

Deterioro de la capacidad de síntesis del Índice de Marginación: una propuesta de índices complementarios

ÓSCAR PELÁEZ HERREROS*

RESUMEN

En este artículo se comprueba que el Índice de Marginación (IM) retiene un porcentaje cada vez menor de la información que trata de resumir. A nivel de municipios, esta cifra ha pasado de 62%, en el año 1990, a 49%, en 2015. La causa es que las variables del IM han ido configurando dos procesos diferenciados: uno vinculado a las carencias en la urbanización y, otro, a las carencias en los servicios de las viviendas. Se propone el cálculo de dos índices complementarios: el de carencias en urbanización y el de carencias en vivienda. Estos índices arrojan resultados coherentes con los de otras investigaciones y permiten identificar de manera específica los municipios y entidad que padecen mayores rezagos en cada aspecto.

Palabras clave: Marginación, carencias en urbanización, carencias en vivienda, análisis factorial.

Clasificación JEL: C43, I31, R11.

* Departamento de Estudios Económicos, El Colegio de la Frontera Norte, sede Tijuana, México. Correo electrónico: opelaez@colef.mx.

ABSTRACT

Decrease in the synthesis capacity of the Marginalization Index: a proposal of complementary indexes

In this paper we verify that the Marginalization Index (IM) retains an increasingly smaller percentage of the information that it attempts to summarize. At municipal level, this figure has fallen from 62% in 1990 to 49% in 2015. The reason is that the variables of the IM have been generating two differentiated processes: one related to the deficiencies in urbanization, and another to the deficiencies in housing services. We propose the calculation of two complementary indexes: the urbanization deficiencies index and the housing deficiencies index. These indexes show consistent results with other investigations and allow the identification of the municipalities and entities that suffer greater deficiencies in each aspect.

Keywords: Marginalization, deficiencies in urbanization, deficiencies in housing, factor analysis.

JEL Classification: C43, I31, R11.

INTRODUCCIÓN

El Índice de Marginación (IM) es una medida sintética de las carencias que padece la población que reside en un determinado territorio. Fue diseñado en la década de los noventa por el Consejo Nacional de Población (Conapo) y calculado para las entidades federativas y municipios de México con datos del XI *Censo General de Población y Vivienda de 1990* (Conapo, 1994). La noción de marginación que respalda la construcción del IM introducida por Coplamar, buscaba “caracterizar aquellos grupos que han quedado al margen de los beneficios del desarrollo nacional y de los beneficios de la riqueza generada, pero no necesariamente al margen de la generación de esa riqueza ni mucho menos de las condiciones que la hacen posible” (1982: 22). Si bien el concepto de “marginación” guarda alguna relación con el de “marginalidad” (Germani, 1973), ambos no deben confundirse, como advierte Cortés (2002 y 2006).

En este tenor, a principios de los noventa, el Programa de las Naciones Unidas para el Desarrollo (PNUD) dio inicio al cálculo y publicación anual del Índice de Desarrollo Humano (IDH) para los

distintos países del mundo (1990). Con fundamento en el enfoque de capacidades de Sen (1983), la popularización del IDH reforzó el cambio de paradigma que llevó a considerar como bienestar no sólo el crecimiento económico, sino también la calidad y las condiciones de vida de las personas, atendiendo a variables que reflejan las necesidades básicas insatisfechas, que soportan las poblaciones y limitan el desarrollo de las capacidades de sus integrantes.

En este contexto, el IM volvió a calcularse con datos de los censos de 2000 y 2010, del conteo de 2005 y de la encuesta intercensal de 2015; se aplicó no sólo a las entidades federativas y municipios, sino también a las localidades y a las áreas geoestadísticas básicas urbanas del país, y se convirtió en un referente para la declaración de las zonas de atención prioritaria (DOF, 2016) junto con el Índice de Rezago Social (IRS) que calcula el Consejo Nacional de Evaluación de la Política de Desarrollo Social (Coneval), mediante una técnica semejante a la del IM.

Debido a esto, tanto el IRS como el IM se expresan como la suma ponderada de las variables estandarizadas que componen cada índice.¹ El IM para el nivel de desagregación territorial de entidades federativas y municipios integra la información de nueve variables observables.² El IRS incluye 11 variables.³ A diferencia del IDH, que ha sido criticado por utilizar ponderaciones fijas iguales a un tercio para cada una de sus tres dimensiones (salud, educación e ingreso), los pesos del IM y el IRS son los correspondientes al primer componente principal de la matriz de

¹ Las variables se estandarizan restando la media a cada observación y dividiendo ese resultado por la desviación estándar.

² Las nueve variables que forman parte del IM a nivel de entidades federativas y municipios son los porcentajes de: 1) población analfabeta de 15 años o más, 2) población sin primaria completa de 15 años, 3) ocupantes en viviendas sin drenaje ni servicio sanitario, 4) ocupantes en viviendas sin energía eléctrica, 5) ocupantes en viviendas sin agua entubada, 6) viviendas con algún nivel de hacinamiento, 7) ocupantes en viviendas con piso de tierra, 8) población en localidades con menos de 5 mil habitantes, y 9) población ocupada con ingresos de hasta 2 salarios mínimos (Téllez *et al.*, 2016: 12).

³ Las once variables del IRS a nivel de entidades federativas y municipios son los porcentajes de: 1) población de 15 años o más analfabeta, 2) población de 6 a 14 años que no asiste a la escuela, 3) población de 15 años y más, con educación básica incompleta, 4) población sin derechohabencia a servicios de salud, 5) viviendas con piso de tierra, 6) viviendas que no disponen de excusado o sanitario, 7) viviendas que no disponen de agua entubada de la red pública, 8) viviendas que no disponen de drenaje, 9) viviendas que no disponen de energía eléctrica, 10) viviendas que no disponen de lavadora, y 11) viviendas que no disponen de refrigerador (Coneval, 2016).

datos, siendo óptimos en el sentido de que maximizan la variabilidad que los índices sintéticos logran retener a partir de las observaciones.

El IM, al igual que el IRS, facilita la comparación de unidades territoriales en un momento dado, resumiendo en un único valor para cada entidad federativa o municipio la información contenida en las nueve variables que lo integran. Debido a ello, aporta conocimiento acerca de si una región padece mayor marginación que otra. Al recurrir a la técnica de componentes principales, se garantiza que la variabilidad de los datos que el índice logra conservar sea la máxima posible. No obstante, la reducción desde las nueve variables observadas a un único valor teórico implica una inevitable pérdida de información, que no sólo es mínima, sino también conocida a partir de los autovalores de la matriz de covarianzas de los datos.

El IM presenta algunas limitaciones, por ejemplo, no considera el número de personas que habita cada unidad territorial (Conapo, 1994: 16). Asimismo, no permite analizar directamente la evolución a lo largo del tiempo de las condiciones de vida de los residentes en un área geográfica específica (Gutiérrez y Gama, 2010: 233; Téllez *et al.*, 2016: 12; Peláez, 2017); sólo admite comparaciones directas de datos correspondientes a un mismo momento. No obstante, tal vez el problema más relevante que ha surgido en los últimos años es el que afecta a su principal virtud, esto es, a la capacidad de resumen que le otorga su sistema de ponderaciones. Si bien los pesos obtenidos mediante la técnica de componentes principales son siempre óptimos por definición, en años recientes la pérdida de información, resultado de la reducción de la dimensionalidad de las nueve variables a un sólo indicador, es cada vez mayor. Esto se debe a que las variables que forman parte del índice, actualmente correlacionan entre sí de manera menos intensa que hace unos años. Al guardar menos similitudes entre ellas, cada una aporta más información no redundante, por lo que es más difícil concentrar en un único indicador.

La solución propuesta por Cortés y Vargas (2013) consiste en prestar atención no sólo al primer componente principal, sino también al segundo, ya que ambos, de manera conjunta, consiguen retener un porcentaje aceptable de la variabilidad total de los datos observados. El problema de su propuesta es que la interpretación del segundo componente principal no es tan clara como el del primero, ya que aquél no pondera todas las variables con pesos positivos, sino también incluye

algunos negativos, dando lugar a ordenaciones poco intuitivas de las unidades territoriales.

El objetivo de este artículo es examinar a detalle la reducción en el porcentaje de varianza total, explicada por el primer componente principal (y consecuentemente por el IM) que ha tenido lugar durante las últimas décadas, y proponer una solución alternativa al análisis del segundo componente principal que facilite la interpretación de los resultados. Con este propósito, en principio, se verifica la progresiva disminución del primer autovalor del análisis factorial y se comentan las diferencias existentes entre el primer y el segundo componente principal. En el apartado dos se discute la solución al problema que plantea el deterioro de la capacidad de síntesis del IM. En el tercer apartado se comprueban las ventajas de esta propuesta. Finalmente, se presentan las conclusiones de la investigación.

1. ESTADO DE LA CUESTIÓN: EVOLUCIÓN DEL AUTOVALOR DEL IM, 1990-2015

Mediante el análisis de componentes principales, nueve variables observables se pueden transformar en nueve componentes teóricas, sin pérdida de información. Los componentes se expresan como combinaciones lineales de las variables originales; se obtienen ordenadas de mayor a menor varianza y están incorrelacionadas entre sí. El caso es que, si las variables observadas correlacionan fuertemente entre ellas, la información que contienen es redundante, pues el primer componente principal tiende a concentrarla, de manera que gran parte de la varianza total es la correspondiente al primer componente. El resto de la varianza se distribuye entre los otros componentes en orden descendente. En esa situación, el primer componente resume de manera adecuada la información originalmente contenida en las nueve variables observadas. Al reducirse la dimensión del problema de nueve variables a una, es más sencillo manejar, comparar e interpretar los valores de las mediciones. Si las unidades territoriales con mayores porcentajes de población analfabeta coinciden con las de altos niveles de población en situación de hacinamiento, en viviendas sin drenaje ni excusado, sin energía eléctrica, entre otras; todas las variables correlacionan fuertemente entre sí y el primer componente principal, que concentra esa información concordante, indicando que las condiciones

de bienestar son peores en esas regiones. El primer componente principal es al que Conapo da el nombre de IM.

El problema con este planteamiento es que, si las variables observadas no correlacionan de manera intensa, la información que contienen no es redundante, sino que cada una aporta detalles específicos y, por tanto, no se puede resumir bien en un componente. En esta situación, el primer componente principal sigue siendo el de mayor varianza, pero ya no representa una proporción tan alta del total.

Es necesario considerar otros componentes para abarcar la riqueza informativa de las distintas variables. Si sólo se atiende a la primera, se pierde gran parte de la variabilidad original. La ventaja de la técnica de componentes principales es que permite cuantificar esa pérdida, valorar la idoneidad de la reducción de dimensiones y tomar una decisión al respecto.

La varianza de cada componente viene dada por el autovalor correspondiente de la matriz de correlaciones de los datos, $\lambda_j, j=1, \dots, 9$. Dado que los componentes principales están incorrelacionados entre sí, su varianza total es la suma de los nueve autovalores. Por tanto, la proporción de la variabilidad total contenida en la j -ésima componente principal se puede calcular como:

$$p_j = \lambda_j / \sum_{j=1}^9 \lambda_j \quad \forall j = 1, \dots, 9.$$

Además, la suma de las varianzas de los nueve componentes debe ser igual a la suma de las varianzas de las nueve variables. Como se trabaja con variables estandarizadas, sus varianzas son unitarias, de modo que:

$$\sum_{j=1}^9 \lambda_j = 9$$

Entonces, la proporción de varianza que aporta cada componente al total se calcula como:

$$p_j = \lambda_j / 9 \quad \forall j = 1, \dots, 9.$$

Cabe esperar que el primer componente concentre la mayor parte de la variabilidad total. De lo contrario, no podrá considerarse como un buen resumen del conjunto de variables observadas.

En este sentido, Cortés y Vargas (2013) encuentran que, con el transcurso del tiempo, el primer componente principal ha ido perdiendo relevancia en favor del segundo y de los subsiguientes componentes. Cada año, el IM explica una menor parte de la variabilidad total de los datos, hasta el punto de que para resumir de forma adecuada la información municipal de 2010 se hace necesario recurrir también al segundo componente principal.

En un sentido semejante, pero para los datos de 2015, Téllez *et al.* (2016: 349) explican: “En el caso de las entidades federativas, los nueve indicadores generaron solo un nuevo componente, factor o medida resumen para conformar el índice de marginación por entidad federativa. En los municipios, en cambio, se obtuvieron dos componentes, sin embargo, solo el primero se consideró como el índice de marginación”. Por su lado, Bracamontes y Camberos (2010: 57) llegan a un resultado casi idéntico para un ejercicio similar con datos de los municipios de Sonora.

Cortés y Vargas (2013) únicamente analizan los datos municipales de 1990, 2000 y 2010. No recurren a información de los conteos de población porque ésta no es estrictamente comparable con la de los censos para el ejercicio que realizan, en el que no sólo estudian la capacidad de resumen de los componentes principales, sino también estiman modelos para verificar la dependencia temporal del IM. Por su parte, Téllez *et al.* (2016) utilizan datos de la encuesta intercensal de 2015.

En esta ocasión, con el fin de conocer con el mayor detalle posible la cuestión de la pérdida de relevancia del primer componente principal, lo que equivale a la progresiva reducción del primer autovalor, resulta conveniente considerar toda la información disponible a pesar de que no sea estrictamente comparable. Por ello, se utilizan los datos de Conapo (2016a y b), a partir de los cuales se pueden calcular los autovalores y autovectores de los análisis de componentes principales, de municipios y entidades federativas, desde 1990 a 2015.

El Cuadro 1 muestra que, efectivamente, el primer autovalor se redujo desde 5.62, en 1990, cuando representaba 62.45% de la variación total de los datos municipales, hasta 4.54, en 2010, concentrando 50.47% de la varianza. En 2015, el primer autovalor es aún menor:

resulta sólo 49.34% de la suma total. Al incluir los datos del conteo de 2005 y de la encuesta intercensal de 2015 junto con los de los censos de 1990, 2000 y 2010, no se observan rupturas en la tendencia a la reducción del primer autovalor y aumento de los siguientes, al contrario, queda más clara esta dinámica.

Cuadro 1
AUTOVALORES DEL ANÁLISIS DE COMPONENTES PRINCIPALES
PARA LOS DATOS DEL IM A NIVEL MUNICIPAL, 1990-2015

Año	1990	2000	2005	2010	2015
λ_1	5.62	5.22	4.86	4.54	4.44
λ_2	0.84	0.97	1.04	1.06	1.12
λ_3	0.69	0.74	0.90	0.92	0.89
λ_4	0.49	0.70	0.73	0.76	0.73
λ_5	0.40	0.48	0.49	0.57	0.59
λ_6	0.31	0.38	0.43	0.43	0.45
λ_7	0.30	0.23	0.26	0.40	0.38
λ_8	0.24	0.20	0.22	0.22	0.31
λ_9	0.12	0.09	0.09	0.09	0.09
Suma	9	9	9	9	9

Fuente: elaboración propia con datos de Conapo (2016a).

El criterio de Kaiser indica que deben retenerse los k primeros componentes, tales que $\lambda k \geq 1$ (1960: 146), esto es, sólo aquellos componentes con varianza mayor a la de las variables originales; para una correcta descripción de los municipios deberían tenerse en cuenta los dos primeros componentes a partir de 2005, y no sólo el primero. Por tanto, se confirma lo señalado por Cortés y Vargas (2013: 7) y Téllez *et al.* (2016: 349).

Cabe comentar que, en el caso de las entidades federativas (Cuadro 2), no ocurre lo mismo que con los municipios. Si bien es cierto que también hay una tendencia a la reducción del primer autovalor, en ningún año el segundo supera la cifra de 0.75. En 1990, el primer componente principal concentraba 78.58% de la varianza total. En 2000, esta participación incluso se amplía hasta el máximo de 79.46%. A partir de ese momento se inicia una progresiva reducción que lleva a que, en 2015, el primer autovalor (6.42) represente 71.35% del total. La diferencia fundamental con el caso de los municipios es que, a pesar del

deterioro de la capacidad de resumen del primer componente principal, éste sigue siendo una buena medida de resumen del conjunto de datos.

Cuadro 2

AUTOVALORES DEL ANÁLISIS DE COMPONENTES PRINCIPALES
PARA LOS DATOS DEL IM A NIVEL ESTATAL, 1990-2015

Año	1990	1995	2000	2005	2010	2015
λ_1	7.07	6.80	7.15	6.80	6.62	6.42
λ_2	0.59	0.75	0.61	0.71	0.67	0.73
λ_3	0.42	0.50	0.37	0.52	0.63	0.63
λ_4	0.32	0.28	0.29	0.37	0.42	0.47
λ_5	0.23	0.19	0.20	0.23	0.27	0.30
λ_6	0.14	0.16	0.14	0.20	0.19	0.19
λ_7	0.11	0.15	0.12	0.09	0.10	0.15
λ_8	0.09	0.12	0.08	0.06	0.07	0.08
λ_9	0.04	0.05	0.03	0.03	0.03	0.03
Suma	9	9	9	9	9	9

Fuente: elaboración propia con datos de Conapo (2016b).

En cualquiera de los casos, la menor capacidad explicativa del primer componente principal se debe a que, con el transcurso del tiempo, las correlaciones entre las nueve variables del IM se han ido debilitando. Cortés y Vargas (2013: 16) señalan tres factores que han contribuido al progresivo debilitamiento de estas correlaciones: i) el tipo de política social del país desde la crisis económica de 1994, ii) la diferencia en las temporalidades de los distintos fenómenos captados por las variables de marginación, y iii) la evolución de la economía mexicana, que ha impedido que la población goce de crecimientos sostenidos en sus ingresos. Como solución a la menor capacidad de síntesis del IM, proponen considerar también el segundo componente principal, que mediría carencias en vivienda, mientras que el primero (el IM), carencias en urbanización. No obstante, los mismos autores explican las dificultades que entraña esta propuesta: “si se retienen dos factores en lugar de solo uno, cabe preguntarse cuál de los factores mide la marginación o bien hay que emplear ambos, con lo cual se complica su aplicación” (Cortés y Vargas, 2013: 10). A ello se suma que la interpretación del segundo componente principal no es tan clara como la del primero. Las ponderaciones de este último (Cuadro 3) son siempre positivas porque se refieren a variables que miden rezagos y que tienen “un claro refe-

rente conceptual como indicadores de la marginación social” (Ávila *et al.*, 2001: 185). El resumen fundamental de las nueve variables es que correlacionan de manera intensa, unas más que otras, pero todas en la misma dirección, porque todas miden carencias. Una vez extraída esa relación, se obtiene el segundo patrón de interrelaciones más importante, reflejado en la estructura de ponderaciones del segundo componente principal (Cuadro 4). Para 2015, casi igual que para 2010, se advierte una fuerte relación positiva entre los indicadores de viviendas sin drenaje ni excusado, sin energía eléctrica y sin agua entubada, al tiempo que se observa una relación negativa de estos con las dos variables de educación, con la variable de ruralidad y con la de ingresos.

Cuadro 3

PONDERACIONES DEL PRIMER COMPONENTE PRINCIPAL, MUNICIPIOS

	1990	2000	2005	2010	2015
1) Población analfabeta	0.154	0.169	0.182	0.193	0.199
2) Sin primaria completa	0.160	0.173	0.182	0.194	0.196
3) Sin drenaje ni excusado	0.146	0.112	0.091	0.100	0.099
4) Sin energía eléctrica	0.127	0.129	0.126	0.127	0.132
5) Sin agua entubada	0.129	0.117	0.125	0.127	0.123
6) Hacinamiento	0.135	0.150	0.159	0.158	0.172
7) Piso de tierra	0.156	0.165	0.173	0.149	0.171
8) Menos de 5,000 habitantes	0.117	0.123	0.131	0.141	0.138
9) Hasta dos salarios mínimos	0.135	0.160	0.167	0.189	0.164

Fuente: elaboración propia con datos de Conapo (2016a).

Cuadro 4

PONDERACIONES DEL SEGUNDO COMPONENTE PRINCIPAL, MUNICIPIOS

	1990	2000	2005	2010	2015
1) Población analfabeta	-0.164	-0.004	-0.103	-0.138	-0.137
2) Sin primaria completa	0.118	-0.197	-0.222	-0.233	-0.231
3) Sin drenaje ni excusado	0.087	-0.247	0.118	0.493	0.446
4) Sin energía eléctrica	-0.491	0.493	0.570	0.523	0.540
5) Sin agua entubada	-0.416	0.566	0.539	0.395	0.326
6) Hacinamiento	-0.145	0.084	-0.020	0.078	0.015
7) Piso de tierra	-0.102	0.179	0.119	-0.045	0.083
8) Menos de 5,000 habitantes	0.682	-0.479	-0.345	-0.319	-0.295
9) Hasta dos salarios mínimos	0.489	-0.320	-0.377	-0.291	-0.360

Fuente: elaboración propia con datos de Conapo (2016a).

Por los aspectos comentados sobre los referentes conceptuales, las correlaciones entre variables y las ponderaciones consecuentes, Ávila *et al.* (2001: 184) igualan el IM al primer componente principal. Como medida de resumen, cabe esperar que las unidades territoriales con mayores valores en el IM sean las que padecen mayores porcentajes de población analfabeta, sin primaria completa, habitantes de viviendas sin drenaje, sin energía eléctrica, sin agua entubada, etcétera.

El segundo componente principal también permite ordenar las unidades territoriales, pero asigna valores mayores a aquellas con altos porcentajes de viviendas sin drenaje ni excusado, sin energía eléctrica y sin agua entubada, al tiempo que, con bajos porcentajes de población analfabeta, se ponderan negativamente carencias como primaria incompleta, con menos de 5 000 habitantes y con salarios insuficientes. Este segundo componente se puede entender como un índice de carencias en vivienda.

No obstante, además de valorar de manera positiva las carencias en las viviendas, penaliza a los municipios rezagados en términos de educación, salarios y distribución de la población. La interpretación del segundo componente resulta bastante más compleja que la del primero, donde todos los rezagos se incluyen con el mismo signo.

Hay un detalle adicional: a partir del análisis de la matriz de correlaciones, Cortés y Vargas (2013) establecen dos grupos de variables: las relativas a la “carencia en vivienda”, que incluyen las viviendas sin electricidad, sin agua entubada, con hacinamiento y con piso de tierra, y agruparon las otras cinco variables del IM en la “carencia en urbanización”. Sin embargo, al examinar las ponderaciones del segundo componente en los años 2010 y 2015 (Cuadro 4), surgen algunas diferencias. Si bien es cierto que el segundo componente puede interpretarse como “carencia en vivienda”, las variables que más la caracterizan son tres: viviendas sin energía eléctrica, viviendas sin drenaje ni excusado, y viviendas sin agua entubada. Las restantes variables apenas tienen peso, incluso negativo.

2. METODOLOGÍA PARA UNA SOLUCIÓN ALTERNATIVA: LOS ÍNDICES DE URBANIZACIÓN Y VIVIENDA

Ante el problema de la pérdida de capacidad de resumen por parte del primer componente principal, lo cual es especialmente grave en

el caso de los municipios, se propuso considerar también el segundo componente para medir carencias en vivienda, mientras que el primero sintetizaría la escasez en urbanización. Para evitar las dificultades que introduce la bidimensionalidad de la marginación, en el presente artículo se recomienda separar en dos grupos las nueve variables que integran el IM y calcular un factor para cada uno de esos grupos de variables más correlacionadas entre sí. A partir de los análisis factoriales anteriores (cuadros 3 y 4), se propone un primer grupo formado por tres variables, concretamente por los porcentajes de ocupantes en viviendas: sin drenaje ni servicio sanitario, sin energía eléctrica, y sin agua entubada. Estas tres variables son las que presentan mayor peso en el segundo componente principal y, por ello, lo definen en los años 2010 y 2015 (Cuadro 4). Además, se trata de las tres variables con menor peso en el primer componente principal durante los mismos años (Cuadro 3).

El segundo grupo de variables lo forman los seis porcentajes restantes: población analfabeta, sin primaria completa, hacinamiento, piso de tierra, residentes en localidades con menos de 5 mil habitantes y ocupados con ingresos de hasta 2 salarios mínimos. Como se indicó, los grupos de variables no coinciden exactamente con los que establecen Cortés y Vargas (2013) a partir del análisis de las matrices de correlaciones, aunque guardan muchas semejanzas, por lo que se puede afirmar que “dentro de la marginación municipal subyacen dos procesos medidos en términos de carencias: urbanización y vivienda” Cortés y Vargas (2013: 17). Algo muy parecido se observa en el segundo componente que Bracamontes y Camberos (2010: 58) calculan para los municipios de Sonora con datos de 2000.

Al aplicar la técnica de componentes principales a las nueve variables, se obtienen nueve componentes, donde el primero de ellos es el IM, pero éste es insuficiente para explicar la riqueza informativa del conjunto de datos. Si la técnica se aplica al grupo de tres variables, se obtienen tres componentes. En el siguiente apartado se comprueba que el primero de estos tres componentes es suficiente para resumir esas variables correlacionadas fuertemente entre sí, y no tanto con las restantes. Además, las ponderaciones de ese primer componente tenderán a ser positivas, facilitando su interpretación al asignar más valor a las unidades territoriales con mayores porcentajes de ocupantes en viviendas: sin drenaje ni servicio sanitario, sin energía eléctrica y

sin agua entubada. A semejanza del IM con las nueve variables, este primer componente puede considerarse como un índice de carencias en los servicios de las viviendas (ICV).

Algo similar ocurre con el grupo de seis variables. Al aplicar el análisis de componentes principales, se obtienen seis componentes, de los cuales el primero tenderá a ser suficiente para resumir la información esencial del conjunto, además de presentar ponderaciones positivas. Este primer componente del grupo de seis variables puede interpretarse como índice de carencias en urbanización (ICU).

Las ventajas de usar la técnica de componentes principales son las siguientes: a) asigna las ponderaciones que hacen máxima la variabilidad de las observaciones que los índices logran retener; b) cuantifica la información perdida al reducir la dimensión de los datos y, en este caso, permite seguir trabajando en el marco de referencia del IM, con observaciones y análisis que abarcan más de dos décadas. Las limitaciones son las habituales de cualquier indicador sintético.

En concreto, Bustos (2011: 170) destaca que la optimalidad de las ponderaciones del análisis de componentes principales es una propiedad esencialmente matemática, basada en la estructura de las correlaciones, por lo que no necesariamente representa la influencia real de las variables en el fenómeno que se está midiendo. No obstante, este es un problema inherente a cualquier medida de resumen. Al respecto de las ponderaciones, Scott y Bloom (1997: 89-90) argumentan que “es difícil comparar (aun intrapersonalmente) la pérdida relativa de bienestar por no tener electricidad con la pérdida por ser analfabeta, mucho más comparar distribuciones en estas dimensiones para poblaciones enteras”.

Siendo conscientes de estas limitaciones, a partir del conjunto de nueve variables originales se pueden obtener el IM, el ICV y el ICU. El cotejo de los resultados de Ávila *et al.* (2001: 194), Anzaldo y Prado (2006: 331) y De la Vega *et al.* (2011) aporta un primer indicio de lo que señalan Cortés y Vargas (2013: 17) sobre los procesos de urbanización y vivienda: “estaban vinculados en los años 1990 y 2000 de modo que eran empíricamente indistinguibles y que se desligaron en la primera década del siglo XXI” De la Vega, (2011: 331). Esto se refleja en la reducción del primer autovalor y en el aumento del segundo, especialmente en el caso de los municipios (Cuadro 1); se indica la progresiva aparición de dos procesos diferenciados donde antes sólo había uno.

Dado este fenómeno, la ordenación de los territorios generada por el ICV se asemejaría más a la del ICU, y ambas a la del IM al inicio del periodo de análisis, aumentando las diferencias con el transcurso del tiempo. La semejanza de las ordenaciones de los índices se puede medir con el coeficiente de rangos de Spearman:

$$\rho = 1 - \frac{6}{n(n^2 - 1)} \sum_{i=1}^n d_i^2$$

que está acotado entre -1 (total discordancia en las ordenaciones) y 1 (total concordancia), que toma el valor 0 cuando las ordenaciones son independientes, y donde n es el número de unidades territoriales (entidades federativas o municipios) y d_i es la diferencia entre los órdenes de uno y otro índice.

A partir de lo comentado, cabe esperar que ρ esté más próximo a la unidad en los primeros años que en los más recientes, cuando los procesos de urbanización y vivienda han empezado a disociarse. Esto mismo podría comprobarse mediante el coeficiente de correlación lineal r , en el que la covarianza se divide por el producto de las desviaciones estándar, y que también está acotado entre -1 y 1 .

La ventaja del coeficiente ρ es que no se limita a relaciones lineales. Su principal inconveniente es que se restringe a comparar ordenaciones, sin valorar la información cardinal de los datos. En cualquier caso, se verificó que la aplicación de uno u otro coeficiente no altera los resultados ni las conclusiones a las que se llega en el presente artículo.

3. RESULTADOS

Los autovalores que se muestran en el Cuadro 5 indican que el primer componente principal de cada grupo es suficiente para retener la información esencial de cada conjunto de variables. Para cualquier año, el único autovalor superior a la unidad es el primero, cumpliendo así el criterio de Kaiser (1960: 146). En el caso de las carencias en los servicios de las viviendas, el primer componente, el ICV, concentra entre 71% y 55% de la variabilidad total, dependiendo del año. Para las carencias en urbanización, el primer componente aglutina entre 70% y 63% de la varianza total de las seis variables de ese grupo.

Cuadro 5

**AUTOVALORES DEL ANÁLISIS DE COMPONENTES PRINCIPALES
PARA LOS DATOS DEL ICV Y DEL ICU A NIVEL MUNICIPAL, 1990-2015**

Año	1990	2000	2005	2010	2015
ICV					
λ_1	2.13	1.77	1.69	1.69	1.66
λ_2	0.50	0.76	0.82	0.77	0.83
λ_3	0.37	0.47	0.48	0.54	0.52
ICU					
λ_1	4.08	4.22	4.10	3.86	3.80
λ_2	0.76	0.77	0.77	0.80	0.79
λ_3	0.46	0.39	0.44	0.60	0.54
λ_4	0.31	0.30	0.36	0.41	0.45
λ_5	0.26	0.23	0.25	0.23	0.34
λ_6	0.12	0.09	0.10	0.09	0.09

Fuente: elaboración propia con datos de Conapo (2016a).

La ventaja fundamental de trabajar con dos índices en vez de dos componentes del IM es que las ponderaciones utilizadas para su construcción a partir de las variables son todas positivas (Cuadro 6). Ello implica que, por ejemplo, el ICV asigne valores más altos a las unidades territoriales con mayores problemas de viviendas sin drenaje, sin energía eléctrica y sin agua entubada. En el segundo componente del IM sucede lo mismo pero, además, se ponderan negativamente las carencias en educación, ingreso y distribución de la población (Cuadro 4). Por ello, la interpretación del ICV y del ICU es más intuitiva que la del segundo componente.

Cuadro 6

PONDERACIONES DEL ICV Y DEL ICU A NIVEL MUNICIPAL, 1990-2015

Año	1990	2000	2005	2010	2015
ICV					
3) Sin drenaje ni excusado	0.39	0.37	0.35	0.41	0.42
4) Sin energía eléctrica	0.39	0.47	0.50	0.49	0.51
5) Sin agua entubada	0.41	0.46	0.47	0.44	0.42
ICU					
1) Población analfabeta	0.21	0.21	0.22	0.23	0.24
2) Sin primaria completa	0.22	0.22	0.22	0.24	0.24
6) Hacinamiento	0.19	0.19	0.19	0.18	0.20
7) Piso de tierra	0.22	0.20	0.20	0.18	0.20
8) Menos de 5 000 habitantes	0.17	0.16	0.16	0.17	0.17
9) Hasta dos salarios mínimos	0.20	0.21	0.21	0.23	0.21

Fuente: elaboración propia con datos de Conapo (2016a).

Los municipios con ICV más alto en 1990, esto es, los que padecían mayores carencias en vivienda eran: San Jacinto Tlacotepec (3.19), Camocuautla (3.05), Huaytlalpan (3.03) y Metlatónoc (3.02). En el otro extremo de la distribución, con los menores ICV del país, se encontraban las delegaciones de Benito Juárez (-1.63), Venustiano Carranza (-1.62) y Cuauhtémoc (-1.62).

Cabe mencionar que tanto el ICV como el ICU, al ser componentes principales, comparten las propiedades del IM; en concreto, para cada conjunto de datos, su media siempre es cero y su desviación estándar igual a la unidad. Los municipios o delegaciones que se encuentran mejor que la media, presentan valores negativos, mientras que aquellos que soportan carencias superiores al promedio, tienen valores positivos más altos cuanto mayores son sus carencias.

Los municipios que en 1990 mostraban un ICU más alto eran: Tehuipango (2.60), Coicoyán de las Flores (2.48) y San Simón Zahuatlán (2.47). Tehuipango también tuvo el IM más alto de ese año: 2.77 (Conapo, 2016a). Coicoyán de las Flores, sin embargo, clasificaba en el tercer lugar en el IM, mientras que San Simón Zahuatlán era el quinto municipio con mayor IM en 1990 (Conapo, 2016a). Debido a que el ICU comparte seis variables con el IM, las ordenaciones que genera son mucho más parecidas a las de este índice que las del ICV.

Esto es así para todos los años, como se ve a continuación con el coeficiente de rangos de Spearman, y es coherente con los análisis realizados por otros autores mencionados. Las carencias en servicios de las viviendas han quedado cada vez más relegadas al segundo componente principal, ignorado por el IM, pero recuperable a través del ICV. El ICU tenderá a arrojar resultados más parecidos a los del IM, ya que sólo contiene la información predominante en este último índice.

Cuadro 7
COEFICIENTES DE RANGOS DE SPEARMAN, MUNICIPIOS

Año:	1990	2000	2005	2010	2015
ICV-ICU	0.797	0.722	0.647	0.639	0.640
ICV-IM	0.901	0.830	0.759	0.751	0.743
ICU-IM	0.978	0.983	0.984	0.984	0.986

Fuente: elaboración propia con datos de Conapo (2016a).

Cuadro 8

COEFICIENTES DE RANGOS DE SPEARMAN, ENTIDADES FEDERATIVAS

Año	1990	1995	2000	2005	2010	2015
ICV-ICU	0.913	0.874	0.890	0.836	0.802	0.787
ICV-IM	0.951	0.921	0.933	0.918	0.899	0.897
ICU-IM	0.988	0.988	0.989	0.976	0.975	0.969

Fuente: elaboración propia con datos de Conapo (2016b).

Los coeficientes de rangos de Spearman más bajos se obtienen al comparar las ordenaciones del ICV con las del ICU (Cuadro 7). Bastantes más semejanzas guardan las ordenaciones generadas por cualquiera de estos dos indicadores con las del IM. Al fin y al cabo, no son más que partes de éste. Destaca especialmente la alta concordancia entre el ICU y el IM, con coeficientes cercanos a la unidad.

Respecto a la evolución de las semejanzas entre índices, resalta que el ICV cada vez se parece menos al ICU y al IM, y estos últimos cada vez arrojan resultados más parecidos entre sí. Esto responde a la progresiva diferenciación de los procesos de urbanización y vivienda durante décadas recientes y que, en última instancia, ha llevado al deterioro de la capacidad de síntesis del IM, que capta mayoritariamente las carencias del proceso de urbanización e ignora las de vivienda.

Si bien la necesidad de considerar un segundo componente principal afecta al análisis municipal y no al estatal, el ICV y el ICU se pueden calcular de igual forma para las entidades federativas del país, donde el IM también ha ido perdiendo capacidad de resumen. En las entidades federativas resulta aún más claro que el primer componente principal de cada grupo es suficiente para aglutinar la información esencial de cada conjunto de variables.

El primer componente de las carencias en servicios de las viviendas, el ICV, retiene entre 81% y 67% de la varianza total. El ICU, por su parte, es más estable y su contribución a la variabilidad total se sitúa entre 84% y 81%. El segundo autovalor más alto es el de las carencias en vivienda durante 2015, que alcanza un valor de 0.66, todavía alejado de 1. Lo mismo que en la desagregación por municipios, las ponderaciones de todas las variables son positivas en ambos índices todos los años. El Distrito Federal es la entidad que obtiene un menor valor tanto en el ICV como en el ICU desde 1990 hasta 2015.

En el extremo opuesto, Chiapas se mantiene como la entidad federativa con mayores carencias en urbanización en cada año del periodo analizado. En cuanto a carencias en vivienda, destaca nuevamente Chiapas, primer lugar en el ICV en 1990 y 1995, pero también Guerrero, que alcanza ese primer lugar en 2000 y lo mantiene hasta la actualidad.

Las ordenaciones generadas por el ICV, el ICU y el IM guardan más semejanzas entre sí (Cuadro 8) que en el caso de los municipios (Cuadro 7). Incluso, el ICV da lugar a ordenaciones muy parecidas a las del ICU. Esto se relaciona con que para las entidades federativas no es necesario tomar un segundo componente principal, a diferencia de lo que ocurre con los municipios a partir de 2005 (Cuadro 1). No obstante, el progresivo deterioro experimentado por la capacidad de resumen del IM también tiene reflejo (Cuadro 8). La concordancia entre el ICV y el ICU se reduce con el transcurso del tiempo; ocurre lo mismo con el IM, aunque en menor medida.

CONCLUSIONES

A lo largo de las últimas décadas, las variables del IM han dado lugar a dos procesos diferenciados. El más relevante de ellos, el que aglutina más variables y dimensiones, es el vinculado a las carencias de la urbanización. El otro proceso, definido por las carencias en servicios de las viviendas, ha sido relegado por el primero y tiene menos peso en el IM, especialmente a nivel de municipios. No obstante, su creciente relevancia como fenómeno independiente explica el mayor peso del segundo componente principal y el progresivo debilitamiento de la capacidad de síntesis del IM.

Con el fin de no tener que considerar simultáneamente dos componentes, dados los problemas operativos y de interpretación que ello implica, se propuso el cálculo de dos índices parciales, uno para cada proceso. Estos arrojan resultados claros y coherentes con los hallazgos de otras investigaciones, lo que permite concluir que, efectivamente, ha habido una disociación de los procesos de urbanización y vivienda, de manera que el IM de los últimos años está captando predominantemente los aspectos vinculados a las carencias en urbanización, sobre todo en el caso de los municipios. Dado que la construcción e interpretación de cada índice es idéntica a la del IM, sólo que para el subgrupo de variables correspondiente, su cálculo permite identificar de manera especí-

fica qué municipios y entidades federativas padecen mayores carencias en urbanización o en vivienda, recuperando parte de la información que el IM está dejando de considerar. También es cierto que los resultados de estos índices parciales aún guardan muchas semejanzas con los del IM, en especial los de urbanización y algo menos los de vivienda.

BIBLIOGRAFÍA

- Anzaldo, Carlos y Minerva Prado (2006), *Índices de marginación, 2005*. México: Conapo.
- Ávila, José Luis; Carlos Fuentes y Rodolfo Tuirán (2001), *Índices de marginación, 2000*. México: Conapo.
- Bracamontes, Joaquín y Mario Camberos (2010), “¿Concentración o convergencia en el crecimiento y desarrollo de Sonora?”, *Frontera Norte*, 22 (44). Tijuana: El Colegio de la Frontera Norte, pp. 41-78.
- Bustos, Alfredo (2011), “Niveles de marginación: una estrategia multivariada de clasificación”, *Realidad, Datos y Espacio: Revista Internacional de Estadística y Geografía*, 2 (1). Aguascalientes: Instituto Nacional de Estadística y Geografía, pp. 169-186.
- Conapo / Conagua (1994), *Desigualdad regional y marginación municipal en México*, 1990. México.
- Conapo (2016a), “Índice de marginación por municipio 1990-2015”, *Datos abiertos del índice de marginación*. México. <http://www.conapo.gob.mx/es/CONAPO/Datos_Abiertos_del_Indice_de_Marginacion> [25 de diciembre de 2016].
- Conapo (2016b), “Índice de marginación por entidad federativa 1990-2015”, *Datos abiertos del índice de marginación*. México. <http://www.conapo.gob.mx/es/CONAPO/Datos_Abiertos_del_Indice_de_Marginacion> [25 de diciembre de 2016].
- Coneval (2016), “Índice de rezago social 2015 a nivel nacional, estatal y municipal”, *Medición de la pobreza: Índice de rezago social*. México. <http://www.coneval.org.mx/Medicion/IRS/Paginas/Indice_Rezago_Social_2015.aspx> [30 de diciembre de 2016].
- Coplamar (1982), *Geografía de la marginación: necesidades esenciales de México*. México: Siglo XXI.
- Cortés, Fernando (2002), “Consideraciones sobre la marginalidad, marginación, pobreza y desigualdad en la distribución del ingreso”, *Papeles de Población* 8 (31). Toluca: UAEM, pp. 9-24.

- Cortés, Fernando (2006), “Consideraciones sobre la marginación, la marginalidad, marginalidad económica y exclusión social”, *Papeles de Población*, 12 (47), Toluca: UAEM, pp. 71-84.
- Cortés, Fernando y Delfino Vargas (2013), “La dependencia temporal de la marginación municipal en México 1990-2010: una tercera mirada al índice de marginación”, documento de trabajo del Programa Universitario de Estudios del Desarrollo. México: UNAM. <<http://www.pued.unam.mx/archivos/opinion/001.pdf>> [30 de noviembre de 2016]
- De la Vega, Sergio; Raúl Romo y Ana L. González (2011), *Índice de marginación por entidad federativa y municipio 2010*, México: Conapo.
- DOF (2016), “Decreto por el que se formula la Declaratoria de las Zonas de Atención Prioritaria para el año 2017”, en *Diario Oficial de la Federación*, 30 de noviembre.
- Germani, Gino (1973), *El concepto de marginalidad. Significado, raíces históricas y cuestiones teóricas, con particular referencia a la marginalidad urbana*. Buenos Aires: Nueva Visión.
- Gutiérrez, Humberto y Viviana Gama (2010), “Limitantes de los índices de marginación de Conapo y propuesta para evaluar la marginación municipal en México”, *Papeles de Población*, 66. Toluca: UAEM, pp. 227-257.
- Kaiser, Henry F. (1960), “The application of electronic computers to factor analysis”, *Educational and Psychological Measurement* 20 (1). Santa Barbara, EE. UU.: University of California, pp. 141-151.
- Peláez, Óscar (2017), “La marginación a lo largo del tiempo: cálculo del Índice de Marginación Absoluta (IMA) para las entidades federativas de México, 1970-2010”, *Economía: Teoría y Práctica*, 46. México: UAM, pp. 115-137.
- PNUD (1990), *Desarrollo humano: informe 1990*. Colombia: Tercer Mundo Editores.
- Scott, John y Erik Bloom (1997), “Criterios de asignación para la superación de la pobreza en México”, *Economía Mexicana*, 6 (1). México: Centro de Investigación y Docencia Económicas, pp. 83-159.
- Sen, Amartya (1983), “Poor, relatively speaking”, *Oxford Economic Papers*, 35 (2). Oxford: University Press Oxford, pp. 153-169.
- Téllez, Yolanda; Rubén Almekjo; Aldo R. Hernández y Raúl Romo (2016), *Índice de marginación por entidad federativa y municipio 2015*. México: Conapo.