las gráficas 9 y 10 se pone a prueba la asociación entre la informalidad y dos variables que nos permiten operacionalizar los niveles de corrupción y el acceso a financiamiento. La primera es el Índice Nacional de Corrupción y Buen Gobierno, realizado por Transparencia Mexicana, y la segunda, el costo de comenzar un negocio, calculado por el Banco Mundial.⁸

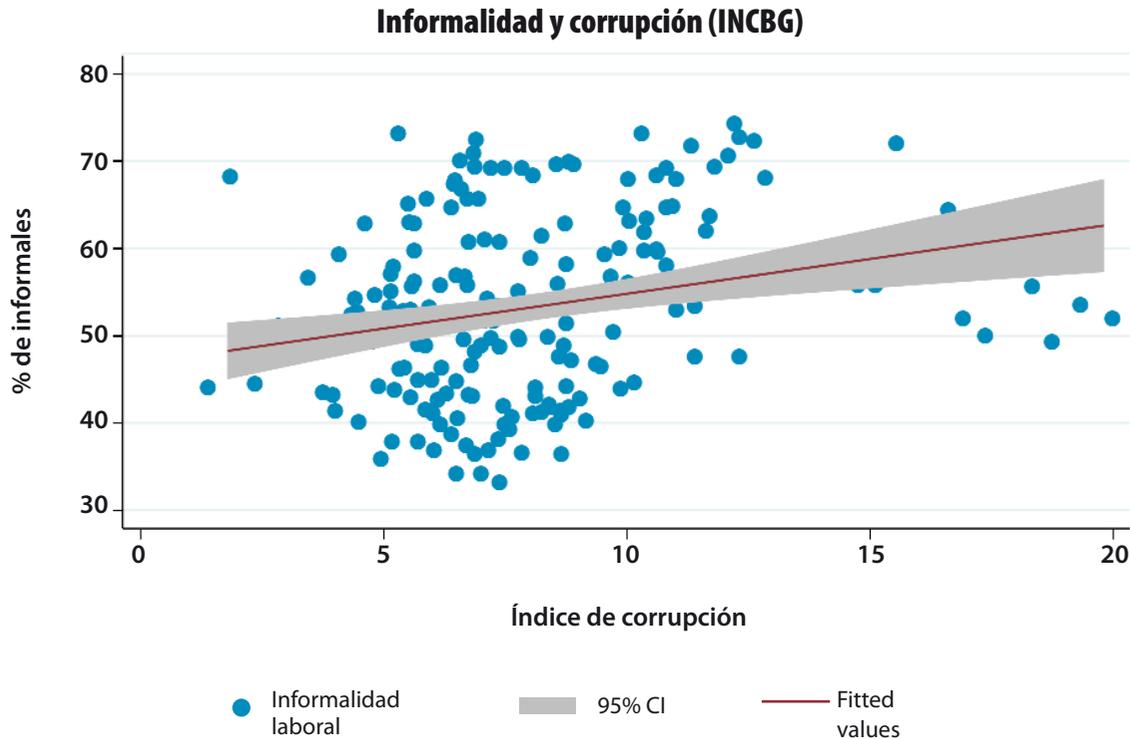
Podemos observar que la relación entre corrupción e informalidad es positiva, tal como señala la literatura (ver gráfica 9). En aquellas entidades con mayores niveles de corrupción, también podemos observar mayor informalidad. Por otro lado, la gráfica 10 indica que en los estados donde comenzar un negocio es más costoso, los niveles de informalidad son más bajos. La importancia que le da la literatura a la relación entre estas variables y la informalidad es la razón por la cual están presen-

tes en nuestro análisis descriptivo; sin embargo, es importante señalar que no están en el análisis de regresión debido al bajo número de observaciones con el que contamos para cada variable. Introducir las en este modelo reduce el número de casos observados de 352 a 66.

Resulta importante puntualizar que en el proceso de investigación se consideró un número importante de variables institucionales, económicas, políticas y culturales en potencia relevantes para la comprensión de las marcadas diferencias en los niveles de informalidad laboral a nivel estatal. Algunas de ellas, mencionadas con frecuencia en fuentes periodísticas y académicas como causales de informalidad laboral (presencia de altos niveles de población indígena, altos índices de marginación social, ausencia de alternancia política o el control político estatal por parte de partidos políticos específicos), fueron descartadas por carecer de significancia estadística en modelos de regresión preliminares.

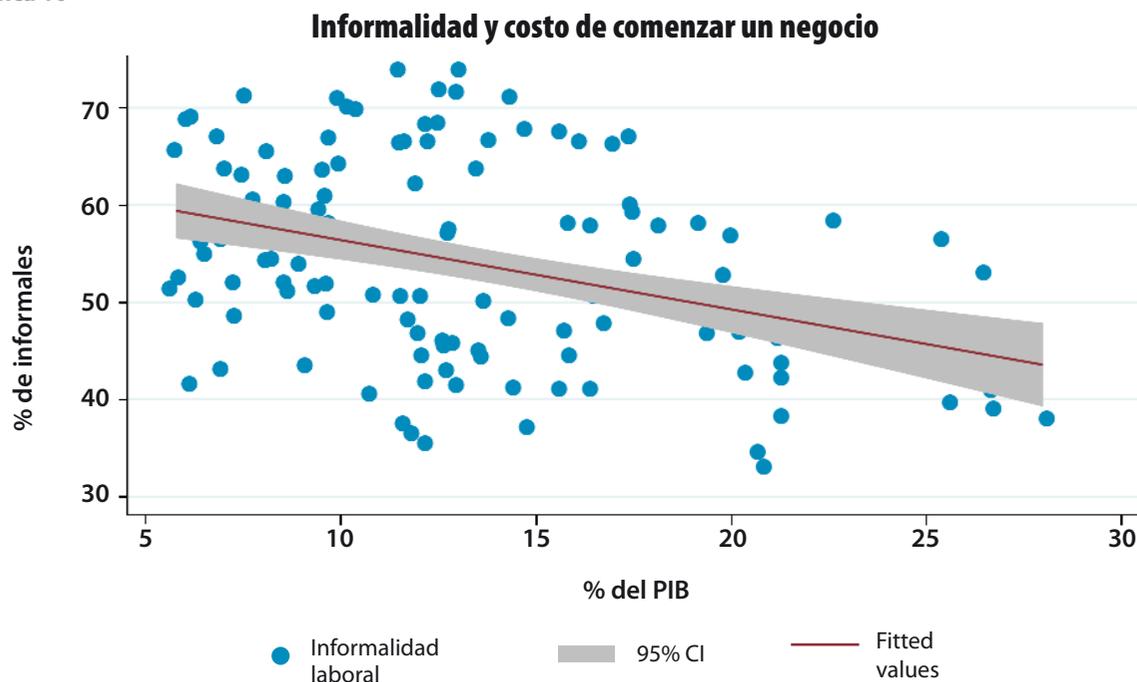
⁸ Detalles sobre cómo están medidas estas variables se encuentran en el cuadro 1.

Gráfica 9



Nota: valores más altos indican mayor corrupción.
Fuente: Transparencia Mexicana, valores ajustados por el IMCO, $n = 0192$.

Gráfica 10



Fuente: Banco Mundial. *Doing Business Database*, $n = 134$.

Modelo y análisis

En primera instancia, utilizamos la técnica de regresión lineal múltiple para explorar los efectos que cada una de nuestras variables de interés tiene sobre la dependiente: la tasa de informalidad laboral (ver cuadro 2). El modelo utiliza efectos fijos para controlar por variables no observables por estado. Se reporta el coeficiente, el error estándar para cada una y la R^2 ajustada de todo el modelo.

Trabajar con datos panel requiere tomar en cuenta las tendencias de las variables en el tiempo. Es importante considerar dichas tendencias para evitar concluir sobre relaciones entre la variable dependiente y las independientes que podrían llegar a ser espurias. Una técnica que permite controlar por dichas tendencias es el modelo de primeras diferencias,⁹ en el cual se calcula la primera

⁹ Se optó por un modelo de primeras diferencias en lugar de uno de variable dependiente rezagada después de realizar pruebas de estacionalidad (Prueba Dickey-Fuller) para la dependiente. Se realizó la prueba para cada uno de los paneles (entidades federativas) y la hipótesis nula se rechazó en 24 de 32 casos, por lo que podemos asumir que no hay estacionalidad en la variable *tasa de informalidad* cuando se toman en cuenta los datos agregados.

diferencia¹⁰ tanto de la variable dependiente como de los regresores. Los resultados se encuentran en el cuadro 3.

En ambos modelos se observan efectos significativos de tres variables: PIB per cápita, afiliación al Seguro Popular y microempresarialidad. La primera se asocia de manera negativa con la informalidad, es decir, a mejor desempeño económico de una entidad, menor el porcentaje de informalidad; las otras dos, por su parte, están asociadas de manera positiva con la tasa de informalidad de los estados. Es de señalarse que las variables significativas conservan su dirección y son similares en magnitud entre los dos modelos.

Por otro lado, Dougherty y Escobar (2013) plantean una relación negativa entre la cualificación de la mano de obra y la informalidad, la cual se puede identificar en los modelos de los cuadros 2 y 3. Como planteamos con anterioridad, consideramos

¹⁰ La primera diferencia de la variable dependiente sería expresada de la siguiente manera:

$$\Delta Y_t = Y_t - Y_{t-1}$$

Cuadro 2

Y = tasa de informalidad		
PIB per cápita (logaritmo natural)	-6.503	***
	(1.38)	
IED	-0.000	
	(0.00)	
Afiliados al Seguro Popular (%)	0.060	***
	(0.01)	
Cualificación de la mano de obra	-0.322	
	(0.21)	
Cualificación de la mano de obra (al cuadrado)	0.004	
	(0.00)	
Microempresarialidad	0.644	***
	(0.05)	
Constante	11.735	*
	(5.62)	
R ²	0.4392	
N	352	

(*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$

la posibilidad de que la relación entre estas dos variables no fuera lineal, por lo cual introducimos la variable *cualificación de la mano de obra* al cuadrado. El resultado del modelo no es suficiente para mostrar un cambio en la dirección de la relación. Si bien los signos de los coeficientes de la variable original y la variable al cuadrado son

Cuadro 3

Y = tasa de informalidad-primera diferencia		
PIB per cápita (logaritmo natural)-primera diferencia	-8.784	***
	(2.39)	
IED-primera diferencia	-0.000	
	(0.00)	
Afiliados al Seguro Popular (%) -primera diferencia	0.060	**
	(0.02)	
Cualificación de la mano de obra-primera diferencia	-0.600	
	(0.36)	
Cualificación de la mano de obra (al cuadrado)-primera diferencia	0.007	
	(0.01)	
Microempresarialidad-primera diferencia	0.521	***
	(0.05)	
Constante	0.013	
	(0.12)	
R ²	0.4348	
N	320	

(*) $p < 0.05$, (**) $p < 0.01$, (***) $p < 0.001$

distintos, solo el primer coeficiente es significativo. Este resultado indica que, de acuerdo con los datos analizados, la relación entre *cualificación de la mano de obra* e informalidad es lineal y negativa; en aquellas entidades con trabajadores más escolarizados hay porcentajes más bajos de informalidad.

Dougherty y Escobar (2013) también encuentran una relación positiva entre la microempresarialidad y la informalidad; al incluir esta variable al modelo podemos ver cómo la R^2 aumenta de forma considerable; sin embargo, ésta resta poder explicativo a la cualificación de la mano de obra.

En cuanto al efecto del porcentaje de afiliación al SP sobre la informalidad, nuestros resultados se alinean con lo que señala la literatura: encontramos que, a mayor penetración del programa, corresponde un aumento en los porcentajes de informalidad a nivel estatal.¹¹ Similar a lo encontrado por Azuara y Marinescu (2011) y Duval Hernández y Smith Ramírez (2011), el impacto del Seguro Popular es pequeño pero significativo y se puede identificar a lo largo de los tres modelos anteriores.

Conclusiones

El análisis presentado en este artículo pretende contribuir a un mejor entendimiento de los factores que impactan la prevalencia de informalidad laboral en la economía mexicana, aprovechando la información disponible acerca de los niveles diferenciales de empleo informal en los mercados laborales a nivel subnacional. Pusimos a prueba el efecto que variables como el PIB per cápita, la IED, el porcentaje de afiliados al SP, la cualificación de la mano de obra y la intensidad de la microempresarialidad tuvieron sobre el porcentaje de informalidad.

Encontramos que las variables que mejor explican las diferencias en los niveles de empleo informal entre entidades federativas son el PIB per cápita, el porcentaje de afiliados al SP y la intensidad de la microempresarialidad. El hecho de que el PIB posea el coeficiente más alto y se relacione de manera negativa con el porcentaje de informales corrobora la posición preponderante que ocupa

esta variable dentro de la literatura sobre determinantes de la informalidad. Llama la atención que la IED no resulta significativa en nuestro modelo, lo que parece apoyar la postura de Blomström & Kokko (2003) y Smarzynska (2004) acerca de las condicionantes necesarias para observar el efecto del capital extranjero en el desarrollo de las empresas locales. Consideramos también que la variable PIB per cápita asume parte importante del efecto que la dimensión económica pudiera tener sobre la variable dependiente.

Por otro lado, encontramos en especial valioso el hallazgo de la relación positiva y significativa entre los porcentajes de afiliados al SP y de informalidad. Si bien la literatura no es contundente sobre la naturaleza de esta relación, nuestros datos sugieren una relación estrecha y positiva entre ellas, en concordancia con los trabajos de Azuara y Marinescu (2011), Aterido *et al.* (2011), Duval Hernández y Smith Ramírez (2011), Bosch *et al.* (2012), Bosch y Campos-Vázquez (2014) y Arenas *et al.* (2015).

Por su parte, el efecto de la microempresarialidad es de suma importancia. Como resultado del análisis bivariado, encontramos que la relación entre esta variable y el porcentaje de informalidad a nivel subnacional es muy fuerte (su coeficiente de correlación es de 0.93). Valoramos como importante mantener esta variable dentro del modelo para no sobrestimar el efecto de las demás variables independientes y confirmamos la estrecha relación entre el tamaño de las empresas y el tipo de empleo que en ellas se desarrolla.

Por último, podemos concluir que resulta urgente contar con mejores y más frecuentes mediciones a nivel subnacional de variables como corrupción, calidad y eficiencia gubernamental, acceso a financiamiento y productividad empresarial. Dichos factores son mencionados de manera recurrente en la literatura como predictores de la informalidad laboral (Gatti y Honorati, 2008; Dougherty y Escobar, 2013; Mishra y Ray, 2013; Dutta, Kar y Roy, 2013); sin embargo, su medición aún es insuficiente, por lo menos a nivel entidad federativa en México.

¹¹ Si bien en la literatura se habla de un efecto negativo sobre la formalidad (empleos formales, probabilidad de laborar en el sector informal), nosotros utilizamos como variable dependiente el porcentaje de trabajadores informales por estado; es por eso que consideramos que la asociación positiva entre los porcentajes de afiliación al SP y de informalidad estatal corresponde a los hallazgos de los autores mencionados.

Fuentes

- Alesina, A.; R. Di Tella y R. MacCulloch. "Inequality and happiness: Are Europeans and Americans different?", en: *Journal of Public Economics*. 88(9-10), 2004, pp. 2009-2042.
- Azuara, O. y I. Marinescu. Informality and the Expansion of Social Protection Programs: Evidence from Mexico. Mimeo, University of Chicago, 2011 (DE) <https://core.ac.uk/download/pdf/12030422.pdf>, consultado el 20 de enero de 2018.
- Arenas, E.; S. Parker; L. Rubalcava y G. Teruel. "Evaluación del programa del Seguro Popular del 2002 al 2005: impacto en la utilización de servicios médicos, en el gasto en salud y en el mercado laboral", en: *El Trimestre Económico*. Vol. LXXXII (4), núm. 328, 2015, pp. 807-845.
- Artazcoz, L.; J. Benach; C. Borrell e I. Cortés. "Social inequalities in the impact of flexible employment on different domains of psychosocial health", en: *Epidemiology and Community Health*. 59, 2005, pp. 761-767.
- Aterido, R.; M. Hallward-Driemeier y C. Pagés Serra. *Does Expanding Health Insurance Beyond Formal Sector Workers Encourage Informality?: Measuring the impact of Mexico's Seguro Popular*. IDB working paper series No. IDB-WS-280. Washington, D.C., Inter-American Development Bank, 2011 (DE) https://publications.iadb.org/handle/11319/4551?locale-attribute=en&locale-attribute=es&scope=123456789/12&thumbnail=false&order=desc&rpp=5&sort_by=score&page=1&query=renewable&group_by=none&etal=0&filtertype_0=country_en&filter_0=Mexico&filter_relational_operator_0>equals, consultado el 20 de enero de 2018.
- Benach, J.; A. Vives; M. Amable; C. Vanroelen; G. Tarafa y C. Muntaner. "Precarious Employment: Understanding an Emerging Social Determinant on Health", en: *Annual Review of Public Health*. 35, 2014, pp. 229-253.
- Bloomström, M. y A. Kokkom. *The Economics of Foreign Direct Investment Incentives*. Centre for Economic Policy Research. Discussion Paper No. 3775. Stockholm School of Economics, 2003 (DE) https://www.researchgate.net/profile/Ari_Kokko/publication/5094506_The_Economics_of_Foreign_Direct_Investment_Incentives/links/0fcfd509287bc3648c000000.pdf, consultado el 20 de agosto de 2018.
- Bosch, M.; M. B. Cobacho y C. Pagés. *Taking stock of eight years of implementation of Seguro Popular in Mexico*. Washington, D.C., Inter-American Development Bank, 2012 (DE) https://pdfs.semanticscholar.org/6e3a/ad979e7a4781acc3f0dcb6242362ccd9aad9.pdf?_ga=2.127968908.1005023539.1535046516-178834022.1535046516, consultado el 20 de enero de 2018.
- Bosch, M. y R. M. Campos-Vázquez. "The trade-offs of social assistance programs in the labor market: The case of the 'Seguro Popular' program in Mexico", en: *American Economic Journal: Economic Policy*. 6(4), 2014, pp. 71-99.
- Dabla-Norris, E.; M. Gradstein y G. Inchauste. "What Causes Firms to Hide Output? The Determinants of Informality", en: *Journal of Development Economics*. 85, 2008, 1-27.
- Dougherty, S. *Legal Reform, Contract Enforcement and Firm Size in Mexico*. OECD Economics Department Working Papers, no. 1042. OECD Publishing, 2013 (DE) <https://www.oecd-ilibrary.org/docserver/5k483jcn8s8-en.pdf?expires=1534872682&id=id&accname=guest&checksum=2F43BE9677143DE4E1F2405B80A18192>, consultado el 20 de enero de 2018.
- Dougherty, S. y O. Escobar. *The Determinants of Informality in Mexico's States*. OECD Economics Department Working Papers, no. 1043. OECD Publishing, 2013 (DE) <http://sean.dougherty.org/econ/papers/informal.pdf>, consultado el 20 de enero de 2018.
- Dutta, N.; S. Kar y S. Roy. "Corruption and persistent informality: An empirical investigation for India", en: *International Review of Economics & Finance*. 27, 2013, pp. 357-373.
- Duval Hernández, R. y R. Smith Ramírez. *Informality and Seguro Popular under Segmented Labor Markets*. Borrador preliminar. Ciudad de México, Centro de Investigación y Docencia Económicas (CIDE), 2011 (DE) http://conference.iza.org/conference_files/worldb2011/3229.pdf, consultado el 20 de enero de 2018.
- Fajnzylber, P. "Informality, Social Protection, and Atipovety Policies", en: Perry, G. E. et al. (eds.). *Informality: Exit and Exclusion*. Washington, D.C., The World Bank, 2007 (DE) <https://openknowledge.worldbank.org/bitstream/handle/10986/6730/400080Informal101OFFICIAL0USE0ONLY1.pdf?sequence=1&isAllowed=y>, consultado el 20 de enero de 2018.
- Gatti, R. y M. Honorati. *Informality among Formal Firms: Firm-level, Cross-country Evidence on Tax Compliance and Access to Credit*. Policy Research Working Paper 4476. The World Bank, 2008 (DE) <https://openknowledge.worldbank.org/bitstream/handle/10986/6444/wps4476.pdf?sequence=1>, consultado el 20 de agosto de 2018.
- Graham, C. y S. Pettinato. "Frustrated Achievers: Winners, Losers, and Subjective Well-Being in New Market Economies", en: *The Journal of Development Studies*. 38(4), 2001, pp. 100-140.
- Haskel, J. E.; S. C. Pereira y M. J. Slaughter. *Does Inward Foreign Direct Investment Boost the Productivity of Domestic Firms?* National Bureau of Economic Research Working Paper 8724. National Bureau of Economic Research, 2002 (DE) <http://discovery.ucl.ac.uk/18102/1/18102.pdf>, consultado el 10 de agosto de 2018.
- INEGI. *Indicadores de ocupación y empleo. Cifras oportunas durante junio de 2018*. Comunicado de prensa núm. 305 del 26 de julio de 2018 (DE) http://www.beta.inegi.org.mx/contenidos/saladeprensa/boletines/2018/iooe/iooe2018_07.pdf, consultado el 10 de agosto de 2018.
- Kuykendall, L. y L. Tay. "Employee subjective well-being and physiological functioning: An integrative model", en: *Health Psychology Open*. 2(1), 2015.

- Loayza, N. V.; L. Servén y N. Sugawara. *Informality in Latin America and the Caribbean*. Policy Research Working Paper Series No. 4888. The World Bank, 2009.
- Markusen, J. R. y A. J. Venables. "Foreign Direct Investment as a Catalyst for Industrial Development", en: *European Economic Review*. 43(2), 1999, pp. 335-356.
- Mishra, A. y R. Ray. *Informality, Corruption and Inequality*. Bath economics research papers, working paper no. 13/10. University of Bath, 2010 (DE) <http://www.bath.ac.uk/economics/research/working-papers/2010-papers/13-10.pdf>, consultado el 10 de agosto de 2018.
- _____ *Informality and corruption*. Bath papers in international development and well-being, working paper no. 21. University of Bath, 2013 (DE) <http://www.bath.ac.uk/cds/publications/bpd21.pdf>, consultado el 10 de agosto de 2018.
- OIT. *Women and men in the informal economy: a statistical picture, employment sector*. Italia, International Labour Office, 2002.
- _____ *World of work report 2014: developing with jobs*. Italia, International Labour Office, 2014.
- Smarzynska, B. "Does Foreign Direct Investment Increase the Productivity of Domestic Firms? In Search of Spillovers through Backward Linkages", en: *The American Economic Review*. 94(3), 2004, pp. 605-627.
- Temkin, B. "The Negative Influence of Labor Informality on Subjective Well-Being", en: *Global Labour Journal*. 7(1), 2016, pp. 69-93.
- Temkin, B. y G. Bensusán. *México en el escenario latinoamericano: informalidad y precariedad*. Presentación en el Foro Internacional: El empleo informal y precario: causas, consecuencias y posibles soluciones. México, 24 y 25 de noviembre de 2014 (DE) <https://es.slideshare.net/ForoEmpleo/mesa-1-graciela-bensusn-y-benjamn-temkin>, consultado el 20 de enero de 2018.
- Temkin, B. y J. Cruz Ibarra. "Las dimensiones de la actividad laboral y la satisfacción con el trabajo y con la vida: el caso de México", en: *Estudios Sociológicos*. 36(108), 2018, pp. 507-538.

Obteniendo indicadores de actividad económica municipal basados en información representativa de los Censos Económicos

Obtaining Municipality-Level Economic Activity Indicators Based on Representative Economic Censuses information

Francisco de Jesús Corona Villavicencio* y Jesús López-Pérez**



* Instituto Nacional de Estadística y Geografía (INEGI), franciscoj.corona@inegi.org.mx

** INEGI, jesus.lopez@inegi.org.mx

Nota: nuestra gratitud por los comentarios y sugerencias realizadas por el personal de la Dirección General Adjunta de Cuentas Nacionales del INEGI, en particular a Francisco Guillén, Lourdes Mosqueda, Leonel García y Jaime Rodríguez; asimismo, se agradece el acceso al Laboratorio de Microdatos del Instituto, en especial a Natalia Volkow; por último, nuestro agradecimiento a Gerardo Leyva Parra, director general adjunto de Investigación por sus valiosas aportaciones.

Con el objetivo de producir información económica a nivel municipal con una temporalidad y desagregación sectorial pertinente —la cual permita a quienes toman decisiones generar políticas económicas—, en este trabajo se presenta una metodología para desarrollar indicadores de actividad económica (IAE) a nivel municipal del 2003 al 2013 para todos los municipios de México, la cual se basa en utilizar la Regla de Combinación de Guerrero y Peña (2003) e información regionalmente representativa de los Censos Económicos (CE) del Instituto Nacional de Estadística y Geografía. Con el fin de tener una medición de la incertidumbre asociada al hacer uso de la metodología propuesta, se construyen IAE estatales, los cuales son comparados con el Producto Interno Bruto de cada entidad publicados por el Sistema de Cuentas Nacionales de México. Se puede concluir que el error de estimación se debe sobre todo a que los CE no incorporan el total de la economía; no obstante, tienden a reflejar de buena manera los principales sectores económicos. Se ilustran los resultados para los municipios económicamente más representativos del país.

Palabras clave: error de estimación; interpoladores; Producto Interno Bruto; Regla de Combinación; Valor Agregado Censal Bruto.

Recibido: 26 de diciembre de 2018
Aceptado: 12 de febrero de 2019

1. Introducción

Hoy en día existe una creciente demanda de información económica con un nivel de desagregación regional y temporal mayor a la publicada de forma oficial. En la actualidad, para muchos indicadores se cuenta con datos que pueden llegar a ser, incluso, solo de nivel nacional y con una temporalidad quinquenal, lo cual imposibilita al sector académico y a quienes toman decisiones realizar políticas económicas con una mejor precisión. Por tal motivo, tener información más desagregada permite, por ejemplo, elaborar un mayor número de análisis económicos y econométricos, lo cual puede traducirse en la implementación de políticas públicas en beneficio de la sociedad. En consecuencia, dado el gran despliegue de recursos económicos, huma-

In order to produce economic information at the municipal level with appropriate timing and sectorial breakdown —which allows decision makers to generate economic-policies— this paper presents a methodology for developing indicators of economic activity (IAE) at the municipal level 2003 to 2013 for all municipalities in Mexico. It is based on the combination rule of Guerrero and Peña (2003) and regionally information representative of the Economic Census (EC) of the National Institute of Statistics and Geography. In order to have a measure of the uncertainty associated with the use of the proposed methodology IAEs at state level are constructed, which are compared to the gross domestic product of each entity published by the System of National Accounts of Mexico. It can be concluded that the estimation error is mainly due to the fact that the EC does not include the total economy; however, they tend to reflect in a good way the main economic sectors. The results are illustrated for the most economically representative municipalities in the country.

Key words: Estimation error; Interpolation techniques; Gross Domestic Product; Combination Rule; Gross Census Value Added.

nos y materiales necesarios para obtener indicadores económicos de manera directa —sea a través de los Censos Económicos (CE) o encuestas— hoy en día ha surgido una gran cantidad de trabajos que implementan técnicas de estimación para la obtención de indicadores con mayor nivel de desagregación.

Con el objetivo de realizar un recuento técnico de lo existente en la literatura relacionada con resolver esta problemática, podemos dividir entre los métodos de estimación de áreas pequeñas (SAE, por sus siglas en inglés) y los de desagregación temporal, cada uno de los cuales permite generar indicadores económicos con las características de desagregación sectorial, geográfica o temporal deseadas.

Dentro del primer grupo de trabajos, podemos mencionar los desarrollados por Battese *et al.* (1988), Casas-Cordero *et al.* (2015), Correa *et al.* (2012), Drew *et al.* (1982), Elbers *et al.* (2003), Marhuenda *et al.* (2018), Molina y Morales (2009), Molina *et al.* (2014), Molina y Rao (2010), así como Torabi y Rao (2014), entre muchos otros, quienes sobre todo generan indicadores de pobreza, desigualdad y fuerza laboral. Cabe mencionar que estos estimadores cuentan con una desagregación regional mayor a los datos publicados de manera oficial, aunque vale la pena destacar que la frecuencia temporal corresponde a la periodicidad con la cual se actualizan estas fuentes de información, por ejemplo, cada cinco años, de acuerdo con los periodos censales e intercensales. Existe la posibilidad de utilizar datos provenientes de registros administrativos, sin embargo, dado que su incorporación como elementos de análisis y estimación es reciente, restringen que estos datos puedan generar indicadores con una temporalidad histórica mayor.

También, vale la pena destacar el trabajo de Suárez-Campos *et al.* (2015) donde se comparan los resultados de la estimación directa y los obtenidos a través de modelos lineales mixtos con el ingreso promedio mensual por trabajo en la vivienda para todos los municipios de México; lo anterior, a partir de los datos proporcionados por la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH) 2010.

Por otra parte, entre las investigaciones enfocadas a desagregar por temporalidad indicadores económicos desde una perspectiva de análisis de series de tiempo, podemos mencionar a Chow y Lin (1971), Denton (1971), Nieto (1998), Guerrero y Nieto (1999), Di Fonzo (1990) y el muy recurrente método utilizado en estadística oficial, conocido como de Denton-Cholette (ver Dagum y Cholette, 2006).

En fechas recientes, Guerrero y Corona (2018a, 2018b) propusieron el uso de técnicas de empalme de series de tiempo, desagregación temporal y retropolación restringida para generar series esta-

tales y trimestrales del Producto Interno Bruto (PIB) por Gran Actividad Económica (GA). El requisito indispensable tanto en los trabajos de SAE como de desagregación temporal es utilizar variables auxiliares, las cuales deben tener un carácter más oportuno en lo que se refiere a su actualización y, a la par, es fundamental que estén correlacionadas con los indicadores que se desea estimar o desagregar.

En este contexto, hay una disyuntiva en las agencias generadoras de información oficial sobre hasta qué punto un dato estimado o desagregado puede ser o no publicado como dato oficial e, incluso, solo como cifra experimental, pues tanto las técnicas de SAE como de desagregación temporal conllevan a la estimación de una variabilidad aludida al método de estimación que, muchas veces, puede generar que lo estimado no sea del todo fiable.

En lo que respecta a los métodos de SAE, Molina (2018) menciona que, aunque no existe un consenso absoluto sobre el coeficiente de variación (CV) permitido —entendido éste como el grado de variabilidad atribuible al método de estimación—, los institutos nacionales de estadística no deben de publicar, en ningún caso, un CV superior a 20 por ciento. Nótese que, en este caso, el CV depende directamente del tamaño de la muestra y se espera que, conforme éste sea más grande, el error atribuible al método de estimación sea más pequeño; no obstante, esto puede implicar un gran uso de recursos económicos y humanos haciendo, en muchos casos, que el cálculo del indicador sea inviable.

En cuanto a los métodos de desagregación temporal, el enfoque puede llegar a ser diferente pues, por ejemplo, el más utilizado en estadística oficial —Denton-Cholette— no genera una medida en relación con la variabilidad atribuible al método de estimación. Guerrero y Corona (2018a) enfatizan en la importancia de estimar la matriz de covarianza en el proceso de desagregación temporal y es esperable que, conforme la variable auxiliar tenga una mayor relación con la variable a desagregar, el error asociado al método de estimación,

será menor. También, es importante considerar que los supuestos que subyacen de los modelos de series de tiempo llevan a tomar en cuenta hechos estilizados y con mucha frecuencia recurrentes en series de tiempo económicas, como los conceptos de no estacionariedad y cointegración. En ambos enfoques metodológicos es importante que los datos validen los supuestos de los modelos implementados para los objetivos particulares.

En lo que concierne a la generación de indicadores de actividad económica (IAE) regional puntuales, con un nivel de desagregación municipal y un sentido temporal anual para México, podemos situar el problema de estimación en un punto intermedio entre las técnicas de SAE y los procedimientos de desagregación temporal pues, por un lado, existe información auxiliar con un nivel de agregación regional mayor, cuya temporalidad es anual (PIB estatal) y, por el otro, hay variables regionales que permitirían desagregar a nivel municipal, provenientes de los CE realizados cada cinco años. En este sentido, nótese que si el PIB municipal es el vector que se desea estimar, podemos hacer uso de información preliminar proveniente de los CE, de tal forma que se puede imponer la restricción de que la suma del PIB municipal sea equivalente al estatal, con un nivel de desagregación regional, sectorial y temporal pertinentes. La formalización de esta idea está dada por Guerrero y Peña (2003) y es conocida como la Regla de Combinación (RC), la cual, dado que permite la inclusión de restricciones contemporáneas (sectoriales o geográficas) y temporales, representa una alternativa coherente para dar solución a este problema.

Algunos trabajos empíricos que anteceden a éste son los de Ebener *et al.* (2005), que resaltan la importancia de la estimación de datos subnacionales para la cuantificación de la pobreza, ya que la agregación a nivel nacional tiende a enmascarar las variaciones subnacionales. Para este propósito, hacen uso de imágenes satelitales para mostrar que el tamaño de superficie iluminada y el porcentaje de la frecuencia de luminosidad pueden predecir el PIB per cápita a nivel nacional y subnacional.

También Henderson *et al.* (2012) proponen un enfoque alternativo y novedoso para medir el crecimiento económico utilizando series de tiempo de luminosidad nocturna proveniente de imágenes satelitales. Con base en este último, han surgido diferentes aportaciones y, para el caso de México, se encuentra el trabajo reciente de Guerrero y Mendoza (2018) y el de Llamosas-Rosas *et al.* (2018). Ambos utilizan la base de datos de la oficina de Administración Nacional Oceánica y Atmosférica para obtener información sobre luminosidad nocturna; el primero para estimar el crecimiento del PIB por países de 1993 al 2008 y, en particular, para México, China y Chile; el segundo, para determinar el crecimiento económico de las principales 15 zonas turísticas de playa en México de 1993 al 2017.

Sin embargo, ninguno de estos trabajos incluye restricciones oficiales de carácter contable. Vale la pena mencionar que el Buró del Censo de los Estados Unidos de América calculó de manera directa el PIB por zona metropolitana del 2002 al 2017, mientras que el Banco Central del Ecuador publica anualmente, desde el 2007, cifras del valor agregado bruto cantonal para las 221 provincias del país, calculadas con base en registros contables.

En esta investigación, el objetivo principal es estimar los IAE para todos los municipios y delegaciones¹ de México del 2003 al 2013 bajo un enfoque de estimación, usando la RC de Guerrero y Peña (2003), basado en el uso de información oficial y recomendaciones realizadas por el Sistema de Cuentas Nacionales de México (SCNM) del Instituto Nacional de Estadística y Geografía (INEGI) para contar con un primer acercamiento sobre la actividad económica a nivel municipal de manera fidedigna, cuya metodología pueda ser replicable y utilizada por los diferentes tomadores de decisiones de política económica.

Dado que no es objetivo del trabajo realizar un análisis estructural de los resultados sino, más bien, con fines metodológicos, dicha metodología pro-

¹ Hoy demarcaciones territoriales de la Ciudad de México.

puesta se valida en una aplicación real donde se puede estimar un error fuera de muestra y también se realiza una aplicación empírica con los datos municipales estimados.

En la segunda sección del artículo se describe la metodología usada para generar IAE municipal utilizando la información de los CE y del SCNM. En la tercera, se presentan los resultados de aplicar la metodología sugerida, primero analizando su funcionamiento utilizando datos estatales, evaluando el error fuera de muestra y, después, se producen IAE para los 20 municipios más importantes de las tres GA y el total. En la cuarta, se revisa la consistencia de los IAE municipales a través de un análisis de la evolución económica de los estados. Por último, en la quinta se muestran las conclusiones y consideraciones finales.

2. Metodología

Aquí, primero se realizan consideraciones de carácter metodológico, las cuales fueron sugeridas por el SCNM del INEGI para obtener IAE que reflejen las condiciones económicas municipales de mejor manera. Después, partiendo del hecho de que los CE se levantan cada cinco años, se establece la metodología para generar IAE municipales.

2.1 Consideraciones metodológicas

Para la formación de los IAE —considerando las sugerencias ya mencionadas—, en primera instancia se delimita trabajar con información del Valor Agregado Censal Bruto (VACB) ya que, técnicamente, representa el PIB antes de impuestos. También, se busca equivalencia y representatividad regional entre la información de los CE y de las cuentas nacionales. Para ello, se toman en cuenta solo sectores de los CE que durante el proceso de recolección de datos reflejan las condiciones económicas de las entidades federativas y de los municipios, ya que después, durante el proceso de integración de la información, se agrupan los datos en unidades económicas concentradoras (matrices, conglomerados,

grupos financieros, etc.) bajo cierta delimitación geográfica, pues no necesariamente el domicilio fiscal es equivalente al lugar geográfico donde se realiza la actividad económica. A manera de ejemplo, en el sector 52. *Servicios financieros y de seguros*, gran parte de su actividad se reporta en la Ciudad de México, ya que ahí se encuentra la mayoría de las empresas corporativas; no obstante, esto no significa que no haya actividad de dicho sector en el resto de las entidades. De esta manera, para la formación de los IAE, los siguientes sectores no son incluidos:

- 23. Construcción.
- 48-49. Transportes, correo y almacenamiento.
- 52. Servicios financieros y de seguros.
- 55. Corporativos.

A la par de ellos, tampoco se considera el 93. *Actividades legislativas, gubernamentales, de impartición de justicia y de organismos internacionales y extraterritoriales*, ya que los CE no reportan información del VACB para este sector.

Para el 11. *Agricultura, cría y explotación de animales, aprovechamiento forestal, pesca y caza*, se recurre a las bases de datos del Sistema de Información Agroalimentaria y Pesquera (SIAP) sobre agricultura y ganadería (SAGARPA, 2018), para complementar la información de los CE, que solo consideran silvicultura y acuicultura. Asimismo, la información proveniente de los CE considerados (INEGI, 2004, 2009 y 2014) está libre de confidencialidad, es decir, se recurrió al Laboratorio de Microdatos del INEGI para producir la información que permite generar los IAE con la representatividad sectorial y regional deseadas.

Cabe anotar que, dada la naturaleza de los CE, los IAE generados están dando señales de la economía relacionadas con las unidades económicas fijas y semifijas, es decir, no consideran a otras, como las del ambulante ni las casas-habitación donde se efectúa una actividad productiva para autoconsumo o se ofrecen servicios que se realizan en otro sitio, etc.; en otras palabras, los CE no consideran el total de la economía. Así, existe una clara distinción

entre un IAE y el PIB, donde este último sí considera toda la estructura de la economía. Para fines de comparación estadística en la medición del error, se hace uso indistinto entre los IAE estimados y los PIB obtenidos a través del SCNM.

2.2 Obteniendo los IAE municipales con datos de los CE

Supongamos un vector de estimaciones preliminares del PIB municipal dado por $W^* = (W_1^*, \dots, W_H^*)'$ con $W_h^* = (W_{1,h}^*, \dots, W_{N,h}^*)'$ para los años censales $h = 1, \dots, H$ con $W_{i,h}^* = \Omega_h Y_{ih}$, donde $\Omega_h = \frac{VACB_{i,h}}{\sum_{i=1}^N VACB_{i,h}}$; es decir, W_h^* son vectores formados por las razones de VACB a nivel municipal para algún sector en específico, dividido entre el VACB del total del estado y multiplicado por $Y_{i,h}$ que, en este caso, representa el PIB estatal solo para los años censales.

Por tal motivo, ya que W^* no está definida para cualquier año $t = 1, \dots, T$, donde se tiene la relación $t = 4h - 3$, es necesario realizar aproximaciones a través de interpoladores —denotados como f_s —, los cuales pueden ser lineales o no. De esta manera, podemos caracterizar al nuevo vector de estimaciones preliminares de la siguiente forma:

$$W = f_s(W^*), \quad (1)$$

donde $W = (W_1', \dots, W_T')'$ con $W_t = (W_{1,t}, \dots, W_{N,t})'$. De manera matemática, si q es un número entero, entonces las estimaciones se denotan como:

$$f_s(W^*) = \begin{cases} W_t' = W_q^{*'} & \text{si } \frac{t+3}{4} = q, \\ W_t' = f_s(W_{q-1}^{*'}, W_q^{*'}) & \text{si } \frac{t+3}{4} \neq q. \end{cases} \quad (2)$$

Es decir, la información inexistente entre $W_{q-1}^{*'}$ y $W_q^{*'}$ se obtiene a través de utilizar una interpolación f_s que, en este particular caso, se le aplica a los cuatro puntos faltantes entre cada año censal. Se denotan tres diferentes f_s , primero se propone el uso del interpolador lineal $S_{L'}$, donde se unen los puntos a través de una proporcionalidad lineal en-

tre los datos faltantes, es decir la utilizada con más frecuencia en trabajos similares, como por ejemplo en Germán-Soto (2005). También, se utiliza el procedimiento sugerido por Forsythe *et al.* (1977), $S_{SP'}$, donde se ajusta una posición cúbica exacta a través de los cuatro puntos en cada extremo de los datos y el algoritmo de Stineman (1980), $S_{ST'}$, donde se computan pendientes más bajas cerca de pasos o picos abruptos en la secuencia de puntos. Para lo anterior, se usa la librería *imputeTS* del programa estadístico R (R Core Team, 2017).

Nótese que este procedimiento trasladado al problema de Guerrero y Corona (2018a, 2018b) correspondería a la fase de desagregación temporal.

En consecuencia, sea Z el PIB municipal para algún sector, podemos establecer la siguiente relación:

$$Z = W + D, \quad (3)$$

donde Z y D tienen dimensiones iguales a W , siendo D un vector de discrepancias con $E(D/W) = 0$. Asimismo, se establece la siguiente relación conocida:

$$Y = cZ, \quad (4)$$

donde $c = (1, \dots, 1)'$ es un vector de dimensión N , de tal forma que la suma de los PIB municipales estimados sean equivalentes al PIB estatal observado.

Siguiendo a Guerrero y Corona (2018b), quienes utilizan la RC para reconciliar cifras preliminares a un total oficial, la estimación óptima resultante es de la forma $\tilde{Z} = W + A(Y - cW)$ con $A = c'(cc')^{-1} = \frac{1}{N}c'$, de manera que se obtiene la solución:

$$\tilde{Z} = W + \frac{1}{N}c'(Y - cW), \quad (5)$$

que satisface la restricción, pues $c\tilde{Z} = cW + \frac{1}{N}cc'(Y - cW) = Y$.

Ahora, en lugar de usar la misma ponderación para todos los municipios para distribuir la discrepancia entre el dato estatal Y y la suma de los valo-

res municipales cW —que es lo que ocurre al usar $\frac{1}{N}c'$ —, se propone utilizar un promedio ponderado donde la ponderación es la proporción del sector de cada municipio en el total estatal. Esto se logra al definir las ponderaciones:

$$p_{it} = W_{it} (cW)^{-1} \text{ para } i = 1, \dots, N, \quad (6)$$

donde $W_{i,t}$ es el valor del sector en el municipio i y el año t , proveniente de la base de datos estatal estimada mediante el uso de interpoladores, mientras que $cW = \sum_{i=1}^N W_{i,t}$ es el total estatal estimado del sector en consideración. Por lo tanto, en la expresión por usar se sustituye al vector $\frac{1}{N}c' = (\frac{1}{N}, \dots, \frac{1}{N})'$ por $p = (p_{1,t}, \dots, p_{N,t})' = W(cW)^{-1}$, con lo que se obtiene la solución buscada:

$$\hat{Z} = W + p(Y - cW), \quad (7)$$

que cumple con la restricción contable, o sea, $c\hat{Z} = cW + cp(Y - cW) = Y$, debido a que $cp = \sum_{i=1}^N W_{i,t} (cW)^{-1} = 1$. En este caso, \hat{Z} representa al IAE municipal. De manera alternativa, se pudo proponer el uso del método RAS de tal forma que, sabiendo que, en teoría, las participaciones $\sum_{t=1}^T \Omega_i = 1$ para toda i , se pueden ajustar las discrepancias contables utilizando dicho método. En este trabajo se eligió mantener el concepto estadístico de la reconciliación de cifras oficiales con base en restricciones contemporáneas, tanto sectoriales como geográficas.

3. Aplicación empírica

3.1 Indicadores de actividad económica estatal: evaluación del error fuera de muestra

Para evaluar el funcionamiento de las técnicas sugeridas en la sección 2.1, se estima Z usando datos estatales, $\hat{Z}_{i,t}^E$, estimando el siguiente estadístico:

$$MAPE = (MAPE_1, \dots, MAPE_N), \quad (8a)$$

$$MAPE_i = \left| \frac{\hat{Z}_{i,t}^E - Y_{i,t}}{\sum_{t=1}^T Y_{i,t}} \right| \frac{1}{T} \quad (8b)$$

donde $MAPE_i$ es la media absoluta del porcentaje de error estimado para la entidad federativa i en algún sector en particular.

Para resumir los resultados, el $MAPE_i$ se presenta como un estadístico ponderado de acuerdo con el peso que tiene cada una de las i entidades, de tal manera que, para cada sector, tenemos lo siguiente:

$$MAPE^p = cpMAPE. \quad (9)$$

Si consideramos las tres GA podemos, asimismo, tener el siguiente error ponderado:

$$MAPE^{p\omega} = \omega (MAPE_{GA1}^p, MAPE_{GA2}^p, MAPE_{GA3}^p), \quad (10)$$

donde $\omega = \frac{1}{PIB} (GA1, GA2, GA3)$. Nótese que, en este caso, podemos utilizar como variable preliminar a $W_{i,h}^* = \Omega_h^* Y_{i,h}$, donde $\Omega_h^* = \frac{PIB_{i,h}}{\sum_{i=1}^N PIB_{i,h}}$, de tal manera que no existe error de interpolación atribuible a $W_{i,h}^*$ a través de la información de los CE. En otras palabras, es el error atribuible solo al método de estimación ya que, para el 2003, 2008 y 2013 no existen errores generados por las diferencias entre los valores de los CE y los provenientes del SCNM. En esencia, se está suponiendo que el VACB es equivalente al PIB. En el segundo caso se considera Ω_h , de tal manera que el error es atribuible tanto al método de interpolación como a las diferencias existentes entre la información proporcionada por CE y la que reporta el SCNM.

En el cuadro 1 se muestran los errores de estimación considerado por GA y para el total ponderado para cada uno de los interpoladores utilizados.

Se puede apreciar que el error ponderado atribuible a los interpoladores es muy pequeño, ya que alcanza un máximo de 0.048 para el S_{sp} en la GA1, obteniendo en todos los casos el menor o igual error para el procedimiento S_L , es decir, para el interpolador lineal, donde se obtiene un $MAPE^{p\omega}$ de 0.020. Cuando se asume $W_{i,h}^* = \Omega_h^* Y_{i,h}$, se puede apreciar que los errores se incrementan en todos los casos; no obstante, el menor sigue siendo cuan-

Cuadro 1

Porcentajes absolutos de error para cada método de interpolación por GA y total ponderado

Método	$W_{i,h}^* = \Omega_h^* Y_{i,h}$				$W_{i,h}^* = \Omega_h Y_{i,h}$			
	GA1	GA2	GA3	Total	GA1	GA2	GA3	Total
S_L	0.045	0.038	0.008	0.020	0.118	0.194	0.188	0.188
S_{SP}	0.048	0.040	0.009	0.021	0.118	0.197	0.197	0.195
S_{ST}	0.047	0.039	0.009	0.021	0.118	0.196	0.196	0.193

do se utiliza el método S_L , que en este caso se obtiene un $MAPE^{Pw} = 0.188$.

En términos económicos, esperaríamos que, si las desigualdades económicas entre las entidades que componen al país se mantienen relativamente igual de heterogéneas entre los municipios que forman las entidades, los errores atribuibles al método de estimación y al uso de los CE serían similares a los obtenidos en la última columna del cuadro 1, dependiendo del tipo de interpolador usado.

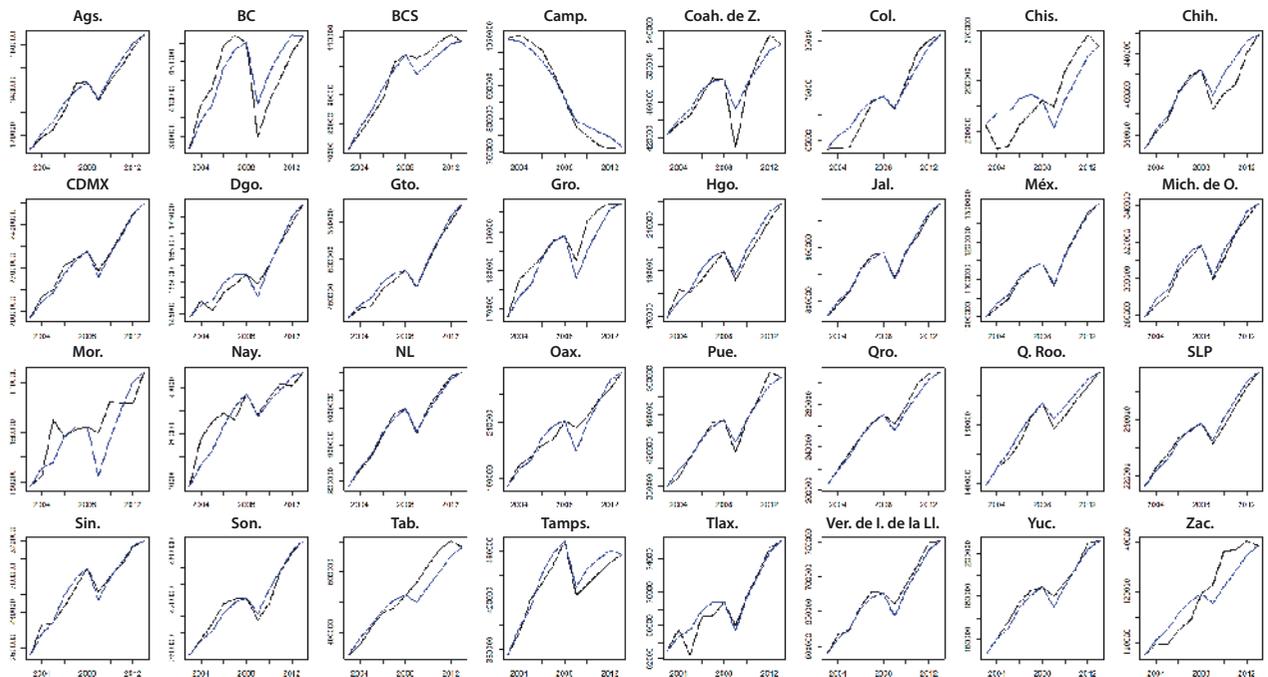
Para fines visuales, la gráfica 1 resume el comportamiento de los IAE estimados a través de S_L

para todos los estados, tomando en cuenta los niveles de la suma de los sectores económicos considerados y asumiendo que $W_{i,h}^* = \Omega_h^* Y_{i,h}$.

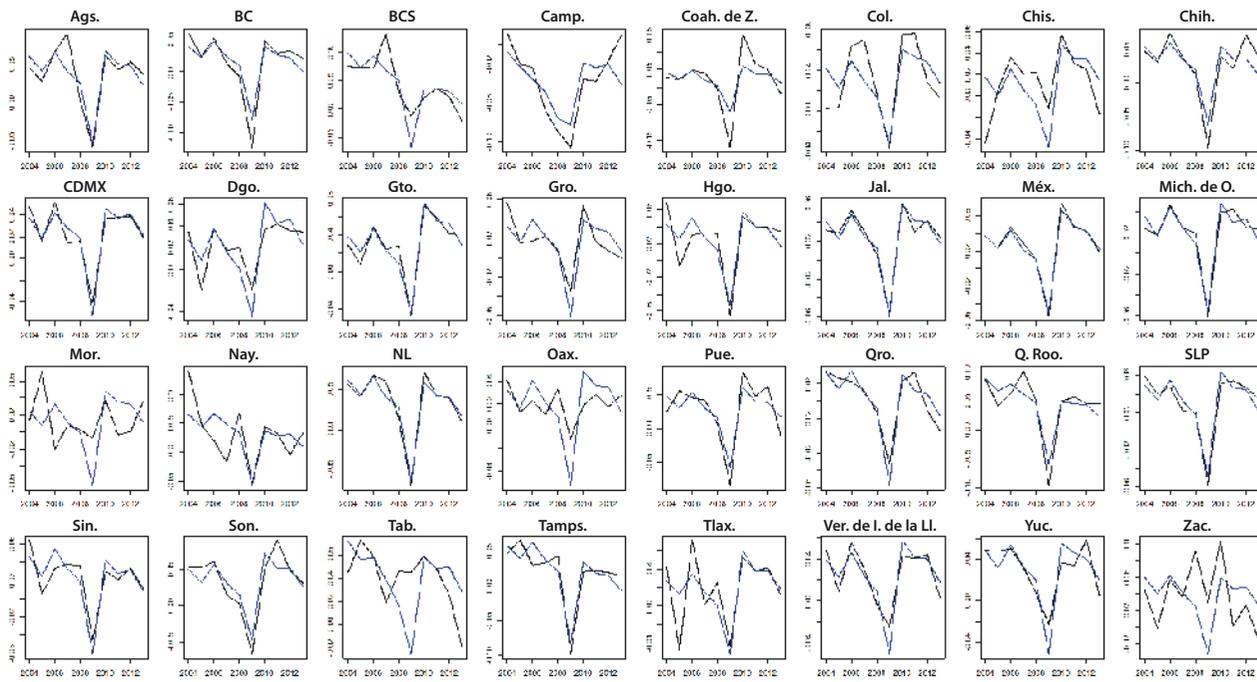
Se puede apreciar que, en todos los casos, los IAE tienen un comportamiento muy similar a los obtenidos por el SCNM, que básicamente son el reflejo de los pequeños errores mostrados en el cuadro 1.

Otro aspecto interesante es verificar si los IAE reflejan de buena forma los cambios anuales de los indicadores reportados en el SCNM. La gráfica 2 presenta los mismos resultados de la 1, pero para las series en cambios porcentuales.

Gráfica 1



Gráfica 2



Igual que en la 1, en la gráfica 2 se puede observar que los cambios porcentuales estimados son parecidos a los observados; solo en algunos casos como Nayarit, Tabasco, Tlaxcala y Zacatecas se muestran ciertas diferencias; no obstante, pueden considerarse como mínimas.

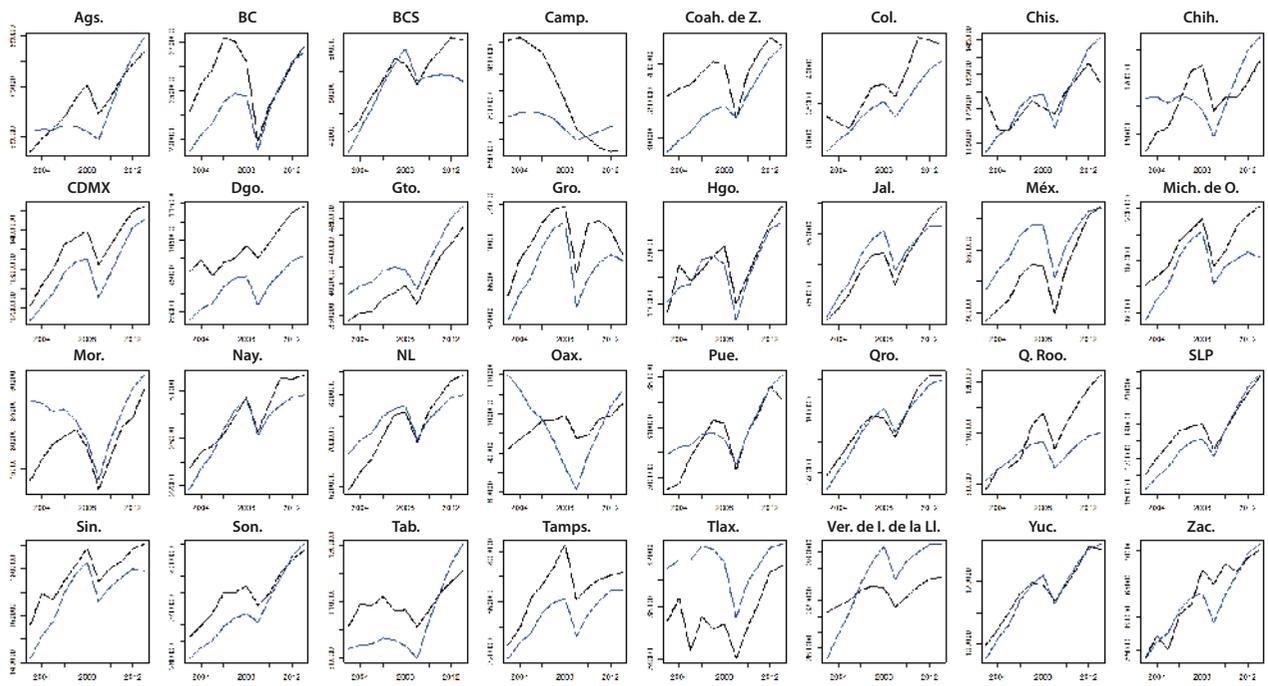
Ahora, es importante denotar los mismos resultados de las dos gráficas anteriores, pero cuando $W_{i,h}^* = \Omega Y_{i,h}$, cuya situación es la que se espera en la práctica. La gráfica 3 muestra los resultados para las series en niveles.

Se puede observar que los IAE ya no reflejan con perfección los movimientos de los indicadores obtenidos a través del SCN. Para fines ilustrativos, se estiman las correlaciones lineales para IAE con su respectivo PIB y la correlación mínima fue obtenida para Oaxaca, donde alcanzó -0.059, aunque el cuantil de 25% ya presenta una correlación de 0.827, siendo la mediana 0.898 y la máxima 0.990, es decir, con excepción de ciertas entidades, el comportamiento a niveles está siendo rescatado de forma razonable; no obstante, nótese que las correlaciones puede ser espurias a menos que las series

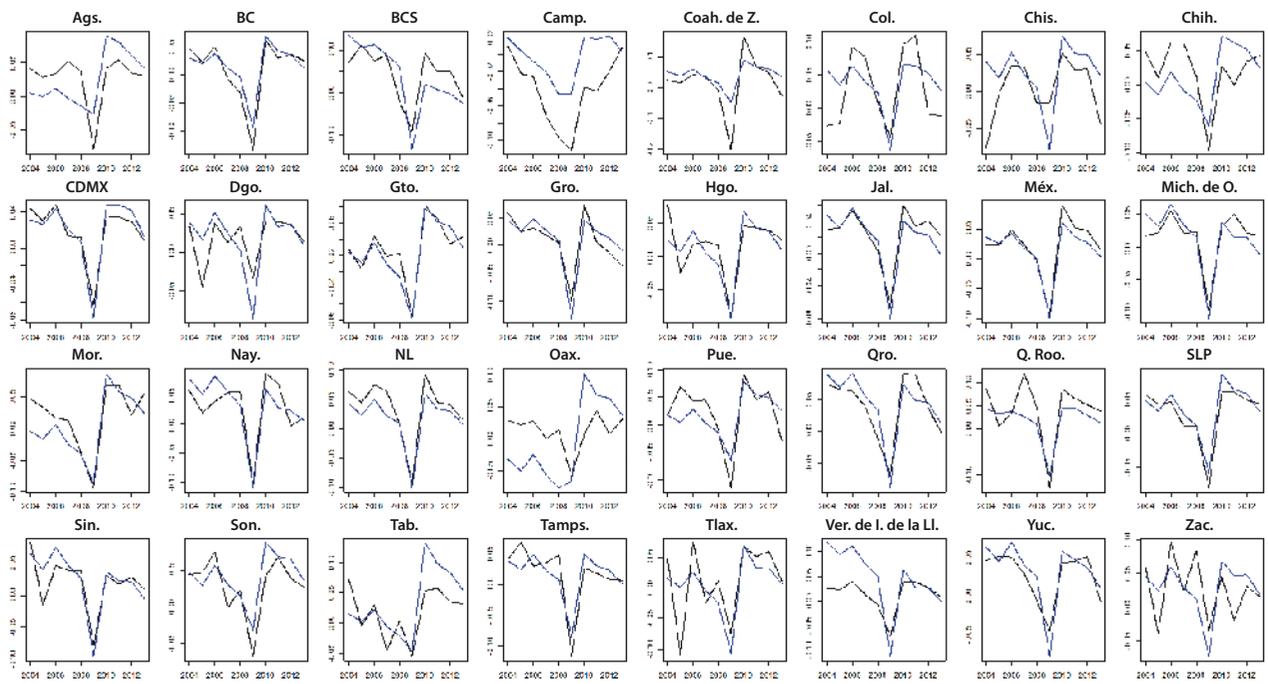
de tiempo sean estacionarias o cointegradas. En estos casos, es complicado argumentar de manera empírica acerca del orden de integración y, más aún, sobre la presencia de cointegración, debido a que la temporalidad de las series es relativamente corta, del 2003 al 2013. Por ende, una mejor forma de denotar la relación entre los IAE y los PIB es a través de las series en cambios anuales, debido a que se espera que éstas sí sean estacionarias, por lo que la correlación sería un mejor indicador de la relación lineal que guardan las estimaciones y las variables consideradas como observadas. Lo anterior se muestra en la gráfica 4.

Se puede apreciar que los IAE reflejan de forma certera los movimientos de sus respectivos indicadores obtenidos a través del SCN. Realizando el mismo ejercicio de correlaciones, se puede apreciar que la mínima se alcanza para Oaxaca, obteniendo un valor de 0.358; sin embargo, se puede ver que es positivo, lo cual nos indica que siguen el mismo sentido en sus movimientos. Ahora, la mediana está centrada en 0.859 y la máxima alcanza un valor de 0.974, por lo que los IAE sí reflejan los principales movimientos de sus respectivos PIB.

Gráfica 3



Gráfica 4



Es importante considerar que este ejercicio es netamente estadístico y, dentro de este contexto, conforme dentro de los municipios el VACB sea equivalente al PIB, los IAE tenderán a reflejar de mejor forma las verdaderas características de la

economía; no obstante, como primer acercamiento, los IAE tienden a darnos buenas señales de las economías municipales cuando en términos matemáticos se satisfaga que $\frac{APIB^{Nac}}{APIB^{Est}} \approx \frac{APIB^{Est}}{APIB^{Mun}}$, es decir, que las variaciones estatales de la economía que se re-

flejan a nivel nacional son similares al efecto que tienen las variaciones municipales en la economía, con su respectiva entidad.

En términos de error, si queremos hacer comparables los resultados con un CV en relación con las técnicas de SAE, podemos apreciar que los resultados indican una cercanía al *benchmark* de 20%; no obstante, no pueden ser directamente comparables pues, como ya se comentó, los IAE no están midiendo el total de la economía.

3.2 Indicadores de actividad económica municipal

Se estiman para todos los municipios² de México, tanto por GA como para el total de las economías, en niveles, miles de pesos y en primeras diferencias logarítmicas. Los resultados se presentan para los 20 municipios con mayor volumen de producción en el 2013, de acuerdo con cada GA y al total municipal. Los resultados completos están disponibles bajo petición.

3.2.1 Gran Actividad Económica 1

La gráfica 5 presenta los resultados en niveles de las series para los 20 municipios más representativos de la GA1 que, como se ha comentado con anterioridad, contemplan la información de los CE (acuicultura y silvicultura) y la proveniente del SIAP (agricultura y ganadería), por lo cual se puede tener certidumbre de que reflejan el total del sector primario.

Se puede apreciar que en los primeros 20 municipios se encuentran sobre todo aquellos que tienen costa, por ejemplo: Mazatlán, Ensenada, Alvarado, Ozulama de Mascareñas, Tapachula, Cabo Corrientes y Huatabampo, entre otros. Su participación con respecto al total de la GA1 es de 28.26 por ciento. La gráfica 6 presenta los mismos resultados de la 5, pero para las primeras diferen-

cias logarítmicas de las series, es decir, los cambios porcentuales anuales de los IAE.

Los resultados más importantes señalan que existen diferencias entre los municipios, lo cual obedece a la coyuntura municipal específica, aunque destacan los picos de San Salvador Huixcolotla y Jonuta, ocurridos en el 2009; de esta forma, el que cuenta con mayor tasa de crecimiento promedio durante el periodo analizado es el primero con 56%, aunque también es el que presenta una mayor desviación estándar equivalente a 122%; por otra parte, el de menor tasa de crecimiento es Ahome, la cual es negativa (-1.4%) y el que tiene una menor dispersión en su tasa de crecimiento promedio es Tapachula con 5%; las tasas de crecimiento promedio exhibidas por los dos municipios más importantes son para Mazatlán con 3.9% y Ensenada con 2.7 por ciento.

3.2.2 Gran Actividad Económica 2

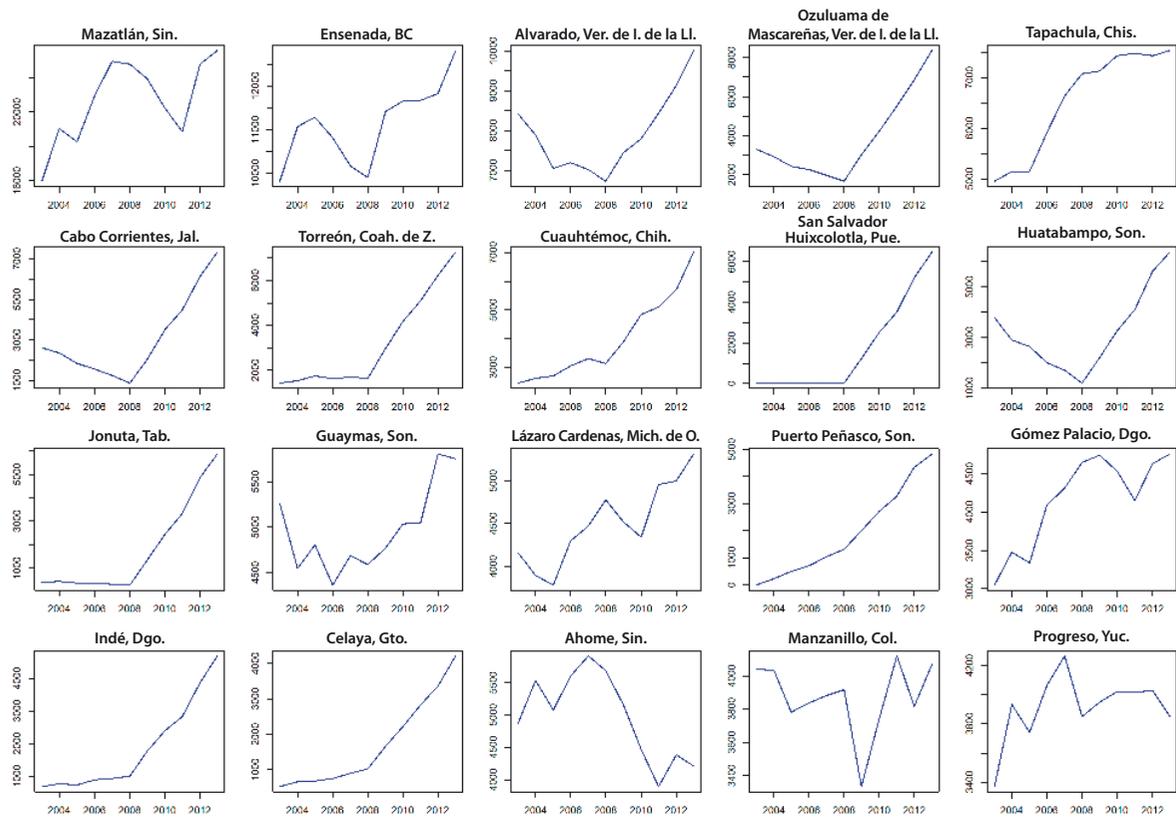
La gráfica 7 presenta los resultados en niveles para los 20 municipios con mayor participación dentro de la GA2, que abarcan a la minería, los energéticos y las manufacturas.

Los primeros dos lugares los ocupan municipios dedicados al petróleo, es decir, Carmen y Paraíso, seguidos de otros notablemente industrializados, como Juárez, Monterrey y Hermosillo. Estos 20 municipios abarcan 49.3% del total de la GA2. Algunos resultados interesantes se observan cuando ocurrió la crisis financiera del 2009, provocando tasas de crecimiento negativas para todos los municipios de orientación manufacturera, con excepción de Reforma y Escobedo. La gráfica 8 muestra estos mismos resultados para los cambios porcentuales de las series.

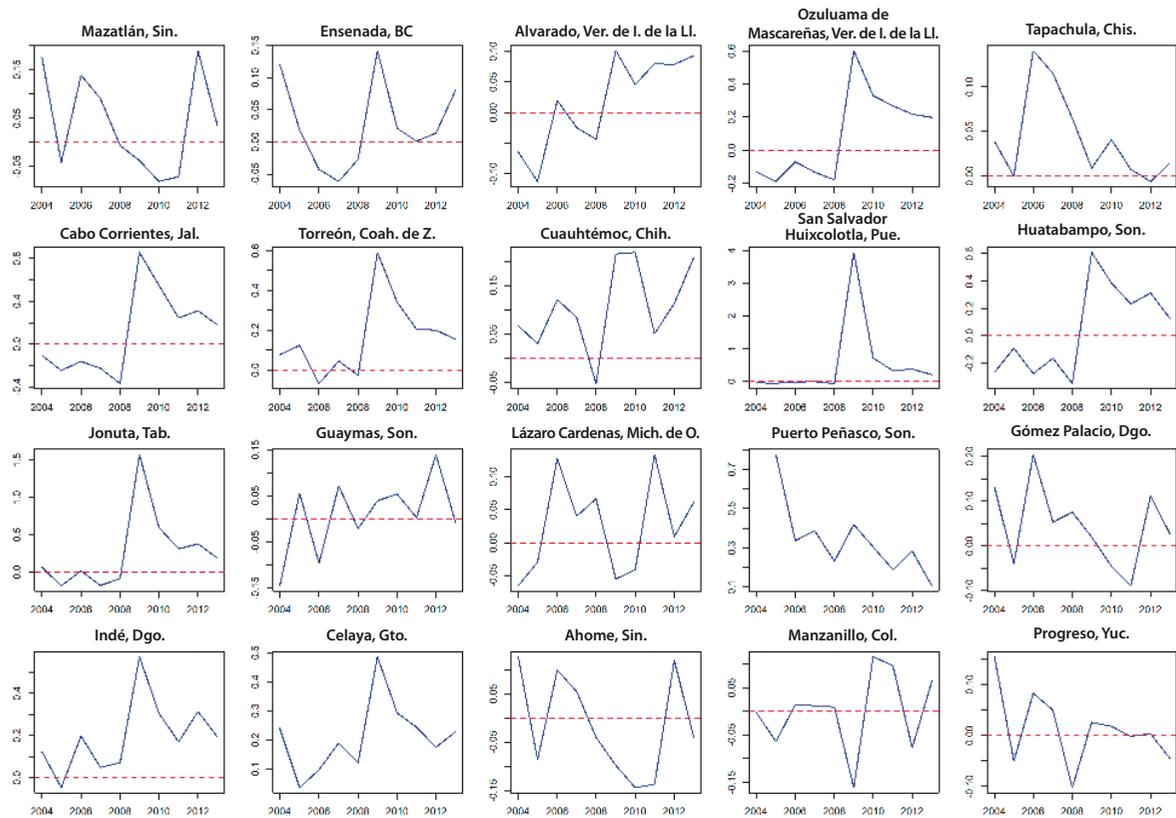
Para el grupo de sectores que componen a la GA2, se aprecia que, para muchos municipios, existe una caída en el año de la crisis económica y financiera del 2009. Los municipios que tienen tasas de crecimiento promedio anuales negativas para todo el periodo son Carmen, Ramos Arizpe, Guadalajara y Reforma con valores equivalentes

² Incluye las delegaciones (hoy demarcaciones territoriales) del entonces Distrito Federal (hoy Ciudad de México).

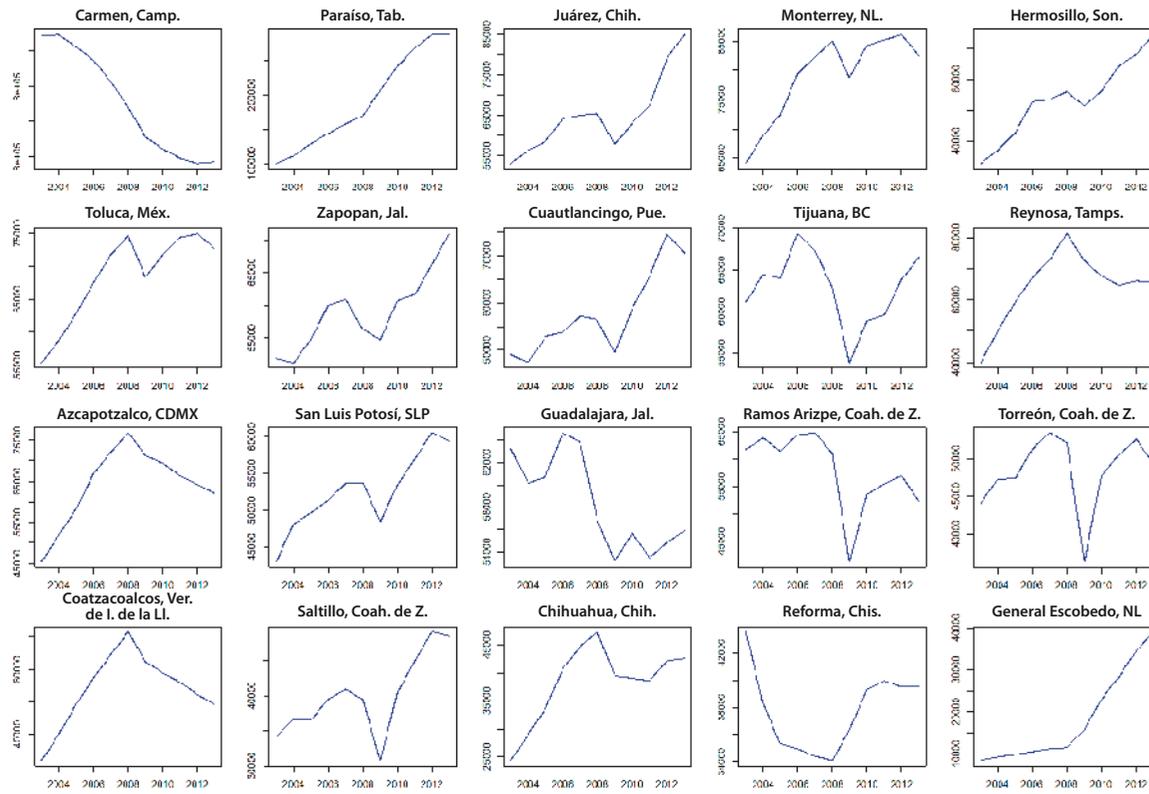
Gráfica 5



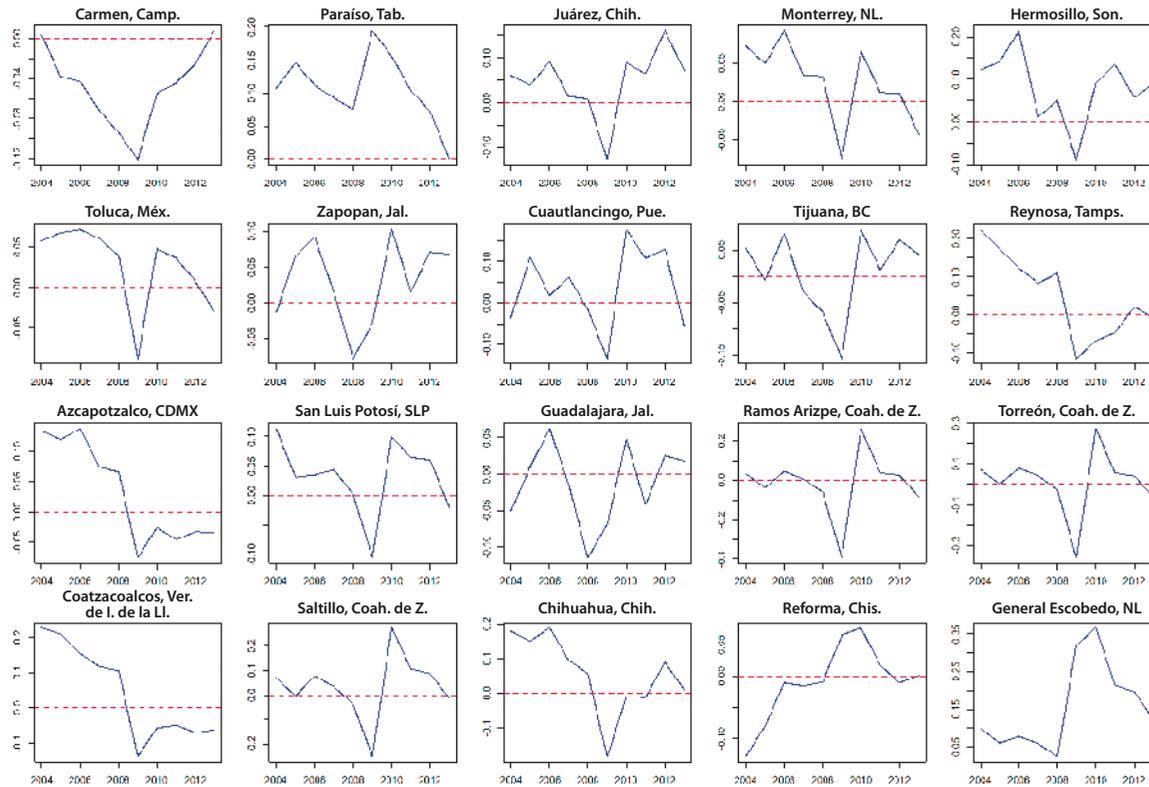
Gráfica 6



Gráfica 7



Gráfica 8



a -4.8, -1.7, -1.24 y -1%, respectivamente. Por otra parte, los que exhiben un mayor crecimiento son General Escobedo (15.4%) y Paraíso (10.6%). En términos de dispersión, es de interés observar que las desviaciones estándar son menores respecto a la GA1, siendo la menor la de Carmen con 4% y la mayor Ramos Arizpe con 16.2%, es decir, dos de los municipios con las tasas más bajas de crecimiento promedio anuales.

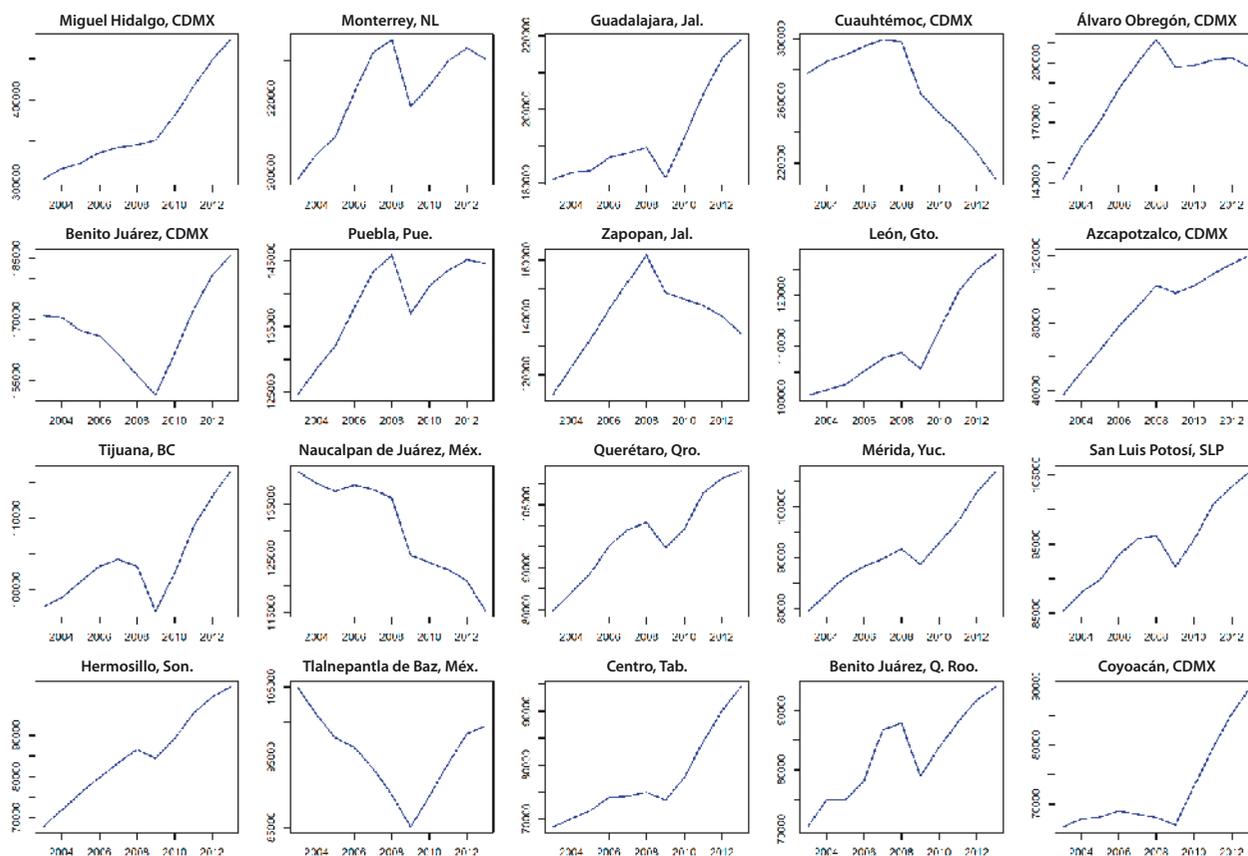
3.2.3 Gran Actividad Económica 3

Para este caso, como se comentó, las cifras estimadas no contemplan los sectores de transporte, servicios financieros ni corporativos, los cuales reportan su información de manera consolidada y es usual que sus matrices estén ubicadas en la Ciudad de México; tampoco consideran la valoración de los servicios que presta el gobierno. La gráfica 9 presenta los resultados para los niveles de las series.

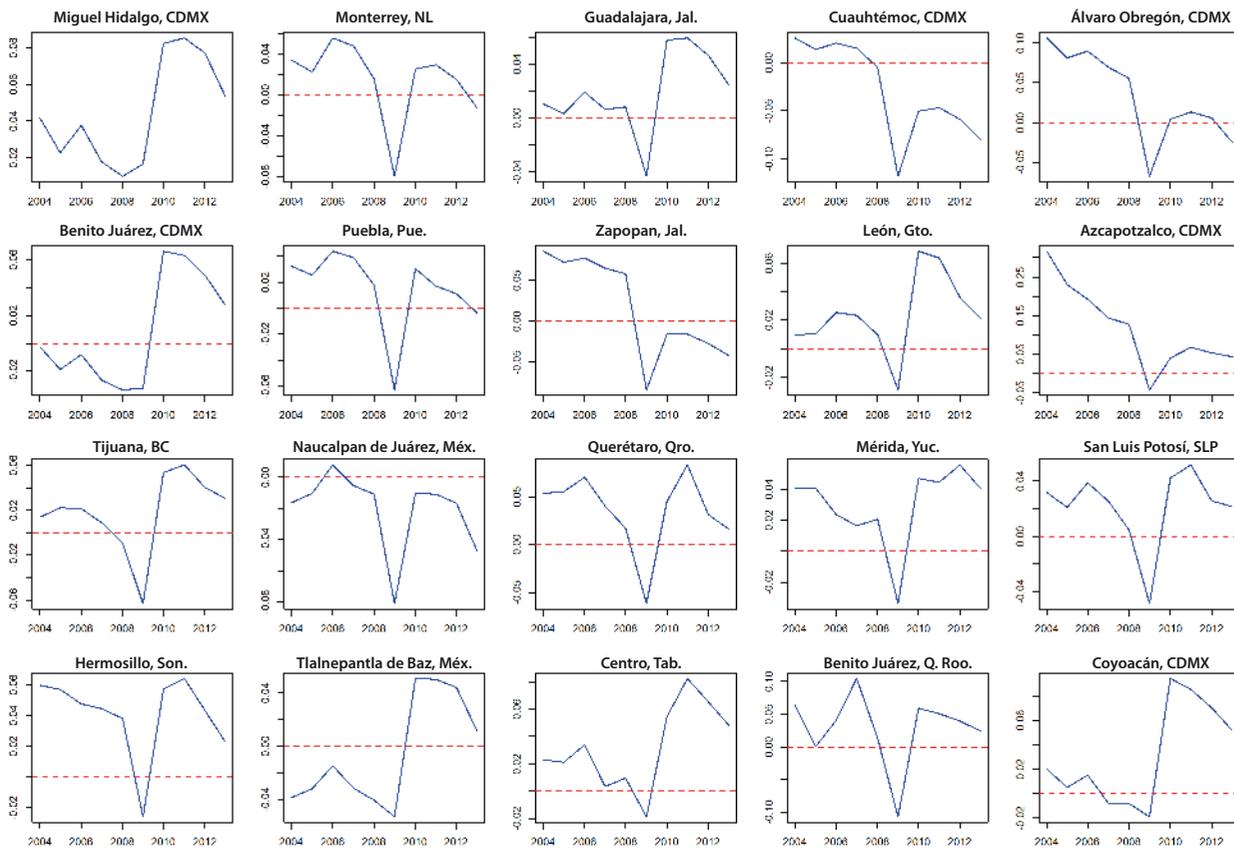
En este caso, los 20 municipios más representativos concentran 40.49% de la actividad económica. Destaca, también, que seis de ellos pertenecen a las hoy demarcaciones territoriales de la Ciudad de México (entonces Distrito Federal), a pesar de que no contempla sectores económicos de importancia. La gráfica 10 presenta los resultados para la diferencia logarítmica de las series.

Se puede apreciar el efecto que tuvo la crisis del 2009 en todos los municipios y delegaciones. Llama la atención el crecimiento promedio anual negativo que presentan la Cuauhtémoc (-2.8%), Naucalpan de Juárez (-2%) y Tlalnepantla de Baz (-0.1%), de la Zona Metropolitana del Valle de México. Por otra parte, es claro observar que la Miguel Hidalgo es la delegación con una mayor tasa de crecimiento promedio anual, equivalente a 11.7 por ciento. Asimismo, la menor dispersión es para Naucalpan de Juárez y la mayor para la Azcapotzalco con valores de 2.5 y 10.7%, respectivamente.

Gráfica 9



Gráfica 10



3.2.4. Total de la actividad económica municipal

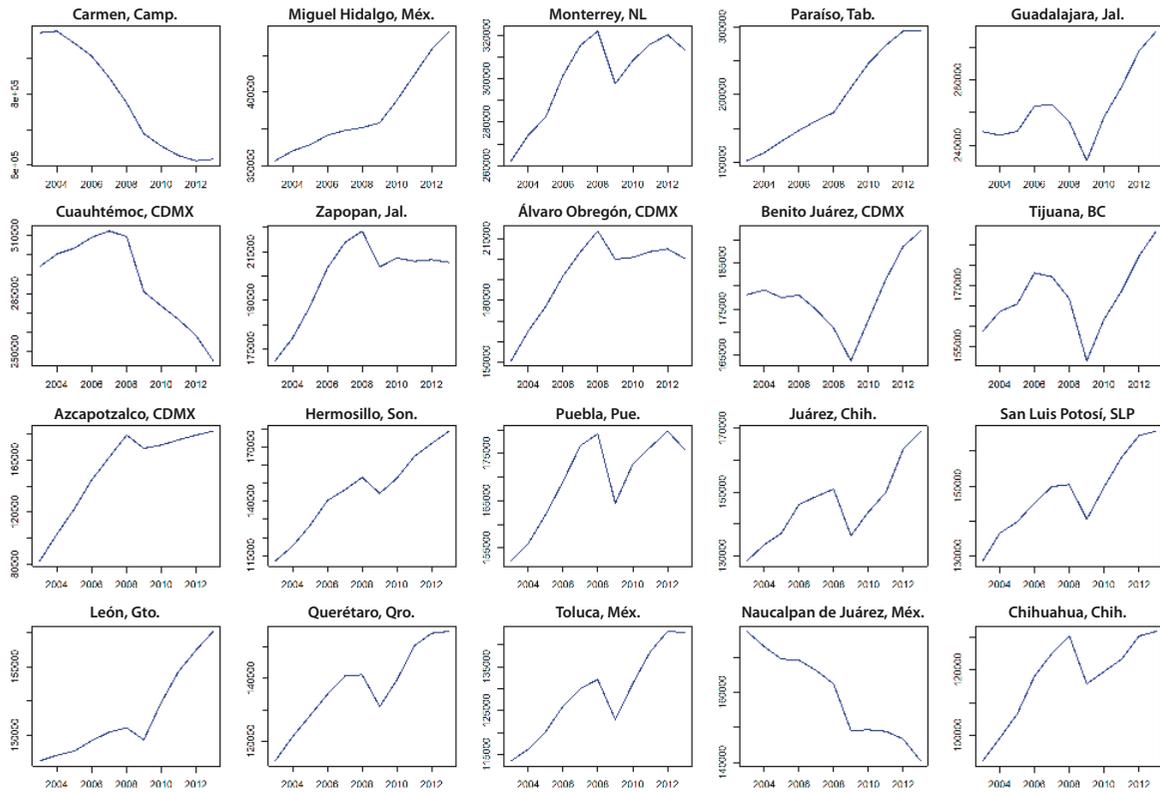
La gráfica 11 presenta los resultados expresados en niveles para el total de la actividad económica municipal, la cual contempla los sectores comentados en la sección 2, es decir, dada la cobertura sectorial considerada en este trabajo, las GA no corresponden a su respectivo total, por lo cual el total estatal o la suma de los municipios no corresponde al publicado por el SCN.

Para el total de la actividad económica, los primeros 20 municipios concentran 38.13 por ciento. Se puede apreciar que algunos destacan en el total por tener una composición económica balanceada al ocupar una posición preponderante tanto en la GA2 como en la GA3, por ejemplo: Monterrey, Guadalajara, Zapopan, Tijuana, Azcapotzalco y Hermosillo. Asimismo, dentro de este grupo, solo tres muestran una tendencia decreciente: Carmen,

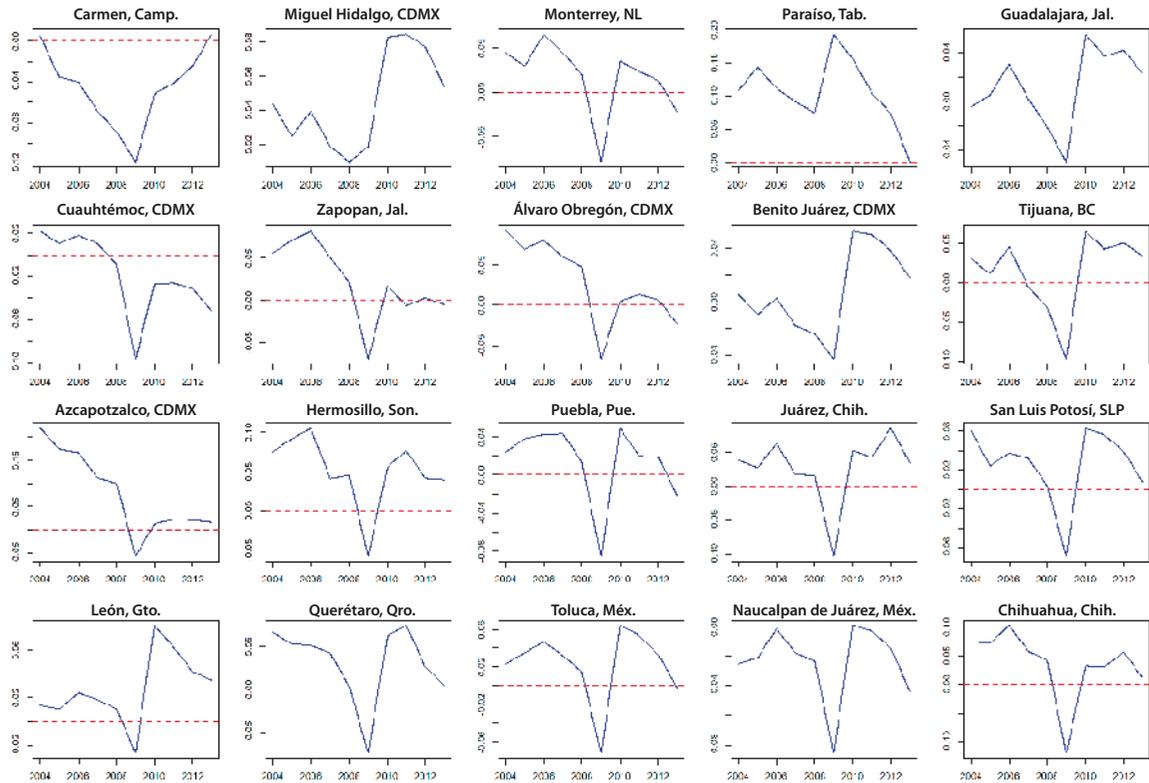
la Cuauhtémoc y Naucalpan de Juárez; mientras que el resto observa un crecimiento constante a pesar de la caída acontecida en el 2009. La gráfica 12 presenta los resultados para los cambios porcentuales de los IAE.

Se puede observar que, para algunos municipios, el efecto de la crisis del 2009 es más importante que para otros, con excepción de Paraíso y la Miguel Hidalgo, que no llegaron a presentar un crecimiento negativo por este evento. Los que presentan tasas de crecimiento promedio anuales negativas son Carmen (-4.6%), Naucalpan de Juárez (-2.3%) y la Cuauhtémoc (-1.8%), mientras que Paraíso es el que más ha crecido con una tasa de 10.6%; es importante comentar que las dispersiones rondan entre 2.6 y 8.9%, siendo la menor la de Naucalpan de Juárez y la mayor la de la Azcapotzalco.

Gráfica 11

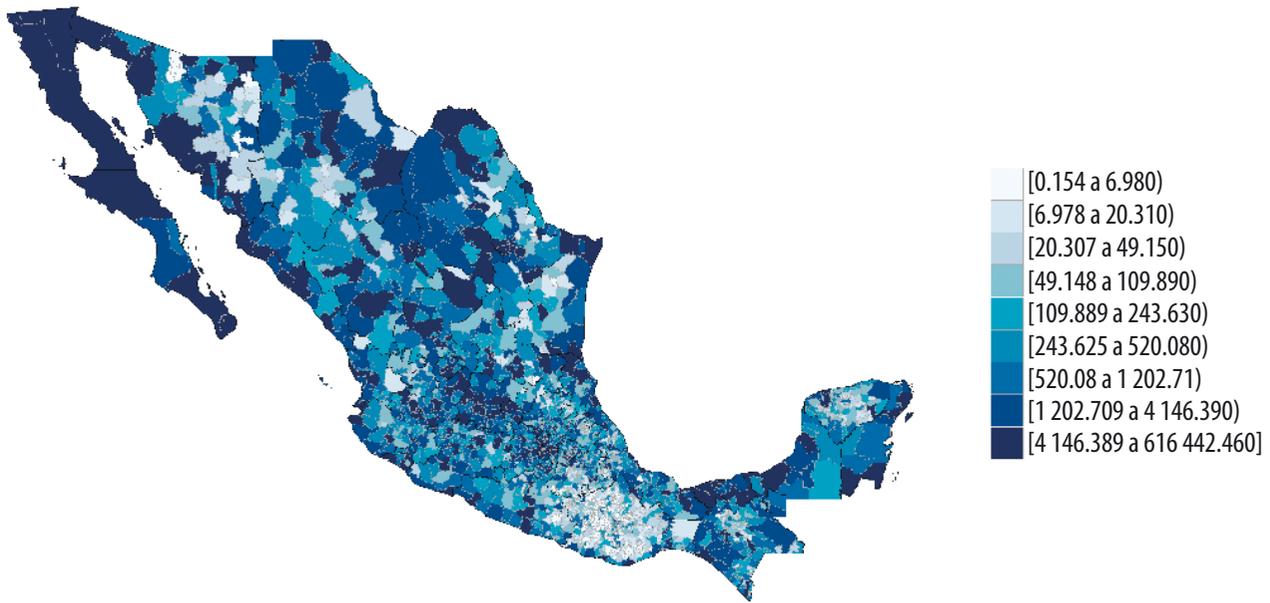


Gráfica 12



Mapa 1

Total de municipios de México



Por último, con el fin de tener una visión geográfica de la actividad económica en el país y motivar, con ello, el análisis regional, en el mapa de la República Mexicana se muestran las cifras de los IAE municipales para el total de la economía en el 2013.

Se puede apreciar la gran heterogeneidad que hay en el país donde, debido a la gran cantidad de municipios, es complicado establecer algún patrón sistemático en el comportamiento económico regional, aunque resalta la presencia del rezago económico en el sureste de México, en particular en lo que se refiere a los municipios de Guerrero, Oaxaca, Puebla y Yucatán.

4. Análisis de crecimiento económico de las entidades federativas a través de los IAE municipales

Una manera de evaluar el funcionamiento estructural de los indicadores estimados utilizando un concepto económico puede ser a través de un

análisis de convergencia económica. Para tener un primer acercamiento con la calidad de los índices desde esta perspectiva, se propone realizar la siguiente regresión lineal tipo panel de datos:

$$y_{i,t} = \beta_0 + \beta_1 g_{y_{i,t}} + \varepsilon_{it} \quad (11)$$

donde β_0 es el intercepto; β_1 , la contribución de $g_{y_{i,t}}$, el crecimiento del índice, sobre los niveles del índice, $y_{i,t}$ y, finalmente, ε_{it} es el coeficiente de error asociado a la estimación. En este caso, el coeficiente β_1 señala una especie de relación *beta-convergencia*, que indica que los municipios con mayor IAE están creciendo a una menor velocidad para el periodo analizado. Es importante señalar que la expresión (11) suele ser estimada en términos per cápita; no obstante, dado que no es objetivo esencial de este trabajo el desentrañar las características de los índices estimados sino, más bien, el proponer una metodología desde una perspectiva estadística, se considera como línea futura estimar dicha regresión expresada en términos per cápita. El cuadro 2 resume los resultados para las 32 entidades federativas.

Cuadro 2

**Resultados de la estimación de β_1
en la regresión $y_{i,t} = \beta_0 + \beta_1 g_{y_{i,t}} + \varepsilon_{it}$
obtenida a través del ajuste de un panel
de datos para cada entidad federativa**

Entidad	β_1	p-valor
Ags.	0.001	0.928
BC	-0.012	0.080
BCS	0.013	0.114
Camp.	-0.010	0.077
Coah. de Z.	0.033	0.000
Col.	-0.002	0.851
Chis.	-0.008	0.006
Chih.	0.003	0.609
CDMX	0.009	0.149
Dgo.	0.028	0.000
Gto.	0.006	0.108
Gro.	-0.015	0.000
Hgo.	-0.001	0.872
Jal.	0.007	0.049
Méx.	0.002	0.484
Mich. de O.	0.008	0.010
Mor.	0.002	0.606
Nay.	0.011	0.046
NL	0.016	0.008
Oax.	-0.009	0.001
Pue.	0.025	0.000
Qro.	0.018	0.004
Q. Roo	-0.027	0.185
SLP	0.008	0.251
Sin.	-0.002	0.725
Son.	0.005	0.268
Tab.	0.002	0.848
Tamps.	0.002	0.567
Tlax.	0.011	0.109
Ver. de I. de la LI.	0.002	0.412
Yuc.	-0.005	0.197
Zac.	0.016	0.005

Nota: en negritas se indica cuando la relación fue significativa a 10 por ciento.

Se puede apreciar que, a 10% de significancia, en Baja California, Campeche, Chiapas, Guerrero y Oaxaca, los municipios con mayor actividad económica están creciendo a una velocidad menor que los que tienen un menor IAE. Cabe anotar que, al estar presente este efecto en Chiapas, Guerrero y Oaxaca, posiblemente esté dando indicios de que los municipios menos pobres puedan estar emparejándose con los más pobres, mientras que, en Campeche, el decaimiento del Complejo Cantarell a partir del 2006, hizo que se presentara dicho efecto. En otras palabras, estaría dando indicios de un proceso de convergencia económica. Para el caso de Baja California, es interesante analizar a detalle la presencia de dicha relación, lo cual queda fuera del alcance de este trabajo.

Para finalizar, en Durango, Jalisco, Michoacán de Ocampo, Nayarit, Nuevo León, Puebla, Querétaro y Zacatecas parece haber un efecto de polarización económica entre los municipios que componen dichas entidades. Para el resto de los estados, la relación no es significativa a 10 por ciento.

5. Conclusiones

Con el fin de obtener una metodología que permita generar IAE municipales con una temporalidad y nivel sectorial pertinente, este trabajo representa la primera investigación formal que propone un enfoque estadístico para tener IAE dada la información existente de los CE, donde fue vital contar con el seguimiento del SCNM, de tal forma que la información utilizada y los indicadores generados sean representativos a nivel municipal. Lo anterior contribuye a llenar un hueco existente en la literatura que, a su vez, permite a quienes toman decisiones generar políticas económicas en bien de los municipios.

Por tal motivo, en este trabajo se generan IAE municipales del 2003 al 2013 utilizando el concepto de RC de Guerrero y Peña (2003) y usando como variables auxiliares indicadores construidos con información representativa y proveniente de los CE. Para verificar el error de estimación, se construyen

IAE estatales con la metodología aquí propuesta. Los resultados obtenidos se comparan con los PIB estatales generados por el SCNM. El error de estimación se calcula en dos formas: primero, suponiendo que la información de los CE es equivalente a la que se utiliza en el SCNM para la construcción del PIB, es decir, que el VACB es equivalente al PIB, para lo cual se obtiene un error de estimación que ronda en 2%; asimismo, considerando el error de estimación proveniente del uso solo de la información de los CE, que consiste sobre todo en incorporar a éste mayor cantidad de información del sector informal de la economía, el error estimado es cercano a 19 por ciento.

Se presentan los resultados para los 20 municipios y delegaciones más importantes de cada GA y el total de la actividad económica. Se puede apreciar que los IAE permiten identificar concentraciones importantes dentro de cada GA para el 2013, cercanas a 50, 40 y 30% para los 20 municipios y delegaciones más importantes dentro de las GA2, GA3 y GA1, respectivamente. Además, los IAE reflejan eventos relevantes de la economía, como la crisis económica y financiera del 2009, además de otros movimientos de interés para cada entidad.

Asimismo, se realiza una regresión de panel de datos para estimar, de forma preliminar, el proceso de convergencia económica, basados en la información municipal generada. Los resultados indican que en Baja California, Campeche, Chiapas, Guerrero y Oaxaca parece existir cierta evidencia de convergencia económica. Para el resto de los casos, o existe polarización económica o los resultados no permiten concluir estadísticamente.

Las cifras elaboradas cuentan con suficiente cobertura temporal y detalle geográfico que pueden ser utilizadas junto con mapas orográficos, climáticos, de vegetación y de asentamientos humanos para la elaboración de futuros análisis de economía espacial.

Como línea futura de trabajo se propone extender los resultados a 1988 y generar los mecanismos apropiados para incorporar información importan-

te del sector informal de la economía no capturada por los CE. También, se pretende realizar un análisis de convergencia económica expresando los IAE en términos per cápita y un estudio de desempeño económico similar a lo propuesto por Brida *et al.* (2011), pero aplicado a los municipios y demarcaciones territoriales de México.

Fuentes

- Battese, G. E.; R. M. Harter & W. A. Fuller. "An error-components model for prediction of county crop areas using survey and satellite data", en: *Journal of the American Statistical Association*. 83(401), 1988, pp. 28-36.
- Brida, J. G.; J. S. Pereyra; M. Puchet Anyul & W. A. Risso. "Regímenes de desempeño económico y dualismo estructural en la dinámica de las entidades federativas de México, 1970-2006", en: *Economía Mexicana*. Nueva época. 22(1), 2011, pp. 101-149.
- Casas-Cordero, C.; J. Encina & P. Lahiri. "Poverty mapping for the Chilean Comunas", en: Pratesi, M. (ed.). *Analysis of Poverty Data by Small Area Estimation: Methods for poverty mapping*. 2015, pp. 379-404.
- Chow G. y A. Lin. "Best linear unbiased interpolation, distribution, and extrapolation of time series by related series", en: *The Review of Economics and Statistics*. 53, 1971, pp. 372-375.
- Correa, L.; I. Molina y J. N. K. Rao. *Comparison of methods for estimation of poverty indicators in small areas*. Unpublished report. 2012.
- Dagum, E. B. y P. A. Cholette. "Benchmarking, Temporal Distribution, and Reconciliation Methods for Time Series", en: *Lecture Notes in Statistics* 186. 2006.
- Denton, F. T. "Adjustment of monthly or quarterly series to annual totals: An approach based on quadratic minimization", en: *Journal of the American Statistical Association*. 66, 1971, pp. 99-102.
- Di Fonzo, T. "The Estimation of M Disaggregate Time Series when Contemporaneous and Temporal Aggregates are Known", en: *The Review of Economics and Statistics*. 72(1), 1990, pp. 178-182.
- Drew, D.; M. P. Singhm y G. H. Choudhry. "Evaluation of small area estimation techniques for the Canadian Labour Force Survey", en: *Survey Methodology*. 8, 1982, pp. 17-47.
- Ebener, S.; C. Murray; A. Tandon & C. C. Elvidge. "From wealth to health: modelling the distribution of income per capita at the sub-national level using night-time light imagery", en: *International Journal of Health Geographics*. 4(1), 2005, p. 5.
- Elbers, C. M.; J. O. Lanjouw y P. Lanjouw. "Micro-level estimation of poverty and inequality", en: *Econometrica*. 71, 2003, pp. 355-364.
- Forsythe, G. E.; C. B. Moler & M. A. Malcolm. *Computer methods for mathematical computations*. 1977.

- Guerrero V. M. y F. Corona. "Retropolating some relevant series of Mexico's System of National Accounts at constant prices: The case of Mexico City's GDP", en: *Statistica Neerlandica*. 72(4), 2018a, pp. 495-519.
- Guerrero V. M. y F. Corona. "Retropolación hasta 1980 de algunas series del Sistema de Cuentas Nacionales de México", en: *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*. 2018b, pp. 98-119.
- Guerrero, V. M. y F. H. Nieto. "Temporal and contemporaneous disaggregation of multiple economic time series", en: *Test*. 8, 1999, pp. 459-489.
- Guerrero, V. M. & J. A. Mendoza. "On measuring economic growth from outer space: a single country approach", en: *Empirical Economics*. 2018, pp. 1-20.
- Guerrero, V. M. & D. Peña. "Combining multiple time series predictors: a useful inferential procedure", en: *Journal of Statistical Planning and Inference*. 116(1), 2003, pp. 249-276.
- Germán-Soto, V. "Generación del Producto Interno Bruto mexicano por entidad federativa, 1940-1992", en: *El Trimestre Económico*. Vol. 72, núm. 287, 2005, pp. 617-653.
- Henderson, J. V.; A. Storeygard & D. N. Weil. "Measuring economic growth from outer space", en: *American Economic Review*. 102(2), 2012, pp. 994-1028.
- Instituto Nacional de Estadística y Geografía (INEGI). *Censos Económicos 2004. Microdatos*. 2004.
- _____ *Censos Económicos 2009. Microdatos*. 2009.
- _____ *Censos Económicos 2014. Microdatos*. 2014.
- Llamas-Rosas, I.; E. R. González & M. S. Bustos. "Medición de la actividad económica en las principales zonas turísticas de playa en México a través de la luminosidad fotografiada desde el espacio", en: *Documentos de Investigación Banco de México*. 2018, p. 10.
- Marhuenda, Y.; I. Molina; D. Morales y J. N. K. Rao. "Poverty mapping in small areas under two-fold nested error regression model", en: *Journal of the Royal Statistical Society. Series A*, en prensa, 2018.
- Molina, I. "Desagregación de datos en encuestas de hogares: metodologías de estimación en áreas pequeñas", en: *Serie de la CEPAL*. Noviembre del 2018.
- Molina, I. y D. Morales. "Small area estimation of poverty indicators", en: *Boletín de Estadística e Investigación Operativa*. 25, 2009, pp. 218-225.
- Molina, I. y J. N. K. Rao. "Small area estimation of poverty indicators", en: *Canadian Journal of Statistics*. 38, 2010, pp. 369-385.
- Molina, I.; B. Nandram y J. N. K. Rao. "Small area estimation of general parameters with application to poverty indicators: a hierarchical Bayes approach", en: *Annals of Applied Statistics*. 8, 2014, pp. 852-885.
- Nieto, F. H. "Ex-post and Ex-ante Prediction of Unobserved Economic Time Series: A Case Study", en: *Journal of Forecasting*. 17, 1998, pp. 35-58.
- Stineman, R. W. "A Consistently Well Behaved Method of Interpolation", en: *Creative Computing*. 6 (7), 1980, pp. 54-57.
- Torabi, M. y J. N. K. Rao. "On small area estimation under a sub-area level model", en: *Journal of Multivariate Analysis*. 127, 2014, pp. 36-55.
- R Core Team. *R: A language and environment for statistical computing*. Vienna, Austria, R Foundation for Statistical Computing, 2017 (DE) <https://www.R-project.org/>
- Secretaría de Agricultura, Ganadería, Desarrollo Rural, Pesca y Alimentación (SAGARPA). *Sistema de Información Agroalimentaria y Pesquera (SIAP)*. 2018.
- Suárez Campos, M. A.; G. Aguilar y R. Mejía-González. "Estimación del ingreso por trabajo en los municipios y las delegaciones de México utilizando técnicas de estimación para áreas pequeñas", en: *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía*. 6(3), 2015, pp. 44-61.

Vinculación longitudinal *de los Censos Económicos 1994-2014 de México*

The Longitudinal Linkage *of Mexico's Economic Census 1994-2014*

Matías Busso,* Óscar Eduardo Fentanes Téllez** y Santiago Levy Algazi***

* Banco Interamericano de Desarrollo (BID), RES, MBUSSO@iadb.org

** BID, VPS, o.fentanes.t@gmail.com

*** BID, VPS, santiagolevy4@gmail.com

Nota: agradecemos al Instituto Nacional de Estadística y Geografía (INEGI) por el acceso a los datos utilizados para este enlace; el enlace presentado en este documento es un trabajo llevado a cabo por personal del Banco Interamericano de Desarrollo y no se considera parte de los registros oficiales del INEGI; los identificadores creados y los datos mencionados pueden ser consultados en el Laboratorio de Microdatos bajo aprobación del INEGI; se agradece, también, a Jesica Torres Coronado y Natalia Volkow por sus valiosos comentarios y sugerencias; las opiniones expresadas en esta publicación son de los autores y no necesariamente reflejan el punto de vista del BID, de su Directorio Ejecutivo ni de los países que representa.

Esta nota técnica describe la metodología para construir una base de datos longitudinal a partir de los Censos Económicos de 1994 hasta 2014. El proceso se basa en un algoritmo que enlaza establecimientos con idéntica o similar ubicación, entidad legal y clase de actividad. Puesto que ya existe un conjunto de identificadores longitudinales para los de 2009 y 2014, estos son utilizados para validar nuestros resultados, obteniendo 90% de precisión. Enlazamos 0.92 millones de establecimientos para los Censos 1994-1999, 1.44 millones para 1999-2004, 1.52 millones para 2004-2009 y 2.15 millones para 2009-2014.

Palabras clave: datos longitudinales; Censos Económicos; INEGI.

Códigos JEL: C81, D21

This technical note describes the methodology to construct a longitudinal dataset using the Economic Censuses of Mexico from 1994 to 2014. The procedure is based on an algorithm that links establishments with identical or significantly similar location, legal entity and kind of activity. Since a set of longitudinal identifiers is already available for the 2009 and 2014 Economic Censuses, it is used to validate our results, obtaining 90% of accuracy. This paper links 0.92 million establishments for the period 1994-1999, 1.44 million for 1999-2004, 1.52 million for 2004-2009, and 2.15 million for 2009-2014.

Key words: Longitudinal Database; Economic Census; INEGI.

JEL Classification: C81, D21



1. Introducción

El Instituto Nacional de Estadística y Geografía (INEGI) lleva a cabo en México los Censos Económicos (CE) quinquenalmente desde 1989. Los CE recolectan información de todos los negocios operando en instalaciones fijas y ubicados en localidades urbanas de más de 2 500 habitantes.

Los Censos 2009 y 2014 introdujeron el identificador Clave Única de Identificación Estadística (CLEE),¹ el cual enlaza longitudinalmente establecimientos de ambos levantamientos censales y los subsecuentes. A pesar de que la CLEE ya puede ser utilizada para estudios longitudinales, no está disponible para ediciones anteriores, limitando el potencial de estas bases de datos.

En esta nota técnica, describimos el proceso de vinculación de los Censos Económicos 1994 hasta 2014. A pesar de que en principio se podría incluir también los de 1989, existen dificultades para armonizar su codificación industrial y geográfica con otras ediciones.

El trabajo de vinculación de los CE 1999 hasta 2014 ya fue descrito en la nota técnica de Busso, Fentanes y Levy (2018). La presente investigación la extiende a los Censos 1994 e incluye en el *Apéndice B* el funcionamiento de un par de comandos para STATA que permiten corregir inflexiones en cadenas de texto para facilitar los trabajos de vinculación con otros conjuntos de registros.

El resto del documento se estructura como sigue: en la sección 2 describimos los datos; la 3 se enfoca en el algoritmo de vinculación de los Censos; en la 4 presentamos los resultados; en la 5, realizamos algunos ejercicios de validación, incluyendo medidas de flujos de trabajadores y entrada y salida de establecimientos; finalmente, en la sección 6, discutimos algunas consideraciones sobre nuestro procedimiento y en la 7 explicamos cómo solicitar el acceso a los identificadores aquí descritos a través del Laboratorio de Microdatos del INEGI.

¹ Fue creada mediante una combinación de procedimientos humanos y computacionales de identificación.

2. Los Censos Económicos de México

2.1 Cobertura

Nuestra fuente de datos son los Censos Económicos. La cobertura temporal es 1994, 1999, 2004, 2009 y 2014.² Utilizamos todas las clases de actividad (cerca de 800 por año) y todas las localidades urbanas disponibles. Ya que la unidad económica de los CE es el establecimiento, la vinculación que aquí presentamos también es a este nivel. El número de establecimientos por periodo se muestra en el cuadro 1.

La cantidad de unidades económicas de los CE se incrementa, principalmente, por el nacimiento de nuevos establecimientos, pero también por ampliaciones en la cobertura, sobre todo, la adición de nuevas localidades reclasificadas como urbanas.

Cuadro 1

Establecimientos	
Censos	Total
1994	2 209 443
1999	2 804 984
2004	3 005 157
2009	3 724 019
2014	4 230 745

2.2 Variables

Para todos los CE, contamos con información detallada que nos permite identificar establecimientos; por ejemplo, entidad legal, nombre de la unidad, claves de localización, año de inicio de actividades, clase de actividad, entre otra. La lista completa se muestra en el cuadro 2.

Los códigos de ubicación de E03 a E07 son codificaciones estandarizadas definidas por el INEGI; las variables E10, E11 y E14 son cadenas de texto capturadas manualmente; E08 es el nombre del establecimiento, por ejemplo, *Minimercado María*;

² Cada edición captura información del año previo, por ejemplo, la edición 2009 contiene información del 2008.

Variables disponibles

Ubicación	Entidad legal	Industria
E03 estado	E01 NIC	E17 clase de actividad
E04 municipio	E02 NOP	
E05 localidad	E08 nombre del establecimiento	
E06 AGEBA ^a	E09 razón social	
E07 manzana	G111 año de inicio de actividades	
E10 calle		
E11 número		
E14 colonia		

^a AGEB significa Área Geoestadística Básica.

E09 es la entidad legal, por ejemplo, *María SA de CV*; si la unidad no pertenece a una entidad legal, E09 reporta el nombre del(la) propietario(a); E17 es la clase de actividad de acuerdo con el Sistema de Clasificación Industrial de América del Norte (SCIAN) para los Censos 1999 hasta 2014. Para los de 1994 y 1999, las clases de actividad se agrupan de acuerdo con la Clasificación Mexicana de Actividades y Productos (CMAP).³

Las variables E01 (NIC)⁴ y E02 (NOP)⁵ son identificadores disponibles para todos los Censos. Estos pueden ser utilizados para vinculación longitudinal para todos los CE, pero solo para un número limitado de establecimientos, generalmente grandes. Para la mayor parte de las unidades económicas, la combinación de códigos NIC-NOP solo puede ser utilizado como un identificador dentro de un levantamiento, pero no longitudinalmente.

3. Vinculación

Su proceso se puede resumir en cinco pasos, como el modelo presentado en Christen (2012) —el enlace se lleva a cabo del I al IV; el V, validación, será discutido en la sección 5—, que son:

³ Los Censos Económicos 1999 incluyen ambas clasificaciones industriales.

⁴ Número de Identificación Censal.

⁵ Número Operativo.

- I. Estandarización: se sustituyen o eliminan caracteres especiales, como acentos y signos de puntuación; también, se armonizan descripciones de los establecimientos, como *Abarrotes* o *Tienda*, que representan el mismo tipo de negocio; asimismo, se corrigen errores en el tipo de entidad legal, como *SA d CV* en lugar de *SA de CV*; en general, se eliminan, estandarizan o sustituyen caracteres en todas las variables capturadas manualmente (no claves estandarizadas), que son propensas a errores de captura.
- II. Indexación (previnculación): se proponen candidatos para vinculación; por ejemplo, si dos establecimientos tienen la misma ubicación en t y $t + 5$, se compara el nombre del(la) propietario(a) o del establecimiento para decidir si es un buen enlace.
- III. Comparación: se usan diferentes estrategias para comparar pares de establecimientos indexados; en general, se utilizan procedimientos de STATA para comparar cadenas de texto.
- IV. Clasificación de enlaces: se asigna un identificador único a los establecimientos enlazados; después, se etiquetan con el fin de ser excluidos en futuras fases del algoritmo —discutido más adelante—; además, se asigna el número de la fase en la que fue enlazado un establecimiento.
- V. Validación: se mide la precisión del algoritmo aplicándolo a los CE 2009 y 2014,

los cuales ya fueron enlazados y validados por el INEGI (ver sección 5).

Para llevar a cabo los pasos I-IV, se define un algoritmo de 10 fases. Todas se basan en las reglas de continuidad definidas por la Organización para la Cooperación y el Desarrollo Económicos (OCDE, 2008). Estas consideran tres factores de continuidad: ubicación, entidad legal y clase de actividad; si alguna unidad económica mantiene al menos dos de tres de un periodo a otro, se considera la misma.

Para ejecutar las 10 fases, se emplea principalmente STATA. En algunas utilizamos el comando *matchit*, escrito por Raffo (2017), el cual compara cadenas de texto y asigna un coeficiente de similitud entre 0 y 1. También, usamos el comando *soundex*, el cual consiste en la primera letra de la cadena de texto seguido de tres dígitos asignados por STATA; estos son los mismos para similares cadenas de consonantes.

Las fases son:

1. Se enlazan los establecimientos con combinación idéntica de NIC y NOP.⁶
2. Se enlazan establecimientos con la misma combinación de estado, municipio, localidad, AGEB, manzana y clase de actividad.
3. Se indexan unidades económicas con la misma combinación de estado, municipio, localidad, AGEB, manzana y número exterior. Luego, se enlazan si tienen un coeficiente de similitud de, al menos, 45% en el nombre del establecimiento y 75% en la entidad legal.⁷
4. Se indexan establecimientos con la misma combinación de estado, municipio, clase de actividad y entidad legal. Luego, se enlazan si tienen un coeficiente de similitud de, al menos, 30% en el nombre de la unidad.
5. Se enlazan establecimientos con la misma combinación de estado, municipio, AGEB y entidad legal.

6. Se indexan los establecimientos con la misma combinación de estado, municipio, localidad, AGEB, manzana y clase de actividad. Luego, se enlazan si tienen el mismo *soundex* en el nombre de la unidad económica y entidad legal.
7. Se enlazan los establecimientos con la misma combinación de estado, municipio, localidad, AGEB, manzana, clase de actividad y año de inicio de actividades.
8. Se enlazan los establecimientos con la misma combinación de estado, municipio, localidad, AGEB, manzana, clase de actividad y número exterior.
9. Se indexan las unidades económicas con la misma combinación de estado, municipio, localidad y AGEB o clase de actividad. Luego, se enlazan si tienen un coeficiente de similitud de, al menos, 65% en el nombre del establecimiento y en entidad legal.
10. Se enlazan los establecimientos con la misma combinación de clase de actividad, nombre del establecimiento y entidad legal.

Siempre que se enlazan unidades económicas de acuerdo con una secuencia de variables se consideran solo aquellas que presenten una combinación única dentro del levantamiento censal; por ejemplo, en la fase 2 enlazamos las que tienen la misma ubicación y clase de actividad en t y $t + 5$; sin embargo, si dos reportan la misma ubicación y clase de actividad en t , no será claro cuál de los dos es el que reapareció en $t + 5$. Para evitar ambigüedades y minimizar errores de vinculación, excluimos estos casos y se intenta en fases futuras enlazar los establecimientos con diferentes combinaciones de variables.

Los valores de los coeficientes de similitud que se requieren en algunas fases fueron determinados de tal forma que logren predecir correctamente al menos 90% de los enlaces (en los CE 2009 y 2014); se puede ser más restrictivo con estos, pero las ganancias en precisión no necesariamente compensan las pérdidas de buenos enlaces por no cumplir con los nuevos criterios. En la sección 5 se detalla la precisión de cada fase.

⁶ Algunos NIC-NOP tiene duplicados. En los CE 1999 eran menos de 400; en los Censos 2004 y 2009, menos de 100; y ninguno en la edición 2014.

⁷ Si el nombre del establecimiento o la entidad legal aparece vacía o reporta SIN NOMBRE, no se considera para el enlace.

4. Resultados

Tras llevar a cabo las 10 fases del algoritmo, se obtienen los resultados mostrados en el cuadro 3. Para cualquier par de CE adyacentes, t y $t + 5$, enlazamos al menos 50% de los establecimientos en t .

El cuadro 4 desglosa los enlaces totales por fase.⁸ Las fases 1 a la 6 son, por mucho, las más importantes, representando al menos 86% de los enlaces para cualquier periodo (y como veremos más adelante, también son las más precisas).

El algoritmo de enlace solo se aplica a pares de CE consecutivos; sin embargo, podemos seguir algunos establecimientos a través de varias ediciones. De acuerdo con los cuadros 3 y 5, de los 2.2 millones de establecimientos registrados en los CE 1994, sobrevivieron 991 mil en los de 1999; de estos, 637 mil reaparecieron en los Censos

2004; para los de 2009, de estos quedaron 433 mil; y, finalmente, en la edición 2014 se capturaron de nuevo 333 mil de ellos. En otras palabras, podemos formar un panel balanceado de 333 mil establecimientos de los levantamientos censales 1994 hasta 2014. Por otra parte, también es posible integrar otro de 1999 hasta 2014 de 675 mil (ver cuadro 4; otras posibilidades se pueden observar en este mismo).

Cuadro 5
Posibles paneles balanceados

Periodos	CE	Establecimientos
3	1994-1999-2004	637 465
	1999-2004-2009	907 992
	2004-2009-2014	1 081 257
4	1994-1999-2004-2009	433 033
	1999-2004-2009-2014	675 379
5	1994-1999-2004-2009-2014	333 041

Cuadro 3

Establecimientos enlazados

t a $t + 5$	Establecimientos en t	Enlaces	% t
1994-1999	2 209 443	991 230	44.9
1999-2004	2 804 984	1 444 584	51.5
2004-2009	3 005 157	1 522 578	50.7
2009-2014	3 724 019	2 154 410	57.9

⁸ La fase 1 no fue aplicada para 2009-2014 porque las variables NIC y NOP son redundantes con la CLEE. Mientras que para periodos anteriores, la fase 1 logra alrededor de 7% de los enlaces, para 2009-2014 representaría 100 por ciento.

Cuadro 4

Porcentaje de enlaces por fase

Periodo	Enlaces	Fase										Total
		1	2	3	4	5	6	7	8	9	10	
1994-1999	991 230	7.7	33.8	25.8	0.8	27.3	3.1	0.1	1.0	0.4	0.0	100.0
1999-2004	1 444 584	7.1	38.2	17.2	15.1	6.5	2.1	0.5	12.5	0.8	0.1	100.0
2004-2009	1 522 578	7.1	40.0	16.9	11.2	8.9	3.0	1.6	10.4	0.8	0.1	100.0
2009-2014	2 154 410	0.0	49.2	14.1	16.8	9.4	2.7	2.8	4.3	0.5	0.1	100.0

5. Validación

La calidad del enlace depende de su cobertura y precisión; ambas características pueden ser evaluadas respondiendo a las siguientes preguntas:

- (i) Cobertura: ¿cuántos establecimientos deben ser enlazados para cada par de CE consecutivos?
- (ii) Precisión: ¿cuál es la probabilidad de que dos establecimientos vinculados sean efectivamente el mismo?

Para responder ambas, comparamos el enlace realizado por el INEGI a través de la CLEE y el que se logra con nuestro algoritmo.

5.1 Cobertura del enlace

El cuadro 6 muestra que 58% de los establecimientos de los CE 2009 pueden ser enlazados con unidades económicas de los Censos 2014 utilizando la CLEE. Por lo tanto, esperaríamos que el algoritmo enlazara un porcentaje similar; así, éste enlaza 2 154 410, es decir, 57.9% de los establecimientos de los CE 2009. El número de enlaces de ambos métodos es virtualmente el mismo, el algoritmo logra 99.7% del monto total que enlaza la CLEE.

Cuadro 6
Enlaces por método, 2009-2014

Método	Establecimientos en 2009	Enlaces	%
CLEE	3 724 019	2 159 804	58.0
Algoritmo	3 724 019	2 154 410	57.9

El algoritmo alcanza el número esperado de enlaces de los CE 2009-2014; sin embargo, esto no responde por completo a la pregunta (i), necesitamos estimar también cuántos establecimientos debe enlazar este para 1994-1999, 1999-2004 y 2004-2009.

Una forma de responder esta pregunta es utilizando el año de inicio de actividades declarado por informantes. Si un establecimiento reporta una edad mayor o igual a 5 en $t + 5$, podría potencialmente ser observado en t . De acuerdo con el cuadro 7, en los Censos 2009, cerca de 1.8 millones de los 3.7 millones de establecimientos reportaron operaciones en el levantamiento censal 2004 o antes según su edad. Como en los CE 2004 capturaron 3 millones, esperaríamos enlazar 59.7% de ellos con unidades de los de la edición 2009. Aplicando el mismo razonamiento, esperaríamos enlazar 51.5% de los establecimientos de 1999 con unidades de 2004 y 53.3% de 1994 con los de 1999.

Cuadro 7

Establecimientos por edad

Edad	1999	2004	2009	2014
Menos de 5	1 627 248	1 561 466	1 928 674	1 806 638
5 o más	1 177 736	1 443 691	1 795 345	2 424 107
Total	2 804 984	3 005 157	3 724 019	4 230 745

Sin embargo, hay dos razones por las que el número de enlaces esperados podría estar sobrestimado si nos basamos solo en el año de inicio de actividades reportado. La primera es que los Censos Económicos expanden su cobertura geográfica en cada edición dado que algunas localidades crecen y comienzan a ser consideradas como urbanas;⁹ la segunda, dado que el año de inicio de actividades es reportado por informantes, este puede ser impreciso. Mediante el identificador CLEE, podemos medir la discrepancia entre los que sobreviven de acuerdo con su edad y los que en efecto son enlazados de los CE 2009 hasta 2014.

En los Censos Económicos 2014 se capturaron 4.2 millones de establecimientos, de los cuales, el INEGI enlazó 2.2 millones mediante la CLEE. Al mismo tiempo, en la edición 2014, 2.4 millones reportaron actividades en los CE 2009 según su edad; en otras palabras, el INEGI enlazó 89.1% de los establecimientos que declararon operaciones en los levantamientos 2009 y 2014. Si asumimos que este grado de discrepancia (por cobertura o error de reporte) se mantiene constante en el tiempo, esperaríamos enlazar solo 89.1% de los que reportan actividades en cualesquiera dos CE según su edad.

En el cuadro 8 se observa que 1.1 millones de establecimientos manifestaron actividades en los Censos 1994 y 1999 de acuerdo con su edad. De estos, esperamos enlazar 89.1%, es decir, poco más de 1 millón. Finalmente, se enlazan 0.99 millones, es decir, una cobertura de 94.5 por ciento. Bajo este mismo razonamiento, se logra una co-

⁹ En 2009, el levantamiento censal se hizo en 2 mil localidades urbanas. Para 2014, se realizó en 3 600; sin embargo, las 1 600 nuevas solo aportaron 12 mil establecimientos adicionales.

bertura de 112.3% para las ediciones 1999-2004, de 95.2% para 2004-2009 y de 99.7% para 2009-2014. Una superior a 100% implica una posible sobrestimación de la supervivencia de las unidades económicas y una inferior a 100%, una sobrestimación de su mortalidad.

5.2 Precisión del enlace

La pregunta (ii) también puede ser respondida tomando como referencia el enlace 2009-2014 hecho por el INEGI mediante la CLEE. Como se mencionó con anterioridad, el algoritmo y la CLEE enlazan de manera virtual el mismo número de establecimientos; sin embargo, esto no significa que ambos métodos enlacen exactamente los mismos. De hecho, se pueden obtener los siguientes cuatro tipos de enlace (o no enlace):

1. Verdadero positivo: enlazado con el algoritmo y la CLEE.
2. Falso positivo: enlazado con el algoritmo, pero no con la CLEE.
3. Verdadero negativo: no enlazado con el algoritmo ni la CLEE.
4. Falso negativo: no enlazado con el algoritmo, pero sí con la CLEE.

El cuadro 9 presenta los porcentajes de establecimientos de los CE 2009 y 2014 según las cuatro posibilidades de enlace descritas. De los enlazados con el algoritmo en el levantamiento 2009, 90% también fue enlazado por la CLEE (89.8% respecto a 2014). En general, el porcentaje de verdaderos positivos puede interpretarse como la precisión

Cuadro 8

Enlaces esperados y realizados

Periodo	Edad ≥ 5	Enlaces esperados	Enlaces	%
1994-1999	1 177 736	1 049 363	991 230	94.5
1999-2004	1 443 691	1 286 329	1 444 584	112.3
2004-2009	1 795 345	1 599 652	1 522 578	95.2
2009-2014	2 424 107	2 159 879	2 154 410	99.7

Cuadro 9

Tipos de enlace por año

Año	No enlazados			Enlazados		
	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
2009	86.0	14.0	100.0	10.0	90.0	100.0
2014	89.1	10.9	100.0	10.2	89.8	100.0

del algoritmo, o bien, como la probabilidad de que dos establecimientos enlazados por el algoritmo sean efectivamente el mismo.

Por otra parte, el porcentaje de falsos positivos fue de 10% para los CE 2009 y 10.2% para la edición 2014. Este es el precio que se paga por obtener un número alto de enlaces. Una forma de disminuirlo es incrementando las restricciones en algunas de las 10 fases del algoritmo; por ejemplo, requiriendo porcentajes de similitud superiores en el nombre del establecimiento o entidad legal. La desventaja al hacer esto es que se incrementará el porcentaje de falsos negativos, dado que algunos establecimientos que antes eran correctamente enlazados ya no cumplirán con los nuevos criterios.

La otra cara de la moneda cuando hablamos de la precisión del algoritmo son los verdaderos y falsos negativos, es decir, los que no fueron enlazados. Dejar fuera a establecimientos que debieron ser enlazados implica una sobrestimación de la salida de unidades económicas del mercado. Según el cuadro 9, 14% de los establecimientos no enlazados en los Censos 2009 sí reaparecieron en el levantamiento 2014. En términos absolutos, no fueron enlazados 220 mil que efectivamente sobrevivieron, pero al mismo tiempo se tiene un número similar de falsos positivos, por lo que la cantidad de enlaces totales se mantiene similar entre la CLEE y el algoritmo (ver cuadro A1 en el Apéndice).

Los porcentajes mostrados en el cuadro 9 son agregados, incluyendo establecimientos de todos los tamaños, clases de actividad y estados de México; además, consideran todas las fases del algoritmo y no todas son igualmente precisas; podemos desagregar los porcentajes para saber si

existe una diferencia sistemática en la precisión del algoritmo dado el tamaño, clase de actividad, estado o fase.

5.3 Precisión por tamaño

Los cuadros 10 y 11 reportan que la precisión del algoritmo (porcentaje de verdaderos positivos) incrementa con el tamaño. Esto significa que si el algoritmo predice que dos establecimientos grandes son el mismo es prácticamente cierto que lo son. El riesgo de falsos positivos es mayor para los pequeños. Por otra parte, el porcentaje de falsos negativos es también mayor para los grandes, es decir, es más difícil enlazarlos. La mortalidad de establecimientos grandes podría estar sobrestimada. Note que los porcentajes agregados se asemejan a los de los establecimientos de menos de cinco trabajadores, que forman 95% del total.

5.4 Precisión por sector de actividad

Los cuadros A2 y A4 en el *Apéndice* muestran los porcentajes de tipos de enlace desglosados por sectores a dos dígitos del SCIAN. En 2009, la clase con el menor porcentaje de verdaderos positivos es 55 Servicios Administrativos; sin embargo, está formada por solo 204 establecimientos, por lo que no tiene un gran impacto en la precisión agregada. El sector manufacturero, comúnmente utilizado en la literatura, es muy preciso, con menos de 10% de falsos positivos. Los sectores 11, 21, 22, 23 y 55 muestran niveles altos de falsos negativos, lo que implica una sobrestimación en la mortalidad de dichas unidades; sin embargo, representan solo 1.2% del levantamiento censal, por lo que tienen poco impacto en la precisión agregada (ver cuadros A3 y A5 en el *Apéndice*). El resto de los sectores presentan poca variación en los porcentajes respecto al agregado.

5.5 Precisión por estado

Los cuadros A6 y A7 en el *Apéndice* muestran los porcentajes de tipos de enlace desglosados por estado.

El de verdaderos positivos no presenta gran dispersión por entidad, manteniéndose en un rango que va de 87 a 92%; esto significa que la precisión del algoritmo es similar para todas las regiones del país. Por otra parte, los falsos negativos presentan mayor dispersión. En particular, el algoritmo sobrestima la mortalidad de unidades en la Ciudad de México (CDMX). Esto podría ocurrir porque es una zona densamente poblada y establecimientos muy similares se concentran en espacios pequeños, creando ambigüedades difíciles de resolver.

Cuadro 10

Tipo de enlace por tamaño, 2009

Trabajadores	No enlazados			Enlazados		
	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
[0-10]	86.3	13.7	100.0	10.1	89.9	100.0
[11-50]	79.6	20.4	100.0	7.6	92.4	100.0
[50-100]	69.4	30.6	100.0	5.3	94.7	100.0
> 100	63.3	36.7	100.0	5.4	94.6	100.0
Total	86.0	14.0	100.0	10.0	90.0	100.0

Cuadro 11

Tipo de enlace por tamaño, 2014

Trabajadores	No enlazados			Enlazados		
	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
[0-10]	89.4	10.6	100.0	10.4	89.6	100.0
[11-50]	83.8	16.2	100.0	8.2	91.8	100.0
[50-100]	74.8	25.2	100.0	5.8	94.2	100.0
> 100	69.0	31.0	100.0	4.9	95.1	100.0
Total	89.1	10.9	100.0	10.2	89.8	100.0

5.6 Precisión por fase

Como se anticipó, no todas las fases del algoritmo tienen la misma precisión (porcentaje de verdaderos

positivos). Como se muestra en el cuadro 12, la precisión es superior a 90% de las fases 2 a 6 (para la 1 será de 100% porque la combinación NIC-NOP es redundante con la CLEE). Las últimas cuatro fases tienen niveles inferiores de verdaderos positivos; sin embargo, forman solo 7.8% de los enlaces y su impacto en los porcentajes acumulados es limitado. Dichas fases se incluyen porque completan la cobertura de enlaces y mantienen los criterios de continuidad de la OCDE.

Cabe mencionar que 10% de falsos positivos no es una asignación aleatoria de enlaces; dicho de otra forma, no se enlazará un establecimiento de Walmart con una pequeña ferretería. Si bien el algoritmo enlaza establecimientos que la CLEE no, estos aun así mantienen gran similitud de ubicación, razón social y clase de actividad. Aunque algunos enlaces no correspondan al mismo establecimiento en la realidad, esto puede no ser problemático para efectos estadísticos ya que se trata de unidades muy similares. Desafortunadamente, no se puede saber para periodos anteriores al levantamiento 2009 cuáles son falsos positivos (al menos no por procedimientos computacionales), solo es posible conocer que podrían ser alrededor de 10% de los enlaces como en 2009-2014.

Cuadro 12

Tipo de match por fase

Fase	F. pos.	V. pos.	Total	% V. pos.
1	0	0	0	-
2	85 527	973 927	1 059 454	91.9
3	17 673	287 094	304 767	94.2
4	29 680	331 617	361 297	91.8
5	16 996	186 513	203 509	91.6
6	5 290	51 866	57 156	90.7
7	6 358	54 913	61 271	89.6
8	45 834	47 076	92 910	50.7
9	6 819	4 376	11 195	39.1
10	777	2 074	2 851	72.7
Total	214 954	1 939 456	2 154 410	90.0

6. Medidas de flujo de trabajadores y establecimientos

Otra forma de evaluar la calidad del algoritmo es estimando con este y la CLEE tanto medidas de flujo de trabajadores como entrada y salida de establecimientos. Para estimar estas medidas (anualizadas), se sigue el método de Miranda y Jarmin (2002), y se definen como sigue.

6.1 Índices de creación y destrucción de empleo

$$JCR = \frac{JC}{X}$$

donde:

$$JC = E_{t+5} - E_t$$

donde E denota el empleo en establecimientos que se expanden y recién nacidos y X es el empleo promedio de t y $t + 5$. El índice de destrucción de empleo (JDR) se calcula análogamente, pero E es el empleo de los establecimientos que se contraen y los que salen del mercado.

6.2 Índice de entrada y salida de establecimientos

$$\text{Índice de entrada} = \frac{ENTRY}{AVG}$$

donde $ENTRY$ es el número de establecimientos entrantes en $t + 5$ y AVG , el número de unidades económicas promedio entre t y $t + 5$. El índice de salida es similar, pero se reemplaza el número de entrantes por el de salientes ($EXIT$):

$$\text{Índice de salida} = \frac{EXIT}{AVG}$$

Las últimas dos filas del cuadro 13 muestran que los índices de entrada y salida tienen, prácticamente, los mismos valores para los dos métodos de enlace; sin embargo, los cálculos a partir de los enlaces del algoritmo sobrestiman ligeramente las ta-

sas de creación (*JC*) y destrucción de empleo (*JD*). Vale la pena notar que para periodos de la misma duración, por ejemplo, 1999-2004 y 2004-2009, o bien, 1999-2009 y 2004-2014, los índices mantienen el orden de magnitud.

Cuadro 13

Índices de entrada, salida, creación y destrucción de empleo

Enlace	Periodo	Entrada	Salida	C. de emp.	D. de emp.
A l i n t e r n o n a l	1999-2004	9.2	15.4	9.3	13.6
	1999-2009	6.6	9.1	6.4	8.4
	1999-2014	4.8	6.6	4.7	6.4
	2004-2009	11.6	12.7	11.3	11.8
	2004-2014	6.9	8.2	6.6	9.8
	2009-2014	9.4	11.2	11.0	13.5
CLEE	2009-2014	9.4	11.1	9.8	12.7

7. Discusión

La primera advertencia sobre este proceso de vinculación es que no toma en cuenta la reorganización de las unidades económicas como fusiones o particiones. Si ocurre un gran número de fusiones de t a $t + 5$, podríamos estar sobrestimando la salida de establecimientos. Por el contrario, si ocurrieron muchas particiones, se sobrestima el nacimiento de unidades. Desafortunadamente, se tiene poca información acerca de estos fenómenos para ser tomados en cuenta en el procedimiento de enlace.

Además, no se realizan ejercicios para tratar de enlazar establecimientos entre CE no consecutivos. Si un establecimiento se reportó en el levantamiento 1999 y reapareció hasta el de 2009 (por inactividad o falta de registro en 2004), este no será enlazado. Este tipo de casos podrían sobrestimar la salida de unidades económicas del mercado.

Por último, no utilizamos tablas de equivalencia para armonizar recodificaciones de clases de actividad y códigos geográficos. A pesar de que los cambios son mínimos entre levantamientos censales consecutivos, podríamos estar clasificando establecimientos como falsos negativos cuando debieron enlazarse.

8. Laboratorio de Microdatos del INEGI

Los Censos Económicos a nivel establecimiento se consideran información confidencial por el INEGI. La única forma de trabajar estos datos es dentro de su Laboratorio de Microdatos, con sede en la Ciudad de México. Si algún investigador está interesado en utilizar los identificadores descritos en este trabajo, tiene que hacer un requerimiento especial y solicitar se incluyan los identificadores del BID en las bases de datos.

Fuentes

- Busso, M., O. Fentanes & S. Levy Algazi. *The Longitudinal Linkage of Mexico's Economic Census 1999-2014*. (No. IDB-TN-01477). Inter-American Development Bank, 2018.
- Christen, P. "A survey of indexing techniques for scalable record linkage and deduplication", en: *IEEE transactions on knowledge and data engineering*, 24 (9), pp. 1537-1555, 2012.
- Jarmin, R. S. & J. Miranda. *The longitudinal business database*. Center for Economic Studies, Working Paper, pp. 2-17, 2002.
- Organización para la Cooperación y el Desarrollo Económicos (OCDE). *Eurostat-OECD manual on business demography statistics*. OCDE, 2008.
- Raffo, J. *Matchit: Stata module to match two datasets based on similar text patterns*. 2017.

Apéndice

A Tablas

Cuadro A1

Tipo de enlace por año

Año	No enlazados			Enlazados		
	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
2009	1 349 264	220 345	1 569 609	214 954	1 939 456	2 154 410
2014	1 850 319	226 016	2 076 335	220 625	1 933 785	2 154 410

Cuadro A2

Tipo de enlace por industria, porcentajes en 2009

Sectores	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
11 Agricultura	74.6	25.4	100.0	1.6	98.4	100.0
21 Minería	71.0	29.0	100.0	3.5	96.5	100.0
22 Utilidades	54.0	46.0	100.0	0.3	99.7	100.0
23 Construcción	86.8	13.2	100.0	4.9	95.1	100.0
31 Manufacturas	86.4	13.6	100.0	8.1	91.9	100.0
32 Manufacturas	87.9	12.1	100.0	9.6	90.4	100.0
33 Manufacturas	88.2	11.8	100.0	8.9	91.1	100.0
43 Mayoristas	86.4	13.6	100.0	9.5	90.5	100.0
46 Minoristas	86.1	13.9	100.0	11.1	88.9	100.0
48 Transportes	87.8	12.2	100.0	11.7	88.3	100.0
49 Transportes	92.8	7.2	100.0	13.4	86.6	100.0
51 Información	89.8	10.2	100.0	10.6	89.4	100.0
52 Finanzas	92.3	7.7	100.0	14.7	85.3	100.0
53 Bienes Raíces	93.9	6.1	100.0	7.6	92.4	100.0
54 Profesionales	91.6	8.4	100.0	13.1	86.9	100.0
55 Gerencia	64.6	35.4	100.0	16.4	83.6	100.0
56 Apoyo	92.7	7.3	100.0	14.8	85.2	100.0
61 Educación	92.0	8.0	100.0	7.6	92.4	100.0
62 Salud	92.6	7.4	100.0	10.7	89.3	100.0
71 Entretenimiento	93.3	6.7	100.0	9.7	90.3	100.0
72 Alimentación	74.4	25.6	100.0	4.0	96.0	100.0
81 Otros	90.2	9.8	100.0	10.1	89.9	100.0
Total	86.0	14.0	100.0	10.0	90.0	100.0

Cuadro A3

Tipo de enlace por industria, establecimientos en 2009

Sectores	V. neg.	F. neg.	Total no enlazados	F. pos.	V. pos.	Total enlazados
11 Agricultura	5 994	2 040	8 034	184	11 225	11 409
21 Minería	1 057	431	1 488	51	1 418	1 469
22 Utilidades	154	131	285	6	2 298	2 304
23 Construcción	9 159	1 389	10 548	395	7 694	8 089
31 Manufacturas	77 899	12 234	90 133	11 819	133 410	145 229
32 Manufacturas	31 899	4 401	36 300	4 664	43 701	48 365
33 Manufacturas	44 322	5 949	50 271	5 945	60 608	66 553
43 Mayoristas	44 323	6 953	51 276	6 363	60 389	66 752
46 Minoristas	587 281	94 567	681 848	117 966	940 708	1 058 674
48 Transportes	6 662	926	7 588	900	6 776	7 676
49 Transportes	1 891	147	2 038	54	349	403
51 Información	7 254	824	8 078	348	2 928	3 276
52 Finanzas	8 224	690	8 914	1 442	8 350	9 792
53 Bienes Raíces	24 875	1 605	26 480	2 107	25 601	27 708
54 Profesionales	36 727	3 364	40 091	5 865	38 739	44 604
55 Gerencia	53	29	82	20	102	122
56 Apoyo	42 372	3 339	45 711	5 220	29 991	35 211
61 Educación	17 443	1 508	18 951	1 860	22 475	24 335
62 Salud	48 133	3 873	52 006	10 084	84 442	94 526
71 Entretenimiento	20 765	1 501	22 266	1 893	17 662	19 555
72 Alimentación	162 661	56 028	218 689	6 944	166 609	173 553
81 Otros	170 116	18 416	188 532	30 824	273 981	304 805
Total	1 349 264	220 345	1 569 609	214 954	1 939 456	2 154 410

Cuadro A4

Tipo de enlace por industria, porcentajes en 2014

Sectores	V. neg.	F. neg.	Total	F. pos.	V. pos.	Total
11 Agricultura	75.9	24.1	100.0	1.4	98.6	100.0
21 Minería	72.3	27.7	100.0	3.6	96.4	100.0
22 Utilidades	68.1	31.9	100.0	0.3	99.7	100.0
23 Construcción	85.4	14.6	100.0	4.5	95.5	100.0
31 Manufacturas	91.3	8.7	100.0	8.5	91.5	100.0
32 Manufacturas	89.6	10.4	100.0	9.6	90.4	100.0
33 Manufacturas	91.7	8.3	100.0	9.3	90.7	100.0
43 Mayoristas	89.8	10.2	100.0	10.1	89.9	100.0
46 Minoristas	88.6	11.4	100.0	11.3	88.7	100.0
48 Transportes	87.8	12.2	100.0	11.7	88.3	100.0
49 Transportes	85.8	14.2	100.0	9.6	90.4	100.0
51 Información	92.0	8.0	100.0	10.2	89.8	100.0
52 Finanzas	94.4	5.6	100.0	14.7	85.3	100.0
53 Bienes Raíces	94.9	5.1	100.0	7.9	92.1	100.0
54 Profesionales	93.8	6.2	100.0	13.4	86.6	100.0
55 Gerencia	83.3	16.7	100.0	24.2	75.8	100.0
56 Apoyo	94.1	5.9	100.0	14.9	85.1	100.0
61 Educación	93.6	6.4	100.0	7.8	92.2	100.0
62 Salud	94.3	5.7	100.0	11.1	88.9	100.0
71 Entretenimiento	94.5	5.5	100.0	9.7	90.3	100.0
72 Alimentación	82.3	17.7	100.0	4.1	95.9	100.0
81 Otros	92.2	7.8	100.0	10.6	89.4	100.0
Total	89.1	10.9	100.0	10.2	89.8	100.0

Cuadro A5

Tipo de enlace por industria, establecimientos en 2014

Sectores	V. neg.	F. neg.	Total no enlazados	F. pos.	V. pos.	Total enlazados
11 Agricultura	6 831	2 168	8 999	165	11 243	11 408
21 Minería	1 130	432	1 562	53	1 417	1 470
22 Utilidades	284	133	417	8	2 296	2 304
23 Construcción	7 667	1 307	8 974	361	7 728	8 089
31 Manufacturas	117 919	11 241	129 160	12 377	132 948	145 325
32 Manufacturas	36 285	4 219	40 504	4 591	43 265	47 856
33 Manufacturas	54 733	4 986	59 719	6 212	60 754	66 966
43 Mayoristas	60 178	6 857	67 035	6 400	56 913	63 313
46 Minoristas	753 618	96 562	850 180	120 426	941 687	1 062 113
48 Transportes	7 721	1 073	8 794	903	6 791	7 694
49 Transportes	958	158	1 116	37	348	385
51 Información	5 532	482	6 014	339	2 985	3 324
52 Finanzas	13 207	788	13 995	1 438	8 328	9 766
53 Bienes Raíces	33 251	1 793	35 044	2 182	25 589	27 771
54 Profesionales	42 038	2 770	44 808	5 969	38 477	44 446
55 Gerencia	194	39	233	30	94	124
56 Apoyo	53 551	3 340	56 891	5 190	29 530	34 720
61 Educación	21 629	1 473	23 102	1 866	21 914	23 780
62 Salud	71 964	4 368	76 332	10 485	84 120	94 605
71 Entretenimiento	28 564	1 664	30 228	1 954	18 210	20 164
72 Alimentación	270 039	57 902	327 941	7 192	166 315	173 507
81 Otros	263 026	22 261	285 287	32 447	272 833	305 280
Total	1 850 319	226 016	2 076 335	220 625	1 933 785	2 154 410

Cuadro A6

Tipo de enlace por estado, porcentajes en 2009

Estado	V. neg.	F. neg.	Total no enlazados	F. pos.	V. pos.	Total enlazados
Aguascalientes	87.6	12.4	100.0	10.7	89.3	100.0
Baja California	87.5	12.5	100.0	12.8	87.2	100.0
Baja California Sur	88.3	11.7	100.0	9.5	90.5	100.0
Campeche	85.8	14.2	100.0	8.5	91.5	100.0
Coahuila de Z.	91.5	8.5	100.0	10.0	90.0	100.0
Colima	85.9	14.1	100.0	8.5	91.5	100.0
Chiapas	84.5	15.5	100.0	11.8	88.2	100.0
Chihuahua	89.1	10.9	100.0	8.8	91.2	100.0
CDMX	79.6	20.4	100.0	10.7	89.3	100.0
Durango	88.7	11.3	100.0	7.6	92.4	100.0
Guanajuato	84.9	15.1	100.0	10.2	89.8	100.0
Guerrero	86.4	13.6	100.0	11.5	88.5	100.0
Hidalgo	86.4	13.6	100.0	10.0	90.0	100.0
Jalisco	86.4	13.6	100.0	9.8	90.2	100.0
México	85.1	14.9	100.0	10.2	89.8	100.0
Michoacán de O.	87.4	12.6	100.0	10.3	89.7	100.0
Morelos	88.2	11.8	100.0	10.8	89.2	100.0
Nayarit	84.7	15.3	100.0	7.1	92.9	100.0
Nuevo León	90.1	9.9	100.0	10.4	89.6	100.0
Oaxaca	84.4	15.6	100.0	9.2	90.8	100.0
Puebla	85.6	14.4	100.0	10.2	89.8	100.0
Querétaro	85.2	14.8	100.0	10.5	89.5	100.0
Quintana Roo	86.2	13.8	100.0	11.7	88.3	100.0
San Luis Potosí	88.3	11.7	100.0	9.1	90.9	100.0
Sinaloa	87.0	13.0	100.0	7.5	92.5	100.0
Sonora	88.0	12.0	100.0	8.1	91.9	100.0
Tabasco	85.4	14.6	100.0	11.9	88.1	100.0
Tamaulipas	89.5	10.5	100.0	9.0	91.0	100.0
Tlaxcala	87.1	12.9	100.0	10.6	89.4	100.0
Veracruz de I. de la Ll.	87.0	13.0	100.0	9.3	90.7	100.0
Yucatán	84.4	15.6	100.0	7.9	92.1	100.0
Zacatecas	89.8	10.2	100.0	7.8	92.2	100.0
Total	86.0	14.0	100.0	10.0	90.0	100.0

Cuadro A7

Tipo de enlace por estado, porcentajes en 2014

Estado	V. neg.	F. neg.	Total no enlazados	F. pos.	V. pos.	Total enlazados
Aguascalientes	90.7	9.3	100.0	10.9	89.1	100.0
Baja California	90.9	9.1	100.0	13.1	86.9	100.0
Baja California Sur	91.9	8.1	100.0	9.7	90.3	100.0
Campeche	88.0	12.0	100.0	8.6	91.4	100.0
Coahuila de Z.	92.0	8.0	100.0	10.2	89.8	100.0
Colima	88.6	11.4	100.0	8.8	91.2	100.0
Chiapas	89.0	11.0	100.0	12.4	87.6	100.0
Chihuahua	90.7	9.3	100.0	8.9	91.1	100.0
CDMX	82.7	17.3	100.0	11.1	88.9	100.0
Durango	91.0	9.0	100.0	7.9	92.1	100.0
Guanajuato	90.2	9.8	100.0	10.8	89.2	100.0
Guerrero	88.1	11.9	100.0	11.5	88.5	100.0
Hidalgo	90.6	9.4	100.0	10.4	89.6	100.0
Jalisco	90.4	9.6	100.0	10.1	89.9	100.0
México	89.0	11.0	100.0	10.6	89.4	100.0
Michoacán de O.	90.0	10.0	100.0	10.4	89.6	100.0
Morelos	89.8	10.2	100.0	10.6	89.4	100.0
Nayarit	89.4	10.6	100.0	7.5	92.5	100.0
Nuevo León	91.2	8.8	100.0	10.2	89.8	100.0
Oaxaca	89.8	10.2	100.0	9.6	90.4	100.0
Puebla	89.5	10.5	100.0	10.6	89.4	100.0
Querétaro	90.2	9.8	100.0	10.8	89.2	100.0
Quintana Roo	89.5	10.5	100.0	11.7	88.3	100.0
San Luis Potosí	90.8	9.2	100.0	9.2	90.8	100.0
Sinaloa	90.5	9.5	100.0	7.7	92.3	100.0
Sonora	89.8	10.2	100.0	8.4	91.6	100.0
Tabasco	88.5	11.5	100.0	12.2	87.8	100.0
Tamaulipas	90.2	9.8	100.0	8.8	91.2	100.0
Tlaxcala	90.6	9.4	100.0	11.0	89.0	100.0
Veracruz de I. de la LL.	88.4	11.6	100.0	9.5	90.5	100.0
Yucatán	88.4	11.6	100.0	8.1	91.9	100.0
Zacatecas	91.3	8.7	100.0	8.0	92.0	100.0
Total	89.1	10.9	100.0	10.2	89.8	100.0

B Códigos de corrección de inflexiones

Instalación

Los comandos se definen en los archivos *estandariza.ado* y *separa.ado*. Para instalarlos, basta depositar los ADO en el directorio de STATA, donde se alojan los comandos. Para identificar este directorio, se puede ejecutar el comando *sysdir list*; la carpeta está etiquetada como PLUS.

Una vez que se depositan los ADO en el directorio PLUS, es necesario definir una hoja de Excel para el comando a utilizar. Para *estandariza.ado*, se necesitan dos columnas, una con las cadenas de caracteres a estandarizar y una segunda columna con su versión estandarizada. Para el comando *separa.ado* solo se necesita una columna, es decir, una que contenga en cada celda las cadenas de caracteres a identificar y separar.

Sintaxis

Comando *estandariza*

`estandariza var, gen() dexcel() sheet() space()`

Opciones:

`gen()` Nombre de la variable que contiene la versión estandarizada de `var`.

`dexcel()` Directorio y nombre de la hoja de Excel. Ejemplo: "C:/.../doc.xlsx".

`sheet()` Hoja dentro del documento de Excel. Ejemplo: "Hoja1".

`space()` Puede tomar los valores `y` y `n`. Si se selecciona `y`, se estandarizan las cadenas de texto solo cuando formen una palabra independiente. Si se selecciona `n`, se estandariza sin importar en qué parte de la cadena de texto aparezca.

Comando *separa*

`gen(*)` Genera dos variables: `*_1` y `*_2`. `*_1` es la parte no especificada en el Excel. `*_2` es alguna de las cadenas de texto especificadas en Excel.

`dexcel()` Directorio y nombre de la hoja de Excel. Ejemplo: "C:/.../doc.xlsx".

`sheet()` Hoja dentro del documento de Excel. Ejemplo: "Hoja1".

`space()` Puede tomar los valores `y` y `n`. Si se selecciona `y`, se estandarizan las cadenas de texto solo cuando formen una palabra independiente. Si se selecciona `n`, se estandariza no importa en qué parte de la cadena de texto aparezca.

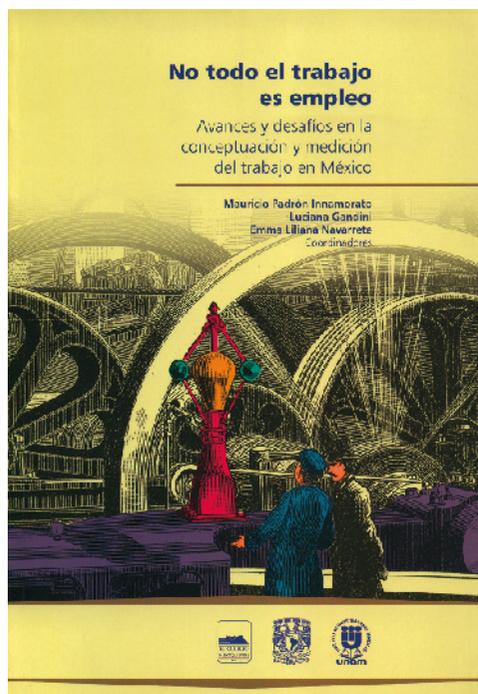
Solo se puede actuar sobre aquello que se mide

If you don't measure it, you can't act on it

Padrón Innamorato, Mauricio; Luciana Gandini y Emma Liliana Navarrete (coords.).

No todo el trabajo es empleo. Avances y desafíos en la conceptualización y medición del trabajo en México.

México, El Colegio Mexiquense, AC-Instituto de Investigaciones Jurídicas de la UNAM, 2017.



Reseña

Adriana Pérez Amador*

“Si no puedes contarlo, no cuenta”¹ es la sentencia anónima que, sin embargo, representa un principio del servicio público de información al que nos apegamos un gran colectivo de investigadores y estadistas que buscamos visibilizar la realidad.

El dato se construye. Una vigilancia epistemológica sobre tal proceso es siempre pertinente, y a ella contribuye la revisión y discusión sobre las aproximaciones teóricas, conceptuales y metodológicas, así como el contexto, con base en las cuales se genera el conocimiento alrededor de ese aspecto de la realidad que el dato busca representar.

* Instituto Nacional de Estadística y Geografía (INEGI), adriana.amador@inegi.org.mx

¹ Padrón *et al.*, 2017, p. 183.

Gran parte del conocimiento sobre el trabajo y los mercados laborales ha partido de datos estadísticos que corresponden a categorías, clasificaciones y conceptos un tanto estáticos frente a una realidad dinámica, cuya transformación da lugar a expresiones donde los datos pueden quedar rebasados bajo los mismos constructos. Por ello, el otrora Grupo Científico de Mercados Laborales de la Sociedad Mexicana de Demografía —ahora constituido como Red Temática Trabajo y Condiciones Laborales (teTra) del Consejo Nacional de Ciencia y Tecnología— llevó a cabo en el 2015 un seminario-taller para examinar la resolución (*Resolución I*) sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo (FT) derivada de la 19.ª Conferencia Internacional de Estadísticos del Trabajo (CIET),² celebrada en Ginebra, Suiza, en octubre del 2013. El fruto de ese examen es la obra coordinada por Mauricio Padrón Innamorato, Luciana Gandini y Emma Liliana Navarrete: *No todo el trabajo es empleo. Avances y desafíos en la conceptualización y medición del trabajo en México*, título por demás sugerente, el cual refuerza la noción de que *trabajo* es lo que engloba toda actividad —remunerada o no— que implica tiempo que puede ser medido, encaminada a la producción de bienes o servicios, independientemente del destino final previsto para éstos.

La noción de que el empleo siempre se trata de trabajo —pero éste no siempre es aquél— es, incluso, algo que ya esboza la principal fuente de información laboral en México: la Encuesta Nacional de Ocupación y Empleo (ENOE), así como otras encuestas especiales que también realiza el Instituto Nacional de Estadística y Geografía (INEGI), como la Nacional de Uso del Tiempo (ENUT).

Esta obra, en la que confluyen 14 investigadores de distintas instituciones, nacionales e internacionales, resulta muy pertinente porque logra plasmar diversas inquietudes alrededor de la nueva forma en que serán abordados conceptos como *trabajo*, *fuerza de trabajo*, *ocupación* y *subutilización* a partir de la *Resolución I*. ¿Cuál es su relevancia? En primer término, se trata del acuerdo —resultado de la revisión del propio concepto de trabajo— que surgió por la necesidad de reconocer como tal a todas las formas de generar bienes y servicios (pasen o no por el mercado) y de ofrecer un marco para su medición, mejorando la comparabilidad de las estadísticas laborales entre los países. Apareció, también, de la necesidad —a la luz de las crisis económicas y financieras a nivel mundial— de hacer frente a las limitaciones del indicador de mayor uso, la tasa de desocupación (o de desempleo abierto) y sus limitaciones para orientar la política pública. Así, la *Resolución I* proporciona medidas de subutilización de la FT que dan mejor cuenta de la carencia de trabajo remunerado requerido por la población y del monto de la población con otras formas de demanda insatisfecha de trabajo remunerado.

2 Mecanismo normativo tripartita en el que confluyen gobiernos (a través de las ONE o ministerios del trabajo), representantes de empleadores y trabajadores.

La *Resolución I*, entonces, amplía de manera simultánea dos horizontes: el de la noción misma de trabajo, que cubre y clasifica modalidades de contribución a la producción y reproducción social (que pasan o no por el mercado) y el de inserción laboral, mismo que no necesariamente transita por la búsqueda activa, sino que conlleva un concepto más amplio, que es el de disponibilidad, el cual deja de privilegiar ciertas formas de proceder de las personas para conseguir empleo, que solo tienen sentido en ciertas fases de su ciclo de vida o en contextos y circunstancias en los que dichos procedimientos pueden resultar eficientes.

Al redefinir el horizonte de inserción laboral, la *Resolución I* establece las poblaciones básicas de referencia: fuerza laboral, fuerza laboral potencial y fuerza laboral ampliada. Al interior de la primera delimita el subconjunto de trabajo vinculado a los mercados laborales: la ocupación por pago o beneficio, así como su relación con el marco conceptual del sistema de contabilidad nacional; es decir, elimina el concepto de Población Económicamente Activa (PEA), restringe el de FT y amplía el del trabajo; y, además de señalar los criterios para definir cada categoría, establece la medición del trabajo en la ocupación y las nuevas medidas de subutilización de la fuerza de trabajo, proponiendo cuatro indicadores denominados Tasas de Subutilización (TSU) desde la 1 hasta la 4. La principal implicación de una nueva definición estadística del *trabajo*, acordada por la comunidad internacional, es que las encuestas nacionales sobre fuerza de trabajo deben ser actualizadas.

Esta publicación consta de seis capítulos y un anexo, además de dos prólogos, uno a cargo de Brígida García (investigadora experta en mercados laborales) y el otro de Juan Trejo (subdirector de Diseño Conceptual de Encuestas de Empleo del INEGI).

En la introducción, los coordinadores exponen el porqué del libro, su propuesta metodológica y el contenido de cada capítulo; donde plantean una pregunta central: ¿cómo se transformará la ENOE? Este cuestionamiento es muy pertinente en la medida en que las ediciones 19 (2013) y 20 (2018) de la CIET son las más importantes y con mayor alcance desde la realizada en 1982, que estableció el estándar de las encuestas de fuerza laboral en el mundo. Esta historia³ se encuentra en el capítulo 1, que aborda "...La importancia de medir la fuerza de trabajo en México, o de cuando la realidad supera a la fuente de información..."; ahí, Emma Liliana Navarrete y Mauricio Padrón hacen un buen recuento de esta necesidad, de la larga tradición en la generación de información estadística con la que cuenta el país gracias a la labor del INEGI y de las diversas CIET que han delineado el diseño conceptual de las encuestas de empleo. La reflexión que de

3 Excluye a la 20.ª CIET porque concluyó en octubre del 2018, posterior a la edición de esta obra.

aquí se deriva es acerca de la exigencia de transformación de las fuentes con el fin de que sus datos no se queden atrás, sino que tengan la capacidad de dar cuenta de las propias transformaciones en el mundo del trabajo, permitiendo una buena representación de esa realidad en constante cambio.

En este recuento, los autores hacen mención del impacto de la ratificación, en abril de 1988, del *Convenio 160 de la Organización Internacional del Trabajo (OIT)* sobre estadísticas del trabajo de 1985; al tratarse de un acuerdo vinculante, compromete a México (con base en los artículos 7 al 9) a compilar estadísticas continuas representativas del país sobre categorías centrales como la PEA, su estructura y distribución para el análisis detallado y como datos de referencia; el empleo, el desempleo y el subempleo visible; y sobre las ganancias medias y horas promedio de trabajo (efectivas y pagadas) para todas las categorías de obreros y empleados; así como principales ramas de actividad.⁴ La implicación más importante de este compromiso fue el lanzamiento de la Encuesta Nacional de Empleo, la cual se levantó de 1988 al 2004 y fue sustituida por la ENOE en el 2005.

Por ello, los siguientes capítulos de la publicación profundizan en los distintos aspectos relacionados con las implicaciones de dicha actualización. La *Resolución I* establece cinco formas excluyentes de trabajo; las cuatro primeras son: trabajo en la ocupación, el destinado a la producción para el autoconsumo, trabajo en formación y el voluntario; además, existen otras formas que no se observan en el ámbito de la población residente en viviendas particulares sino, más bien, entre quienes habitan en las colectivas. La intención que guía la revisión de todos los autores es dar evidencia del impacto que tendrían las modificaciones del cuestionario de la ENOE al incorporar el nuevo marco conceptual, con énfasis en las poblaciones susceptibles de ser *invisibilizadas*.

En el capítulo 2, Nina Castro, Ana Escoto y Edith Pacheco exponen los riesgos de exclusión de poblaciones vulnerables derivados de los cambios de definición o refinamiento del concepto de ocupación, que ahora hace referencia a "...trabajo a cambio de remuneración o ganancias...".

En el apartado 3, Sergio Gaxiola, Clara Márquez y Valeria Montoya analizan el impacto en los indicadores que buscan dar cuenta del nivel de la subutilización de la FT, principalmente los desocupados, y su caracterización; abordan, además, las otras medidas de subutilización que, más allá de la desocupación y subocupación, pretenden dar cuenta

4 También, hubo el compromiso de cumplir con los artículos 11, 12, 14 y 15, los cuales implican el compilar información sobre el costo de la mano de obra respecto a las principales ramas de actividad económica, para el cálculo del Índice de Precios al Consumidor, lesiones profesionales y enfermedades profesionales para todas las ramas de actividad y sobre conflictos del trabajo para todas las ramas de actividad.

de la fuerza de trabajo potencial (FTP), que es la población que, estando fuera de la FT, mantiene un vínculo con el mercado laboral, ya sea que realicen acciones de búsqueda —a pesar de no tener disponibilidad inmediata por encontrarse, generalmente, inserta en el trabajo de producción para el autoconsumo—, o bien, se trate de población disponible, pero que por razones de mercado no lleva a cabo acciones de búsqueda.

Visibilizar de manera adecuada los vínculos con el mercado de trabajo a través del deseo y la disponibilidad representa un reto importante para las fuentes de información que, al mismo tiempo, deben poner especial cuidado en que la manera de preguntar al respecto no induzca una respuesta socialmente deseable o se cree una intención donde no la había.

Sobre las otras formas de trabajo reconocidas por la OIT para su medición por separado (el de producción para autoconsumo, en formación no remunerado y voluntario) tratan los capítulos cuatro, cinco y seis.

En el 4, Nelson Florez y Edith Pacheco, si bien apuntan de qué tamaño sería la exclusión de trabajadores en la producción de bienes para autoconsumo, también hacen recuento de lo que sí se visibiliza, pero que antes quedaba sin tomarse en cuenta.

En el 5, Liliana Estrada Quiroz presenta los elementos básicos que permiten medir, en el marco de la 19.^a CIET, el trabajo en formación no remunerado, en el cual se adquiere experiencia, competencias y también transferencias, imbricado en el ámbito educativo, del hogar y las empresas, a raíz de la reforma laboral del 2012.

En el último capítulo, Viridiana Sosa, Patricia Román y Luciana Gandini señalan que, no obstante, los avances recorridos en su conceptualización y medición, aún queda camino por transitar y la 19.^a CIET resulta una oportunidad. Así, en este apartado se encuentra la respuesta a: ¿qué debería encontrarse en un módulo de trabajo voluntario?, así como sus antecedentes.

En un anexo final está el cuestionario básico actual de la ENOE y la *Resolución I* de la 19.^a CIET.

El proceso de transición de la ENOE hacia un nuevo modelo de encuesta de fuerza de trabajo en México ya ha iniciado: aunado al *Convenio 160* y con base en su marco normativo, el INEGI debe analizar y adecuar al ámbito nacional las recomendaciones que emitan los organismos internacionales para la generación de información; por ello, a raíz de la resolución de la 19.^a CIET, en el Instituto se han tenido que tomar decisiones a corto y largo plazo para emprender, con mucho cuidado, ese camino de cambio, que se estima concluirá en el 2022.

Los trabajos iniciaron en el 2017 y en abril de ese año se estableció el plan de colaboración con la OIT enfocado en la transición a un nuevo modelo de encuesta laboral. Para ello, se integró un grupo técnico que involucra tanto las áreas de operación de encuestas en hogares, como al área de Investigación del INEGI, que ha tenido participación en los comités de redacción de las resoluciones de la 19.^a y la 20.^a CIET. Desde ese mismo año, se inició una serie de pruebas cognitivas. La primera ronda, de julio del 2017, acompañada por la OIT, se orientó en contenidos de la 19.^a CIET (y algunos de la 20.^a). Las entrevistas estuvieron centradas en perfiles: desocupados, subocupados, desalentados, personas no disponibles para trabajar con deseo o interés de trabajar, trabajadores sin pago (familiares y no familiares) y trabajadores en producción de bienes para el autoconsumo. El informe y análisis de esa primera prueba cognitiva fue remitido a la OIT (Ginebra) en diciembre del 2017.

En la reunión subsecuente se entregaron ejemplares de esta obra a los colegas de la OIT y, durante la primera quincena de septiembre —en vísperas de la realización de la 20.^a CIET, celebrada del 10 al 19 de octubre en Ginebra, Suiza— se hizo una nueva entrega del informe a la OIT acerca de las siguientes rondas de marzo y junio del 2018.

Por su parte, la cuarta ronda, de noviembre del 2018, vino acompañada de un módulo sobre medición de la ocupación inadecuada en relación con habilidades, impulsada por la OIT, y calidad del ambiente laboral fomentada por la OCDE.

Como lo señala Brígida García en el prólogo de la obra, se trata de un llamado a la reflexión en torno a las implicaciones de la *Resolución I*, considerando sus riesgos y también sus oportunidades. Con ella se abrió el debate en México sobre la

necesidad de dar cuenta del volumen y las condiciones de la fuerza de trabajo y la comparabilidad internacional. Esta publicación ayuda a comprender cómo, a raíz de esta resolución, se orienta a la integración de un subsistema de estadísticas del trabajo que visibilicen todas las formas de trabajo y, además, la pluriocupación y la pluriactividad, en virtud de que, aun cuando se trate de conceptos distintos, no son mutuamente excluyentes para la población que las desempeña, lo cual implica una adecuada visibilización de la población con más de una ocupación y de todas aquellas actividades productivas dentro de la frontera de producción, con énfasis en el trabajo de producción para autoconsumo y más allá de este umbral.

Es de suma importancia que la comunidad académica esté al tanto de una resolución tan impactante como ésta, lo cual fortalece el diálogo con la oficina nacional de estadística de una manera sin precedente, pues antes las resoluciones solo parecían dirigidas al funcionariado de estas áreas, cuya tarea tradicional era implementar primero y comunicar después. Aquí, la interlocución y la comprensión de lo que está en juego ha enriquecido el proceso de cambio desde su arranque.

Como bien se da cuenta, todo lo que se encuentra en la invisibilidad está al margen del derecho y no es regulado, con los costos sociales y de ingobernabilidad que ello implica. Así, esta obra es útil para comprender los antecedentes, el contexto y el impacto que tendría la adopción de la *Resolución I*. Su lectura es obligada para los estudiosos del trabajo, para quienes estamos inmersos en el proceso de generación de información —y, más aún, en la transición de la ENOE—, así como para los involucrados en el diseño, implementación y evaluación de políticas públicas, reconociendo que solo se actúa sobre aquello que se mide.

Por un logro sostenido en el tiempo de capacidades estadísticas

For a sustained achievement of statistical capacities

Dargent, Eduardo, Gabriela Lotta, José Antonio
Mejía y Gilberto Moncada.

*¿A quién le importa saber? La economía política de la
capacidad estadística en América Latina.*

Washington, D. C., Banco Interamericano de Desarrollo, División de Innovación para
Servir al Ciudadano, 2018.



Reseña

Ximena A. Clark*

¿A quién le importa que un país desarrolle una alta capacidad para producir estadísticas de calidad, oportunas y relevantes en la toma de decisiones de la sociedad? ¿De qué factores depende que una oficina nacional de estadística (ONE) tenga la suficiente autonomía técnica para responder a las demandas de los usuarios de información tanto internos como externos al país? ¿Por qué algunos países desarrollan un sistema estadístico nacional (SEN) altamente técnico y coordinado por el ente rector, la ONE, mientras que otros de características similares no logran tal nivel de desarrollo estadístico? ¿Cuánta conciencia existe, a nivel de hacedores de política, de la importancia de elaborar políticas públicas con base en evidencia y cómo ésta se traduce en el fortalecimiento de la capacidad estadística de un país? ¿Cómo afectan las crisis de diferentes tipos a la capacidad estadística de una nación? Éstas y muchas otras preguntas relacionadas son las que aborda el libro.

* Consultora independiente en institucionalidad estadística, ximenaclark@gmail.com

A través de un relato fluido y motivador, basado en el estudio de casos para 10 países de América Latina (Argentina, Bolivia, Brasil, Colombia, Ecuador, El Salvador, Guatemala, México, Perú y República Dominicana) que se sistematiza mediante un marco metodológico diseñado especialmente, los autores nos introducen en la compleja tarea de responder esas preguntas.

El análisis que nos presentan resulta ser pionero en *definir y sistematizar* los diferentes factores y el tipo de influencia que éstos ejercen en la capacidad estadística de un país, la cual contempla "...la habilidad de producir información (relevante, oportuna y de calidad), el uso que se hace de esos datos y el alineamiento entre la oferta y la demanda de estadísticas...". Más explícitamente, en lo que respecta a su definición formal, realizan una exhaustiva revisión de la literatura, destacando la referencia que hacen a Beccaria (2017),¹ quien la define como: "...la existencia de una estructura o sistema permanente que disponga de los recursos necesarios para generar de manera sostenida datos estadísticos relevantes y de calidad, y para difundirlos adecuada y oportunamente...". Asimismo, citan a Alessandro (2017)² para describir los elementos esenciales que la caracterizan: "...(i) que la oficina estadística cuente con suficientes recursos presupuestarios, humanos y tecnológicos para su trabajo; y (ii) que el marco institucional del sector proteja la independencia técnica y favorezca la coordinación de las distintas entidades productoras de estadística oficial...". También, destaca la mención de Taylor (2016),³ quien identifica tres dimensiones de la capacidad estadística: "...1) Las capacidades del sector de estadística de ofrecer los productos de su trabajo de manera eficiente y eficaz; 2) La autonomía de las instituciones, o el grado en que el sistema estadístico es capaz de producir esta información y hacerla pública sin interferencias de las autoridades y los políticos; y 3) Los recursos humanos y presupuestarios disponibles, que sean suficientes para producir las estadísticas necesarias y periódicas para la toma de decisiones...".

Al respecto, cabe mencionar que el fortalecimiento de la capacidad estadística de los países ha cobrado especial relevancia en el actual contexto internacional, donde se evidencia una creciente conciencia por parte de muchos de éstos por lograr un desarrollo sustentable e inclusivo. Este desafío, materializado hoy en día en los 17 Objetivos de Desarrollo Sustentable de la *Agenda 2030* (con sus más de 230 indicadores para monitorear el progreso de los países), se traduce a la vez en una fuerte demanda por fortalecer la capacidad estadística de los SEN y, en particular, de las ONE de las diferentes naciones.

Por otra parte, la importancia de generar conciencia sobre el valor público de las estadísticas, promoviendo políticas adecuadas de difusión y uso de éstas, que a su vez incentiven una gestión pública basada en evidencia y resultados, constituye otra motivación para fomentar el fortalecimiento estadístico de un país. Más aún, en el contexto actual en que abundan las bases de datos de diverso origen, volumen y calidad, resulta esencial que este fortalecimiento y mayor capacidad estadística incorporen aspectos de calidad, coordinación, acceso y uso de fuentes alternativas de datos, así como autonomía técnica y financiera de las oficinas de estadística como parte de sus elementos esenciales, apoyados además por los avances tecnológicos disponibles hoy en día.

Dicho lo anterior, uno esperaría encontrar en cada país un amplio y transversal apoyo por parte de las autoridades de gobierno para fortalecer su SEN y, en particular, la ONE. Sin embargo, la realidad revela otra situación, presentando una alta heterogeneidad en las capacidades estadísticas de las naciones de la región (10 casos analizados) de América Latina. Aquella se materializa en marcadas diferencias en cuanto a los recursos humanos, financieros y tecnológicos con los que cuentan los respectivos sistemas estadísticos de los países, así como en el nivel de autonomía y la habilidad de las ONE para ejercer sus funciones de rectoría y coordinación del SEN, entre otras.

¿Por qué se dan esas diferencias? ¿Qué factores las explican? Estas preguntas son, en palabras simples, la motivación del libro y, para responderlas, el análisis comparativo que se deriva de estos 10 casos propone abordar, de manera muy integral, las distintas dimensiones que pueden explicar las diferencias. Para ello, combina una evaluación cualitativa con un indicador cuantitativo —elaborado por el BID— que incorpora variadas dimensiones de la capacidad estadística de un país: el Índice de Capacidad Estadística Nacional (ICEN). Éste se basa en la definición provista por Beccaria (2017) y consiste en la medición de cuatro dimensiones principales: a) recursos (financieros, humanos, tecnológicos y físicos) que posee una ONE para producir información, b) normas y prácticas institucionales (aplicables a la ONE y al SEN, en particular en el ámbito de coordinación), c) metodologías (acorde con estándares internacionales y sujetas a control de calidad) y d) difusión (transparente, oportuna y con resguardo de la confidencialidad de la información) a los diferentes usuarios.

De acuerdo con la medición del ICEN para el periodo 2015-2016 en los 10 países del estudio, México, Colombia y Brasil lideran el *ranking* en capacidad estadística, mientras que Argentina, Bolivia, El Salvador y Guatemala presentan los menores puntajes. No obstante, más interesante que el *ranking* propiamente tal, es la correlación que se evidencia entre los niveles de capacidad estadística y aquéllos de capacidad estatal (medidos estos últimos por indicadores como el Índice de Desarrollo de Servicio Civil, del BID, y el Índice de Efectividad del Gasto); incluso más, la existencia de casos donde el nivel de la capacidad estadística supera el nivel esperado acorde con

1 Beccaria, L. *Capacidad estadística: una propuesta para su medición y aproximación práctica*. Washington, D. C., Banco Interamericano de Desarrollo, 2017.

2 Alessandro, M. *La economía política de la capacidad estadística: una aproximación práctica*. Washington, D. C., Banco Interamericano de Desarrollo, 2017.

3 Taylor, M. *The Political Economy of Statistical Capacity: A Theoretical Approach*. Washington, D. C., Banco Interamericano de Desarrollo, 2016.

la capacidad estatal (Colombia, Perú, Ecuador y México) da cuenta de la importancia de otros factores que influyen en dicha capacidad estadística (colaboración con organismos internacionales y/o alianzas estratégicas con otras entidades, por ejemplo). Lo anterior permite vislumbrar un “potencial propio de desarrollo de capacidad” que poseen las ONE, que no depende solo de las capacidades del Estado, y que las asociaría a las llamadas “islas de eficiencia” o “instituciones con niveles técnicos y de autonomía superiores al promedio general del Estado”.

Por su parte, el análisis de estudio de casos de las 10 naciones incorpora una breve descripción de la historia estadística de cada una (y su respectiva ONE), para luego destacar aquello que marca el fortalecimiento o debilitamiento estadístico experimentado. Es así como en Argentina (INDEC) se destacan los costos de la intromisión del poder Ejecutivo en la capacidad estadística. En Bolivia (INE), los Objetivos de Desarrollo del Milenio (con fuerte foco en los programas de reducción de la pobreza) y la descentralización se presentan como catalizadores de ésta. A su vez, en Brasil (IBGE) se enfatiza la importancia de construir legitimidad dentro y fuera del Estado, destacando iniciativas como la Escuela Nacional de Ciencias Estadísticas, la integración de áreas geográficas y estadísticas, además de la consolidación de relaciones con organismos gubernamentales, con la prensa, la sociedad civil y los organismos internacionales, dando lugar a lo que se denomina una *autonomía enraizada*, entre otras. En el caso de Colombia (DANE), la influencia de áreas tecnocráticas y demandas internacionales se traducen en oportunidades de fortalecimiento estadístico. Por su lado, en Ecuador (INEC) se destacan los beneficios y peligros de un presidente (de la nación) interesado en el SEN y en El Salvador (DIGESTYC) se presenta la polarización política como barrera para la reforma de la ONE. En el caso de Guatemala (INE), el análisis se orienta hacia la disyuntiva de tener un sistema de partidos políticos débiles y volátiles y el impacto que ello tiene para promover una reforma estadística. Por otra parte, en México (INEGI) se resaltan las crisis (económico-financieras), sumadas a la influencia de actores internacionales y el proceso de democratización como oportunidades de fortalecimiento de la ONE. Por último, en Perú (INEI) se observa un gradual fortalecimiento estadístico, pero cuya continuidad no está asegurada y en República Dominicana se destaca la existencia de un liderazgo continuo (por parte de la ONE) que enfrenta, sin embargo, vetos al interior del Estado.

A partir de los insumos analíticos provistos por la medición del ICEN y de los estudios de casos, y en su esfuerzo por sistematizar los diferentes factores que influyen en las diferencias de capacidad estadística de los países, los autores identifican tres niveles de análisis: 1) los procesos históricos comunes, 2) los factores de economía política que influyen en la capacidad estadística y 3) las acciones (estratégicas) desarrolladas por los líderes de las ONE y funcionarios públicos para fortalecer la institución.

Con respecto a los procesos históricos comunes, se indica que en general han favorecido una mayor capacidad estadística en la región. Éstos incluyen, principalmente, las reformas del Estado, las exigencias por una mayor transparencia y los apuros fiscales. Dentro de las primeras, se consideran, por ejemplo, los procesos de democratización y descentralización, así como los programas de reducción de la pobreza, todos los cuales han requerido de más y mejores estadísticas.

Con respecto a los factores de economía política, que constituyen el núcleo central del análisis, los autores desarrollan un marco conceptual específico. Para ello, realizan un análisis que se sustenta, por un lado, en la literatura sobre la capacidad del Estado, los factores que explican su desarrollo y su vinculación con el desarrollo de capacidades estadísticas y, por el otro, en la importancia de las estadísticas para las políticas públicas. Asimismo, hacen mención a la literatura que establece “...una relación positiva entre el desarrollo de sistemas estadísticos de calidad y el desarrollo de capacidades estatales...”. Más aún, se identifica un potencial círculo virtuoso entre el desarrollo de capacidades estadísticas y el desarrollo de un Estado más profesional, donde las primeras sirven para diseñar e implementar mejores políticas públicas, basadas en evidencia, y a la vez para monitorear el resultado de éstas y el uso eficiente de los recursos públicos. En otras palabras, esto daría lugar a “...un proceso de interdependencia e influencia mutua entre la construcción de las capacidades del Estado y la capacidad estadística...”.

En la práctica, sin embargo, los Estados no siempre están en condiciones de fortalecer las capacidades estadísticas de sus SEN, en parte por las debilidades propias que exhiben. A su vez, no siempre tienen la disposición o interés en fortalecerlas debido a que éstas luego permiten un mayor monitoreo de las actividades del propio Estado o porque sencillamente no son prioridad y no enfrentan presiones reales para que lo sean, de ahí que sea solo un *potencial* círculo virtuoso.

En cuanto a la importancia de las estadísticas para las políticas públicas y su relación con factores de economía política que determinan la capacidad estadística, los autores revelan que hay escasa literatura al respecto. Por lo mismo, y con base en su análisis previo de casos e indicadores, introducen un marco conceptual que identifica tales factores; entre éstos se cuentan: 1) la existencia de áreas técnicas en las instituciones del Estado que requieren datos de calidad para realizar su labor; 2) la demanda por parte de organismos externos u organizaciones internacionales de estadísticas confiables, de calidad y oportunas de un país, las cuales en ocasiones condicionan a su vez la ayuda y/o colaboración prestada por éstos; 3) el rol que juegan altas autoridades, en particular los presidentes de la nación, para promover una política pública basada en evidencia, con estadísticas de calidad; 4) el efecto (positivo o negativo) que imprimen las crisis económicas y políticas en un determinado momento; y 5) la presencia de una comunidad de usuarios, en particular de una sociedad

civil, académicos y empresas que ejerzan una demanda permanente por estadísticas de calidad.

Si bien estos cinco factores se asocian en forma natural a una mayor capacidad estadística del país, y en particular de la respectiva ONE, cabe señalar algunas legítimas aprehensiones que indican los autores. El primero de éstos pudiera, eventualmente, influir de manera negativa en la capacidad estadística de la ONE, si es que estos sectores técnicos especializados en las instituciones del Estado son los encargados también de producir estadísticas y perciben el fortalecimiento de la ONE como una amenaza a su labor; tal es el caso de los bancos centrales que elaboran las cuentas nacionales del país y la tensión que se genera cuando se analiza el traspaso de esta medición a las ONE.

En cuanto a la demanda de organizaciones internacionales, si bien ésta se traduce normalmente en una significativa mejora de las capacidades estadísticas del país, existe el riesgo de que la imposición de su agenda no asegure necesariamente un fortalecimiento institucional sostenible.

En el caso de la influencia de los presidentes (o altas autoridades públicas) para promover una mayor capacidad estadística, resulta importante que ésta vaya acompañada de los necesarios cambios institucionales, para no terminar siendo un apoyo simplemente *episódico* o incluso para evitar que se revierta.

Con respecto a los dos últimos factores de economía política, las crisis económicas y/o políticas pueden tener impactos negativos o positivos en la capacidad estadística. En consecuencia, para aprovechar de manera eficaz las oportunidades que éstas presentan, resulta relevante establecer alianzas estratégicas con actores nacionales que demanden estadísticas de calidad en forma permanente (academia, sociedad civil, prensa y gremios empresariales, entre otros).

Por otra parte, cabe enfatizar que los factores de economía política identificados en este marco conceptual requieren ser complementados con un marco institucional que asegure independencia técnica y/o autonomía profesional y de gestión para las instituciones integrantes del sistema estadístico nacional. En particular, en el caso de la ONE —candidata natural para ejercer la rectoría y coordinación del SEN—, los marcos legales deben asegurar las condiciones para que ésta pueda ejercer eficaz y eficientemente estas funciones. Asimismo, aspectos asociados a una política de disseminación y difusión de estadísticas, didáctica y transparente a la vez, que promueva una verdadera alfabetización estadística en la población, resultan claves para completar la estrategia de fortalecimiento de las capacidades estadísticas

de un país, toda vez que van generando una demanda y uso permanente por información estadística.

Por último, el tercer nivel de análisis identificado por los autores consiste en las acciones de los líderes de las ONE y los mismos funcionarios para fortalecer la institución. Se resalta la importancia de que las mejoras en las capacidades estadísticas de la ONE surjan, además, de un proceso de fortalecimiento interno: en la medida en que estas mejoras vayan acompañadas de una adecuada gestión del cambio, tendrán mayor probabilidad de ser institucionalizadas, y las oportunidades a las que se enfrenten las ONE serán mejor aprovechadas.

La combinación de factores de economía política que incentivan una mayor capacidad estadística, como los ya mencionados, con un adecuado proceso de fortalecimiento institucional interno bien puede ser lo que marque la diferencia entre los avances que presenta una ONE comparada con otra entre países con similares niveles de desarrollo. Para esto, las diferentes acciones estratégicas de los líderes institucionales —como la generación de alianzas con usuarios de la información estadística, internos como externos, o la actualización del marco estadístico institucional— puede resultar clave. En particular, esta interacción y, más aún, las alianzas estratégicas permiten ir consolidando una demanda permanente por estadísticas de calidad que constituyen, por un lado, un incentivo a seguir fortaleciendo las capacidades estadísticas y, por el otro, una potencial protección a la ONE y al SEN frente a eventuales intenciones de interferencia política que atente contra esa capacidad estadística.

En síntesis, esta obra nos presenta un acabado y sistematizado análisis de los diversos procesos y factores que influyen en el fortalecimiento de las capacidades estadísticas de los países, permitiéndonos entender las marcadas diferencias en los niveles de desarrollo estadístico en la región de América Latina. A través de tres niveles de análisis (procesos históricos, factores de economía política relevantes a las capacidades estadísticas y factores relacionados con la construcción de institucionalidad), los autores proveen una verdadera guía práctica y estratégica a la vez para aquellas naciones en proceso de fortalecer sus capacidades estadísticas, entendido esto, además, como un proceso de mejora continua. Frente a los actuales desafíos que enfrentan los SEN (tecnológicos y de innovación, de seguridad de la información y de mayor coordinación, entre otros), se enfatiza, asimismo, la importancia de establecer formas de protección informal —como las alianzas estratégicas con los diversos usuarios de información y con organizaciones internacionales— de manera tal que lograr un fortalecimiento de capacidades sea sostenible en el tiempo.

Colaboran en este número

Gerardo Leyva

Nació en México en 1967. Es licenciado en Economía por la Universidad Autónoma de Aguascalientes (UAA) y tiene una maestría en la misma disciplina por el Instituto Tecnológico Autónomo de México (ITAM) y otra en Ciencia Regional por la *Cornell University*, donde también obtuvo el Doctorado con Especialización en Crecimiento y Desarrollo Económicos; además, cuenta con el Diplomado en Psicología Positiva por la Universidad Iberoamericana. En el ámbito laboral, ha impartido cursos de Teoría Económica en varias universidades y tiene 22 años de experiencia profesional en el Instituto Nacional de Estadística y Geografía (INEGI); ha sido analista, asesor de tres presidentes de esta institución, director de Censos Económicos, director general adjunto de Estadísticas Económicas y, a partir del 2009, director general adjunto de Investigación, desde donde ha impulsado la incorporación de diversas innovaciones temáticas y metodológicas relacionadas con la medición del bienestar subjetivo, la distribución del ingreso, la productividad, el equilibrio general computable, los indicadores de difusión y la generación de información a partir de *big data*, entre otras. Fue integrante del Grupo de Expertos en Medición de la Pobreza de la ONU (conocido como Grupo de Río) y del Comité Técnico para la Medición de la Pobreza en México. Es miembro de los comités de Estudios Económicos del Instituto Mexicano de Ejecutivos de Finanzas (IMEF), del indicador IMEF del Entorno Económico Empresarial y del de Coyuntura de la Asociación Nacional de Tiendas de Autoservicio y Departamentales (ANTAD), así como del Consejo Asesor Técnico del Centro de Estudios Económicos del Sector Privado (CEESP). Participa en los consejos editoriales de las revistas *Políticas Públicas* de la Escuela de Graduados en Administración Pública y Política Pública (EGAP) del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), *Coyuntura Demográfica* de la Sociedad Mexicana de Demografía (SOMEDE) e *Investigación Económica* de la Universidad Nacional Autónoma de México (UNAM) y es editor técnico de *Realidad, Datos y Espacio Revista Internacional de Estadística y Geografía* del INEGI.

Contacto: gerardo.leyva@inegi.org.mx

Olinca Páez

Nació en México en 1978. Es licenciada en Economía por la Universidad Veracruzana y maestra en Demografía por El Colegio de México (COLMEX) y cursó dos diplomados en el Centro de Investigación y Docencia Económicas (CIDE) Región Centro: uno sobre Género, Sexualidad y Derecho y otro en Gobierno, Gestión y Políticas Públicas. Es investigadora desde el 2003 y ha publicado sobre temas variados de economía y demografía para un público diverso, apoyándose de técnicas estadísticas y econométricas. Recientemente se ha enfocado en el estudio del curso de vida de las mujeres, la desigualdad en el uso del tiempo y sus efectos sobre el bienestar subjetivo. En el 2012 obtuvo el segundo lugar del Premio Nacional de Investigación Social y de Opinión Pública, que otorga el Centro de Estudios Sociales y de Opinión Pública de la Cámara

de Diputados, y en el 2018 fue beneficiaria del *Trust Fund for Statistical Capacity Building* del Banco Mundial. Desde septiembre del 2014 es subdirectora de Investigación de Información Econométrica en el INEGI.
Contacto: olinca.paez@inegi.org.mx

Ana Ruth Escoto Castillo

De nacionalidad salvadoreña. Es doctora en Estudios de Población por El COLMEX y maestra en Población y Desarrollo por la Facultad Latinoamericana de Ciencias Sociales (FLACSO), sede México. En la actualidad, es profesora asociada en el Centro de Estudios Teóricos y Multidisciplinarios en Ciencias Sociales de la Facultad de Ciencias Políticas y Sociales de la UNAM. Se interesa por investigar el bienestar de la población; en el presente, analizando los procesos de desigualdad y exclusión en los mercados laborales latinoamericanos y la organización y estructura de los hogares y en el futuro, a través del estudio de la relación del cambio climático y la distribución de ingresos, el consumo energético de los hogares y sus implicaciones ambientales. Es candidata al Sistema Nacional de Investigadores.

Contacto: ana.escoto@politicas.unam.mx

Landy Sánchez Peña

De nacionalidad mexicana. Es doctora en Sociología por la Universidad de Wisconsin-Madison, con una subespecialidad en Geografía. En la actualidad, es profesora-investigadora del Centro de Estudios Demográficos, Urbanos y Ambientales de El COLMEX y miembro del Sistema Nacional de Investigadores con Nivel II. Ha publicado trabajos a niveles nacional e internacional de temas relacionados con ajuste económico, desigualdad, así como sobre población y medio ambiente. Sus proyectos de investigación actuales giran en torno al consumo energético de los hogares, la demografía de la desigualdad y las intersecciones entre cambio climático y bienestar de la población.

Contacto: lsanchez@colmex.mx

Benjamín Temkin

De nacionalidad mexicana. Es doctor en Ciencia Política por la Universidad de Columbia. En el ámbito laboral, es profesor-investigador en FLACSO México, donde también coordina la Maestría y Especialidad en Política y Gestión Energética y Medioambiental. Entre sus líneas de investigación se encuentran la informalidad laboral, el estudio de actitudes políticas y su relación con el bienestar subjetivo.

Contacto: temkin@flacso.edu.mx

Jorge Cruz Ibarra

De nacionalidad mexicana. Es pasante de la carrera de Comunicación en la Universidad Panamericana. Trabaja como profesor de asignatura y es asistente de investigación en FLACSO México; entre sus intereses de investigación se encuentra la relación entre el comportamiento político y la opinión pública.

Contacto: jmcruz@up.edu.mx

Francisco de Jesús Corona Villavicencio

De nacionalidad mexicana. Es licenciado en Economía por la Universidad Autónoma de Baja California (UABC), maestro en Estadística Aplicada por el ITESM y doctor en Economía y Métodos Cuantitativos por la Universidad Carlos III de Madrid (UC3M). En la actualidad, es investigador en el INEGI y sus líneas de investigación están relacionadas con el análisis econométrico y pronóstico de series de tiempo; también, tiene una en *Sport Analytics*. Pertenece al Sistema Nacional de Investigadores, Nivel Candidato.

Contacto: franciscoj.corona@inegi.org.mx

Jesús López-Pérez

De nacionalidad mexicana. Es licenciado en Economía por el ITESM y maestro en Estadística Aplicada en la misma institución. En la actualidad, es subdirector de Investigación en el INEGI en temas relacionados con el análisis econométrico de series de tiempo. Anteriormente, ocupó diversos cargos en áreas de administración y análisis de riesgo crediticio en instituciones financieras del sector público y privado.

Contacto: jesus.lopezp@inegi.org.mx

Matías Busso

Nació en Argentina. Es doctor en Economía por la Universidad de Michigan. En el ámbito laboral, es economista líder del Departamento de Investigación del Banco Interamericano de Desarrollo (BID), investigador invitado del Centro de Estudios Distributivos, Laborales y Sociales (CEDLAS), miembro del comité ejecutivo del Red de Pobreza y Desigualdad de Asociación de Economía de América Latina y el Caribe (LACEA, por sus siglas en inglés). Su investigación utiliza teoría y evidencia empírica para analizar el diseño de políticas públicas más efectivas en áreas relacionadas con el mercado laboral, economía urbana, educación y productividad. Ha publicado artículos de investigación en *American Economic Review* y en *The Review of Economics and Statistics*, entre otros medios.

Contacto: MBUSSO@iadb.org

Óscar Eduardo Fentanes Téllez

De nacionalidad mexicana. Cuenta con la Licenciatura en Economía por la Universidad de Toulouse 1 y la Maestría en Economía por el CIDE; actualmente, es estudiante del Doctorado en Economía en la *Toulouse School of Economics*, Francia. Se desempeñó como asistente de investigación en el BID (2014-2018), colaborando de cerca en al menos 10 proyectos de investigación y contribuyendo con la parte empírica del libro de Santiago Levy *Esfuerzos mal recompensados. La elusiva búsqueda de la prosperidad en México* en el 2018. Gran parte de su labor se ha llevado a cabo en el Laboratorio de Microdatos del INEGI, donde ha colaborado en la vinculación de bases de datos, como los Censos Económicos.

Contacto: o.fentanes.t@gmail.com

Santiago Levy Algazi

Nació en México. Realizó sus estudios de Maestría y Doctorado en Economía en la Universidad de Boston. En el terreno laboral, tiene una

gran trayectoria en la Academia, sector público y organismos internacionales: fungió como vicepresidente de Sectores y Conocimiento del BID (2008-2018), fue director general del IMSS (2000-2005) y subsecretario de Hacienda (1994-2000), donde se convirtió en el principal arquitecto del reconocido Programa Progresá Oportunidades; cargos previos incluyen el de presidente de la Comisión Federal de Competencia y director del Programa de Desregulación Económica en la entonces Secretaría de Comercio y Fomento Industrial (hoy de Economía); además, ha ocupado posiciones académicas en el ITAM, la Universidad de Boston y el Instituto para el Desarrollo Económico. Es autor de más de 75 artículos y capítulos de libros en temas como reducción de pobreza, competitividad, energía, entre otros, y entre sus diversos libros destacan *Buenas intenciones, malos resultados* (2008) y *Esfuerzos mal recompensados. La elusiva búsqueda de la prosperidad en México* (2018).

Contacto: santiagolevy4@gmail.com

Adriana Pérez Amador

De nacionalidad mexicana. Es licenciada en Política y Gestión Social por la Universidad Autónoma Metropolitana, maestra en Demografía por El COLMEX y cursó el Doctorado en Ciencias Sociales y Humanidades en la UAA. Ha sido docente en la UAA, donde también ha colaborado en proyectos de investigación, así como en el propio COLMEX; en la actualidad, es subdirectora de Investigación A, adscrita a la Dirección General Adjunta de Investigación del INEGI.

Contacto: adriana.amador@inegi.org.mx

Ximena A. Clark

Nació en Concepción, Chile. Es ingeniera comercial por la Universidad de Concepción y cuenta con dos maestrías en Economía, una por la Universidad de Georgetown (1996, dictada en Santiago, Chile) y la otra por la de Boston (1998, en Estados Unidos de América). En el ámbito laboral, en la actualidad se desempeña como consultora independiente en materias de institucionalidad estadística para el BID y oficinas de estadística; anteriormente, ejerció como directora del Instituto Nacional de Estadísticas (INE) de Chile (marzo del 2014-abril del 2018), asumiendo como principales desafíos la realización del Censo de Población y Vivienda 2017 y el fortalecimiento de la institucionalidad estadística nacional; previo a este cargo, trabajó en la Gerencia de Informaciones Estadísticas del Banco Central de Chile (como jefa del Grupo Microdatos, 2012-2014) y en el Ministerio de Economía de Chile (como jefa de la División de Empresas de Menor Tamaño y como asesora económica, 2007-2011); también, fue consultora y economista del Banco Mundial (2000-2005) y consultora del *Harvard International Institute of Development (HIID)* y del *Center of International Development (CID)* de la Universidad de Harvard (1998-2000).

Contacto: ximenaclark@gmail.com

Política y lineamientos editoriales

REALIDAD, DATOS Y ESPACIO REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA es una publicación cuatrimestral que sirve de enlace entre la generación de la información estadística y geográfica oficial y la investigación académica para compartir el conocimiento entre especialistas e instituciones con propósitos similares.

Se publicarán sólo artículos inéditos y originales relacionados con la situación actual del uso y aplicación de la información estadística y geográfica a nivel nacional e internacional.

Es una revista técnico-científica, bilingüe, cuyos trabajos son arbitrados por pares (especialistas), bajo la metodología doble ciego, con los siguientes criterios de evaluación: trabajos inéditos, originalidad, actualidad y oportunidad de la información, claridad en la definición de propósitos e ideas planteadas, cobertura de los objetivos definidos, estructura metodológica adecuada y congruencia entre la información contenida en el trabajo y las conclusiones.

El resultado del proceso de dictaminación se comunica por correo electrónico y contempla tres variantes: recomendado ampliamente (con modificaciones menores), recomendado (pero condicionado a modificaciones sugeridas) y no recomendado (rechazado). Dos dictámenes aprobados, se notifica al autor que se publica y se envía a corrección de estilo; un aprobado y uno rechazado, se le solicita realizar cambios; y dos rechazados, se notifica la no publicación.

Indizaciones y registros

- LATINDEX Catálogo (Sistema Regional de Información en Línea para Revistas Científicas de América Latina, el Caribe, España y Portugal).
- CLASE (Citas Latinoamericanas en Ciencias Sociales y Humanidades).
- REDIB (Red Iberoamericana de Innovación y Conocimiento Científico).

Lineamientos para publicar

Se publicarán trabajos en español e inglés: artículos de investigación, revisión y divulgación; ensayos; metodologías; informes técnicos; comunicaciones cortas; reseñas de libros; revisiones bibliográficas y estadísticas, entre otros.

1. El artículo —o cualquier otro tipo de escrito de los mencionados— deberá entregarse con una carta dirigida al editor responsable de REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA en la que se proponga el texto para su publicación, que se declare que es inédito y que no ha sido postulado de manera paralela en otro medio. Asimismo, deben incluirse los datos completos del(os) autor(es), nacionalidad(es), institución(es) de adscripción y cargo(s) que ocupa(n), domicilio(s) completo(s), correo(s) electrónico(s) y teléfono(s). Esto debe dirigirse a la atención de la M. en C. Virginia Abrin Batule, virginia.abrin@inegi.org.mx (tel. 5278 10 00, ext. 1161).
2. El trabajo se debe presentar en versión electrónica (formato *Word* o compatible) con: a) extensión no mayor de 20 cuartillas; b) letra Helvética, Arial o Times de 12 puntos y c) interlineado de 1.5 líneas. El material adicional al texto se requiere por separado: a) las imágenes, con resolución de 300 ppp y un tamaño no menor a 17 centímetros de base (ancho) en formato JPG o TIF —no remuestrear (ampliar) imágenes de menor resolución—; si son líneas o mapas, deben entregarse en formato vectorial (EPS o Ai), en caso de incluirse imágenes en mapa de bits, incrustarlas o enviarlas con el nombre con el cual se creó el vínculo (conservando los requerimientos de resolución y tamaño estipulados); y para fotografías, éstas no deben ser menores a 5 megapíxeles; b) las fórmulas o expresiones matemáticas tienen que elaborarse con el editor de ecuaciones propio de *Microsoft*[™], pero en caso de usar *software* de terceros, incluir en la entrega PDF testigo en el cual figuren exactamente cómo deben representarse; c) las gráficas, que incluyan el archivo en *Excel* con el cual se desarrollaron o, en su defecto, la imagen JPG legible, de origen, en alta resolución; y d) los cuadros, que sean editables, no se deben insertar como imagen.
3. La colaboración debe incluir: título del trabajo (en español e inglés o viceversa); resúmenes del trabajo en español e inglés (que no excedan de un párrafo de 10 renglones); palabras clave en español e inglés (mínimo tres, máximo cinco); bibliografía u otras fuentes; así como breve(s) semblanza(s) del(os) autor(es) que no exceda(n) de un párrafo de cinco renglones y que incluya(n) nacionalidad(es), grado(s) académico(s), principal(es) experiencia(s) profesional(es), adscripción(es) laboral(es) actual(es) y dirección(es) electrónica(s) de contacto.
4. Las referencias bibliográficas u otras fuentes deberán presentarse al final del artículo de la siguiente manera: nombre(s) del(os) autor(es) comenzando por el(los) apellido(s); título de la publicación con cursivas (si se trata de un artículo, debe estar entrecomillado, seguido de coma y la preposición en con dos puntos y, enseguida, el título de la revista o libro donde apareció publicado, con cursivas); país de origen; editorial; lugar y año de edición; página(s) consultada(s). En el caso de las fuentes electrónicas (páginas web) se debe seguir el mismo orden que en las bibliográficas, pero al final se pondrá entre paréntesis DE (dirección electrónica), la fecha de consulta y la liga completa. Se tienen que omitir aquellas que se mencionen como notas a pie de página. Si se aplica la opción de incluir en cuerpo de texto la referencia de nombre de autor y año de la fuente consultada entre paréntesis, sí deben aparecer todas las referencias mencionadas.

Página electrónica: <http://rde.inegi.org.mx>

Editorial Guidelines and Policy

REALITY, DATA AND SPACE INTERNATIONAL JOURNAL OF STATISTICS AND GEOGRAPHY is a four-monthly publication that connects statistics and geographic official information with academic research in order to share knowledge among specialists and institutions with similar aims.

We will publish only original and unpublished articles related to the current use and appliance of statistical and geographical information at both national and international levels.

It is a technical-scientific and bilingual magazine, with articles previously peer-reviewed by specialists under a double-blind methodology with the following evaluation criteria: unpublished works, originality, information related to opportunity and current affairs, we expect clarity in the definition of aims and ideas stated, defined objectives coverage, accurate methodological structure and coherence between the information of the paper as well as its conclusions.

The result of the paper-assessment process is delivered by email, and it involves three possibilities: fully recommended (with slight modifications), recommended (on condition of suggested modifications) and not recommended (i.e. rejected). When there are two reports of approval, the author gets notified that his/her paper will be published and it is sent to a style editing process. When one report approves the paper for publication and another one rejects it, the author is requested to make some changes for the text to be published. If the text submitted receives two non-favourable reports, the author is notified that the text will not be published.

Index and Registers

- LATINDEX Catalogue (Online Regional Information System for Scientific Journals from Latin America, the Caribbean, Spain and Portugal).
- CLASE (Latin American Quotations in Humanities and Social Sciences)
- REDIB (Latin American Net of Innovation and Scientific Knowledge)

Publishing Guidelines

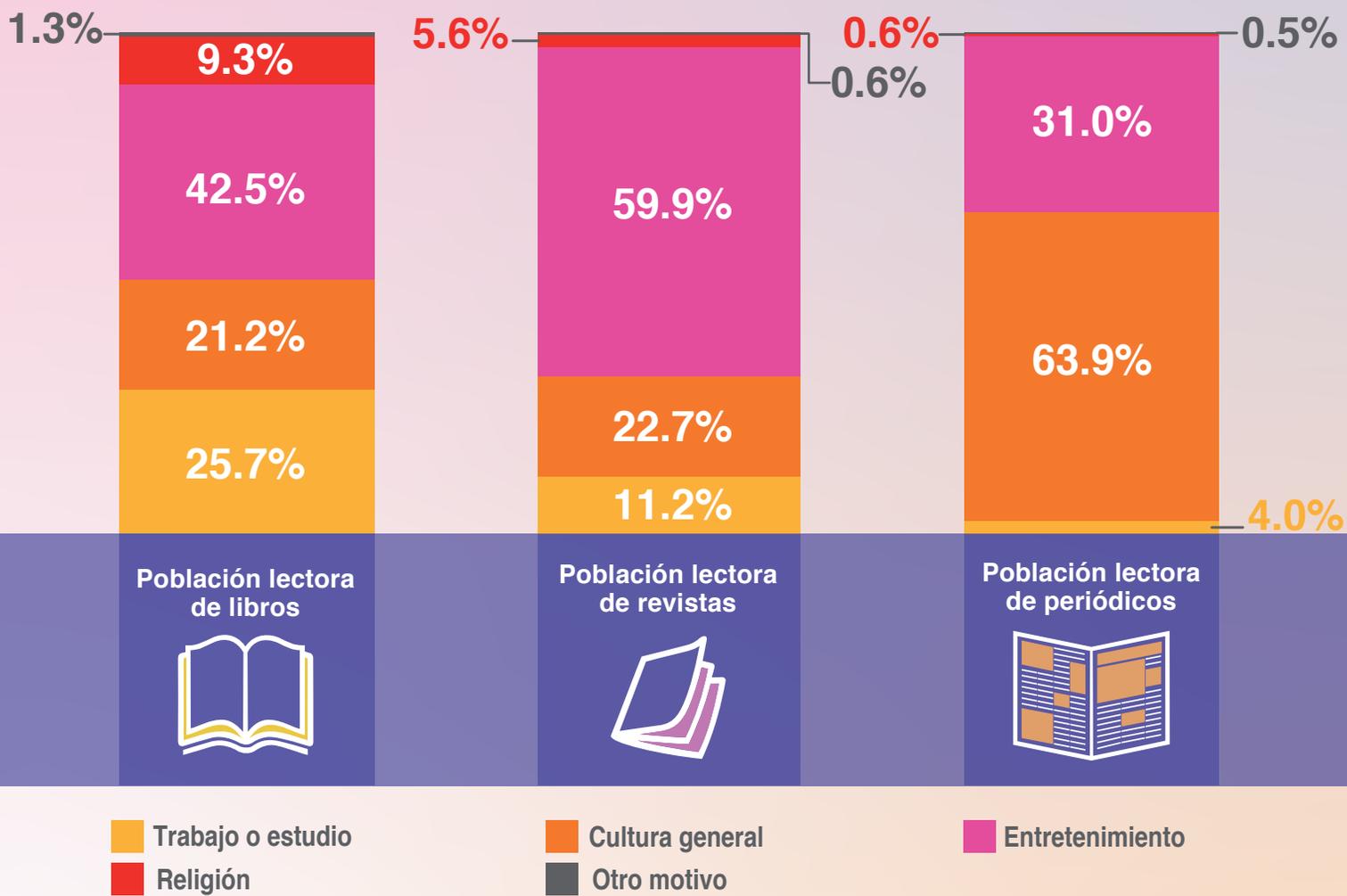
Articles will be published in Spanish or English: research, revision and scientific-spreading articles; methodologies; technical reports; short texts; book reviews; and bibliographical and statistical revisions, among others.

1. The article —or any other kind of text from those aforementioned— must be delivered with an attached letter addressed to the chief editor of Reality, Data and Space. International Statistics and Geography Magazine in which the text intended for publication will be submitted. There it must be stated that the text has not been published, and that it has not been submitted for publication in any other media. The names in full of the authors must be included, as well as their nationalities, adscription institutions, postal address, e-mail address, and telephone numbers. This must be addressed to MSc Virginia Abrin Batule, Virginia.abrin@inegi.org.mx (tel (+52) (55) 52.78.10.00, extension 1161).
2. The article must be submitted in an electronic version (a Microsoft Word file or a compatible one) with the following format: a) the text should not exceed the 20 pages of length; b) typography must be Helvetic, Arial or Times (12 points); and c) there should be a 1.5 line spacing in each paragraph. Additional material to the text will be delivered separately: a) images with a resolution of 300 ppp and no smaller than 17 cm width will be delivered in format JPG or TIF —please do not amplify images with lower resolution—. If the added materials are lines or maps, these must be delivered in vectorial format (EPS or Ai). If there are images in bits map, these must be embedded or attached with the name of the original file with which the link was created (keeping the resolution and size requirements above stated). As regards to photographs, these should not be inferior as 5 megapixels; b) mathematical expressions or formulae have to be created with the equations editor by Microsoft[™], but in case of using third-parties software, please attach a witness PDF in which the exact representation of mathematical formulae or expressions is contained; c) graphics must include the Excel file in which they were created or a legible image in the original JPG format in high resolution; and d) charts must be editable, and must not be inserted as images.
3. The text must include the following: the article's title (both in English and Spanish); the abstract of the article—both in English and Spanish (not longer than a 10-line paragraph); key words—both in English and Spanish (three as minimum and five as maximum); bibliography and other sources; as well as brief biographical sketches of the authors not exceeding a five-line paragraph each including nationalities, academic titles, main professional experiences, current work-related affiliations, and electronic addresses for the authors to be contacted.
4. Bibliographical references and other sources must be included at the end of the article in the following way: author's name (Surname first), and publication's title (in italics). If it is an article, the title must be in quotation marks followed by a comma and the preposition "en" with semicolon (in Spanish), then it should appear the title of the book or magazine in which the article was published (in italics); country of origin; publishing house, edition year, and consulted pages. As regards to electronic sources (web pages) the same order of the bibliographical references must be followed, but at the end the word "EA" (as for Electronic Address) ("DE" in Spanish) must be added within parenthesis followed by consultation date and the complete reference link. Those web links referred previously as footnotes, must be omitted in this section. However, if the name of the author and the year of the consulted source were included in the main body of the text within parenthesis, all these must be included as part of the bibliographical references.

Webpage: <http://rde.inegi.org.mx>



¿Qué nos motiva a leer?



Nota: información estadística sobre el comportamiento lector de la población de 18 años y más de edad que reside en áreas de 100 mil y más habitantes.

Fuente: INEGI. *Módulo sobre Lectura 2019.*

Conociendo México

01 800 111 46 34
www.inegi.org.mx
atencion.usuarios@inegi.org.mx





INEGI Informa



