

Inferencia bayesiana y tablas de mortalidad en México

Manuel Mendoza Ramírez, Alberto Contreras Cristán y Eduardo Gutiérrez Peña

Análisis de estadísticas del INEGI sobre residuos sólidos urbanos

José Paul Carrasco Escobar y José Luis Ángel Rodríguez Silva

Desigualdades entre entidades en materia de tecnologías de información y comunicación en México

Wilfrido Ruiz Ochoa

Percepción remota para la estimación de población en áreas geoestadísticas básicas

José Luis Silván Cárdenas, Jorge Alberto Montejano Escamilla y Mauricio Pablo Cervantes Salas

Revisión metodológica de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003, 2006 y 2011

José Antonio Mejía Guerra

Resolución I de la XIX Conferencia Internacional de Estadísticos del Trabajo y su impacto en la estadística laboral en México

Rodrigo Negrete Prieto y Tomás Ramírez Reynoso



Humedales en México

Son áreas de transición entre los sistemas acuáticos y terrestres; la convención internacional de RAMSAR los considera "...sistemas clave de sustento de la vida en el planeta, conjuntamente con las tierras agrícolas y los bosques". En general, desempeñan una función esencial en el mantenimiento de poblaciones de especies silvestres al servir de hábitat de fauna y flora variadas. Cerca de un tercio de las especies silvestres amenazadas, en peligro o raras identificadas viven en los humedales. Además de mitigar los efectos de las crecidas por lluvias extraordinarias, reducir la erosión y recarga de acuíferos, protegen a las poblaciones humanas contra las inundaciones y mantienen o mejoran la calidad del agua, entre otros beneficios.

En el país existen 142 humedales protegidos (sitios RAMSAR) con una superficie de poco más de 8.8 millones de hectáreas, las cuales están expuestas al incremento de la actividad agropecuaria, el crecimiento urbano y la desecación para cambio de uso del suelo.

El INEGI, por encargo de la Comisión Nacional del Agua (CONAGUA), realiza la actualización del *Inventario Nacional de Humedales de México escalas 1:250 000 y 1:50 000*.

Marismas nacionales en Nayarit. 13 de febrero de 2014. Landsat 8.

Conociendo México

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

 INEGI Informa

 @INEGI_INFORMA

 **INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA**

Contenido

Inferencia bayesiana y tablas de mortalidad en México Manuel Mendoza Ramírez, Alberto Contreras Cristán y Eduardo Gutiérrez Peña	4
Análisis de estadísticas del INEGI sobre residuos sólidos urbanos José Paul Carrasco Escobar y José Luis Ángel Rodríguez Silva	18
Desigualdades entre entidades en materia de tecnologías de información y comunicación en México Wilfrido Ruiz Ochoa	36
Percepción remota para la estimación de población en áreas geoestadísticas básicas José Luis Silván Cárdenas, Jorge Alberto Montejano Escamilla y Mauricio Pablo Cervantes Salas	50
Revisión metodológica de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003, 2006 y 2011 José Antonio Mejía Guerra	72
Resolución I de la XIX Conferencia Internacional de Estadísticos del Trabajo y su impacto en la estadística laboral en México Rodrigo Negrete Prieto y Tomás Ramírez Reynoso	92
Colaboran en este número	132

INSTITUTO NACIONAL DE ESTADÍSTICA Y GEOGRAFÍA

Presidente del Instituto

Eduardo Sojo Garza-Aldape

Vicepresidentes

Enrique de Alba Guerra

Mario Palma Rojo

Rolando Ocampo Alcántar

Félix Vélez Fernández Varela

Dirección General de Estadísticas Sociodemográficas

Miguel Juan Cervera Flores

Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia

Adrián Franco Barrios

Dirección General de Estadísticas Económicas

José Arturo Blancas Espejo

Dirección General de Geografía y Medio Ambiente

Carlos Agustín Guerrero Elemen

Dirección General de Integración, Análisis e Investigación

Enrique Jesús Ordaz López

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

Norberto de Jesús Roque Díaz de León

Dirección General de Vinculación y Servicio Público de Información

Alberto Manuel Ortega y Venzor

Dirección General de Administración

Froylán Rolando Hernández Lara

Contraloría Interna

Marcos Benerice González Tejeda

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

Departamento de Diseño Editorial / INEGI

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 Plataforma Open Access de Revistas Científicas Electrónicas Españolas y Latinoamericanas *e-Revist@s*.

REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, Vol. 6, Núm. 1, enero-abril 2015, es una publicación electrónica cuatrimestral editada por el Instituto Nacional de Estadística y Geografía (INEGI), Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276, Aguascalientes, Aguascalientes, entre la calle INEGI y Avenida del Lago, Avenida Paseo de las Garzas, México. Tel. (55) 52781069. Correo Electrónico: rde@inegi.org.mx. Página electrónica: http://www.inegi.org.mx/RDE/rde_14/rde_14.html

Editor responsable: Enrique Jesús Ordaz López. Reservas de Derechos al Uso Exclusivo No. 04-2014-101310563200-203, ISSN: 2395-8537, ambos otorgados por el Instituto Nacional del Derecho de Autor. Responsable de la última actualización de este Número, Dirección de Vinculación con Instituciones Académicas, M. en C. Virginia Abrín Batule, Avenida Héroe de Nacozari Sur 2301, Fraccionamiento Jardines del Parque, 20276 Aguascalientes, Aguascalientes, entre la calle INEGI y, Avenida del Lago, Avenida Paseo de las Garzas, México, correo electrónico: virginia.abrin@inegi.org.mx, fecha de última modificación: abril de 2015.

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. Así mismo, 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.

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

CONSEJO EDITORIAL

Enrique de Alba Guerra

Presidente del Consejo

Fernando Cortés Cáceres

Profesor Emérito de FLACSO
PUED de la UNAM

Gerardo Bocco Verdinelli

Universidad Nacional Autónoma de México

Ignacio Méndez Ramírez

Universidad Nacional Autónoma de México

Juan Carlos Chávez Martín del Campo

Banco de México

José Ramón Narro Robles

Universidad Nacional Autónoma de México

Lidia Bratanova

UNECE Statistical Division

Manuel Ordorica Mellado

El Colegio de México, AC

María del Carmen Reyes Guerrero

Centro de Investigación en Geografía y
Geomática "Ing. Jorge L. Tamayo", AC

José Antonio de la Peña Mena

Centro de Investigación en Matemáticas, AC

Rodolfo de la Torre García

Programa de las Naciones Unidas
para el Desarrollo

Tonatiuh Guillén López

El Colegio de la Frontera Norte, AC

Víctor Manuel Guerrero Guzmán

Instituto Tecnológico Autónomo de México

Walter Radermacher

Statistical Office of the European Communities

Editorial

Este año damos inicio al volumen 6 de REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, publicación que ha estado abierta a temas de todas las disciplinas interesadas en la información estadística y geográfica que genera el INEGI, a partir de la cual resultan interesantes artículos evaluados por una mesa crítica de especialistas.

Entre ellos contamos, desde el inicio de la Revista, y sus antecesoras, con la colaboración incondicional y permanente del doctor Eduardo Rodríguez-Oreggia y Román, a quien queremos rendir con estas líneas un sentido homenaje a pocos meses de su partida.

El primer artículo, *Inferencia bayesiana y tablas de mortalidad en México*, describe el tratamiento bayesiano que subyace en la producción de las tablas de mortalidad que publicó la autoridad regulatoria mexicana en el 2000 y que son legalmente obligatorias para el cálculo de las reservas de las compañías de seguros de vida que operan en el país.

Análisis de estadísticas del INEGI sobre residuos sólidos urbanos presenta las características principales de este fenómeno y la aplicación de las técnicas estadísticas de análisis de componentes principales y conglomerados, así como una interpretación de los resultados derivados, mismos que pueden servir como guía en el proceso de toma de decisiones en políticas públicas en la materia.

Más adelante, en el artículo *Desigualdades entre entidades en materia de tecnologías de información y comunicación en México* se propone y construye un índice estatal de desarrollo de éstas como instrumento para el seguimiento a nivel subnacional de la estrategia digital nacional.

Percepción remota para la estimación de población en áreas geoestadísticas básicas es el siguiente trabajo en el cual se presentan los resultados de un estudio para evaluar la posibilidad de realizar esta labor en tres delegaciones del Distrito Federal mediante datos de altimetría LIDAR, imágenes multiespectrales e información del catastro.

También, en este número se da a conocer un análisis comparativo general sobre los esfuerzos de recolección de estadísticas de violencia contra la mujer que se han hecho en el INEGI a través del artículo titulado *Revisión metodológica de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003, 2006 y 2011*.

Para concluir, se presenta *Resolución I de la XIX Conferencia Internacional de Estadísticos del Trabajo y su impacto en la estadística laboral en México*, un ensayo que muestra los cambios en la estadística del trabajo en México a partir de las recomendaciones internacionales más recientes dirigidas al diseño, producción y difusión de la información proveniente de las encuestas de fuerza laboral (encuestas de empleo) en el mundo entero.

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

Inferencia bayesiana y tablas de mortalidad en México

Manuel Mendoza Ramírez, Alberto Contreras Cristán
y Eduardo Gutiérrez Peña

El estudio de la mortalidad, en particular la producción de tablas de mortalidad, juega un papel central en la industria de los seguros. Durante un tiempo, la estimación de las tasas de mortalidad de los asegurados se abordó ignorando la incertidumbre implícita en el problema. En fecha reciente, sin embargo, el análisis estadístico ha cobrado relevancia en el tema. En este artículo se describe el tratamiento bayesiano que subyace en la producción de las tablas de mortalidad que publicó la autoridad regulatoria mexicana en el 2000 y que son legalmente obligatorias para el cálculo de las reservas de las compañías de seguros de vida que operan en el país.

Palabras clave: análisis predictivo, inferencia bayesiana, modelos lineales, tablas de mortalidad.

Recibido: 1 de septiembre de 2014

Aceptado: 8 de diciembre de 2014

Nota: los autores agradecen el apoyo del Sistema Nacional de Investigadores y de la Asociación Mexicana de Cultura, AC. La investigación se hizo posible gracias al Programa UNAM-DGAPA-PAPIIT, Proyecto IN106114.

Mortality tables play a key role in life insurance. However, for a long time, the estimation of mortality rates has been conducted without reference to the statistical nature of the problem. Nowadays, an appropriate risk-management strategy requires all sources of uncertainty to be considered in the evaluation of these and other financial systems. Thus, statistical models have become increasingly relevant. In this paper, the Bayesian analysis of linear models is described as the tool that the authorities in Mexico used to produce the mandatory mortality tables currently in use by the insurance industry. These ideas are illustrated with data from the Mexican insurance sector.

Key words: Bayesian inference, linear models, mortality tables, predictive analysis.



Day Of The Dead Festival Held In Oaxaca, Mexico/Richard Ellis/Getty Images

1. Introducción

El estudio de la mortalidad es relevante en distintas disciplinas, la Medicina y la Demografía son sólo dos ejemplos; otra área donde juega un papel fundamental es la de las Ciencias Actuariales, donde su correcta medición es crucial para mantener la solvencia económica de los sistemas de seguros de vida y de pensiones. Por un largo tiempo, los estudios actuariales ignoraron la incertidumbre implícita en las inferencias sobre la mortalidad.

Así, la producción de tablas de mortalidad —arreglos de probabilidades de muerte (en el

plazo de un año) como función de la edad— se enfrentaba como un problema determinístico en el que la tabla se obtenía eliminando las irregularidades de las tasas de mortalidad observadas mediante un procedimiento de suavizamiento, habitualmente conocido como graduación. El resultado es una tabla que corresponde a una curva suave que describe la tendencia general de las tasas respecto a la edad sin considerar la variabilidad respecto a la curva.

Para este proceso de suavizamiento se han utilizado principalmente dos enfoques. El primero, no paramétrico, obtiene los valores de la tabla optimizando una función que penaliza tanto su discre-

pancia respecto a las tasas observadas como su falta de suavidad. El ejemplo más conocido de este tipo de graduación es el método de Whittaker-Henderson (ver London, 1985). Otra alternativa considera, de inicio, que la tabla proviene de una curva que pertenece a una familia paramétrica cuya suavidad está determinada *a priori* y la minimización tiene lugar sólo sobre la componente de discrepancia con los datos observados (Carriere, 1992, por ejemplo). En conexión con esta segunda variante, se han establecido algunos modelos célebres, como la familia Gompertz-Makeham, a partir de la cual se han desarrollado otras familias más complejas y flexibles (ver Heligman y Pollard, 1980, por ejemplo). En cualquier caso, con estos procedimientos, las tasas suavizadas o ajustadas suelen utilizarse como las *verdaderas* para efectos del cálculo de primas y reservas, sin considerar incertidumbre alguna.

Esta situación ha evolucionado de manera paulatina. En una primera fase, diversos autores exploraron la relación de la graduación determinística con algunas técnicas de inferencia estadística, en especial desde una perspectiva bayesiana; por ejemplo, Kimeldorf y Jones (1967) suponen que el vector de tasas observadas sigue una distribución normal multivariada cuya media es el vector de tasas *verdaderas*. Adoptan una matriz de covarianzas diagonal conocida y una distribución inicial (*a priori*) normal, centrada en una tabla graduada previa y cuya matriz de covarianzas establece una estructura de correlación específica. Como estimador, se propone la moda *a posteriori*. De esta manera, la maximización de la densidad final es equivalente a la minimización de una forma cuadrática que puede identificarse con la función objetivo de un procedimiento de graduación. En particular, la estructura de correlación *a priori* correspondería con un requerimiento de suavización. Esta aproximación incluye distintas formas de graduación como casos particulares, especialmente el muy conocido método de Whittaker (1923). La idea después fue generalizada en distintas direcciones; sin embargo, por algunos años, el problema siguió abordándose como determinístico.

En fecha más reciente, el enfoque moderno de riesgos financieros ha contribuido, al menos parcialmente, para que, a partir de la década de los 80 del siglo pasado, se consideren todas las fuentes de incertidumbre que pueden tener impacto en los resultados de los sistemas de seguros. Conceptos como *valor en riesgo* y *capital de solvencia*, entre otros, dependen de un análisis estadístico completo que incluya, en particular, las desviaciones del monto de las reclamaciones respecto a su monto esperado. De esta manera, un creciente número de autores ha recurrido a los métodos estadísticos con este propósito. En este sentido, Forfar *et al.* (1988), Renshaw (1991) y Haberman y Renshaw (1996) son pioneros.

En el ámbito del sector financiero, es interesante observar que las agencias gubernamentales que supervisan a las empresas con el fin de garantizar su solvencia tienen impacto en el desarrollo técnico del sector. En el caso de los seguros de vida, con independencia de cómo fijen las compañías sus precios, el supervisor establece una tabla de mortalidad de uso obligatorio para el cálculo de las reservas. En México, la Comisión Nacional de Seguros y Fianzas (CNSF) produce tablas a partir de la información que obtiene de todas las empresas que operan en el país y, para contribuir a la solvencia, sobrestima las probabilidades de muerte. Las tablas se actualizan cuando se observan cambios relevantes en los patrones de la mortalidad y es conveniente notar que tienen un carácter predictivo, ya que se diseñan para proteger al sector frente a desviaciones de la mortalidad *futura* respecto a las tasas esperadas.

En este artículo se describe el proceso seguido en los años recientes para que la CNSF contase con tablas de mortalidad obligatorias para los seguros de vida en México sustentadas en un análisis estadístico. La sección 2 describe el modelo que se utilizó para producir la tabla que la CNSF publicó en el 2000. En la 3 se aborda el proceso de revisión y actualización de las tablas que implica el estudio del fenómeno en una escala diferente de la empleada en el ajuste del modelo. Por último, en la 4 se incluyen algunos comentarios finales y se discuten posibles extensiones.

2. Análisis de datos y propuesta de modelo

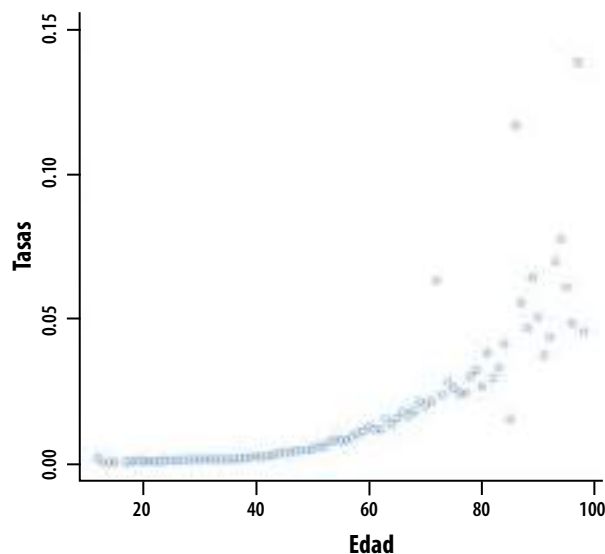
Desde que inició operaciones, hace más de 100 años, el sector de seguros de vida en México utilizó tablas de mortalidad adaptadas de otros países. Sólo en 1967, la autoridad publicó la primera de origen mexicano con información propia, de 1962 a 1967 (*Experiencia mexicana 62-67*). Ésta se usó hasta 1989, cuando fue reemplazada con la *Experiencia mexicana 82-89* (EM 82-89). Las dos tablas se produjeron con técnicas de graduación determinística ajustando curvas de Gompertz-Makeham a las tasas de mortalidad agregadas para todos los años observados. A partir de este resultado se obtuvieron las tablas *básicas*, que describen la tendencia de la mortalidad como función de la edad, y las *modificadas*, que se obtienen a partir de las primeras e incluyen un margen de protección (sobrestimación de las tasas de mortalidad) y, por lo tanto, conducen a un mayor nivel de reservas para hacer frente a la siniestralidad. En 1999 se consideró necesario actualizarlas y, para este fin, se utilizó la información disponible durante el periodo 1991-1998. El procedimiento, en esa ocasión, fue enteramente estadístico y se describe en el resto de esta sección.

2.1 Información 1982-1989

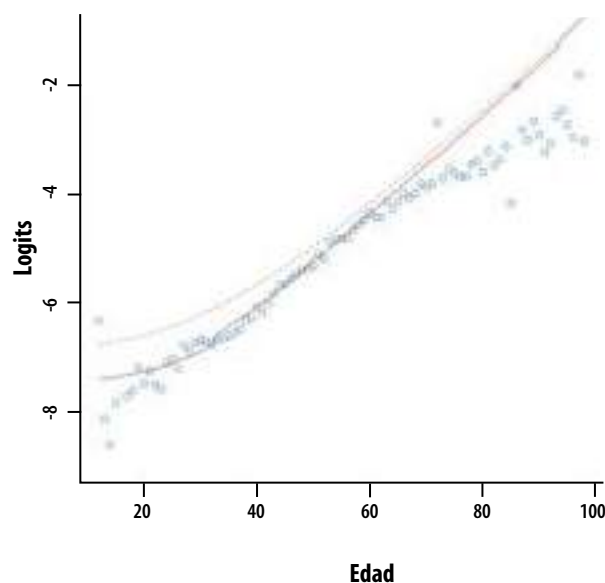
Para contar con un referente contra el cual comparar los resultados del tratamiento propuesto, éste se ilustrará con los datos de 1982 a 1989. En la gráfica 1 se exhiben las tasas de mortalidad por edad observadas en ese periodo. La tasa q_x corresponde a la edad x , y se calcula como el cociente del número de muertes correspondientes a esa edad, D_x , respecto al número de personas expuestas, E_x , de entre las cuales ocurren esas muertes. De esta forma, $q_x = D_x / E_x$. La base de datos contiene, en total, información proveniente de más de 6.6 millones de personas expuestas, distribuida en el intervalo de edades 12-99 años y sólo se cuenta con una tasa por cada edad ya que tanto las muertes como los expuestos fueron agregados sobre los ocho años. En la gráfica 2 se

pueden observar las tasas y las tablas, básica y modificada, en la escala logit: $Y = \ln \{q_x / (1 - q_x)\}$. En la gráfica son claros el comportamiento lineal de los logits y el hecho de que las dos tablas sobrestiman la tendencia de los datos.

Gráfica 1
Tasas de mortalidad observadas (1982-1989)



Gráfica 2
Tasas, tablas básica (sólida) y modificada (trazos); escala logit (1982-1989)



2.2 El modelo

El análisis estadístico para la producción de las tablas descansa en la adopción de un modelo probabilístico que describa apropiadamente la relación entre las tasas y la edad, incorporando la variabilidad inherente en el fenómeno. Para tal fin, es necesario seleccionar un modelo de probabilidad condicional F tal que $Y \sim F(\cdot | X)$ con Y la tasa de mortalidad observada (o una transformación de la misma), mientras que X representa la edad correspondiente. Como ya se ha comentado, Kimeldorf y Jones (1967) suponen F normal con varianza conocida y Y la tasa observada, sin transformación alguna. Otros autores, como Renshaw (1991) y Haberman y Renshaw (1996), utilizan un modelo binomial o Poisson para el número de muertes, y la relación con la edad se describe vía un modelo lineal generalizado.

Sin duda, existe una variedad de posibilidades; sin embargo, con la idea de mantener el análisis lo más simple posible y dada la relación que sugiere la gráfica 2, el análisis realizado para la CNSF en 1999 considera el ajuste bayesiano de un modelo de regresión lineal simple para describir los logits en términos de la edad. El análisis exploratorio de la base de datos original y una descripción preliminar del análisis se pueden encontrar en una serie de reportes técnicos (Mendoza *et al.*, 1999a, 1999b y 2000). El modelo que se considera establece entonces que, condicional en la edad X , Y sigue una distribución normal $N(Y | \beta_0 + \beta_1 X, \sigma^2)$. Incorporando el supuesto usual de independencia, la distribución conjunta de los logits observados está, en general, dada por:

$$Y | X, \beta, \sigma^2 \sim N(\mathbf{Z}\beta, \sigma^2 \mathbf{I}) \quad (1)$$

donde la dimensión de Y es n , el número de edades diferentes en los datos; \mathbf{Z} es una matriz ($n \times p$) de funciones conocidas de X , y β, σ^2 son dos parámetros desconocidos. En el caso de la tabla 1982-1989, los datos incluyen edades de 12 a 98 años con la excepción de la edad 16, así que $n = 86$ y \mathbf{Z} incluye sólo un término constante y a la propia edad X .

Antes de proceder al ajuste del modelo, es conveniente tener en cuenta algunas consideraciones.

Primero, la idea de una tabla modificada es que incluya un margen de protección contra las variaciones *anuales* de las tasas respecto a su valor estimado. Con los datos del periodo 1982-1989, estas variaciones no son observables en virtud de que las tasas disponibles agregan la información de ocho años. De cualquier manera, la propuesta para seleccionar la tabla modificada se ilustrará con estos datos para comparar los resultados con la tabla 82-89 correspondiente. Un segundo tema es el de la varianza. El modelo supone que este parámetro es constante a lo largo de las edades; sin embargo, en las gráficas 1 y 2 se puede observar que la variabilidad es mayor para las edades menores a 18 años y superiores a 70 años. De los 6.6 millones de expuestos disponibles, los datos correspondientes a estas edades extremas representan únicamente 2.73%; de esta forma, para ajustar el modelo, se podrían eliminar esos datos y proceder con las 56 edades restantes. Otra alternativa sería un ajuste por mínimos cuadrados ponderados que incorpore las diferencias en las varianzas; sin embargo, si se utilizan todos los datos y se mantiene el supuesto de homoscedasticidad, el resultado es más simple y conservador. Con esta opción, la varianza que se supone común subestima la variabilidad en las edades extremas, pero la sobrestima a lo largo de las edades que van de 18 a 70 años, es decir, la varianza se sobrestima para los logits de las tasas que corresponden a 97.27% de los expuestos. El efecto neto es que las bandas de pronóstico son más amplias para esas edades y, en consecuencia, la tabla modificada tendrá un mayor margen de protección en esa zona que, por su volumen de siniestros potenciales, compensa de manera sobrada la subestimación en 2.73% de la población restante.

2.3 Análisis bayesiano del modelo

Para el análisis del modelo, se adopta una distribución inicial mínimo informativa $p(\beta, \sigma^2) \propto 1/\sigma^2$ que conduce a la distribución *a posteriori*:

$$p(\beta | \mathbf{Y}, \mathbf{Z}, \sigma^2) \sim \text{Normal}(\beta | \hat{\beta}, \sigma^2 (\mathbf{Z}^t \mathbf{Z})^{-1})$$

$$p(\sigma^2 | \mathbf{Y}, \mathbf{Z}) = \text{IG}(\sigma^2 | \frac{1}{2}(n-p), \frac{n}{2} \hat{\sigma}^2),$$

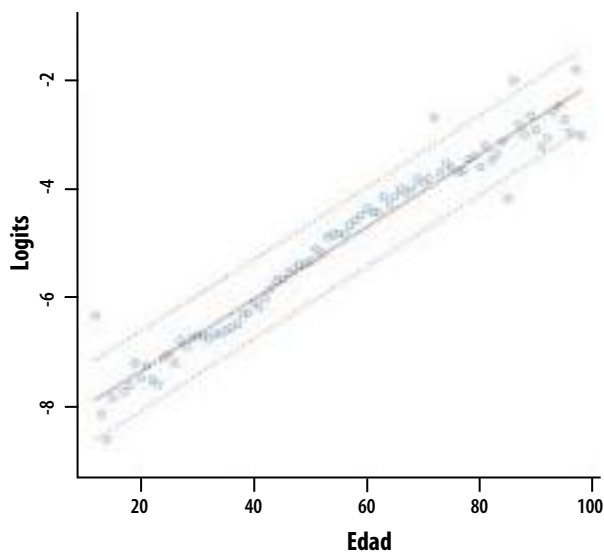
donde $IG(\cdot|\alpha, \eta)$ es una distribución Gamma inversa con parámetros α y η , mientras que $\hat{\beta}$ y $\hat{\sigma}^2$ denotan los estimadores de máxima verosimilitud de los parámetros (ver, por ejemplo, Bernardo y Smith, 2000). Ahora, si X_f es un vector de edades $m \times 1$ para el cual se pretende obtener pronósticos del correspondiente vector de logits Y_f , la distribución predictiva *a posteriori* está dada por:

$$p(Y_f | X_f, Y, Z) = St(Y_f | Z_f \hat{\beta}, \hat{\sigma}^2 \{Z_f (Z^t Z)^{-1} Z_f^t + I\}, n - p), \quad (2)$$

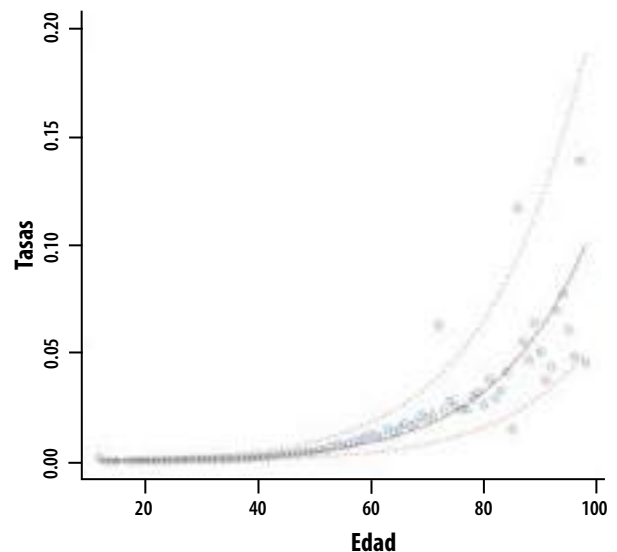
donde Z_f es la matriz correspondiente a X_f , $\hat{\sigma}^2$ es el estimador insesgado usual de σ^2 y $St(\cdot|\mu, \Sigma, \nu)$ representa una distribución *t* de Student con parámetro de localización μ , matriz de escala Σ y ν grados de libertad. En el caso del modelo de regresión lineal simple, $Z_{i1}=1$ y $Z_{i2}=X_i$, ($i=1, \dots, n$). Además, $(n-p)=84$ de manera que, para propósitos prácticos, esta distribución predictiva puede aproximarse con una normal multivariada. La gráfica 3 muestra los logits con la función media y la banda predictiva de 0.95 de probabilidad. Aplicando la transformación inversa, $q = \exp(Y) / (1 + \exp(Y))$, se obtienen las curvas correspondientes en la escala original de las tasas. Éstas aparecen en la gráfica 4 y, en virtud de que la transformación es creciente,

conservan el respectivo contenido de probabilidad. Más aún, la distribución predictiva de los logits es simétrica, de manera que su media coincide con su mediana; por lo tanto, en la escala de las tasas, la tabla básica se puede definir como la curva mediana correspondiente. Las curvas en la gráfica 4

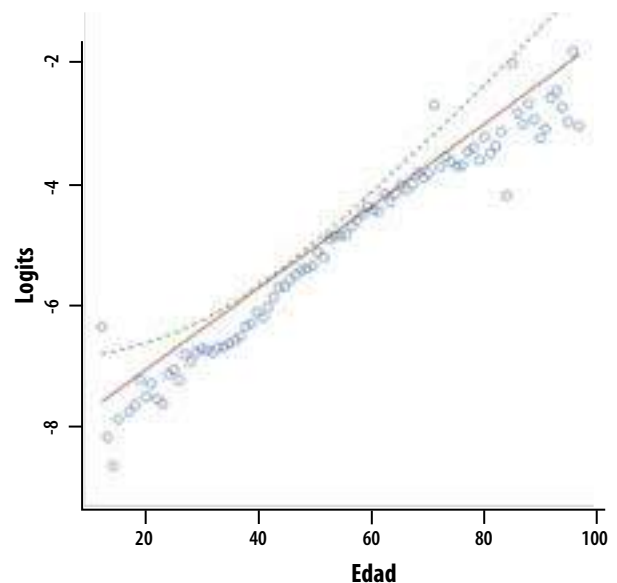
Gráfica 3
Media predictiva y banda de probabilidad 0.95



Gráfica 4
Modelo ajustado a las tasas y banda de probabilidad 0.95



Gráfica 5
Tabla modificada EM 82-89 (puntos) y curva de cuantiles 0.80 (sólida)



corresponden a tres cuantiles (0.025, 0.500 y 0.975, en ese orden). A partir del ajuste, es natural proponer, como nueva tabla modificada, la curva con los cuantiles $(1-\alpha)$ para un valor razonablemente pequeño de α . Para fines de comparación, en las gráficas 5 y 6 aparecen las curvas de los cuantiles 0.80 y 0.99 del modelo con la tabla modificada EM 82-89.

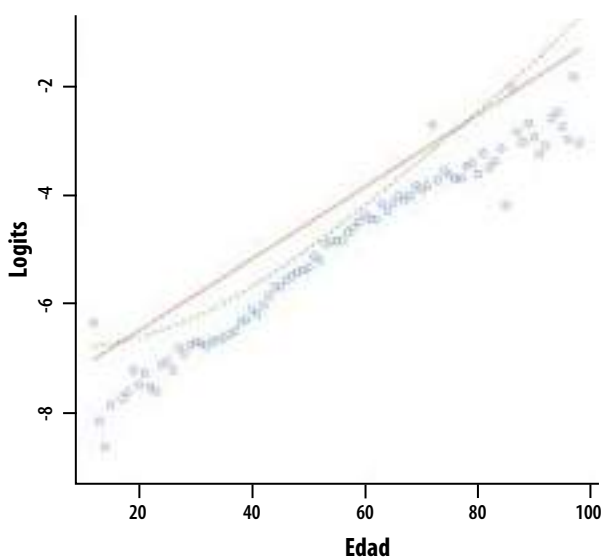
Un primer aspecto que puede destacarse es que la EM 82-89 no ofrece un margen de protección uniforme, en el sentido probabilístico, para las distintas edades y, en todo caso, la probabilidad de excesos es menor a 0.2 para todas las edades. Si se compara con la curva de cuantiles 0.99, la probabilidad de observar excesos es mayor a 0.01 para edades entre 20 y 70 años y menor en los extremos, es decir, ofrece más protección frente a desviaciones de la mortalidad en los grupos de jóvenes y adultos mayores en comparación con el grupo de adultos entre 20 y 70 años que, por otra parte, constituyen la mayor parte de la población asegurada (aproximadamente 97%).

Otro tema es la selección del valor apropiado de α . Las recomendaciones internacionales en materia de administración de riesgos financieros habi-

tualmente sugieren valores entre 0.05 y 0.01. En este caso, sin embargo, este requerimiento puede resultar excesivo porque se impone de manera simultánea en todas y cada una de las edades. En la realidad, si el límite previsto por la tabla modificada se sobrepasa en algunas edades, el efecto en la solvencia del sistema puede ser irrelevante si se compensa con una disminución de la mortalidad en el resto de la población asegurada. En particular, esto puede ocurrir si el grupo de edades para las que se presenta el exceso constituye sólo una pequeña fracción del total de asegurados. Para determinar un valor de α que atienda a las recomendaciones internacionales, pero que efectivamente mida el riesgo de insolvencia por exceso *global* de la mortalidad, se propone elegir la curva de cuantiles a partir de la distribución predictiva del número *total* de muertes para una población asegurada específica. La idea es seleccionar la tabla modificada como la curva de probabilidades de muerte (cuantiles) que produzca un número esperado de muertes que, en la distribución predictiva del total de muertes, corresponda al cuantil del nivel recomendado. La población asegurada puede ser la misma que dio origen a los datos para el ajuste o cualquier otra; por ejemplo, el ajuste podría emplear toda la información del sector, pero la tabla se podría determinar para cada compañía de acuerdo con su perfil de exposición.

Gráfica 6

Tabla modificada EM 82-89 (puntos) y curva de cuantiles 0.99 (sólida)



2.4 Distribución predictiva del número total de muertes

Sea $\mathbf{E}_p^t = (E_{x_1}, E_{x_2}, \dots, E_{x_r})$ el vector con el perfil de exposición. Si $\mathbf{X}_p^t = (X_1, X_2, \dots, X_r)$ es el vector de edades, \mathbf{Y}_p , el de logits y \mathbf{q}_p , el de tasas, todos correspondientes a este perfil, a partir de la distribución predictiva $p(\mathbf{Y}_p | \mathbf{X}_p, \mathbf{Y}, \mathbf{Z})$ en (2) se obtiene $p(\mathbf{q}_p | \mathbf{X}_p, \mathbf{Y}, \mathbf{Z})$ y también se puede derivar $p(S_p | \mathbf{X}_p, \mathbf{Y}, \mathbf{Z})$, la distribución predictiva de $S_p = \mathbf{E}_p^t \mathbf{q}_p$ (el número total de muertes asociadas al perfil de exposición \mathbf{E}_p en conjunto con el vector de tasas de mortalidad \mathbf{q}_p). En la práctica, es fácil simular \mathbf{Y}_p tantas veces como se requiera para obtener una muestra arbitrariamente gran-

de de S_p , que se puede utilizar para aproximar la distribución predictiva de interés. El resultado, con los datos del periodo 1982-1989, cuando se utilizan los datos empleados en el ajuste, se exhibe en la gráfica 7. Ahí, $E(S_p) = 23\ 719$, que sólo difiere en 0.83% del número de muertes observado en esa población (23 918). Si la estimación de este valor se produce con la mediana (23 663), el error relativo resulta 1.11%; en este contexto, el modelo ajustado a las tasas de mortalidad observadas está bien calibrado en el sentido de que aproxima razonablemente el número total de muertes. El cuadro 1 muestra algunos cuantiles en la cola derecha de esta distribución predictiva.

Considerando de nuevo el problema de selección de la curva de cuantiles $q_{(1-\alpha)}$ para la tabla modificada, el cuadro 2 exhibe el valor del producto $E_p^t q_{(1-\alpha)}$ para algunos valores de α . Es claro que, para alcanzar el cuantil 0.95 de la distribución de S_p , es suficiente seleccionar $(1-\alpha) = 0.70$ y que, con la curva de cuantiles 0.74, el número de muertes ya excede el cuantil 0.995 de S_p . En con-

Cuadro 1

Cuantiles de la distribución predictiva de S_p (1982-1989)

Probabilidad	0.900	0.950	0.990	0.995	0.999
Cuantil	25 729	26 345	27 564	28 026	29 012

Cuadro 2

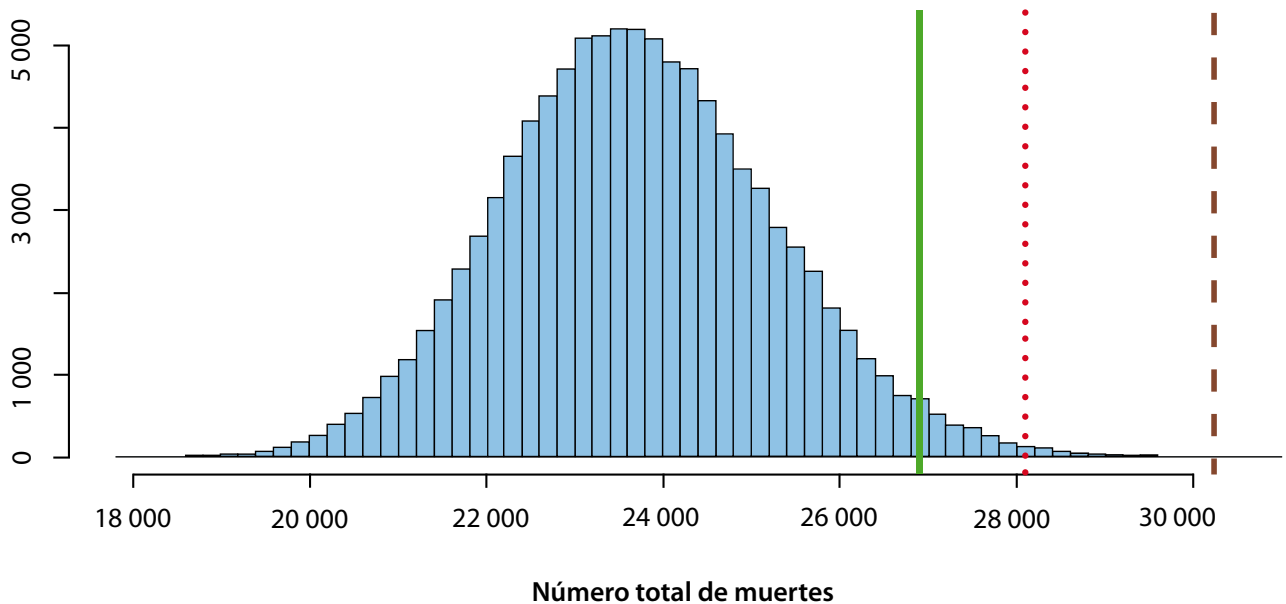
Número total de muertes para algunas curvas de cuantiles (1982-1989)

$1-\alpha$	0.60	0.70	0.74	0.80	0.90
Muertes	24 348	26 902	28 104	30 228	35 524

clusión, es razonable proponer $q_{(1-\alpha)}$ como tabla modificada con $\alpha \in (0.25, 0.30)$. Para fines comparativos, es posible comprobar que si la tabla EM (82-89) se trata como una curva de cuantiles, el número esperado de muertes que produce es 37 180, un valor por completo fuera de rango.

Gráfica 7

Distribución de S_p y valores para las curvas de cuantiles 0.7 (sólida), 0.74 (puntos) y 0.8 (trazos)



2.5 Resultados para los datos 1991-1998

A diferencia de los datos del periodo 1982-1989, la información de la década de los 90 se colectó en forma desagregada; por lo tanto, fue posible calcular las tasas observadas por separado para cada uno de los ocho años disponibles. Después de un primer análisis exploratorio, se determinó que no toda la información era útil. En algunos años, las tasas correspondientes a las edades mayores a 75 años eran muy bajas y, presumiblemente, producto de un registro incorrecto. Por último, el estudio se llevó a cabo con 540 observaciones en lugar de las $88 \times 8 = 704$ de inicio consideradas. La distribución de estos 540 datos se presenta en el cuadro 3.

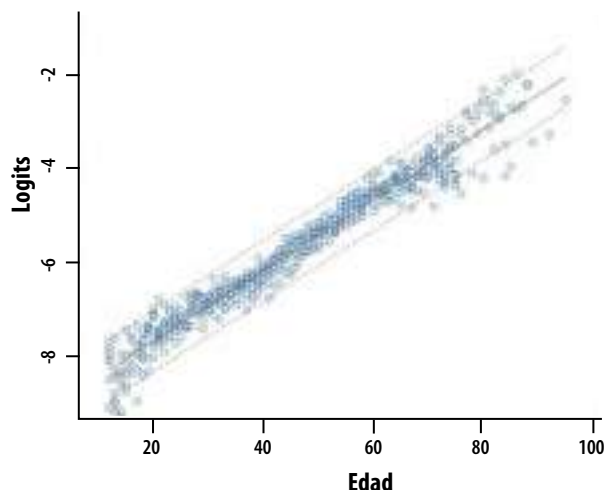
Cuadro 3
Número de observaciones para cada grupo de edad en el periodo 1991-1998

	1991	1992	1993	1994	1995	1996	1997	1998
12-30	17	17	19	19	18	18	18	18
31-45	15	15	15	15	15	15	15	15
46-60	15	15	15	15	15	15	15	15
61-75	15	15	15	15	15	15	15	15
76-99	0	0	13	0	0	0	13	10
	62	62	77	64	63	63	76	73

En este caso, los datos de 1991 a 1998 incluyen 21 746 231 expuestos y 54 564 muertes. Los logits se presentan en la gráfica 8. Como en el caso de los datos previos, es clara una tendencia lineal y el coeficiente de determinación es 0.96. Por otra parte, la variabilidad de nuevo es mayor para las tasas menores a 20 y mayores a 75 años, aunque este fenómeno se atenúa en este caso en virtud de la variabilidad entre años. Ante esta condición, y como se discutió en la sección 2.2, se adoptó una varianza constante como medida conservadora. El modelo se ajustó y los resultados se

emplearon, con el perfil de exposición de todos los datos, para obtener la distribución predictiva del número total de muertes correspondientes a este perfil. El resultado se muestra en la gráfica 9.

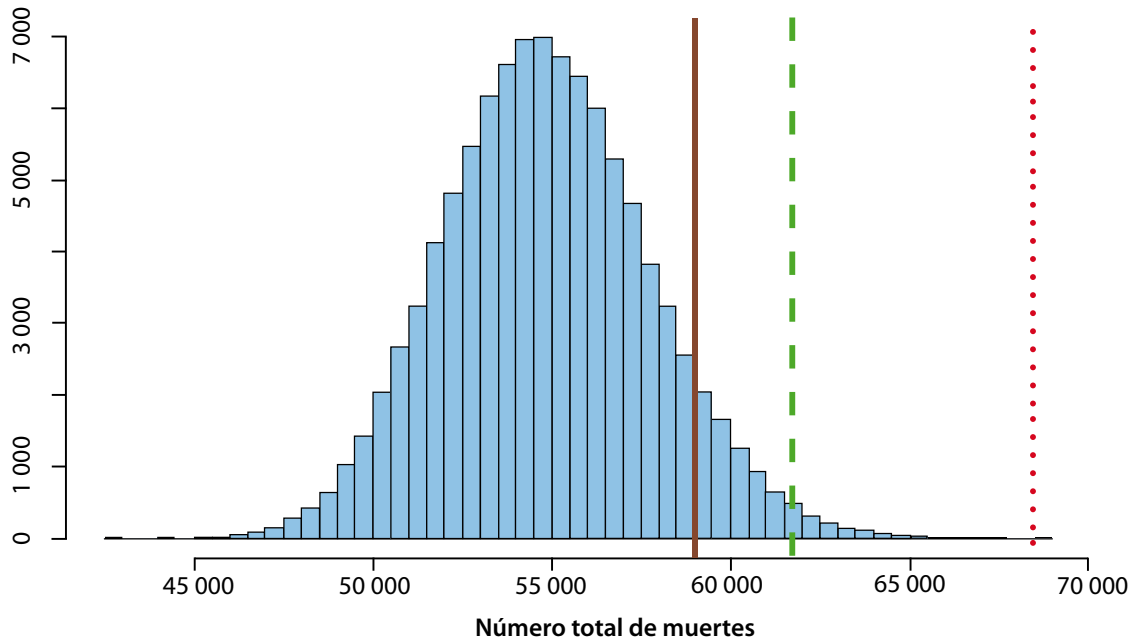
Gráfica 8
Modelo ajustado y banda predictiva 0.95 (datos 1991-1998)



La media y la mediana de esta distribución resultan 54 846 y 54 766. Si se comparan con el número observado de muertes (54 564), se tiene una sobrestimación con errores relativos de 0.5 y 0.3%, en ese orden. Adicionalmente, del modelo se sigue que S_p se encuentra en el intervalo (49 411, 60 750) con probabilidad 0.95. El cuadro 4 muestra algunos cuantiles de S_p . Para determinar el nivel de protección que ofrecen las curvas de cuantiles del modelo, el total de muertes ($E_p^t q_{(1-\alpha)}$) se calculó para algunos valores seleccionados de $(1-\alpha)$. Los resultados se muestran en el cuadro 5. El análisis conjunto de los cuadros 4 y 5 sugiere que cualquier curva con $(1-\alpha) > 0.7$ corresponde a un número de muertes que las observaciones futuras sólo rebasarían con una probabilidad menor a 0.05. De hecho, la CNSF adoptó como tabla obligatoria la curva de cuantiles 0.80, que reduce la probabilidad de exceso por debajo de 0.001. Esta tabla se denominó CNSF 2000-I y se publicó en el 2000, la cual, con los datos que le dan origen, se muestra en las gráficas 10 (escala logit) y 11 (escala de las tasas).

Gráfica 9

Distribución de S_p y valores para las curvas de cuantiles 0.6 (sólida), 0.7 (trazos) y 0.8 (puntos)



Cuadro 4

Cuantiles de la distribución predictiva de S_p (1991-1998)

Probabilidad	0.900	0.950	0.990	0.995	0.999
Cuantil	58 598	59 753	61 900	62 763	64 448

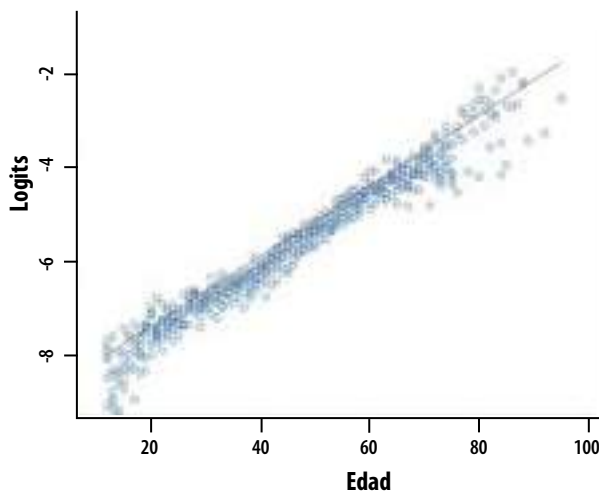
Cuadro 5

Número total de muertes para algunas curvas de cuantiles (1991-1998)

$1-\alpha$	0.60	0.70	0.80	0.90
Muertes	56 503	61 720	68 436	78 996

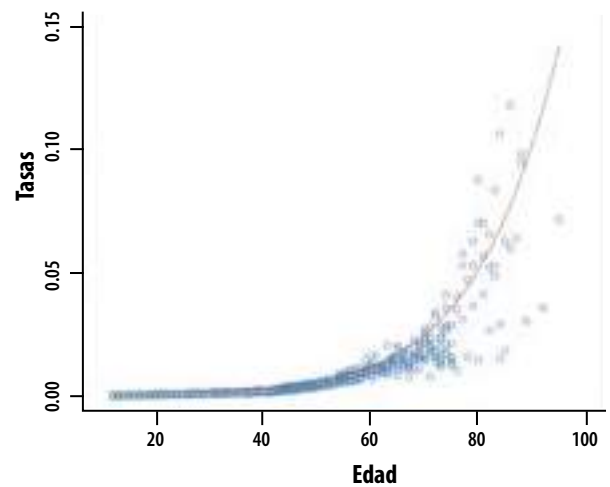
Gráfica 10

Datos de la década de los 90 y tabla final (escala logit)



Gráfica 11

Datos de la década de los 90 y tabla final (escala de las tasas)



3. Proceso de revisión y actualización de las tablas

En el 2005, fue posible tener acceso a una nueva base de datos de mortalidad que produjo la CNSF con información de 1999 al 2003. Los datos incluyen 397 tasas de mortalidad, cuya distribución aparece en el cuadro 6. Con esa información se evaluó si la tabla CNSF 2000-I debía actualizarse. En virtud de la naturaleza estadística de ésta, como primer paso se verificó si las nuevas tasas de mortalidad observadas podían considerarse compatibles con la distribución predictiva producida a partir de los datos de la década de los 90. La gráfica 12 superpone los logits, tanto para cada año como en forma agregada, con la correspondiente banda predictiva de probabilidad 0.95 derivada del modelo vigente. La mayoría de los datos son capturados por las bandas predictivas de manera que, en efecto, son compatibles con el modelo en un sentido operacional preciso: se presentan en una zona contemplada como posible pronóstico del modelo; sin embargo, una característica notable es que, para las edades entre 30 y 60 años, donde se

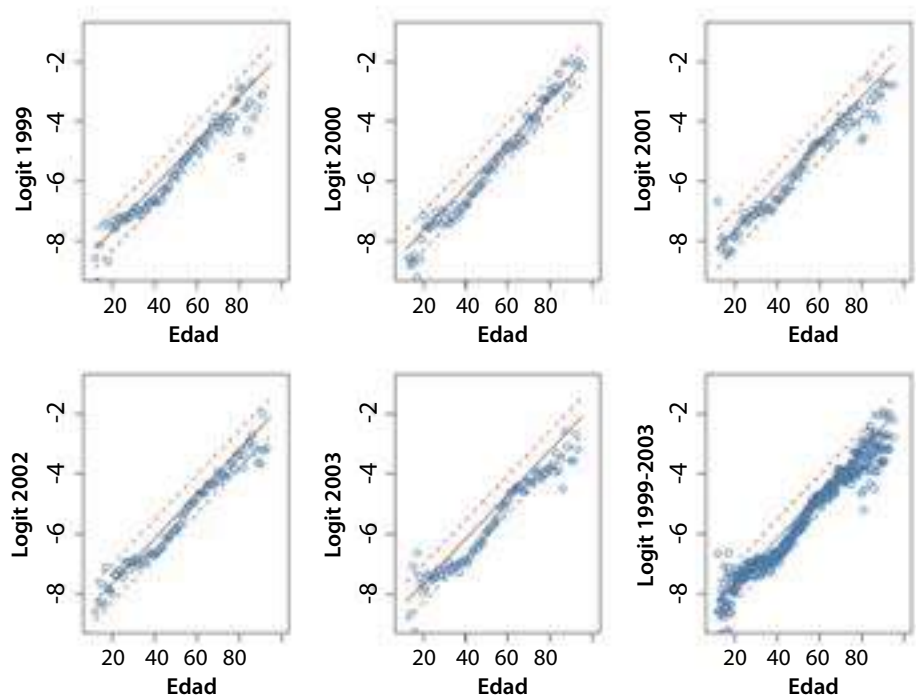
encuentra la mayor parte de la población asegurada, las tasas observadas típicamente se encuentran por debajo de la curva mediana. Este patrón sugiere que el modelo basado en los datos de la década de los 90 sobrestima la mortalidad más reciente. De manera incidental, la disminución de la mortalidad ha sido objeto de estudio por distintos autores (Lee y Carter, 1992; Bengtsson y Keilman, 2003, por ejemplo).

Cuadro 6
Número de observaciones por grupo de edad en el periodo 1999-2003

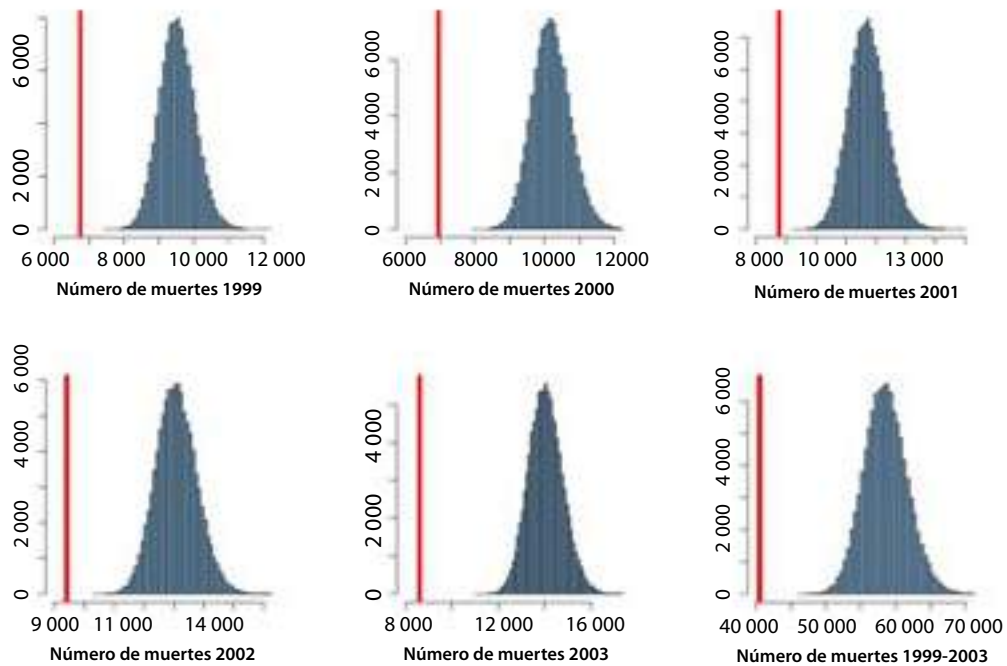
	1999	2000	2001	2002	2003
12-30	18	18	19	18	17
31-45	15	15	15	15	15
46-60	15	15	15	15	15
61-75	15	15	15	15	15
76-99	14	16	17	18	17
	77	79	81	81	79

Gráfica 12

Banda predictiva 0.95 y datos 1999-2003



Distribuciones predictivas y número de muertes observadas (línea sólida)



Cuadro 7

Número total de muertes y pronósticos del modelo 1991-1998

	1999	2000	2001	2002	2003
Muertes observadas	6 747	6 955	8 772	9 422	8 619
Media predictiva	9 530	10 185	11 728	13 087	14 044
Error relativo	41.25%	46.44%	33.70%	38.88%	62.94%

Con el propósito de establecer la magnitud de la eventual sobrestimación, se calculó la distribución predictiva para el total de muertes de cada año en el periodo 1999-2003 utilizando la distribución predictiva de las tasas derivada del modelo de la década de los 90 y el perfil de exposición de cada uno de esos cinco años. El análisis también se puede producir con el perfil agregado de expuestos de los cinco años para obtener la distribución predictiva para el número de muertes agregadas sobre todo el periodo. Estas distribuciones predictivas pueden contrastarse con el número de muer-

tes efectivamente observado en cada año considerado (o en el periodo completo) para decidir si es razonable mantener el modelo vigente o debe producirse uno nuevo que incorpore los datos más recientes. Los resultados muestran que, por ejemplo, el número total de muertes sobre el periodo completo 1999-2003 (40 515) es claramente sobrestimado por la media predictiva correspondiente (58 574) basada en los datos 1991-1998. El error relativo en este caso es de 44.6%, y el intervalo de pronóstico con probabilidad 0.99 (51 129, 67 076) fracasa cuando intenta capturar el valor verdadero. Conclusiones similares se obtienen cuando se analiza por separado cada año del periodo 1999-2003. Estos resultados se exhiben en la gráfica 13. Como complemento, el cuadro 7 muestra (para cada año) el valor observado, la media de la distribución predictiva y el error relativo que se produce si la media se utiliza para estimar ese valor verdadero (observado). El análisis arroja evidencia sustancial en favor del reemplazo de la tabla vigente y ésta fue la recomendación que se estableció como resultado del estudio.

4. Comentarios finales

La tabla de mortalidad que la autoridad reguladora mexicana publicó en el 2000 fue producida a partir de un modelo estadístico que describe la tendencia de las tasas de mortalidad respecto a la edad del asegurado, tomando en cuenta la variabilidad de las observaciones anuales. Con este mecanismo es posible definir una tabla de mortalidad como la curva formada por los cuantiles de orden $(1-\alpha)$ de la distribución *predictiva* de las tasas para cada una de las edades de interés. Así, la curva con las medianas constituye una tabla *básica* (o *central*) en el sentido de que, para cualquier edad, se puede asegurar que la correspondiente tasa anual futura resultará mayor (menor) al valor en la tabla con probabilidad exactamente 0.5. En forma similar, una curva de cuantiles correspondiente a un valor de $(1-\alpha)$ cercano a 1 es una tabla conservadora (*modificada*) que ofrece protección frente a desviaciones adversas (al alza) de la mortalidad futura. El tratamiento estadístico permite medir en términos probabilísticos el grado de protección que es el mismo en cada una de las edades. La condición es demandante en virtud de que se aplica de manera simultánea para todas las edades. Para evitar que conduzca a un requerimiento excesivo en la constitución de reservas, el nivel $(1-\alpha)$ se fija de acuerdo con la proyección que implica en la escala del total de muertes en la población asegurada.

Para decidir el eventual reemplazo de una tabla por una actualizada, se evalúa la compatibilidad de la nueva información disponible con la capacidad predictiva del modelo estadístico ajustado. Esta evaluación también se lleva a cabo en la escala del total de muertes de la población asegurada. Sólo cuando el modelo no es capaz de pronosticar de manera satisfactoria el total de muertes observadas en la nueva información, se propone reemplazar el modelo. El procedimiento es claro y simple.

El ajuste del nuevo modelo se puede efectuar con distintas estrategias. Es posible formar una base de datos que integre todos los datos (viejos y nuevos) para utilizar el procedimiento propuesto con esa información. Otra opción es ajustarlo sólo

con los actualizados. La primera opción responde a la idea de propiciar una transición más suave entre las tablas de mortalidad que produce la autoridad, pero existe la posibilidad de que el nuevo modelo resulte obsoleto relativamente pronto. Si, de manera alternativa, sólo se utiliza la información más reciente, el cambio en las tablas podría ser más dramático, pero el modelo resultante respondería mejor a la información nueva. En cualquier caso, la decisión debe considerar los aspectos de orden estadístico, pero también el impacto de los resultados en la práctica.

Otro asunto se refiere a la selección de la función que relaciona a las tasas con la edad. En el caso considerado en este artículo, tanto los datos de la década de los 80 como los de los 90 exhiben una tendencia que puede ser descrita satisfactoriamente con una regresión lineal; sin embargo, la gráfica de los datos del periodo 1999-2003 sugiere la conveniencia de un modelo diferente. En general, la función se puede determinar con un enfoque convencional, aplicando las ideas habituales de selección de modelos en regresión, o se pueden utilizar modelos más robustos, basados en *splines*, por ejemplo, o incluso los semiparamétricos. Estas ideas ya han sido exploradas y los resultados se reportarán en un trabajo futuro.

Es oportuno señalar que si bien el propósito de este artículo es mostrar un método para producir las tablas de mortalidad que, en el marco de la administración de riesgos moderna, requiere el sector de los seguros, el análisis de los datos que se presenta tiene algunas implicaciones que pueden ser de interés. Específicamente, se observa que las tasas de mortalidad de la población que cuenta con un seguro de vida en México disminuyeron en el periodo 1999-2003, si se comparan con las del lapso 1991-1998. Más aún, esa disminución es en especial notable en las edades que van de 25 a 60 años. Esta tendencia apunta, a lo largo de las siguientes décadas, a una reducción de las reservas y de las primas correspondientes. Es también de observarse que si la misma tendencia se manifiesta en los sistemas de pensiones, el efecto será el opuesto. Una mayor longevidad implicará un

esfuerzo adicional del sistema para cubrir las pensiones en la forma prevista. En cualquier caso, la disminución de la mortalidad que se observa coincide con los resultados de otros estudios (Ordorica, 2010 y García y Ordorica, 2012, por ejemplo).

Es importante reconocer, sin embargo, que la población bajo estudio en este artículo excluye a los menores de 12 años y a muchos ciudadanos que por su situación económica no tienen acceso a un seguro de vida ni a otros beneficios de salud que podrían mejorar sus condiciones de vida. En consecuencia, es razonable conjeturar que el descenso observado de las tasas de mortalidad es mayor al de la población abierta. Un estudio que se ocupe específicamente de esa comparación está fuera de los objetivos de este trabajo pero, sin duda, es de interés demográfico. De hecho, el tema abre la posibilidad para ensayar el método propuesto, más allá del ámbito actuarial, en el área de la Demografía.

Por último, el carácter predictivo de las tablas que aquí se proponen es de utilidad en el corto plazo. Si bien hasta ahora el sector asegurador ha actualizado las tablas después de periodos que oscilan entre 10 y 20 años, la sugerencia es que, con el método propuesto, la revisión y el eventual reemplazo se lleven a cabo con mayor frecuencia para, de esa forma, acortar el horizonte de pronóstico y, en consecuencia, aumentar la confianza en las estimaciones. Por otra parte, en el caso de los estudios demográficos de naturaleza general, es frecuente que se produzcan pronósticos de mortalidad con un horizonte de, por ejemplo, 50 años. Para servir a ese propósito, el modelo que aquí se ha propuesto tendría que incorporar la dinámica de cambio en el tiempo. Éste es otro tema que actualmente es objeto de investigación y será reportado en otro artículo.

Fuentes

- Bernardo, J. M. y A. F. M. Smith. *Bayesian Theory*. Chichester, Wiley, 2000.
- Bengtsson, T. y N. Keilman (eds.). *Perspectives on Mortality Forecasting: I. Current Practice*. Stockholm: Swedish National Social Insurance Board, 2003.
- Carriere, J. F. "Parametric models for life tables", en: *Transactions of Society of Actuaries*, 44, 77-99, 1992.
- Forfar, D. O., J. J. Mc Cutcheon y A. D. Wilkie. "On graduation by mathematical formula", en: *Journal of the Institute of Actuaries*, 115, 1-135, 1988.
- García, V. M. y M. Ordorica. "Proyección estocástica de la mortalidad mexicana por medio del método de Lee-Carter", en: *Estudios Demográficos y Urbanos*, 27(2), 409-448, 2012.
- Haberman, S. y A. E. Renshaw. "Generalized Linear Models and Actuarial Science", en: *Journal of the Royal Statistical Society, Series D*, 45(4), 407-436, 1996.
- Heligman, L. y J. H. Pollard. "The age pattern of mortality", en: *Journal of the Institute of Actuaries*, 107, 49-80, 1980.
- Kimeldorf, G. S. y D. A. Jones. "Bayesian graduation", en: *Transactions of the Society of Actuaries*, 19(54), 66-112, 1967.
- Lee, R. D. y L. R. Carter. "Modeling and Forecasting U. S. Mortality", en: *Journal of the American Statistical Association*, 87, 659-671, 1992.
- London, D. *Graduation: The Revision of Estimates*. Winsted and Abington, CT: ACTEX Publications, 1985.
- Mendoza, M., A. M. Madrigal y E. Martínez. "Análisis exploratorio de la Información del seguro de vida en México". Documento de trabajo 76. México, Comisión Nacional de Seguros y Fianzas, 1999a.
- _____. "Tablas de mortalidad CNSF 2000-I y 2000-G. información del seguro de vida en México". Documento de trabajo 77. México, Comisión Nacional de Seguros y Fianzas, 1999b.
- _____. "Modelos estadísticos de mortalidad. Análisis de datos 1991-1998". Documento de trabajo 80. México, Comisión Nacional de Seguros y Fianzas, 2000.
- Ordorica, M. "Las proyecciones de la población hasta la mitad del siglo XXI", en: *Los Grandes Problemas de México. Población*. García, B. y M. Ordorica (eds.). México, El Colegio de México, pp. 29-52, 2010.
- Renshaw, A. E. "Actuarial graduation practice and generalised linear and non-linear models", en: *Journal of the Institute of Actuaries*, 118, 295-312, 1991.
- Whittaker, E. "A method of graduation based on probability", en: *Proceedings of the Edinburgh Mathematical Society*, 41, 63-75, 1923.

Análisis de estadísticas del INEGI *sobre residuos sólidos urbanos*

José Paul Carrasco Escobar
y José Luis Ángel Rodríguez Silva



Las estadísticas del medio ambiente son de importancia crucial para comprender el entorno natural. Esto es particularmente cierto en el caso del comportamiento de los residuos sólidos urbanos (RSU), los cuales afectan diversos sistemas biológicos naturales y, en casos extremos, pueden causar profundas afectaciones al entorno. Con la finalidad de garantizar el tener más y mejor información acerca de los RSU, el Instituto Nacional de Estadística y Geografía (INEGI) de México ha realizado un esfuerzo importante por acopiar y difundir datos en dicha materia. El objetivo del presente documento es presentar los elementos fundamentales y las características principales del fenómeno de los RSU, así como la aplicación de las técnicas estadísticas de análisis de componentes principales y de análisis de conglomerados a esos datos y una interpretación de los resultados derivados, mismos que pueden servir como guía en el proceso de toma de decisiones en políticas públicas en la materia. En particular, se muestra que el fenómeno es básicamente bidimensional y que las entidades federativas se aglomeran, según el fenómeno de los RSU, en estratos bien delimitados, presentándose así un desequilibrio geográfico y de infraestructura en el tema.

Palabras clave: análisis de componentes principales, análisis de conglomerados, INEGI, residuos sólidos urbanos.

Recibido: 17 de septiembre de 2014

Aceptado: 9 de diciembre de 2014

1. Antecedentes de los residuos sólidos urbanos (RSU)

Este concepto, que en términos generales coincide con el de *basura doméstica y comercial usual* (esto es, que no sea de manejo especial o de tipo peligroso), es uno de los fenómenos que más impacto tiene sobre el medio ambiente y sus recursos. Lo anterior es cierto pues, por un lado, debido a la creciente demanda que tienen los seres humanos sobre sus satisfactores (que muchos de ellos al final se convierten en RSU), se tiene un mayor consumo de los recursos naturales y, en consecuencia, una afectación a los mismos; por el otro, al hacer la disposición final de los RSU, y debido tanto al gran volumen generado como al hecho de que en diversas ocasiones no se cumple con los estándar

Environmental statistics are decisive for understanding the natural environment. This is particularly true in the case of the behavior of Municipal Solid Waste (MSW), which affects various natural biological systems and, in extreme cases, can cause deep damages to the environment. In order to make sure that we have more and better information about the MSW, the National Institute of Statistics and Geography (INEGI) has made a significant effort by collecting and disseminating information on this subject. The aim of this paper is to present the key elements and main characteristics of the phenomenon of MSW, as well as the application of two statistical techniques to these data: Principal Component Analysis and Cluster Analysis, following an interpretation of derivative results, which could serve as a guide in the process of decision making of public policies in this area. In particular, it is shown that the phenomenon is basically a two-dimensional one, and that the states are clustered, like the phenomenon of MSW, into separate strata, thus presenting a geographic and infrastructural imbalance in this area.

Key words: Principal Component Analysis, Cluster Analysis, INEGI, municipal solid waste.

res y normas al respecto, se hace una afectación al medio ambiente. Esto puede implicar, entre otras situaciones, graves problemas en materia de salud pública, utilización de espacios y recursos, además de mayor contaminación ambiental, por mencionar algunas.

Son numerosos los ejemplos en los cuales, por una falta de cuidado en alguno de los múltiples procesos del sistema de los RSU, se tienen consecuencias de gran impacto en el medio ambiente. En el *Informe de la situación ambiental del medio ambiente en México* —Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT), 2012— se señala que existen profundas relaciones entre la generación indiscriminada y mal manejo de los RSU y el cambio climático, el adelgazamiento de la capa

de ozono, la creciente contaminación de suelos y cuerpos de agua, así como la proliferación de fauna nociva y la transmisión de enfermedades.

Conviene, en estos momentos, proporcionar la definición oficial de los RSU, tomada del artículo 5, fracción XXXIII, de la *Ley General para la Prevención y Gestión Integral de los Residuos* (LGPGIR)¹ (Secretaría de Gobernación, 2014), que los define como: "...los generados en las casas habitación, que resultan de la eliminación de los materiales que utilizan en sus actividades domésticas, de los productos que consumen y de sus envases, embalajes o empaques; los residuos que provienen de cualquier otra actividad dentro de establecimientos o en la vía pública que genere residuos con características domiciliarias, y los resultantes de la limpieza de las vías y lugares públicos, siempre que no sean considerados por esta ley como residuos de otra índole...".

La administración del proceso de los RSU, conocido como *Gestión integral de los residuos sólidos urbanos*, consta de un intrincado sistema de partes íntimamente relacionadas unas con otras. En sus

aspectos más simples, implica tres grandes etapas: Generación → Recolección → Disposición final. A su vez, se dividen en varias subetapas, algunas de las cuales se muestran en la figura 1.

La administración pública federal, encabezada en materia de políticas públicas de gestión ambiental por la SEMARNAT, ha llevado a cabo esfuerzos para generar un marco legal para la gestión de los RSU en México, apoyada en la LGPGIR.

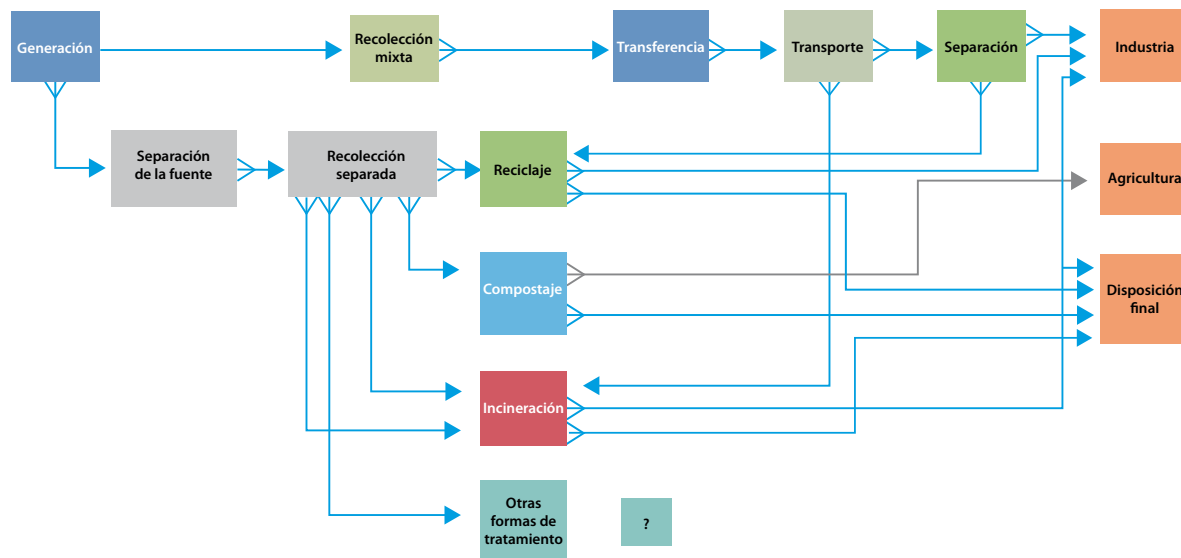
Más aún, en el *Plan Nacional de Desarrollo 2007-2012*, en su Eje 4 asociado a la sustentabilidad ambiental, se detalló en el apartado 4.7 la finalidad del así llamado Programa Nacional para la Prevención y Gestión Integral de los Residuos que, entre otras acciones, en su objetivo 4.2.1, se comprometió a lograr el manejo integral de los residuos mediante la aplicación de los instrumentos, acciones y estrategias contempladas en el marco legal de esos años.

En uno de los primeros ejercicios al respecto, la SEMARNAT, mediante el entonces Instituto Nacional de Ecología (INE) que ahora es el Instituto Nacional de Ecología y Cambio Climático (INECC), llevó a cabo en el 2006 el primer diagnóstico básico de los residuos que, según el artículo

¹ Tal vez el documento normativo más importante en la materia; entró en vigor el 8 de octubre de 2004 y su reglamento fue publicado en el *Diario Oficial de la Federación* (DOF) el 30 de noviembre de 2006, con entrada en vigor 30 días posteriores a su publicación.

Figura 1

Diagrama de flujo de un sistema de manejo de residuos sólidos diferenciado



Fuente: rediseño propio inspirado en SEMARNAT, 2007.

25 de la LGPGIR es "...el estudio que considera la cantidad y composición de los residuos, así como la infraestructura para manejarlos integralmente...". Estos esfuerzos se encaminaban a tratar de establecer una evaluación confiable acerca de la verdadera problemática de los RSU en el país, para así generar información y estadísticas que permitieran guiar las políticas públicas en la materia. Se generaron datos muy interesantes, determinando que la generación de RSU se estimó, para el 2006, en 94 800 toneladas diarias, equivalente a 34.6 millones de toneladas anuales, con una composición aproximada de 53% de materia orgánica; 28%, de residuos potencialmente reciclables y 19%, de no aprovechables. Se estimó que se recolectaba 87% de los residuos generados y, de éstos, 64% se enviaba a los 88 rellenos sanitarios y 21 sitios controlados; lo restante se depositaba en tiraderos a cielo abierto y en sitios sin control.

Uno de los grandes problemas de este diagnóstico fue que se realizó, en estricto sentido, mediante un esquema de muestreo en algunas entidades federativas elegidas para tal fin. Con esto, no se tenía un marco de muestreo inicial del cual partir, que es indispensable para llevar a cabo estudios muestrales por medio de encuestas. Más aún, varios de los procedimientos para formar las cifras anteriores se basaban en una diversidad de criterios y estimaciones que trataban de extrapolarse a todo el país, situación en sí controversial, pues la realidad del fenómeno de los residuos podría ser distinta debido a la heterogeneidad de las regiones que lo integran.

Con la finalidad de aprovechar su experiencia y estructura organizacional, el INEGI se sumó desde el 2009 a la SEMARNAT para colaborar en el esfuerzo de diagnosticar la situación que guardaba el país en este tema. Mediante una serie de reuniones y trabajo conjunto, se formó la propuesta de un cuestionario sobre RSU, el cual se aplicó a una muestra de municipios (unidad de observación) de cada entidad federativa del país en el 2010 generando, así, los primeros resultados a nivel nacional, para dar aprovechamiento a los registros administrativos en materia de RSU.

Los esfuerzos anteriores contribuyeron a la formación de un cuestionario sobre RSU el cual, finalmente, se decidió incorporar como módulo del operativo de los censos nacionales de Gobiernos Municipales y Delegacionales 2011 (con año de referencia 2010) y 2013 (con año de referencia 2012). De ahí se generaron una serie de productos que han sido fundamentales para conocer con un mayor nivel de detalle y actualidad las cifras de los RSU, mismas que serán descritas en el tercer apartado de este documento.

Para cerrar esta primera sección, cabe mencionar que la información sobre RSU ha adquirido una gran relevancia y como ejemplo, el 7 de noviembre de 2014 se publicó en el DOF (INEGI, 2014a) el llamado *Acuerdo por el que se aprueba la inclusión al Catálogo Nacional de Indicadores de un conjunto de indicadores clave en materia de emisiones y residuos*, en el cual se establece la adición en dicho *Catálogo*, en el marco del Sistema Nacional de Información Estadística y Geográfica, de los siguientes indicadores:

- Porcentaje de municipios con disposición adecuada de RSU.
- Proporción de la población con acceso a la recolección de residuos.

El hecho de que estos dos indicadores se hubiesen propuesto y aprobado implica que esta información se considera básica para la formación de políticas públicas en la materia tendientes a contar, de manera general, con un mejor sistema de gestión integral de los RSU, situación en sí misma de importancia. Además, dado que la construcción de los dos indicadores será definida en la forma y términos que determine el Comité Técnico Especializado de Información sobre Emisiones, Residuos y Sustancias Peligrosas, las variables definidas en este estudio (ver tercer apartado) pueden contribuir a su diseño conceptual.

Una aportación relevante de los resultados derivados del presente estudio es la asociada al poder explicativo que guardan las técnicas de la estadística multivariada en relación con la determinación de factores más relevantes de algunas de las estadísticas oficiales en materia de medio ambiente. La

emergencia de significados desde los datos —pero no de los datos en sí mismos— hace de este trabajo una fuente adecuada para el conocimiento del fenómeno de los RSU y puede ser utilizada para un mejor entendimiento del fenómeno y, así, poder profundizar en él, favoreciendo el desarrollo de respuestas respecto a lo que está ocurriendo en tal problemática y el porqué, sin la necesidad de abarcar cada una de las dimensiones del problema, sino enfocándose a las que estadísticamente tienen mayor influencia. Además, los resultados de estudios de este tipo permiten construir teoría, conceptos, hipótesis y proposiciones partiendo de forma directa de los datos y no de los supuestos a priori, de otras investigaciones o de marcos teóricos existentes. La teoría generada se desarrolla de manera inductiva a partir de un conjunto de datos y esto implica un empare de ésta con la realidad del objeto de estudio. Ello contrasta con una teoría derivada deductivamente de alguna otra gran teoría, sin la ayuda de datos y que podría, por lo tanto, no cuadrar con la realidad.

Por otro lado, esta investigación muestra una manera alternativa para el aprendizaje de métodos estadísticos en el contexto de los datos oficiales. El aprendizaje moderno de la estadística está cada vez más ligado al contexto de los datos y a la atención de los requerimientos de quienes toman decisiones basadas en información confiable y de calidad. La disposición de estadísticas oficiales sobre los temas sociales, económicos y de los registros de las funciones administrativas y de control de los gobiernos es cada vez más común y su aprovechamiento es una labor propia tanto de especialistas como de otros profesionales que cuentan con formación estadística. Así, este documento puede verse también como parte de un estudio de caso sobre aprendizaje y evaluación del uso de técnicas estadísticas multivariadas, así como los resultados de su aplicación a datos oficiales sobre el medio ambiente.

2. Marco teórico de dos técnicas estadísticas multivariadas

A continuación, se presentan los aspectos básicos de éstas, las cuales se aplicaron a una serie de estadís-

ticas de los RSU. Es un resumen que sirve de referencia inmediata al lector, aunque al interior de la descripción de cada técnica se da una bibliografía que es relativamente sencilla de conseguir, por si se busca tener un mayor grado de conocimiento al respecto.

2.1 Análisis de componentes principales (ACP)

Pertenece al conjunto de técnicas de la Estadística multivariada que son de tipo descriptivo, esto es, los objetivos que en principio persigue el ACP no son de tipo inferencial o, lo que es lo mismo, pretende más bien describir un cierto conjunto multivariado de observaciones, más que generalizar o inferir sobre alguna población.

Supóngase que se tiene un conjunto de datos multivariados, de n registros u observaciones y p variables, los cuales se disponen en forma tabular de la siguiente forma:

$$X = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix}$$

Lo que busca la técnica de ACP es formar un nuevo conjunto de variables, que son combinaciones lineales de las primeras, y que son designadas como Z_1, Z_2, \dots, Z_p , las cuales no tienen correlación lineal entre sí y la primera de ellas es la más importante en el sentido de que es la que tiene más variabilidad, de ahí el adjetivo de *principal*, para luego pasar a la segunda en orden de importancia, y así sucesivamente:

$$\text{var}(Z_1) \geq \text{var}(Z_2) \geq \cdots \text{var}(Z_p).$$

Para que la técnica de ACP tenga relevancia en un problema particular, debe ser el caso que las variables originales tengan altos niveles de correlación entre ellas; de hecho, en el caso de que dichas variables no estén correlacionadas, el ACP no tiene ningún efecto en lo absoluto.

El empleo de la metodología de ACP involucra, matemáticamente, la resolución de un problema de maximización, lo cual implica la resolución de un problema de descomposición espectral, tópicos sobre los cuales no se entrará en detalle en este documento. El lector interesado puede consultar Johnson (2000, pp. 93-146), Johnson y Wichern (2002, pp. 426-476) y Manly (1984, pp. 59-71).

Dentro de los resultados que ofrece el ACP se encuentran los siguientes:

- Escribiendo las componentes principales:

$$CP_1 = a_{11}X_1 + a_{12}X_2 + \dots + a_{1p}X_p = \text{componente principal 1}$$

$$CP_2 = a_{21}X_1 + a_{22}X_2 + \dots + a_{2p}X_p = \text{componente principal 2}$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots \quad \vdots$$

$$CP_p = a_{p1}X_1 + a_{p2}X_2 + \dots + a_{pp}X_n = \text{componente principal } p$$

Los coeficientes a_{ij} (conocidos como las *cargas* de las componentes principales) nos permiten valorar la importancia que las variables originales tienen sobre las componentes principales que, junto con el signo de tales cargas, brindan potencialmente un método para identificar de forma cualitativa el significado de las componentes principales, de las cuales hay tantas como variables originales, aunque tienen la interesante propiedad de que la correlación entre cualquier par de componentes principales distintas entre sí es 0, por lo que la dirección multidimensional de cada componente principal tiene un sentido que explica una dimensión propia del fenómeno y de ahí que si se determina que, por ejemplo, se requieren dos componentes principales para explicar 80% de la varianza, podremos decir que el fenómeno es básicamente bidimensional.

- Mediante la determinación numérica de los valores de las componentes principales es posible llevar a cabo un ordenamiento de los registros, según su impacto en las componentes que, para fines prácticos, casi siempre se toman sólo las primeras de ellas. Quedará explicado este procedimiento al momento que se aplique en los datos de RSU bajo consideración.

2.2 Análisis de conglomerados (AC)

Es una poderosa herramienta que nos permite la formación de grupos de observaciones o de variables con la finalidad de asociar aquellos que tengan la mayor *afinidad* entre sí. Es, quizá, uno de los recursos estadísticos más utilizados en diversas ramas, como Sociología, Economía, mercadotecnia, etc., debido a que uno de los fines que tienen los investigadores es el de buscar patrones de afinidad al interior de la información que pretenden describir y explotar.

La técnica comienza con la elaboración de una matriz de distancias, que consta de un conjunto matricial de números que definen el grado de cercanía que tienen los individuos entre sí. Debido a que existen potencialmente muchas formas de definir estas distancias —siempre y cuando cumplan con los requisitos de la técnica—, también hay una gran diversidad de métodos de conglomerados. A su vez, según la forma en que se aglomeran de forma iterativa los individuos (o las variables), pueden distinguirse diversos métodos; dentro de los más utilizados son los llamados *jerárquicos*. En tal sentido, es difícil mencionar, sin ambigüedad alguna, cuál es el mejor método en algún problema dado, pues hay una gran cantidad de criterios al respecto; sin embargo, tal como se comentó arriba, de cualquier forma son de gran utilidad práctica.

Uno de los resultados gráficos más notables del AC son los llamados *dendogramas*, los cuales son figuras que permiten visualizar la forma de aglomeración de nuestros objetos de interés; son gráficas que se tendrá la oportunidad de utilizar e interpretar en la sección 3.3.

El lector interesado puede consultar mayores detalles de las técnicas de análisis de conglomerados en Johnson (2000, pp. 319-396), Johnson y Wichern (2002, pp. 668-745) y Manly (1984, pp. 100-113).

Antes de cerrar este apartado, algunas palabras: existe una diversidad de *software* estadístico que permite llevar a cabo tanto el ACP como el AC, y entre los que han demostrado en la práctica mayores bondades son el Minitab® 17, IBM

SPSS® 22 y Statística® 12; sin embargo, también hay un *software* libre, el R (release 3. 1. 12), el cual es muy eficaz y cuenta con una gran cantidad de librerías especializadas, así como una comunidad internacional dispuesta a brindar apoyo en problemas técnicos que pudieran presentarse. Aunque existe un pequeño precio al respecto, y es el hecho de tener que llevar a cabo un cierto segmento de programación, lo cual, en algunos usuarios, implicaría la inversión de más tiempo para familiarizarse con dicho entorno.

3. Aplicación de las ACP y AC a un subconjunto de datos de RSU

Tal y como se ha señalado, el INEGI ha realizado un gran esfuerzo por acopiar información de RSU; hasta el momento se han logrado generar datos de gran relevancia, pues antes de este ejercicio de estadística oficial no se tenía información al nivel de detalle con la que ahora se cuenta y mucha de ella provenía de estimaciones (SEMARNAT, 2012).

Dentro de la información que en materia de RSU se genera en el INEGI se tienen los siguientes temas (INEGI, 2014b):

- Centros de acopio:
 - Materiales valorizables, según tipo, para 2010, por entidad federativa.
 - Materiales valorizables, según tipo, para 2012, por entidad federativa.
- Estaciones de transferencia:
 - Procesos reportados de RSU, para 2010, por entidad federativa.
 - Procesos reportados de RSU, para 2012, por entidad federativa.
- Estudios:
 - Composición de RSU en el periodo 2008-2010, por entidad federativa.
 - Composición de RSU, para 2012, por entidad federativa.
 - Generación de RSU en el periodo 2008-2010, por entidad federativa.
 - Generación de RSU, para 2012, por entidad federativa.
- Programas locales:
 - Gestión integral de RSU en el periodo 2008 a 2010, por entidad federativa.
 - Gestión integral de RSU, para 2012, por entidad federativa.
- Recolección:
 - Promedio diario, según tipo de recolección, para 2010, por entidad federativa.
 - Promedio diario, para 2010, según entidad federativa.
 - Promedio diario, para 2010, según municipio y delegación.
 - Promedio diario, para 2012, según entidad federativa.
 - Promedio diario, para 2012, según municipio y delegación.
 - Tipo de vehículo, para 2010, por entidad federativa.
 - Tipo de vehículo, para 2010, por municipio y delegación.
 - Tipo de vehículo, para 2012, por entidad federativa.
 - Tipo de vehículo, para 2012, por municipio y delegación.
- Disponibilidad de servicios para RSU:
 - Para 2010, según entidad federativa.
 - Para 2010, según municipio y delegación.
 - Para 2012, según entidad federativa.
 - Para 2012, según municipio y delegación.
- Reglamentación en cuanto a servicios de recolección y disposición final, para 2010.
- Tipos de sitios de disposición final de los RSU, para 2010.

Así, el INEGI pone a disposición de los usuarios en general un vasto conjunto de datos acerca del fenómeno de los RSU. Para el ejercicio mostrado en este trabajo, los autores acopiaron diversas variables asociadas de los RSU, según está dispuesta la información en el sitio del Instituto en internet (www.inegi.org.mx) para determinar niveles de asociación multivariados en distintas fases de la gestión integral de residuos. Conviene señalar que no se consideraron todas las variables que se encuentran en el sitio pues algunas de ellas no contaban con información para todas las entidades federativas, por ejemplo, para variables de

PET. Así, en total, se consideraron 20 variables de proporción de la siguiente manera:

- VAR_1 = proporción de municipios y delegaciones con centros de acopio respecto al total de municipios en la entidad federativa.
- VAR_2 = proporción del total municipios y delegaciones con centros de acopio respecto al total nacional de centros de acopio.
- VAR_3 = proporción de municipios y delegaciones con estaciones de transferencia, sólo almacenamiento temporal y traspaso respecto al total de municipios en la entidad federativa.
- VAR_4 = proporción de municipios y delegaciones con estaciones de transferencia, compactación y/o selección de materiales respecto al total de municipios en la entidad federativa.
- VAR_5 = proporción de municipios y delegaciones con estudios sobre la generación de residuos sólidos urbanos respecto al total de municipios en la entidad federativa.
- VAR_6 = proporción de municipios y delegaciones con programas locales orientados a la gestión integral de los residuos sólidos urbanos respecto al total de municipios en la entidad federativa.
- VAR_7 = proporción de municipios con estaciones de transferencia respecto al total de municipios de la entidad federativa.
- VAR_8 = proporción del total de municipios con estaciones de transferencia respecto al total nacional.
- VAR_9 = proporción de vehículos utilizados para la recolección de residuos sólidos urbanos con compactador respecto al total en la entidad federativa.
- VAR_{10} = proporción de vehículos utilizados para la recolección de residuos sólidos urbanos con caja abierta respecto al total en la entidad federativa.
- VAR_{11} = proporción de otros vehículos utilizados para la recolección de residuos sólidos urbanos y otro tipo de vehículos respecto al total en la entidad federativa.
- VAR_{12} = proporción del total de municipios y delegaciones con disponibilidad de servicios

relacionados con los residuos sólidos urbanos respecto al total en la entidad federativa.

- VAR_{13} = proporción de municipios y delegaciones con disponibilidad de servicios de sólo recolección y disposición final de residuos sólidos urbanos respecto al total en la entidad federativa.
- VAR_{14} = proporción de municipios y delegaciones con disponibilidad de servicios de recolección, disposición final y tratamiento de residuos sólidos urbanos respecto al total en la entidad federativa.
- VAR_{15} = proporción del promedio diario de pesaje de residuos sólidos urbanos recolectados respecto al total en la entidad federativa.
- VAR_{16} = proporción del promedio diario por vehículo/capacidad/viajes de residuos sólidos urbanos recolectados respecto al total en la entidad federativa.
- VAR_{17} = proporción del promedio diario de otros métodos de obtención de residuos sólidos urbanos recolectados respecto al total en la entidad federativa.
- VAR_{18} = proporción de municipios y delegaciones con servicio de recolección y disposición final con reglamento relacionado con los residuos sólidos urbanos respecto al total en la entidad federativa.
- VAR_{19} = proporción de sitios de disposición final tipo relleno sanitario reportados como destino de los residuos sólidos urbanos respecto al total en la entidad federativa.
- VAR_{20} = proporción de sitios de disposición final tipo tiradero a cielo abierto reportados como destino de los residuos sólidos urbanos respecto al total en la entidad federativa.

Los siguientes hechos son relevantes en cuanto a la naturaleza y selección de estas variables:

- Los datos se presentan en el *Anexo 1*.
- En primer lugar, la justificación de tomar proporciones y no los valores numéricos en absoluto es debido a que México es un país de gran heterogeneidad, misma que se manifiesta parcialmente en el hecho de que varias entidades federativas cuentan con números

- muy distintos de municipios —y delegaciones, en el caso del Distrito Federal (DF)—, cada uno con sus propias peculiaridades. Así, por ejemplo, el DF tiene 16 delegaciones de las cuales todas cuentan con estudios sobre la generación de residuos, mientras que Jalisco, con 125 municipios, realizó 29 estudios de tal naturaleza. Si no se llevara a cabo la proporción (con 100% para el caso del DF y de 23.2% en el de Jalisco), la aplicación sobre números absolutos de las técnicas estadísticas que se emplean más adelante podría llevarnos a resultados confusos.
- Todas las variables bajo estudio se especificaron en el sentido de *coadyuvar a la mejor atención del problema de los RSU*, esto es, para tener una mayor potencia en la interpretación de varias de las cantidades derivadas del ACP y un mejor nivel de asociación en el AC. La literatura recomienda el hecho de que las variables se manifiesten en *solamente un sentido*, ya sea en la mejora del problema (por ejemplo, más y mejores plantas de transferencia) o bien, todas en un sentido de empeoramiento del problema (e. g., mayor generación per cápita de RSU). Éste es, finalmente, un artilugio matemático para que la interpretación sea lo más eficiente posible y que no se entremezclen de manera contextual sentidos de variables que puedan estar en sentidos contrarios u opuestos, situación que los métodos estadísticos no podrían determinar por sí solos y que es competencia del analista.
 - Más aún, para no crear correlaciones espurias, se eliminaron los totales cuando se tenían las variables de desagregación y, a su vez, se dejaron sólo las mínimas indispensables sin pérdida de información en el sentido de que con las que se consideran en el análisis sea posible reproducir la base de datos original.
 - Las variables analizadas fueron perturbadas de forma aleatoria alrededor de su valor, con la función *jitter* de R, y los resultados fueron prácticamente los mismos a los que serán comentados en las siguientes secciones, por lo cual no parece ser que sean demasiado sensibles los resultados a pequeños cambios alea-

torios en las cifras, situación que robustece en sí las conclusiones de este trabajo.

- Los datos sólo se tomaron por entidad federativa, pues a pesar de que hay resultados a un nivel más fino —esto es, al de municipio—, no todas las variables de interés se encuentran desagregadas a tal grado. A su vez, para tener resultados a nivel del país, parece apropiado realizar el análisis a nivel de entidad.
- Por último, el único año analizado fue el 2012. A pesar de que hay información para años anteriores, el alcance de este trabajo se acota sólo para éste, debido a que no se pretendió un estudio longitudinal, sino el aprovechamiento de la información actual.

Una de las razones contextuales —quizá la más relevante— para llevar a cabo un análisis cruzado con diversas variables de RSU, tal y como se señala, es que varias entidades federativas pueden, quizá, tener una buena gestión en algún rubro, aunque tal vez no tanto en otro. Así, el formar un análisis multivariado de esta naturaleza puede aportar al fortalecimiento de una mejor comprensión global del fenómeno, así como coadyuvar en la toma de decisiones en materia de políticas públicas.

3.1 Análisis de correlaciones

Al llevar a cabo un análisis de correlaciones, se observó que del total de $19 \times 20/2 = 190$ correlaciones de variables en sus cruces (eliminando las correlaciones triviales), y al contar el número de correlaciones que, en valor absoluto, fueron mayores a 0.7, se encontró que estas últimas fueron 35, por lo que hay una proporción no trivial de casi 18.42% que gozan de dicha propiedad. Si el valor de umbral de 0.7 disminuye a 0.5 —el cual, para ciertos análisis de estadística espacial se considera un valor alto (e. g. Bivand, Pebesma y Gómez-Rubio, 2013)—, la proporción se incrementa a casi 25 por ciento. Más aún, existen correlaciones muy altas, por ejemplo, entre VAR_{11} y VAR_{16} , de un orden de 0.99. La tabla completa de correlaciones se presenta en el *Anexo 2*.

Estas consideraciones acerca de las correlaciones justifican, en parte, el que sea razonable llevar a cabo tanto el ACP como el AC.

3.2 Aplicación del ACP

En el cuadro que aparece enseguida se presentan los valores de las varianzas explicadas por cada componente principal. Se observa que utilizando la *regla de más de 70% de varianza explicada*, serían necesarias tres componentes principales para explicar nuestro fenómeno de 20 dimensiones. Si exigimos un poco menos y empleamos la regla de *un mínimo de 60% de varianza explicada*, bastaría con dos componentes principales. Así, entonces, parece claro que la problemática de los RSU tiene de dos a tres dimensiones y no el número total de variables de proporción que están bajo análisis. La anterior afirmación se refuerza al considerar el diagrama de sedimentación, mostrado en la gráfica 1, en la cual el *codo* de la misma se presenta

en la segunda componente principal, motivo por el cual nos inclinamos a la consideración de que la dimensionalidad más importante es de 2, en el sentido de que 61.7% de la varianza acumulada puede explicarse con tal número de componentes principales.

Escribamos ahora las dos primeras componentes principales:

$$Z_1 = 0.294VAR_1 + 0.303VAR_2 + 0.205VAR_3 + 0.024VAR_4 + 0.229VAR_5 + 0.110VAR_6 + 0.265VAR_7 + 0.211VAR_8 + 0.282VAR_9 + 0.165VAR_{10} + 0.295VAR_{11} + 0.008VAR_{12} - 0.301VAR_{13} + 0.302VAR_{14} + 0.082VAR_{15} + 0.295VAR_{16} + 0.302VAR_{17} + 0.096VAR_{18} - 0.005VAR_{19} - 0.174VAR_{20}$$

$$Z_2 = -0.047VAR_1 - 0.145VAR_2 + 0.328VAR_3 + 0.227VAR_4 + 0.119VAR_5 - 0.010VAR_6 + 0.261VAR_7 - 0.014VAR_8 - 0.110VAR_9 - 0.322VAR_{10} - 0.075VAR_{11} + 0.323VAR_{12} - 0.111VAR_{13} - 0.060VAR_{14} + 0.020VAR_{15} - 0.184VAR_{16} + 0.028VAR_{17} + 0.391VAR_{18} + 0.464VAR_{19}$$

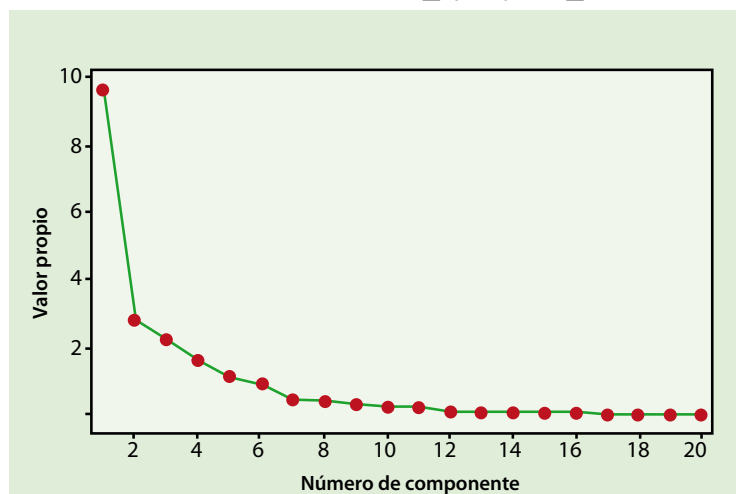
Valores propios, proporción y proporción acumulada de las CP sobre datos de RSU

	CP ₁	CP ₂	CP ₃	CP ₄	CP ₅	CP ₆	CP ₇	CP ₈
Valor propio	9.5721	2.7735	2.2142	1.5967	1.1128	0.9018	0.4585	0.3828
Proporción	0.479	0.139	0.111	0.08	0.056	0.045	0.023	0.019
Acumulada	0.479	0.617	0.728	0.808	0.863	0.909	0.931	0.951

Fuente: elaboración propia con resultados generados en Minitab®.

Gráfica 1

Sedimentación de Var₁, ..., VAR₂₀



Fuente: elaboración propia con resultados generados en Minitab®.

$-0.301VAR_{20}$

Se observa que la primera componente principal adquiere sus mayores valores en las variables 1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 13, 14, 16 y 20. Al analizar las primeras, notamos que éstas se encuentran involucradas con la infraestructura (vehículos y medios para pesar los RSU) en la gestión integral de los RSU. A su vez, los signos negativos se presentan en las variables 13 y 20, que se asocian sobre todo con la disposición final de los residuos, aunque los coeficientes son relativamente pequeños. Así, podemos asociar a esta componente una interpretación *de infraestructura de gestión de los RSU, salvo disposición final*. Ésta sería la componente principal que más impacto tiene en general en el fenómeno de los RSU.

De manera similar, en la segunda componente principal se contrasta en general una parte conceptual, de planeación, con respecto a la infraestructura u operación. Así, podríamos pensar que esta segunda componente representa la *planeación en contra de la operación en la gestión de los RSU*.

Ahora, la gráfica 2 corresponde a la puntuación

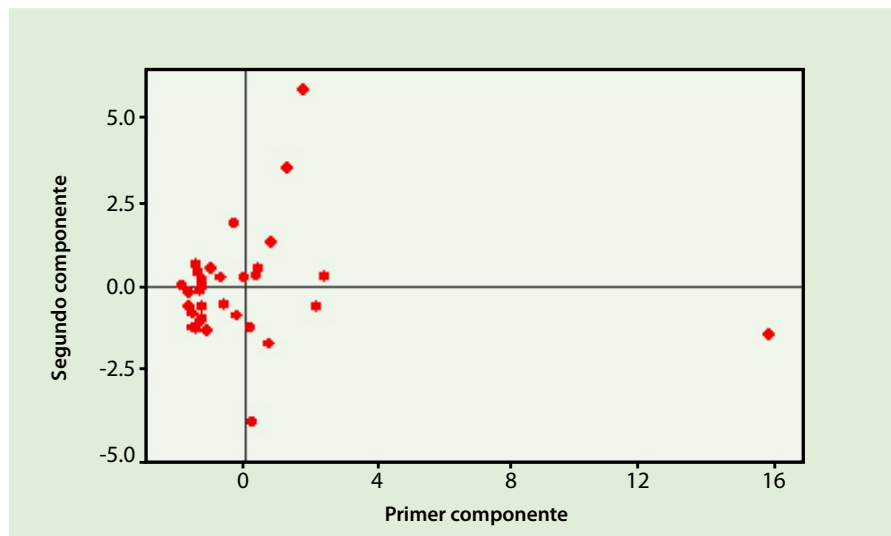
de las dos primeras componentes principales, y en ella se aprecia de manera muy marcada el punto extremo derecho, que corresponde al DF, el cual tuvo el mayor valor. Otros dos valores que son destacables son el punto superior, correspondiente a Baja California, con un gran valor en la segunda componente, y Oaxaca con el menor valor en esta última componente principal.

La gráfica 3 nos proporciona información, según el ACP, acerca de la afinidad de las variables, destacando que la número 13 parece ser la que marca una dimensionalidad distinta con respecto a las restantes, así como la 20 y la 10, aunque éstas en menor medida. Por otro lado, las variables 4, 12 y 19 tienen una gran afinidad en tal sentido.

Por último, se presenta el orden de las entidades federativas según la primera componente principal: DF, Baja California, Aguascalientes, Querétaro, Jalisco, Morelos, estado de México, Nuevo León, Quintana Roo, Guanajuato, Veracruz de Ignacio de la Llave, Oaxaca, Michoacán de Ocampo, Puebla, Tlaxcala, Hidalgo, Colima, Chihuahua, Baja California Sur, Tabasco, Sinaloa, Coahuila de Zaragoza, Chiapas, Tamaulipas, San Luis Potosí,

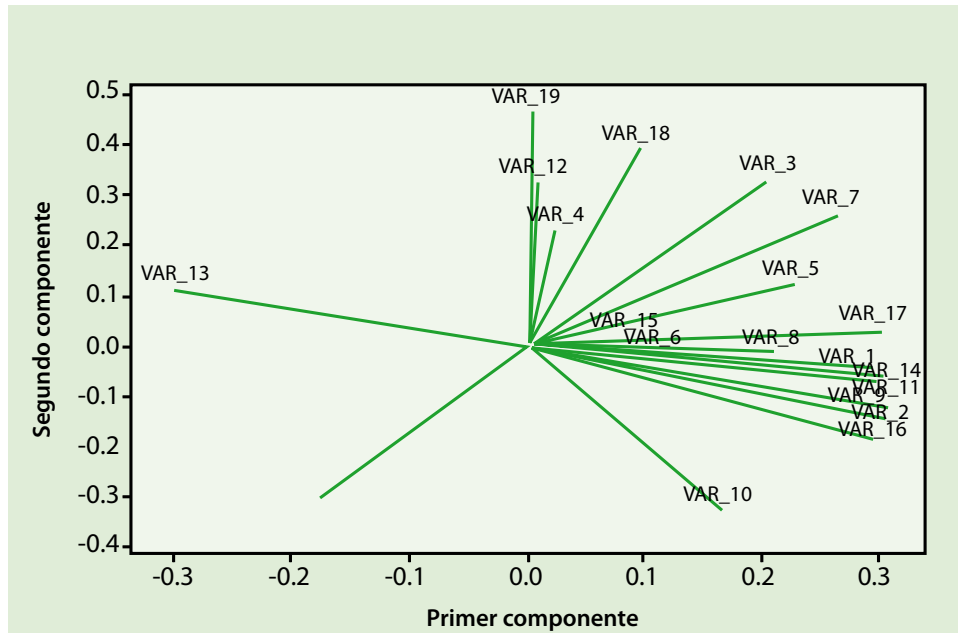
Gráfica 2

Puntuación de Var_1, \dots, VAR_{20}



Fuente: elaboración propia con resultados generados en Minitab®.

Cargando gráfica de Var_1, ..., VAR_20



Fuente: elaboración propia con resultados generados en Minitab®.

Durango, Sonora, Zacatecas, Campeche, Nayarit, Guerrero y Yucatán. De este ordenamiento se observa que las cinco primeras tienen el mayor desarrollo en *infraestructura de gestión de los RSU, salvo disposición final*, mientras que las que tienen mayor área de oportunidad al respecto son las cinco últimas.

Realizando el ordenamiento, ahora con respecto a la segunda componente principal, los resultados son: Baja California, Aguascalientes, DF, Quintana Roo, Querétaro, Jalisco, Guanajuato, Morelos, estado de México, Colima, Nuevo León, Chihuahua, Coahuila de Zaragoza, Tlaxcala, Baja California Sur, Sinaloa, Tabasco, Campeche, Durango, Hidalgo, Puebla, Michoacán de Ocampo, Nayarit, Sonora, Zacatecas, Veracruz de Ignacio de la Llave, San Luis Potosí, Tamaulipas, Chiapas, Yucatán, Guerrero y Oaxaca. Así, en la dimensión de la *planeación en contraste con la operación en la gestión de los RSU*, las cinco entidades federativas que parecen tener mayor desarrollo son las cinco primeras y

las menos favorecidas en tal dimensión son las últimas cinco.

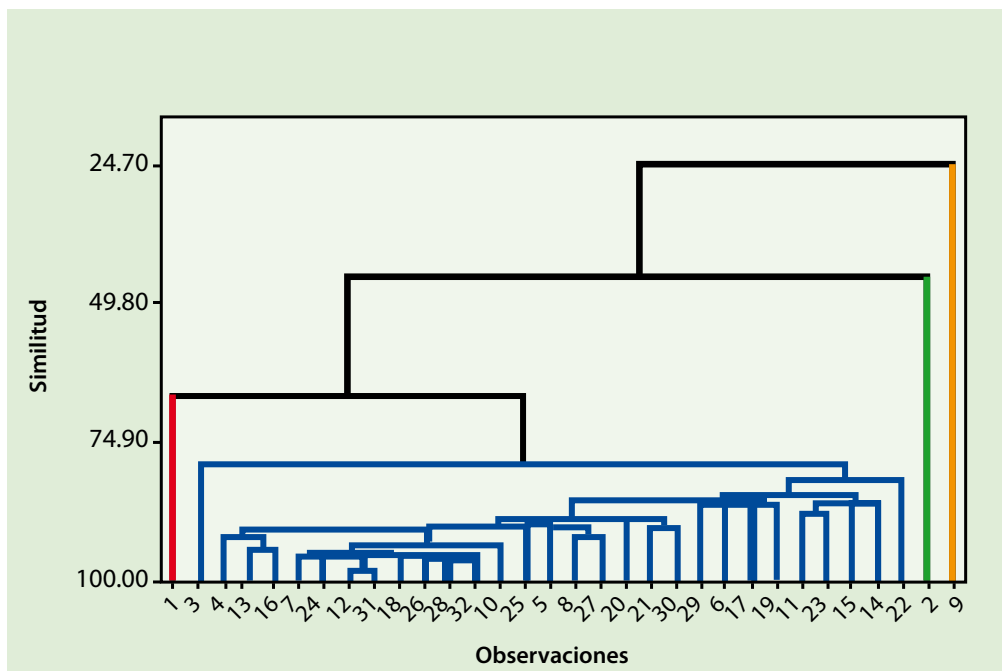
3.3 Empleo del AC

En la gráfica 4 se observa el *dendograma* de asociación mediante enlace simple y distancia euclidiana de las entidades federativas, en el cual se distinguieron cuatro conglomerados o colecciones de observaciones. En el primer grupo, de manera aislada, se encuentra el DF; en el segundo, Baja California; en el tercero, Aguascalientes; y en el cuarto, el resto de entidades federativas. Esto respalda, hasta cierto punto, las afirmaciones comentadas en la sección previa del ACP.

En la gráfica 5 se muestran los cuatro grupos de variables, en las cuales se observan tres conglomerados con una sola variable cada uno de ellos: la 20, que corresponde a la *proporción de sitios de disposición final tipo tiradero a cielo abierto reportados como destino de los residuos sólidos*

Gráfica 4

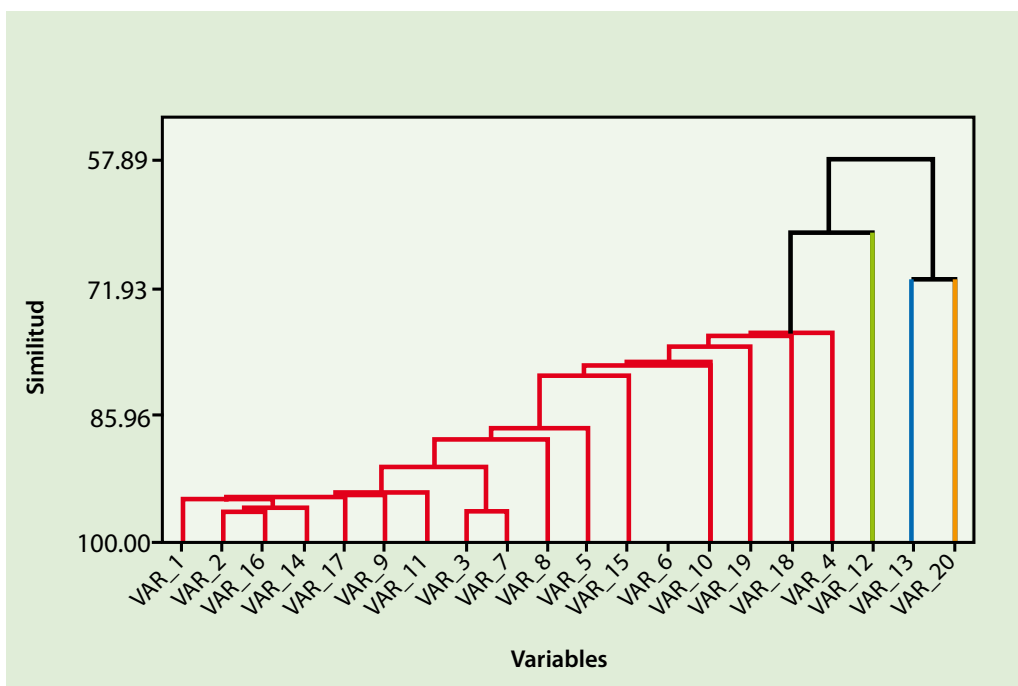
Dendrograma
Enlace simple, distancia euclidiana



Fuente: elaboración propia con resultados generados en Minitab®.

Gráfica 5

Dendrograma
Enlace simple, distancia de coeficiente de correlación



Fuente: elaboración propia con resultados generados en Minitab®.

urbanos respecto al total en la entidad federativa; la 13, que es la *proporción de municipios y delegaciones con disponibilidad de servicios de sólo recolección y disposición final de residuos sólidos urbanos respecto al total en la entidad federativa*; y la 12, asociada a la *proporción del total de municipios y delegaciones con disponibilidad de servicios relacionados con los residuos sólidos urbanos respecto al total en la entidad federativa*. El cuarto grupo aglutina a todas las demás; esto es, hasta cierto punto, coincidente con el resultado del ACP, en el cual se contrasta la *infraestructura de gestión de los RSU, salvo disposición final*, que parece marcar una de las diferencias más importantes en este conjunto de datos.

4. Conclusiones, reflexiones y comentarios

Del resultado de la aplicación del ACP y del AC se observa que hay dos dimensiones fundamentales en la gestión de la basura, esto es, la *infraestructura de gestión de los RSU, salvo disposición final*, por un lado, y la *planeación en contraste con la operación en la gestión de los RSU*, por el otro. Por supuesto, la designación de estos nombres de dimensiones fue una decisión por parte de los autores y podría estar sujeta a otro tipo de reflexiones de otros investigadores; sin embargo, lo destacable aquí es que de un fenómeno de 20 dimensiones se generan dos más que son combinaciones lineales de las primeras y que resultan ser las que marcan de manera preponderante el fenómeno de los RSU, con lo cual se simplifica de manera conceptual su comprensión.

A su vez, el hecho de que las entidades federativas se hubiesen contrastado bajo conglomerados en el conjunto formado por el DF, Baja California y Aguascalientes por un lado y las demás entidades

federativas por el otro, nos dice que, a pesar de que se han dado avances en la gestión de los RSU (básicamente en estas tres entidades referidas), existen áreas de oportunidad muy importantes, sobre todo en las del sur del país.

El análisis por supuesto podría no parar aquí y sería posible aplicar otro tipo de herramientas estadísticas que podrían complementar el mostrado en este documento.

Fuentes

- Bivand, R., E. Pebesma y V. Gómez-Rubio. *Applied Spatial Data Analysis with R*. Second ed. New York, Springer, 2013.
- INEGI. "Acuerdo por el que se aprueba la inclusión al Catálogo Nacional de Indicadores de un conjunto de indicadores clave en materia de emisiones y residuos", en: *Diario Oficial de la Federación*. México, DF, 2014a. Consultado en http://www.dof.gob.mx/nota_detalle.php?codigo=5367561&fecha=07/11/2014
- _____. *Censo Nacional de Gobiernos Municipales y Delegacionales 2013. Módulo 6. Residuos sólidos urbanos*. México, INEGI, 2014b. Consultado en <http://www3.inegi.org.mx/sistemas/temas/default.aspx?s=est&c=21385>
- Johnson, D. *Métodos multivariados aplicados al análisis de datos*. México, DF, International Thomson, 2000.
- Johnson, R. y D. Wichern. *Applied Multivariate Statistical Analysis*. Fifth ed. New Jersey, United States of America, 2002.
- Manly, B. *Multivariate Statistical Methods. A Primer*. London, UK, Chapman & Hall, 1984.
- Secretaría de Gobernación. "Ley General para la Prevención y Gestión Integral de los Residuos", en: *Diario Oficial de la Federación*. México, DF, 4 de junio de 2014. Consultado en http://www.diputados.gob.mx/LeyesBiblio/pdf/263_040614.pdf
- Secretaría de Medio Ambiente y Recursos Naturales (SEMARNAT). *Guía metodológica para el levantamiento y evaluación de información para efectuar el diagnóstico de la gestión integral de los RSU*. México, SEMARNAT, 2007.
- _____. *Informe de la situación del medio ambiente, 2012. Compendio de estadísticas ambientales. Indicadores clave y de desempeño ambiental*. México, SEMARNAT, 2012. Consultado en http://app1.semarnat.gob.mx/dgeia/informe_12/pdf/Cap7_residuos.pdf

Datos de variables de proporción utilizadas en el análisis

EF	TM	VAR_1	VAR_2	VAR_3	VAR_4	VAR_5	VAR_6	VAR_7	VAR_8	VAR_9	VAR_10	VAR_11	VAR_12	VAR_13	VAR_14	VAR_15	VAR_16	VAR_17	VAR_18	VAR_19	VAR_20
Ags.	11	0.091	0.015	0.182	0.182	0.818	0.182	0.273	0.035	0.009	0.003	0.025	1.000	0.909	0.091	0.013	0.003	0.000	0.636	0.091	0.000
BC	5	0.000	0.000	1.200	0.000	0.000	0.000	0.600	0.053	0.030	0.004	0.038	1.000	1.000	0.000	0.039	0.000	0.189	0.600	1.000	0.000
BCS	5	0.200	0.001	0.000	0.000	0.400	0.200	0.000	0.000	0.009	0.002	0.000	1.000	1.000	0.000	0.000	0.019	0.000	0.400	0.400	1.200
Camp.	11	0.000	0.000	0.000	0.000	0.182	0.000	0.000	0.000	0.005	0.003	0.000	1.000	1.000	0.000	0.005	0.010	0.000	0.364	0.091	0.909
Coah.	38	0.000	0.000	0.000	0.000	0.000	0.026	0.000	0.000	0.024	0.010	0.041	1.000	1.000	0.000	0.035	0.009	0.000	0.553	0.132	0.737
Col.	10	0.000	0.000	0.000	0.000	0.200	0.100	0.000	0.000	0.005	0.005	0.000	1.000	1.000	0.000	0.009	0.005	0.000	0.300	0.200	0.400
Chis.	118	0.025	0.011	0.025	0.000	0.008	0.025	0.025	0.027	0.021	0.028	0.027	0.975	0.949	0.025	0.012	0.035	0.016	0.102	0.051	0.856
Chih.	67	0.015	0.001	0.045	0.000	0.015	0.090	0.045	0.027	0.039	0.016	0.026	1.000	0.955	0.045	0.064	0.017	0.002	0.313	0.269	0.701
DF	16	0.875	0.480	0.813	0.000	1.000	0.188	0.750	0.115	0.184	0.102	0.329	1.000	0.000	1.000	0.038	0.298	0.537	0.375	0.000	0.000
Dgo.	39	0.000	0.000	0.051	0.000	0.077	0.026	0.026	0.018	0.009	0.015	0.008	1.000	1.000	0.000	0.013	0.006	0.000	0.103	0.179	0.744
Gto.	46	0.109	0.007	0.000	0.000	0.174	0.065	0.000	0.000	0.027	0.064	0.123	1.000	0.913	0.087	0.061	0.019	0.024	0.543	0.391	0.522
Gro.	81	0.025	0.007	0.012	0.000	0.012	0.062	0.012	0.009	0.015	0.031	0.017	1.000	0.988	0.012	0.017	0.034	0.000	0.062	0.025	0.963
Hgo.	84	0.095	0.021	0.000	0.000	0.143	0.143	0.024	0.018	0.021	0.041	0.011	1.000	0.976	0.024	0.010	0.021	0.050	0.333	0.083	0.845
Jal.	125	0.256	0.046	0.104	0.008	0.232	0.160	0.112	0.124	0.077	0.069	0.036	1.000	0.888	0.112	0.070	0.074	0.028	0.656	0.256	0.608
Méx.	125	0.048	0.023	0.080	0.000	0.016	0.040	0.104	0.133	0.121	0.105	0.049	0.984	0.952	0.032	0.155	0.083	0.037	0.544	0.224	0.488
Mich.	113	0.115	0.038	0.009	0.018	0.088	0.150	0.053	0.027	0.022	0.092	0.044	1.000	0.929	0.071	0.036	0.042	0.008	0.292	0.062	0.912
Mor.	33	0.121	0.005	0.303	0.000	0.030	0.030	0.303	0.088	0.012	0.025	0.009	1.000	0.848	0.152	0.013	0.014	0.000	0.091	0.121	0.424
Nay.	20	0.000	0.000	0.000	0.000	0.100	0.150	0.000	0.000	0.011	0.004	0.000	1.000	1.000	0.000	0.006	0.021	0.000	0.150	0.050	0.850

Datos de variables de proporción utilizadas en el análisis

EF	TM	VAR_1	VAR_2	VAR_3	VAR_4	VAR_5	VAR_6	VAR_7	VAR_8	VAR_9	VAR_10	VAR_11	VAR_12	VAR_13	VAR_14	VAR_15	VAR_16	VAR_17	VAR_18	VAR_19	VAR_20
NL	51	0.039	0.002	0.118	0.000	0.000	0.020	0.118	0.053	0.037	0.012	0.002	1.000	0.784	0.216	0.068	0.008	0.000	0.275	0.157	0.471
Oax.	570	0.116	0.078	0.007	0.000	0.023	0.086	0.005	0.035	0.027	0.135	0.025	0.847	0.802	0.046	0.013	0.030	0.008	0.070	0.012	0.704
Pue.	217	0.028	0.045	0.028	0.000	0.018	0.051	0.037	0.071	0.042	0.027	0.038	0.986	0.963	0.023	0.085	0.017	0.000	0.051	0.037	0.387
Qro.	18	0.278	0.075	0.167	0.000	0.167	0.056	0.167	0.027	0.020	0.013	0.009	1.000	0.778	0.222	0.024	0.004	0.011	0.333	0.500	0.389
Q. Roo	10	0.200	0.007	0.100	0.000	0.200	0.200	0.100	0.009	0.013	0.006	0.000	1.000	1.000	0.000	0.030	0.003	0.007	0.600	0.400	0.400
SLP	58	0.086	0.020	0.017	0.000	0.017	0.052	0.017	0.009	0.016	0.019	0.027	0.983	0.983	0.000	0.022	0.019	0.000	0.086	0.086	0.793
Sin.	18	0.056	0.004	0.000	0.000	0.222	0.056	0.000	0.000	0.024	0.012	0.007	1.000	1.000	0.000	0.043	0.010	0.003	0.222	0.278	0.778
Son.	72	0.028	0.002	0.028	0.014	0.014	0.083	0.028	0.027	0.027	0.020	0.031	1.000	0.986	0.014	0.034	0.008	0.002	0.125	0.111	0.931
Tab.	17	0.000	0.000	0.059	0.000	0.000	0.000	0.059	0.009	0.019	0.013	0.008	1.000	0.941	0.059	0.001	0.033	0.009	0.353	0.118	0.765
Tamps.	43	0.000	0.000	0.023	0.000	0.093	0.047	0.023	0.009	0.033	0.004	0.000	0.953	0.930	0.023	0.036	0.029	0.001	0.116	0.116	0.860
Tlax.	60	0.033	0.002	0.000	0.000	0.017	0.100	0.000	0.000	0.012	0.009	0.005	1.000	1.000	0.000	0.004	0.016	0.000	0.083	0.017	0.050
Ver.	212	0.061	0.106	0.028	0.000	0.028	0.094	0.028	0.062	0.057	0.052	0.033	0.962	0.934	0.028	0.022	0.087	0.065	0.142	0.085	0.566
Yuc.	106	0.000	0.000	0.019	0.000	0.019	0.085	0.009	0.018	0.012	0.044	0.021	1.000	0.981	0.019	0.016	0.006	0.000	0.066	0.047	0.981
Zac.	58	0.017	0.001	0.000	0.000	0.052	0.103	0.000	0.000	0.017	0.014	0.011	0.983	0.966	0.017	0.006	0.019	0.001	0.138	0.121	0.862
Sum.	2 457	0.079	1.000	0.039	0.002	0.059	0.082	0.042	1.000	1.000	1.000	1.000	0.956	0.908	0.048	1.000	1.000	1.000	0.198	0.097	0.669

Nota: se presentan los valores de las proporciones de las variables consideradas en el presente estudio a tres dígitos, aunque el análisis fue llevado a cabo con 12 dígitos de precisión, por lo cual al reproducir el ejercicio es posible que surjan pequeñas discrepancias.

Fuente: elaboración propia con datos provenientes de INEGI (2014).

Tabla de correlaciones de las variables de proporción bajo análisis

	VAR_1	VAR_2	VAR_3	VAR_4	VAR_5	VAR_6	VAR_7	VAR_8	VAR_9	VAR_10	VAR_11	VAR_12	VAR_13	VAR_14	VAR_15	VAR_16	VAR_17	VAR_18	VAR_19	VAR_20
VAR_1	1.00	<u>0.40</u>	<u>0.45</u>	0.01	<u>0.75</u>	<u>0.51</u>	<u>0.66</u>	0.09	0.14	0.07	0.26	0.04	<u>-0.90</u>	<u>0.91</u>	0.00	0.24	<u>0.40</u>	0.25	<u>-0.02</u>	<u>-0.37</u>
VAR_2	<u>0.40</u>	1.00	0.17	-0.04	0.24	0.15	0.25	<u>0.92</u>	<u>0.95</u>	<u>0.92</u>	<u>0.98</u>	-0.22	<u>-0.43</u>	<u>0.40</u>	<u>0.88</u>	<u>0.98</u>	<u>0.98</u>	-0.04	-0.15	-0.15
VAR_3	<u>0.45</u>	0.17	1.00	0.04	0.34	-0.04	<u>0.93</u>	0.05	0.04	-0.03	0.12	0.14	<u>-0.49</u>	<u>0.51</u>	-0.03	0.07	0.33	0.34	<u>0.57</u>	-0.62
VAR_4	0.01	-0.04	0.04	1.00	<u>0.54</u>	0.33	0.19	-0.02	-0.05	-0.05	-0.03	0.08	0.00	0.02	-0.04	-0.05	-0.05	0.33	-0.09	-0.34
VAR_5	<u>0.75</u>	0.24	0.34	<u>0.54</u>	1.00	<u>0.61</u>	<u>0.57</u>	-0.01	0.03	-0.04	0.15	0.18	<u>-0.66</u>	<u>0.69</u>	-0.08	0.12	0.26	<u>0.41</u>	-0.04	-0.40
VAR_6	<u>0.51</u>	0.15	0.04	0.33	<u>0.61</u>	1.00	0.12	0.01	0.03	0.03	0.08	0.07	-0.26	0.27	-0.03	0.09	0.11	0.26	-0.11	-0.07
VAR_7	<u>0.66</u>	0.25	0.04	0.19	<u>0.57</u>	0.12	1.00	0.07	0.07	-0.01	0.17	0.16	<u>-0.71</u>	<u>0.74</u>	-0.03	0.13	<u>0.37</u>	<u>0.36</u>	<u>0.36</u>	-0.68
VAR_8	0.09	<u>0.92</u>	0.05	0.07	0.07	0.01	0.07	1.00	<u>0.99</u>	<u>0.98</u>	<u>0.95</u>	-0.21	-0.12	0.08	<u>0.99</u>	<u>0.97</u>	<u>0.90</u>	-0.02	-0.07	-0.07
VAR_9	0.14	<u>0.95</u>	0.04	-0.05	0.03	0.03	0.07	<u>0.99</u>	1.00	<u>0.98</u>	<u>0.98</u>	-0.21	-0.17	0.13	<u>0.98</u>	<u>0.99</u>	<u>0.93</u>	-0.02	-0.09	-0.05
VAR_10	0.07	<u>0.92</u>	-0.03	-0.05	-0.04	0.03	-0.01	<u>0.98</u>	<u>0.98</u>	1.00	0.96	-0.30	-0.10	0.05	<u>0.98</u>	<u>0.97</u>	<u>0.89</u>	-0.07	-0.12	0.01
VAR_11	0.26	<u>0.98</u>	0.12	-0.03	0.15	0.08	0.17	<u>0.95</u>	<u>0.98</u>	0.96	1.00	-0.18	-0.29	0.26	<u>0.94</u>	<u>0.99</u>	<u>0.96</u>	-0.01	-0.10	-0.11
VAR_12	0.04	-0.22	0.14	0.08	0.18	0.07	0.16	-0.21	-0.21	-0.30	-0.18	1.00	0.08	0.08	-0.19	-0.21	-0.14	0.33	0.23	-0.10
VAR_13	<u>-0.90</u>	<u>-0.43</u>	<u>-0.49</u>	0.00	<u>-0.66</u>	-0.26	<u>-0.71</u>	-0.12	-0.17	-0.17	-0.10	0.08	1.00	<u>-0.99</u>	-0.03	-0.27	<u>-0.44</u>	-0.08	0.16	0.42
VAR_14	<u>0.91</u>	<u>0.40</u>	<u>0.51</u>	0.02	<u>0.69</u>	0.27	<u>0.74</u>	0.08	0.13	0.05	0.26	0.08	<u>-0.99</u>	1.00	0.00	0.23	<u>0.41</u>	0.13	-0.12	-0.44
VAR_15	0.00	<u>0.88</u>	-0.03	-0.04	-0.08	-0.03	-0.03	0.00	0.00	0.00	-0.19	0.08	-0.03	0.00	1.00	<u>0.95</u>	<u>0.86</u>	-0.02	-0.04	-0.02

Anexo 2

Tabla de correlaciones de las variables de proporción bajo análisis

	VAR_1	VAR_2	VAR_3	VAR_4	VAR_5	VAR_6	VAR_7	VAR_8	VAR_9	VAR_10	VAR_11	VAR_12	VAR_13	VAR_14	VAR_15	VAR_16	VAR_17	VAR_18	VAR_19	VAR_20	Concluye
VAR_16	0.24	<u>0.98</u>	0.07	-0.05	0.12	0.09	0.13	<u>0.97</u>	<u>0.99</u>	<u>0.97</u>	<u>0.99</u>	-0.21	-0.27	0.23	<u>0.95</u>	1.00	<u>0.96</u>	-0.04	-0.14	-0.06	
VAR_17	<u>0.40</u>	<u>0.98</u>	0.33	-0.05	0.26	0.11	<u>0.37</u>	<u>0.90</u>	<u>0.93</u>	<u>0.89</u>	<u>0.96</u>	-0.14	-0.44	<u>0.41</u>	<u>0.86</u>	<u>0.96</u>	1.00	0.04	-0.02	-0.22	
VAR_18	0.25	-0.04	0.34	0.33	<u>0.41</u>	0.26	<u>0.36</u>	-0.02	-0.02	-0.07	-0.01	0.33	-0.08	0.13	-0.02	-0.04	0.04	1.00	<u>0.55</u>	-0.36	
VAR_19	-0.02	-0.15	<u>0.57</u>	-0.09	-0.04	-0.11	<u>0.36</u>	-0.07	-0.09	-0.12	-0.10	0.23	0.16	-0.12	-0.04	-0.14	-0.02	<u>0.55</u>	1.00	-0.28	
VAR_20	<u>-0.37</u>	-0.15	<u>-0.62</u>	-0.34	<u>-0.40</u>	-0.07	<u>-0.68</u>	-0.07	-0.05	0.01	-0.11	-0.10	<u>0.42</u>	<u>-0.44</u>	-0.02	-0.06	-0.22	<u>-0.36</u>	-0.28	1.00	

Nota: presenta las correlaciones entre las variables de proporción; aquellos valores que han sido subrayados presentan un valor p menor a 0.05.

Fuente: elaboración propia con resultados generados en Minitab® y R.

Desigualdades entre entidades en materia de tecnologías de información y comunicación en México

Wilfrido Ruiz Ochoa

General Views Of Shoppers During El Buen Fin/ Mexico's Black Friday/Bloomberg/Getty Images



En el documento se propone y construye un índice estatal de desarrollo de las tecnologías de información y comunicación (TIC) como instrumento para el seguimiento a nivel de entidad de la estrategia digital nacional. A manera de validación preliminar, se revisa su consistencia con el comportamiento competitivo y la dinámica de innovación de las entidades federativas.

Entre los estados más rezagados se observan problemas serios en materia de infraestructura y capital humano para el desarrollo de estas tecnologías. Por su parte, el posicionamiento competitivo y la capacidad de innovación de las entidades punteras podrían potenciarse a partir de una mayor inversión en capital humano y penetración de las TIC en ámbitos estratégicos.

Palabras clave: brecha digital, competitividad regional, innovación tecnológica.

Recibido: 15 de septiembre de 2014

Aceptado: 28 de noviembre de 2014

Introducción y objetivos

Las tecnologías de información y comunicación (TIC) representan un grupo convergente de tecnologías vinculadas con la microelectrónica, la informática, las telecomunicaciones, la televisión, la radio y la optoelectrónica (Castells, 1999) que se han constituido en pieza clave no sólo para el impulso de la productividad, sino también para la planificación, organización y gestión de procesos. Su influencia sobre el desarrollo económico y bienestar social resulta innegable,¹ siendo su accesibilidad y manejo elemental para ascender por la senda de la competitividad y el crecimiento, así como para mejorar las condiciones de vida. Han trastocado prácticamente todos los ámbitos de desenvolvimiento del ser humano, por lo que su impacto ha resultado transversal, multidimensional, creciente y sinérgico, contribuyendo a reforzar el proceso globalizador en sus distintas vertientes.

¹ Entre los que han abordado estas cuestiones, se encuentran Katz *et al.* (2013), De la Fuente (2008) y Valderrama y Castillo (2011). En International Telecommunications Union (ITU, 2012) se presentan estudios de caso para diversos países, donde se mide el impacto de las TIC sobre el producto interno bruto (PIB) y el crecimiento económico.

In this document is proposed a state index for information and communication technologies development (TICs), as a tool for monitoring at state level the National Digital Strategy. The consistency of results is reviewed, through the relationship between the index and competitive and innovation state patterns.

It is found that the reasons for regional backwardness are problems related to the lack of infrastructure and human capital for development of TICs among the states. Meanwhile, competitive positioning and innovation capacity of the leading states could be improved, increasing investment in human capital and TICs penetration in strategic areas.

Keywords: digital divide, regional competitiveness, technological innovation.

A pesar de su exponencial crecimiento y difusión, el rezago que vienen arrastrando algunos países del tercer mundo en materia de las TIC es enorme, lo cual los ha dejado fuera del mapa de las comunicaciones y de las oportunidades derivadas de las aplicaciones de estas tecnologías. En este sentido, México es un país en transición, en relación con el avance que ha logrado el resto del mundo. Katz, Koutroumpis y Callorda (2013) lo ubican así a partir de la posición que ocupa en el índice de digitalización (ID) que han elaborado.²

Cuando se integran diversos indicadores en forma simultánea para valorar el grado de desarrollo de las TIC —como la infraestructura para su uso, la intensidad con la que se utilizan y las capacidades de la población para acceder a ellas y manejarlas—, se encuentra que el índice de desarrollo de las TIC (IDT), desarrollado por la Unión Internacional de Telecomunicaciones (2013:8), ubicó en el 2012 a México en el lugar número 83 respecto a 121 países.

² En Katz *et al.* (2013) no se muestra la posición puntual de México respecto a los 184 países para los que se calculó el ID. Siguiendo una escala de 0 a 100, este indicador permite, sin embargo, clasificar los países en: limitados (ID menor a 20), emergentes (ID entre 20 y 35), transicionales (ID entre 35 y 50) y avanzados (con un ID mayor a 50).

Entre las 10 naciones latinoamericanas más pobladas, nuestro país sólo supera a Perú. En este sentido, se encuentran más desarrollados en materia de TIC Chile (51), Argentina (53) y Colombia (77). Más de 60 lugares separan a nuestro país de Estados Unidos de América (EE.UU.) y Canadá, que ocupan los lugares 17 y 20, en ese orden.

Pero no sólo en el plano internacional se observa que México se encuentra rezagado en materia de TIC. En lo interno, y en un sentido amplio, la brecha digital³ mexicana resulta sumamente desigual entre regiones y entidades; además —como señalan Katz *et al.* (2013)—, enfrenta como principales problemas: brecha de demanda en banda ancha,⁴ baja tasa de innovación digital, reducido nivel de competencia en telecomunicaciones y altos costos de fricción en el ámbito regulatorio.

Frente a lo anterior, la estrategia digital del actual gobierno plantea que México deberá llegar a la media del índice de digitalización de los países que forman la Organización para la Cooperación y el Desarrollo Económicos (OCDE), el cual fue de 59.2 puntos porcentuales en 2011 (GR, 2013). El reto es enorme, dado que los mexicanos se ubican en último lugar entre las naciones de esta organización con una puntuación de 37.05 que, aunque es ligeramente superior al promedio latinoamericano (32), se encuentra muy por debajo al que alcanzaron EE.UU. (62) y Canadá (61) en 2011. Según la estrategia digital nacional, la meta para el 2018 es que nuestro país alcance un ID semejante al de Chile (45.33).

Seguramente, una de las dificultades para avanzar en el acortamiento de la brecha digital inter-

3 De acuerdo con Eurostat, el concepto de brecha digital se refiere a las desigualdades entre aquellas comunidades que tienen accesibilidad a internet y pueden hacer uso del mismo y las que no (EC, 2014b). Pero en una concepción más amplia, tales desigualdades consideran tanto el desarrollo y acceso a todas las nuevas tecnologías de la información y la comunicación como a la capacidad para utilizarlas. Concebida en este último sentido, esta brecha no sólo se produce entre países, también entre hogares, regiones, segmentos socioeconómicos y sectores económicos (ALADI, 2003).

4 El concepto de banda ancha acepta diversas definiciones dependiendo de los umbrales de velocidad de transmisión de información que permite. De acuerdo con la *Ley Federal de Telecomunicaciones*, se entiende por la misma "...el acceso de alta capacidad que permite ofrecer diversos servicios convergentes a través de infraestructura de red fiable, con independencia de las tecnologías empleadas, cuyos parámetros son actualizados periódicamente..." (GR, 2014:1).

nacional de nuestro país reside en que la política federal no dispone aún de una estrategia regional diferenciada que incorpore al diseño de la misma condicionantes vinculadas con las profundas diferencias regionales que existen en el país (Ruiz Ochoa, 2008). Además, dentro de los ámbitos en los que pretende fomentar el desarrollo de las tecnologías de la información, no se han establecido medidas para potenciar su impacto sobre la capacidad de innovación y el crecimiento. Ésta es una asignatura que requiere atenderse, pues las TIC podrían formar parte de los múltiples habilitadores que se necesitan para hacer posible el desarrollo de eficientes sistemas regionales de innovación, sobre lo cual se requiere avanzar para afrontar el evidente rezago del país. En el 2011, por ejemplo, el Foro Económico Mundial (WEF, 2011) colocó a México en la posición 81 respecto a 142 países en el *ranking* mundial de la innovación.

Dicho lo anterior, este trabajo tiene como objetivo elaborar un índice de desarrollo de las tecnologías de información y comunicación (IDTIC) a nivel estatal que, al ser desagregable en sus componentes, permita identificar por entidad federativa las debilidades y fortalezas en la materia. Con ello, se pretende proponer un instrumento de seguimiento del desarrollo de las TIC por entidad, que pudiera considerarse de manera eventual para valorar las metas de la estrategia digital nacional. El ámbito de análisis de este índice es el social.

Enfoque metodológico

De acuerdo con el *Partnership* para la medición del desarrollo de las TIC —que es una comisión auspiciada por la Organización de las Naciones Unidas (ONU)—, las estadísticas disponibles cruciales para formular políticas y estrategias concernientes al crecimiento y desarrollo posibilitado por estas tecnologías son muy limitadas, en especial en los países en desarrollo (*Partnership*, 2010).

No obstante, a nivel global se cuenta con diversos esfuerzos y experiencias en la elaboración de índices internacionales sobre el grado de desarrollo de las tecnologías de la información; entre ellos, desta-

can cuatro: *Networked Readiness Index* (NRI) del Foro Económico Mundial (WEF, 2013), el índice de digitalización (ID) de Katz *et al.* (2013), el *e-readiness rankings* (ERR) del *Economist Intelligence Unit* (EIU, 2011) (que perdió continuidad a partir del 2014) y el IDT (ya citado) de la ITU (2013).⁵

Los índices NRI y ERR son los más completos en términos del número de variables que consideran: 53 y 39 respectivamente; ambos incluyen componentes de entorno; el primero contempla el ambiente regulatorio, político, innovador y de negocios propicio para el desarrollo de las TIC, en tanto que el ERR considera el ambiente social, cultural, educativo, legal y gubernamental. Algunos de ellos —que se ponderan en menor medida en comparación con los directamente relacionados con el desarrollo de las TIC—, son difíciles de cuantificar a escala subnacional, además de que para determinadas variables no siempre aplica la diferenciación a ese nivel.

El índice de digitalización de Katz *et al.* —utilizado en la actualidad por el gobierno de México para medir los avances de la estrategia digital nacional— y el IDT del ITU incluyen un menor número de variables (23 y 11, en ese orden) que se vinculan en forma más directa con el desarrollo de las TIC e integran como único componente de entorno las habilidades y el capital humano. En la formulación del índice de Katz *et al.* se encuentran indicadores que no están disponibles por entidad, como el costo y las tarifas para acceder a los servicios de telefonía e internet, o la velocidad e inversión en banda ancha.

A escala subnacional se dispone de un menor número de experiencias en la elaboración de índices sobre desarrollo de las TIC. En el caso de México, por ejemplo, se han elaborado, en el mejor de los casos, subíndices sobre la penetración de estas tecnologías como componentes de la capacidad de innovación (Ruiz, 2008 y FCCyT, 2012). Dada esta restricción, se decidió partir de las variables contempladas por

⁵ Además, se han elaborado el *Global Cloud Scorecard* (BSA, 2013), un índice sobre el desarrollo del cómputo en la nube y el *Digital Opportunity Index* (DOI) de la ITU (2008), que valora el grado de adopción de nuevas tecnologías vinculadas con la banda ancha y el acceso a internet mediante dispositivos móviles. Ambos se descartan como referentes por no reflejar el desarrollo integral de todas las TIC.

el IDT, al cual se le agregaron otras más que son reconocidas como clave para las TIC por parte del *Partnership* (2010).

Así, el IDTIC se descompone en un conjunto de variables que reflejan la infraestructura para el acceso a estas tecnologías y la disponibilidad y calidad del capital humano para manejarlas, formando, en conjunto, el primer componente de este índice.

Se consideró, además, un componente de resultados que se integra por variables que muestran la intensidad de uso y grado de acceso a las TIC. Por último, y a diferencia de la metodología de la UIT, se agrega a este componente el grado de penetración en ámbitos estratégicos, que han sido propuestos en fecha reciente por la Comisión Económica para América Latina y el Caribe (CEPAL) para agregar indicadores complementarios sobre el cumplimiento de los Objetivos de Desarrollo del Milenio (Balboni *et al.*, 2011); entre ellos se encuentran la penetración de las TIC en el sector educativo, en la gestión gubernamental, en las empresas y en el sector salud. Este último no se considera en los hechos, debido a la carencia de información.

Componentes del desarrollo de las TIC

Se consideraron 25 variables, de las cuales 12 forman el componente de infraestructura y capital humano y 13, el de resultados, estando el índice balanceado en este sentido (ver cuadro 1). De

Cuadro 1
Componentes y factores que forman el IDTIC

Componentes y factores	Número de variables
Componentes de capital humano e infraestructura	
Habilidades y capital humano	4
Infraestructura para el uso de TIC	8
Componentes de resultados	
Penetración en áreas estratégicas	7
Intensidad de uso de las TIC	6
Total de variables	25

manera más detallada, en el cuadro 2 se muestran las variables del primer componente y en el cuadro 3, las del segundo, indicándose los factores y las fuentes de la información.

Se han incluido tanto variables relativizadas o promedio que, al tener en cuenta el peso poblacional u otra característica del estado referido, se expresan en términos per cápita, como variables flujo que muestran simplemente la distribución porcentual interestatal entre estados de una determinada variable. Con esto último se pretende incorporar, también, el grado de concentración de una determinada característica en el contexto interestatal.

Dentro del componente de infraestructura y capital humano, se visualizan como variables flujo el número total de teléfonos públicos por localidad, que pudieran ser accesibles a población procedente de diversas zonas geográficas, y la distribución interestatal de las viviendas particulares que disponen de todas las modalidades de TIC. Por ser su número limitado, interesa más su peso en el contexto interestatal.

Cuadro 2 Continúa
Variables que forman el componente de capital humano e infraestructura del índice estatal de desarrollo de las TIC

Factores y variables	Año	Fuente
Habilidades y capital humano para el aprovechamiento de las TIC*		
1. Tasa de alfabetización de la población de 15 años y más	2010	INEGI (2011)
2. Años promedio de escolaridad de la población de 15 años y más	2010	INEGI (2011)
3. Puntaje promedio de Pisa en ciencias, matemáticas y lectura	2009	INEE (2011)
4. Participación de la población de 15 a 55 años de edad con años aprobados en carreras, en la población total	2010	INEGI (2011)
Infraestructura para el uso de las TIC**		
5. Suscriptores de telefonía fija por cada 100 habitantes	2010	COFETEL (2013)
6. Suscriptores de telefonía celular por cada 100 habitantes	2010	COFETEL (2013)

Cuadro 2 Concluye
Variables que forman el componente de capital humano e infraestructura del índice estatal de desarrollo de las TIC

Factores y variables	Año	Fuente
7. Porcentaje de hogares que disponen de computadora	2010	INEGI (2011a)
8. Porcentaje de hogares que disponen de internet	2010	INEGI (2011a)
9. Porcentaje de hogares que disponen de televisión	2010	INEGI (2011a)
10. Porcentaje de hogares que disponen de radio	2010	INEGI (2011a)
11. Número promedio de teléfonos públicos por localidad	2010	COFETEL (2013)
12. Distribución interestatal de viviendas particulares habitadas que disponen de todas las TIC	2010	INEGI (2012)

* Factor contemplado, en lo general, por los cuatro índices de desarrollo de las TIC internacionales citados.

** Exceptuando las variables 11 y 12, las demás son consideradas como clave por el *Partnership* (2010).

Cuadro 3 Continúa
Variables que forman el componente de resultados del índice estatal de desarrollo de las TIC

Factores y variables	Año	Fuente
Penetración de las TIC en áreas estratégicas***		
13. Distribución interestatal de usuarios de computadora desde sitios públicos	2010	INEGI (2011a)
14. Usuarios de computadora con propósito escolar, por cada 100 alumnos registrados desde la primaria hasta el nivel superior	2010	INEGI (2011a)
15. Índice de gobierno electrónico estatal*	2011	Sandoval et al. (2011a)
16. Usuarios frecuentes de comercio electrónico, en relación con el flujo total de internautas que lo utilizan como medio de compra o pago	2010	INEGI (2011a)

Variables que forman el componente de resultados del índice estatal de desarrollo de las TIC

Factores y variables	Año	Fuente
17. Participantes que utilizan el comercio electrónico para compras en el extranjero, en relación con el flujo total de internautas que acceden a él para compras	2010	INEGI (2011a)
18. Distribución interestatal del comercio electrónico total	2010	INEGI (2011a)
19. Usuarios que utilizan la computadora en su trabajo, por cada 100 personas ocupadas	2010	INEGI (2011a)
Intensidad de uso de las tecnologías de la información***		
20. Usuarios de internet por cada 100 habitantes de 5 años y más	2010	INEGI (2011a)
21. Usuarios de computadoras por cada 100 habitantes de 5 y más	2010	INEGI (2011a)
22. Usuarios que utilizan internet de manera frecuente, por cada 100 habitantes de 5 años y más**	2010	INEGI (2011a)
23. Usuarios que utilizan computadoras de manera frecuente, por cada 100 habitantes de 5 años y más**	2010	INEGI (2011a)
24. Distribución interestatal de usuarios de internet	2010	INEGI (2011a)
25. Distribución interestatal de usuarios de internet, que acceden mediante dispositivos móviles	2010	INEGI (2011a)

* Refleja la evaluación de los 31 portales de los gobiernos estatales y del Distrito Federal, en cuanto a información, interacción, transacción, integración y participación (ver Luna *et al.*, 2011).

** Se consideran usuarios frecuentes quienes utilizan internet o las computadoras todos los días o, al menos, una vez a la semana.

*** Exceptuando las variables expresadas mediante distribución interestatal, el resto se contemplan en lo general como variables clave por el *Partnership* (2010).

En relación con las variables del componente de resultados, se consideró la penetración de las TIC en áreas estratégicas, como la distribución interestatal de usuarios de internet desde sitios públicos y de los de comercio electrónico. Además de estas variables flujo, se incorporan algunas que reflejan

de manera relativizada la presencia de las TIC en el ámbito laboral, comercial, educativo y en la función gubernamental. En cuanto a la intensidad de uso de las TIC, se incluye a los usuarios de internet y de computadoras, así como el acceso a la red mediante dispositivos móviles.

Como principal fuente de información se utilizaron los resultados de la Encuesta Nacional sobre la Disponibilidad y Uso de las Tecnologías de la Información en los Hogares (ENDUTIH) 2011, con representatividad a escala de entidad federativa (INEGI, 2011).⁶ De manera complementaria, se usó el Sistema de Información Estadística de Mercados de Telecomunicaciones (SIEMT) de la COFETEL (2013) y datos tanto del Censo de Población y Vivienda 2010 (INEGI, 2011) como del Instituto Nacional para la Evaluación de la Educación (INEE, 2011).

Puede observarse que la mayoría de las variables hacen referencia a datos del 2010, lo cual no le resta mérito al presente documento, pues las desigualdades regionales respecto a tecnologías de información y comunicación siguen patrones estructurales difíciles de transformar en el corto plazo.

Elaboración del índice

Para normalizar las variables se parte del método *Mini-Max* (ver Fuentes *et al.*, 2013), que consiste en lo siguiente:

$$S_i = \frac{X_i^j - \min X_i^j}{X_i^j \max - X_i^j \min} \quad (1)$$

donde S_i es una de las n variables normalizadas correspondientes al desempeño de la entidad i . A la variable X correspondiente al atributo j de la entidad i se le resta su mínimo en la serie de las 32 entidades analizadas. El resultado es dividido por la diferencia entre el valor máximo y mínimo de la serie. Este procedimiento se repite para toda variable j , es decir, $\forall j = 1, \dots, 32$.

⁶ A partir del 2010, el tamaño de la muestra de esta encuesta "...aumentó en la medida necesaria para generar resultados a nivel estatal..." (ENDUTIH, 2011 a:1).

La variable normalizada adquirirá mayor valor en la medida en que su distancia al máximo sea menor. De esta forma, se obtiene una escala comparable entre variables normalizadas acotada entre 0 y 1, cuya sumatoria se lleva a cabo de la siguiente manera: si partimos de que un conjunto de n variables forman un factor f del desarrollo de las TIC, entonces la sumatoria ponderada de las S_i variables que lo forman está dada por:

$$f_1^i = \sum_{i=1}^n \alpha_i S_i \quad (2)$$

El ponderador α se define como $1/n$ a efecto de que el factor 1 refleje el promedio de los valores que adquieren las variables que lo forman en la entidad i . Este procedimiento se repite para todas las variables que pertenezcan al factor analizado, manteniendo el mismo ponderador para cada una de ellas. Esto es: $\forall i \in f_1, \alpha_1 = \alpha_2 \dots = \alpha_n$. Ahora bien, debido a que el índice está formado por dos componentes, se elaboran dos subíndices para cada uno de ellos de la siguiente forma:

$$I_g^i = \sum_{m=1}^k \gamma_m^i f_m^i \quad (3)$$

El subíndice g , referido al primer o segundo componente de una entidad i , estará definido en la ecuación (3) como el promedio de los valores que adquieren los k factores considerados. Como en los casos anteriores, los ponderadores son iguales: $\forall i \in f_m$ y $\gamma_1 = \gamma_2 = \dots = \gamma_m$.

Finalmente, el índice de desarrollo de las TIC para una entidad i agrega los subíndices de cada componente de la siguiente forma:

$$I_t^i = \emptyset I_1^i + \delta I_2^i \quad (4)$$

ponderando el primero con \emptyset y el segundo con δ ; esta ponderación puede ser igual para cada componente o diferenciada. En el presente trabajo se optó por el primer tipo de ponderación, marcando en este sentido una diferencia con el IDT.

Posicionamiento estatal en materia de desarrollo de las TIC

A partir de la distribución interestatal del IDTIC calculado y siguiendo la regla de los cuartiles, se clasificaron las entidades en niveles bajo, medio-bajo, medio-alto y alto en cuanto a desarrollo de las TIC. En el cuadro 4 puede observarse que 47% de la población de México reside en entidades que se encuentran en niveles bajo y medio-bajo de desarrollo relativo de las TIC, por encontrarse por debajo de la media nacional en el IDTIC. Las mayores disparidades se encuentran en los extremos de la distribución de este índice, entre los estratos de bajo y muy alto desarrollo.

Cuadro 4

Estratos del grado de desarrollo de las tecnologías de información entre las entidades mexicanas y representatividad poblacional

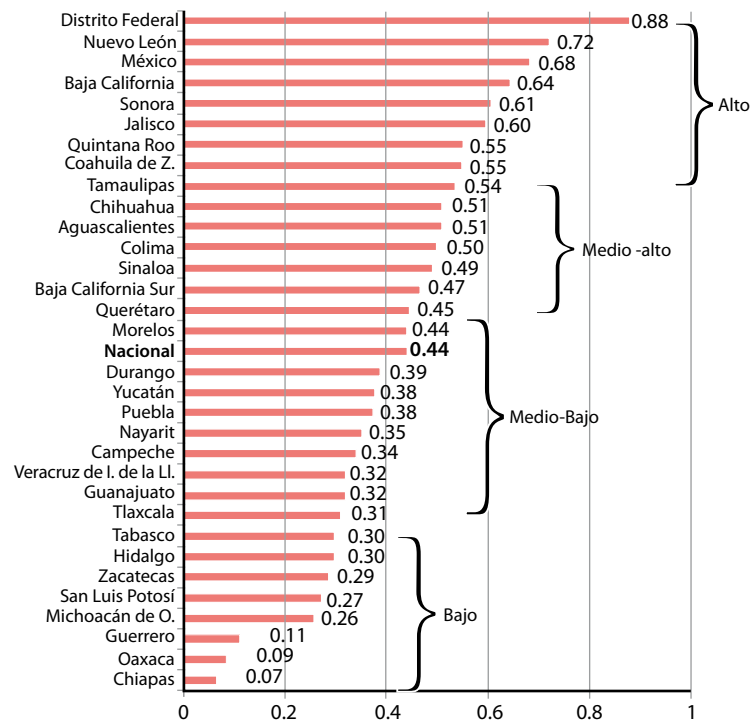
Estratos de desarrollo de las TIC*	Rango del IDTIC	Promedio del IDTIC	Varianza del IDTIC	Porcentaje de población
Bajo	0.07 a 0.30	0.21	0.010	22.61
Medio-bajo	0.31 a 0.44	0.36	0.002	24.41
Medio-alto	0.45 a 0.54	0.49	0.001	12.16
Alto	0.55 a 0.88	0.65	0.012	40.81
Nacional	0.07 a 0.88	0.44	0.034	100

* Los estratos fueron definidos a partir de la identificación de los cuartiles.

Fuente: elaboración propia con base en el comportamiento interestatal del IDTIC.

En la gráfica 1 se muestra un ordenamiento de los estados de mayor a menor desarrollo de las TIC. Resulta clara la preponderancia de la capital y el estado de México, que son las entidades más pobladas y con mayor desarrollo de estas tecnologías. Pero no tan lejos de ellas se encuentran las seis que colindan con EE.UU., en particular Nuevo León, Baja California, Sonora y Coahuila de Zaragoza, que se considera cuentan con un desarrollo relativamente alto de TIC.

Posición de las entidades federativas en el ID de las TIC, 2010



Fuente: elaboración propia.

La franja norte, junto con el occidente (centrado en Jalisco) y el centro concentran una proporción importante de la industria del país que, seguramente, influye en los requerimientos de tecnologías de información y comunicación. En contraste, la región sur —formada por Chiapas, Oaxaca y Guerrero— es la que presenta el menor nivel de desarrollo de TIC, lo cual coincide también con el hecho de que son las entidades que han alcanzado el menor nivel de ingreso per cápita relativo (Ruiz Ochoa, 2008).

Análisis de debilidades y fortalezas

El índice nos permite descomponerlo en los cuatro factores que lo forman y revisar sus características diferenciadas. En el cuadro 5 puede observarse, por ejemplo, que el promedio del capital humano disponible para desarrollar estas tecnologías representa el factor del IDTIC que alcanza el mayor promedio a nivel nacional, cuestión que se acentúa entre los estratos alto y medio-alto, lo

Cuadro 5

Valores promedio y grado de concentración de los factores del IDTIC según estrato de desarrollo de las TIC

Estratos/ factores	Capital humano	Infraestructura	Penetración	Intensidad de uso	IDTIC
Promedio					
Bajo	0.29	0.22	0.22	0.11	0.21
Medio-bajo	0.48	0.34	0.33	0.28	0.36
Medio-alto	0.65	0.47	0.38	0.47	0.49
Alto	0.73	0.59	0.61	0.68	0.65
Nacional	0.54	0.42	0.42	0.38	0.44
Grado de concentración (índice de Gini)					
Bajo	0.32	0.28	0.16	0.32	0.24
Medio-bajo	0.09	0.1	0.14	0.11	0.06
Medio-alto	0.03	0.05	0.13	0.06	0.03
Alto	0.09	0.12	0.08	0.08	0.08
Nacional	0.21	0.23	0.24	0.33	0.24

Fuente: elaboración propia con base en los subíndices estimados que se reportan en el Anexo.

cual subraya la importancia de la educación para abatir la brecha digital regional.

Es justamente entre las entidades más rezagadas donde se observa la mayor desigualdad en cuanto al grado de madurez alcanzado por los factores que permiten avanzar en materia de TIC, así lo indica el índice de Gini, sobre todo en el caso de disponibilidad y calidad del capital humano e infraestructura.

En el cuadro 6 se han ordenado de mayor a menor las entidades federativas de acuerdo con el valor que adquiere su IDTIC. Comparando éste con la ordenación alcanzada por sus factores, encontramos que los estados de México, Baja California y Jalisco, aun cuando ocupan el tercero, cuarto y sexto lugar, respectivamente, en el desarrollo de las TIC, presentan rezagos relativos en cuanto a disponibilidad y calidad de capital humano, donde se sitúan en el noveno, décimo y decimose-

Cuadro 6

Continúa

Posición relativa de las entidades mexicanas, de acuerdo con el valor del IDTIC y de los factores que lo forman

Estado/componentes	Factores del índice				Índice global IDTIC
	Primer componente		Segundo componente		
	Calidad y disponibilidad de capital humano	Infraestructura para TIC	Penetración estratégica de las TIC	Utilización de las TIC	
Distrito Federal	1	1	2	1	1
Nuevo León	2	2	4	2	2
México	10	3	1	3	3
Baja California	9	4	3	4	4
Sonora	6	7	6	5	5
Jalisco	12	5	5	7	6
Quintana Roo	15	14	9	6	7
Coahuila de Z.	3	11	8	10	8
Tamaulipas	13	13	7	8	9
Chihuahua	8	10	16	9	10
Aguascalientes	4	9	18	14	11
Colima	5	12	17	11	12
Sinaloa	11	17	10	13	13
Baja California Sur	7	6	30	12	14
Querétaro	14	15	15	17	15
Morelos	17	8	21	15	16
Nacional	19	16	14	16	17
Durango	16	18	20	20	18
Yucatán	23	21	13	19	19

Posición relativa de las entidades mexicanas, de acuerdo con el valor del IDTIC y de los factores que lo forman

Estado/componentes	Factores del índice				Índice global IDTIC
	Primer componente		Segundo componente		
	Calidad y disponibilidad de capital humano	Infraestructura para TIC	Penetración estratégica de las TIC	Utilización de las TIC	
Puebla	24	23	11	21	20
Nayarit	18	20	27	22	21
Campeche	21	26	26	18	22
Veracruz de I. de la Ll.	29	28	12	25	23
Guanajuato	28	19	19	26	24
Tlaxcala	20	29	29	23	25
Tabasco	25	30	23	24	26
Hidalgo	27	25	22	27	27
Zacatecas	22	22	25	29	28
San Luis Potosí	26	27	28	28	29
Michoacán de O.	30	24	24	30	30
Guerrero	32	31	31	32	31
Oaxaca	31	32	33	31	32
Chiapas	33	33	32	33	33

Fuente: elaboración propia con base en el Anexo.

gundo sitio. Por su parte, en los estratos medios se encuentra un conjunto de entidades que no han logrado posicionarse debido a que el grado en que utilizan las TIC y la forma en que penetran éstas es insuficiente: Aguascalientes, Colima, Baja California Sur y Morelos.

Entre los estados que se ubican por abajo del promedio nacional podemos distinguir a los de la región sur, donde se observan limitaciones en casi todos los componentes. También, observamos entidades cercanas a la frontera sur (Campeche, Yucatán, Veracruz de Ignacio de la Llave y Guanajuato) que, no obstante que se encuentran por debajo de la media nacional, alcanzan mejor posicionamiento en las áreas estratégicas. De esta manera, la descomposición

del índice pudiera resultar orientativa para estimular de manera dirigida el desarrollo de las TIC por entidad.

Explorando la relación entre innovación, competitividad y TIC

Las tecnologías de la información y comunicación son un habilitador central que facilita la innovación y, a partir de este vínculo, la competitividad (Ruiz, 2008). Entonces, a manera de validación preliminar, se revisa la consistencia entre el comportamiento de estos fenómenos y el del IDTIC. No se busca demostrar una relación de causalidad, sino acercarnos a una validación de su consistencia, comparándolo con los índices de competitividad del Instituto

Tecnológico y de Estudios Superiores de Monterrey (Campos y Naranjo, 2011) y del Instituto Mexicano para la Competitividad (IMCO, 2013) por una parte, así como con los de innovación estatal de Aregional (2010) y del Foro Consultivo, Científico y Tecnológico (FCCyT, 2012), por otra.⁷

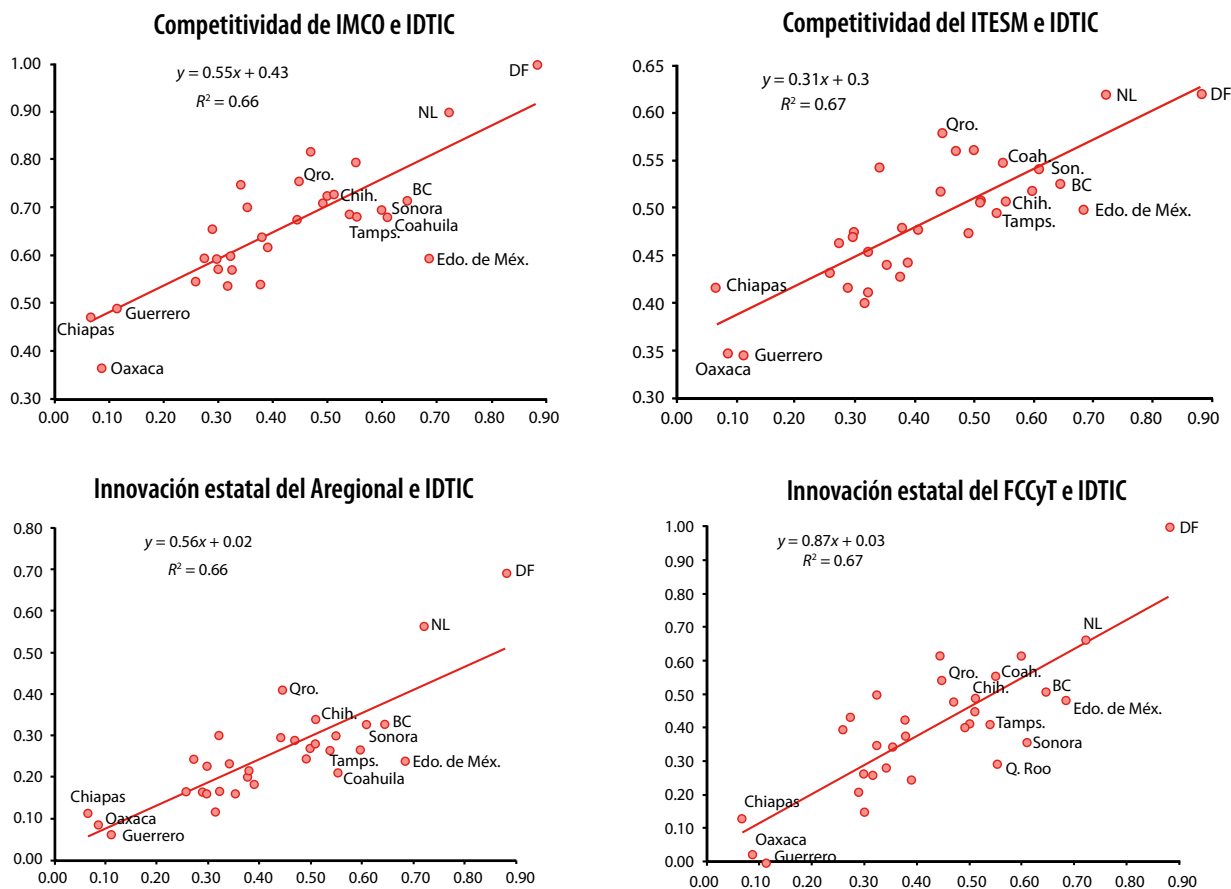
En la gráfica 2 puede observarse una relación positiva con una baja dispersión de entidades entre los pares de índices citados, con un grado de ajuste de casi 70 por ciento. La mayoría de las del norte se encuentran por debajo de la línea de tendencia de ambas relaciones (ver gráfica 1), escapando sólo

⁷ Existen, desde luego, otras opciones para valorar la consistencia del índice, como pudiera ser un análisis de sensibilidad a otras metodologías de cálculo, tal como sugiere y lleva a cabo la ITU (2013: 213).

Nuevo León —la más industrializada y competitiva de la región—, lo cual pudiera evidenciar que la dinámica de innovación en entidades donde predomina la industria maquiladora de ensamble influye en que la incidencia de las TIC sobre los perfiles competitivos sea más débil que en las que han seguido una industrialización de mayor anclaje interno.

Se evidencia la coincidencia entre el rezago innovador o competitivo y el desarrollo de las TIC entre las entidades del sur más pobres. En especial en relación con Oaxaca, Guerrero y Chiapas. Se muestra, además, una reducida dispersión de las entidades a lo largo de la línea de tendencia por lo que, en general, podemos considerar que el comportamiento del IDTIC resulta consistente.

Gráfica 2
Análisis relacional entre los índices de innovación y competitividad y de desarrollo de las TIC*
(sobre el eje Y, competitividad e innovación estatal y en el X, IDTIC)



* Los indicadores han sido normalizados, de manera que no superen el 1.
Fuente: elaboración propia.

Conclusiones

Con el fin de monitorear las metas planteadas en la estrategia digital nacional, México dispone de un índice de digitalización cuyos pilares (aunque pertinentes) no se encuentran *desdoblados* de forma regional (GR, 2013: 43). Es cierto que los pilares de asequibilidad —relativos a los costos y las tarifas para acceder a internet y telefonía— y de capacidad en banda ancha no son factibles de valorar de manera regional, pero sí los de accesibilidad, utilización, capital humano y confiabilidad que contempla.

Al respecto, el IDTIC aquí propuesto pudiera vincularse a ciertos factores de estos pilares o componentes con el fin de que la estrategia digital pueda dar seguimiento a metas estatales. Ésta sería una forma concreta de contribuir a los retos mayúsculos que se ha planteado el país en cuanto al desarrollo de las TIC.

De acuerdo con el índice elaborado, se concluye que 47% de la población de México reside en entidades muy dispares, que se encuentran en niveles bajo o medio-bajo de desarrollo de TIC. Resulta clara, además, la preponderancia tecnológica de la capital nacional y el estado de México, pero no tan lejos de ellas las de las entidades del norte que pueden considerarse con un desarrollo relativamente alto de TIC; la franja norte se encuentra muy interrelacionada con la del sur de EE.UU. y, junto con el occidente y el centro, concentra una proporción importante de la industria del país, lo cual seguramente influye en los requerimientos de tecnologías de información y comunicación. En contraste, la región sur —formada por Chiapas, Oaxaca, Guerrero, Tabasco y Michoacán de Ocampo— presenta el menor nivel de desarrollo de estas tecnologías, lo cual coincide con su bajo ingreso per cápita relativo. De hecho, es entre ellas donde se observa la mayor desigualdad en cuanto al grado de madurez de los factores que permiten avanzar en materia de TIC, sobre todo, en lo que respecta a disponibilidad y calidad de capital humano e infraestructura.

Dentro de los ámbitos en los que se pretende fomentar el desarrollo de las tecnologías de la información no se han establecido medidas para potenciar su impacto sobre la capacidad de innovación y la competitividad regional. Debe identificarse, por ejemplo, la amplitud regional de los nodos tecnológicos de las TIC, integrarlos a la dinámica de los *clusters* sectoriales del país y fomentarlos en aquellas zonas desconectadas de las redes de comunicación e información.

Fuentes

- Aregional institucional (Aregional). *Índice de innovación estatal (IIE)*. Serie Innovación Regional. Año 10, núm. 31. México, DF, 2010.
- Asociación Latinoamericana de Integración (ALADI). *La brecha digital y sus repercusiones*. Secretaría General. Estudio 157. Julio del 2003. Consultado en http://fcbosque.org/index/phocadownloadpap/aladi_brecha_digital-es.pdf, en noviembre del 2014.
- Balboni, Mariana, Cesar Crisancho, Pauline Stockins y Daniel Taccari. *Propuesta de indicadores complementarios para el monitoreo de los ODM: indicadores de acceso a las tecnologías de la información y las comunicaciones*. Serie de Estudios Estadísticos. Núm. 74. Santiago de Chile, CEPAL, 2011.
- Business Software Alliance (BSA). *Global Cloud Computing Scorecard. A Clear Path to Progress*. EE.UU., Galexia Consulting, BSA, 2013.
- Campos Serna, M. y E. Naranjo Priego. *La competitividad de los estados mexicanos*. Monterrey, Nuevo León, Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM) y EGAP, 2011. Consultado en <http://sitios.itesm.mx/webtools/competitividad/competitividad.html>, el 7 de septiembre de 2014.
- Castells, M. *La era de la información. Economía, sociedad y cultura. La sociedad red*. Vol. 1. México, DF, Siglo XXI Editores, 1999.
- Comisión Federal de Telecomunicaciones (COFETEL). *Sistema de Información Estadística de Mercados de Telecomunicaciones (SIEMT)*. México, DF, 2013. Consultado en <http://siemt.cft.gob.mx/SIEM/#!prettyPhoto/15>, el 3 de agosto de 2014.
- De la Fuente, A. *Inversión en TIC y productividad: un breve panorama y una primera aproximación al caso de las regiones españolas*. Barcelona, España, 2008. Consultado en <http://pareto.uab.es/wp/2009/76309.pdf>, el 10 de agosto de 2014.
- Economist Intelligence Unit (EIU). *Digital economy rankings 2010. Beyond e-readiness*. The Economist-IBM.
- European Commission (EC). *Estadísticas sobre la sociedad de la información*. Eurostat, 2014a. Consultado en http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Information_society_statistics/es

- Information society statistics at regional level. Eurostat, 2014b. Consultado en http://epp.eurostat.ec.europa.eu/statistics_explained/index.php/Information_society_statistics_at_regional_level#Rankings_of_selected ICT_indicators
- Foro Consultivo, Científico y Tecnológico (FCCyT). *Ranking nacional de ciencia, tecnología e innovación 2011*. México, DF, Talleres Imagen Maestra, 2012. Consultado en http://www.foroconsultivo.org.mx/libros_editados/ranking_de_cti_2011.pdf, el 9 de marzo de 2013.
- Fuentes Flores, N., M. Ramírez y A. Díaz-Bautista. *Índice de competitividad de las entidades federativas mexicanas 2012*. Tijuana, Universidad Autónoma de Baja California-El Colef, 2013.
- Gobierno de la República (GR). "Ley Federal de Telecomunicaciones y Radiodifusión", en: *Diario Oficial de la Federación*. México, DF, julio del 2014. Consultado en http://www.diputados.gob.mx/LeyesBiblio/pdf/LFTR_140714.pdf, el 27 de noviembre de 2014.
- _____. *Estrategia digital nacional*. México, DF, 2013. Consultado en <http://cdn.mexicodigital.gob.mx/EstrategiaDigital.pdf>, el 3 de agosto de 2013.
- Instituto Mexicano para la Competitividad (IMCO). *Índice de competitividad estatal 2012*. México, DF, IMCO, 2013. Consultado en http://imco.org.mx/indice_de_competitividad_estatal_2012/, el 6 de agosto de 2013.
- International Telecommunications Union (ITU). *Measuring Information Society*. Place des Nations, CH-1211. Geneva, Switzerland, 2013.
- _____. *Impact of broadband on the economy, Telecommunication Sector*. Broadband Series. Geneva, Switzerland, 2012. Consultado en http://www.itu.int/ITU-D/treg/broadband/ITU-BB-Reports_Impact-of-Broadband-on-the-Economy.pdf, el 5 de enero de 2014.
- _____. *World Information Society Report 2007. The Digital Opportunity Index (DOI)*. Place des Nations, CH-1211. Geneva, Switzerland, 2008.
- Instituto Nacional de Estadística, Geografía e Informática (INEGI). *Censo de Población y Vivienda 2010*. Aguascalientes, Ags., México, INEGI, 2011.
- _____. *Estadísticas sobre disponibilidad y uso de tecnologías de información y comunicación en los hogares 2011*. Aguascalientes, Ags., México, INEGI, 2012. Consultado en http://www.inegi.org.mx/prod_serv/contenidos/espanol/bvinegi/productos/encuestas/especiales/endutih/ENDUTIH2011.pdf, en noviembre del 2013.
- Instituto Nacional de la Evaluación de la Educación (INEE). *Base de datos PISA 2009*. México, DF, 2011. Consultado en <http://www.inee.edu.mx/index.php/bases-de-datos/bases-de-datos-pisa/bases-de-datos-pisa-2009>, el 4 de septiembre de 2013.
- Katz, R., P. Koutroumpis y F. Callorda. "The Latin American path towards digitization", en: *Info*. Vol. 15, núm. 3, 2013, pp. 6-24.
- Luna, Dolores E., J. Ramón Gil-García, Luis F. Luna y Rodrigo Sandoval. *Índice de gobierno electrónico estatal: la medición 2010*. Cuadernos de trabajo del CIDE. Núm. 264. México. DF, diciembre del 2011. Consultado en <http://www.cide.edu/publicaciones/status/dts/DTAP%20264.pdf>
- Partnership on Measuring ICT for Development (*Partnership*). *Indicadores clave sobre TIC*. ITU, OECD, UNCTAD, CEPAL, Eurostat. Noviembre del 2010.
- Ruiz Durán, C. "México: geografía económica de la innovación", en: *Comercio Exterior*. Vol. 58, núm. 11. México, DF, noviembre del 2008, pp. 756-768. Consultado en http://revistas.bancomext.gob.mx/rce/magazines/120/1/756_ClementeRuiz.pdf, el 3 de marzo de 2013.
- Ruiz Ochoa, W. "Exploración de largo plazo del rezago regional y de las disparidades interestatales del PIB per cápita en México", en: *Análisis Económico*. Núm. 54, vol. XXIII. México, DF, 2008, pp. 77-102.
- Sandoval Almazán, R., D. Luna, J. Gil-García y L. Luna Reyes. "Índice de gobierno electrónico estatal 2011", en: *Política Digital en Línea*. Enero del 2012. Consultado en <http://www.politicadigital.com.mx/?P=leernoticia&Article=21552>, el 3 de abril de 2013.
- Valderrama Santibáñez, A. L. y O. Neme Castillo. "Efecto de las tecnologías de la información y comunicaciones (TIC) en las exportaciones manufactureras en México", en: *Economía*. Vol. 8, núm. 24. UNAM, 2011, pp. 99-122.
- World Economic Forum (WEF). *The Global Information Technology Report 2013*. Growth and Jobs in a Hyperconnected World. INSEAD-WEF. Insight Report. Switzerland, 2013. Consultado en <http://www.weforum.org/gitr>, el 10 de julio de 2014.
- _____. *The Global Competitiveness Report, 2010-2011*. Geneva, Switzerland, 2011. Consultado en <http://www3.weforum.org/docs/>, el 12 de julio de 2014.

Anexo

Continúa

Descomposición del IDTIC en sus factores, 2010

Estado y componentes	Factores del índice				Índice global
	Calidad y disponibilidad de capital humano	Infraestructura para las TIC	Penetración estratégica de las TIC	Utilización de las TIC	IDTIC
Aguascalientes	0.72	0.50	0.36	0.46	0.51
Baja California	0.65	0.58	0.63	0.71	0.64
Baja California Sur	0.67	0.54	0.20	0.46	0.47

Descomposición del IDTIC en sus factores, 2010

Estado y componentes	Factores del índice				Índice global
	Calidad y disponibilidad de capital humano	Infraestructura para las TIC	Penetración estratégica de las TIC	Utilización de las TIC	IDTIC
Campeche	0.49	0.30	0.24	0.33	0.34
Coahuila de Z.	0.73	0.49	0.51	0.47	0.55
Colima	0.67	0.48	0.37	0.47	0.50
Chiapas	0.00	0.06	0.18	0.03	0.07
Chihuahua	0.66	0.49	0.38	0.51	0.51
Distrito Federal	1.00	0.94	0.72	0.86	0.88
Durango	0.56	0.37	0.31	0.31	0.39
Guanajuato	0.39	0.36	0.34	0.19	0.32
Guerrero	0.11	0.11	0.19	0.04	0.11
Hidalgo	0.42	0.30	0.29	0.18	0.30
Jalisco	0.63	0.55	0.58	0.62	0.60
México	0.65	0.59	0.77	0.72	0.68
Michoacán de O.	0.33	0.32	0.27	0.10	0.26
Morelos	0.56	0.50	0.30	0.40	0.44
Nayarit	0.55	0.35	0.24	0.27	0.35
Nuevo León	0.88	0.63	0.63	0.75	0.72
Oaxaca	0.12	0.06	0.09	0.07	0.09
Puebla	0.43	0.32	0.46	0.29	0.38
Querétaro	0.60	0.44	0.38	0.37	0.45
Quintana Roo	0.59	0.44	0.51	0.66	0.55
San Luis Potosí	0.42	0.29	0.23	0.14	0.27
Sinaloa	0.63	0.41	0.46	0.46	0.49
Sonora	0.67	0.52	0.55	0.69	0.61
Tabasco	0.43	0.26	0.29	0.22	0.30
Tamaulipas	0.61	0.47	0.53	0.55	0.54
Tlaxcala	0.53	0.26	0.22	0.24	0.31
Veracruz de l. de la Ll.	0.35	0.28	0.43	0.22	0.32
Yucatán	0.44	0.34	0.42	0.31	0.38
Zacatecas	0.47	0.33	0.25	0.11	0.29
Nacional	0.54	0.42	0.42	0.39	0.44

Fuente: elaboración propia con base en la metodología indicada en el texto.

Percepción remota

para la estimación de población en áreas geoestadísticas básicas

José Luis Silván Cárdenas, Jorge Alberto Montejano Escamilla
y Mauricio Pablo Cervantes Salas

French Testis Range Ship 'Monge' On January 4Th, 2001, France/Gamma-Rapho/Getty Images



Nota: el estudio fue financiado con recursos del Fondo Sectorial CONACYT-INEGI (proyecto núm. 187593); los autores reconocen la aportación de investigadores y estudiantes que estuvieron involucrados en algunas etapas del proyecto y cuya labor contribuyó a la exitosa culminación del mismo.

Se presentan los resultados de un estudio para evaluar la posibilidad de estimar la población en áreas geoestadísticas básicas (AGEB) de tres delegaciones del Distrito Federal (DF) mediante datos de altimetría LIDAR, imágenes multiespectrales y datos del catastro. En una primera etapa se establecieron relaciones empíricas entre población y vivienda usando los datos censales del 2000 y 2010; después, se emplearon indicadores derivados de catastro y percepción remota para estimar el número de viviendas. Entre los modelos de estimación de población evaluados, el que usa incrementos relativos permitió estimaciones de población con una mediana de error relativo de 5% cuando se emplearon las viviendas del censo. Al estimar el número de viviendas con base en información de percepción remota y de catastro, este error se incrementó con un valor de mediana por debajo de 15% para la mejor composición de modelos. El mayor error de estimación se obtuvo para la delegación Cuauhtémoc, donde se presume que la población se aloja en otros usos de suelo aparte del uso exclusivamente habitacional y donde, además, la distribución espacial del uso de suelo está muy agrupada, hecho que demandó un modelo con más información que no fue posible extraer mediante percepción remota.

Palabras clave: estimación de población, estimación de viviendas, percepción remota, LIDAR.

Recibido: 15 de septiembre de 2014

Aceptado: 15 de diciembre de 2014

Introducción

El volumen total de personas que forman una población es un insumo básico para diversos procesos de análisis previos al diseño de políticas públicas y de planes de ordenamiento territorial. Desafortunadamente, los censos de población y vivienda, que son la fuente más confiable para obtener dicha información, se realizan cada 10 años debido a los altos costos y tiempos requeridos, y aunque son escalonados con los conteos de po-

This article presents the results from a study aimed to estimate population on basic geo-statistical areas (census tracts) along three districts of Mexico City through altimetry LIDAR data, high-spatial resolution multispectral imagery, and cadastral data. In a first stage, empirical relationships between population and housing units were developed based on data from the 2000 and 2010 censuses. Indicators derived from remote sensing and cadastral data were used to estimate housing units. Among the population estimation models tested, the one employing relative population increments yielded estimates with 5% of median relative error when the number of housing units was extracted from census. When estimating housing units with remote sensing, the errors increased but with a median value under 15% for the best composite model. The highest population estimation error was observed for Cuauhtémoc district, where not only population seemed to live in more than one land use type, but also the distribution of land uses was rather clustered, which demanded a model with much more information that could not be extracted from remote sensing.

Key words: population estimation, housing estimation, remote perception, LIDAR.

blación para alcanzar un intervalo de muestreo¹ de cinco años, la frecuencia resulta insuficiente, sobre todo en áreas territoriales con una acelerada dinámica poblacional. En estos casos, una estimación inicial rápida para años no censales supone una ventaja importante en la toma de decisiones. Por lo tanto, en esta investigación se planteó, como hipótesis general, que los senso-

¹ Cabe anotar que los conteos no cuentan todo el universo, como en el caso de los censos.

res montados en plataformas aéreas o satelitales podrían aportar una solución consistente y de bajo costo.

La estimación de población en áreas pequeñas tiene una multiplicidad de metodologías que, por su naturaleza, han sido agrupadas en métodos basados en funciones matemáticas y en variables sintomáticas o indicadores. El primer grupo produce estimaciones de población que suelen ser muy útiles para calcular la población de áreas pequeñas que no cuentan con estadísticas vitales en periodos intercensales (González y Torres, 2012). Es recomendable que este grupo de métodos se aplique en lapsos que no excedan de 15 años y que el control de ajustes haga empleo de proyecciones de población (González y Torres, 2012, p. 108). Los métodos basados en variables sintomáticas, llamados de esa forma porque se sustentan en encontrar información disponible de variables que se encuentran relacionadas de alguna manera al volumen y cambios de una población (González y Torres, 2012, p. 105). Entre las variables que se han empleado destacan los registros de consumidores de servicios básicos (agua y electricidad), las matrículas escolares, los registros electorales y las estadísticas de construcciones habitacionales (González y Torres, 2012, pp. 115-116).

El ejercicio que aquí se presenta es una variación de los métodos de variables sintomáticas, pues busca asociar el número total de habitantes con ciertos aspectos observables de la población mediante el uso de percepción remota. Por su parte, la estimación de población a través de datos de percepción remota y sistemas de información geográfica (SIG) no es una novedad. No obstante, un vistazo a la literatura disponible sobre el tema revela que la gran mayoría de los estudios ha empleado datos provenientes de sensores de resolución baja (alrededor de 1 km por pixel), como DMSP y AVHRR (Yang, Yue y Gao, 2013; Sutton, P., 2000; Sutton, Roberts, Elvidge y Melj, 1997; Dobson, Bright, Coleman, Durfee y Worley, 2000) y, sobre todo, de resolución moderada (entre 10 y 100 metros por pixel), como los sensores de la serie *Landsat* (Azar, Engstrom, Graesser y Comenetz,

2013; Wu y Murray, 2007; Lu, Weng y Li, 2006; Harvey, 2002; Iisaka, 1982). Dentro del primer grupo se encuentran aquellos estudios que se enfocan a la estimación de población a nivel regional y global, mientras que en el segundo están los que hacen estimaciones a nivel de ciudades y regiones. También, es interesante notar que, en una escala regional, el principal indicador de población suele ser el área de los asentamientos, en tanto que en una escala de ciudad o localidad se tienden a emplear los de áreas de superficie impermeable, tipos de coberturas o usos de suelo.

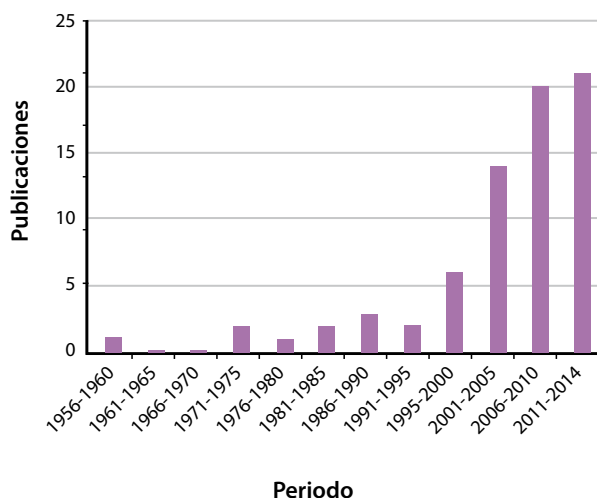
Los estudios de estimación de población para áreas pequeñas (como manzanas o AGEb) incluyen, por lo general, el uso de imágenes de resolución espacial alta (de alrededor de 1 metro por pixel), entre las que sobresale la fotografía aérea, que cuenta con una larga tradición que ha evolucionado de los formatos analógicos a los digitales. En México se tiene como antecedente de su uso para este fin un estudio realizado por El Colegio de México (COLMEX), AC (Ordorica e Ibarra, 2002), en el cual se estimó la población de la Villa Milpa Alta mediante el ajuste del filtro de Kalman a series históricas de población (1970-2000) y área del asentamiento para las distintas fechas de los censos. En ese caso, el área se extrajo delineando sobre fotografías aéreas los límites del asentamiento urbano.

Con el lanzamiento de *IKONOS* en 1999, considerado como el primer satélite civil de resolución espacial alta, se inició una etapa de exploración de metodologías automatizadas para estimar población en áreas pequeñas. De esto da cuenta un incremento exponencial en el número de publicaciones sobre el tema a partir de entonces (ver gráfica 1). Este acelerado aumento también se explica por la amplia disponibilidad de herramientas de análisis espacial en los SIG, los cuales permitieron integrar datos vectoriales con datos en formato *raster*, que es el nativo de las imágenes.

Entre los sensores satelitales multiespectrales de alta resolución espacial que se consideran en esta categoría, se encuentran *IKONOS* (1 m de

Gráfica 1

Número de publicaciones sobre el tema de estimación de población mediante PR y SIG por intervalo de cinco años, salvo el último*



* En total, se consultaron 71 publicaciones en 11 revistas, cuatro libros, cuatro congresos y una tesis.

resolución en la pancromática), *QuickBird* (0.6 m), *WorldView-2* (0.5 m) y *Geoeye-1* (0.4 m). De las 71 publicaciones consultadas, 13 reportaron el uso de imágenes satelitales de alta resolución espacial, donde *IKONOS* y *QuickBird* destacan como las de mayor uso, ya sea solas o en combinación con otro tipo de datos. Este número se duplica si se consideran estudios que emplearon fotografías aéreas, datos de catastro o mapas topográficos detallados con información de edificios.

Si bien la estimación de población en áreas pequeñas se lleva a cabo de forma típica a partir de dos enfoques (mediante el conteo de unidades de viviendas y la determinación de áreas de viviendas), en fechas recientes, en algunos países (como Japón) se han empleado datos de catastro actualizados para estimar población utilizando no sólo el área y número de edificios, sino también el número de niveles (Lwin y Murayama, 2009).

Con el advenimiento del sensor LIDAR se pudieron hacer mediciones en tres dimensiones, tanto del terreno como de los elementos sobre él. No obstante que los escáneres por láser se

crearon en la década de los 90, su aplicación para la estimación de población es muy reciente (la referencia más antigua data del 2010). En este tipo de modelos se suele incorporar el volumen de edificios como variable sintomática, de tal forma que ya no sólo es necesario el conteo o delineado de áreas construidas, sino también la estimación de su altura y/o el número de niveles (Silván-Cárdenas, L. Wang, Rogerson, Feng y Kamphaus, 2010; Qiu, Sridharan y Chun, 2010; Sridharan y Qiu, 2013; Lwin y Murayama, 2011).

Con estos antecedentes, se planteó un proyecto cuyo objetivo fue desarrollar y probar una metodología de estimación de población en las áreas geostatísticas básicas empleando imágenes de alta resolución espacial y datos de altimetría LIDAR para la ciudad de México. El problema se abordó a partir del planteamiento de dos hipótesis interrelacionadas:

- El número de viviendas de un área refleja la población total del área.
- El número de viviendas se puede estimar a partir de datos de percepción remota.

Mientras que la primera asume de manera implícita que la población que no habita en viviendas es una fracción muy pequeña del total poblacional (en todo caso, es proporcional a la población que sí lo hace), la segunda asume que la manifestación física de cada vivienda es cuantificable. De ser el caso, la población podría estimarse a partir del número de viviendas, el cual se calcularía a su vez a partir de los espacios habitacionales detectados con percepción remota.

El resto del documento describe el desarrollo y prueba de modelos empíricos que intentan probar o refutar las dos hipótesis. En primer lugar, se introducen ambos modelos; en segundo, se describe el área de estudio y el procesamiento de los datos de catastro y de percepción remota para el cálculo de variables sintomáticas; después, se presentan y analizan los resultados de la evaluación de los modelos y, por último, se dan las conclusiones y reflexiones derivadas del estudio.

Modelos de estimación de población

Para probar la primera hipótesis, se desarrollaron cinco modelos regresivos de población con número de viviendas como variable independiente (ver tabla 1).

Los modelos $P1$ y $P2$ relacionan de forma directa el número de unidades de vivienda con la población total en una misma fecha. Mientras $P1$ establece una relación lineal entre las variables, el modelo $P2$ establece una relación logarítmica que se deriva de la ley de crecimiento alométrico. En Biología, la ley de crecimiento alométrico establece que el incremento relativo de un organismo es proporcional al incremento relativo de un órgano. En el contexto geográfico, se ha mostrado que la población tiende a obedecer una relación similar con el tamaño del asentamiento (Nordbeck, 1965; Nordbeck, 1971), la cual se puede hacer más específica al considerar el número de viviendas en lugar del tamaño del asentamiento. Si, además, se considera que las variables son continuas, se puede plantear la relación en función de derivadas, lo que resulta en una ecuación diferencial cuya solución expresa una relación lineal entre los logaritmos de las variables. El modelo $P2$ presenta

una ligera variación de la ley de crecimiento alométrico porque se suma 1 a las variables para evitar logaritmos con 0.

Los modelos $P3$, $P4$ y $P5$ representan otras variaciones que relacionan el incremento de la población con el aumento de unidades de viviendas pero, a diferencia de $P2$, éstos mantienen el carácter discreto de los incrementos; por ejemplo, el $P5$ relaciona los crecimientos normalizados por la suma de las poblaciones de las dos fechas, lo que equivale a un incremento relativo a la población promedio del periodo.

Modelos de estimación de viviendas

Para probar la hipótesis dos, se desarrollaron 10 modelos, los cuales resultan de combinar dos tipos de relación (con y sin transformación de variables) con cuatro tipos de variables independientes, las cuales son: 1) número de edificios con uso de suelo exclusivamente habitacional, 2) área habitable² de edificios con uso de suelo habitacional, 3) área habitable de edificios con usos de suelo seleccionados óptimamente y 4) área habitable total sobre

² Área de la planta por el número de niveles.

Tabla 1

Modelos regresivos de estimación de población

ID	Expresión	Descripción
$P1$	$P = aV + b$	Relación lineal entre población total (P) y número de viviendas (V)
$P2$	$P = b(V + 1)^a - 1$	Relación lineal entre logaritmos de $P + 1$ y $V + 1$
$P3$	$P = P_0 + a(V - V_0) + b$	Relación lineal entre los incrementos de población ($P - P_0$) y vivienda ($V - V_0$)
$P4$	$P = b(P_0 + 1) \left(\frac{V + 1}{V_0 + 1} \right)^a - 1$	Relación lineal entre incrementos de logaritmos de $P + 1$ y $V + 1$.
$P5$	$P = \frac{(1 + a + b)V + (1 - a + b)V_0}{(1 - a - b)V + (1 + a - b)V_0} P_0$	Relación lineal entre incrementos normalizados por $P + P_0$ y $V + V_0$

Nota: las literales a y b en las expresiones matemáticas representan parámetros del modelo que deben ser estimados; en contraste, P_0 y V_0 son valores iniciales conocidos de población y vivienda, por ejemplo, del censo previo.

todos los usos de suelo disponibles. Las expresiones explícitas de cada modelo se proporcionan en la tabla 2, junto con una breve descripción.

Los modelos V1 y V3 están inspirados en los tradicionales que relacionan el número de viviendas y el volumen construido, en ese orden. No obstante, en lugar de emplear el volumen, se emplea el área de la planta multiplicada por el número de niveles, cantidad denominada como área habitable. Así, si la altura de entresijos fuera la misma para todos los edificios, el área habitable resultaría ser proporcional al volumen envolvente del edificio.

En nuestro caso, esta altura no es constante, tal como se muestra más adelante. Los modelos V2 y V4 corresponden a las versiones con transformación logarítmica de las variables, como extensión de la ley de crecimiento alométrico.

Los V5 y V6 son, en principio, modelos con múltiples variables que representan el área habitable de edificios con distintos usos de suelo. Con ello, se presupone que la población puede vivir en edificios con uso de suelo reportado en catastro distinto al exclusivamente habitacional. Más aún, en lugar de seleccionar *a priori* los usos de suelo

Tabla 2

Modelos de estimación de viviendas en términos de características de edificios y usos de suelo

D	Expresión	Descripción
V1	$V = aN_H + b$	Relación lineal del total de viviendas (V) con el número de edificios exclusivamente habitacionales (N_H)
V2	$V = b(N_H + 1)^a - 1$	Relación lineal entre $N_H + 1$ y $V + 1$ en escala logarítmica
V3	$V = aS_H + b$	Relación lineal entre el total de viviendas y el área habitable de edificios exclusivamente habitacionales (S_H)
V4	$V = b(S_H + 1)^a - 1$	Relación lineal entre $S_H + 1$ y $V + 1$ en escala logarítmica
V5	$V = \sum_{I=1}^K a_I S_I + b$	Relación lineal con área habitable de edificios para K usos de suelo seleccionados óptimamente*
V6	$V = b \times \prod_{I=1}^K (S_I + 1)^{a_I} - 1$	Relación lineal en la escala logarítmica con área habitable de edificios para K usos de suelo seleccionados óptimamente*
V7	$V = aS + b$	Relación lineal con área habitable total (S)
V8	$V = b(S + 1)^a - 1$	Relación logarítmica con área habitable total
V9	$V = V_0 + a(S - S_0) + b$	Relación lineal entre incrementos $S - S_0$ y $V - V_0$
V10	$V = V_0 + (a\sqrt[3]{(S - S_0)} + b)^3$	Relación lineal entre la raíz cúbica de los incrementos

*Los usos de suelo óptimos fueron seleccionados usando el método de regresión lineal por paso (Efroymson, 1960).

Nota: las literales *a* y *b* en las expresiones matemáticas representan parámetros del modelo que deben ser estimados; en contraste, V_0 y S_0 representan los valores iniciales de vivienda y área habitable.

donde la población podría estar alojada, éstos se determinan a partir de una muestra estadística. Esto se hizo mediante el método de selección de variables conocido como regresión lineal por pasos (Efroymson, 1960). Se trata de un algoritmo iterativo que incorpora y retira variables del modelo en función de su desempeño. Hay dos parámetros claves que controlan la entrada y salida de variables del modelo: 1) la probabilidad de entrada y 2) la probabilidad de salida, las cuales son de significancia estadística de una variable. En el caso del modelo V6, primero se debe expresar el modelo en forma lineal vía logaritmos y después se aplica la regresión lineal por paso a las variables transformadas.

Los modelos V7 a V10 asumen que la información de uso de suelo no es relevante para la estimación de población. Éste puede ser el caso cuando, en el área de estimación, existe una mezcla muy homogénea de usos de suelos, de tal forma que el área habitable de la totalidad de los edificios guarda una proporción directa con el área habitable de edificios donde se aloja la población. Mientras los modelos V7 y V8 son estáticos, como todos los anteriores, los V9 y V10 son dinámicos porque relacionan los incrementos de viviendas con aumentos del área habitable. Por su parte, los V8 y V10 corresponden a las versiones con transformación de variables. En el caso de V10 se emplea la raíz cúbica en lugar de logaritmos; esto fue inspirado por la interpretación del exponente de una relación alométrica que, de acuerdo con Nordbeck (1971), correspondería al cociente de las dimensiones de las variables. Por ejemplo, el exponente b , en una relación alométrica de la forma $A = aP^b$ entre el área A de un asentamiento y la población total P , por lo general se aproxima a $2/3$. Se presume que esto se debe a que mientras el área es una variable de dos dimensiones, la población representa un volumen de tres (Nordbeck, 1971). En otras palabras, la relación que guardaría la raíz cúbica de la población con la raíz cuadrada del área del asentamiento es en esencia lineal. En el caso del modelo V10, tanto el aumento del número de viviendas como el incremento del área habitable son variables volumétricas, por lo que

las raíces cúbicas representan variables unidimensionales que estarían relacionadas de forma lineal.

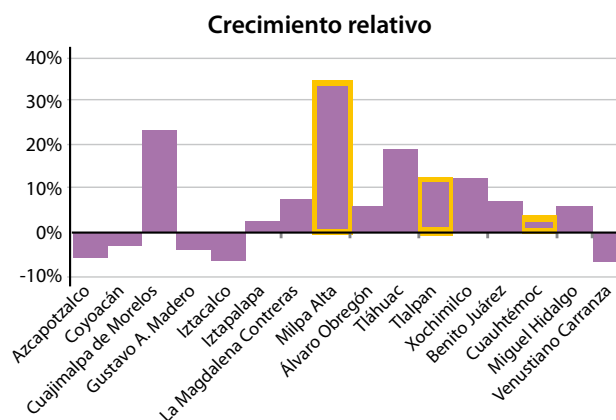
Selección de áreas de estudio y preparación de los datos

Con el fin de probar los modelos en condiciones de densidad y ritmo de crecimiento diversos, se seleccionaron tres delegaciones representativas de un núcleo densamente poblado con bajo crecimiento relativo, un área periurbana con crecimiento moderado y una rural con crecimiento acelerado: Cuauhtémoc, Tlalpan y Milpa Alta, respectivamente. La gráfica 2 muestra el incremento relativo en la población entre el 2000 y 2010 para las 16 delegaciones del DF. Se resaltan en amarillo las que corresponden a las delegaciones seleccionadas para el estudio, aquéllas con decremento poblacional se descartaron debido a que, se presume, la dinámica de la poblacional no estaría ligada a la del crecimiento urbano de esas demarcaciones, sino a otras específicas, como: la especialización del uso del suelo, el encarecimiento del suelo intraurbano y la posterior gentrificación,³ además del crecimiento suburbano, entre otras.

3 Proceso de transformación urbana en el que la población original de un sector o barrio deteriorado es desplazada de manera progresiva por otra de un mayor nivel adquisitivo, a la vez que se renueva.

Gráfica 2

Crecimiento relativo de la población en las delegaciones del DF durante el periodo 2000-2010



Fuente: basado en datos del XII Censo General de Población y Vivienda 2000 y del Censo de Población y Vivienda 2010.

Una vez seleccionadas las áreas de estudio, se obtuvo la geometría de AGEB para cada una de ellas, así como los datos de población y vivienda de los censos del 2000 y 2010 (INEGI, 2010 y 2000). Éstos fueron ligados a la geometría de AGEB en un sistema de información geográfica. Para ello, se tomó como base la geometría del 2010 debido a que es la más completa y precisa, por lo que todos los atributos se anexaron a ella. El ligado de los datos se hizo vía las claves de AGEB, de tal forma que a las que no existían en el 2000 se les asignaron valores de 0 en los atributos de población y vivienda del 2000. Además, se encontraron dos AGEB en el 2000, una en Tlalpan y otra en Milpa Alta, que se subdividieron en el 2010; en este caso, se hizo una asignación del dato del 2000 proporcional al área de las AGEB del 2010.

Cuantificación del área habitable

Además de anexar a la geometría base los atributos para el total de la población y el número de viviendas, se anexaron también otros para la cantidad de edificios y el área habitable por categoría de uso de suelo, los cuales fueron calculados a partir de datos del catastro y de percepción remota.

Los datos del catastro, con actualización al 2000, incluyeron polígonos de área construida con el número de niveles, así como polígonos de predio con la clave de uso de suelo. De la clave de uso de suelo, sólo se utilizó la primera letra que representa el uso de suelo principal, de tal forma que se ignoraron variaciones de un mismo uso o usos mezclados. Fue así como se generó una clasificación de predios en 27 categorías de uso de suelo. El área habitable se calculó primero por polígonos de área construida y luego se agregó por polígonos de predio. En el nivel de predio se combinó con la información de uso de suelo para después ser agregada a nivel de AGEB para cada uso de suelo por separado y para el total. El mismo procedimiento se siguió para el número de edificios.

Ya que no se contó con una versión del catastro actualizada al 2010, se generó una capa de áreas

construidas que se detectaron a partir de datos de altimetría LIDAR e imágenes multiespectrales, todos adquiridos a finales del 2007. Aunque la fecha de adquisición de los datos de PR no coincidió con los datos del censo del 2010, se estimó suficiente debido a que, en promedio, los edificios se construyen en 12 meses y se habitan de dos a cinco años para asegurar rentabilidad inmobiliaria. A partir de los datos LIDAR, se estimó el número de niveles y se probó la viabilidad para actualizar la información de uso de suelo con base en características de los edificios y otras variables derivadas. Los resultados de estas pruebas fueron publicados de forma previa (Silván-Cárdenas, Almazán-González y Couturier, 2014). La extracción del uso de suelo mostró ser muy difícil, incluso con una taxonomía relativamente simple, por lo que en este estudio se consideró sólo la información del uso de suelo disponible en el catastro del 2000. A partir de esta información se generaron los indicadores de área habitable y número de edificios, tal como se hizo con los polígonos del catastro.

Extracción de áreas construidas

Ésta consistió en una serie de pasos que se describen de forma breve en esta sección. En el primero se generó un modelo digital de alturas (MDA) a partir de la nube de puntos LIDAR, siguiendo una metodología previamente publicada (Silván-Cárdenas y Wang, 2014; Silván-Cárdenas, J., 2013). En el segundo se empleó la imagen multiespectral *QuickBird* para identificar áreas de vegetación a partir del índice normalizado de vegetación (NDVI, por sus siglas en inglés). Las áreas de vegetación son excluidas del MDA, con lo que se obtiene un modelo digital de alturas de edificios.

En el tercer paso, con el fin de delimitar las áreas construidas, se estimó la magnitud del gradiente del MDA mediante el filtro de Sobel empleado para la detección de bordes en el procesamiento digital de imágenes. De forma previa se había desarrollado una metodología de extracción de áreas construidas a partir del modelo digital de alturas en formato *raster* (Silván-Cárdenas y Wang, 2011).

En ese método, los píxeles son clasificados en edificado y no edificado, basado sólo en medidas texturales, lo que por lo general produce una máscara de edificios con contornos complejos y ruidosos. En cambio, en esta investigación, primero se aplicó el método de segmentación por cuencas a la magnitud del gradiente del MDA. Luego, se simplificaron los segmentos a partir de una serie de criterios empíricos antes de ser convertidos a polígonos. Por ejemplo, los segmentos con altura menor a 2 metros se eliminan, ya que éstos, por norma, no podrían constituir área habitable.⁴ Asimismo, los segmentos aislados de áreas menores a 16 metros cuadrados se eliminan por un razonamiento similar.⁵ Por otro lado, si estos segmentos pequeños son interiores a otros más grandes, no se eliminan sino que se unen

con éstos, ya que los primeros suelen representar estructuras en azoteas, como: tinacos, tanques estacionarios, pequeñas bodegas, etcétera.

Por último, la segmentación resultante se convirtió a formato vectorial para ser incorporada en el SIG (ver imagen 1). A los polígonos así generados se le anexaron los valores de área y de altura media, así como el número de niveles que se estimó usando el modelo que se describe a continuación.

Estimación del número de niveles

A partir de alturas estimadas por LIDAR para las áreas construidas según el catastro, se construyó un modelo que permitió estimar el número de niveles de edificios. Para ello, se analizó el diagrama de dispersión entre el número de niveles del catastro y las alturas estimadas. Dada la discrepancia temporal entre ambos parámetros, uno debería esperar cierta dispersión, pero también una tendencia predominante debido a que, de manera presumible, la gran mayoría de áreas construidas no sufrieron

- 4 El art. 99 del *Reglamento de Construcciones para el Distrito Federal (RCDF)* vigente señala que la altura mínima para cualquier circulación en cualquier edificio no deberá de ser menor a 2.10 m de altura.
- 5 A pesar de que el art. 5.º del RCDF define una vivienda como un espacio habitable que contiene como mínimo 24 m², al final del mismo se establece que: "Se considerará vivienda mínima la que tenga, cuando menos, una pieza habitable y servicios completos de cocina y baño...". El art. 9.º transitorio considera que una pieza habitable tiene como mínimo 7 m²; una cocina, 3 m²; y baño sin límites preestablecidos. Así, una vivienda mínima podría tener cuando menos 12 m²; sin embargo, es poco probable que tenga estas dimensiones dadas las necesidades adicionales espaciales, como patio de servicio/iluminación.

Imagen 1

Ejemplo de extracción de edificios a partir de la segmentación del modelo digital de alturas



Nota: a la izquierda se muestra la segmentación y a la derecha, los polígonos generados sobrepuestos en el modelo digital de alturas.

cambios en el lapso de ocho años; por lo tanto, el modelo que describe la relación entre las variables debe ajustarse precisamente a los puntos alojados alrededor de dicha tendencia (ver gráfica 3).

Para esta investigación, se ajustaron de forma visual tres segmentos lineales en tres distintos rangos de alturas (línea azul en la gráfica 3); el conjunto de los segmentos describe, precisamente, el modelo de estimación, el cual se puede expresar mediante la siguiente ecuación:⁶

$$n = \begin{cases} \left\lceil \frac{2h}{5} \right\rceil, & \text{si } h < 15 \\ \left\lceil \frac{h+3}{3} \right\rceil, & \text{si } 15 \leq h < 45 \\ \left\lceil \frac{h+19}{4} \right\rceil, & \text{si } h \geq 45 \end{cases}$$

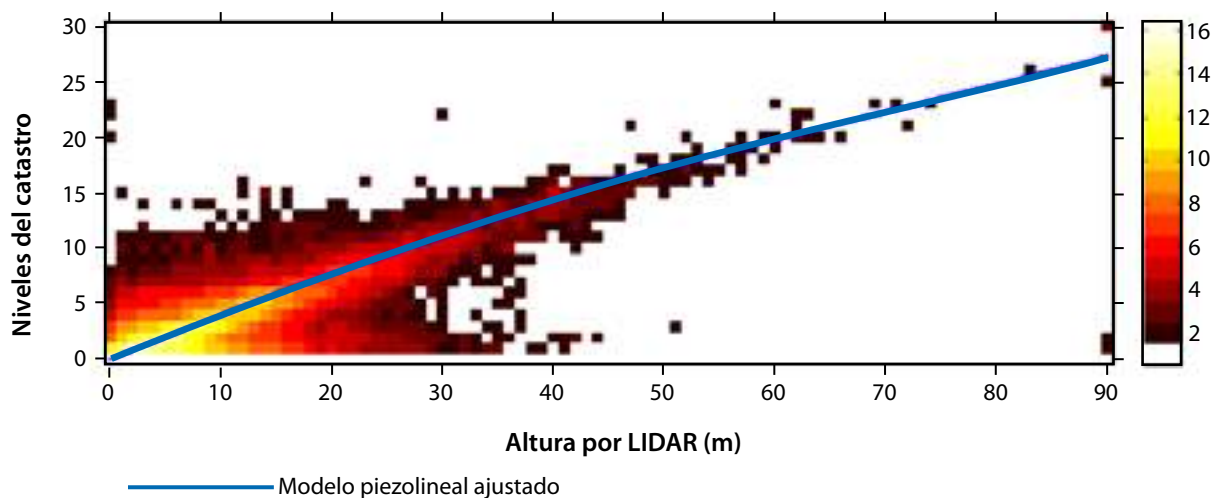
donde n y h denotan el número de niveles y la altura, en ese orden, y los corchetes significan redondeo al entero más cercano.

⁶ El modelo se refinó a tres segmentos a partir de un modelo propuesto de forma previa (Silván-Cárdenas, Almazán-González y Couturier, 2014).

Es interesante notar que los coeficientes del modelo indican que la altura de entepiso para edificios de cinco pisos o menos es de 2.5 metros, mientras que para edificios de entre seis y 15 pisos es de 3 metros y para más de 15 pisos, de 4 metros. Además, las constantes que se suman representan los niveles que no contribuyen a la altura del edificio, esto es, los de sótanos. Así, los edificios de entre seis y 15 pisos tienen al menos un piso bajo el nivel del terreno, mientras que los de más de 15 cuentan con alrededor de cinco pisos bajo el nivel del terreno. Estos valores no reflejan necesariamente los rangos de valores de altura libre y niveles de sótanos establecidos en el *Reglamento de Construcciones para el Distrito Federal* (Arnal, 2005),⁷ ya que resultan de una observación estadística sobre una gran cantidad de edificios de diferente tipo y que fueron construidos con un amplio rango de valores posibles.

⁷ Los requerimientos mínimos de altura se encuentran especificados en el capítulo 2 (*Habitabilidad, accesibilidad y funcionamiento*) de las *Normas Técnicas Complementarias para el Proyecto Arquitectónico* del RCDF (Arnal, 2005), destacándose que el rango mínimo para todos los tipos de edificaciones oscila entre 2.10 m y 3 m libres, sin contar los casos donde el director responsable de obra (DRO) determina la altura mínima. La altura máxima, exceptuando los casos donde interviene el DRO, es de 3.60 m para efecto de la contabilización del número de niveles. En caso de exceder esta altura, se tomará como equivalente a dos niveles construidos para efecto del conteo de elevadores y clasificación de usos y destinos.

Gráfica 3
Diagrama de dispersión entre altura y número de niveles de pisos de edificios de las tres delegaciones



Medidas de desempeño de los modelos

Tanto para los modelos de estimación de población como para los de viviendas se definieron métricas para cuantificar la bondad de ajuste y la capacidad de generalización. Éstas incluyen:

1. El coeficiente de determinación (R^2), el cual mide el porcentaje de varianza de la variable dependiente que está determinada por la(s) variable(s) independiente(s).
2. La mediana del error absoluto (MAE), que proporciona una medida de tendencia central del error absoluto, al tiempo que es robusto ante datos anómalos.
3. La mediana del error absoluto relativo (MARE), la cual proporciona una medida robusta de la distribución del error en términos proporcionales a los valores de la variable estimada.

Para cada delegación, se escogieron de manera aleatoria a lo sumo 50 AGEBS, pero no más de dos terceras partes del total, las cuales fueron empleadas como muestra de calibración, mientras que el resto se emplearon como muestra de validación. El estadístico R^2 se calculó sobre la muestra de calibración, mientras que las métricas MAE y MARE se calcularon sobre la muestra de validación. Para el cálculo de MARE, se excluyeron las divisiones entre 0 para evitar indeterminación.

Resultados

Las tablas de atributos con claves de AGEBS se importaron en Excel y luego se calcularon los incrementos absolutos y relativos, y demás transformaciones, de las variables requeridas para *linealizar* los modelos (ver tablas 1 y 2). La estimación de parámetros de todos los modelos, con excepción de V_5 y V_6 , se hizo a partir de la regresión simple en Excel 2010 (Microsoft Corp.). Para el caso de los modelos V_5 y V_6 , se aplicó la implementación de la regresión lineal por paso (*stepwisefit*) disponible en el *ToolBox Statistics* de MATLAB (*The Mathworks Inc.*). En estos casos se emplearon los valores de

probabilidad de entrada y salida de variables de 0.01 y de 0.02, respectivamente. Para el cálculo final, las estimaciones de los modelos fueron redondeadas al entero más cercano y los valores negativos reasignados con 0.

Estimación de población con viviendas del censo

Los valores de las medidas de desempeño de cada modelo se presentan en las tablas 3, 4 y 5 para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta, respectivamente. También, muestran los coeficientes (A y B) de cada modelo *linealizado*.⁸ En términos del ajuste, se observó que los modelos P_1 y P_2 ajustaron bastante bien los datos de las tres delegaciones ($R^2 > 0.91$), el modelo P_3 obtuvo el peor ajuste en los tres casos ($R^2 < 0.85$), mientras que los modelos P_4 y P_5 sólo ajustaron de manera aceptable los datos de Tlalpan y Milpa Alta. Estas observaciones sugieren que el enfoque incremental implícito en estos modelos resulta menos apropiado para la Cuauhtémoc. En otras palabras, la dinámica demográfica en esta delegación parece no estar en sincronía con la dinámica de la vivienda. Para entender esto, basta recordar que la Cuauhtémoc perdió población residente a partir del sismo de 1985 (de 1980 al 2010 perdió 34.74%) y sólo del 2000 al 2010 ha comenzado a ganarla, pero en una proporción muy baja. De forma adicional, hubo un crecimiento de establecimientos comerciales, bodegas y oficinas (sobreespecialización del suelo). Este cambio de giro se ha dado muchas veces fuera del marco legal, propiciando que no esté consignado en los registros del catastro, es decir, muchos de los usos del suelo que dicen habitacional no corresponden a vivienda, sino a oficinas, bodegas o comercios. Del mismo modo, ha sucedido en la colonia Condesa (del 2000 al 2010 perdió 10.35% de su población); para el 2010, 15.74% de la vivienda estaba deshabitada o era usada para un giro diferente al habitacional (Salinas Arreortua, 2013).

⁸ A y B denotan los coeficientes de un modelo lineal genérico de la forma $y = Ax + B$. Éstos corresponden con los parámetros a y b de los modelos a través de transformaciones no lineales en los casos necesarios. Así, por ejemplo, mientras que para el modelo $V_1 A = a$ y $B = b$, para $V_2, A = a$ y $B = \log b$, etcétera.

Tabla 3

Resultados de las medidas de desempeño de los modelos de estimación de población para la delegación Cuauhtémoc

Modelo	A	B	R ²	MAE	MARE
P1	2.6832	-140.73	0.9191	263	10%
P2	1.0966	0.09222	0.9673	301	11%
P3	1.4318	-456.47	0.5048	196	7%
P4	0.9366	-0.1289	0.7895	105	4%
P5	0.92	-0.2817	0.8207	89	4%

Nota: los datos en negritas representan los mejores valores de desempeño, muestra de calibración con 50 AGEb, muestra de validación con 103 AGEb.

Tabla 4

Resultados de las medidas de desempeño de los modelos de estimación de población para Tlalpan

Modelo	A	B	R ²	MAE	MARE
P1	3.6574	-264.84	0.9597	214	9%
P2	1.0799	0.28329	0.9779	206	9%
P3	3.1437	-492.21	0.7748	188	7%
P4	1.259	-0.1172	0.9799	131	6%
P5	1.0912	-0.2302	0.942	124	5%

Nota: los datos en negritas representan los mejores valores de desempeño, muestra de calibración con 50 AGEb, muestra de validación con 156 AGEb.

Tabla 5

Resultados de las medidas de desempeño de los modelos de estimación de población para Milpa Alta

Modelo	A	B	R ²	MAE	MARE
P1	3.6319	24.6158	0.9912	54	4%
P2	0.993	0.5826	0.9968	58	4%
P3	2.4333	-159.79	0.8519	134	11%
P4	1.1353	-0.1101	0.9572	73.5	9%
P5	1.1222	-0.2437	0.9572	58	5%

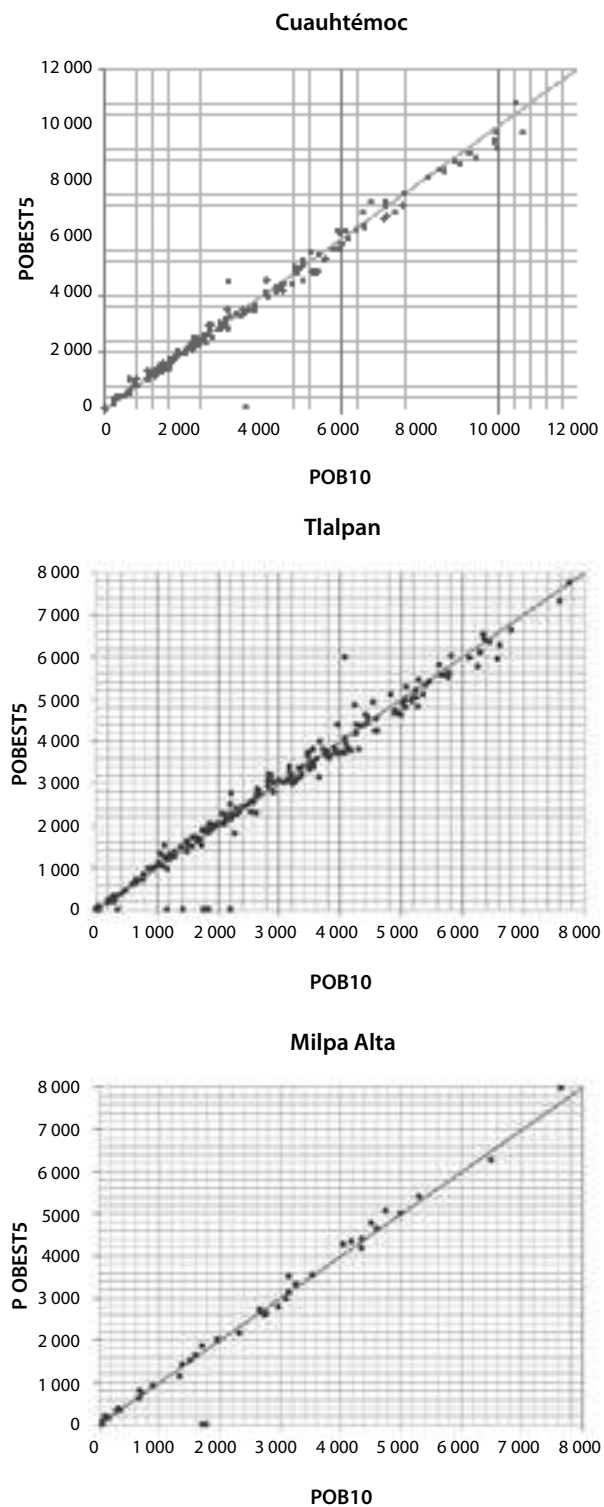
Nota: los datos en negritas representan los mejores valores de desempeño, muestra de calibración con 21 AGEb, muestra de validación con 10 AGEb.

No obstante, en términos de la capacidad de generalización, medida a través de las métricas del error de estimación MAE y MARE, los resultados fueron distintos. Los modelos P4 y P5 presentaron valores de errores relativamente más bajos para las

tres delegaciones, mientras que P1 y P2 lo hicieron sólo para Milpa Alta. Esta aparente contradicción en realidad indica que los modelos que ajustaron mejor la muestra de calibración estimaron peor la muestra de validación y viceversa.

Gráfica 4

Diagramas de dispersión entre población observada y población estimada con el modelo P5 para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta



Nota: la línea a 45° indica estimación perfecta.

Para propósitos de estimación, se prefiere el modelo P5 porque fue el que mejor estimó la población y el que resultó ser prácticamente insensible al cambio de densidad de la población, manteniendo un error de 5% en términos relativos y de entre 58 y 124 personas por AGEb en términos absolutos. La gráfica 4 muestra los diagramas de dispersión entre los valores observados y los estimados mediante este modelo. La gran mayoría de puntos caen alrededor de la línea de estimación perfecta, con lo cual se puede asumir que se valida la hipótesis 1. No obstante, hay casos anómalos que no cumplen con ella; por ejemplo, en el caso de la delegación Cuauhtémoc, el punto sobre la línea horizontal (vea el panel superior izquierdo de la gráfica 4) y corresponde a una AGEb en Tlatelolco que presentó una reducción en el número de viviendas (de 1 210 en el 2000 a sólo 70 viviendas en el 2010) y, en contraste, su población creció de forma ligera de 3 281 a 3 596 personas en el mismo periodo. Evidentemente, se trata de un caso anómalo que deberá analizarse con mayor cuidado. De manera adicional, hubo casos que no fueron representados de forma apropiada por el modelo, los cuales corresponden a las AGEb de reciente creación que se encontraron en Tlalpan y Milpa Alta, donde no fue posible cuantificar un incremento relativo, por lo que el modelo arrojó valores nulos de población.

Estimación de viviendas mediante datos de edificios

La prueba de los modelos V1 a V6 se hizo con datos del catastro del 2000, mientras que la prueba de los modelos V7 a V10 empleó los datos de PR del 2007. Esto fue necesario debido a que el primer grupo de modelos requiere información de uso de suelo confiable con la cual sólo se contó para el 2000. Además, los modelos V8 y V10, que son incrementales, requirieron tanto los datos del catastro como los derivados de PR. Los resultados obtenidos para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta se muestran en las tablas 6, 7 y 8, respectivamente.

Tabla 6

Resultados del ajuste de los modelos de estimación de vivienda a los datos de la delegación Cuauhtémoc

Modelo	A	B	R2	MAE	MARE
V1	0.93	632.87	0.27	357	41%
V2	0.41	1.96	0.50	308	41%
V3	0.01	498.98	0.34	334	37%
V4	0.57	0.21	0.46	259	35%
V5	0.86 (B)	339.38	0.69	245	29%
	0.01 (H)				
	0.34 (U)				
V6	0.08 (H)	2.0508	0.54	275	35%
	0.15 (I)				
V7	0.003	511.08	0.52	419	30%
V8	0.51	0.38	0.52	416	34%
V9	0.001	433.89	0.12	254	20%
V10	0.02	6.84	0.22	170	18%

Nota: los usos de suelo óptimos para el modelo 5 fueron: espacios abiertos (B), habitacional (H) y espacios públicos (U), mientras que para el modelo 6 fueron: habitacional (H) e industrial (I); los datos en negritas representan los mejores valores de desempeño; muestra de calibración con 50 AGEb, muestra de validación con 103 AGEb.

Para los modelos que incorporan la información de uso de suelo, se observó una clara distinción de la delegación Cuauhtémoc de las otras dos; en el caso de ésta, el modelo V₅ fue el que presentó mejores cifras de ajuste y error. Este modelo incorpora una selección óptima de usos de suelo, los cuales resultaron ser habitacional, espacios abiertos y públicos. Es interesante notar que todos los usos seleccionados tienen una contribución positiva en la estimación de las viviendas, lo cual sugiere que el uso de suelo podría estar desactualizado en el sentido de que ciertos espacios abiertos son, en realidad, de uso habitacional o, incluso, que son irregulares. Esto es menos probable en el caso de los espacios públicos y, por el valor relativamente pequeño del coeficiente, es posible que se hayan censado algunas viviendas en espacios públicos.

Por otro lado, en el caso de las delegaciones Tlalpan y Milpa Alta, el uso de suelo óptimo para el modelo V₅ fue sólo habitacional, lo que indica que la relación entre las viviendas y los edificios es mucho más simple que en la Cuauhtémoc. Más aún, los modelos sobresalientes para estas delegaciones fueron los V1 y V2, los cuales relacionan el número de edificios con el de viviendas. Esto sugiere que hay poca variabilidad en el tamaño de los edificios de vivienda en estas delegaciones, lo cual podría ser porque predominan las casas unifamiliares sobre los edificios multifamiliares.

En el caso de los modelos que no requieren información de uso de suelo, se ajustó mejor el modelo V8 seguido del V7, sobre todo en Milpa Alta. El primero se ajustó, incluso, mejor que los

Tabla 7

Resultados del ajuste de los modelos de estimación de viviendas a los datos de la delegación Tlalpan

Modelo	A	B	R2	MAE	MARE
V1	0.61	296.21	0.61	206	30%
V2	0.67	0.94	0.52	164	34%
V3	0.00	399.36	0.32	256	32%
V4	0.45	0.65	0.51	228	35%
V5	0.00 (H)	535.84	0.08	249	32%
V6	0.14 (H)	2.13	0.21	232	33%
V7	0.001	285.71	0.45	258	29%
V8	0.96	-2.51	0.63	268	36%
V9	0.0005	82.58	0.25	85	10%
V10	0.06	1.57	0.24	76	10%

Nota: para los modelos 5 y 6, el uso óptimo fue habitacional (H); los valores en negritas indican el mejor desempeño; muestra de calibración con 50 AGEb, muestra de validación con 156 AGEb.

que incorporaron el uso de suelo, salvo para la delegación Cuauhtémoc, donde el ajuste del modelo V8 ($R^2 = 0.52$) fue menor que el del V5 ($R^2 = 0.69$). Lo anterior sugiere que la información de uso de suelo mejora el ajuste sólo si la demarcación presenta una fuerte sectorización en el uso de suelo, como es el caso en la Cuauhtémoc, donde se encuentran corredores comerciales, colonias habitacionales, etcétera.

En contraste al nivel de ajuste, la capacidad de generalización de los modelos ocurre a la inversa,

es decir, los modelos V9 y V10 presentaron errores más bajos que los V7 y V8. En términos del error relativo, destaca el desempeño consistentemente superior del V10 con las tres delegaciones ($MARE \leq 18\%$). Esta aparente contradicción indica en realidad que los modelos V7 y V8 tienden a sobreajustar los datos, lo cual no ocurre con los V9 y V10.

En términos de la capacidad de generalización, el desempeño del modelo V10 resultó superior incluso que el mejor modelo con información de uso

Tabla 8

Resultados del ajuste de modelos de estimación de viviendas a los datos de la delegación Milpa Alta

Modelo	A	B	R2	MAE	MARE
V1	0.56	231.76	0.61	133	16%
V2	0.67	0.89	0.72	127	17%
V3	0.01	310.55	0.44	240	25%
V4	0.61	-0.06	0.75	183	25%
V5	0.01 (H)	310.55	0.44	240	25%
V6	0.61 (H)	-0.06	0.76	183	25%
V7	0.01	-3.09	0.93	48	26%
V8	0.83	-1.33	0.96	42	17%
V9	0.00	51.91	0.71	86	17%
V10	0.12	1.12	0.75	95	15%

Nota: para los modelos 5 y 6, el uso óptimo fue habitacional (H); los valores en negritas indican el mejor desempeño; muestra de calibración con 21 AGEb, muestra de validación con 10 AGEb.

de suelo. La gráfica 5 presenta los diagramas de dispersión entre las estimaciones del modelo V10 y los valores de vivienda del censo. Las AGEB que presentaron desviaciones significativas respecto a la línea de estimación perfecta fueron sobre todo aquéllas donde se observó un crecimiento en sentido contrario al del área habitable. De forma más específica, los errores de estimación estuvieron correlacionados de manera negativa al incremento de número de viviendas, con valores para el coeficiente de correlación de Pearson de -0.6, -0.8 y -0.6 para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta, respectivamente. Así, a mayor decremento en el número de viviendas hay, por lo general, un mayor error de estimación.

Estimación de población con viviendas estimadas

De los resultados reportados con anterioridad es claro que la composición P5-V10 define la mejor opción para la estimación de la población. No obstante, en esta investigación se probó el desempeño de las combinaciones del modelo P5 con V7, V8, V9 y V10; estos cuatro modelos se eligieron porque: 1) son los que incorporaron los datos de percepción remota y 2) son los que obtuvieron el mejor desempeño en términos del error absoluto. Los errores de estimación de estas pruebas se presentan en la tabla 9.

Tabla 9

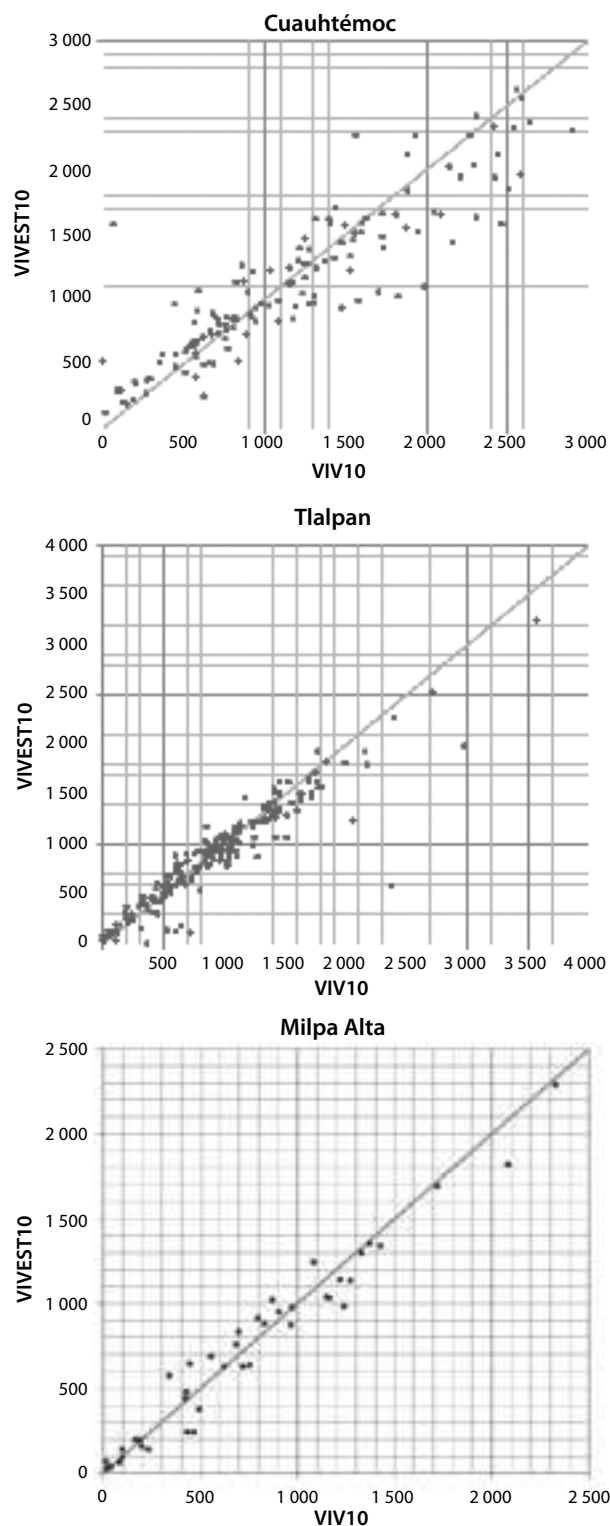
Desempeño de la estimación de población del 2010 con el modelo P5, pero estimando viviendas con los modelos V7 a V10 y con datos de PR

Delegación	Viviendas	MAE	MARE
Cuauhtémoc	V7	818	30%
	V8	915	34%
	V9	524	20%
	V10	317	15%
Tlalpan	V7	936	35%
	V8	983	39%
	V9	317	12%
	V10	288	12%
Milpa Alta	V7	291	15%
	V8	293	16%
	V9	287	15%
	V10	309	14%

Nota: los errores se calcularon sobre todas las AGEB, sin importar si fueron usadas en la calibración de los modelos.

Gráfica 5

Diagramas de dispersión entre viviendas observadas y estimadas mediante el modelo V10 para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta



Nota: la línea a 45° indica estimación perfecta.

Como era de esperarse, los errores de estimación de viviendas se transfirieron a la estimación de población, incrementando con respecto a los observados cuando se emplearon los valores censales del número de viviendas. El peor desempeño se observó para modelos no incrementales (V7 y V8), que alcanzaron valores relativos de hasta 39%, mientras que el mejor se observó para los incrementales (V9 y V10), los cuales presentaron errores por debajo de 18% en términos relativos y de entre 288 y 317 personas en absolutos. Consistente con el desempeño observado para los modelos de estimación de vivienda, el mejor para la estimación de la población fue la composición P5-V10, exhibiendo errores relativos de 15% para la delegación Cuauhtémoc, 14% para Milpa Alta y 12% para Tlalpan.

La distribución espacial del error relativo de las estimaciones hechas mediante la composición P5-V10 para las delegaciones Cuauhtémoc, Tlalpan y Milpa Alta se muestran en las imágenes 2, 3 y 4, respectivamente; en éstas se puede observar que diferentes tipos de errores se concentran en diferentes áreas. Al compararlos con las de uso de suelo, se observó cierta concordancia con la distribución del uso de suelo habitacional. En específico, se observó que las AGEB con predominancia de uso de suelo habitacional presentaron una subestimación entre 10 y 50%, mientras que aquéllas con predominancia de uso de suelo no habitacional se sobrestimó la población en diferentes proporciones. En general, los errores se deben a las mismas causas que en el caso de estimación de vivienda.

Imagen 2

Error relativo de la estimación de población en AGEB de la delegación Cuauhtémoc

Las AGEB Cuauhtémoc

Error relativo

- Subestimado más de 100%
- Subestimado entre 50 y 100%
- Subestimado entre 10 y 50%
- Preciso en +/- 10%
- Sobrestimado entre 10 y 50%
- Sobrestimado entre 50 y 100%
- Sobrestimado más de 100%

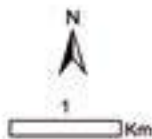


Imagen 3

Error relativo de la estimación de población en AGEB de la delegación Tlalpan

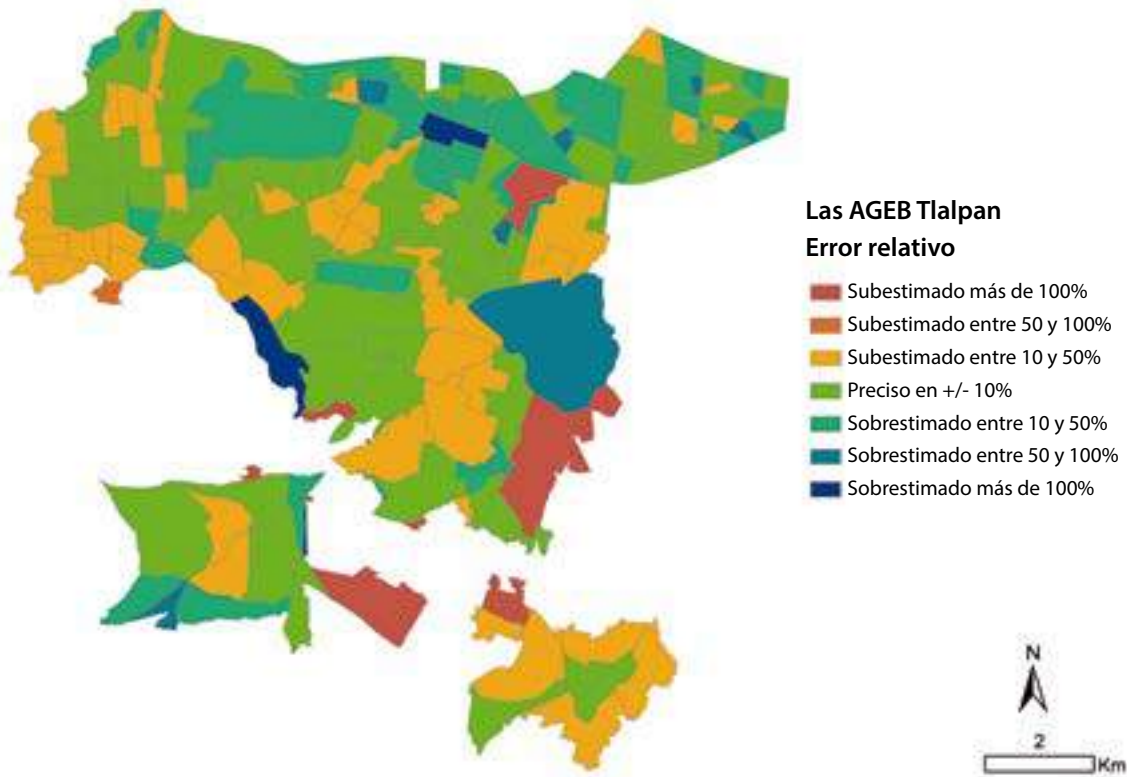
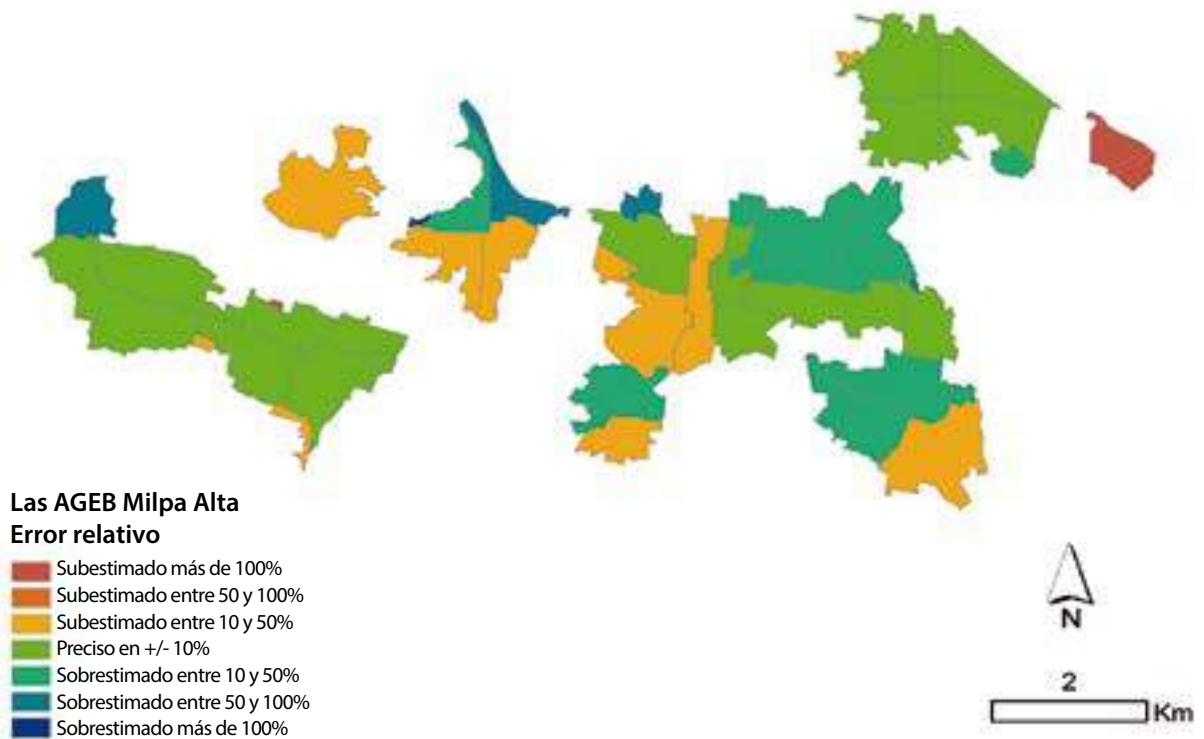


Imagen 4

Error relativo de la estimación de población en AGEB de la delegación Milpa Alta



Discusión

El presupuesto ejercido por el INEGI para el Censo de Población y Vivienda 2010 ascendió a 4 437.96 millones de pesos (mdp); si este costo se distribuye entre el número total de variables censadas (75), el costo aproximado de conocer sólo totales de población es de unos 60 mdp; por su parte, la adquisición de LIDAR involucra dos costos, uno fijo por traslado del sensor (por 21 mil dólares) más otro por cobertura (150 dólares/km², en promedio), que es más de 10 veces el costo de las imágenes multispectrales (13 dólares/km²). Aunque el costo de traslado del sensor suele amortiguarse si se sobrepasa una cobertura mínima, es claro que un levantamiento de todo el territorio nacional resulta impráctica e innecesaria, sobre todo porque 72% de la población se encuentra en áreas urbanas que representan sólo 0.6% de éste (Grajales, 2011). Así, considerando una cobertura conservadora de 30 mil km² a nivel nacional (1.5% del territorio), el costo estimado del levantamiento LIDAR con las imágenes rondaría los 70 mdp lo que, evidentemente, es inviable, pues a esto habría que aumentar los costos asociados al procesamiento de los datos. En todo caso, el costo-beneficio de las adquisiciones de este tipo de datos debe verse en términos de la multiplicidad de usos que se les puede dar en ámbitos como el urbano, agrícola, forestal y ambiental. Más aún, es de esperarse que, a medida que la tecnología se popularice, los costos de adquisición tiendan a bajar. En tanto eso ocurre, resulta útil evaluar los niveles de confiabilidad que se pueden alcanzar con estas metodologías.

Queda claro que para estimar la población ha de estimarse de forma adecuada la vivienda para reducir el margen de error en la estimación. Por su parte, las pruebas hechas mostraron que la estimación de viviendas a partir de datos de edificios resultó ser muy compleja. En este caso, se desarrollaron hasta 10 modelos, de los cuales resultaron mejores aquéllos que incorporaron un enfoque incremental, es decir, que requirieron información de edificios y viviendas tanto para una fecha previa como para la de la estimación.

En estos casos, el nivel de error más bajo fue de 10% en términos relativos. No obstante, se debe destacar que las características de los datos alimentados en estos casos no fue similar para ambas fechas, pues para la primera la información de edificios se extrajo del catastro, mientras que para la segunda se sacó de datos de percepción remota. Presumiblemente, esto debió ser un factor para que el nivel de error de las estimaciones no resultara aún más bajo.

Otro factor fue, sin duda, la incapacidad de incorporar la información de uso de suelo en el caso incremental. Esto se infiere porque los modelos no incrementales que incorporaron la información de uso de suelo disponible en el catastro mostraron que ésta es de gran importancia para estimar las viviendas en áreas con uso de suelo fuertemente sectorizado, como en el caso de la delegación Cuauhtémoc. Previamente, se habían realizado pruebas para detectar edificios de viviendas a partir de sus características físicas, como tamaño y distribución (Silván-Cárdenas, Almazán-González y Couturier, 2014). Si bien los resultados fueron poco favorables, debido a que muchas edificaciones presentan uso de suelo mezclado, se intuye que aun con una detección perfecta de edificios de vivienda se presenta la dificultad de que la población se aloje en otros usos de suelo distinto al habitacional, lo cual hace que la relación área habitable-viviendas sea más errática. Por esta razón, en esta investigación se desarrollaron modelos que no consideran la información de uso de suelo, los cuales probaron ser útiles en ciertos casos, en específico, en áreas donde la mezcla de uso de suelo es más homogénea y donde el crecimiento en altura no es tan significativo, como en Milpa Alta. Para este caso, resultó suficiente emplear como indicador de viviendas el número de edificios con uso habitacional reportado en catastro (modelos V1 y V2) o el área habitable de todos los edificios, sin importar su uso de suelo (modelos V7 y V8).

Dos fenómenos que no fueron considerados al momento del desarrollo de los modelos y que, de manera eventual, podrían explicar parte del

margen de error son el hacinamiento y la desocupación de la vivienda. El primero lo definió el INEGI para el Censo de Población y Vivienda 2010 como el "Porcentaje de viviendas con más de 2.5 ocupantes por dormitorio..." (INEGI, 2010). Para la Sociedad Hipotecaria Federal (SHF), este concepto implica la existencia de dos o más hogares por vivienda (SHF, 2012, p. 1). El fenómeno del hacinamiento, como indicador de marginalidad, supone el comprometer la privacidad de los habitantes de una vivienda, propiciando espacios no adecuados para el desarrollo de las actividades esenciales para las personas. Según el Consejo Nacional de Población (CONAPO), "Ninguna fuente de información (incluido el censo) proporciona datos sobre el tamaño del espacio físico y las dimensiones de los dormitorios y de la vivienda, para relacionarlas con el número de personas que los ocupan..." (CONAPO, 2011, p. 21), hecho que, por sí mismo, constituye un obstáculo para el correcto desarrollo de los modelos aquí planteados, en especial los basados en la relación entre el área habitable y el número de viviendas. Si la población está en función del número de viviendas, y las viviendas están en función del área habitable, deberíamos poder conocer el porcentaje del área habitable que se encuentra en condiciones de hacinamiento para diferenciarla de las demás, de tal modo que el modelo fuera diferenciado para zonas con mayor probabilidad de hacinamiento y para otras con un comportamiento constante en este indicador.

Tan sólo en el DF, hacia el 2011, se estimaba que alrededor de 35% del total de viviendas tenía algún grado de hacinamiento, provocado directamente por el fenómeno poblacional del desdoblamiento, que "...lleva a que donde antes vivía una pareja con uno o dos hijos, ahora también [lo hacen] nueras, yernos, nietos o abuelos..." (Llanos, 2012). A pesar de que estos datos son matizados por el CONAPO —según esta fuente, el DF presentaba 34.8% de viviendas con algún nivel de hacinamiento hacia el 2005, mientras que para el 2010 esta cifra se ubicó en 26.1% (2011: 21)—, lo cierto es que casi una de cada tres viviendas presenta esta situación y que al año, al menos 35 mil

familias nuevas buscan satisfacer sus necesidades de vivienda y, cuando no son satisfechas, se integran a hogares ya formados (INVI, 2011).

En términos del modelo incremental, lo anterior podría significar que de un periodo a otro, las familias que se han desdoblado y no han ocupado o comprado una vivienda, estarían abonando al problema de la subestimación de la población que reside en una vivienda, problema reportado en las imágenes 7, 8 y 9.

De modo análogo, se plantea que el fenómeno del abandono o la desocupación de la vivienda podría ser un factor que esté influyendo en la sobrestimación de la población observada con anterioridad, al estimar un crecimiento poblacional derivado de un aumento de viviendas que, de facto, no están habitadas. Diversas fuentes consultadas (INEGI, 2010; Sánchez y Salazar, 2011) coinciden en que alrededor de 8% del total de viviendas en el DF se encuentra en esta situación, hecho que tampoco fue considerado en los modelos incrementales. A pesar de ser una de las tasas más bajas del país y de que siempre existirá en el mercado un porcentaje positivo de desocupación derivado de procesos y tiempos naturales del mercado inmobiliario (la absorción de la vivienda en el tejido urbano), puede ser un fenómeno que distorsione los resultados esperados.

Conclusiones

En este artículo hemos dado cuenta de diversos modelos que, de manera eventual, podrían ser utilizados para estimar de forma rápida un cambio en la cantidad de población en áreas pequeñas con base en información captada por sensores montados a bordo de plataformas aéreas y satelitales. La evidencia encontrada apunta a que es factible estimar la población en áreas pequeñas a partir del número de viviendas, el cual se puede inferir, a su vez y en ciertos casos, a partir de observaciones de percepción remota. Para determinar hasta qué punto es importante el estado de densificación del área donde se quiere estimar la población, se

corrieron pruebas para tres delegaciones representativas de densidades alta, media y baja. Se encontró que es posible construir un modelo de estimación de población en términos de viviendas, cuyo desempeño es casi independiente de la densidad de población (modelo P5); en otras palabras, si el número de viviendas se pudiera determinar con mucha precisión, entonces es posible hacer una estimación de la población con un error relativo de no más de 5 por ciento.

No obstante, al estimar el número de viviendas con datos de percepción remota y del catastro, la cota superior del error de estimación de la población incrementó hasta 15 por ciento. Hay que considerar que este nivel de error se obtiene cuando se hace el ajuste en dos etapas: primero para estimar la población mediante datos de viviendas reales y luego para estimar las viviendas a partir del área habitable. Es de esperar que el error pueda ser menor si se hace un ajuste directo entre área habitable y población, esto a expensas de complicar la optimización de parámetros ya que se imposibilitaría la regresión lineal simple.

Al observar la distribución espacial de los errores de estimación de población a nivel de AGEb, se notó un agrupamiento de los tipos de errores (sobrestimación y subestimación) en función de la concentración de uso de suelo habitacional. Un análisis más a fondo de los errores de estimación sugiere que la distribución espacial de los mismos refleja la heterogeneidad espacial de fenómenos socioeconómicos, como el abandono y el hacinamiento, de tal forma que resulta necesario conocer la distribución espacial de dichas variables para poder reducir los errores de estimación poblacional obtenidos en esta investigación.

Fuentes

Arnal, S. *Reglamento de Construcciones para el Distrito Federal*. México, Trillas, 2005.

Azar, D.; R. Engstrom; J. Graesser y J. Comenetz. "Generation of fine-scale population layers using multi-resolution satellite imagery and geospatial data", en: *Remote Sensing of Environment*. 130, 2013, pp. 219-232.

CONAPO. *Índice de marginación por entidad federativa y municipio 2010*. Consultado en http://www.conapo.gob.mx/work/models/CONAPO/indices_margina/mf2010/CapitulosPDF/1_4.pdf el 10 de septiembre de 2014.

Dobson, J. E.; E. A. Bright; P. R. Coleman; R. C. Durfee y B. A. Worley. "LandScan: A Global Population Database for Estimating Populations at Risk", en: *Photogrammetric Engineering & Remote Sensing*. 66(7), 2000, pp. 849-857.

Efroymson, M. "Multiple regression analysis", en: Ralston, A. y H. Wilf (eds.). *Mathematical Methods for Digital Computers*. New York, Wiley, 1960, pp. 191-203.

González, L. y E. Torres. "Estimaciones de población en áreas menores en América Latina: Revisión de métodos utilizados", en: Cavenaghi, S. *Estimaciones y proyecciones de población en América Latina: desafíos de una agenda pendiente*. Brasil, Asociación Latinoamericana de Población, 2012.

Grajales, S. G. *Estado de las ciudades de México 2011*. México, SEDESOL-Organización de las Naciones Unidas, 2011.

Harvey, J. T. "Estimating census district populations from satellite imagery: Some approaches and limitations", en: *International Journal of Remote Sensing*. 23(10), 2002, pp. 2071-2095.

Iisaka, J. "Population estimation from Landsat imagery", en: *Remote Sensing of Environment*. 12(4), 1982, pp. 259-272.

INEGI. *XII Censo General de Población y Vivienda*. Consultado en www.inegi.org.mx en el 2013.

_____. *Censo de Población y Vivienda 2010*. Consultado en www.inegi.org.mx en el 2013.

Instituto de Vivienda del Distrito Federal (INVI). *Evaluación anual de resultados 2010-2011*. Consultado en http://www.cgma.df.gob.mx/work/sites/cgma/docs/reportes/evaluanuales_resultados/EvalAnualINVI20121212.pdf en el 2014.

Llanos, R. "Predomina hacinamiento en un tercio de viviendas de la capital", en: *La Jornada*. 22 de febrero de 2012, p. 36.

Lu, D., Q. Weng y G. Li. "Residential population estimation using a remote sensing derived impervious surface approach", en: *International Journal of Remote Sensing*. 27(16), 2006, pp. 3553-3570.

Lwin, K. K. y Y. Murayama. "Estimation of building population from LIDAR derived digital volume model", en: *Spatial analysis and modeling in geographical transformation process*. 2011, pp. 87-98.

_____. "A GIS approach to estimation of building population for micro-spatial analysis", en: *Transactions in GIS*. 13(4), 2009, pp. 401-414.

Nordbeck, S. "The law of allometric growth", en: *Michigan Inter-University Community of Mathematical Geographers, Discussion Paper 7*. 1965.

_____. "Urban Allometric Growth", en: *Geografiska Annaler. Series B, Human Geography*. 53(1), 1971, pp. 54-67.

Ordorica, M. y V. Ibarra. "Aplicación del filtro de Kalman para realizar estimaciones de población en áreas pequeñas", en: Negrete Salas, M. E. *et al.*

- Entre fenómenos físicos y humanos*. México, El Colegio de Mexico, 2002, pp. 201-217.
- Qiu, F.; H. Sridharan y Y. Chun. "Spatial Autoregressive Model for Population Estimation at the Census Block Level Using LIDAR-derived Building Volume Information"; en: *Cartography and Geographic Information Science*. 37(3), 2010, pp. 239-257.
- Salinas Arreortua, L. A. *Transformaciones urbanas en el contexto neoliberal*. Consultado en http://contested-cities.net/CCmadrid/wp-content/uploads/sites/3/2013/11/2013_CC_LSalinas_Condesa.pdf el 9 de diciembre de 2014.
- Sánchez, L. y C. Salazar. "Lo que dicen las viviendas deshabitadas sobre el censo de población 2010", en: *Coyuntura Demográfica*, 1. Noviembre de 2011.
- Sociedad Hipotecaria Federal (SHF). *México: rezago habitacional, demanda de vivienda 2012 y bono demográfico*. Consultado en <http://www.shf.gob.mx/estadisticas/EstudiosVivienda/Documents/Rezago%20habitacional%20y%20Demanda%20de%20vivienda%202012.pdf>
- Silván-Cárdenas, J. "A multiscale erosion operator for discriminating ground points in LIDAR point clouds", en: Carrasco-Ochoa, J., J. Martínez-Trinidad, J. Salas-Rodríguez y G. S. Baja (eds.). *Pattern Recognition. LNCS 7914*. Querétaro, Springer Berlin Heidelberg, 2013, pp. 213-223.
- Silván-Cárdenas, J. y L. Wang. "Extraction of Buildings Footprint from LIDAR Altimetry Data with the Hermite Transform", en: *Pattern Recognition-MCPR'11 Proceedings of the Third Mexican conference on Pattern recognition*. 6718. Cancún, Mexico, Springer-Verlag Berlin Heidelberg, 2011, pp. 314-321.
- _____. "Multiscale approach for ground filtering from Lidar altimetry measurements", en: Weng, Q. *Scale Issues in Remote Sensing*. EE.UU., John Wiley y Sons, 2014, p. 18.
- Silvan-Cárdenas, J., J. Almazán-González y S. Couturier. "Remote Identification of Housing Buildings with High-resolution Remote Sensing", en: *Lecture Notes in Computer Sciences*. 8495. Cancún, Springer, 2014, pp. 380-390.
- _____. "Remote Identification of Housing Buildings with High-resolution Remote Sensing", en: *Lecture Notes in Computer Sciences*. 8495. Cancún, Springer, 2014, pp. 380-390.
- Silván-Cárdenas, J., L. Wang, C. Wu, P. Rogerson, T. Feng y B. Kamphaus. "Assessing fine-spatial-resolution remote sensing for small-area population estimation", en: *International Journal of Remote Sensing*. 31(21), 2010, pp. 5605-5634.
- Sridharan, H. y F. Qiu. "A Spatially Disaggregated Areal Interpolation Model Using Light Detection and Ranging Derived Building Volumes", en: *Geographical Analysis*. 45(3), 2013, pp. 238-258.
- Sutton, P. "Census from Heaven: an estimate of the global population using night-time satellite imagery", en: *International Journal and Remote Sensing*. 22, 2000, pp. 3061-3079.
- Sutton, P., D. Roberts, C. Elvidge y H. Melj. "A Comparison of Nighttime Satellite Imagery and Population Density for the Continental United States", en: *Photogrammetric Engineering & Remote Sensing*. 63(11), 1997, pp. 1303-1313.
- Wu, C. y A. T. Murray. "Population Estimation Using Landsat Enhanced Thematic Mapper Imagery", en: *Geographical Analysis*. 39(1), 2007, pp. 26-43.
- Yang, X., W. Yue y D. Gao. "Spatial improvement of human population distribution based on multi-sensor remote-sensing data: an input for exposure assessment", en: *International Journal of Remote Sensing*. 35(15), 2013, pp. 5569-5583.

Revisión metodológica de la Encuesta Nacional **sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003, 2006 y 2011**

José Antonio Mejía Guerra

Este trabajo tiene como objetivo presentar un análisis comparativo general sobre los esfuerzos de recolección de estadísticas de violencia contra la mujer que se han hecho en el Instituto Nacional de Estadística y Geografía (INEGI) en la forma de tres entregas de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares en los años 2003, 2006 y 2011. Se complementa con cuatro anexos.

Palabras clave: violencia contra la mujer, violencia familiar, violencia en el hogar, violencia doméstica, violencia laboral, violencia escolar, ENDIREH, encuestas en hogares.

Recibido: 19 de septiembre de 2014
Aceptado: 7 de enero de 2015

Nota: la elaboración de este documento no hubiera sido posible sin las significativas aportaciones de Eva Gicela Ramírez y Margarita Romero; el autor agradece también la colaboración de Clara Mantilla y el apoyo de Rosalba Jiménez.

The objective of this document is to present a comparative analysis of the data collection efforts that the National Institute of Statistics and Geography (INEGI) has undertaken to generate information on violence against women. Namely the three rounds of the National Survey on the Dynamics of Household Relations (ENDIREH) in the years 2003, 2006 and 2011. It is complemented with four annexes.

Key words: violence against women, family violence, violence in the household, domestic violence, workplace violence, school violence, ENDIREH, household surveys.



Obama's Halting of Deportations for Immigrants Is Too Late For Some /The Washington Post/Getty Images

Introducción

La violencia contra la mujer —ya sea en su entorno familiar, laboral o escolar— es una problemática social reconocida; sin embargo, para evidenciarla y dimensionarla, es necesario contar con instrumentos de recolección de información altamente especializados y complejos, como las encuestas en hogares. En el caso de la violencia contra la mujer, la dificultad no es sólo metodológica y logística, también es social, dado el bagaje cultural asociado al fenómeno, como el machismo y la estructura patriarcal de la sociedad, por mencionar algunos ejemplos.

En este artículo se hace un análisis comparativo general sobre los esfuerzos de recolección de estadísticas de violencia contra la mujer que el INEGI ha generado mediante tres entregas de la Encuesta

Nacional sobre la Dinámica de las Relaciones en los Hogares en los años 2003, 2006 y 2011. Al tratarse de una temática en la que a nivel mundial hay pocas recomendaciones metodológicas para su medición, se han dado cambios importantes en los contenidos y marcos conceptuales de las tres encuestas. Este documento busca ser un insumo útil para el proceso de preparación y diseño de la siguiente ENDIREH.

Antecedentes

En México, la recolección de información sobre violencia contra la mujer se remonta a la década de los 90. El primer esfuerzo fue el realizado por el Centro de Investigación y Estudios Superiores en Antropología Social (CIESAS) en 1994, que se denominó Encuesta sobre Organización Doméstica; en ese mismo año se levantó la Encuesta sobre

Organización Familiar por parte del Grupo de Educación Popular con Mujeres (GEM), AC; en 1998, El Colegio de México (COLMEX) aplicó la Encuesta sobre Dinámica Familiar; en 1999, el INEGI incursionó en esta temática con la Encuesta sobre Violencia Intrafamiliar, la cual tuvo una cobertura sólo para el Área Metropolitana de la Ciudad de México; y en el 2003, el Instituto Nacional de Salud Pública hizo un ejercicio de recolección de información a las usuarias de los servicios públicos de salud, la Encuesta Nacional sobre Violencia contra las Mujeres (INEGI, 2012a). Todos estos esfuerzos tuvieron una cobertura geográfica limitada.

En vista de la gran demanda por contar con información sobre violencia doméstica contra la mujer, el INEGI, en coordinación con el Instituto Nacional de las Mujeres (INMUJERES) y el Fondo de Desarrollo de las Naciones Unidas para las Mujeres (UNIFEM, ahora ONU Mujeres), realizó en el 2003 el levantamiento de la primera ENDIREH, con cobertura nacional y representatividad para algunas entidades federativas.

Posteriormente, en el 2006, se hizo el segundo levantamiento de la Encuesta, el cual incluyó otras formas y espacios de violencia contra la mujer (no sólo la que sucede en el hogar), además de tener representatividad para las 32 entidades federativas del país.

En el 2011 se llevó a cabo un tercer operativo de la ENDIREH con algunas adecuaciones metodológicas, pero manteniendo la cobertura nacional y para las 32 entidades federativas.

Las Naciones Unidas, en su publicación *Directrices para la producción de estadísticas sobre la violencia contra la mujer: encuestas estadísticas* (Naciones Unidas, 2013)¹ indica que el mejor instrumento para captar información completa sobre la violencia contra las mujeres son las encuestas en hogares. Ese documento menciona que son apenas poco más de 40 países en el mundo los que han realizado una en-

cuesta en hogares con este tema con cobertura nacional, y que en muy pocos casos las encuestas se han hecho de manera periódica y con información comparable en el tiempo, ambos elementos esenciales para lograr que los datos recolectados puedan ser utilizados para informar el diseño y seguimiento de políticas públicas.

La búsqueda de fuentes alternativas que suplan, o al menos complementen, la información que se recolecta con las encuestas en hogares es común a muchas temáticas. En muchos casos, la respuesta esperada o deseada es un mayor aprovechamiento estadístico de los registros administrativos; en general, la transición no es fácil dadas las limitaciones de cobertura y contenido que éstos tienen. En el caso particular de los registros administrativos sobre violencia contra las mujeres, es difícil pensar el poder dejar de hacer encuestas y utilizarlos como fuente, la razón es que hay un alto porcentaje de episodios que no se denuncian (por múltiples razones, por ejemplo, desconfianza de las autoridades, miedo al agresor, vergüenza, etc.) y más aún, existe una gran cantidad de tipos de violencia en contra de las mujeres que no están tipificados como delito, por lo que nunca aparecerán en ningún registro, y sólo podrán ser captados mediante encuestas en hogares.

Las ENDIREH de un vistazo

2003

Este segmento es un extracto de la sección de antecedentes del documento *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003. Características metodológicas* (INEGI, 2004).

La ENDIREH 2003 surgió de una solicitud realizada por el INMUJERES al INEGI para llevar a cabo una encuesta con representatividad nacional sobre la violencia en pareja. En conjunto con la solicitud, el INMUJERES entregó un documento en el que se planteaba el soporte conceptual de la Encuesta, sus objetivos, la temática y las variables a incluirse en el cuestionario. Fue respon-

¹ Este documento será publicado formalmente por Naciones Unidas, pero actualmente se encuentra una versión en su sitio en internet.

sabilidad del INEGI plasmar esos requerimientos en forma de preguntas en éste.

Para iniciar las actividades, se creó un grupo de trabajo en el que, además, se incorporó el UNIFEM; entre sus tareas estuvo el definir la integración del cuestionario, delimitar la población objetivo, buscar el interés de las entidades federativas para ampliar la cobertura de la muestra y formar una red de apoyo para aquellas mujeres que lo solicitaran. Un total de 11 entidades federativas contribuyeron con recursos para la ampliación de la muestra y contar con información representativa para cada una de ellas, además de los resultados a nivel nacional.

Las responsabilidades centrales del INEGI fueron la preparación del cuestionario, el diseño estadístico, la estrategia de campo, el reclutamiento y la capacitación del personal, así como el levantamiento, la captura y el procesamiento de los datos. A diferencia de otro tipo de encuestas en hogares, y siguiendo las mejores prácticas internacionales, se acordó contar con terapias de descarga para las entrevistadoras; este proceso quedó a cargo del INMUJERES y de sus contrapartes estatales.

Los resultados de la Encuesta tuvieron representatividad nacional y para las 11 entidades federativas que proporcionaron recursos financieros para la ampliación de la muestra en su ámbito territorial.

2006

Los siguientes párrafos son un extracto de las secciones de introducción y antecedentes del documento *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2006*. Marco conceptual (INEGI, 2007).

La información de la ENDIREH 2003 fue de gran utilidad para el gobierno, la sociedad civil y la Academia, por lo que se decidió volver a levantar una encuesta similar revisando su materia, universo y cobertura temática, de lo cual surgió una versión más completa al incluir otras formas de violencia contra la mujer, así como otros ámbitos y agre-

sores. El esquema de colaboración entre el INEGI, INMUJERES y UNIFEM repitió lo hecho en el 2003.

En esta edición de la ENDIREH se contó, además, con un apoyo importante de otras instancias, entre ellas: la Comisión Especial para Conocer y Dar Seguimiento a las Investigaciones Relacionadas con los Femicidios en la República Mexicana y a la Procuración de Justicia Vinculada, así como la Fiscalía Especial para la Atención de Delitos Relacionados con Actos de Violencia contra las Mujeres (FEVIM).

La ENDIREH 2006 tuvo cobertura nacional y, por primera vez, también para cada una de las 32 entidades federativas. Esta cobertura territorial y nivel de desagregación se consiguieron desde el inicio, pues ya no fue necesario solicitar recursos a éstas para ampliar la muestra.

2011

La información que se presenta a continuación proviene de la sección de introducción del documento *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2011*. Marco conceptual (INEGI, 2012a).

El tercer levantamiento de la ENDIREH surgió del gran impacto que la información de las dos rondas anteriores tuvo y en el deseo de dar continuidad a la disponibilidad de información estadística sobre el tema. Esta encuesta volvió a tener representatividad nacional y para cada una de las 32 entidades federativas; se le hicieron algunas modificaciones metodológicas, pero en general se buscó mantener la comparabilidad con la ENDIREH 2006.

De nueva cuenta, el INMUJERES participó en el diseño conceptual de la Encuesta, y hubo un amplio involucramiento por parte de la Cámara de Diputados. Al estar ya propiamente formado el Sistema Nacional de Información Estadística y Geográfica (SNIEG), se buscó la participación de distintas instancias de la administración pública general por medio del Comité Técnico Especializado en Información con Perspectiva de Género (CTEIPG).

De este vistazo en el tiempo se puede decir que, como en cualquier ejercicio de recolección de datos, el dilema al que se enfrenta el productor de información ante la necesidad de generar una nueva observación en la serie es: comparabilidad vs. innovación, es decir, se debe optar por mantener la continuidad metodológica y conceptual en la serie o se debe preferir la inclusión de elementos nuevos que mejoren el instrumento, aun cuando esto se haga a costa de la continuidad de la serie. Este dilema es particularmente complicado cuando se trata de temas en los que existe poca experiencia práctica y cuando tampoco se cuenta con recomendaciones metodológicas a nivel internacional. Esta limitación en estándares internacionales no es producto de la falta de interés en el tema sino, más bien, de su complejidad y de las diferencias culturales entre naciones, y aun dentro de un mismo país.

Estas complicaciones se plasmaron en el trabajo realizado por el Grupo de Amigos de la Presidencia de la Comisión Estadística de Naciones Unidas, al que se le solicitó preparar un manual técnico para la medición de la violencia contra la mujer; de ahí resultaron las directrices de las Naciones Unidas, documento que fue presentado en la 44.ª sesión de dicha comisión en el 2013, en el cual fue necesario incluir recomendaciones metodológicas para la medición de los diferentes tipos de violencia contra la mujer específicos de ciertas regiones del mundo, mismas que fueron agregadas como complementarias y opcionales al núcleo básico de indicadores de violencia contra las mujeres recomendado.

Cambios entre las rondas de la ENDIREH

En las tres se han realizado modificaciones importantes, que van desde adecuaciones al instrumento de recolección hasta los objetivos de la Encuesta. Todas se han hecho con el fin de mejorar la calidad de la información que se genera, pero no se ha elaborado un análisis detallado del impacto que éstas han tenido. En este documento se hace mención de algunos de los cambios más relevantes.

En el sentido más estricto, la comparación de los marcos conceptuales es más valiosa al realizarla entre las encuestas del 2006 y el 2011, pues ambas fueron levantamientos con una visión amplia del fenómeno de violencia contra la mujer, mientras que la ENDIREH 2003 se circunscribió a la violencia doméstica ejercida por la pareja conviviente; sin embargo, se subrayan algunas de las modificaciones realizadas con respecto a la primera entrega de la ENDIREH.

Entre las ENDIREH 2003 y 2006 se dio un cambio fundamental, que fue el de la población objetivo de la Encuesta: en la ronda del 2003 fueron las mujeres de 15 años o más con pareja residente en el hogar, y el tema central fue la violencia doméstica a manos de la pareja conviviente; en el 2006 se consideró a las mujeres de 15 años o más casadas o unidas, alguna vez unidas o solteras, es decir, se amplió el universo a todas las mujeres mayores de 15 años; y en la ENDIREH 2011 se mantuvo la misma población objetivo. Tanto en el 2006 como en el 2011 se trabajó con tres cuestionarios: uno para mujeres casadas o unidas, otro para mujeres alguna vez unidas y uno más para las solteras (nunca unidas).

Otro cambio importante entre la ronda del 2003 y la del 2006 fue la ampliación de los ámbitos en los que se da la violencia y el consecuente incremento en los potenciales agresores. En el 2006 se agregó la violencia en el trabajo, en la escuela y en otros ámbitos sociales, además de incluir a otros potenciales agresores (además de la pareja). Este esquema se mantuvo en la ENDIREH 2011.

También, hubo modificaciones en la selección de la mujer a entrevistar en cada vivienda visitada. En el 2003 se eligió a una por vivienda que cumpliera con las siguientes características: 15 años de edad o más, ser residente habitual en ésta y vivir en la actualidad con su pareja, independientemente de su situación conyugal. En el 2006 se escogió a una casada por unidad; si había mujeres solteras se aplicaba el módulo que les correspondía, y el de alguna vez unidas se utilizaba siempre y cuando no

hubiera mujeres casadas en la vivienda. En todos los casos, debía tratarse de mujeres de 15 años o más, que fueran residentes habituales de la vivienda. En caso de existir dos o más mujeres con el mismo estado conyugal se elegía a aquella cuya fecha de cumpleaños fuera la inmediata posterior a la fecha de la primera visita. En la ENDIREH 2011 se dio otro cambio en la metodología, pues se entrevistaron a todas las mujeres de 15 años y mayores residentes habituales de la vivienda, aplicando a cada una el cuestionario correspondiente a su situación conyugal.

Una diferencia importante entre las tres rondas de la ENIDREH es el tamaño de muestra. En el 2003 se calculó una de 20 160 viviendas considerando que los resultados fueran representativos a nivel nacional; al haber 11 entidades federativas que sufragaron los costos necesarios para que los resultados fueran representativos para ellas, la muestra se amplió a 57 230. En el 2006, desde el diseño se consideró tener representatividad a nivel nacional, urbano, rural y para las 32 entidades federativas, por lo que se estimó una muestra de 128 mil viviendas. En el 2011 se siguió ese mismo esquema.

Los periodos de levantamiento de la información han sido similares, entre los meses de octubre y noviembre: en el 2003 fue del 20 de octubre al 14 de noviembre, en el 2006 se realizó del 9 de octubre al 3 de noviembre y en el 2011 tuvo lugar entre el 3 de octubre y el 11 de noviembre.

Los cambios mencionados y otras variaciones (tamaño de muestra, etc.) se resumen en la tabla que se incluye como anexo 1 de este documento, la cual fue preparada por Margarita Romero.

El cuestionario también ha tenido modificaciones en su diseño, desde el cambio en el fraseo de algunas preguntas hasta la inclusión de nuevas preguntas, pasando por modificaciones en las opciones de respuesta a la misma pregunta. Debido a que los cambios se han dado de manera simultánea, es difícil distinguir cuál ha sido su impacto en los resultados obtenidos, es decir, qué variaciones en los niveles de

un indicador responden a cambios en la realidad y cuáles son efectos de las alteraciones del instrumento y del trabajo de campo. En el anexo 2 se detalla la cobertura temática en cada una de las encuestas y el número de preguntas para cada uno de los temas; la matriz fue preparada por Margarita Romero.

En general, las rondas del 2006 y 2011 son consistentes en cuanto a los elementos incluidos en los cuestionarios, lo que permite realizar una tipología y clasificación de eventos violentos que son comparables en términos generales. Los detalles de los tipos y clasificaciones de actos violentos en contra de las mujeres que fueron captados en las ENDIREH 2006 y 2011 se incluyen en la matriz que comprende el anexo 3, la cual fue preparada por Eva Gicela Ramírez, subdirectora de Estadísticas Sociales y de Violencia contra la Mujer; en ésta se incluyen tanto los tipos de violencia como los lugares en los que se llevó a cabo la agresión.

Existe un tema que ha tenido sólo algunas modificaciones entre las tres rondas, el referente a la *relación actual* de pareja, del cual se deriva el indicador de prevalencia de la violencia de pareja en los últimos 12 meses. Como ya se mencionó, este tipo de violencia fue el objetivo central de la ENDIREH 2003, mientras que en las rondas del 2006 y 2011 se amplió el foco para incluir violencia de otros agresores y en otros ámbitos diferentes al hogar. Las tasas de prevalencia de la violencia reciente (12 meses) de pareja entre mujeres casadas o unidas (de 15 años y más) ha variado de un nivel de 46.6% en el 2003 a 40% en el 2006, llegando a 27.3% en el 2011. Se requiere un análisis detallado de patrones y perfiles sociodemográficos para entender mejor el origen de la disminución en la tasa de prevalencia.

Para tener una mejor idea de los niveles y las tendencias de las tasas e indicadores sobre violencia contra las mujeres que se obtienen de las ENDIREH, es necesario hacer un análisis más profundo no sólo de los marcos conceptuales, sino también del trabajo de campo, captura de la información y de la base de datos, lo cual va más allá del alcance de este documento, pero es esencial

realizarlo en vista del proceso de planeación de la siguiente Encuesta, que se dará por vez primera en un contexto institucional distinto al de las tres ediciones anteriores dentro del INEGI. Es probable que lo más adecuado sea que el análisis detallado sea responsabilidad de expertos externos que puedan adentrarse en los marcos conceptuales, cuestionarios, experiencias de campo, bases de datos y vivencias de los participantes de las tres ENDIREH existentes.

Impacto de los cambios

Algunas de las variaciones que se han dado entre las rondas de la ENDIREH tienen impactos claros y explicables; por ejemplo, el cambio entre la edición del 2003 y la del 2006 en lo relativo a las mujeres objeto de la Encuesta implicó que en el 2003 pudiera verse como un subconjunto de la Encuesta del 2006 (y de la del 2011), ya que al expandirse el universo más allá de las mujeres de 15 años y más con pareja residente en el hogar, a las de 15 años y más casadas o unidas, alguna vez unidas y solteras se convirtió a la ENDIREH del 2006 en una encuesta de violencia contra la mujer, y no sólo en una de violencia doméstica contra la mujer; además, está la inclusión de otros agresores y la violencia en otros espacios, como: la escuela, el trabajo, etcétera.

Por otro lado, el cambio entre la ENDIREH 2006 y la 2011 en la selección de la mujer a entrevistar, que implicó que en el 2011 se entrevistara a todas las mujeres de 15 años y más residentes en el hogar, debió estudiarse con mayor detenimiento. En términos numéricos, partiendo de una muestra del mismo tamaño (128 mil viviendas), la decisión de encuestar a todas las mujeres impactó el número de entrevistadas, particularmente en la categoría de *alguna vez unidas* (ver detalle en el anexo 1), al pasar de 15 733 a 27 203.

De acuerdo con el documento de directrices de Naciones Unidas, esto puede tener un impacto en la tasa de respuesta y en la calidad de la información, ya que es posible que, al enterarse del contenido

de la encuesta después de que se ha entrevistado a la primera mujer de la vivienda, el resto decida no participar, lo cual podría generar un sesgo si es que la selección de la primera mujer a entrevistarse en cada vivienda no se hace de manera sistemática previendo la posibilidad de atrición descrita.

Un ejercicio básico de comparación entre los datos del 2006 y 2011 (Romero, Margarita, 2012) muestra que no hay diferencias significativas en las tasas de incidencia de violencia entre las viviendas con una sola entrevista y aquéllas con dos o más, sugiriendo que la entrevista de todas las mujeres en el hogar no implicó un sesgo como el que las recomendaciones internacionales temen que pueda existir (ver anexo 4); sin embargo, es necesario realizar un análisis más profundo de los resultados.

Otro cambio importante entre el 2006 y 2011 es la forma en la que se preguntó sobre agresiones sufridas por las mujeres por parte de personas distintas a su pareja. En el 2006, las preguntas se organizaron por espacio físico (trabajo, escuela y otros —calle, transporte público, etc.—), pero sin ligar de manera implícita al agresor con el espacio de la agresión. Además, fue establecido el último año como periodo de referencia para agresiones conectadas a su relación laboral, la época de estudiante como referencia temporal para ataques ligados con el ambiente escolar y toda la vida para agresiones en otros espacios.

Por su parte, el cuestionario del 2011 implementó una matriz de tipos de agresión cruzados con agresor y lugar en el que ocurrieron, y tenían dos espacios de tiempo: alguna vez en la vida y en el último año. Esto permitió establecer de manera explícita el lugar de la agresión, aunque no quedaba claro si el vínculo entre la mujer agredida y el agresor estaba dado por el tipo de relación entre ambos o por el espacio en el que se dio el ataque. La estructura de la matriz utilizada en el 2011 tuvo la estructura que posteriormente sería planteada en las *Directrices para la producción de estadísticas sobre la violencia contra la mujer: encuestas estadísticas*.

Recomendaciones internacionales

El proceso de diseño conceptual, planeación y ejecución del trabajo de campo, así como el análisis de la información de las ENDIREH 2003, 2006 y 2011, estuvo basado en los documentos metodológicos, experiencias de países y estudios académicos disponibles en sus respectivos periodos de diseño, además de que en el 2006 y 2011 se contó con la experiencia de encuestas anteriores; sin embargo, a diferencia de otras temáticas, en el caso de la captación de información de violencia contra las mujeres utilizando encuestas en hogares no existía una recomendación metodológica que tuviera el respaldo de Naciones Unidas. No obstante, como ya se comentó, a partir del 2013 se cuenta con el documento *Directrices para la producción de estadísticas sobre la violencia contra la mujer: encuestas estadísticas*.

Este documento es producto de más de tres años de trabajo detonado por un mandato de la Asamblea General de las Naciones Unidas (resolución 61/143), lo que se tradujo en la conformación de un Grupo de Amigos del Presidente en el seno de la Comisión Estadística de las Naciones Unidas (39.ª sesión, 2008). Este grupo estuvo encabezado por México, tuvo como secretariado técnico a la División de Estadística de las Naciones Unidas (UNSD, por sus siglas en inglés) y lo integraron representantes de las oficinas de estadística de Australia, Bangladesh, Botsuana, Bulgaria, Canadá, Chile, Costa Rica, Egipto, Ghana, Tailandia, Turquía y los Estados Unidos de América.

Las directrices fueron preparadas para proveer a los países elementos para generar un grupo básico de indicadores de violencia contra la mujer utilizando encuestas en hogares. El documento incorpora recomendaciones que van desde el diseño conceptual de la encuesta hasta el muestreo, el trabajo de campo y la presentación de resultados.

Asimismo, indican que el mejor instrumento para lograr tener información de calidad de este tema son las encuestas en hogares. Sugiere que, dentro de lo posible, éstas sean específicas sobre

el rubro y no módulos o preguntas en otros ejercicios de recolección de información. Subraya que la capacitación de las encuestadoras es aún más importante que en otras encuestas debido a la sensibilidad del tema.

Su ejecución conlleva consideraciones éticas y de seguridad mayores que otras encuestas. No es suficiente cumplir con los principios fundamentales de la ética estadística, hay que ir más allá dado lo sensible del tema, pues las mujeres entrevistadas no están hablando sólo de su entorno, expresan experiencias traumáticas muy concretas vividas por ellas. Por otro lado, la temática de la encuesta implica tomar consideraciones extraordinarias para garantizar la seguridad de la mujer entrevistada y de la entrevistadora. Este tipo de encuestas implican, además, contar con información para referir a las mujeres que lo soliciten a servicios e instituciones de apoyo, así como tener personal especializado para apoyar psicológicamente a las entrevistadoras, dada la carga emocional que conlleva una encuesta de esta naturaleza.

La relevancia de la información que se obtiene de encuestas especializadas de violencia contra la mujer se basa en la cantidad de información complementaria que producen, es decir, no son sólo una cuenta de experiencias de violencia sino que, además, incluyen datos sociodemográficos de la víctima, las circunstancias en las que se da la violencia y el acceso (o falta de) a servicios legales o de salud. Asimismo, permiten captar tipos de violencia, como la psicológica y la económica, que no es posible estimar por otros medios.

También, se subraya la importancia de que la oficina de estadística (recolectora de la información) trabaje de manera muy cercana con las instituciones que serán usuarias de la información, para garantizar que la encuesta cumpla con todos los requerimientos conceptuales y metodológicos para hacerla una herramienta útil.

En términos generales, la ENDIREH cubre los requerimientos marcados en los párrafos anteriores, aunque fue hasta la ronda del 2006 de la Encuesta

en la que se amplió la cobertura temática para incluir violencia en otros espacios (distintos al hogar) y por agresores diferentes a la pareja o expareja.

El futuro

En el caso del diseño y planeación de la siguiente ENDIREH, el INEGI cuenta con el beneficio de la experiencia de las tres rondas anteriores y, además, con las *Directrices para la producción de estadísticas sobre la violencia contra la mujer: encuestas estadísticas*.

En términos organizativos, dentro del INEGI es un gran paso el que ahora la responsabilidad de diseño y análisis de la ENDIREH esté en un área especializada en la temática de la violencia contra las mujeres, esto facilita la continuidad y especialización del personal involucrado. El siguiente reto institucional es lograr que la ENDIREH deje de ser una encuesta *especial* y se convierta en una *tradicional*, con una cobertura geográfica y periodicidad definidas, y no sujeta a elementos no técnicos o a la disponibilidad de recursos presupuestales.

Por otro lado, el reto es *hacer un corte de caja* de lo logrado hasta ahora con las tres rondas de la ENDIREH, sistematizar lo aprendido (tanto lo positivo como lo negativo) y plasmar esas lecciones en la siguiente edición (marco conceptual, diseño de cuestionario, trabajo de campo, etcétera). Este documento pretende ser un insumo inicial que sirva para lograr este objetivo; sin embargo, corresponde al área responsable del diseño de la siguiente ENDIREH hacer un análisis con mayor profundidad e integrar lo que podría ser la versión definitiva de esta encuesta.

Además, sería conveniente tener un análisis externo, por parte de expertos en temas de violencia contra la mujer, que permita tener una visión sin sesgos de lo logrado con las ENDIREH y de los retos a futuro para esta encuesta en México. Este análisis podría aprovecharse para generar un estudio de lecciones aprendidas en la ejecución de encuestas sobre violencia contra la mujer en México, documento que sería de gran utilidad en otros países

que no han tenido la experiencia de contar con tres rondas de este tipo de encuestas, y que se beneficiarían del aprendizaje positivo y negativo del trabajo realizado por el INEGI.

En el documento *Violencia de género en las parejas mexicanas. Análisis de resultados de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2006*, los autores lo plantean así: "Lo ideal sería establecer una periodicidad en el levantamiento de información de cinco años, que permita evidenciar cambios estructurales en la prevalencia de la violencia contra la mujer y de los factores asociados a la misma, más allá de fluctuaciones atribuibles al uso de instrumentos diferentes. En ese sentido, además de abogar por un levantamiento periódico de encuestas sobre violencia, insistimos en la necesidad de la adopción de un instrumento fundamentalmente homogéneo, fruto de una evaluación rigurosa como los que se han empleado hasta ahora..." (Castro y Casique, 2008).

Para esto, es importante considerar las tres rondas de la ENDIREH como instrumentos de aprendizaje que le permitirán al INEGI, y al país, consolidar una herramienta de recolección de datos de primer nivel sobre el tema de violencia contra las mujeres.

Fuentes

Castro, Roberto e Irene Casique (coords.). *Violencia de género en las parejas mexicanas. Análisis de resultados de la Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares 2006*. México, Instituto Nacional de las Mujeres y Centro Regional de Investigaciones Multidisciplinarias-UNAM, 2008, 176 pp. Consultado en septiembre del 2013 en: http://cedoc.inmujeres.gob.mx/documentos_download/100925.pdf.

INEGI. *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2003. Características metodológicas*. México, INMUJERES, UNIFEM, INEGI y PNUD, 2004, 147 pp. Consultado en octubre del 2013 en: <http://www.inegi.org.mx/est/contenidos/espanol/metodologias/encuestas/hogares/metodendireh.pdf>.

_____. *Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2006. Marco conceptual*. México, INEGI, 2007, 33 pp. Consultado en octubre del 2013 en: http://www.inegi.org.mx/est/contenidos/espanol/metodologias/encuestas/hogares/ENDIREH06_marco_concep.pdf.

“ENDIREH-2006’S Achievements and limitations in determining indicators for measuring violence against women in Mexico”, en: *Working Paper 7*. INEGI, 2007a. Documento preparado por Eva Gicela Ramírez Rodríguez. Subdirección de Estadísticas Sociales y de Violencia contra las Mujeres. Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia, INEGI, y presentado en el *Expert Group Meeting on Indicators to Measure Violence against Women*, co-organizado por UNDAW, UNECE y UNSD en colaboración con ECA/ECLAC/ESCAP/ESCWA, Ginebra, Suiza, del 8 al 10 de octubre. Consultado en septiembre del 2013 en: <http://www.unece.org/fileadmin/DAM/stats/documents/ece/ces/ge.30/2007/mtg1/wp.7.e.pdf>.

Encuesta Nacional sobre la Dinámica de las Relaciones en los Hogares (ENDIREH) 2011. Marco conceptual. México, INEGI, 2012a, 24 pp. Consultado en octubre del 2013 en: <http://www.inegi.org.mx/est/contenidos/Proyectos/Encuestas/Hogares/especiales/endireh/endireh2011/default.aspx>.

Comentarios preliminares sobre algunos resultados ENDIREH-2011. México, INEGI, 2012b. Documento interno de trabajo preparado por Eva Gicela Ramírez Rodríguez. Subdirección de Estadísticas Sociales y de Violencia contra las Mujeres, Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia, México. INEGI. Consultado en agosto del 2013.

Análisis estadístico para determinar si el haber entrevistado a todas las mujeres de 15 años y más residentes en las viviendas seleccionadas en la ENDIREH 2011 provocó disminución en la declaración de violencia de pareja. México, INEGI, 2012c. Documento interno de trabajo preparado por Margarita Romero Centeno. Dirección de Estadística. Dirección Regional Oriente, INEGI. Consultado en agosto del 2013.

Algunas anotaciones para estimar las prevalencias de violencia contra las mujeres, por tipo y modalidad a partir de los resultados de la ENDIREH-2011 y su comparación con la ENDIREH-2006. México, INEGI, 2013a. Documento interno de trabajo preparado por Eva Gicela Ramírez Rodríguez. Subdirección de Estadísticas Sociales y de Violencia contra las Mujeres, Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia, INEGI. Consultado en agosto del 2013.

Nota técnica sobre la estimación de la prevalencia total de violencia contra las mujeres a lo largo de sus vidas. México, INEGI, 2013b. Documento interno de trabajo preparado por Eva Gicela Ramírez Rodríguez. Subdirección de Estadísticas Sociales y de Violencia contra las Mujeres, Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia, INEGI. Consultado en agosto del 2013.

Naciones Unidas. *Intensification of efforts to eliminate all forms of violence against women (Resolution adopted by the General Assembly on December 19 2006) A/RES/61/143*. Versión en inglés. Naciones Unidas, 2006, 7 pp. Consultado en octubre del 2013 en: http://www.un.org/en/ga/search/view_doc.asp?symbol=A/RES/61/143&Lang=E.

Directrices para la producción de estadísticas sobre la violencia contra la mujer: encuestas estadísticas. Versión en inglés. Naciones Unidas, 2013. 395 pp. Consultado en octubre del 2013 en: http://unstats.un.org/unsd/gender/docs/guidelines_VAW.pdf.

Naciones Unidas. Comisión de Estadística. *Informe sobre el 39.º periodo de sesiones (26 al 29 de febrero de 2008) E/2008/24-E/CN.3/2008/34*. Versión en español. Naciones Unidas, 2008, 44 pp. Consultado en octubre del 2013 en: <http://unstats.un.org/unsd/statcom/doc08/Report-Spanish.pdf>.

Anexo 1

Continúa

Características operativas

Característica	2003	2006	2011
Unidad de observación	Viviendas seleccionadas y sus hogares.	Viviendas seleccionadas y sus hogares.	Viviendas seleccionadas y sus hogares.
Población objeto de estudio	La población objeto de estudio son las mujeres de 15 años y más con pareja residente en el hogar.	Mujeres de 15 años y más casadas o unidas; alguna vez unidas y solteras.	Mujeres de 15 años y más casadas o unidas; alguna vez unidas y solteras.
Representatividad	<ul style="list-style-type: none"> Nacional. Estatad (sólo para 11 entidades: Baja California, Coahuila de Zaragoza, Chiapas, Chihuahua, Hidalgo, Michoacán de Ocampo, Nuevo León, Quintana Roo, Sonora, Yucatán y Zacatecas). 	<ul style="list-style-type: none"> Nacional. Nacional urbano. Nacional rural. Estatad. 	<ul style="list-style-type: none"> Nacional. Nacional urbano. Nacional rural. Estatad.

Características operativas

Característica	2003	2006	2011
Esquema de muestreo	<ul style="list-style-type: none"> • Probabilístico. • Estratificado. • Trietápico. • Por conglomerados. 	<ul style="list-style-type: none"> • Probabilístico. • Estratificado. • Bietápico. • Por conglomerados. 	<ul style="list-style-type: none"> • Probabilístico. • Estratificado. • Bietápico. • Por conglomerados.
Marco de la muestra	Marco Nacional de Viviendas 2002.	Marco Nacional de Viviendas 2002.	Marco Nacional de Viviendas 2002.
Selección de la muestra	Por entidad, dominio y estrato.	Por entidad, dominio y estrato.	Por entidad, dominio y estrato.
Tamaño de la muestra	20 160 viviendas (para las 11 entidades, se amplió la muestra a 4 mil viviendas, por lo que el total nacional fue de 57 230)	128 mil viviendas (4 mil por entidad).	128 mil viviendas (4 mil por entidad)
Periodo de levantamiento	Del 20 de octubre al 14 de noviembre de 2003.	Del 9 de octubre al 3 de noviembre de 2006.	Del 3 de octubre al 11 de noviembre de 2011.
Estrategia de selección	Se elige a una mujer por vivienda y debe cumplir con las siguientes características: tener 15 años de edad o más, ser residente habitual de la vivienda y vivir actualmente con su pareja, independientemente de su situación conyugal.	Se elige a una mujer casada por vivienda; en caso de haber mujeres solteras se aplica también un módulo de solteras; el módulo de alguna vez unidas se aplica siempre y cuando no haya casadas en la vivienda; todas ellas deben cumplir con las siguientes características: tener 15 años de edad o más, ser residente habitual de la vivienda. En caso de existir dos mujeres con el mismo estado conyugal, se elige a aquella que su fecha de cumpleaños sea el inmediato posterior a la fecha de la primera visita.	Se aplica un cuestionario por cada mujer de 15 años o más residente de la vivienda; el instrumento a aplicar es de acuerdo con su estado conyugal.
Instrumentos de captación	General y secciones para mujeres casadas o unidas.	General y secciones para mujeres casadas o unidas. Módulo para mujeres divorciadas, separadas y viudas. Módulo para mujeres solteras	General y secciones para mujeres casadas o unidas Cuestionario para mujeres alguna vez unidas (divorciadas, separadas y viudas). Cuestionario para mujeres nunca unidas (solteras)
Mujeres entrevistadas			
Casadas	34 184	83 159	87 169
Alguna vez unidas	-	15 773	27 203
Solteras	-	26 444	28 300
Total	34 184	125 376	142 672
Dato expandido			
Casadas	19 471 972	21 631 993	24 569 503
Alguna vez unidas	-	3 551 020	7 227 737
Solteras	-	8 092 748	7 424 805
Total	19 471 972	35 756 378	39 222 045

Fuente: Romero Centeno, Margarita con base en las ENDIREH 2003, 2006 y 2011 (INEGI).

Cobertura temática y número de preguntas

Temas	2003			2006			2011		
	Casadas o unidas	Alguna vez unidas	Solteras	Casadas o unidas	Alguna vez unidas	Solteras	Casadas o unidas	Alguna vez unidas	Solteras
Características de la vivienda							6		
Características de la vivienda e identificación de hogares	12			9					
Residentes y hogares							3		
Datos sociodemográficos	18			14			11		
Introducción (verificación de estado conyugal, ámbito laboral y otros ingresos)	19								
Verificación de estado conyugal de mujeres de 15 años y más							5		
Lengua indígena y otros ingresos							7	7	5
Ámbito laboral, escolar y social				25	23	11	14	14	14
Familia de origen y situación patrimonial				15	15	5	8	8	
Familia de origen	20								
Relación con los hijos	9								
Situación patrimonial									2
Vida en pareja	7			19	16	1	13	13	6
Tensiones y conflictos	7			5	5	1	3	3	2
Relación actual o anterior	22			24	32	5	24	32	11
Decisiones	1			2			2		
Aportes económicos y disponibilidad de recursos				5					
Aportes económicos	1								
Disponibilidad de recursos	6				2		3	3	

Concluye

Cobertura temática y número de preguntas

Temas	2003			2006			2011		
	Casadas o unidas	Alguna vez unidas	Solteras	Casadas o unidas	Alguna vez unidas	Solteras	Casadas o unidas	Alguna vez unidas	Solteras
Libertad personal	5			1			1		1
Opinión sobre los roles masculinos y femeninos	9			4	4	4	3	3	3
Recursos sociales	5			2	2		2	2	
División del trabajo en el hogar	1						1	1	1
Mujeres de 60 años y más				7	7		8	8	8
Total	142			132	106	27	114	94	53

Fuente: Romero Centeno, Margarita con base en las ENDIREH 2003, 2006 y 2011 (INEGI).

Anexo 3

Continúa

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2006

Tipo/clase de violencia	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	En la escuela (sólo quienes declararon haber asistido a la escuela)	De cualquier otra persona (de la comunidad)	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia actual (todas)	En la familia de origen
Violencia física						
Violencia física	<ul style="list-style-type: none"> • Empujado o jalado. • Amarrado. • Pateado. • Aventado objetos. • Tratado de asfixiar o ahorcar. • Golpeado con las manos o con un objeto. • Agredido con cuchillo o arma. 	<ul style="list-style-type: none"> • Agresiones físicas. 		<ul style="list-style-type: none"> • Agresiones físicas. 	<ul style="list-style-type: none"> • Golpeado o agredido físicamente. 	<ul style="list-style-type: none"> • La golpeaban. • Presenció golpes entre integrantes de su familia.
Violencia sexual						
Intimidación sexual			<ul style="list-style-type: none"> • Le hicieron sentir miedo de sufrir un ataque o abuso sexual. • Recibido piropos groseros u ofensivos sobre su cuerpo o de tipo sexual. 			

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2006

Tipo/clase de violencia	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	En la escuela (sólo quienes declararon haber asistido a la escuela)	De cualquier otra persona (de la comunidad)	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia actual (todas)	En la familia de origen
Acoso sexual		<ul style="list-style-type: none"> Propuestas de relaciones sexuales a cambio de calificaciones. Castigos o represalias por negarse a las pretensiones de carácter sexual. 		<ul style="list-style-type: none"> Insinuaciones o propuestas de relaciones sexuales a cambio de mejoras en el trabajo. Represalias por negarse a las pretensiones de carácter sexual. 		
Abuso sexual	<ul style="list-style-type: none"> Le exige tener relaciones sexuales no deseadas. La obliga a hacer actos sexuales no deseados. Usa la fuerza física para obligarla a tener relaciones sexuales. 	<ul style="list-style-type: none"> Manoseos y caricias no consentidas. Relaciones sexuales forzadas. 	<ul style="list-style-type: none"> Manoseos y caricias no consentidas. Relaciones sexuales forzadas. Actos sexuales forzados por dinero. 	<ul style="list-style-type: none"> Manoseos y caricias no consentidas. Relaciones sexuales forzadas. 	<ul style="list-style-type: none"> Manoseos y caricias no consentidas. Relaciones sexuales forzadas. 	
Violencia emocional o psicológica						
Indiferencia	<ul style="list-style-type: none"> La ignoran, no la toman en cuenta, no le brindan atención. Le deja de hablar. 				<ul style="list-style-type: none"> La ignoran, no la toman en cuenta, no le brindan atención. 	
Degradación	<ul style="list-style-type: none"> Avergüenza o menosprecia. Le dice que la engaña. Ofensas por incumplimiento de obligaciones domésticas. 	<ul style="list-style-type: none"> Humillaciones y degradación. La ignoran o le hacen sentir menos por ser mujer. 		<ul style="list-style-type: none"> Humillaciones y degradación. La ignoran o le hacen sentir menos por ser mujer. 	<ul style="list-style-type: none"> Insultos u ofensas. Humillaciones y menosprecio. 	<ul style="list-style-type: none"> La insultaban u ofendían. Presenció insultos u ofensas entre integrantes de su familia.
Intimidación	<ul style="list-style-type: none"> Le hace sentir miedo. La vigila, espía o persigue. Destruye, tira o esconde cosas. Entrado o intentado entrar a su casa por la fuerza o sin su consentimiento (desde la separación o divorcio). 					

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2006

Tipo/clase de violencia	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	En la escuela (sólo quienes declararon haber asistido a la escuela)	De cualquier otra persona (de la comunidad)	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia actual (todas)	En la familia de origen
Aislamiento	<ul style="list-style-type: none"> • Encierro y prohibición de salir o de recibir visitas. • Poner en contra suya a hijos y/o pariente. • Controla o domina sus movimientos (solteras). 				<ul style="list-style-type: none"> • Encierro. 	
Amenazas	<ul style="list-style-type: none"> • Amenaza con irse, dañarla, quitarle a los hijos o correrla. • Amenaza con arma. • Amenaza con matarla, matarse él o matar a los hijos. 				<ul style="list-style-type: none"> • Amenazas de golpes o correrla de casa. 	
Violencia económica y patrimonial						
Despojo y coerción	<ul style="list-style-type: none"> • Se apropia o le quita dinero o bienes. • Le quitó sus pertenencias (en la separación o después de la separación). 		<ul style="list-style-type: none"> • Quitado o robado bienes o propiedades. • Quitado o robado documentos. • Poner a nombre de otro sus bienes o propiedades. 		<ul style="list-style-type: none"> • Le quitan dinero o le obligan a entregarlo. 	
Control o chantaje económico	<ul style="list-style-type: none"> • Reclamos por cómo gasta el dinero. • Es codo o tacaño con los gastos para la casa. • Le prohíbe trabajar o estudiar. 					
Incumple responsabilidad económica	<ul style="list-style-type: none"> • No cumple con dar el gasto o amenaza con no darlo. • Se gasta el dinero que se necesita para la casa. • No cumple con el acuerdo establecido en la separación o divorcio. 					

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2006

Tipo/clase de violencia	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	En la escuela (sólo quienes declararon haber asistido a la escuela)	De cualquier otra persona (de la comunidad)	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia actual (todas)	En la familia de origen
Discriminación laboral				<ul style="list-style-type: none"> • Solicitaron prueba de embarazo para ingresar. • Despido, no contratación o disminución del salario por embarazarse. • Menor salario que a un hombre que hace lo mismo. • Menos prestaciones que un hombre. • Menor oportunidad para ascender que un hombre. • Despido, no contratación o disminución del salario por edad o estado civil. 		

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2011

	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	De cualquier otra persona diferente a la pareja o esposo	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia de origen
Violencia física				
Violencia física	<ul style="list-style-type: none"> • Empujado o jalado. • Amarrado. • Pateado. • Aventado objetos. • Golpeado con las manos o con un objeto. • Tratado de asfixiar o ahorcar. • Agredido con cuchillo o arma. 	<ul style="list-style-type: none"> • La han agredido físicamente. 		<ul style="list-style-type: none"> • La golpeaban. • Presenció golpes entre integrantes de su familia.
Violencia sexual				
Intimidación sexual		<ul style="list-style-type: none"> • Le han hecho sentir miedo de ser atacada o abusada sexualmente. • Le han dicho piropos o frases de carácter sexual que le molesten u ofendan. • La han obligado a mirar escenas o actos sexuales (exhibicionistas, pornografía). 		
Acoso sexual		<ul style="list-style-type: none"> • Insinuaciones o propuestas de relaciones sexuales a cambio de mejoras en el trabajo. • Represalias o castigos por negarse a las propuestas de índole sexual. 		
Abuso sexual	<ul style="list-style-type: none"> • Le exige tener relaciones sexuales no deseadas. • La obliga a hacer actos sexuales no deseados. • Usa la fuerza física para obligarla a tener relaciones sexuales. 	<ul style="list-style-type: none"> • Tocado o manoseado sin su consentimiento. • Obligado a tener relaciones sexuales. • Obligado a realizar actos sexuales por dinero. 		
Violencia emocional o psicológica				
Indiferencia	<ul style="list-style-type: none"> • La ignoran, no la toman en cuenta, no le brindan atención. • Le deja de hablar. 	<ul style="list-style-type: none"> • La han ignorado, no la toman en cuenta. 		
Degradación	<ul style="list-style-type: none"> • Avergüenza o menoprecia. • Le dice que la engaña. • Ofensas por incumplimiento de obligaciones domésticas. 	<ul style="list-style-type: none"> • La han humillado o denigrado. 		<ul style="list-style-type: none"> • La insultaban u ofendían. • Presenció insultos u ofensas entre integrantes de su familia.
Intimidación	<ul style="list-style-type: none"> • Le hace sentir miedo. • La vigila, espía o persigue. • Destruye, tira o esconde cosas. • Entrado o intentado entrar a su casa por la fuerza o sin su consentimiento (desde la separación o divorcio). 			
Aislamiento	<ul style="list-style-type: none"> • Encierro y prohibición de salir o de recibir visitas. • Poner en contra suya a hijos y/o pariente. • Controla o domina sus movimientos (solteras). 			

Tipología y clasificación de los eventos de violencia captados por la ENDIREH 2011

	Del actual o último esposo o pareja (las solteras, sólo quienes han tenido novio o pareja)	De cualquier otra persona diferente a la pareja o esposo	En el trabajo (sólo quienes trabajaron el año anterior)	En la familia de origen
Amenazas	<ul style="list-style-type: none"> • Amenaza con irse, dañarla, quitarle a los hijos o correrla. • Amenaza con arma. • Amenaza con matarla, matarse él o matar a los hijos. 	<ul style="list-style-type: none"> • La han amenazado con correrla de su casa. 		
Violencia económica y patrimonial				
Despojo y coerción	<ul style="list-style-type: none"> • Se apropia o le quita dinero o bienes. • Le quitó sus pertenencias (en la separación o después de la separación). • Le han quitado bienes o propiedades. • Le han quitado documentos que prueben que usted es propietaria de algún bien. • Le han obligado a poner a nombre de otra persona las cosas o propiedades que usted compró o heredó. 	<ul style="list-style-type: none"> • Le han quitado bienes o propiedades. • Le han quitado documentos que prueben que usted es propietaria de algún bien. • Le han obligado a poner a nombre de otra persona las cosas o propiedades que usted compró o heredó. 		
Control o chantaje económico	<ul style="list-style-type: none"> • Reclamos por cómo gasta el dinero. • Es codo o tacaño con los gastos para la casa. • Le prohíbe trabajar o estudiar. 			
Incumple responsabilidad económica	<ul style="list-style-type: none"> • No cumple con dar el gasto o amenaza con no darlo. • Se gasta el dinero que se necesita para la casa. • No cumple con el acuerdo establecido en la separación o divorcio. 			
Discriminación laboral			<ul style="list-style-type: none"> • Solicitaron prueba de embarazo para ingresar. • Despido, no contratación o disminución del salario por embarzarse. • Menor salario que a un hombre que hace lo mismo. • Menos prestaciones que un hombre. • Menor oportunidad para ascender que un hombre. • Despido, no contratación o disminución del salario por edad o estado civil. 	

Fuente: INEGI. *Algunas anotaciones para estimar las prevalencias de violencia contra las mujeres, por tipo y modalidad a partir de los resultados de la ENDIREH-2011 y su comparación con la ENDIREH-2006.* México, INEGI, 2013a. Documento interno de trabajo preparado por Eva Gicela Ramírez Rodríguez. Subdirección de Estadísticas Sociales y de Violencia contra las Mujeres, Dirección General de Estadísticas de Gobierno, Seguridad Pública y Justicia, INEGI.

Anexo 4

Comparación de tasas de prevalencia de la violencia de pareja en viviendas con dos o más entrevistas para las ENDIREH 2006 y 2011

En este anexo se presenta el trabajo de análisis realizado por Margarita Romero de la Dirección de Estadística de la Dirección Regional Oriente del INEGI.

Para tratar de determinar si el tipo de operativo de campo implementado en la ENDIREH 2011 afectó la declaración de violencia de pareja, es necesario analizar los resultados de la Encuesta para aquellas viviendas en las que se aplicaron más de tres entrevistas, o dos, pero en una situación distinta a la de la ENDIREH 2006. En total, en la ENDIREH 2011 hubo 27.4% de las viviendas en esta situación.

Para realizar el análisis, se dividió a las viviendas en tres grupos: aquéllas donde se aplicó una combinación de entrevistas similar a la del 2006 (casada; alguna vez unida; soltera; casada y soltera; alguna vez unida y soltera); en las que se realizaron dos entrevistas, pero con una combinación que no se tuvo en el 2006 (dos casadas; dos alguna vez unidas; dos solteras; casada y alguna vez unida); y aquéllas con aplicación de tres o más entrevistas.

La distribución de las mujeres entrevistadas, de acuerdo con estos grupos, se aprecia en el cuadro 1.

Es decir, del total de mujeres entrevistadas, 63.7% vive en viviendas donde se aplicó una o dos entrevistas en una situación igual a la del 2006; 13.3% proviene de viviendas donde se aplicaron dos entrevistas, pero en una situación distinta a la de 2006; por último, 23% de las mujeres entrevistadas viven en viviendas donde se aplicaron tres o más entrevistas.

Una mirada rápida a esta distribución arroja visos de consistencia: las casadas o unidas viviendo mayoritariamente en hogares nucleares (71.7%); ¿dos solteras sin casadas o alguna vez unidas?, son, en su mayoría, chicas jóvenes no necesariamente familiares compartiendo vivienda, y es el grupo por mucho minoritario de las solteras: 4.5%; etcétera.

En los cuadros 2 y 3 se muestran los porcentajes de violencia a lo largo de la relación y en los últimos 12 meses para el 2006, 2011, en general, y 2011 según los tres grupos mencionados. Incluyen el *no especificado* prorrateado. En ningún caso, el grupo de mujeres donde se aplicaron tres o más entrevistas por vivienda afectó significativamente el promedio general obtenido para el total en la ENDIREH 2011.

La conclusión es que el haber entrevistado a todas las mujeres de la vivienda no afectó la declaración de prevalencia de la violencia de pareja, ni a lo largo de la relación ni en el último año; las diferencias son de poca magnitud entre los distintos grupos, a veces a la baja, a veces a la alta; además, como en el caso de las casadas, el grupo similar al del 2006, por ser el de mayor peso, es el que impuso el promedio general.

Cuadro 1

	2011 = 2006 (%)	2011 ≠ 2006 dos entrevistas (%)	2011 ≠ 2006 tres o más entrevistas (%)
Total	63.7	13.3	23.0
Casadas o unidas	71.7	13.7	14.6
Alguna vez unidas	53.1	24.6	22.3
Nunca unidas	51.7	4.5	43.8

Cuadro 2

Violencia de pareja a lo largo de la relación

	2006	2011 (total de entrevistas)	2011 dividido en tres grupos		
			2011 = 2006	2011 ≠ 2006 dos entrevistas	2011 ≠ 2006 tres o más entrevistas
Total	43.3	47.0	46.6	51.4	45.4
Casadas o unidas	46.7	44.9	44.8	45.5	44.5
Alguna vez unidas	61.6	64.3	63.6	65.1	65.2
Nunca unidas	26.1	37.2	37.8	36.8	36.5

Cuadro 3

Violencia de pareja en los últimos 12 meses

	2006	2011 (total de entrevistas)	2011 dividido en tres grupos		
			Viviendas donde se trabajó igual que en el 2006	Viviendas con dos entrevistas, exceptuando combinaciones iguales al 2006*	Viviendas con tres o más entrevistas
Total	-	27.4	28.4	26.7	24.9
Casadas o unidas	40.2	33.6	33.5	34.9	32.9
Alguna vez unidas	31.8*	11.4	10.4	13.0	12.0
Nunca unidas	-	22.5	23.0	16.6	22.5

* Este porcentaje se refiere a la violencia ejercida después de la separación, la que no necesariamente coincide con la del último año. En el 2006, para las alguna vez unidas y para las solteras, no se captó la violencia del último año.

Resolución I de la XIX Conferencia Internacional de Estadísticos del Trabajo y su impacto en la estadística laboral en México

Rodrigo Negrete Prieto y Tomás Ramírez Reynoso

Mexico Prices Rise as Inflation Remains Above Central Bank Limit/Bloomberg/Getty Images



Nota: los comentarios y puntos de vista de los autores no necesariamente reflejan los de la institución en la que laboran; los autores agradecen a Lilia Guadalupe Luna Ramírez, Mario Moreno Moreno y Efraín Muñoz Valadez su colaboración, tanto en el proceso de investigación como en la integración del artículo.

Este ensayo ilustra los cambios que pueden tener lugar en la estadística del trabajo en México a raíz de las recomendaciones que se desprenden de la XIX Conferencia Internacional de Estadísticos del Trabajo (CIET), mismas que van dirigidas en especial al diseño, producción y difusión de la información proveniente de las encuestas de fuerza laboral (de empleo) en el mundo entero. Para ello, se utilizan los datos que genera la Encuesta Nacional de Ocupación y Empleo (ENOE), pero bajo el nuevo marco y construcción de indicadores que introduce la *Resolución I* de la Conferencia. Al ilustrar hacia dónde llevan estas recomendaciones, se analizan los aspectos en los que se logra claridad y en los que no, así como conceptos en los cuales se requiere proceder con cautela. Se trata, entonces, de someter a evidencia empírica lo que pretende esta resolución y reflexionar, asimismo, lo que ello demandará para la implementación de este tipo de encuestas a futuro.

Palabras clave: Organización Internacional del Trabajo (OIT), Conferencia Internacional de Estadísticos del Trabajo (CIET), Sistema de Cuentas Nacionales, frontera de la producción, trabajo, ocupación, empleo, desocupación abierta (desempleo), subocupación (subempleo), fuerza de trabajo, indicadores de subutilización, fuerza de trabajo potencial, fuerza de trabajo ampliada, población económicamente activa (PEA), Encuesta Nacional de Ocupación y Empleo (ENOE).

Introducción

No debe sorprender que la Organización Internacional del Trabajo (OIT) haya sido fundada como parte de los acuerdos del *Tratado de Versalles* de 1919 que puso fin a la Primera Guerra Mundial. El problema de cómo integrar a las masas a millones de soldados movilizados a la vida ciudadana de nueva cuenta, no era un asunto menor y *la cuestión obrera* volvía a convertirse en una preocupación central a raíz de la revolución bolchevique de 1917 con sus secuelas en la revolución espartaquista en Alemania y la autoproclamada República Soviética Húngara, mismas que, aunque no sobrevivieron

The aim of this article is to show what changes can be made in Mexico's statistics on work and employment, after considering the resolutions of the XIX International Conference of Labour Statisticians (ICLS). These recommendations are mainly related to the design, production and dissemination of the information obtained through the labour force surveys in the world. To achieve this, data generated by the National Survey on Occupation and Employment (ENOE, in spanish) will be used, but this time, under the guidelines of the 1st Resolution of the above-mentioned conference. While showing where these recommendations lead us to, those aspects where clarity is or is not achieved will be analysed, as well as those concepts in which it is very important to act upon very carefully. In a word, this is about putting into empirical evidence what the resolution is aiming at reflecting at the time on what would be required in order to implement this kind of surveys in the future.

Key words: International Labour Organization (ILO), International Conference of Labour Statisticians (ICLS), System of National Accounts (SNA), SNA's Production Boundary, Work, Labour, Occupation, Employment, Unemployment, Underemployment, Labour Force, Labour Underutilization Rates, Potential Labour Force, Extended Labour Force, Economically Active Population, National Survey on Occupation and Employment (ENOE).

como en Rusia, eran un aviso de que la paz de las trincheras no prometía lo mismo en las ciudades. La economía de guerra de pleno empleo había concluido con el silencio de los cañones, pero la sombra del desempleo masivo con una ingobernabilidad aparejada se proyectaba ahora sobre Europa. Era claro que el fin de la *Gran Guerra* no cerraba un capítulo, más bien abría una nueva era de incertidumbre.

Cuantificar, de entrada, los trabajos disponibles y de qué clase son éstos, inventariar los distintos tipos de remuneraciones y cuánta población perceptora había fue una de las primeras preocupaciones

de la OIT, y la tarea no podía iniciar si no había un acuerdo en cómo definir y cómo clasificar. Por ello, en el otoño de 1923, el organismo convocó en Ginebra, Suiza, a la I Conferencia Internacional de Estadísticos del Trabajo (CIET). En 1925, durante la segunda de ellas, se hizo un primer intento por definir el concepto de desempleo de una manera precisa, lo cual obligaba, a su vez, a definir al universo en el que éste incidía, que es la fuerza laboral. A partir de entonces se han celebrado 19 CIET, donde se han construido los marcos conceptuales y definiciones para cuantificar toda una diversidad de asuntos directa o indirectamente relacionados a la problemática laboral, como el trabajo infantil; las horas trabajadas y los periodos de ausencia; los ingresos, los costos y la productividad; las huelgas y los paros; la cobertura de los contratos colectivos y la seguridad social; los índices de precios; el empleo informal; y las clasificaciones de tipos de ocupaciones y sectores de actividad, así como las de accidentes y enfermedades de trabajo (ver anexo 1).

Las CIET se han convertido en el foro de normalización y estandarización estadística a nivel mundial relativa a las estadísticas del trabajo, y la oficina estadística de la OIT, la instancia a cargo tanto de su organización y de la elaboración de los documentos de discusión como del seguimiento de sus recomendaciones, dada su tarea de compilar estadísticas de todos los países miembros del organismo.

El tema de la medición de la población económicamente activa (PEA) —fuerza laboral—, del desempleo y subempleo, aunque apareció de forma temprana como era de esperar, no ha sido finiquitado en una sola conferencia y, en realidad, algunas de las CIET han revisado las recomendaciones de las anteriores, ya sea a la luz de los problemas y dificultades no previstos que surgieron en el primer intento de aplicación de los conceptos acordados y/o por el surgimiento de nuevas técnicas de captación, todo ello acompañado de fenómenos históricos, como el surgimiento del llamado estado benefactor después de la Segunda Guerra Mundial en el mundo industrializado y,

más reciente, la globalización: nuevas realidades que obligan a definir con más rigor cuestiones que parecían simplemente intuitivas.

En 1925, cuando por primera vez se trató de abordar el tema del desempleo, todavía no habían madurado las técnicas de muestreo y, con ello, la posibilidad de aplicar encuestas probabilísticas para inferir la magnitud de personas en tal situación. Fue el Departamento de Comercio y Trabajo (después nombrado *Bureau of Labor Statistics* o BLS) de Estados Unidos de América (EE.UU.), la primera institución en el mundo en incursionar en estas técnicas y metodologías¹ y no fue sino hasta 1982 que la OIT consideró que se había acumulado suficiente experiencia en el tema, aunado al desarrollo y difusión de la tecnología informática, como para proponer un marco conceptual para la medición de la fuerza de trabajo, el empleo y el desempleo, privilegiando como fuente de captación encuestas probabilísticas en los hogares, mismas que tenían que ser diseñadas y aplicadas por instituciones especializadas de estadística.

La XIII Conferencia (1982) fue una de las más influyentes en la historia de las CIET. El hecho de que sus recomendaciones fueran, a su vez, desarrolladas, detalladas y precisadas por el *staff* técnico de la OIT en lo que, en la comunidad de los estadísticos del trabajo se le conoce como *El libro púrpura* (*The Purple Book*),² constituyó un gran impulso para estandarizar los conceptos de las

1 El poder del muestreo aleatorio sobre otras formas de muestreo y los teoremas que soportan esa superioridad fueron desarrollados en la década de los 30 por P. Chandra Mahalanobis, mientras que Jerzy Neyman le dio fundamento a la utilización de intervalos de confianza también en esa década. A su vez, Margaret Martin y Morris Hansen del *Census Bureau* de EE.UU. fueron los primeros en intentar llevar estas técnicas a la práctica para medir cuestiones de ingresos en la población y, ulteriormente, Cal Dedrick y Fred Stephan para medir el desempleo. Sin embargo, estos pioneros no dejaron de toparse con la dificultad de procesar los cuestionarios recopilados y justificar, así, la eficiencia del muestreo en términos operativos, dado que las tecnologías computacionales comenzaron a despegar después de la Segunda Guerra Mundial, en tanto que la difusión de éstas tardaría aún más. Por mientras, se avanza hasta lograr una primera sistematización de las técnicas de muestreo con el muy influyente libro de Hansen y Hurwitz (*Sample Survey Methods and Theory*, 1953). Para una visión panorámica de la historia del desarrollo estadístico durante el siglo pasado, ver: Salsburg, David. *The Lady Tasting Tea: How Statistics Revolutionized Science in the Twenty Century*. 2001.

2 Manual elaborado por Hussmanns, Mehran y Verma (conceptuólogo, experto en procesos de producción estadística y experto muestrista, respectivamente) y cuyo nombre oficial es *Employment, Unemployment and Underemployment: an ILO manual on concepts and methods*.

encuestas en hogares donde ya se aplicaban,³ así como su implementación donde no habían sido introducidas.

Con el transcurso de los años, tanto la XIII CIET como la oficina estadística de la OIT fueron, de algún modo, víctimas de su propio éxito. Al volverse la tasa de desocupación (TD) el indicador de referencia alrededor del cual se centraba el diseño estadístico de las encuestas, tal parecía que la razón de ser de éstas era sólo su cuantificación y que el resto de la información captada sería contextual. Fue así que la TD se volvió, de facto, el indicador focal, pese a los mismos intentos de la OIT a través de las CIET de introducir otros indicadores en los años subsiguientes.⁴ Esa centralidad de la TD se volvió problemática entre los países en desarrollo que más de cerca seguían las recomendaciones. Una de las paradojas que comenzó a detectarse fue que la TD medida de manera estricta podía registrar menores niveles que en países desarrollados; ello comenzó a dejar en evidencia, por lo menos, dos supuestos detrás del concepto: 1) que las naciones tienen un nivel de modernidad homogéneo en el que operan mercados laborales estructurados y 2) que operan instituciones de amplia cobertura como la seguridad social (incluido el seguro de desempleo), de modo que los individuos calibran y ajustan sus expectativas comportándose como buscadores activos y sistemáticos de empleo asalariado para su inserción laboral a partir de estas condiciones. El problema que de ahí deriva es que la TD, en especial en los países en desarrollo, más que ser una medida del déficit de empleos que enfrenta la PEA, se convierte en la medida de un comportamiento específico frente a ese déficit, en el cual es más factible que incurra el segmento

3 En México se comenzó a incursionar en las encuestas en hogares para la medición del empleo y desempleo en 1972 con la Encuesta Nacional de Hogares (ENH), seguida de la Encuesta Continua de Mano de Obra (ECMO) y posteriormente por la Encuesta Continua Sobre Ocupación (ECSO) en 1977, pero fue la Encuesta Nacional de Empleo Urbano (ENEU) en 1983 la que tuvo lugar después de la XIII CIET y, por ende, la primera en ser influida por sus lineamientos, así como la primera a cargo del recién creado Instituto Nacional de Estadística, Geografía e Informática (INEGI), hoy Instituto Nacional de Estadística y Geografía. Antes de la ENEU, las encuestas en hogares, sobre todo en países en desarrollo, seguían el denominado modelo Atlántida, promovido por el *Census Bureau* de EE.UU.

4 Hacia mediados de la década de los 90, la OIT intentó la promoción de los denominados *Key Indicators of the Labour Market (KILM)*, que incluían, además de la TD, la tasa de empleo en el sector informal, así como la de subempleo, entre otros.

moderno de la PEA que, no necesariamente, ha de ser el único en el paisaje o el que predomina.⁵ Esto, incluso, se volvió más evidente cuando en muchas naciones en desarrollo se dio el paso de encuestas de cobertura sólo urbana a otras con inclusión de áreas rurales para tener una representatividad nacional.⁶

Todo lo anterior llevó a que algunos países relajaran su concepto de desempleo para obtener TD de magnitudes *creíbles*. Así, siendo que la definición de desocupación o desempleo abierto tiene tres condiciones: 1) no tener un vínculo laboral, 2) estar disponible para trabajar y 3) estar en situación de búsqueda de trabajo,⁷ no pocos países del África Subsahariana e, incluso, algunos de América Latina, omitieron el tercer criterio como requisito para contabilizar. Al eliminarlo —lo que conceptualmente hablando es una omisión clave porque la acción de búsqueda es lo que actualiza o hace presente la presión ejercida en el mercado laboral—, se obtenían magnitudes de TD mayores, con el inconveniente de que presentaban un comportamiento estacionario en el tiempo, reflejando poco o nada los efectos del ciclo económico, que es una de las mayores fortalezas de la TD rigurosamente construida. Lo anterior, a su vez, introdujo un factor de desorden en la comparabilidad internacional. El INEGI nunca optó por esta ruta fácil de relajamiento del concepto, lo que le ha obligado, de manera

5 Para una explicación a fondo de por qué la TD muestra un comportamiento contraintuitivo en áreas atrasadas, sin que ello necesariamente invalide su utilidad analítica, ver Negrete, Rodrigo. "El indicador de la polémica recurrente: la tasa de desocupación y el mercado laboral en México", en: *Realidad, Datos y Espacio. Revista Internacional de Estadística y Geografía*. Vol. 2., núm 1, 2011.

6 En México se realizó un primer intento de extender la medición del empleo y desempleo a áreas rurales con la Encuesta Nacional de Empleo Rural (1988) a manera de prueba piloto que daría lugar a la Encuesta Nacional de Empleo (ENE), misma que comenzaría a levantarse una vez al año desde 1995 y, trimestralmente, a partir del 2000, mientras que con la ENOE (que entró en operación en el 2005, reemplazando a la ENEU/ENE), la información de cobertura nacional se difunde cada mes como preliminar y trimestral, como definitiva. En esto, México se adelantó a muchos países en desarrollo cuyo paso de encuestas urbanas a nacionales (es decir, incluyendo representatividad rural) se dio aproximadamente entre 10 y 15 años después. Hoy en día, incluso las encuestas de fuerza de trabajo de Brasil (PME) y de Argentina (EPH), siguen siendo urbanas, aunque en parte ello se entiende por la extensión territorial de ambas naciones. La paradoja de TD bajas en áreas rurales alejadas de centros urbanos ha sido algo relativamente nuevo con lo que se han topado varios países.

7 Este encadenamiento de condiciones (considerada como la definición propiamente estadística de desocupación o desempleo abierto) aparece por primera vez formulada en el párrafo 10 de la página 4 de la *Resolución I* de la XIII CIET (Ginebra, Suiza, 18-19 de octubre de 1982) y se abunda en ella en la página 97 de *The Purple Book* de la OIT (*supra*, nota 2).

recurrente, a explicar el sentido de un indicador que, en primera instancia, no resulta intuitivo a medios de comunicación y público en general, a menos que se manejen ciertas nociones de teoría económica.⁸

Pero el problema de que un comportamiento (búsqueda activa de empleo asalariado) sea algo específico y no necesariamente generalizado también puede darse a futuro, por lo que ha sido una preocupación no sólo de las naciones desarrolladas inmersas en un proceso de envejecimiento sino, en general, de todas (desarrolladas o no) en proceso de transición demográfica: la población de adultos maduros y mayores, una vez que se va perdiendo la seguridad en el empleo ante el escenario de competir con población más joven por puestos de trabajo, puede asumir de entrada que no tiene oportunidades y, por ende, no competir, es decir, no presentar el comportamiento de búsqueda activa, pese a tener una necesidad de (re)inserción laboral. En general, cabe pensar en un monto no despreciable de población que se autoderrota frente a las reglas del mercado laboral y más susceptible de incurrir en las trampas de la pobreza. Se trata de segmentos que, simplemente, no son visibilizados por la posición que, de facto, guarda la TD como indicador estelar en el análisis económico, ello en una época y momento (poscrisis global 2008) que reclama ya no supeditar lo social (sea en términos analíticos o programáticos) a lo meramente económico.

La problemática que se va esbozando, entonces, es cómo no diluir en materia de indicadores laborales lo económico en lo social, como tampoco lo social en lo económico; lo primero es incurrir en el desorden, lo segundo, en una miopía progresiva.

En paralelo, un aspecto no menos relevante —a propósito de visibilización en los últimos 10 ó 15 años— ha sido la fuerte influencia del enfoque

⁸ Ver Negrete, nota 5. El llamar a la TD *el indicador de la polémica recurrente* no es gratuito; el esfuerzo de comunicación que ha hecho el INEGI al respecto a lo largo de los años —décadas ya— ha sido considerable. Muchas oficinas de estadística en el mundo ni siquiera lo han intentado y se limitan a señalar que siguen las definiciones internacionales, como si el mensaje implícito fuera “sí no os parece quejaos con la OIT”.

de género en las estadísticas y en los mecanismos implícitos de clasificación que a ellas subyacen. La vindicación de las tareas domésticas no remuneradas, por ejemplo, ha obligado a reconsiderar el concepto mismo de trabajo; si éste debe quedar acotado o no a una relación laboral o a un contexto de transacciones de mercado. Ello ha vuelto necesario reconsiderar el uso corriente de un término con algo que dé cuenta que el proceso de reproducción social tiene momentos de mercado y no de mercado; que se necesita un mapa más amplio para entender un término como *trabajo*. La OIT se percata que en 90 años de conferencias internacionales nunca hubo la necesidad de definir un término que se daba por sentado, pero que ahora necesita explicitarse, uno que hay que llevar más allá de una acepción común, no inmune en modo alguno a cierto sesgo en el entender *qué es trabajar*. Si *trabajo* es algo que va más allá de cierto territorio que antes se creía era el todo, ahora se necesita una nueva cartografía del todo y las partes: los territorios que no se miraban, al ahora mirarse cambian la idea misma que se tenía de un mundo.

Por todo lo anterior, es que 30 años después de celebrada la XIII CIET, el Consejo de Administración de la OIT autorizó la convocatoria del 2 al 11 de octubre de 2013 de la XIX CIET con un doble mandato:⁹

- Revisión de las normas internacionales vigentes en materia de estadísticas de la PEA, el empleo, el desempleo y el subempleo, incluidas las medidas de subutilización de la fuerza de trabajo como complemento de la tasa de desempleo.
- Examinar las actividades estadísticas pasadas y previstas desde su última reunión en noviembre del 2008.

En el presente ensayo nos centraremos en su *Resolución 1* (relacionada con el primer mandato), dado que las restantes se refieren ya sea a la

⁹ Ver OIT. *XIX Conferencia Internacional de Estadísticos del Trabajo. Proyecto de informe*. Página 1, donde se cita la resolución de la 313.ª Reunión de Consejo en Ginebra, Suiza, en marzo del 2012.

conformación o avances de grupos de trabajo sobre diversos temas.¹⁰ Dicho en pocas palabras, la *Resolución I* se centra en revisar lo que estuvo vigente desde 1982, de donde se desprende por su propio peso que la XIX CIET es la más importante de los últimos 30 años.

No está de más mencionar que, aunque en marzo del 2012 el Consejo de la OIT instruye a que se convoque de manera formal la Conferencia a celebrarse al siguiente año, la oficina estadística del organismo había formado, desde diciembre del 2009, con integrantes de organismos y oficinas estadísticas invitadas (entre ellas el INEGI), un grupo de discusión sobre las reformas a implementar en el marco conceptual, a la par que organizó reuniones regionales preparatorias con invitación abierta a todas las oficinas estadísticas de los países de la región en cuestión. Asimismo, a finales de enero del 2013, también en Ginebra, el *staff* estadístico (secretariado técnico) de la OIT y el grupo de discusión sostuvieron reuniones con agrupaciones de trabajadores y empleadores para explicar el contenido de los borradores de los documentos que serían presentados ese otoño en la Conferencia. Todo esto se menciona porque la oficina estadística de la OIT no había procedido con tantas etapas preparatorias en CIET previas, lo que habla de la relevancia para el organismo de la XIX Conferencia Internacional de Estadísticos del Trabajo¹¹ que se realizó del 2 al 11 de octubre de 2013 en Ginebra, Suiza, con la asistencia de 105 países,¹² 11 organismos internacionales, representantes de organizaciones

patronales y sindicales de seis naciones, cinco organizaciones no gubernamentales más un frente de liberación (Palestina).¹³

El objetivo enunciado de la *Resolución I* es en torno al cual se estructura este ensayo.¹⁴ Primero se expondrá a qué conduce la definición de trabajo que introdujo la Conferencia y a explicar su sentido; se hablará, asimismo, de los cambios de terminología y lo que conlleva; a continuación, la exposición se centrará en los nuevos indicadores de subutilización de fuerza de trabajo que propone la OIT a través de la CIET y, por último, se hará una reflexión acerca de lo que (a nuestro juicio) consigue esta resolución, pero también sobre incertidumbres y puntos que la oficina estadística de la OIT, al editar nuevos manuales para las oficinas estadísticas, tendría que examinar más a fondo. Todo esto se ilustrará con información proveniente de la Encuesta Nacional de Ocupación y Empleo (ENOE) que es la encuesta de fuerza laboral de México. Cuando no se trate de una serie, es decir, cuando sea un dato fijo —y salvo indicación de lo contrario—, el momento en el tiempo al que queda referido será el segundo trimestre del 2014.

Concepto de trabajo y el nuevo marco para identificar sus distintas modalidades

Como ya se mencionó, la XIX CIET (por primera vez en la historia de las conferencias) se vio en la necesidad de definir el concepto *trabajo*. Antes

10 *Resolución II*, acerca de la labor de estadísticas sobre trabajo forzoso; *Resolución III*, de la labor futura en materia de estadísticas sobre las cooperativas; *Resolución IV*, acerca de la labor futura en materia estadística sobre migración laboral; *Resolución V*, del funcionamiento de la Conferencia Internacional de Estadísticos del Trabajo y la actualización de su Reglamento.

11 No menos relevante en términos contextuales es que el secretariado técnico de la OIT, a cargo de la Conferencia y de los textos de discusión, correspondió a una nueva generación dentro del organismo, pues miembros destacados del anterior *staff* estadístico de la OIT, que tuvieron una labor fundamental en la elaboración tanto de *The Purple Book* como de los manuales de medición de la informalidad, pasaron al retiro.

12 La delegación mexicana estuvo formada por funcionarios de la Dirección General de Investigación de Estadísticas del Trabajo de la Secretaría del Trabajo y Previsión Social (Omar Rodríguez Alarcón y Eloísa Pulido Jaramillo) y del INEGI (Rodrigo Negrete Prieto), quienes garantizaron la presencia del país en las tres instancias de discusión de la XIX CIET: 1) reuniones de comité para comentar el proyecto de resolución; 2) en el comité de redacción (*Draft Committee*) que se reunía al final de la jornada diaria y 3) plenarios en las que se resumían los comentarios al proyecto de resolución, así como la aprobación de otros documentos o iniciativas presentadas por la OIT.

13 La ausencia más destacada en la XIX CIET fue la de EE.UU., cuya asistencia se vio afectada por el llamado *Shut Down* presupuestal, al no alcanzarse al respecto en ese otoño un acuerdo entre el Congreso y el Poder Ejecutivo estadounidense. Sin embargo, la representante del BLS de este país (Anne E. Polivka) formó parte del grupo de discusión de los documentos que a la postre fueron presentados en la XIX CIET.

14 A continuación, se cita el punto número 1 del apartado *Objetivos y ámbito de la aplicación*:

“La presente resolución tiene por objeto establecer normas sobre estadísticas del trabajo que sirvan para orientar a los países en la actualización e integración de los programas estadísticos que existen en este campo. Define el concepto estadístico de trabajo a título de referencia e incluye conceptos, definiciones y directrices de carácter operativo para: a) Subconjuntos distintos de actividades productivas, denominadas formas de trabajo; b) Clasificaciones de la población según su situación en la fuerza de trabajo y su forma de trabajo principal y; c) Medidas de subutilización de la fuerza de trabajo”.

OIT. XIX CIET. *Resolución I: sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo*. Ginebra, p. 2.

de continuar, se hace inevitable citar el párrafo 6 de la resolución *in extenso*:

“El trabajo comprende todas las actividades realizadas por personas de cualquier sexo y edad con el fin de producir bienes o prestar servicios para el consumo de terceros o para uso final propio.

- “a) El trabajo se define independientemente de la legalidad de la actividad y de su carácter formal o informal.
- “b) El trabajo excluye las actividades que no entrañan la producción de bienes o servicios (por ejemplo la mendicidad y el robo), las actividades de cuidado personal (por ejemplo la higiene y el aseo personales) y las actividades que no pueden ser realizadas por terceros para el beneficio de una persona (por ejemplo dormir, aprender y las actividades para el entretenimiento propio).
- “c) El concepto de trabajo está en conformidad con la frontera **general** [lo marcado con negritas es nuestro] de la producción tal como se define en el Sistema de Cuentas Nacionales 2008 (SCN 2008) y su concepto de unidad económica que distingue entre:
 - “i. Unidades de mercado (es decir, sociedades, cuasi-sociedades y empresas de mercado no constituidas como sociedades);
 - “ii. Unidades no de mercado (es decir, administración pública e instituciones sin fines de lucro) y;
 - “iii. Hogares que producen bienes o servicios para uso final propio.
- “d) El trabajo puede realizarse en cualquier tipo de unidad económica...”.¹⁵

Lo primero que se puede observar aquí es el carácter descriptivo y no normativo del concepto (inciso a), dado que trabajo es todo lo que abona a la producción de bienes y servicios, cualesquiera que sean éstos y la manera como se generen. Lo segundo que cabe subrayar es que, si bien estamos ante

¹⁵ *Ibid.*, p. 2, párrafo 6.

una definición (de forma deliberada) muy amplia de trabajo, no toda actividad lo es. El denominado principio de tercera persona establece esa diferencia: si otro lo puede hacer por mí, es trabajo, pero si la acción es intransferible, no lo es. Se puede ir, incluso, más allá de los ejemplos de lo que se excluye en el inciso b: la participación en un acto cívico, religioso o político (votar, asistir a una asamblea, una marcha, una manifestación, una ceremonia o procesión, etc.) son actividades que, si los hace otro en lugar del interesado, pierden todo sentido y propósito para este último: se disuelve su significado al momento mismo de pensar que un tercero se haga cargo en nombre mío. Lo mismo aplica cuando uno se dedica a los estudios. En contraste, todo lo que es sustituible por un tercero, aunque no sea una actividad de mercado, en principio da lugar a algo que potencialmente deriva en esa dirección; el trabajo doméstico no remunerado o servicios que se generan para el propio hogar lo ejemplifican: siempre podrán realizarse fuera del mercado pero, al mismo tiempo, es concebible que los realice un tercero y, por ello, existirá un mercado de servicio doméstico a la espera de que alguien acuda a solicitarlo. Una cosa es que la XIX CIET no confunda el trabajo con el mercado de trabajo, pero ello no quiere decir que sean cosas desvinculadas: el principio de tercera persona establece el eslabón entre ambos, la conexión que permite llamarlos a ambos *trabajo*. El principio de la tercera persona opera porque el bien o servicio de por medio no es difuso o indeterminado; si, en cambio, yo asisto a una manifestación, se podrá argüir que contribuyo al fortalecimiento de la ciudadanía, pero ello no es lo suficientemente preciso como para que se le considere un servicio cuantificable. No hay un mercado de servicios a la vida cívica participativa. Lo que no es trabajo no tiene un equivalente al que se le pueda establecer un precio de mercado,¹⁶ no proyecta tal sombra.

Muchos actos de mendicidad se disfrazan de servicios, pero guardan el común denominador

¹⁶ Los quehaceres domésticos, así como los cuidados de niños, enfermos o ancianos, sí tienen un equivalente de mercado, es justo la razón por la cual se pueden calcular las denominadas cuentas satélite de estas actividades al imputárseles en alguna medida un valor, aunque dichas actividades hayan sido realizadas sin transacciones de mercado de por medio.

que su valor de mercado no se les puede precisar porque, de entrada, nunca existió demanda para ellos; aunque detonen un flujo de ingresos, dicho flujo no encaja en el concepto de transacción, sino de transferencia unilateral de recursos de un hogar a otro, como ocurre en un robo o asalto. Hablando en términos jurídicos, serán actos muy distintos,¹⁷ pero no puede decirse de un servicio de por medio, de ahí que no sean trabajo y, por ende, no toda percepción de ingresos lo supone. Ésta es la precisión implicada en el inciso b: así como no toda percepción de ingresos supone trabajo, no toda actividad lo es. Hay trabajos sin ingresos e ingresos sin trabajo, pero un trabajo sin ingresos o remuneraciones siempre tendrá un equivalente que los tenga. Eso es lo que implica hablar de una **frontera general** del Sistema de Cuentas Nacionales (SCN), que es el horizonte de todo trabajo concebible y que no hay que confundir con el referente más específico de **frontera de producción** del SCN contenida en la primera (más adelante se abundará sobre esta distinción).

17 Cabe observar, sin embargo, que muchas formas de mendicidad y sobre todo la disfrazada (servicios no solicitados o seudoservicios) tienen un ingrediente de extorsión.

Los incisos c y d establecen todos los ámbitos posibles en los que tiene lugar el trabajo: las unidades económicas pueden ser empresas constituidas o no, instituciones y hogares. Así, el concepto de trabajo es preciso (no todo es trabajo) y general al mismo tiempo. A partir de estos elementos se puede armar el mapa conceptual de las distintas modalidades de trabajo.

El primer objetivo que se persigue es dejar de pensar que todo trabajo es algo que ocurre en la **frontera de producción** del SCN, pues hasta antes de la XIX CIET, su noción era por entero dependiente de esta circunscripción.

En el cuadro 1 se observa, de entrada, la relación entre la frontera de producción del SCN y la general, esta última (en color más claro) desborda a la primera. La frontera de producción contiene todo lo referente a bienes (de mercado o no), así como de servicios de mercado, pero lo que ya no contiene es la prestación de servicios que generen los hogares, sea para ellos mismos (quehaceres domésticos y cuidados a sus enfermos o vulnerables) o para otros a los que apoyan (por ejemplo,

Cuadro 1

Marco conceptual para las estadísticas del trabajo

Destino previsto de la producción	Para uso propio final (autoconsumo)		Para su consumo por terceros				
	Forma de trabajo	Trabajo de producción para el autoconsumo		Trabajo en una ocupación (trabajo por remuneración o beneficio)	Trabajo en formación no remunerado	Otras formas de trabajo*	Trabajo voluntario
Servicios		Bienes	En unidades de mercado y no de mercado				En hogares productores
Relación con el SCN 2008			Dentro de la frontera de producción del SCN			Bienes	Servicios
	Dentro de la frontera general del SCN						

* Trabajo obligatorio no remunerado para terceros.

los servicios de cuidados que presta el hogar de los padres a los hogares de los hijos cuando se les encarga el cuidado de los nietos). Detrás de esta provisión de servicios (autogenerados o de hogar a hogar) hay un trabajo, pero ya no uno contenido en la frontera de la producción. No hay que olvidar que estos servicios los podría generar un tercero que cobrara por ello; si así lo hiciera, ese servicio quedaría circunscrito a la frontera de producción y abonaría al producto interno bruto (PIB), habiendo una verdadera transacción de por medio, no importando que el servicio fuera prestado de manera formal o informal. Lo que rebasa dicha frontera si bien no abona al PIB por no mediar una transacción ya sea con el trabajo mismo o con sus productos, no por ello deja de ser trabajo, cuya equivalencia en términos de mercado puede contabilizarse en una cuenta satélite.

El esquema, entonces, permite separar el trabajo que es contabilizado en el PIB del que no lo es, y obsérvese que la condición de formal o informal, legal o ilegal, no es lo que determina la contabilización en el PIB, sino el hecho de que no tienen lugar verdaderas transacciones que, sin embargo, tienen un equivalente de mercado.

Habiendo claridad entre lo que corresponde a cada frontera, se establecen cinco formas de trabajo mutuamente excluyentes;¹⁸ algunas todo lo que generan queda circunscrito a la frontera de producción (y, por ende, toda su actividad reflejada en el PIB), mientras que hay otras formas de trabajo que, dependiendo de lo que generen, pueden estar dentro o fuera de la frontera de producción, pues si dan lugar a servicios para el hogar u otros hogares sin mediación de transacciones, rebasan la frontera de producción y caen en la frontera general. La XIX CIET pone mucho énfasis en que lo que distingue a estas cinco modalidades es el destino de la producción y el tipo de transacción u objetivo de la misma (percibir o no ingresos monetarios). Se tiene entonces:

- 1) El trabajo de producción para autoconsumo. Su producción de bienes es cuantificable dentro

¹⁸ OIT. XIX CIET. Resolución... , op. cit., p. 3, párrafo 7.

de la frontera de producción, pero la de servicios (quehaceres domésticos, cuidados) cae fuera de ésta y sólo queda contenida dentro de la frontera general. En lo que a la generación de bienes se refiere, si bien éstos no pasan por una transacción, siempre queda abierta la posibilidad de un excedente sujeto a transacción de mercado, cosa que no ocurre con los servicios para el propio hogar que se agotan o consumen en sí mismos al momento que se generan.

- 2) Trabajo en una ocupación donde toda la producción de bienes y servicios supone tarde o temprano una transacción de mercado que da lugar a un ingreso monetario.
- 3) El trabajo en formación no remunerado donde los bienes y servicios generados son objeto de transacciones.¹⁹
- 4) El trabajo voluntario que sí genera bienes cae dentro de la frontera de la producción, pero si genera servicios al igual que el inciso 1, la rebasa.
- 5) Otras formas de trabajo, como el forzado, que puede ser impuesto por el Estado o por particulares. Aquí, si bien no hay transacción al momento de hacer uso del recurso laboral —no hay un *quid pro quo*—, sí puede haber transacciones con los productos o servicios en los que se utilizan tales recursos laborales.

De este modo se logra un objetivo de especificidad, es decir, romper —aún dentro de la frontera de producción— la equivalencia de que todo trabajo que abona al PIB (es decir, que tiene lugar dentro de la frontera de producción) forma parte del mercado laboral. Una transacción de mercado laboral es muy distinta de otro tipo de transacciones o acuerdos. En el mercado laboral siempre hay un *quid pro quo* entre sujetos libres en el que el

¹⁹ Podría pensarse que en estos casos no se cumpliría el principio de tercera persona en la naturaleza de la transacción (al menos desde el punto de vista del aprendiz), pero éste, de hecho, se encuentra inserto en un tipo de relación laboral, así como en un proceso de generación de bienes y servicios bien definidos y especificables, y en la medida en que trabajan para alguien, ese alguien puede sustituirlos: de un modo u otro producen servicios laborales reemplazables y esos servicios los podría proveer alguien más a cambio de una paga o transacción de mercado. Por lo demás, los insumos que proporcionan —horas de trabajo— forman parte de la productividad asociada al PIB y, por ello, quedan circunscritos a la frontera de producción del SCN.

oferente percibe un ingreso monetario, pero antes de la XIX CIET simplemente se daba por hecho que si la actividad abonaba al PIB, en automático es asunto del mercado laboral.

La ocupación (inciso 2) se refiere entonces al trabajo dependiente (como lo es el asalariado) o al independiente (cuentas propias, empleadores) que, o se orienta a una transacción de mercado o simplemente pasa por ella. Cabe señalar que, aun cuando un asalariado trabaja para una unidad económica no orientada al mercado (como lo puede ser una institución gubernamental o una privada sin fines de lucro), desde el momento que intercambia su tiempo por un pago, tanto él como su unidad empleadora participan de una transacción de mercado.

Esta distinción tiene como consecuencia más relevante en términos cuantitativos el separar la agricultura de subsistencia (que ahora cae en la categoría 1) de la 2 u ocupación por remuneración o beneficio (es decir, de la que participa de un modo u otro del mercado laboral). Antes, ambas quedaban unificadas dentro de la denominada PEA, pero ya no más. La racionalidad o el objetivo de esta separación se explicará más adelante a propósito de la construcción de indicadores, lo mismo que las consecuencias, no del todo previstas, que ello puede generar.

Antes de seguir adelante, es importante subrayar que si bien tanto los que encabezan la producción de autoconsumo como sus trabajadores familiares auxiliares salen sumariamente de la fuerza de trabajo (que es la PEA menos este grupo), no es así en el caso de los trabajadores familiares auxiliares de un negocio familiar o informal.²⁰ El razonamiento es que aún, y pese a que no perciben de forma directa una remuneración monetaria, el sentido de su actividad es que su hogar lo perciba porque claramente la unidad económica está volcada a transacciones de mercado.

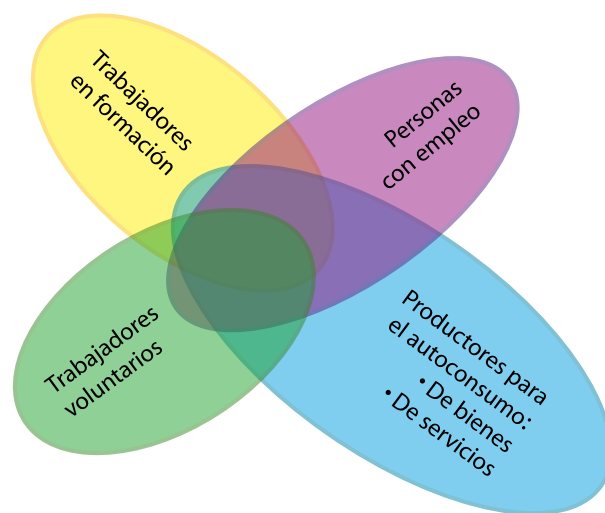
20 OIT. XIX CIET. Resolución..., op. cit., para ver lo que incluye el concepto de ocupación o trabajo por remuneración o beneficio ver p. 8, párrafo 30.

Se han establecido, así, las distintas modalidades de trabajo que, aunque conceptualmente excluyentes, un individuo las puede combinar y éstas pueden ser múltiples, como se muestra en la figura 1. Un caso frecuente sería el de una persona con empleo que también realiza quehaceres del hogar y que, además, se da tiempo para ayudar prestando algún servicio a otros hogares. Esta combinación permite ilustrar algunas de las preocupaciones del enfoque de género sobre las asimetrías de demanda de tiempo en una sociedad.

Desde el punto de vista estadístico debieran, en principio, identificarse todos los bloques de población en edad de trabajar,²¹ según las modalidades

21 Ha de observarse que, si bien en su definición, *trabajo* no queda circunscrito a edad alguna, la aplicación de este marco para la generación continua de estadísticas está acotado a la población en edad de trabajar. La XIX CIET no define una edad mínima, pero sí establece que, en caso de que el país en cuestión ya tenga un programa de recolección de estadísticas de trabajo infantil, dicho límite "...deberá tener en cuenta la edad mínima para acceder al mercado laboral, con las excepciones previstas en la legislación nacional, o la edad en que cesa la enseñanza obligatoria...". *Ibid.*, p. 15, párrafo 65. En ese mismo párrafo se señala, por cierto, que "No debería establecerse un límite de edad máxima...". En el caso de México, al contar la ENOE con un módulo bianual enfocado al trabajo infantil, la interpretación de la norma fijaría este marco para los 15 años de edad en adelante, de acuerdo con la última reforma de la *Ley Federal del Trabajo*. Cabe observar, asimismo, que de 15 años en adelante es la norma de armonización que la OCDE aplica a las estadísticas del trabajo de los países miembros.

Figura 1
Conceptos distintos, pero actividades no mutuamente excluyentes



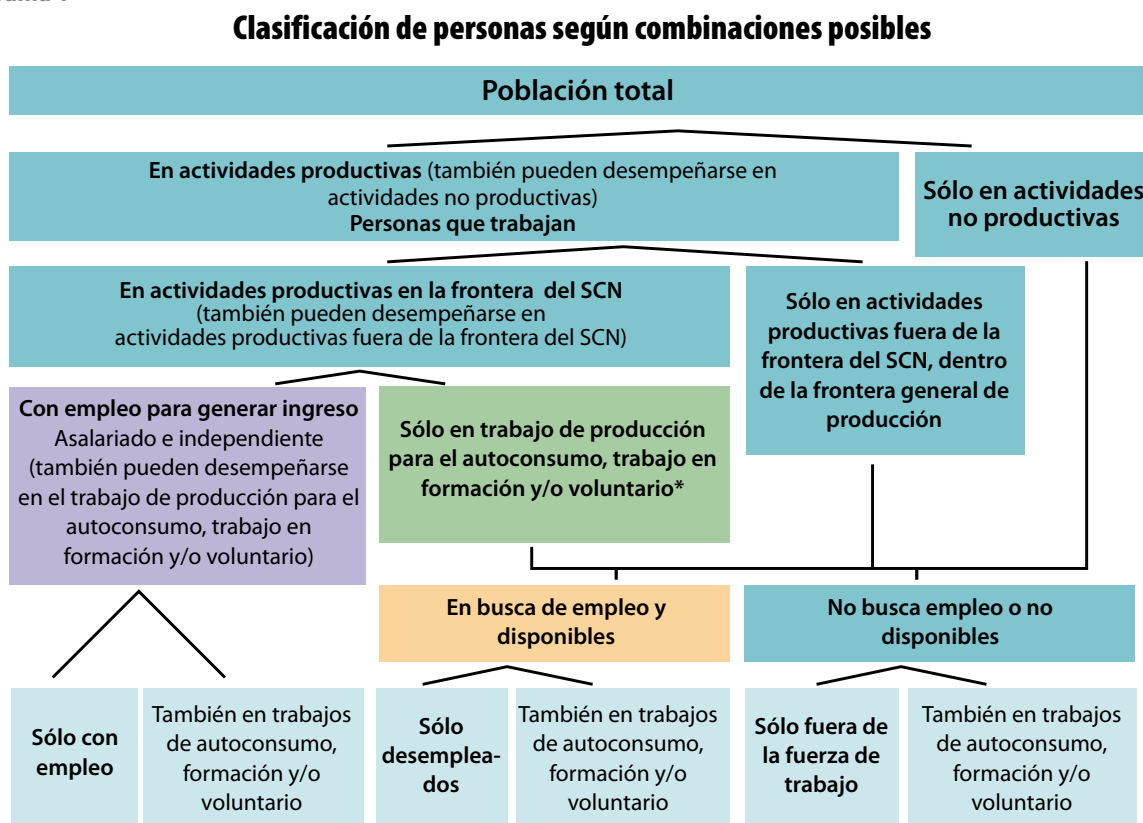
que combinan y su intención de insertarse en el mercado laboral propiamente dicho (ver diagrama 1). Al hacerlo se deberían, asimismo, identificar con plena claridad la masa de horas aportadas en el ámbito del mercado de trabajo, en el de la frontera de la producción y en el de la general del SCN, sucesivamente.

Lo anterior implica que las actuales encuestas de fuerza laboral deberán derivar en un subsistema de encuestas del trabajo con módulos para captar la población involucrada en las distintas modalidades, aunque no necesariamente con la misma periodicidad. Si bien la identificación de los segmentos de población que se dedican en exclusivo a modalidades de trabajo distintas de la ocupación o empleo pudiera ser problemática para encuestas probabilísticas, no sería imposible detectar la combinación de modalidades o roles para el grueso de la población.

De hecho, la ENOE ya capta algo de esto, dado que al final del *Cuestionario de ocupación y empleo* hay una sección de preguntas sobre la realización de actividades distintas a las que tienen lugar dentro de la frontera de la producción del SCN, como: tiempo dedicado a cuidados del hogar, realizar compras y trámites, llevar o trasladar a algún miembro del hogar, quehaceres domésticos o prestar servicios gratuitos a la comunidad.

Al adoptar este marco y reorganizar a partir de él lo que ya capta la ENOE, lo primero que resalta es que del total de la población que genera bienes o servicios (es decir, que trabaja) más de 40% lo hace sólo bajo modalidades que quedan fuera de la frontera de producción. La proporción correspondiente a mujeres asciende a casi 58%, esto es, más de la mitad de las que generan bienes y/o servicios no serían visibilizadas de otro modo salvo bajo este esquema (ver gráfica 1).

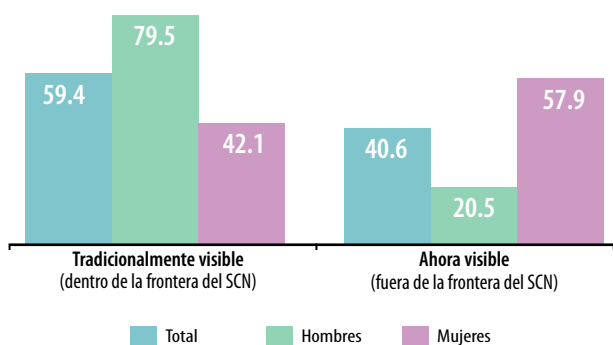
Diagrama 1



* Nuevo tratamiento basado en la propuesta de revisión del concepto de empleo.

Gráfica 1

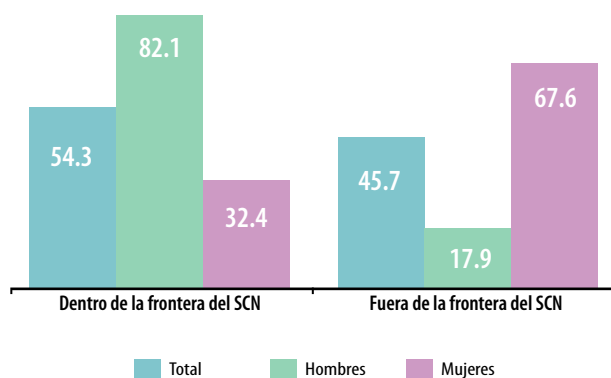
Población en edad de trabajar: porcentaje de población que contribuye al proceso de reproducción social dentro y fuera de la frontera del SCN



En términos de masa de horas trabajadas bajo todas las modalidades, 45.7% son suministradas fuera de la frontera de producción del SCN, pero en el caso de las mujeres se trata de más de dos terceras partes del volumen total del tiempo aportado de trabajo (ver gráfica 2).

Gráfica 2

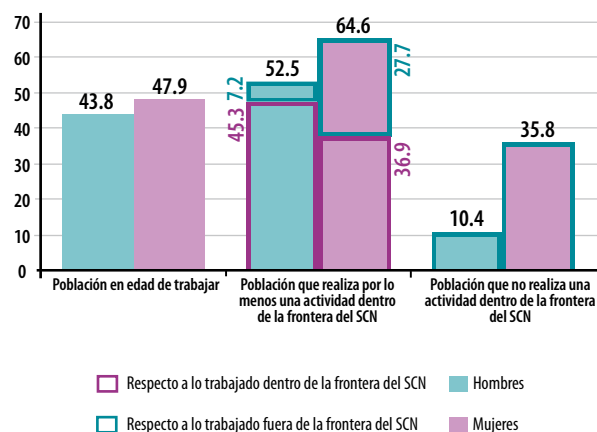
Masa de horas trabajadas: porcentaje dentro y fuera de la frontera de producción del SCN



Visto por horas per cápita para la población de 15 años o más, el acumulado da menos de 44 horas para la población masculina y casi 48 para la femenina; sin embargo, si nos enfocamos en la población que participa de la ocupación de mercado, como resultado de la combinación de modalida-

Gráfica 3

Promedio de horas de trabajo per cápita



des de trabajo se tiene que mientras la población masculina registra 52.5 horas a la semana, la femenina, 64.6, casi una jornada y media de diferencia por semana (ver gráfica 3).²²

Ahora bien, los cambios comentados que introduce el marco conceptual también obligan a modificar la terminología en la identificación de las poblaciones tradicionales en las que se subdivide la población en edad de trabajar. Para subrayar que al colocar en otros casilleros tanto la población dedicada a la producción de bienes de autoconsumo (predominantemente agricultura de subsistencia) como a los aprendices no remunerados y similares (trabajo en formación), se obtiene un contenido más decantado de población involucrada en transacciones de mercado, el término población económicamente activa cambia entonces a fuerza de trabajo (o fuerza laboral); por su parte, se renuncia al término población no económicamente activa (PNEA) para ser ahora población fuera de la fuerza de trabajo, dado que, al menos una parte de ella, al realizar otras modalidades de trabajo fuera del ámbito de las transacciones de mercado

²² Seguramente, las encuestas de uso del tiempo podrían registrar niveles más elevados, sin embargo, otras como la ENOE, lo pueden hacer de manera más ágil o frecuente y con una captación más clara de cómo se da la inserción laboral entre la población que combina actividades. Por su parte, los clasificadores de las encuestas de uso del tiempo (sea bajo la vertiente ICATUS o la CAUTAL) tendrán que revisar su esquema a la luz de la XIX CIET para que se puedan hacer con nitidez las distinciones conceptuales que ésta demanda.

(como los cuidados y quehaceres del hogar), también forman parte de la economía (ver diagrama 2). El cambio terminológico no deja de ser afortunado si se tiene presente que, después de todo, el término *economía* deriva del griego *oikos*, que significa casa.

A su vez, una porción de lo que ahora es la población fuera de la fuerza de trabajo o subconjunto conocido como población disponible, se integra a una nueva categoría ahora denominada fuerza de trabajo potencial (no ocupados disponibles para trabajar pero sin búsqueda + no ocupados buscadores de trabajo pero sin disponibilidad).²³ Una tercera subpoblación surge de la combinación de poblaciones dentro y fuera de la fuerza de trabajo

²³ El término *disponible* es la manera corta de decir *disponibilidad para comenzar un trabajo*, esto es, que la persona podría dedicarse a ello en breve. El caso de individuos que buscan un trabajo, pero que no estarían disponibles para tomarlo, es raro en nuestros países y se acuñó pensando que, en algunas naciones desarrolladas, es extendida la práctica entre jóvenes de inscribirse en una bolsa de trabajo durante el periodo escolar (búsqueda) no estando, sin embargo, disponibles para comenzar sino hasta las vacaciones de verano. Por su parte, la población con alguna necesidad de laborar y disponible para hacerlo, pero que por alguna razón no está en situación de búsqueda, en nuestros países sería 99% de la ahora denominada fuerza de trabajo potencial con una presencia importante de grupos vulnerables. El cómo interpretar en términos operativos el concepto de disponibilidad se discutirá en el tramo final de este ensayo ya que, a nuestro juicio, es uno de los puntos débiles de la XIX CIET.

que quieren integrarse al ámbito de las transacciones de mercado (ocupación) con lo que se forma un nuevo concepto horizonte o abarcador: fuerza de trabajo ampliada. Cada una de estas poblaciones jugará un papel relevante en la construcción de indicadores que se comentarán en las siguientes secciones. El cuadro 2 resume un antes y un después en cuanto a la manera de denominarlas.

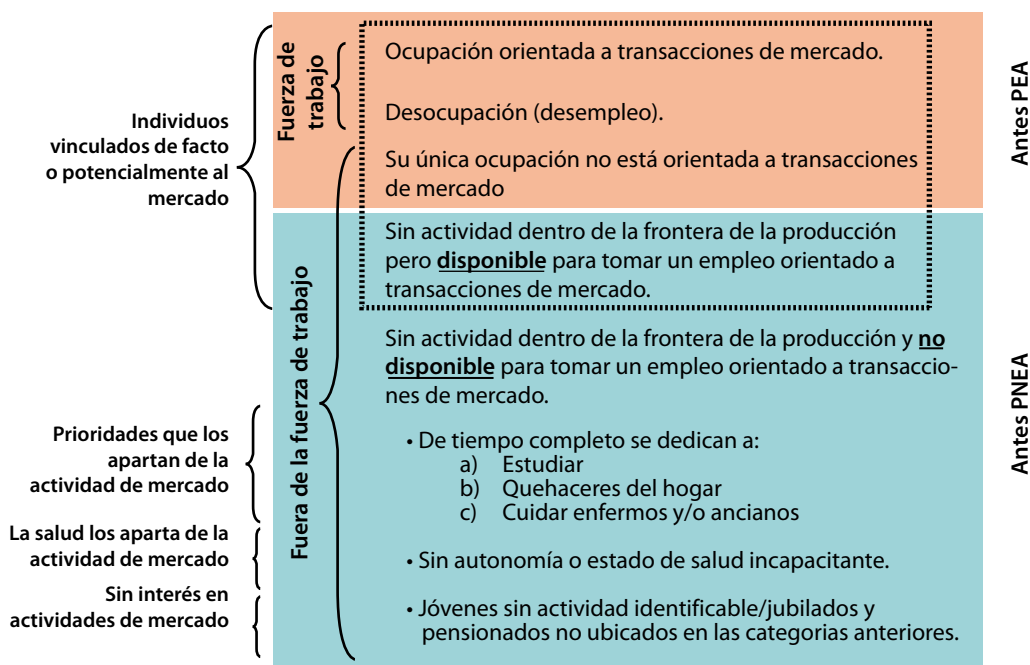
Cuadro 2

Nuevos términos

Antes	Después
Población económicamente activa (PEA)	Población en la fuerza de trabajo (PFT)
Población no económicamente activa (PNEA)	Población fuera de la fuerza de trabajo
Población disponible	Fuerza de trabajo potencial (FTP)
-	Fuerza de trabajo ampliada (PFT + FTP)
Población en edad de trabajar (PET)	Población en edad de trabajar (PET)

Diagrama 2

Nueva delimitación de las poblaciones

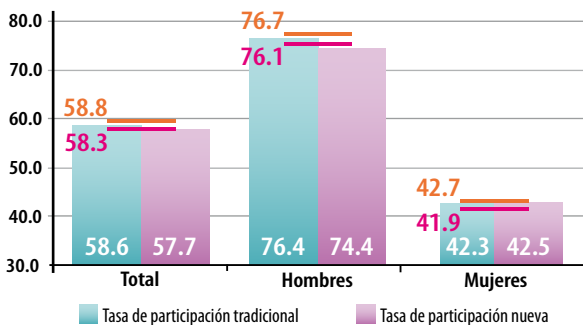


Tasa de participación

Como ya se dijo en la sección anterior, los cambios conceptuales traen aparejados ajustes en las magnitudes respectivas de lo que hasta antes de la XIX CIET quedaba denominada como PEA y ahora se acota a fuerza de trabajo. El primer indicador al que esto afecta es al que se conoce como tasas de participación (porcentaje de la PEA en la población en edad de trabajar). ¿De qué magnitud estaríamos hablando en el caso de México? La gráfica 4 muestra que el efecto se aproximaría a un punto porcentual de disminución (números en la base de las barras), lo que, es probable, sería estadísticamente significativo si se toman en cuenta los intervalos de la estimación (números en la parte superior de las barras); la disminución en la población masculina sería mayor (dos puntos porcentuales) si se considera que en la agricultura de subsistencia —misma que se deduce o resta de la PEA— hay una presencia mayoritaria de este sexo. Por ello, en el caso de las mujeres, no se aprecia una reducción (incluso, hay un ligero incremento) que se explica porque la barra de la izquierda está calculada contra una población en edad de trabajar (PET) de 14 años y más, mientras que la de la derecha respecto a la de 15 y más que será el nuevo estándar para México (ver nota al pie 21).

Gráfica 4

Tasas de participación tradicional de 14 años y más y nueva tasa de participación

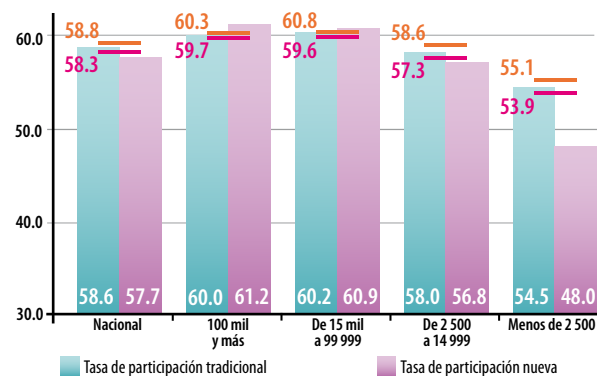


Al mirar por rango de localidad (la ENOE tiene cuatro dominios de representatividad al respecto), lo que se aprecia de entrada en la gráfica 5 es la gran diferencia que resulta para el ámbito de localidades dispersas de menos de 2 500 habitantes, con una reducción de casi 6.5 puntos porcentuales

(valores en la base de las barras); la otra disminución es para las localidades de menos de 15 mil habitantes (superando el punto porcentual). En las poblaciones más grandes, en lugar de disminución se da un incremento por el cambio a 15 años como la edad mínima considerada de la PET, cosa que se corrobora al observar la gráfica 5 bis en la que cada par de barras está calculada como porcentaje de la población de 15 años y más. Controlando ahora por el factor edad, se aprecia una disminución con el nuevo cálculo en los dominios mayores a 15 mil habitantes, menos por la sustracción de la población enfocada a la producción de bienes para el autoconsumo que por la de los trabajadores en formación no remunerados. Para las áreas más rurales y dispersas, la diferencia es, incluso, mayor cuando el referente es la misma edad.

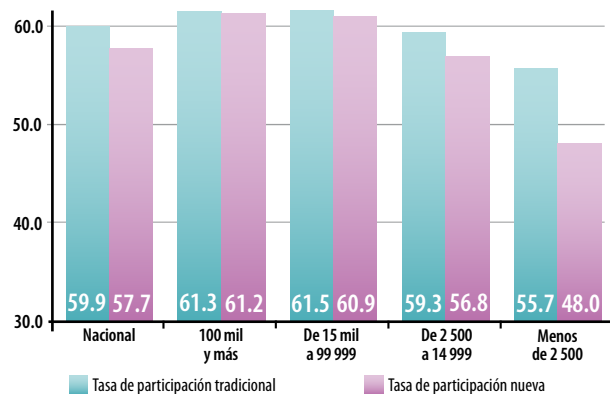
Gráfica 5

Tasas de participación tradicional de 14 años y más y nueva tasa de participación, por tamaño de localidad



Gráfica 5 bis

Tasas de participación tradicional de 15 años y más y nueva tasa de participación, por tamaño de localidad



Indicadores de subutilización

Una vez establecido el marco conceptual del trabajo y sus modalidades, el otro objetivo fundamental de la *Resolución I* es impulsar la construcción de indicadores de subutilización (tasas SU) de poblaciones, de modo tal que la tasa de desocupación abierta o su sucedáneo (SU1) sea una entre otras y no *La Tasa* —con mayúsculas— frente a la cual cualquier otro indicador es un mero complemento.

Como ya se vio, el haber ampliado el concepto de trabajo obliga, al mismo tiempo, a definir con mayor precisión qué ámbito del mismo corresponde a las transacciones de mercado y qué otras modalidades a transacciones distintas o a ninguna (estas últimas en la frontera general, pero fuera de la de producción). La clave en los indicadores SU será tratar de captar a toda la población en edad de trabajar, que procure insertarse en la modalidad específica de transacciones de mercado. *Procurar* significa aquí hacerlo por dos vías; por la de búsqueda activa de trabajo o por la de disponibilidad para trabajar y, quien tiene tal intención, parte o de no realizar trabajo alguno (pensemos, por ejemplo, en estudiantes de tiempo completo o quienes apenas concluyeron sus estudios o también alguien que perdió su empleo y no se involucra en ninguna otra tarea o rol salvo la de proveedor del hogar) o de realizar algún otro tipo de trabajo **siempre y cuando sea distinto** de la modalidad que involucra transacciones de mercado.

Quienes cumplan con esas condiciones (es decir, no estar insertos en un trabajo por remuneración o beneficio, pero que ahí quieren dirigirse) formarán de un modo u otro el numerador de estas tasas.

Es importante entender aquí el significado de esto a la luz de lo que modifica en la vieja lógica de 1982, según la cual trabajo y mercado laboral eran sinónimos. Bajo el principio de las denominadas reglas de prioridad,²⁴ si la persona ya contribuía con tiempo de trabajo a la frontera de producción del SCN se le consideraba ocupada, lo cual signifi-

caba que, en lo que respecta a la tasa de desocupación abierta, estaba presente en el denominador (por estar referido el indicador a algo que acontece en la PEA), pero no en el numerador de la tasa, aunque estuviera en situación de búsqueda y disponible para tomar trabajo remunerado. En la XIX CIET, por su parte, al diferenciar modalidades dentro de esa frontera, la regla de prioridad se circunscribe sólo a una vertiente de trabajo (la de transacciones de mercado o por remuneración o beneficio) de modo que si, por ejemplo, tenemos campesinos en la agricultura de subsistencia (trabajo generador de bienes para el consumo propio) y éstos se encuentran en situación de búsqueda y/o disponibilidad de inserción al mercado laboral, pueden formar ahora parte del numerador de la tasa, porque su modalidad de trabajo (autoconsumo) no tiene prioridad sobre su intención de inserción en el mercado laboral. En suma, ahora pueden ser considerados desocupados (desempleados abiertos) de haber tal intención.

Por ello, en la *Resolución I* de la XIX CIET, si bien la condición de ocupación de mercado (aportar algún insumo laboral a ese proceso) mantiene su prioridad sobre estar en situación de búsqueda/disponibilidad, esta última situación tiene prioridad, a su vez, a la hora de clasificar a los individuos sobre cualquier otra condición de trabajo, es decir, sobre el hecho de si la persona realiza otra modalidad distinta a la de mercado laboral (dentro o fuera de la frontera de la producción) como, desde luego, sobre cualquier otra actividad no asimilable al trabajo.²⁵

El diagrama 3 muestra todas las subpoblaciones que van a jugar un papel en las cuatro tasas de subutilización que considera la *Resolución I* de la Conferencia y cómo se combinan en los numeradores de cada tasa:

- Desocupación. Son todas aquellas personas en edad de trabajar que no estaban ocupadas (es decir, que no aportaban insumo laboral alguno al trabajo en contexto de tran-

²⁴ Hussmanns, Mehran y Verma. *Op. cit.*, pp. 38-40.

²⁵ Ver OIT. XIX CIET. *Resolución...*, *op. cit.*, sección *Clasificación de la población en edad de trabajar*, p. 5, párrafos 14, 15 y 16.

sacciones de mercado), que habían llevado a cabo actividades de búsqueda de un puesto de trabajo durante un periodo reciente especificado y que estaban disponibles en la actualidad para ocupar un puesto de trabajo en caso de que existiera oportunidad de hacerlo.²⁶

- Subocupación. Este concepto se asimila al de subocupación por insuficiencia de tiempo de trabajo (más adelante se abundará al respecto) y se refiere a todas aquellas personas ocupadas que durante el periodo de referencia deseaban trabajar horas adicionales; estaban disponibles para hacerlo y cuyo tiempo de trabajo era inferior a un valor umbral especificado de horas.²⁷
- Fuerza de trabajo potencial. Todas las personas en edad de trabajar que, durante el periodo de referencia corto, no estaban ocupadas ni desocupadas y que: a) habían llevado actividades de búsqueda sin estar disponibles a la fecha (ver nota al pie 23) o b) no llevaron actividades de búsqueda,

pero deseaban un puesto de trabajo y estaban actualmente disponibles.²⁸

En las secciones siguientes analizaremos cómo se construyen cada una de estas tasas, los resultados que darían para México en función del diseño actual de la ENOE y en qué medida logran o no sus objetivos propuestos.

Diagrama 4

Tasa de desocupación

$$SU1 = \frac{\text{Desocupación}}{\text{Fuerza de trabajo}} \times 100$$

Tasa SU1 (desocupación abierta)

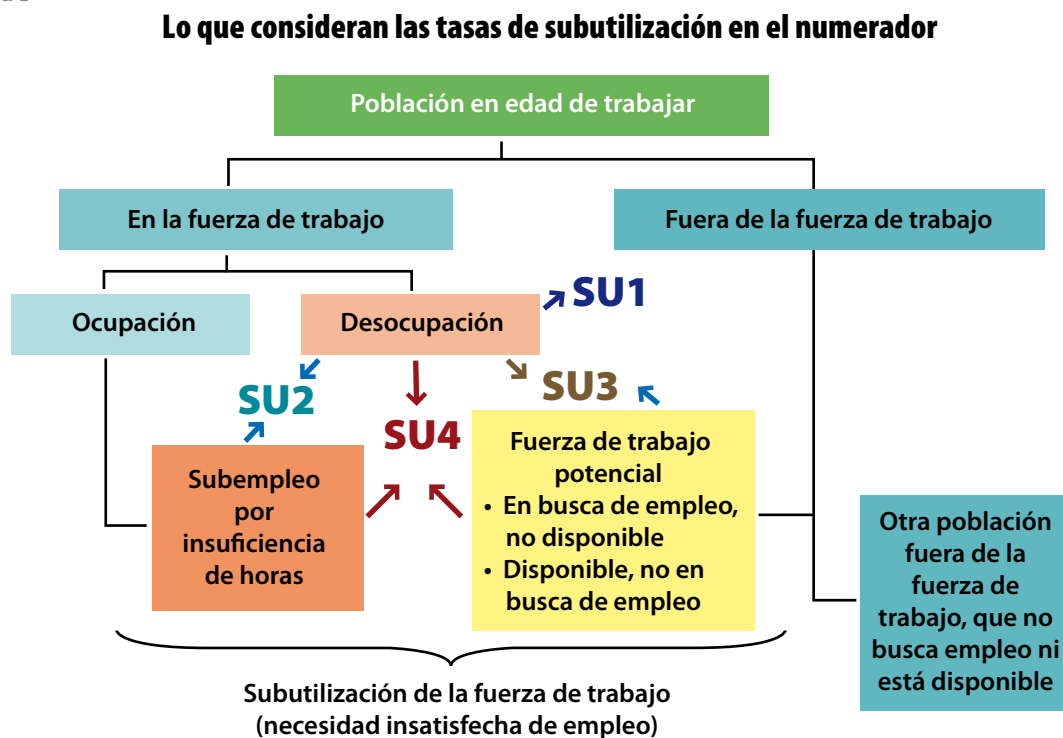
La denominada tasa de subutilización 1 (SU1) será casi idéntica a la tasa de desocupación abierta, vigente en la actualidad en la mayoría de los países

26 Ibid., sección *Desocupación*, pp. 11-12, párrafo 47.

27 Ibid., sección *Subocupación por insuficiencia de tiempo de trabajo*, p. 11, párrafo 43.

28 Ibid., sección *Fuerza de trabajo potencial (entrantes)*, p. 13, párrafo 51.

Diagrama 3



desarrollados,²⁹ pero distinta en magnitud a la que registran naciones en desarrollo. Ello es así porque la agricultura de subsistencia (en tanto modalidad predominante de trabajo de producción para el autoconsumo que formaba parte de la PEA) se deduce de ésta, de modo que el denominador resultante (fuerza de trabajo) queda como una magnitud menor de forma perceptible y, por ende, la tasa ha de incrementarse.

De esta manera es como la OIT procura resolver la paradoja de que ahí en donde hay más pobreza la TD tiende a ser menor, cosa que se percibe al comparar regiones heterogéneas dentro de una nación o, también, naciones de distinto nivel de desarrollo (siempre y cuando hayan seguido las mismas reglas de construcción de la TD). Es de subrayar que el *staff* técnico de la OIT consideró correctamente que esto nada tiene que ver con el famoso criterio de clasificar como ocupado a quien trabajó, al menos, una hora en la semana de referencia y, sin duda, hubo consenso en la XIX CIET de que tal referente no tendría por qué modificarse³⁰ ya que siendo la hora trabajada la unidad del insumo laboral en el SCN, sería lógicamente inconsistente que haya aportación al PIB de los que, por otra parte, fueran susceptibles de ser clasificados como desocupados con un criterio más relajado. También, se ha demostrado de forma empírica que la noción de que hay una cantidad significativa de individuos trabajando sólo una hora diaria en el mercado laboral tiene poca importancia cuantitativa y, por ende, poco efecto si se les dejara de incluir en la ocupación.³¹ La solución que busca la OIT a esta controversia es: 1) por la vía conceptual y 2) por la inclusión, en todo caso, de la insuficiencia de horas trabajadas en la tasa SU2, mas no por la relajación de un criterio de demarcación entre la ocupación y la desocupación abierta.

Sin embargo, para que las TD (ahora SU1) rompieran con la paradoja de niveles más bajos en zonas

de rezago, debiera tener un efecto más por la vía del numerador que por un ajuste al denominador. Visto en términos de quienes se dedican a la agricultura de subsistencia, esto significaría que se presentaran como buscadores sistemáticos de empleo en el mercado laboral: inclusión en teoría factible en el numerador porque hay que recordar que las nuevas reglas de prioridad subrayarían esa condición por encima del hecho de que estén realizando un trabajo no circunscrito en el mercado laboral.

Las pruebas empíricas muestran, sin embargo, que éste no sería el caso: sólo una fracción de individuos dedicados a estas actividades se comportan como buscadores sistemáticos en sus localidades de residencia. Ello es así porque, de entrada, no hay mercados laborales permanentes y estructurados al alcance y, asimismo, los individuos tienen un conocimiento completo de sus posibilidades de inserción en su localidad y los alrededores, de modo que no tiene sentido comportarse como buscadores de trabajo.

Aun cuando este segmento poblacional se comportara como buscador sistemático, ello no bastaría para acercar la TD de países menos desarrollados (como México) a la de los países desarrollados. Un elemento no considerado por el *staff* de la OIT es que la TD, más que ser un indicador de rezago o atraso, es uno de mercado laboral estrictamente hablando que, como tal, puede ser el reflejo de una modalidad de éste, propensa a calibraciones más por el lado de los salarios y la calidad del empleo en general que por su volumen, es decir, que para economías que en el pasado han experimentado *shocks* recurrentes acompañados de draconianos programas de ajuste, el laboral puede comportarse más como un mercado que se adapta con mayor intensidad por el lado de los precios que por el de las cantidades, como ha sido el caso de México en las últimas décadas.³²

Por todo ello, lo primero que se puede observar es que el incremento que presentaría la SU1 con respecto a la TD difundida para un mismo periodo (ver gráfica 6) sería todo menos espectacular, pues pasaría de 4.87 a 5.04. El margen con respecto al in-

29 La magnitud de la tasa no habrá de modificarse en la mayoría de las naciones desarrolladas porque lo que se deduce de la PEA es tan poco que ésta y fuerza de trabajo son magnitudes casi equivalentes. Sólo para aquellos países como Alemania y Suiza donde el aprendizaje laboral sin remuneración ha sido una vía institucionalizada de inserción laboral, su eliminación del denominador cuantitativamente pudiera tener algún efecto.

30 OIT. XIX CIET. Resolución... *op. cit.*, sección *Ocupación*, p. 7, párrafo 27, inciso a.

31 Para el caso de México, ver Negrete. "El indicador de la polémica...", *op. cit.*

32 *Ibid.*

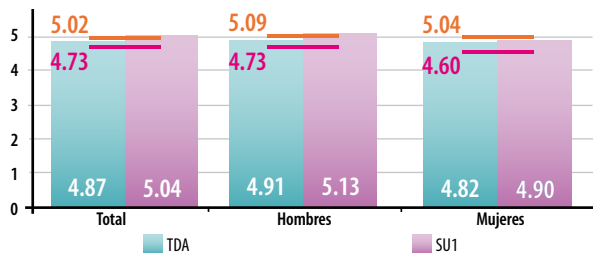
tervalo superior de la estimación (5.02) es tan pequeño que apenas resulta estadísticamente significativo. El mayor efecto de ajuste ocurre por la eliminación del denominador de la agricultura de subsistencia, más que por su inclusión en el numerador. La gráfica 7 muestra, por su parte, que no se rompe con la tendencia del indicador a vincularse más a lo urbano que a lo rural: se trate de TD o de SU1, el nivel disminuye en ambos casos conforme queda referido a localidades más pequeñas y, definitivamente, el menor de todos se registra para poblaciones menores de 2 500 habitantes, es decir, las más rurales, con todo y que ahí la diferencia entre la TD publicada y lo que sería la SU1 es la mayor observada (2.7 vs. 3.13).

Visto para todo el periodo que cubre la ENOE (ver gráfica 8), el promedio de las diferencias sería de poco más de una décima de punto porcentual, con un máximo que se acercaría a dos y un mínimo de seis centésimas, aunque hay que observar que las mayores diferencias ocurren después del 2009 y no antes. De cualquier modo, al ser tan pequeñas, no es de sorprender que la tendencia sea la

misma. De hecho, para el ámbito urbano de más de 100 mil habitantes (ver gráfica 9) casi se trata de series idénticas, y sólo para el dominio de muestra representativo de localidades rurales dispersas se llegan a registrar diferencias que rebasan las seis décimas de punto porcentual (ver gráfica 10).

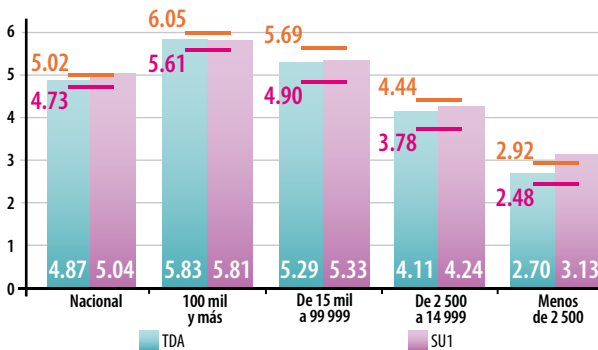
Gráfica 6

Comparativo de la tasa de desocupación abierta (TDA) y SU1



Gráfica 7

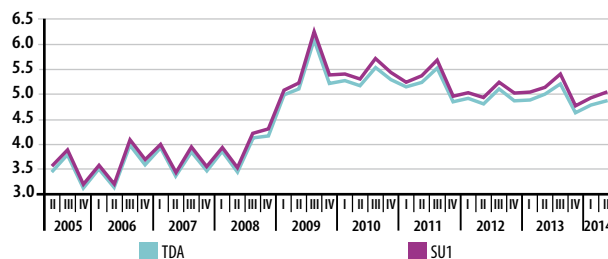
Comparativo de la TDA y SU1, por tamaño de localidad



Gráfica 8

Evolución de la TDA y SU1: ámbito nacional, 2005-2014

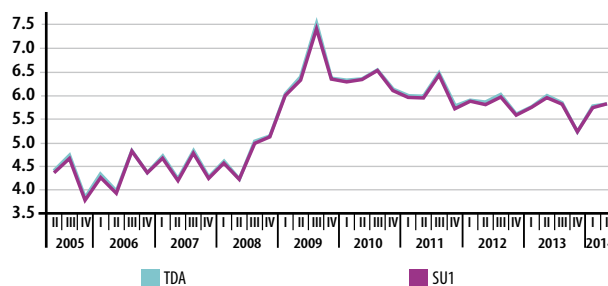
Promedio de las diferencias: 0.12; diferencia máxima: 0.19; diferencia mínima: 0.06; coeficiente de correlación: 0.9997.



Gráfica 9

Evolución de la TDA y SU1: localidades de 100 mil y más habitantes, 2005-2014

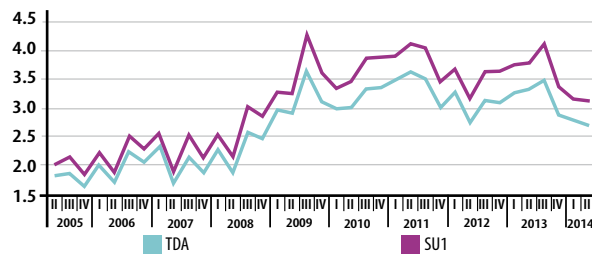
Promedio de las diferencias: -0.04; diferencia máxima: -0.01; diferencia mínima: -0.09; coeficiente de correlación: 0.9998.



Gráfica 10

Evolución de la TDA y SU1: localidades de menos de 2 500 habitantes, 2005-2014

Promedio de las diferencias: 0.40; diferencia máxima: 0.64; diferencia mínima: 0.17; coeficiente de correlación: 0.9964.

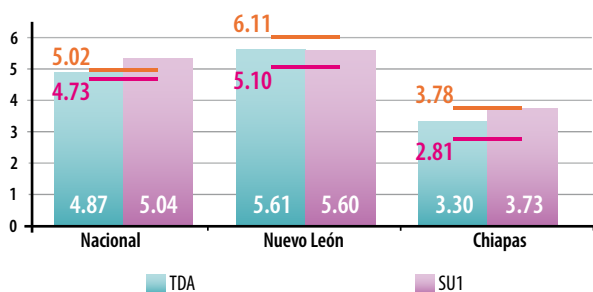


Un contraste que cabría hacer en un país tan heterogéneo como México sería entre una entidad federativa de economía moderna y otra en la que lo rural sigue teniendo un peso específico considerablemente mayor que en el promedio del país: Nuevo León, con un PIB per cápita (ajustado por poder de paridad de compra) similar al de la República Checa, y Chiapas —con uno a medio camino entre el de Guatemala y El Salvador— se prestan para explorar cómo se comportaría el indicador después de pasar de TD a SU1. La gráfica 11 muestra que el cambio no alcanza a revertir el hecho de que ahí

donde existen mercados laborales estructurados (Nuevo León) el indicador seguirá siendo mayor que en Chiapas. De hecho, en la primera entidad no sufre cambio tangible alguno, mientras que en la segunda, si bien es palpable el incremento (de 3.30 a 3.73), todavía queda contenido en el intervalo de la estimación. Generalizando para todas las entidades del país, la gráfica 12 presenta el ordenamiento a que da lugar TD frente al que daría SU1: las entidades que cambian de posición relativa son las de diferente coloración. SU1 reducirá la varianza a nivel nacional al cerrar un poco las brechas, pero no más que eso.

Gráfica 11

Comparativo de la TDA y SU1: ámbito nacional, Nuevo León y Chiapas

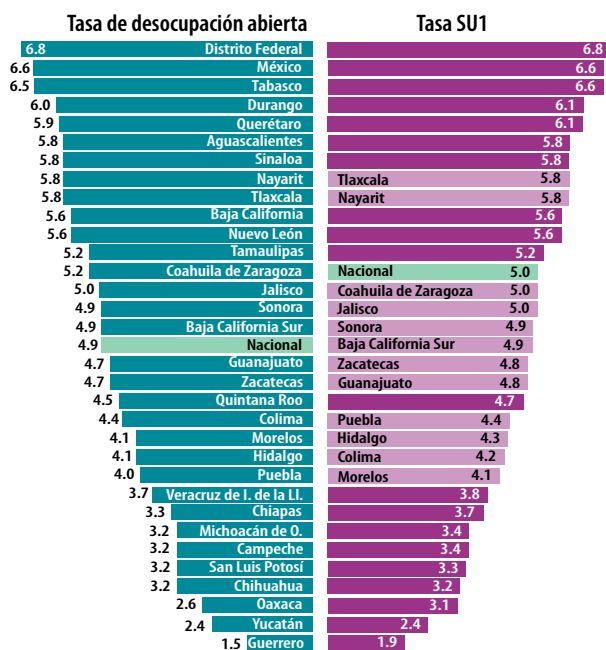


Tasa de subocupación

Antes de pasar a la tasa SU2, misma donde se integran en el numerador la desocupación abierta y la subocupación, resulta necesario detenerse en las modificaciones que sufre este último concepto y analizarlo por separado. El de subocupación es probable que sea el indicador más revisado por las CIET a lo largo de su historia sin llegar a una solución del todo satisfactoria tanto para las oficinas nacionales de estadística como para organismos internacionales, como la OCDE. La XI CIET (1966) abordó el tema bajo el enfoque de subempleo visible e invisible, conceptos que fueron derogados por la XVI (1998), la cual se centró en el subempleo por criterio de tiempo (*time-related underemployment*). A su vez, el problema con esta CIET es haber apuntado a una noción en la que se contabilizan personas que quieren trabajar más de lo que les permite su ocupación actual, pero sin suministrar reglas operativas precisas de cómo implementarlo, quedando ello sujeto a una libre interpretación. Así, la OCDE —que concentra y recopila estadísticas de todo tipo de sus países miembros, incluyendo, desde luego, laborales—, consciente de los problemas para adoptar el concepto de una manera práctica, ha optado hasta la fecha simplemente por el de empleo a tiempo parcial como un indicador más seguro en sus publicaciones.³³

Gráfica 12

Comparativo de la TDA y SU1 por entidad federativa



33 Tal es el caso de OECD. *Labour Force Statistics; Employment Outlook and Education at Glance*, así como en *Family database*.

La XIX CIET trató de ser más puntual al respecto. De acuerdo con ella, es fundamental que se contabilicen: 1) a personas que desean trabajar horas adicionales y 2) siempre y cuando el número de horas en su ocupación actual sea inferior a un umbral especificado.³⁴ Este último criterio hace más preciso o más restrictivo el concepto según se quiera ver.

En el caso de la tasa de subocupación que venía difundiendo la ENOE quedarían sólo los casos de personas que laboran menos de 35 horas a la semana (menos de lo que acumula quien labora jornadas completas), amén de depurarse de entre quienes cumplen esta condición, los dedicados a la agricultura de subsistencia, dado que el indicador —al igual que SU1— se calcula como porcentaje de quienes están circunscritos al trabajo por paga o remuneración (es decir, al trabajo que se da en un contexto de transacciones de mercado).

Es de destacar, entonces, que quedarían fuera del indicador todos aquellos cuya jornada rebasa el umbral, pero que dan señales de que su actividad es baja o poco intensa y que la ENOE hasta ahora contempla: situaciones como la de quien está al frente de una tienda de conveniencia en una coyuntura con muy pocos clientes (semivacía de actividad), no importando que haya mantenido abierto el negocio de sol a sol.³⁵ Incluso, ésta es una situación que llega a presentarse entre asalariados que cumplen un horario de ocho horas diarias, en especial en el sector privado, previo a un recorte o despido en las semanas subsecuentes. La recomendación al respecto de la XIX CIET parece no percatarse que, después de todo, los asalariados no deciden la extensión de su jornada o ésta no se puede adaptar tan fácilmente a

las circunstancias: así como hay algún grado de rigidez salarial (*sticky wages*), también la hay con respecto a los horarios de trabajo.

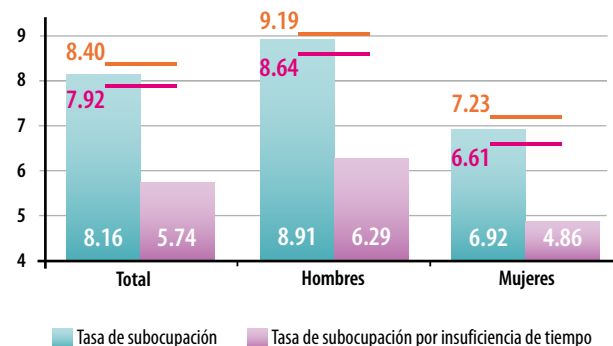
La eliminación de casos como éstos para apegarse ahora a la XIX CIET se traduce en una significativa disminución del nivel de la tasa de subocupación de 8.16 a 5.74 puntos a nivel nacional (ver gráfica 13); la disminución sería aún mayor entre la población masculina, quizá porque en ella hay una mayor propensión a jornadas que traspasan el umbral que entre la población femenina.

Por otra parte, la tasa de subocupación de la ENOE —a diferencia de la de desocupación— sí tenía la peculiaridad de ser mayor en áreas rurales que urbanas. Por fortuna, esta propiedad no se pierde al haberse eliminado los casos que incluía por arriba del umbral. De hecho, cabe observar que es en áreas rurales donde la disminución es menor (ver gráfica 14) por el hecho de que la eliminación de casos en el numerador casi se compensa con la desaparición en el denominador de la agricultura de subsistencia. Es importante decir al respecto que la única posibilidad de figurar en los indicadores SU que queda a quienes se dedican a esta actividad es por la vía de mostrar disponibilidad de trabajar (lo que daría un lugar en el numerador en SU3 y SU4), pero el problema de cómo interpretar esa disponibilidad y resolverla de manera operativa no es menor (punto que se abordará en la parte final de este ensayo).

34 El valor de umbral puede fijarse ya sea en el que marca la frontera entre trabajo a tiempo completo y parcial; en los valores medianos o modales de las horas habitualmente trabajadas de todas las personas ocupadas o en las normas de tiempo de trabajo "... tal cual se especifican en la legislación o en las prácticas nacionales..."; ver OIT. *XIX CIET. Resolución...*, op. cit., sección *Subocupación por insuficiencia de tiempo*, p. 11, párrafo 43, inciso c.

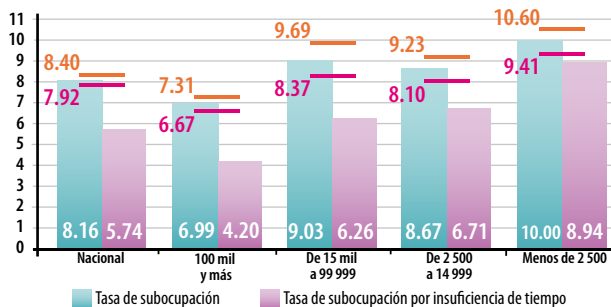
35 El cuestionario temático de la ENOE con la pregunta 5a indaga de manera explícita si en la semana de referencia la persona que labora tuvo poco trabajo que hacer y en la pregunta 5b, si ello es motivo de inquietud o preocupación, esto último para descartar situaciones en las que la poca actividad es lo habitual (como en el trabajo de vigilancia) o en los que la poca actividad es lo normal en ese momento del año.

Gráfica 13
Comparativo de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo



Gráfica 14

Comparativo de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo, por tamaño de localidad

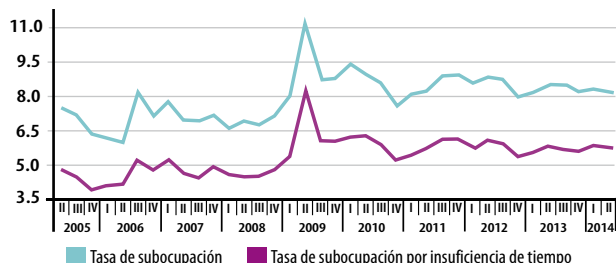


Vista la serie en su conjunto (ver gráfica 15), en promedio se tendría una disminución cercana a los 2.6 puntos porcentuales pudiendo alcanzar, en algún momento, un máximo de disminución que supere los tres puntos. La tendencia, sin embargo, apenas se vería modificada.

Gráfica 15

Evolución de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo: ámbito nacional, 2005-2014

Promedio de las diferencias: -2.59; diferencia máxima: -3.15; diferencia mínima: -1.90; coeficiente de correlación: 0.9793.



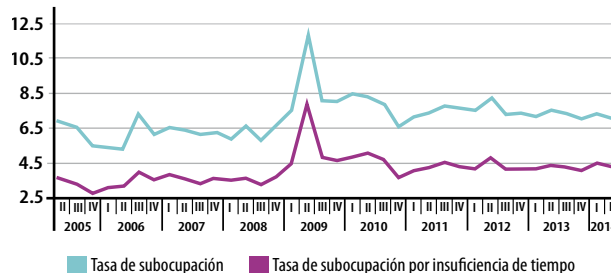
Vista la serie histórica por dominios de muestra, y al revés de lo que sucede en la comparación entre la TD y SU1, las diferencias mayores habrán de registrarse en las áreas urbanas de más de 100 mil habitantes, siendo más moderadas en las áreas rurales dispersas (ver gráficas 16 y 17).

El contraste entre Nuevo León y Chiapas corrobora que es en la entidad de economía moderna donde la disminución sería más acusada mientras que en la más atrasada los niveles siguen siendo mayores y menor la disminución de la tasa.

Gráfica 16

Evolución de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo: localidades de 100 mil y más habitantes, 2005-2014

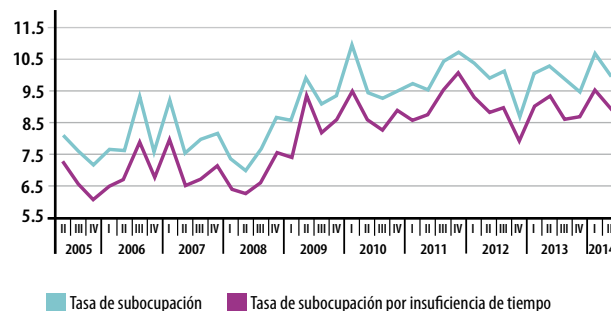
Promedio de las diferencias: -3.02; diferencia máxima: -3.70; diferencia mínima: -2.14; coeficiente de correlación: 0.9709.



Gráfica 17

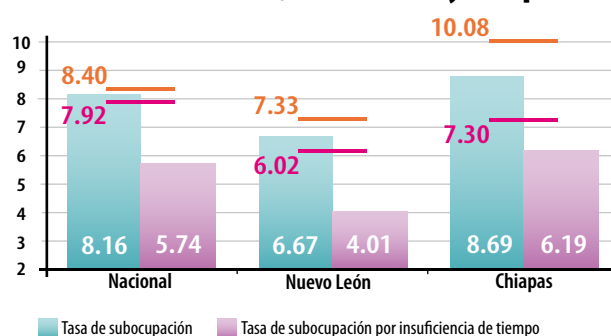
Evolución de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo: localidades de menos de 2 500 habitantes, 2005-2014

Promedio de las diferencias: -0.99; diferencia máxima: -1.50; diferencia mínima: -0.48; coeficiente de correlación: 0.9815.



Gráfica 18

Comparativo de la tasa de subocupación y tasa de subocupación por insuficiencia de tiempo: ámbito nacional, Nuevo León y Chiapas



SU2

Ésta es el otro indicador promovido por la XIX CIET que se calcula con respecto a la fuerza de trabajo, es decir, referido a la PEA menos la población en formas de trabajo no de mercado dentro de la frontera de producción del SCN (ver diagrama 5). Unifica en el numerador a todos los que muestran disponibilidad para invertir tiempo en un trabajo de mercado en dos vertientes: 1) quienes no hayan trabajado ni una hora en la semana o no tengan vín-

Diagrama 5

Tasa compuesta de desocupación y subempleo por insuficiencia de horas

$$SU2 = \frac{\text{Subempleo por insuficiencia de horas} + \text{Desocupación}}{\text{Fuerza de trabajo}} \times 100$$

culo laboral alguno al momento,³⁶ presentando un comportamiento de buscadores activos (desocupados abiertos) y 2) quienes trabajaron por debajo de un umbral de horas y que, en su caso, querrían invertir **más** tiempo de trabajo de mercado de lo que lo pueden hacer en la actualidad (subocupados por insuficiencia de tiempo de trabajo).

Esta misma idea de presentar una suerte de *continuum* entre la población que procura insertarse en el mercado laboral con un grado de ocupación nulo o cero con aquella otra con un número reducido de horas, la ENOE ya lo manejaba en un indicador denominado tasa de ocupación parcial y desocupación (TOPD1) en la que simplemente se añadían al numerador quienes trabajaron menos de 15 horas en la semana de referencia, no importando si desearían trabajar más que eso o no. La diferencia, ahora, de involucrar a un concepto como el de la subocupación por una parte expande el umbral hasta las 35 horas, pero por la otra, es más exigente al procurar que haya una disposición para trabajar más de lo que se labora (en el

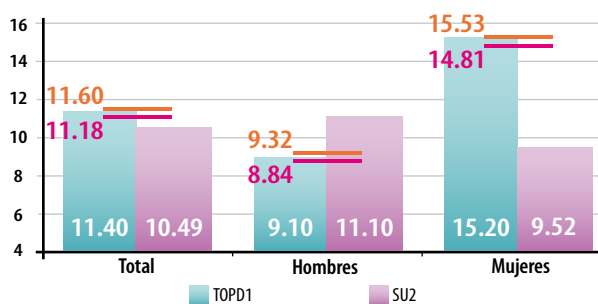
³⁶ Es decir, que si no trabajaron en la semana de referencia, ello fue por no tener trabajo y no por estar de vacaciones laborales o por cualquier otra razón de ausencia en el periodo que se pueda presentar entre quienes tienen un empleo.

contexto de las transacciones de mercado). No hay que perder de vista que el concepto de subocupación amplía y restringe al mismo tiempo y el efecto neto de ello se verá al compararlo con TOPD1. Con esto no se quiere decir en modo alguno que, por arrojar magnitudes menores, SU2 sea un indicador menos deseable o deficiente: después de todo, puede tener más sentido. Sólo no hay que olvidar los factores que podrían explicar la diferencia con respecto al indicador que habrá de reemplazar, sobre todo cuando se está introduciendo un concepto de subocupación que, además, elimina situaciones consideradas en la subocupación que difundía la ENOE, tal como se explicó en el apartado anterior (en todo caso, es ahí donde pudiera haber controversia).

De esta forma, en la gráfica 19 se puede apreciar que SU2 es un indicador que presenta un nivel ligeramente por debajo de TOPD1 (11.40 vs. 10.49), sin ser la reducción significativa al no aproximarse al límite inferior de la estimación TOPD1. Pese a lo anterior, no deja de resaltar que el nuevo indicador arrojaría niveles más altos para la población masculina, pero menores de forma considerable para la femenina. Tal parece que TOPD1 involucraba un monto considerable de mujeres con periodos de actividad de mercado muy cortos, sin que ello fuera necesariamente una situación no deseada o en contra de su interés (porque, por ejemplo, permite en varios casos acomodar a la carga doméstica a mujeres que contribuyen de manera auxiliar a negocios familiares, sin necesidad de involucrarse más allá de las 15 horas semanales),

Gráfica 19

Comparativo de la tasa de ocupación parcial y desocupación (TOPD) y SU2



mientras que en SU2, al tener a bordo el concepto de subocupación, implica de una manera u otra que el tiempo de trabajo corto es una situación no deseada. La división por sexo del trabajo algo tiene que ver en que sea una situación menos tolerada por la población masculina (que se supone más involucrada en el trabajo de mercado que en el que se realiza fuera de él) y, por ende, menos dada a desempeñar un papel auxiliar (desde el cual no tiene mucho caso rendir más horas). En suma, SU2 elimina muchos casos de contribución auxiliar, mismos que no están distribuidos uniformemente entre los sexos.

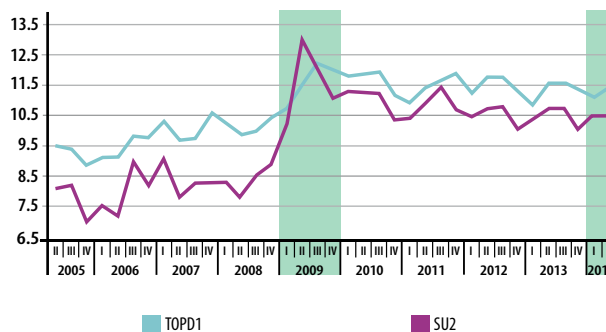
Visto por dominios de muestra (tamaños de localidad), queda claro que, pese a la eliminación del concepto de subocupación de casos de jornada normal o prolongada pero de baja intensidad laboral, la propiedad de presentar los mayores niveles en las áreas rurales dispersas se mantiene y ello no deja de reflejarse en SU2, donde el porcentaje de fuerza de trabajo contabilizado es de 11.79 contra 9.77 en las áreas más urbanizadas o representativas de localidades de 100 mil y más habitantes (ver gráfica 20), aunque no hay que dejar de tener presente que SU2 excluye a la que gravita en torno a la agricultura de subsistencia.

En promedio, SU2, a lo largo de todo el periodo, se habrá de ubicar para el país en su conjunto en un nivel inferior a un punto porcentual a TOPD1. No deja de llamar la atención, sin embargo, que durante la recesión del 2009, SU2 habría sido más

sensible a esa coyuntura y, por ende, registrado niveles mayores, cosa que por lo demás ya no sucede en el lapso de desaceleración del 2014 (ver gráfica 21).

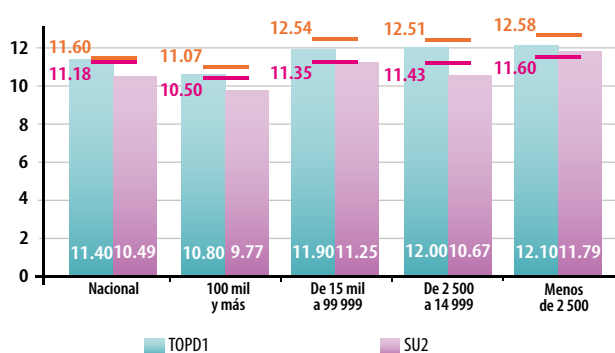
Gráfica 21
Evolución de la TOPD y SU2: ámbito nacional, 2005-2014

Promedio de las diferencias: -1.05; diferencia máxima: 1.63; diferencia mínima: -2.38; coeficiente de correlación: 0.9232.

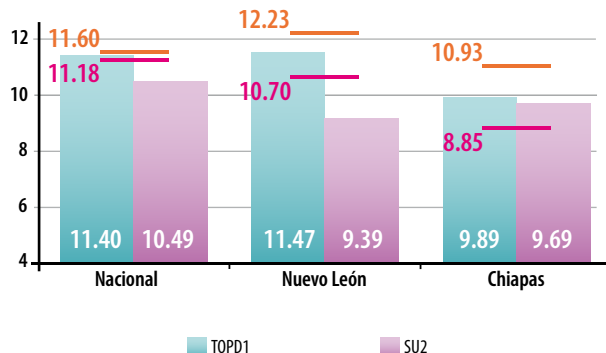


En cuanto a lo que sucede con regiones contrastantes en su nivel de desarrollo (Nuevo León y Chiapas), SU2 habría de registrar un nivel un tanto superior en la de mayor rezago (9.69), aunque parece que ello no sería sistemático si se toma en cuenta que el promedio nacional es aún mayor (10.49). Por lo demás, el cambio del indicador de TOPD1 a SU2 pareciera incidir mucho más en Nuevo León (ver gráfica 22), donde el tiempo parcial voluntario puede ser una práctica más extendida (misma que SU2 filtra).

Gráfica 20
Comparativo de la TOPD y SU2, por tamaño de localidad



Gráfica 22
Comparativo de la TOPD y SU2: ámbito nacional, Nuevo León y Chiapas



SU3

La tercera tasa que plantea la XIX CIET ya no es una de subutilización de fuerza de trabajo, sino de subutilización de algo que incluye a la fuerza de trabajo potencial y, al hacerlo, se denomina fuerza de trabajo ampliada (ver diagrama 6). La idea central aquí es que hay alguna continuidad entre quienes no están en el mercado de trabajo, pero procuran insertarse en él de forma activa (desocupados abiertos) y entre quienes no están en el mercado de trabajo y quisieran insertarse en él, pero carecen de algún otro rasgo que caracteriza el comportamiento de los desocupados abiertos.³⁷

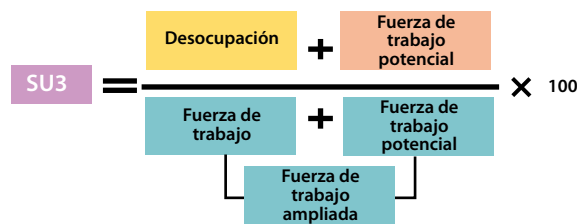
Como se recordará, la fuerza de trabajo potencial incorpora población que reúne dos de las tres condiciones que definen a la población desocupada abierta. Al igual que ésta, suponen un grado de ocupación nulo y contempla dos situaciones: a) disponibilidad para realizar trabajo en un contexto de transacciones de mercado, pero sin comportamiento de búsqueda o b) no disponibilidad al momento, pero presentando comportamiento de búsqueda. Normalmente, la abrumadora mayoría de casos corresponden al primer inciso (ver nota al pie 23).

Así, la denominada población disponible es clave en esta construcción. Las razones para no estar buscando de forma activa un empleo pueden decir mucho de la población de la que se está hablando. La ENOE, por ejemplo, identifica razones como: no hay trabajo en su especialidad, oficio o profesión; en su localidad no hay trabajo o sólo se realiza en ciertas temporadas del año; no cuenta con la escolaridad, los papeles o la experiencia necesaria para

37 Esta otra idea de continuidad entre desocupados y otros grupos de población no ocupados en el mercado de trabajo (digamos, una continuidad en dirección contraria a la que plantea SU2), se había plasmado en la ENEU (antecesora de la ENOE) en un indicador denominado tasa de desempleo abierta alternativa (TDAA). Sin embargo, esta idea no daba un indicador interesante porque añadía a los desocupados población disponible para trabajar entendiendo por disponibilidad sólo a los *desalentados* (*discouraged*), es decir, a aquellos que habían buscado un trabajo, pero habían terminado por desistir. El diseño del cuestionario de la ENEU no se planteaba que podría haber población con necesidad de trabajar que sencillamente no busca trabajo por una serie de cuestiones que tienen que ver con la vulnerabilidad o marginalidad. Fue con el cuestionario de la ENOE que el horizonte de lo que puede ser población disponible se amplió de forma considerable lo cual permite, entre otras cosas, el ejercicio que aquí se muestra.

Diagrama 6

Tasa compuesta de desocupación y fuerza de trabajo potencial



realizar un trabajo; piensa que por su edad o por su aspecto no lo aceptarán en un trabajo; la inseguridad pública o el exceso de trámites lo desalientan a iniciar una actividad. No sorprende, entonces, que se trate de población al margen del mercado laboral o en la que predominen grupos vulnerables. Es importante señalar, a su vez, que la ENOE antes de preguntar por las razones tiene el cuidado de preguntar si, pese a que la persona no ha estado buscando trabajo (de mercado), tiene necesidad de uno.³⁸

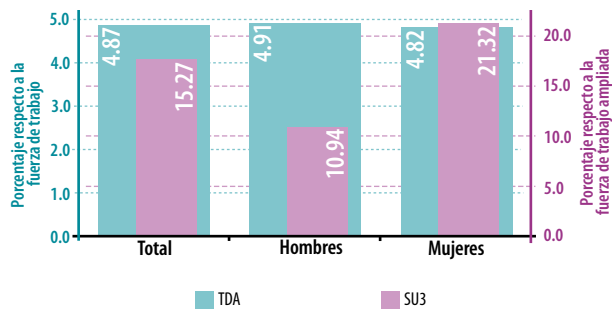
Dado lo anterior, es de esperarse que un indicador en el que influye tanto ese grupo poblacional (que las circunstancias lo sitúan fuera de alcance, por así decirlo, del mercado laboral), involucre una proporción considerablemente mayor de mujeres con respecto a hombres que las dos tasas SU previas. Haciendo uso de una doble escala para contrastar con la tasa de desocupación abierta, la gráfica 23 muestra que mientras para este indicador las proporciones son equilibradas, con SU3 se involucra a un porcentaje de población femenina casi del doble que la masculina (21.32 contra 10.94). Para toda la población, el nivel del indicador se sitúa en 15.27 de la fuerza de trabajo ampliada, es decir, de la población o que ya se encuentra en el mercado laboral (ocupados) o que quiere participar en él mostrando un comportamiento activo (desocupados abiertos) o pasivo (fuerza de trabajo potencial dominada por población disponible). Al interpretar la gráfica, hay que tomar en cuenta que TDA y SU3 no se refieren al mismo denominador,

38 La pregunta en concreto es la 2f: ¿Actualmente tiene necesidad de trabajar?, a la que sigue otra sobre las razones de por qué, teniendo esa necesidad, no ha estado buscando trabajo (2g).

pero se muestran de manera conjunta para ver su comportamiento diferencial respecto a las poblaciones masculina y femenina, así como más adelante se hace lo propio entre dominios de muestra y regiones.

Gráfica 23

Comparativo de la tasa de desocupación (TD) y SU3

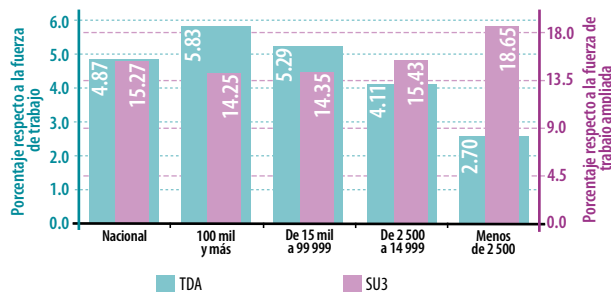


Visto por dominios de muestra o tamaños de localidad (ver gráfica 24), queda claro que SU3 presenta el comportamiento opuesto a la TDA, con los mayores niveles en las localidades rurales dispersas (18.65 frente al 14.25 de las localidades urbanas de más de 100 mil habitantes).

La gráfica 25 se centra sólo en la población femenina y deja claro que en localidades rurales dispersas el indicador recoge hasta un tercio de lo que sería fuerza de trabajo ampliada, mientras que en el dominio de muestra representativo de localidades mayores a 100 mil habitantes, la magnitud es casi la mitad que registran las áreas rurales dispersas. El indicador, ciertamente, visibiliza algo que con anterioridad estaba fuera de foco.

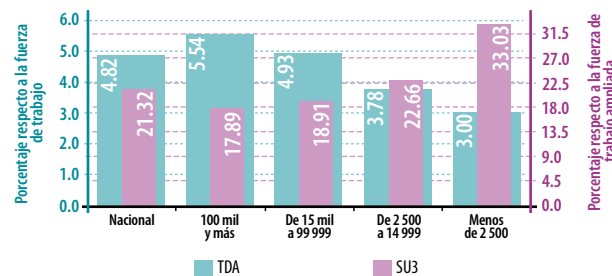
Gráfica 24

Comparativo de la TD y SU3, por tamaño de localidad



Gráfica 25

Comparativo de la TD y SU3, por tamaño de localidad: mujeres

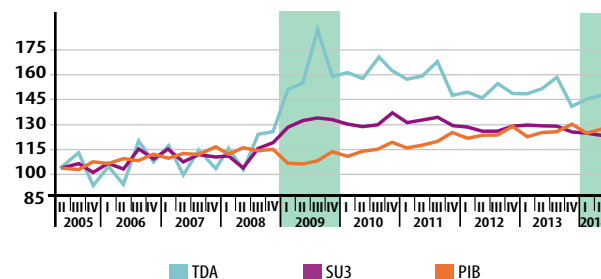


Sin embargo, desde la óptica de una serie en la que se presentan distintas coyunturas, lo que se gana en niveles se pierde un tanto en sensibilidad frente al ciclo económico. En ese sentido, SU3 no puede competir con SU1 (o con la tasa de desocupación abierta dado el caso). En la gráfica 26, y dadas las diferencias de escala entre SU3, la TDA y el PIB a precios constantes, las tres magnitudes se transforman en un índice base 2005 para observar su comportamiento conjunto. Durante la crisis del 2009, la respuesta de SU3 fue menos pronunciada que la de TDA y lo sería aún menos si excluyéramos del indicador a los desocupados, quedándonos sólo con la población disponible; pero la capacidad de respuesta del indicador se diluyó del todo en el 2014 cuando el escenario ya no fue de recesión sino de desaceleración.

Gráfica 26

Evolución de la TD, SU3 y PIB: ámbito nacional, 2005-2014

Índices segundo trimestre del 2005 = 100

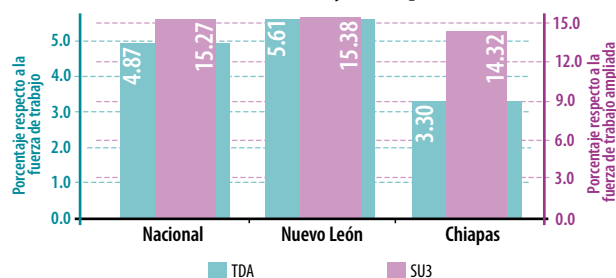


Lo anterior no quiere decir necesariamente que el indicador (SU3) esté mal planteado, sino que dice algo que no dice SU1 y viceversa. El que los indicadores se designen ahora como SU es una manera como la OIT y la XIX CIET tratan de transmitir la idea de que son indicadores con el mismo nivel de relevancia.

Por último, en el contraste de ámbitos desarrollados y rezagados (ver gráfica 27), llama la atención que, pese a ser sensible a la situación que se presenta en localidades rurales dispersas, ello no se trasluce en que entidades como Chiapas registren tasas mayores a Nuevo León. La condición de disponibilidad puede presentar también algún carácter regional que vaya más allá del binomio rural-urbano.

Gráfica 27

Comparativo de la TD y SU3: ámbito nacional, Nuevo León y Chiapas

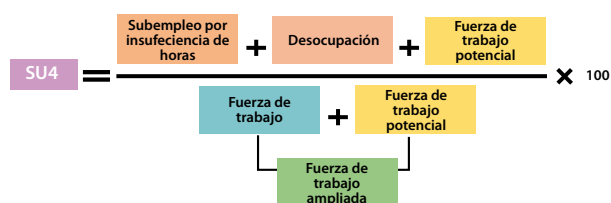


SU4

Ésta es el agregado mayor que combina en el numerador todas las situaciones hasta ahora descritas; es como un espectro que comprende desde una inserción débil (subempleo por insuficiencia de horas) a un intento de inserción activo (desocupación abierta) y al final, a una mera posibilidad de inserción (fuerza de trabajo potencial). Al igual que SU3, el indicador se calcula como porcentaje de la población en edad de trabajar que está en el mercado de trabajo, participa de sus mecanismos o desearía estar ahí más allá de sus circunstancias actuales (ver diagrama 7).

Diagrama 7

Tasa de subutilización en su máxima agregación

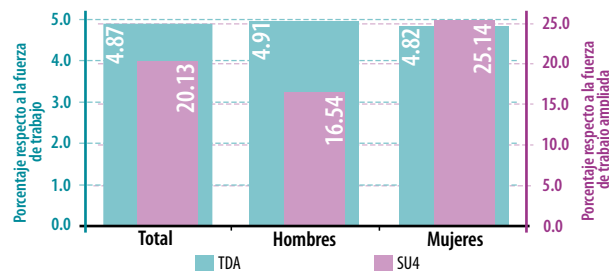


Su efecto diferenciado por sexo (ver gráfica 28) conserva las propiedades de SU3, es decir, registraría porcentajes mayores entre las mujeres que en los hombres, en tanto que por regiones (ver gráfica 29) seguiría siendo mayor en las localidades rurales dispersas, captando hasta más de 37% de la fuerza de trabajo ampliada de las mujeres (ver gráfica 30).

En general, y para toda la población, estaríamos hablando de un indicador que involucra a más de una quinta parte de toda la fuerza de trabajo que participa o tiene interés de participar del mercado de trabajo (fuerza de trabajo ampliada) y, por lo pronto, no hay un referente internacional con qué compararlo.

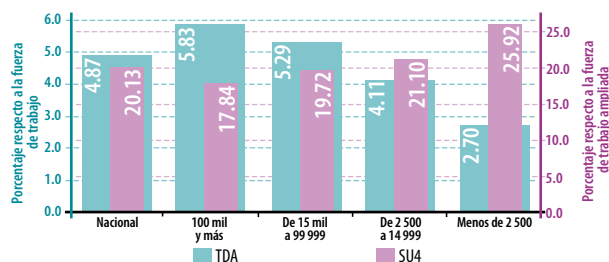
Gráfica 28

Comparativo de la TD y SU4



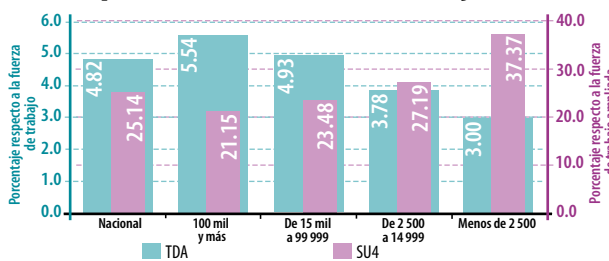
Gráfica 29

Comparativo de la TD y SU4, por tamaño de localidad



Gráfica 30

Comparativo de la TD y SU4, por tamaño de localidad: mujeres



La tentación, por supuesto, sería comparar ese nivel con las TDA que se registran en otras partes del mundo pero, una vez más, no hay que confundir niveles con comportamientos. Un indicador así de agregado como lo muestra la gráfica 31 no dejará de tener menos sensibilidad a las distintas fases del ciclo económico (desaceleración, recesión, recuperación, aceleración).

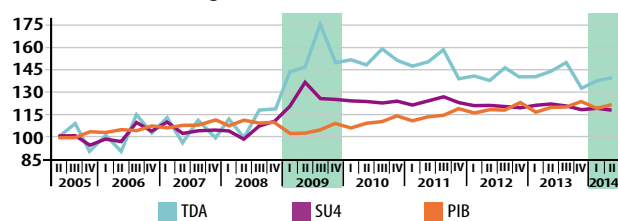
Es así que, no por ser el más agregado, pretende ser el santo grial de los indicadores. Aquí hay *trade-offs* de modo que lo que se gana en un aspecto se pierde en alguna medida en otro. No dejaremos de subrayar que cada una de estas tasas estén al mismo nivel de relevancia y la etiqueta SU facilita impulsar esa idea, de modo que el término que se ha posicionado de forma histórica no desplace a los demás.

Por último, en la comparación entre regiones contrastantes (ver gráfica 32), si bien la SU4 de Chiapas supera a la de Nuevo León, no alcanza a situarse por encima del promedio nacional. Esto es un indicio de que este indicador no puede sustituir a otros ya utilizados de precariedad laboral que sí son sistemáticos en cuanto a su correspondencia con una configuración geográfica del atraso.

Gráfica 31

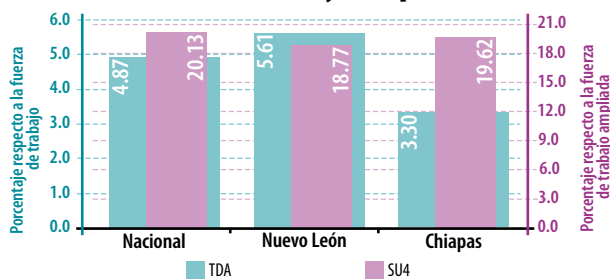
Evolución de la TD, SU4 y PIB: ámbito nacional, 2005-2014

Índices segundo trimestre del 2005 = 100



Gráfica 32

Comparativo de la TD y SU4: ámbito nacional, Nuevo León y Chiapas

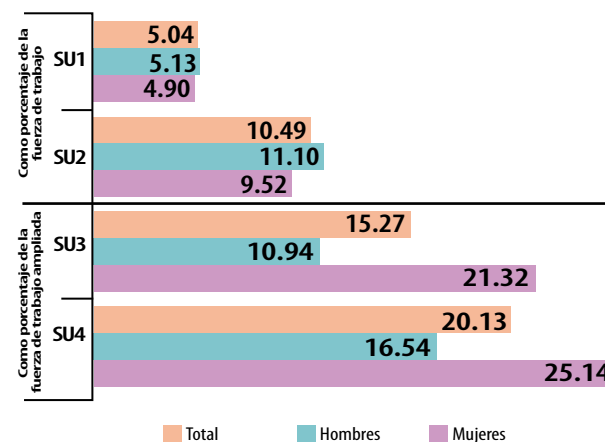


Indicadores en panorámica

Lo que se tiene, entonces, es que los indicadores referidos a la fuerza laboral o de trabajo mantienen su sensibilidad frente al ciclo económico (SU1) por ser casi lo mismo que la tasa de desocupación abierta, a la par que en el caso de SU2 se añade la propiedad de que se registran niveles más altos en las localidades rurales dispersas. Pese a ello, ni uno ni otro indicador visibilizan más a grupos vulnerables como las mujeres, tal como lo hacen SU3 y SU4, que rebasan la frontera de la fuerza de trabajo para recoger situaciones menos obvias (ver gráfica 33). Esta visibilización, sin embargo, no hace que estos indicadores sean idóneos para un análisis de coyuntura como tampoco para calibrar los instrumentos clásicos (política monetaria, cambiaria y fiscal) frente al ciclo económico. El balance general es que la secuencia SU1 a SU4 no pretende ser el recorrido hacia un indicador óptimo sino ser, más bien, una caja de herramientas con la que, dependiendo de lo que se quiere analizar, se haga uso de una de ellas.

Gráfica 33

Resumen de tasas SU, por sexo



Hay, por lo demás, una limitación inherente a estos indicadores que se observa en la secuencia de mapas (1 al 6), cada uno de los cuales muestra cuatro colores (dado que las 32 entidades federativas fueron organizadas en cuartiles) del valor mayor (en rojo) al menor (en verde) del indicador, de modo que cada color corresponde a un grupo de ocho

entidades. El mapa 1 corresponde a la TDA: en él se observa la paradoja del desempleo más de una vez enunciada a lo largo de este ensayo. El cuartil con los valores más bajos incluye entidades económica y socialmente rezagadas del sur y suroeste del

país. En ese grupo sólo desentonan los estados de Chihuahua y San Luis Potosí en el segundo trimestre del 2014. Dado que TDA y SU1 son muy similares, esta última (ver mapa 2) exhibe casi la misma configuración.

Mapa 1

TDA



Mapa 2

SU1



Para el caso de SU2 (ver mapa 3), el patrón regional de los cuartiles parece diluirse mucho más, habiendo pocas zonas de coloración continua, pero lo que llama la atención en el mapa de SU3 (ver mapa 4) es que las entidades federativas del atraso (Chiapas, Oaxaca, Guerrero e, incluso, Michoacán de Ocampo) se presentan todas en colores distintos, es decir, pertenecen a cuartiles diferentes.

Con SU4 (ver mapa 5), esta situación no cambia mucho (a lo más, Michoacán de Ocampo y Oaxaca caen en un mismo cuartil), pero no deja de sorprender que Guerrero esté entre las entidades con el nivel más bajo del indicador.

Ahora bien, esto no sucede con un indicador de precariedad laboral que, por tradición, ha difundido el INEGI desde la ENEU y que se sigue haciendo con

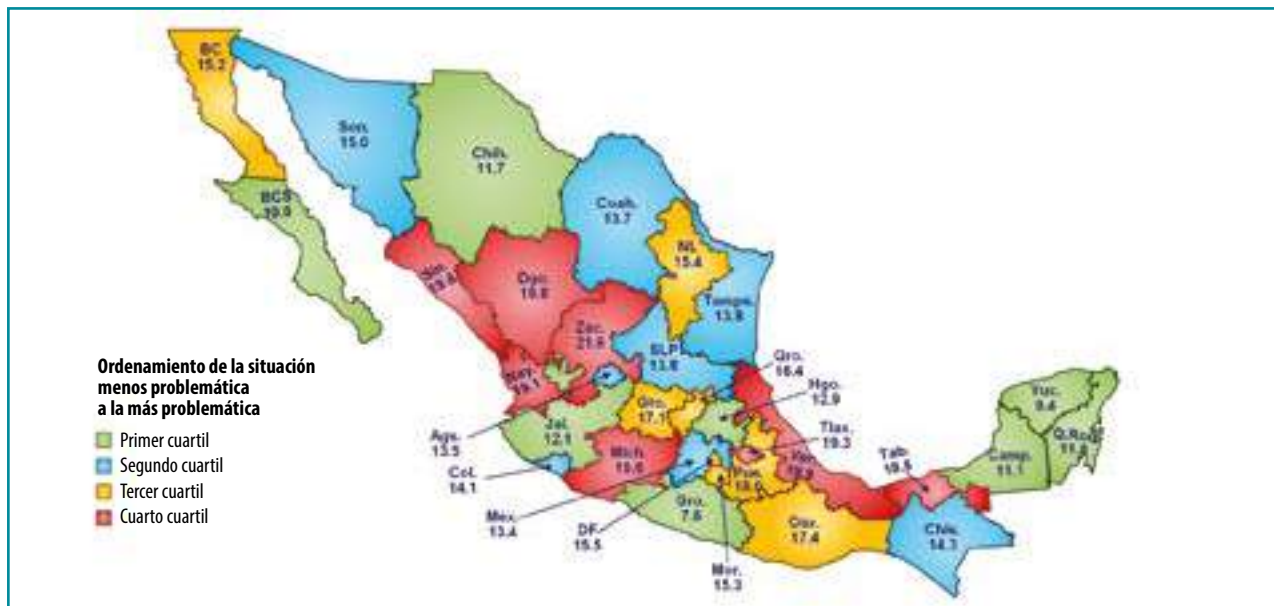
Mapa 3

SU2



Mapa 4

SU3



la ENOE (ver mapa 6), que es la denominada tasa de condiciones críticas de ocupación (TCCO). Este indicador, que alude a una combinación insatisfactoria de tiempo y/o ingresos (reúne en el numerador a quienes trabajan jornadas incompletas; los que trabajan jornadas completas, pero percibiendo menos de un salario mínimo, y los que perciben entre

1 y 2 salarios mínimos laborando 48 horas o más, es decir, con sobrejornada), corresponde de forma inequívoca en el mapa al patrón regional de la pobreza y de la mayor presencia de grupos indígenas en el país. La razón es simple: casi 70% de lo que capta este indicador se refiere a la población vinculada a la agricultura campesina de subsistencia.

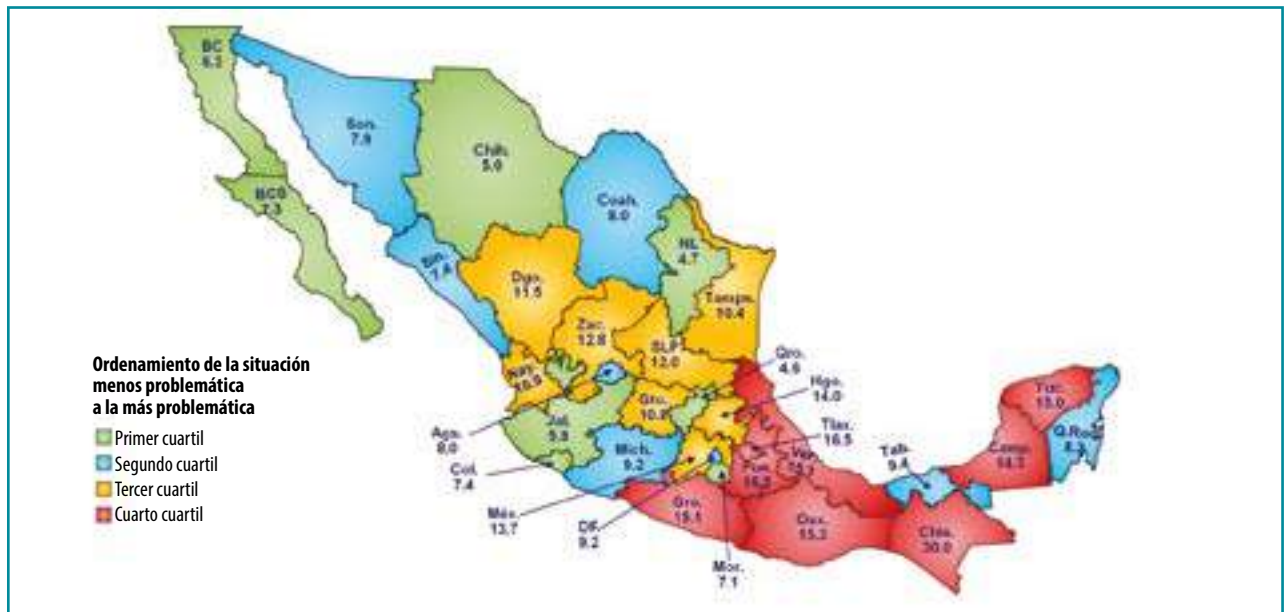
Mapa 5

SU4



Mapa 6

TCCO



La manera de vincular a quienes se dedican a estas actividades a las tasas SU, supone o que deben presentar un comportamiento de búsqueda activo (que es más bien excepcional, como ya se argumentó en la sección dedicada a SU1) o por manifestarse por tomar un trabajo de mercado a la brevedad, tal como interpreta esta noción la OIT. Ninguno de esos caminos son recorridos en estos casos, lo que deja en términos de los indicadores a esta población en una especie de limbo conceptual. Es por ello que indicadores de precariedad, como TCCO o la tasa de informalidad laboral (TIL) que publica la ENOE, son irrenunciables en este contexto, pues ambas integran a esta población en sus agregados. La diferencia con las tasas SU es que ambas (TCCO y TIL) quedan referidas a lo que sería un denominador distinto: toda la población (y no sólo la que se inserta en trabajo de mercado) que participa de la frontera de la producción del Sistema de Cuentas Nacionales.

Visión crítica del concepto *disponibilidad*, impulsado por la OIT en la XIX CIET

En la construcción que se ha seguido hasta ahora de los indicadores SU, en especial de aquellos que involucran población disponible que no ejerce una búsqueda activa de empleo, es importante no pasar por alto el criterio de disponibilidad (disposición para laborar en el mercado de trabajo), según las razones que aduce la población no ocupada para buscar uno. Algunas de éstas dejan en claro que no se busca trabajo porque no es la prioridad del individuo (estudiantes de tiempo completo o también la dedicación a tiempo completo a la crianza de la familia en cierta fase del ciclo de vida del hogar) o simplemente la situación por la que atraviesa (recuperación de una enfermedad o padecimiento crónico e invalidante para laborar), pero hay otras (como ya se mencionó en la sección dedicada a SU3) que nos dicen que el individuo estima que la integración laboral queda fuera de su alcance pese a tener necesidad de ello y es ahí desde donde la ENOE infiere su disponibilidad.

Sin embargo, la idea de disponibilidad que impulsó la OIT en la XIX Conferencia (expresada en las últimas líneas del párrafo 54 de la sección *Fuerza de trabajo potencial*)³⁹ merece un comentario, ya que básicamente apuesta a detectarla en términos de si la persona desearía tomar un trabajo en caso de ofrecérselo uno o en todo caso, se presta a esa interpretación.⁴⁰

Ir por esta vía tiene considerables riesgos. Lo primero sería generar una suerte de efecto Heisenberg en el proceso de la entrevista, es decir, alterar lo que se mide por el acto mismo de la medición, en este caso, al introducir expectativas o inducir deseos. Por lo demás, preguntar si la persona tomaría un trabajo a la brevedad debe enfrentar la posibilidad de que a quien se interroga demande saber de qué trabajo se está hablando, con qué paga, horario, ubicación y hasta qué tipo de tareas implica. Esta necesidad de retroalimentación del informante no se presenta entre los buscadores activos porque el trabajo que han estado buscando es, después de todo, uno acotado por la realidad del mercado laboral que se enfrenta, no así entre aquellos que no se encuentran buscando uno. Malo pues, si la persona requiere de más precisiones del entrevistador; más malo aún si responde sin parpadear teniendo en mente un trabajo del todo fuera de sus posibilidades, como uno calificado cuando la persona no lo está. Es imposible saber qué tiene en mente el informante cuando se le pone ante preguntas de este tenor.

La pregunta, asimismo, resulta fuera de contexto para alguien afectado por un paro técnico en su industria o para el agricultor afectado por el mal tiempo; ellos no están pensando en tomar un

39 "A fin de identificar a las dos categorías de la fuerza de trabajo potencial, así como a los no buscadores dispuestos, las preguntas para determinar si las personas llevaron a cabo 'actividades de búsqueda' y si estaban 'actualmente disponibles' se deberían hacer a todas las personas 'no ocupadas' en el periodo de referencia corto. **La pregunta para determinar si la persona desea un puesto de trabajo** debe hacerse sólo a las personas que no realizaron 'actividades de búsqueda...'. OIT. XIX CIET. Resolución... , op. cit., p. 13, párrafo 54 (lo marcado con negritas es nuestro).

40 Como un ejemplo de pregunta contrafactual que los asesores de la OIT han inducido en otros países está, por ejemplo, la 2a del *Cuestionario individual de la Encuesta Permanente de Hogares (EPH)* que se levanta en Argentina: *Si la semana pasada conseguía un trabajo, ¿podía empezar a trabajar ya?*, más elocuente aún es la 8 del cuestionario de fuerza de trabajo inducida en la Gran Encuesta Integrada de Hogares (GEIH) en Colombia: *¿Desea conseguir un trabajo remunerado o instalar un negocio?*

trabajo, sino retomar el que tienen, por lo que esta filosofía de implementar la exploración de disponibilidad tampoco ha de trasladarse al concepto de subocupación. En estos casos, las preguntas sobre por qué se ha suspendido la actividad (en la medida en la que permitan diferenciar que ello ha sido por un contexto que afecta dicha actividad y no por motivos personales), debieran ser más que suficientes para determinar la condición de disponibilidad sin introducir contrafactuales o ponerse a especular con el informante sobre lo que haría o no.

Aparte de lo anterior, hay una implicación no menor en esto para la logística de una encuesta de fuerza laboral que puede repercutir, incluso, en su diseño estadístico: la mayoría de ellas en el mundo tiene que recurrir —en no pocas ocasiones— a un informante indirecto en la vivienda (un miembro del hogar que contesta por otro) para que no se incremente la tasa de no respuesta, pues localizar a la persona específica del hogar obliga a retornar varias veces, lo que incrementa los costos de operación de manera notable (de hecho, uno de los mayores costos de una encuesta en hogares es poner al entrevistador a la puerta de la vivienda). Los retornos, a su vez, retardan la integración de la muestra y, por ende, la difusión oportuna de lo que se captó en ella. Sin embargo, al introducir un sondeo o una pregunta hipotética sobre si se tomaría un trabajo a la brevedad requiere —más que cualquier otra pregunta de un cuestionario de fuerza de trabajo— que el informante directo se haga cargo de la respuesta; uno indirecto no podría decir, en muchos casos, si la persona de la que se está hablando —misma que no ha estado en un proceso de búsqueda de trabajo— tomaría un empleo de inmediato o no. Una aproximación así se podría plantear en esquemas de levantamiento telefónicos⁴¹ o por internet, pero en países donde no hay una cobertura completa de estos servicios es un reto mayor.

41 Con todo, los países que optan por la entrevista telefónica (Alemania, España y Reino Unido, por citar algunos), necesariamente su primer contacto lo establecen de manera presencial y sólo las ulteriores visitas, por esta vía, así que en el primer contacto o visita tendrían que enfrentar, de todos modos, el tener que hablar con el informante directo. Por lo demás, la vía telefónica ha tenido que enfrentar cada vez más el competir con llamadas de bancos y tarjetas de crédito, así como promociones de todo tipo, por la atención del informante.

Hay, pues, algunos puntos en la *Resolución I* de la XIX CIET en la que lo conceptual no puede dissociarse de las consideraciones de las operaciones de campo y, en general, de implementación de encuestas en hogares.

Conclusiones

La *Resolución I* de la XIX CIET, a partir de una reconsideración de la forma de vincularse con el marco conceptual del SCN, se ha visto obligada a definir el horizonte de lo que es trabajo, al mismo tiempo, de una manera más amplia y más precisa con la finalidad de fundamentar sus distintas modalidades y lo que las distingue entre sí. El trabajo orientado a transacciones de mercado (ocupación por remuneración o beneficio) ya no es más sinónimo de trabajo, como tampoco lo es de insumo laboral para la frontera de la producción del Sistema de Cuentas Nacionales; es sólo un caso específico al respecto. Esto permite visibilizar grupos de población, así como cuantificar la masa de horas trabajadas, ya sea por separado o por combinaciones de actividades. La ENOE, al responder a este marco, ilustra en el camino las posibilidades de visibilizar la aportación del trabajo femenino al proceso de reproducción social de forma integral. En general, las demandas del nuevo marco —más allá de la implementación de módulos o encuestas sobre modalidades de trabajo distintas para crear un subsistema de encuestas del trabajo— no están del todo fuera del alcance de las encuestas en hogares de fuerza laboral, como este ejercicio con información de la ENOE lo demuestra.

Asimismo, es un mérito de la *Resolución I* que su planteamiento de indicadores se hizo bajo la consideración de que (en especial para los grupos vulnerables de la población) su necesidad de inserción en el trabajo remunerado no necesariamente pasa por los mecanismos o engranajes que establecen la interacción entre oferta y demanda de servicios laborales. Es correcto, asimismo, el planteamiento de que un solo indicador de subutilización de fuerza laboral en modo alguno puede decirlo o resumirlo todo. La denominación SU es

acertada para evitar la tentación de que uno solo de esos indicadores sea focal y los demás coreografía. Las tasas SU que se mantienen referidas al concepto de fuerza laboral son pertinentes para el análisis económico, como las que van más allá lo son para el análisis social, visibilizando en específico a la población femenina con necesidad de opciones laborales y, en particular, en localidades rurales y dispersas. Aun así, no se supera del todo la paradoja de que los mayores niveles de los indicadores no se registran de manera sistemática en las zonas de mayor atraso y pobreza. Ello, en buena medida, obedece a que no es fácil reintegrar a los indicadores a la población ligada a la agricultura de subsistencia una vez que en primera instancia ésta ha quedado fuera de la fuerza de trabajo. Por lo anterior, indicadores de precariedad laboral, como la tasa de condiciones críticas de ocupación (TCCO) que difunde la ENOE, o indicadores con una fuerte impronta estructural, como la tasa de informalidad laboral (TIL), deben quedar posicionados al mismo nivel que las TSU.

Si bien el marco conceptual en general es muy pertinente, no lo son ciertos criterios que promueven la identificación de la población con necesidad de inserción laboral (para contabilizarla como fuerza de trabajo potencial) o para laborar más horas (en el caso del subempleo) con preguntas contrafactuales sobre si se hubiera tomado o se tomaría un trabajo ofrecido en un periodo determinado. Tal parece que las implicaciones de ello no son entendidas en su totalidad como tampoco los riesgos que introduce esa filosofía de implementación del criterio de disponibilidad, pues no es lo mismo cuantificar una población en la que, de forma inadvertida, se inducen expectativas o deseos que una en necesidad de trabajar. El ejercicio con la ENOE, dada su profunda exploración de la población fuera de la fuerza laboral, muestra que es posible identificar una que no se manifiesta en los mercados de trabajo, pero con necesidad de inserción a partir de las explicaciones que los individuos dan de no estar en proceso de búsqueda. El objetivo de visibilización se consigue de este modo, mientras que, introducir escenarios hipotéticos o contrafactuales en la entrevista con este mismo propósito,

puede tener costos mayores, además de propiciar una gran incertidumbre en la interpretación de lo que así se obtenga.

Fuentes

- DANE. *Gran Encuesta Integrada de Hogares (GEIH). Formulario de fuerza de trabajo*. Colombia, 2013.
- INDEC. *Encuesta Permanente de Hogares (EPH). Cuestionario individual*. Argentina, 2014.
- INEGI. *Encuesta de Ocupación y Empleo (ENOE). Cuestionario de ocupación y empleo (COE). Base de datos 2005-2014*. México, INEGI, 2014.
- _____. *40 años de encuestas de fuerza laboral en México*. México, INEGI, 2014.
- Hussmanns, R.; F. Mehran y V. Verma. *Employment, Unemployment and Underemployment: an ILO manual on concepts and methods*. Geneva, International Labour Office, 1990.
- Negrete, Rodrigo. "El indicador de la polémica recurrente: la tasa de desocupación y el mercado laboral en México", en: *Realidad, Datos y Espacio. Revista Internacional de Estadística y Geografía*. Vol. 2, núm. 1, enero-abril del 2011.
- _____. "¿Por qué han sido bajas las tasas de desempleo en México?", en: *Notas. Revista de Información y Análisis*. Núms. 14 y 15. México, 2001.
- OIT. *XIX Conferencia Internacional de Estadísticos del Trabajo. Informe de la discusión*. Ginebra, 2-11 de octubre de 2013.
- _____. *Resolución I: resolución sobre las estadísticas del trabajo, la ocupación y la subutilización de la fuerza de trabajo*. Ginebra, 2013.
- _____. *Relevant resolutions and guidelines adopted by the International Conference of Labour Statisticians (ICLS)*. Ginebra, 2013.
- Salsburg, David. *The Lady Tasting Tea: how statistics revolutionized science in the twentieth century*. New York, Freeman, Holt and Company, 2001.

Anexos

Anexo 1

Subjects considered by International Conferences of Labour Statisticians

* Resulted in a resolution which represents the current ILO recommendations on the subject. Subsequently, some issues in a specific resolution may have been the subject of later resolutions (for example, those relating to underemployment and income from employment).

Guidelines endorsed.

Tópicos del trabajo abordados por las CIET

	ICLS	Year
Unemployment statistics	2 nd	1925
Employment and unemployment statistics	6 th	1947
Employment and unemployment statistics	8 th	1954
Measurement of underemployment	9 th	1957
Progress in the measurement of underemployment, particularly in less developed countries	10 th	1962
Measurement of underemployment	11 th	1966
Labour force, employment, unemployment and underemployment*	13 th	1982
Implications of employment promotion schemes on the measurement of employment and unemployment#	14 th	1987
Measurement of underemployment*	16 th	1998
Persons on extended absences from work#	16 th	1998
Measurement of employment in the informal sector		
Employment in the informal sector	14 th	1987
Statistics of employment in the informal sector*	15 th	1993
Child labour statistics		
Methodological considerations	16 th	1998
Measurement of hours of work and absence from work		
Statistics of hours of work*	10 th	1962
Statistics of paid holidays	13 th	1982
Statistics of absence from work	14 th	1987
Concept and measurement of absence from work	15 th	1993
Income from employment, wages and labour cost		
Statistics of wages and hours of labour	1 st	1923
International comparisons of real wages	2 nd	1925
International comparisons of real wages	4 th	1931
Convention concerning statistics of wages and hours of work	5 th	1937
Wages and payroll statistics	7 th	1949
International comparisons of real wages: a study of methods	8 th	1954
Developments relating to Convention 63	8 th	1954
Wages statistics	9 th	1957
Progress in the application of Convention 63	10 th	1962
Statistics of labour cost*	11 th	1966

Tabla 1

Tópicos del trabajo abordados por las CIET

Continúa

	ICLS	Year
Statistics of wages and employee income*	12 th	1973
Revision of the ILO October Inquiry on Occupation Wages	13 th	1982
Income from employment	15 th	1993
Measurement of income from employment*	16 th	1998
Occupational injuries and diseases		
Statistics of industrial accidents	1 st	1923
Statistics of industrial accidents	6 th	1947
Standardization of statistics of industrial injuries and occupational diseases	7 th	1949
Methods of statistics of occupational diseases	8 th	1954
Statistics of employment injuries	9 th	1957
	10 th	1962
Statistics of occupational injuries	13 th	1982
	16 th *	1998
Strikes and lockouts		
Statistics of strikes and lockouts	3 rd	1926
Industrial disputes: statistics of strikes	14 th	1987
Statistics of strikes, lockouts and other forms of industrial action*	15 th	1993
Collective agreements		
Statistics of collective agreements*	3 rd	1926
Labour productivity		
Statistics of labour productivity	7 th	1949
Developments in labour productivity statistics	8 th	1954
Social security		
Social security statistics: development and uses*	9 th	1957
Progress in the field of social security statistics	10 th	1962
Household income and expenditure		
Family budget surveys	3 rd	1926
Methods of family living surveys	7 th	1949
Family budgets and consumer prices	9 th	1957
Scope, method and uses of family expenditure surveys*	12 th	1973
Consumer price indices		
Cost-of-living index numbers	2 nd	1925
	6 nd	1947

Tabla 1

Concluye

Tópicos del trabajo abordados por las CIET

	ICLS	Year
Special problems in the computation of consumer price index numbers	10 th	1962
Consumer price indices*	14 th	1987
Integrated system for labour statistics		
Framework for a labour accounting system	15 th	1993
Classification of industries and occupations		
	1 st	1923
Classification of industries and occupations	2 nd	1925
	3 rd	1926
International Standard Classification of Occupations (ISCO)	7 th	1949
ISCO minor groups	8 th	1954
ISCO major, minor and unit groups	9 th	1957
	11 th	1966
Revision of ISCO	13 th	1982
	14 th	1987
Classification according to status in employment		
International classification according to status in employment (ICSE)	9 th	1957
Revision of ICSE*	15 th	1993
Status in employment (ICSE-93)	16 th	1998
Presentation and dissemination of labour statistics		
International coding of labour statistics	13 th	1982
Dissemination practices for labour statistics#	16 th	1998

Anexo 2

Tabla 2

Resumen de tasas tradicionales y nuevas, por entidad federativa

Entidad federativa	Tasas tradicionales				Tasas propuestas			
	TDA	TS	TCCO	SU1	SU2	SU3	SU4	TS insuf. tiempo
Estados Unidos Mexicanos	4.87	8.16	11.42	5.04	10.49	15.27	20.13	5.74
Aguascalientes	5.82	2.83	8.04	5.84	7.74	13.49	15.24	2.02
Baja California	5.64	4.23	6.32	5.63	8.34	15.15	17.59	2.87
Baja California Sur	4.93	15.20	7.26	4.89	13.25	10.03	17.94	8.79
Campeche	3.19	6.12	14.26	3.33	7.46	11.13	14.93	4.28
Coahuila de Zaragoza	5.23	8.65	8.01	5.21	10.04	13.68	18.08	5.09
Colima	4.40	8.08	7.42	4.24	10.12	14.06	19.34	6.14
Chiapas	3.30	8.69	29.97	3.73	9.69	14.32	19.62	6.19
Chihuahua	3.18	1.80	4.97	3.19	4.58	11.67	12.93	1.43
Distrito Federal	6.83	6.88	9.20	6.79	10.38	15.53	18.78	3.85
Durango	6.00	7.23	11.52	6.15	10.85	19.81	23.82	5.00
Guanajuato	4.73	14.54	10.82	4.80	13.61	17.05	24.73	9.26
Guerrero	1.53	6.54	15.07	1.86	6.98	7.64	12.45	5.21
Hidalgo	4.08	11.68	13.98	4.35	13.77	12.89	21.47	9.85
Jalisco	4.98	8.65	5.80	5.03	11.00	12.10	17.62	6.28
México	6.56	6.30	13.74	6.65	10.54	13.38	17.00	4.18
Michoacán de Ocampo	3.19	8.90	9.20	3.40	10.52	19.56	25.49	7.36
Morelos	4.14	5.29	7.08	4.12	8.11	15.29	18.81	4.15
Nayarit	5.76	15.60	10.86	5.76	14.88	19.06	26.89	9.68
Nuevo León	5.61	6.67	4.67	5.60	9.39	15.38	18.77	4.01
Oaxaca	2.58	9.93	15.32	3.12	10.86	17.37	23.97	7.99
Puebla	4.03	5.35	16.50	4.35	8.37	17.98	21.42	4.20
Querétaro	5.92	2.24	4.57	6.13	7.29	16.44	17.47	1.23
Quintana Roo	4.52	8.59	8.30	4.69	10.19	11.79	16.88	5.77
San Luis Potosí	3.18	6.26	12.05	3.38	8.06	13.83	18.01	4.85
Sinaloa	5.79	12.63	7.43	5.80	14.49	19.58	27.00	9.22
Sonora	4.95	13.08	7.88	4.94	12.84	14.96	22.03	8.31
Tabasco	6.54	9.70	9.36	6.62	13.29	18.55	24.36	7.14
Tamaulipas	5.23	15.69	10.36	5.19	15.03	13.81	22.76	10.38
Tlaxcala	5.76	22.38	16.51	5.78	21.55	19.31	32.82	16.74
Veracruz de Ignacio de la Llave	3.66	4.88	15.13	3.84	8.40	19.93	23.73	4.75
Yucatán	2.37	7.81	14.95	2.41	8.55	9.43	15.13	6.30
Zacatecas	4.73	14.46	12.83	4.81	15.54	21.64	30.48	11.27

Tabla 3

Resumen de tasas tradicionales y nuevas, serie 2005-2014

Periodo		Tasas tradicionales			Tasas propuestas				
		TDA	TS	TCCO	SU1	SU2	SU3	SU4	TS insuf: tiempo
2005	I	3.85	8.85	15.04	3.91	9.52	12.82	17.91	5.84
	II	3.48	7.49	14.64	3.56	8.12	13.01	17.13	4.73
	III	3.79	7.24	14.08	3.89	8.18	13.31	17.18	4.46
	IV	3.13	6.34	13.31	3.21	7.02	12.67	16.11	3.94
2006	I	3.52	6.18	13.12	3.61	7.55	13.23	16.78	4.09
	II	3.14	6.01	11.90	3.19	7.17	12.91	16.48	4.11
	III	4.00	8.21	12.08	4.10	9.04	14.28	18.69	5.15
	IV	3.59	7.05	11.82	3.68	8.22	13.65	17.72	4.72
2007	I	3.93	7.75	11.54	4.01	9.01	14.34	18.80	5.21
	II	3.33	6.95	10.94	3.42	7.87	13.46	17.45	4.61
	III	3.84	6.94	11.31	3.96	8.22	13.93	17.75	4.43
	IV	3.46	7.16	10.89	3.55	8.22	13.74	17.91	4.84
2008	I	3.87	6.65	10.45	3.94	8.34	13.86	17.81	4.59
	II	3.43	6.90	10.67	3.52	7.84	12.94	16.83	4.47
	III	4.12	6.80	10.51	4.22	8.52	14.39	18.23	4.49
	IV	4.16	7.17	10.57	4.29	8.87	14.83	18.90	4.78
2009	I	4.99	7.99	12.21	5.08	10.16	15.92	20.42	5.35
	II	5.10	11.12	12.88	5.22	13.13	16.35	23.33	8.35
	III	6.14	8.77	12.40	6.31	11.93	16.52	21.53	6.00
	IV	5.21	8.81	12.07	5.38	11.10	16.36	21.42	6.04
2010	I	5.28	9.38	12.33	5.40	11.29	16.06	21.28	6.22
	II	5.16	8.93	12.01	5.30	11.24	15.86	21.13	6.26
	III	5.55	8.58	11.60	5.73	11.27	16.05	20.99	5.88
	IV	5.28	7.59	11.96	5.43	10.33	16.91	21.21	5.18
2011	I	5.14	8.15	11.97	5.24	10.37	16.20	20.74	5.41
	II	5.23	8.30	11.10	5.36	10.78	16.43	21.22	5.73
	III	5.53	8.90	11.17	5.69	11.48	16.57	21.69	6.14
	IV	4.84	8.92	10.95	4.95	10.74	15.95	21.07	6.09
2012	I	4.91	8.56	12.15	5.03	10.46	15.85	20.66	5.72
	II	4.80	8.87	11.63	4.92	10.71	15.54	20.68	6.09
	III	5.11	8.70	11.90	5.24	10.81	15.65	20.60	5.88
	IV	4.86	7.99	11.64	5.01	10.09	15.97	20.47	5.35
2013	I	4.89	8.19	11.80	5.04	10.32	16.04	20.71	5.56
	II	5.00	8.53	12.03	5.13	10.67	15.97	20.88	5.84
	III	5.22	8.49	12.16	5.41	10.76	15.95	20.70	5.65
	IV	4.62	8.17	11.92	4.77	10.07	15.49	20.20	5.57
2014	I	4.80	8.34	11.70	4.94	10.49	15.38	20.32	5.84
	II	4.87	8.16	11.42	5.04	10.49	15.27	20.13	5.74

Tabla 4

Tasas calculadas contra el total de población de 15 años o más de edad, por entidad federativa

Entidad federativa	Tasas propuestas				
	TD	SU1	SU2	SU3	SU4
Estados Unidos Mexicanos	2.93	2.91	6.05	9.87	13.02
Aguascalientes	3.44	3.43	4.55	8.64	9.76
Baja California	3.53	3.51	5.20	10.52	12.21
Baja California Sur	3.29	3.25	8.80	7.04	12.59
Campeche	1.98	1.96	4.41	7.15	9.59
Coahuila de Zaragoza	3.25	3.22	6.19	9.27	12.24
Colima	2.98	2.83	6.76	10.46	14.39
Chiapas	1.87	1.87	4.85	8.06	11.05
Chihuahua	1.86	1.85	2.66	7.42	8.23
Distrito Federal	4.22	4.18	6.38	10.53	12.73
Durango	3.46	3.44	6.06	12.96	15.59
Guanajuato	2.88	2.86	8.11	11.66	16.92
Guerrero	0.91	0.85	3.18	3.70	6.03
Hidalgo	2.45	2.43	7.69	7.90	13.16
Jalisco	3.08	3.07	6.71	7.98	11.62
México	3.94	3.94	6.25	8.55	10.86
Michoacán de Ocampo	1.91	1.91	5.91	13.21	17.21
Morelos	2.41	2.39	4.69	10.02	12.32
Nayarit	3.69	3.61	9.32	13.90	19.62
Nuevo León	3.46	3.44	5.76	10.53	12.85
Oaxaca	1.53	1.47	5.13	9.61	13.27
Puebla	2.46	2.43	4.67	11.71	13.95
Querétaro	3.21	3.21	3.81	9.66	10.27
Quintana Roo	3.07	3.04	6.61	8.27	11.84
San Luis Potosí	1.86	1.86	4.44	8.53	11.11
Sinaloa	3.46	3.42	8.55	13.54	18.67
Sonora	3.20	3.19	8.28	10.79	15.88
Tabasco	3.82	3.80	7.62	12.19	16.02
Tamaulipas	3.20	3.15	9.12	9.22	15.19
Tlaxcala	3.56	3.43	12.81	13.40	22.78
Veracruz de Ignacio de la Llave	1.95	1.95	4.27	12.16	14.48
Yucatán	1.50	1.47	5.20	6.18	9.91
Zacatecas	2.63	2.53	8.18	13.84	19.49

Tabla 5

Tasas calculadas contra el total de población de 15 años o más de edad, serie 2005-2014

Periodo	Tasas propuestas					
		TD	SU1	SU2	SU3	SU4
2005	I	2.27	2.23	5.43	8.05	11.25
	II	2.05	2.02	4.62	8.20	10.79
	III	2.26	2.22	4.68	8.44	10.89
	IV	1.87	1.85	4.04	8.08	10.28
2006	I	2.11	2.09	4.37	8.51	10.79
	II	1.88	1.85	4.15	8.29	10.59
	III	2.42	2.39	5.27	9.31	12.18
	IV	2.18	2.14	4.80	8.87	11.52
2007	I	2.36	2.32	5.21	9.29	12.19
	II	2.00	1.97	4.55	8.68	11.25
	III	2.30	2.28	4.72	8.93	11.38
	IV	2.11	2.08	4.82	9.00	11.74
2008	I	2.33	2.29	4.85	8.98	11.54
	II	2.08	2.05	4.57	8.37	10.89
	III	2.47	2.44	4.92	9.31	11.79
	IV	2.49	2.46	5.08	9.54	12.16
2009	I	2.96	2.92	5.84	10.34	13.26
	II	3.04	3.00	7.54	10.63	15.17
	III	3.73	3.68	6.96	10.81	14.09
	IV	3.17	3.14	6.47	10.79	14.12
2010	I	3.14	3.11	6.50	10.41	13.80
	II	3.11	3.08	6.53	10.37	13.82
	III	3.35	3.31	6.51	10.42	13.62
	IV	3.12	3.08	5.85	10.91	13.68
2011	I	3.03	2.98	5.91	10.43	13.36
	II	3.12	3.09	6.22	10.73	13.86
	III	3.33	3.29	6.64	10.84	14.19
	IV	2.95	2.89	6.27	10.53	13.91
2012	I	2.93	2.90	6.04	10.33	13.47
	II	2.93	2.89	6.29	10.27	13.68
	III	3.12	3.07	6.34	10.30	13.57
	IV	2.93	2.89	5.81	10.40	13.33
2013	I	2.91	2.89	5.92	10.40	13.43
	II	3.03	2.99	6.22	10.50	13.73
	III	3.17	3.14	6.23	10.40	13.50
	IV	2.82	2.79	5.89	10.20	13.30
2014	I	2.87	2.85	6.05	9.97	13.17
	II	2.93	2.91	6.05	9.87	13.02

Colaboran en este número

Manuel Mendoza Ramírez

Es actuario, maestro en Estadística e Investigación de Operaciones y doctor en Ciencias, todos los grados obtenidos en la Universidad Nacional Autónoma de México (UNAM). Hoy en día, es profesor de tiempo completo del Departamento de Estadística del Instituto Tecnológico Autónomo de México, donde ha sido director del Programa de Maestría en Administración de Riesgos y jefe del Departamento. Ha sido presidente de la Asociación Mexicana de Estadística y, entre otras distinciones, es miembro electo de la Academia Mexicana de Ciencias y del *International Statistical Institute*. Es, también, miembro del Sistema Nacional de Investigadores. Su área de especialidad es la Estadística Bayesiana.

Contacto: mendoza@itam.mx

Alberto Contreras Cristán

Es actuario por la UNAM y doctor en Estadística por el *Imperial College*, Londres. Actualmente, se desempeña como investigador en el Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas IIMAS-UNAM. Es miembro del Sistema Nacional de Investigadores. Sus áreas de interés son la inferencia estadística, las series de tiempo y la Estadística Bayesiana.

Contacto: alberto@sigma.iimas.unam.mx

Eduardo Gutiérrez Peña

Es maestro en Estadística por la UNAM y doctor en Estadística por el *Imperial College*, Londres. En la actualidad, es investigador titular en el Departamento de Probabilidad y Estadística del IIMAS-UNAM. Ha sido jefe del Departamento de Probabilidad y Estadística del IIMAS-UNAM, presidente de la Asociación Mexicana de Estadística y miembro de la mesa directiva de la *International Society for Bayesian Analysis*. Algunas de las distinciones académicas que tiene son: ganador del premio *Jan Tinbergen Award*, otorgado por el *International Statistical Institute* y miembro electo de ese mismo instituto. Pertenece al Sistema Nacional de Investigadores. Sus áreas de especialización son Estadística y Probabilidad con énfasis en inferencia bayesiana.

Contacto: eduardo@sigma.iimas.unam.mx

José Paúl Carrasco Escobar

Es candidato a doctor en Matemática Educativa en el Instituto Politécnico Nacional. Es egresado de la Maestría en Enseñanza de las Matemáticas con la salida terminal en Educación Matemática de la Universidad Autónoma del Estado de Morelos y licenciado en Matemáticas Aplicadas y Computación por la Universidad Nacional Autónoma de México. En la actualidad se desempeña como jefe del Departamento de Diagnóstico en la Dirección General de Estadísticas Sociodemográficas del INEGI y como docente del área de Matemática en diversos niveles educativos.

Contacto: paul.carrasco@inegi.org.mx

José Luis Ángel Rodríguez Silva

Es doctor en Planeación y Liderazgo Educativo por la Universidad del Valle de México; candidato a doctor tanto en Ciencia y Tecnología, con orientación en Matemáticas Aplicadas de la Universidad de Guadalajara, como en Matemática Educativa en el Instituto Politécnico Nacional; es egresado de las maestrías en Ciencias en Estadística Oficial del Centro de Investigación en Matemáticas, en Ciencias en Ingeniería Industrial del Instituto Tecnológico de Aguascalientes y en Finanzas y Negocios de la Universidad Autónoma de Aguascalientes; además, es máster universitario en Ética para la Construcción Social de la Universidad de Deusto, España, y cuenta con una Especialidad en Estadística de la Universidad de Guanajuato; la Licenciatura en Ciencias Físico-Matemáticas la obtuvo en la Universidad Michoacana de San Nicolás de Hidalgo. Actualmente se desempeña como subdirector de Acopio de Información de la Dirección General de Geografía y Medio Ambiente del INEGI y como docente de cátedra del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM), Campus Aguascalientes.

Contacto: angel.rodriguez@inegi.org.mx

Wilfrido Ruiz Ochoa

Es licenciado en Economía por la Universidad Autónoma de Baja California (UABC); obtuvo la Maestría en Economía Industrial en El Colegio de la Frontera Norte (Colef), así como el Diplomado en Economía Teórica y el Doctorado en Desarrollo Económico Cuantitativo en la Universidad Autónoma de Madrid, España. Actualmente, es profesor-investigador del Departamento de Estudios Económicos del Colef (desde 1994) y miembro del Sistema Nacional de Investigadores, nivel I, y funge como coordinador del Observatorio de la Coyuntura Económica Fronteriza (OCEF), con un consejo consultivo formado por profesores investigadores del Colef, ITESM, UNAM, Universidad Autónoma de Coahuila y UABC, entre otras instituciones. Forma parte de los consejos editoriales de las revistas *Frontera Norte* (del Colef) y *Economía y Sociedad* (de la Universidad Michoacana). Ha publicado cerca de 30 artículos en revistas arbitradas; entre sus campos de especialidad está el tema de las desigualdades regionales.

Contacto: wruiz@colef.mx

José Luis Silván Cárdenas

Es doctor en Ciencias de Información Geográfica por la Universidad Estatal de Texas (2009), así como ingeniero en Computación (1998) y maestro en Ingeniería Eléctrica (2002) por la Universidad Nacional Autónoma de México. Realizó una estancia postdoctoral en la Universidad Estatal de Búfalo, donde coordinó un proyecto sobre estimación poblacional. Del 2002 a la fecha ha estado adscrito como investigador titular C del Centro de Investigación en Geografía y Geomática "Ing. Jorge L. Tamayo", AC (CENTROGEO), donde actualmente coordina el Programa de Posgrado. Durante su carrera, ha recibido varios reconocimientos, destacando la medalla Alfonso Caso (UNAM, 2002) y el *J. Warren Nystrom Award* (AAG, 2011). Es miembro del Sistema Nacional de Investigadores (SNI), nivel 1, desde el 2011 y coautor en varias publicaciones sobre los temas de extracción de información geográfica mediante imágenes ópticas y LIDAR.

Contacto: jlsilvan@centrogeo.org.mx

Jorge Alberto Montejano Escamilla

Es doctor en Urbanismo por la Universitat Politècnica de Catalunya (Tesis Excelente Cum-Laude) y arquitecto por la Universidad Iberoamericana. En la actualidad, es profesor-investigador titular B en el CENTROGEO. Es miembro del SNI en calidad de candidato. Sus líneas actuales de investigación son metropolización territorial; crecimiento urbano; movilidad; espacio público; lógicas y modelos de localización, vivienda social y sistemas de información geográfica aplicados a la planeación y a los estudios urbanos.

Contacto: jmontejano@centrogeo.edu.mx

Mauricio Pablo Cervantes Salas

Es biólogo por la Facultad de Ciencias de la UNAM (1993), maestro en Demografía y doctor en Estudios de Población por El Colegio de México (2002-2010). Del 2013 a la fecha, ha estado adscrito como investigador del CENTROGEO, donde actualmente es investigador asociado B. Ha enfocado su investigación en estudios socioambientales y sobre esa línea de investigación es que ha desarrollado sus últimas publicaciones.

Contacto: mcervantes@centrogeo.org.mx

José Antonio Mejía Guerra

Es maestro en Políticas Públicas por la Universidad de Georgetown y en Economía por la Universidad George Washington; cuenta con las licenciaturas en Derecho por la Universidad Autónoma de Nuevo León y en Economía por el Instituto Tecnológico y de Estudios Superiores de Monterrey. Del 2008 al 2012 se desempeñó como vicepresidente de la Junta de Gobierno del Instituto Nacional de Estadística y Geografía. Antes, laboró durante 13 años en el Banco Interamericano de Desarrollo (BID) en Washington, D.C. Actualmente, trabaja como Especialista Líder en Modernización del Estado en el BID.

Contacto: jose.antonio.mejia@gmail.com

Rodrigo Negrete Prieto

Es economista egresado de la UNAM. En la actualidad se desempeña como investigador especialista del INEGI en encuestas en hogares, indicadores de mercado laboral y el marco conceptual para la medición estadística del trabajo. Ha sido miembro de distintos grupos de trabajo internacionales que abordan la temática y publicado diversos artículos en revistas especializadas, así como también en *Fractal*, *Nexus* y *Este País* sobre temas de interés general.

Contacto: rodrigo.negrete@inegi.org.mx

Tomás Ramírez Reynoso

Es economista por la Facultad de Economía de la UNAM. Se especializó en el INEGI en los procesos de validación de encuestas y generación de indicadores de empleo, así como en los de estadísticas vitales y sociodemográficas. Su experiencia como generador la amplió a la de usuario en los estudios de población emprendidos en el Consejo Estatal de Población del Estado de Aguascalientes entre 1996 y el 2009; es autor de diversos artículos y publicaciones. En la actualidad, es director del área de encuestas en hogares de contenido socioeconómico en el INEGI, entre las que destacan la Encuesta Nacional de Ocupación y Empleo (ENOE), así como la Encuesta Nacional de Ingresos y Gastos de los Hogares (ENIGH).

Contacto: tomas.ramirez@inegi.org.mx

**Lineamientos para publicar en
REALIDAD, DATOS Y ESPACIO.
REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA**

Los trabajos presentados a REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA deberán tratar temas de interés relativos a la situación actual de la información estadística y geográfica.

Sólo se reciben para su posible publicación trabajos inéditos, en español o inglés. Por ello, es necesario anejar una carta dirigida al editor de REALIDAD, DATOS Y ESPACIO. REVISTA INTERNACIONAL DE ESTADÍSTICA Y GEOGRAFÍA, en la que se proponga el artículo para su publicación y se declare que es inédito y que no se publicará en otro medio. En esta carta deben incluirse los datos completos del autor o autores, institución, domicilio completo, correo electrónico y teléfono. El envío de los artículos debe dirigirse a la atención de la M. en C. Virginia Abrín Batule, virginia.abrin@inegi.org.mx (tel. 5278 10 00, ext. 1161).

Los trabajos se tienen que presentar en versión electrónica (formato *Word* o compatible), en la cual se incluyan las imágenes, gráficas y cuadros (en el formato de los programas con que fueron generados y en archivos independientes, tales como Adobe Photoshop, Adobe Illustrator, TIF, EPS, PNG o JPG, con una resolución de 300 dpi y en un tamaño de 13 x 8 cm). Las expresiones y/o algoritmos, enviarlas con el formato anterior. Se sugiere una extensión de 15 cuartillas, tipo de letra Helvética, Arial o Times de 12 puntos e interlineado de 1.5 líneas.

Los artículos deben incluir: título del trabajo, nombre completo del autor o autores, institución donde trabaja y cargo que ocupa, teléfonos, correo electrónico, breve semblanza del autor o autores (que no exceda de un párrafo de cinco renglones), 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) y bibliografía u otras fuentes.

Las referencias bibliográficas deberán presentarse al final del artículo de la siguiente manera: nombre del autor comenzando por el o los apellidos; título del artículo (entrecomillado); título de la revista o libro donde apareció publicado (en cursivas); editor o editorial; lugar y año de edición. En el caso de las fuentes electrónicas (páginas *Web*) se seguirá el mismo orden que en las bibliográficas, pero al final entre paréntesis se pondrá DE (dirección electrónica), la fecha de consulta y la liga completa. Omitir las que se mencionen como notas a pie de página.

Todos los artículos recibidos serán sometidos a evaluación y el proceso de dictaminación será de acuerdo con la metodología de doble ciego (autores y dictaminadores anónimos).

**GUIDELINES FOR PUBLISHING IN
REALITY, DATA AND SPACE.
INTERNATIONAL JOURNAL OF STATISTICS AND GEOGRAPHY**

The papers submitted to Reality, Data and Space. International Journal of Statistics and Geography, must deal with issues of interest relating to state-of-the-art statistical and geographical information.

Only unpublished works, in English or Spanish will be accepted for possible publication. Therefore, it is required to attach a letter addressed to the Publisher of Reality, Data and Space. International Journal of Statistics and Geography, proposing the article for publication and stating it is unpublished material and it will not be published in any other way. The letter must include the full details of the author or authors, institution, full address, e-mail and telephone number. The dispatch of the articles should be directed to the attention of the M. C. Virginia Abrín Batule, virginia.abrin@inegi.org.mx (tel. 5278 10 00, ext. 1161).

Contributions must be submitted in electronic format (Word format or compatible), containing the images, charts and tables (in the original format of the software they were created on, and in separate files, such as Adobe Photoshop, Adobe Illustrator, TIF, EPS, PNG or JPG, with a resolution of 300 dpi and a 13 x 8 cm size of). The equations and or the algorithm send it in the same form. An extension of 15 pages, Helvetica, Arial or Times 12 points typeface, and a spacing of 1.5 lines is suggested.

The articles should include: title, full name of the author or authors, institution where he/she works and her/his position, phone, e-mail, a brief biography of the author or authors (not exceeding a 5 lines paragraph), summaries of the work, in English and Spanish (not exceeding a 10 lines paragraph), keywords, in English and Spanish (minimum 3, maximum 5) and bibliography reference list.

Bibliographical references must appear at the end of the article as follows: Author's name beginning with the surname; article's Title (in quotation marks); Title of the magazine or book where it was published (in italics); Publisher or editorial; house and year of the edition. In the case of electronic sources (Web pages) it will be used the same arrangement as for bibliographical references, but it will be followed by the mention DE (dirección electrónica, in Spanish) between brackets, the date of consultation and the full link.

All contributions received will be subject to evaluation and the approval process will be carried according to the methodology of double-anonymity (anonymous authors and adjudicators).

Utiliza el Directorio Estadístico Nacional de Unidades Económicas

DENUE
INTERACTIVO
01/2015



Ahora puedes:

- Consultar la información actualizada de poco más de 4.9 millones de negocios activos en el país.
- Descargar de manera gratuita todo el *Directorio* o sólo la información seleccionada de acuerdo con tu interés.
- Descargar la aplicación para dispositivos móviles (*Android* e *IOS*).
- Consultarlo en el *Mapa Digital de México (MDM)* para asociar la información de los negocios a otros datos.

Conociendo México

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

 INEGI Informa

 @INEGI_INFORMA



**INSTITUTO NACIONAL
DE ESTADÍSTICA Y GEOGRAFÍA**

Resultados oportunos

Los **Censos Económicos 2014** captaron un total de **5.7 millones** de establecimientos en el país y **29.9 millones** de personas ocupadas en ellos.

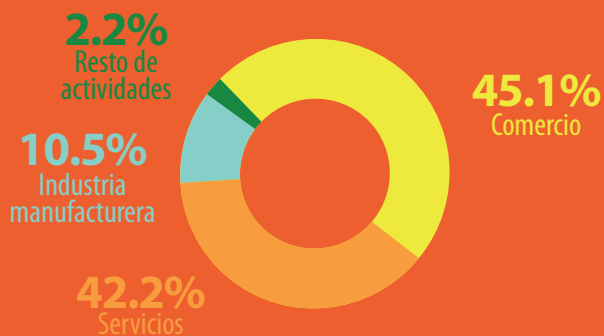


El **Distrito Federal** y el **estado de México** concentran los **mayores** porcentajes de **personal ocupado**:

14.5% el Distrito Federal y **9.3%** el estado de México.

La **mitad** de las **personas ocupadas** **trabajan** en sectores de **servicios**.

De cada **100 establecimientos** en el país, **45** se dedican al **comercio** y **42**, a los **servicios**.



El **43%** de los **establecimientos** tiene una **edad de cinco años o menos**.

La **longevidad** de los negocios está relacionada con el **tamaño** de los mismos:

a mayor número de personas ocupadas, más posibilidades de alcanzar una mayor edad.

En nuestra economía predominan los **negocios pequeños**: **94 de cada 100 establecimientos** tienen como máximo **10 personas ocupadas**.



Fuente: INEGI, en: www.inegi.org.mx.

Conociendo México

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

 INEGI Informa

 @INEGI_INFORMA

